**GENOME WIDE ASSOCIATION STUDIES TO IDENTIFY RICE SALT-TOLERANCE MARKERS**

**Juan Patishtan**

**Tom N Hartley**

**Raquel Fonseca de Carvalho**

**Frans JM Maathuis**

Department of Biology, University of York, York YO10 5DD United Kingdom

Corresponding author: Frans J.M. Maathuis

Department of Biology; University of York

York YO10 5DD, United Kingdom.

Tel: +44 1904 328652

Fax: +44 1904 328505

email: frans.maathuis@york.ac.uk

**ABSTRACT**

Salinity is an ever increasing menace that affects agriculture world-wide. Crops such as rice are salt sensitive but its degree of susceptibility varies widely between cultivars pointing to extensive genetic diversity which can be exploited to identify genes and proteins that are relevant in the response of rice to salt stress. We used a diversity panel of 306 rice accessions and collected phenotypic data after short (6h), medium (7d) and long (30d) salinity treatment (50 mM NaCl). A genome wide association study (GWAS) was subsequently performed which identified around 1200 candidate genes from many functional categories but this was treatment period dependent. Further analysis showed the presence of cation transporters and transcription factors with a known role in salinity tolerance and those that hitherto were not known to be involved in salt stress. Localisation analysis of single nucleotide polymorphisms (SNPs) showed the presence of several hundred non-synonymous SNPs (nsSNPs) in coding regions and earmarked specific genomic regions with increased numbers of nsSNPs. It points to components of the ubiquitination pathway as important sources of genetic diversity that could underpin phenotypic variation in stress tolerance.

**KEYWORDS:** GWAS; Genome Wide Association Study; *Oryza sativa*; membrane transport; salinity, rice.

**INTRODUCTION**

There are an estimated 40,000 or more varieties of rice (Kushwaha, 2016) within the two main cultivated and 21 wild *Oryza* species (Londo *et al.*, 2006). Cultivation of *Oryza sativa* and *Oryza glaberrima* occurs all over the world making rice the second largest crop in terms of global tonnage.

*O. sativa* is the main cultivated species. It derives from the wild species *Oryza rufipogon (Kim et al., 2016)* and its domestication probably occurred between 12,000 and 7000 years ago. The two main *O. sativa* subspecies are *indica* and *japonica*. Both of these can be further divided into several subpopulations (Wang *et al.*, 2016; McCouch *et al.*, 2016; Huang *et al.*, 2012).

Genomic mapping is a useful approach in plant breeding to locate regions that are important for particular traits (Grotewold *et al.*, 2015). Quantitative trait loci (Pandit *et al.*, 2010) analyses, typically point to specific chromosomal sub-regions whereas the more recently developed genome wide association studies (GWASs) can identify precise chromosomal locations with the resolution to define specific genes or even polymorphisms within coding regions (Korte and Farlow, 2013; Si *et al.*, 2016). In general, these approaches employ statistical formalisms to determine the strength of the association between a genotype and phenotype and as such help inform breeding by providing molecular markers, or by identifying genes and alleles that contribute to specific traits.

In rice, GWAS has revealed candidate genes that underlie traits such as yield (Liang *et al.*, 2016), flowering time (Zhao *et al.*, 2011), root morphology (Biscarini *et al.*, 2016), panicle architecture (Bai *et al.*, 2016), as well as grain size, length and shape (Si *et al.*, 2016; McCouch *et al.*, 2016; Feng *et al.*, 2016). Where biotic and abiotic stress are concerned, a small number of candidate genes has been associated with blast pathogen resistance (Zhu *et al.*, 2016), response to drought (Huang *et al.*, 2010), ozone (Ueda *et al.*, 2015) and aluminium (Famoso *et al.*, 2011).

Salinity creates an abiotic stress that greatly diminishes rice yield across the globe. Though some wild varieties have relatively high levels of tolerance, most elite cultivars show high salt sensitivity and it is imperative that the extensive genetic diversity of this trait is exploited to combine high yield with salt resistance (Mohanty *et al.*, 2002; Gregorio *et al.*, 2002). Only a very limited number of studies has applied GWAS strategies to unravel the molecular mechanism that generates tolerance (Kumar *et al.*, 2015). To further detail and expand the genetic basis that governs rice responses to salt stress, we applied GWAS to a rice diversity panel of ~300 cultivars (McCouch *et al*., 2010) that represents all major subpopulations and geographic areas. We deliberately applied a level of salinity (50 mM) that imposes little or moderate stress in tolerant rice while growth depression in sensitive lines is severe. By treating plants for a short (6 hour), medium (7 days) and long (30 days) period we hypothesise that we would discover genes that contribute to temporary variability in rice responses but also ones that play a more constitutive role. GWAS identified genes that were previously shown to play a role in salt tolerance but also discovered many new genes, particularly those involved in ubiquitination dependent protein degradation. The outcomes will help understand the physiological basis of salinity tolerance and provide suitable markers for future breeding efforts.

**MATERIALS AND METHODS**

**Plant material and growth conditions**

All experiments were carried out in a glasshouse with 32/28oC day/night temperature with 50-60% RH and 12h day/night period. For GWAS, seeds of 306 rice varieties (see Table 1, Suppl. data) were germinated in clay based terragreen substrate. At 15 days old, 1 seedling of each variety was transferred to 8L boxes in a randomised configuration, and exposed to hydroponic standard medium (SM; Yoshida *et al.*, 1976). Medium was changed each week. After 30 days growth on SM, plants were exposed to either control or salt stress conditions for 6 hours (short term), 7 days (medium term) or 30 days (long term) durations. Salinisation of the medium was achieved by the addition of 50 mM NaCl. The entire procedure was repeated two more times to yield three data points for each treatment and variety derived from three biological replicates that were grown between August 2014 and July 2015.

For pot grown plants, 15d old seedlings were transplanted to 8 cm pots containing compost (John Innes). Plants were grown for another 15 days before treatment which consisted of watering with tap water (50 ml/pot/day) or a same volume of tap water containing 50 mM NaCl. Treatment lasted for 30 days and three plants per accession per treatment were grown. Shoot fresh and dry weights were determined.

Initial fresh and final fresh and dry weights were recorded and relative growth rate (RGR) was calculated as (lnW2-lnW1)/(t2-t1) where W1 and W2 are start and end weight measured at t1 and t2 (in days). For measurements of tissue K+ and Na+ flame photometry was applied as described in Ahmad et al. (2016). Briefly, fresh tissues were dried at 80 °C after which tissue was extracted in 10 ml of 20 mM CaCl2. Extracts were analysed using a flame photometer. Growth data and tissue cation data were obtained from three independent biological replicates for each experiment and treatment condition.

**Phenotype data analyses**

Phenotypic data were analysed using ANOVA (SAS 9.3 package). Tukey’s honest significance test (Tukey HSD; P<0.05) was used to test differences between the means of the data sets. Correlation analyses were carried out using the SAS 9.3 platform (SAS Institute, 2011).

**GWAS protocol**

GWAS analysis was carried out using the R package GenABEL (Aulchenko et al., 2007) for 306 diverse rice cultivars. To analyse SNP calls, a high-density rice array of 700,000 SNPs was used derived from the sequencing of wild and domesticated rice accessions at ~ 7 x genome coverage (McCouch et al., 2016). SNPs were removed from the analysis if their minor allele frequency was less than 5% across the panel or the SNP genotype was unknown for > 20% of the cultivars. This typically resulted in ~200,000 remaining SNPs. The GenABEL package uses the Family-Based Score Test for Association (FASTA) and a mixed linear model (MLM) to correct false positive associations derived from population structure, using the top four principal components calculated from the kinship matrix as covariate (Chen and Abecasis, 2007). A false discover rate of 5% was used to correct for multiple testing.

**Identification of candidate genes and polymorphisms**

To identify genes that were potentially associated with the measured phenotype parameters, two different approaches were followed, using phenotypic data from short term (ST, 6h), medium term (MT, 7days) and long term (LT, 30days) treatments. For each treatment period, tissue cation data, and tissue cation data combined with growth data, were used as input traits (see Suppl. Table 4 for GWAS analyses, listing a total of 28 sets of data). Firstly, for each association signal a genomic area of 200 kb around the most significant (i.e. the highest –log10 P value) SNP call was defined (this size is based on the average linkage disequilibrium in rice, e.g. Zhao et al., 2015). This genomic region was then interrogated using the Phytozome-11 rice genome browser (<https://phytozome.jgi.doe.gov/pz/portal.html>) to identify all genes belonging to that particular association peak. All genes with coding regions within the 200kb window were included, however those annotated as '(retro)transposon', 'hypothetical' or 'unknown' were excluded from the analysis.

In a second approach, all SNPs that exceeded the significance criterion (Suppl. Table 4) were assessed for location in coding regions of candidate genes. If present in coding regions, the potential impact on protein composition of each SNP was subsequently determined, using the 'Allele finder' facility (http://rs-bt-mccouch4.biotech.cornell.edu/AF/)(McCouch et al., 2016) **Gene Ontology analysis**

Candidate gene lists from short, medium and long term treatments were submitted to the 'AgriGO' ontology enrichment facility (http://bioinfo.cau.edu.cn/agriGO/analysis.php which is based on a Fisher's exact test and a Yekutieli multitest adjustment, using a 5% false positive detection threshold.

**RESULTS AND DISCUSSION**

**Growth characteristics**

For medium and long term growth in SM, the RGR of rice cultivars varied by a factor of around 2-2.5. Figure 1A shows data for long term RGR which varied from ~2.7 to 6.6% day-1 and very similar rates (2.9-7.2%) were obtained for medium term growth (Data not shown). Salinisation (50 mM NaCl) of the nutrient solution caused mild to severe growth reduction with RGR ranging from 0.026-7.3% for medium term treatment (Fig. 1B). Again, RGRs for long term saline growth were very similar to those for medium term salt treatment, ranging from 0.02 to 4.9%. On the basis of these growth characteristics, relatively tolerant and sensitive accessions are listed in Suppl. Table 2. The growth data show two important characteristics. Firstly, growth variability expands drastically when plants become salt stressed (from around 2.5-fold to around 25-fold variability between lowest and highest rates). Secondly, the lack of difference between medium and long term suggests that in both non-saline and saline conditions, growth of rice stabilises within ~7 days of treatment to a level that subsequently changes little, at least up to 30 days. Furthermore, the data show that 50 mM NaCl causes severe stress in sensitive lines (RGR approaching zero) while very little negative impact is seen in the most tolerant lines, suggesting it is an appropriate concentration to test germplasm for genetic variation.

RGRs provide a good indicator of vigour and to test whether this property was consistent across treatments we carried out correlation analyses between control and treatment RGR values. No significant correlations were detected indicating that RGR values from control and saline treatments are not related. This clearly indicates that growth characteristics of accessions are greatly affected by salinity and vigour in control conditions is not a good predictor of tolerance. Furthermore, RGRs are not necessarily convenient parameters to express tolerance. We therefore analysed relative change in RGR for each accession, expressed as '% reduction' by comparing non-saline and saline RGR. For medium term treatment, RGR-reduction was ~2 to 24% (Fig. 2A) while in the long term experiment RGR-reduction was ~7 to 99% (Fig. 2B). Thus, for both treatments large differences in stress between cultivars occurred with some lines being hardly affected by NaCl exposure while growth virtually stopped in others.

On the basis of hydroponics growth properties, the 5 most tolerant and 5 most sensitive accessions (Suppl. Table 2) were identified and grown in salinised soil to assess if the growth phenotypes were consistent across different media. Figure 3 shows that for the sensitive lines, growth reductions are significantly smaller in soil compared to hydroponics. Although these pot experiments are not equivalent to field trials, the results do confirm that differences in tolerance are largely maintained irrespective of whether rice was grown in liquid medium or in soil.

To establish whether growth patterns had any subpopulation related characteristics, RGRs and RGR-reductions were separately analysed for the two major subpopulations of indica and japonica varieties. Average RGR and RGR-reduction in all growth conditions were not significantly different between these subpopulations , irrespective of whether medium term or long term treated plants were concerned.

**K+ and Na+ tissue concentrations**

Although yield is agronomically highly significant, it is not necessarily informative where mechanisms of salt tolerance are concerned. Furthermore, growth is the culmination of multiple physiological processes that subsume the activity of many hundreds of genes. Given the many potential factors contributing to this trait (both physiological and environmental), many markers that are identified on the basis of associations with growth traits will be non-specific in the context of tolerance to salt stress. In contrast, the use of tissue cation levels, especially Na+ could help point to more specific phenotype-genotype associations. We therefore measured root and shoot Na+ and K+ levels in tissues from all control and salinity treated plants.

Table 1 shows root and shoot K+ and Na+ (rootK, shootK, rootNa and shootNa) levels for control and each of the 3 treatment experiments. In roots, 6h salt exposure caused a drastic increase in Na+ levels (range 110-719, average 334 µmol gDW-1) but during the same period only a very small proportion of Na+ was translocated to the shoot (range 8-58, average 25). The latter property may stem from the 'excluder' eco-physio-type of rice which generally leads to preferential accumulation of salt in roots and prevention of high Na+ in photosynthetic tissues. After 7d treatment with 50 mM NaCl, root tissue Na+ accumulation increased to an average of 502 µmol gDW-1 (range 180-1213) while it still remained relatively low in shoots (average 312 µmol gDW-1). Long term NaCl exposure did not essentially alter this pattern in roots with long term Na+ levels at an average 457 µmol gDW-1 but do suggest a breakdown of regulatory systems to prevent Na+ translocation to the shoot in many cultivars. In the period from 7 to 30 days, the average shoot Na+ more than doubled to around 700 µmol gDW-1. Interestingly, even after this prolonged exposure, around 10 accessions managed to limit shoot Na+ to values below 100 µmol gDW-1, a value that roughly equates to 10 mM on FW basis (assuming a 10-fold FW-1:DW-1 ratio).

Tissue K+ is increased to some extent by salinisation with the exception of root K+ in the long term treatment which dropped from ~440 to ~315 µmol gDW-1. However, it is evident that variation in tissue K+ values between accessions remains fairly constant and is approximately 5-fold at most. In contrast, values of tissue Na+ fluctuate by a factor of almost 100 suggesting a far larger genetic diversity for this trait compared to tissue K+ contents.

Exposure to saline conditions will immediately cause Na+ influx into root tissues, especially since the electrochemical gradient for Na+ is high and inward. It has been argued that uptake of inorganics like Na+ and Cl- and, to a lesser extent K+, is a relatively energy efficient manner of osmotic adjustment to saline conditions (Yeo *et al.*, 1991; Cramer, 2002). Thus, influx of Na+ and Cl- may be beneficial in the early stages but, due to build up of ion toxicity, a risky strategy in the longer run. To assess how tissue cation levels relate to tolerance (either in the form of final plant weight, RGR or expressed as RGR-reduction), correlations were calculated between various parameters and those that gave significant values are presented in Suppl. Table 3. No correlations were detected in the short term experiment between tissue cation levels and/or growth parameters. However, after medium term salinisation, moderate (0.4>r<0.59) but significant (p<0.01) correlations could be found between shootK and shootNa. Furthermore, Fig. 4A shows the presence of a correlation between rootK and RGR for long term salinity treatment while a clear and strong negative correlation between shootNa and growth was also present (Fig. 4B) which varied between -0.5 and -0.62 depending on the way growth is expressed. Similar correlations were found when growth was expressed in other ways, e.g. as RGR-reduction or as final plant weight (Suppl. Table 3). After long term salinisation, additional moderate correlations were detected which showed a negative coefficient in the case of shootK and growth.

A strong negative correlation between shootNa and growth during salinity has been observed before in the case of rice (e.g. Platten et al. 2013). However, the correlation between shootK and shootNa combined with the presence of a negative correlation between shootK and growth (as observed for both the medium and long term trials) suggests tolerance is relatively low in lines that show elevated shootK and that during salinity some of the ionic toxicity may actually derive from K+. This notion is at odds with the frequently reported concept that high tissue K+ levels promote salt tolerance (Tester and Davenport, 2003).

**Association mapping**

The diversity panel used in this study largely consists of commercial cultivars that have been selected for high yields. Nevertheless, both the growth (Figs. 1 and 2) and tissue cation concentration results (Table 1) show that there is considerable diversity between lines regarding tolerance, irrespective of whether it is based on RGR, RGR-reduction or when considering rootK and shootNa, parameters that closely correlate with salt tolerance. In turn, these findings strongly suggest that considerable genetic diversity is present within the panel which may underpin the variation in these traits.

To test this notion, phenotypic and genotypic data were interrogated for the occurrence of significant associations. In order to avoid the impact of generic growth effects, growth data (RGR, RGR-reduction and final mass) on their own were not used as parameters for these analyses. Instead, data derived from tissue cation concentrations on their own or in combination with growth parameters (during saline growth) were entered as traits and used to scrutinise the entire complement of SNPs. Using these conditions and a 5% false discovery rate threshold, association signals were obtained. Figure 5 shows an example Manhattan plot for shootNa data from the medium term trial. A clear association signal is visible located at the beginning of chromosome 6.

In combination, GWAS analyses of the three trials, using various traits and trait combinations, led to a total of ~1900 significant SNPs and ~ 2300 candidate genes (Suppl. Tables 4 and 5). The latter reduced to ~950 when duplicates, transposons and non-annotated genes were removed (Suppl. Table 6). The contribution of various trials (short, medium and long term) and treatments to generate significant SNP signals varied greatly: when analysing tissue K+ levels obtained from plants grown in SM (control treatment), no significant signals were obtained. Combinations of tissue K+ and growth traits similarly did not yield any signals in control treatment that exceeded the significance threshold.

In the case of short term salt exposure only tissue data and plant weights were obtained because 6h was too short to reliably determine (relative) growth rates. Analyses using root and shoot cation levels on their own revealed 7 association peaks with ~55 significant SNPs in the case of shootNa but none for rootK, rootNa or shootK. However, when the same parameters were used in combination with plant weight data, 20 association signal peaks (Supp. Table 7) were found which subsumed a further ~400 SNPs. A total of 600 candidate genes was found in the 200 kb windows flanking the top SNP of each of these 20 association peaks.

 Phenotypic data from medium term NaCl exposure were entered in various combinations and generated a total of ~25 significant association signals composed of ~1000 SNPs (Suppl. Table 5 and 6). Searches in the surrounding genomic areas identified ~750 candidate genes (Suppl. Table 5). As seen for short term trials, there was a wide variation between traits in generating association signals: for example, rootK in combination with RGR data associated with ~20 candidate genes whereas shootNa and rootNa data associated with >100 candidate genes.

GWAS output on the basis of long term NaCl phenotypic data showed ~26 association signals across all chromosomes except 9 (Suppl. Table 6). After long term exposure, many 'tissue+RGR' and 'tissue+RGR-reduction' combinations of phenotype data produced overlapping association peaks which suggests that RGR and RGR-reduction values from salinised plants are highly correlated. The association peaks covered ~550 SNPs and led to the identification of ~800 candidate genes.

After removal of genes encoding unknown/hypothetical proteins, gene ontology analysis (http://bioinfo.cau.edu.cn/agriGO/analysis\_precheck.php?session=697416059) showed that the total complement of around 950 candidate genes was derived from many functional classes but was enriched in categories in a time dependent manner (Fig. 6). The short treatment is characterised by cell death, electron transport and photosynthesis. After medium duration treatment, processes to do with protein synthesis (amino acid activation) and (posttranslational) protein modulation in the form of phosphorylation (kinases) and protein breakdown (ligation) become prominent. Only after long term treatment is there a significant enrichment in transport of lipids and ions. Both transcriptional and posttranslational regulation remain prominent in this phase. In all, these results suggest that during the early phases, cell death and readjustment of energy generation are the main responses while large scale protein modification (including protein degradation, a function that is highly over represented in candidate genes that were identified in all three trials, see below) only occurs after several days. The absence of ion transport as an enriched function in the short and medium term treatments is somewhat counterintuitive but could suggest that natural variation in transport proteins is limited. Comparison between trials showed that 105 candidate genes were shared between short and medium term treatments, 128 between medium and long term treatments and 129 candidate genes were identical between short and long term treatments. A total of 51 candidate genes was found in all three trials (Suppl. Table 8). The recurrence of genes in short, medium and long term treatments could signify that they have roles in salt stress response in all growth phases and therefore have great potential. But, it could equally point to genes and processes that play generic roles during stress and are not specifically related to salt stress. Interestingly, out of the 51 genes listed in Suppl. Table 8, there are at least four that are involved in ubiquitination: rice contains ~30 Spk1 genes (0.15% of its genome) and 3 of these (~6%) emerged in all three trials, pointing to a high degree of overrepresentation. Spk1 proteins are components of the SCF (Skp-cullin-F-box) complex, the multi-protein ubiquitin ligase that catalyses the ubiquitination of proteins destined for proteasomal degradation. F-box proteins like FBX289 contribute to the specificity of SCF by binding to target proteins and then associate to the Skp1 component of the complex.

Since we were particularly interested in trying to unravel the genetic basis for the role of tissue cation (K+ and Na+) levels in relation to salt tolerance, candidate genes were further curated with a focus on (a) K+ and Na+ transporters (b) membrane transport, and (c) regulatory proteins that could be involved in modulating cation transport functions. Apart from K+ and Na+ transporters, specific genes (membrane transporters, kinases, phosphatases, transcription factors, etc.) were included only if evidence from literature and/or data from public expression databases pointed to a potential role in salinity/drought tolerance.

Amongst the ~70 remaining identified genes (Table 2), 5 candidates have previously been shown to be involved in responses to dehydration and primarily consist of transcription factors and one LEA type chaperone. Many of these are also responsive to other types of abiotic stress (Genevestigator; Hruz *et al.*, 2008). Interestingly, the two heat stress related transcription factors show relatively higher expression in salt sensitive RILs compared to tolerant RILs (Genevestigator; Hruz *et al.*, 2008) which could be the result of polymorphisms at this locus. Several anion transporters appear in Table 2: CLCs may be involved in NO3- or Cl- transport and, like malate channels, could be relevant for balancing large amounts of Na+ that are accumulated in the vacuole. Sulphur is indispensable in the generation of reducing power; modulation of sulphate transport may be a corollary of increased oxidative stress which requires increased reducing power in the form of SH groups.

Eight kinases and two phosphatases were found but in all cases details are lacking regarding their function. Salt, drought and ABA responses have all been shown to significantly alter the phosphorylation status of many proteins (Chitteti and Peng, 2007). This implies the activity of specific kinases and phosphatases but with the rice genome encoding hundreds of genes that encode these enzymes, there is a major task in unravelling networks that regulate salt stress responses. Regulatory pathways may be instigated by Ca2+ signals: it has been suggested that Ca2+ signalling is important in the response of plants to salinity. In most studies, the recorded Ca2+ signal is primarily caused by osmotic effects but the large number of Ca2+ related candidate genes could support a role of Ca2+ signalling in salt/drought. Expression of several rice H+:Ca2+ exchangers from the CAX family has been shown to correlate to salinity tolerance of different cultivars (Senadheera *et al.*, 2009). Vacuolar Ca2+ may contribute to osmotic balancing but CAXs may also be important in restoring Ca2+ homeostasis after signalling. Two Ca2+ permeable channels were revealed for which evidence points to a role in salt tolerance. One (OsCA1;4) is a plasma membrane located channel which, in analogy to its Arabidopsis counterpart, could function as a osmosensor. OSCAs show a modular architecture with a 6-9 transmembrane helices region, a low-complexity region, a coiled-coil region, and the DUF221 domain. OSCA1;4 is expressed in many tissues (Genevestigator; Hruz *et al.*, 2008). No expression variation was detected between salt sensitive and salt tolerant RILs (Genevestigator; Hruz *et al.*, 2008) but transcript levels differed in the offspring of many crosses (e.g. Zhenshan\*Minghui) and in comparative studies between accessions (e.g. IR64 and Nipponbare). This shows there is cultivar related variability at this locus which may underlie phenotypic variation. The tonoplast localised non selective cation channel TPC1 was also discovered from the GWAS analysis. Though non-selective, the electrophysiological properties of TPC1 channels prevent a role in Na+ transport but recent literature suggests that TPC1 has a role in salt induced long distance Ca2+ signalling (Choi et al., 2014). Potential downstream components of salt/drought related Ca2+ signalling were present in the form of calmodulin binding proteins and Ca2+/calmodulin related kinases in the form of two CAMK/CIPK isoforms).

Four peptide and amino acid transporters were identified. Interestingly, transcriptional changes in response to salt/drought of these were, without exception, down, often with values of more than 5-fold. RIL lines derived from a cross between Zhenshan and Mingui show considerable variability in expression of several these proteins, confirming the notion of genetic diversity at these loci.

Many publications have shown the extensive transcriptional changes that occur in all plant tissues after exposure to salinity. Numbers of transcripts that show a significant response (either up- or downregulated) typically run in the many hundreds or, more often, several thousand. Identifying relevant transcription factors is therefore of great importance because it provides a putative shortcut to manipulating large numbers of genes simultaneously. Transcription factors of the bZIP, MYB and WRKY subfamilies have been shown to play a role in abiotic stress in many plant species, including drought and salinity stress in rice (Wang *et al.*, 2015; Hossain *et al.*, 2010; Ding *et al.*, 2014). Indeed, several showed large fold changes in response to salt/drought with the MYB transcription factor Os01g19330 being down by a factor >10 in response to drought and showing variation within Zhenshan\*Mingui derived RILs.

A large group of transporters was present in the curated list with proteins that contribute to uptake and distribution of various cations. ZIPs have been shown to catalyse uptake of micronutrients such as Zn and Fe (Chen *et al.*, 2008). Though a link between ionic aspects of salinity stress and transition metal homeostasis has been postulated (Maathuis, 2006) there is as yet little hard evidence to substantiate it. Ammonium transporters maybe important for several reasons: salinity may cause accumulation of nitrogen compounds in rice, including ammonium (Zhao *et al.*, 2008). Furthermore, the biosynthesis of compatible solutes like proline requires large amounts of reduced N. Thus, both ammonium nitrate transport may require adjustment during salt stress.

A relatively large number of K+ transporters was identified. This is not surprising since a close link has been found between the chemically similar K+ and Na+ cations (Benito *et al.*, 2014; Maathuis *et al.*, 2014). Indeed, one of the main toxicity aspects of Na+ is believed to originate in the displacement of K+ as cofactor by Na+ in essential enzymatic activity (Bertorello *et al.*, 1991; Tester and Davenport, 2003). Furthermore, high levels of external Na+ can inhibit K+ uptake and hence scupper osmotic rebalancing (Horie *et al.*, 2011). OsKAT1 was previously shown to be relevant for salt tolerance of rice cultured cells (Obata *et al.*, 2007), possibly in its capacity to prevent K+ deficiency during salinity. In Arabidopsis, the K+ channel AKT2,3 is upregulated during salinity (Maathuis, 2006). AKT2,3 has been postulated to function in phloem mediated recirculation of K+ (Maathuis *et al.*, 2014) from shoot to root, an essential part of plant K+ homeostasis which is likely to be affected by salinity, especially in relation to root/shoot partitioning of Na+ and K+. HAK9 is significantly downregulated in drought and salinity while HAK11 and 25 show upregulation in drought conditions (Genevestigator; Hruz *et al.*, 2008). HAK9 and 11 show contrasting expression patterns with HAK9 primarily located in shoot tissue and HAK11 in roots and floral tissues. HAK25 is ubiquitously present. These K+ carriers may provide auxiliary roles where K+ uptake and distribution are concerned (Chen *et al.*, 2015; Nieves-Cordones *et al.*, 2016).

The role of rice HKT transporters in Na+ uptake and distribution is well established (Ren *et al.*, 2005). For HKT1;5, involved in retrieval of Na+ from the xylem, allelic variation has been shown to underlie important differences between tolerant and sensitive cultivars (Platten *et al.*, 2013). Tolerance is associated with greater HKT1;5 selectivity, in turn determined by four amino acid substitutions in the cytoplasmic loop domains of the protein (Rubio *et al.*, 1999), which results in less shoot Na+ accumulation. Only one HKT isoform (HKT1;3) emerged from our GWAS. In roots, OsHKT1;3 is expressed in peripheral layers and in the stele (Rosas-Santiago *et al.*, 2015; Jabnoune *et al.*, 2009). Within stelar tissue it is predominantly located in the phloem whereas in shoots, HKT1;3 is found in vasculature and in bulliform cells which are involved in leaf rolling (Véry *et al.*, 2014). It is yet to be determined what the physiological role of OsHKT1;3 is but from its prominent presence in bulliform cells a role in generating turgor changes to control leaf rolling in response to water stress is possible. Furthermore, presence in stelar tissues implies a contribution to long distance partitioning of Na+.

Candidate genes with potential roles in cation homeostasis further included non-selective channels of the CNGC (CNGC9) and glutamate-like receptor (GLR) families. These could participate in uptake and distribution of K+, Na+ and Ca2+. Some Arabidopsis CNGCs have been shown to be involved in Na+ uptake in roots and in Na+ distribution (Maathuis, 2006; Leng *et al.*, 2002), but evidence for a similar role in rice is as yet absent.

**The search for causative SNPs**

GWAS is generally capable of pinpointing more precise genomic loci than for example QTL-based methods, but resolution is still often a problem. Large, significant signals can easily cover extensive genomic regions each encompassing dozens of significant SNPs that may be related to multiple candidate genes. As shown above, subsequent use of bioinformatics and literature can help narrow down large gene collections and increase confidence of specific candidate genes. Alternatively, the presence of SNPs in coding sequences, especially non-synonymous SNPs (nsSNPs) with allele frequencies that correlate with the trait of interest, provides another route to assess the likely relevance of candidate genes. nsSNPs have been shown to be able to cause significant phenotypic variation. For example, a single SNP in an arsenate reductase caused several fold changes in accumulated arsenic while an nsSNP in the vacuolar K+ channel AtTPK1 affected its Ca2+ dependence (Hartley and Maathuis, 2015). Out of the ~2000 significant SNPs that were obtained from all treatments and trait combinations, ~310 were located in intronic regions and ~290 in exons though this number is raised to ~400 if transposons and hypothetical proteins are included. The latter implies that 15-20% of all significant SNPs are located in coding sequence, a percentage that is around 10 times higher than expected by chance assuming that ~2% of eukaryotic genomes consists of coding sequence.

In the curated list of candidate genes, our GWAS did not identify any nsSNPs but a total of 7 synonymous SNPs (sSNPs) was detected of which 5 were in exons and 2 resided in introns. Two sSNPs were present in RLCK348, a ubiquitously expressed 994 amino acid proteins with two kinase domains (Vij *et al.*, 2008). Its transcription is significantly downregulated by drought and salinity conditions but drastically increased in response to cold and heat (Genevestigator; Hruz *et al.*, 2008), suggesting a role in multiple abiotic stresses. No different expression levels were detected between tolerant and sensitive RILs but large differences in response to drought were observed for example between IRAT109 and Zhenshan97. OSCA1;4 (a putative hyperosmolality-gated calcium-permeable channel) and an 'early dehydration response' gene contained one sSNP while the calmodulin binding protein Os11g44680, contained a synonymous SNP in exon 4 and one in exon 6 (Yuan *et al.*, 2014). Interestingly, its transcriptional regulation mirrors that of RLCK348 with significant downregulation in drought and salinity conditions but drastically increased transcript levels in response to cold and heat (Genevestigator; Hruz *et al.*, 2008). At this locus too, there is considerable transcript variability between accessions, especially between rice *indica* (IR24, IR29, IR64) and *japonica* populations (Nipponbare, Li-Jiang-Xin-Tuan-Hei-Gu) in response to drought and temperature stress.

In these 4 cases, genome sequence SNPs only caused synonymous substitutions. Although the protein changing effects of nsSNPs provide a more direct route to explaining phenotypic variation, the role of sSNPs is by no means insignificant. For example, sSNPs have the capacity to affect mRNA splicing, and can alter the penetrance of concurring mutations elsewhere in the gene. sSNPs in human pathology have been shown to impact on mRNA stability and translation and as such contribute to variation in dopamine receptors (Duan *et al.*, 2003). It has even been argued that sSNPs that lead to rare codons can affect protein folding by creating translational pauses which may be vital for correct protein conformation (Hunt *et al.*, 2009).

Amongst the exonic SNPs, ~155 were nsSNPs and led to either amino acid substitutions or premature stop codons (Suppl. Table 9). Several genes contained multiple nsSNPs while in one gene, OsFBX289, four nsSNPs were identified. FBX289 is an F-box domain containing protein that is involved in protein degradation via the ubiquitination complex. It is also potentially significant that FBX289 appeared in all three trials, just as many other proteins that are components of the E3 ligase complex that binds proteins for degradation such as the Skp1 isoforms. The genomic region where FBX289 is located (chromosome 8, ~17.4 Mb) contains 9 nsSNPs in approximately 300 kb. This, and a second region in chromosome 6 (~17.5 Mb) with 7 nsSNPs suggest they are highly relevant in salt tolerance.

**GWAS versus QTLs**

A high number of QTLs related to salinity tolerance has been identified. Saltol-1 is a major QTL associated shoot Na/K ratios which maps to chromosome 1 (Fig 7; ~9.3Mb – 16.4Mb) (Bonilla P 2002; Soda et al. 2013). OsHKT1;5 (positioned @11.45 Mb) was identified as a major contributor to Saltol and is involved in Na+ retrieval from the xylem (Ren et al. 2005). GWAS identified a major association signal in its vicinity (~11.1 Mb). Another well studied gene within Saltol is the lectin domain containing protein SalT (Os01g24710; Claes et al., 1990) which was also identified in this study (Suppl Table 5). Figure 7 shows many more examples of GWAS signals that overlap with previously identified QTLs which suggests consistency between studies and methodology.

**In conclusion**

Rice cultivars show a remarkably wide degree of growth responses to salinity which strongly suggests that there is large genetic variation in this species that can be exploited. Insights into the contribution of particular genes may help us to unravel tolerance mechanisms and, via breeding or engineering, can inform which alleles are beneficial. Previous genomics studies (e.g. QTL analyses, microarray studies) and mutation screens have already identified hundreds of genes that are involved in responses to salt stress but there are likely many others still to be discovered. Furthermore, exactly how genetic polymorphisms that are present within populations affect protein function and phenotypes is largely obscure. GWAS analyses can throw light on both of these unknowns, especially when using appropriate populations and high resolution genotyping (McCouch *et al.*, 2016).

We applied GWAS to a diversity panel that represents rice accessions from all areas of the globe and is therefore likely to contain a high degree of genetic variability. A total of around 950 genes was identified which belonged to various functional categories but was overrepresented in GO classifications of transport (especially cation transport), hydrolase activity and transcriptional regulation (Fig. 6). By using three treatment periods we were able to compare genes that likely play a role in early, medium or late responses. Such comparisons showed that the overall functional categories (e.g. cation transport) are similar for each treatment period but individual genes are not. Furthermore, whereas cation transporters do not appear at all, a group of specific genes related to ubiquitination was consistently found in all three treatments.

**REFERENCES**

Ahmad, I., Mian, A. & Maathuis, F. J. (2016) Overexpression of the rice AKT1 potassium channel affects potassium nutrition and rice drought tolerance. *Journal of experimental botany,* 67**,** 2689-2698.

Alexandrov, N., Tai, S., Wang, W., Mansueto, L., Palis, K., Fuentes, R. R., …, Ulat, V. J. (2015) SNP-Seek database of SNPs derived from 3000 rice genomes. *Nucleic acids research* 43, 1023-1027.

Ammar, M. H. M., Awadhesh Pandit, R. K. Singh, S. Sameena, M. S. Chauhan, A. K. Singh, P. C. Sharma, Kishor Gaikwad, T. R. Sharma, Trilochan Mohapatra, and N. K. Singh. 2009. 'Mapping of QTLs Controlling Na+, K+ and CI− Ion Concentrations in Salt Tolerant Indica Rice Variety CSR27', *Journal of Plant Biochemistry and Biotechnology*, 18: 139-50.

Aulchenko, Y. S., De Koning, D. J. & Haley, C. (2007) Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics,* 177**,** 577-585.

Bai, X., Zhao, H., Huang, Y., Xie, W., Han, Z., Zhang, B., ..., Guo, Z. (2016). Genome-Wide Association Analysis Reveals Different Genetic Control in Panicle Architecture Between and Rice. *The Plant Genome* 9, 2.

Benito, B., Haro, R., Amtmann, A., Cuin, T. A. & Dreyer, I. (2014) The twins K+ and Na+ in plants? *Journal of Plant Physiology,* 171**,** 723-731.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)***,** 57, 289-300.

Bertorello, A. M., Aperia, A., Walaas, S. I., Nairn, A. C. & Greengard, P. (1991) Phosphorylation of the catalytic subunit of Na+, K(+)-ATPase inhibits the activity of the enzyme. *Proceedings of the National Academy of Sciences,* 88**,** 11359-11362.

Biscarini, F., Cozzi, P., Casella, L., Riccardi, P., Vattari, A., Orasen, G., ..., Perrini, R. (2016) Genome-Wide Association Study for Traits Related to Plant and Grain Morphology, and Root Architecture in Temperate Rice Accessions. *PLOS ONE*, 11, e0155425.

Bonilla P, Dvorak J, Mackill D, Deal K, Gregorio G. 2002. 'RFLP and SSL mapping of salinity tolerance genes in chromosome 1 of rice (Oryza sativa) using recombinant inbred lines. ', *Philippine Agricultural Scientist*, 85: 8.

Chen, G., Hu, Q., Luo, L., Yang, T., Zhang, S., Hu, Y., ..., Yu, L. (2015) Rice potassium transporter OsHAK1 is essential for maintaining potassium‐mediated growth and functions in salt tolerance over low and high potassium concentration ranges. *Plant, cell & environment*, 38, 2747-2765.

Chen, W., Feng, Y. & Chao, Y. (2008) Genomic analysis and expression pattern of OsZIP1, OsZIP3, and OsZIP4 in two rice (Oryza sativa L.) genotypes with different zinc efficiency. *Russian Journal of Plant Physiology,* 55**,** 400-409.

Chitteti, B. R. & Peng, Z. (2007) Proteome and phosphoproteome differential expression under salinity stress in rice (Oryza sativa) roots. *Journal of Proteome Research,* 6**,** 1718-1727.

Choi, W.G., Toyota, M., Kima, S., Hilleary, S., Gilroy, S. (2014) Salt stress-induced Ca2+ waves are associated with rapid, long-distance root-to-shoot signaling in plants. *Proceedings of the National Academy of Sciences,* 111, 6497-6502.

Claes, B., R. Dekeyser, R. Villarroel, M. Van den Bulcke, G. Bauw, M. Van Montagu, and A. Caplan. 1990. 'Characterization of a rice gene showing organ-specific expression in response to salt stress and drought', *Plant Cell*, 2: 19-27.

Cramer, G. R. (2002) Sodium-calcium interactions under salinity stress. *Salinity: Environment-Plants-Molecules,* 205-227.

Ding, Z. J., Yan, J. Y., Xu, X. Y., Yu, D. Q., Li, G. X., Zhang, S. Q. & Zheng, S. J. (2014) Transcription factor WRKY46 regulates osmotic stress responses and stomatal movement independently in Arabidopsis. *The Plant Journal,* 79(1), 13-27.

Duan, J., Wainwright, M. S., Comeron, J. M., Saitou, N., Sanders, A. R., Gelernter, J. & Gejman, P. V. (2003) Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Human molecular genetics,* 12**,** 205-216.

Famoso, A. N., Zhao, K., Clark, R. T., Tung, C. W., Wright, M. H., Bustamante, C., ..., Kochian, L. V. (2011) Genetic architecture of aluminum tolerance in rice (Oryza sativa) determined through genome-wide association analysis and QTL mapping. *PLOS Genetics,* 7 (8)**,** e1002221.

Feng, Y., Lu, Q., Zhai, R., Zhang, M., Xu, Q., Yang, Y., ..., Wang, S. (2016) Genome wide association mapping for grain shape traits in indica rice. *Planta***,** 244, 819-830.

Ghomi, K., B. Rabiei, H. Sabouri, and A. Sabouri. 2013. 'Mapping QTLs for traits related to salinity tolerance at seedling stage of rice (Oryza sativa L.): an agrigenomics study of an Iranian rice population', *OMICS*, 17: 242-51.

Gregorio, G., Senadhira, D., Mendoza, R., Manigbas, N.,Roxas, J. & Guerta, C. (2002) Progress in breeding for salinity tolerance and associated abiotic stresses in rice. *Field Crops Research,* 76**,** 91-101.

Grotewold, E., Chappell, J. & Kellogg, E. A. (2015) The Shifting Genomic Landscape in Plant Genes. *Genomes and Genetics*, John Wiley and Sons, Ltd, Chichester, U.K. 17-44.

Hartley, T. N. & Maathuis, F. J. (2015). Allelic variation in the vacuolar TPK1 channel affects its calcium dependence and may impact on stomatal conductance. *FEBS letters*, 590, 1, 110-117.

Horie, T., Brodsky, D. E., Costa, A., Kaneko, T., Schiavo, F. L., Katsuhara, M. & Schroeder, J. I. (2011) K+ transport by the OsHKT2;4 transporter from rice with atypical Na+ transport properties and competition in permeation of K+ over Mg2+ and Ca2+ ions. *Plant physiology,* 156**,** 1493-1507.

Hossain, M. A., Lee, Y., Cho, J.-I., Ahn, C.-H., Lee, S.-K., Jeon, J.-S., ..., Kang, H. (2010) The bZIP transcription factor OsABF1 is an ABA responsive element binding factor that enhances abiotic stress signaling in rice. *Plant molecular biology,* 72**,** 557-566.

Hossain, H., M. A. Rahman, M. S. Alam, and R. K. Singh. 2015. 'Mapping of Quantitative Trait Loci Associated with Reproductive-Stage Salt Tolerance in Rice', *Journal of Agronomy and Crop Science*, 201: 17-31.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., ..., Widmayer, P. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Advances in bioinformatics,* vol 2008, Article ID 420747, 5 pages.

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., ..., Li, C. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature genetics,* 42**,** 961-967.

Huang, X., Zhao, Y., Wei, X., Li, C., Wang, A., Zhao, Q., ..., Li, W. (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature genetics,* 44**,** 32-39.

Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M. & Kimchi-Sarfaty, C. (2009) Silent (synonymous) SNPs: should we care about them? *Single Nucleotide Polymorphisms: Methods and Protocols***,** 578, 23-39.

Islam, M. R.; Salam, M. A.; Hassan, L.; Collard, B. C. Y.; Singh, R. K.; Gregorio, G. B. 2011. 'QTL mapping for salinity tolerance at seedling stage in rice', *Emirates Journal of Food & Agriculture (EJFA*, 23: 9.

Jabnoune, M., Espeout, S., Mieulet, D., Fizames, C., Verdeil, J.-L., Conéjéro, G., ..., Rodríguez-Navarro, A. (2009) Diversity in expression patterns and functional properties in the rice HKT transporter family. *Plant physiology,* 150**,** 1955-1971.

Kim, H., Jung, J., Singh, N., Greenberg, A., Doyle, J. J., Tyagi, W., ..., Chung, J. W. (2016) Population Dynamics Among six Major Groups of the *Oryza rufipogon* Species Complex, Wild Relative of Cultivated Asian Rice. *Springer* *Rice,* 9**,** 56.

Korte, A. & Farlow, A. (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant methods,* 9**,** 29.

Koyama, M. L., A. Levesley, R. M. Koebner, T. J. Flowers, and A. R. Yeo. 2001. 'Quantitative trait loci for component physiological traits determining salt tolerance in rice', *Plant Physiol*, 125: 406-22.

Kumar, V., Singh, A., Mithra, S. A., Krishnamurthy, S., Parida, S. K., Jain, S., ..., Tiwari, K. K. (2015) Genome-wide association mapping of salinity tolerance in rice (Oryza sativa). *DNA Research*, 22 (2), 133-145.

Kushwaha, U. K. S. (2016) Rice. Black Rice. Research, History and Development*.* Springer International Publishing, 1-19.

Leng, Q., Mercier, R. W., Hua, B. G., Fromm, H. & Berkowitz, G. A. (2002) Electrophysiological analysis of cloned cyclic nucleotide-gated ion channels. *Plant Physiology,* 128**,** 400-410.

Liang, S., Wu, L., Ren, G., Zhao, X., Zhou, M., Mcneil, D. & Ye, G. (2016) Genome-wide association study of grain yield and related traits using a collection of advanced indica rice breeding lines for irrigated ecosystems. *Field Crops Research*,193, 70-86.

Lin, H. X., M. Z. Zhu, M. Yano, J. P. Gao, Z. W. Liang, W. A. Su, X. H. Hu, Z. H. Ren, and D. Y. Chao. 2004. 'QTLs for Na+ and K+ uptake of the shoots and roots controlling rice salt tolerance', *Theor Appl Genet*, 108: 253-60.

Lin, H., Yanagihara, S., Zhuang, J., Senboku, T., Zheng, K. & Yashima, S. 1997. 'Identification of QTL for salt tolerance in rice via molecular markers', *Zhongguo shuidao kexue*, 12: 6.

Londo, J. P., Chiang, Y. C., Hung, K. H., Chiang, T.Y. & Schaal, B. A. (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences,* 103**,** 9578-9583.

Maathuis, F. J. (2006) The role of monovalent cation transporters in plant responses to salinity. *Journal of Experimental Botany,* 57**,** 1137-1147.

Maathuis, F. J., Ahmad, I. & Patishtan, J. (2014) Regulation of Na+ fluxes in plants. *Frontiers in Plant Science,* 5, 467.

Mccouch, S. R., Wright, M. H., Tung, C.-W., Maron, L. G., Mcnally, K. L., Fitzgerald, M., ..., Singh, N. (2016) Open access resources for genome-wide association mapping in rice. *Nature communications,* 7.

Mohammadi, R., M. S. Mendioro, G. Q. Diaz, G. B. Gregorio, and R. K. Singh. 2013. 'Mapping quantitative trait loci associated with yield and yield components under reproductive stage salinity stress in rice (Oryza sativa L.)', *Journal of Genetics*, 92: 433-43.

Mohanty, A., Kathuria, H., Ferjani, A., Sakamoto, A., Mohanty, P., Murata, N. & Tyagi, A. (2002) Transgenics of an elite indica rice variety Pusa Basmati 1 harbouring the codA gene are highly tolerant to salt stress. *Theoretical and Applied Genetics,* 106**,** 51-57.

Nieves-Cordones, M., Ródenas, R., Chavanieu, A., Rivero, R. M., Martinez, V., Gaillard, I. & Rubio, F. (2016) Uneven HAK/KUP/KT Protein Diversity Among Angiosperms: Species Distribution and Perspectives. *Frontiers in plant science,* 7.

Obata, T., Kitamoto, H. K., Nakamura, A., Fukuda, A. & Tanaka, Y. (2007) Rice shaker potassium channel OsKAT1 confers tolerance to salinity stress on yeast and rice cells. *Plant Physiology,* 144**,** 1978-1985.

Pandit, A., Rai, V., Bal, S., Sinha, S., Kumar, V., Chauhan, M., ..., Gautam, R. (2010) Combining QTL mapping and transcriptome profiling of bulked RILs for identification of functional polymorphism for salt tolerance genes in rice (Oryza sativa L.). *Molecular Genetics and Genomics,* 284**,** 121-136.

Peiter, E., Maathuis, F. J., Mills, L. N., Knight, H., Pelloux, J., Hetherington, A. M. & Sanders, D. (2005) The vacuolar Ca2+-activated channel TPC1 regulates germination and stomatal movement. *Nature*, 434, 404-408.

Platten, J. D., Egdane, J. A. & Ismail, A. M. (2013) Salinity tolerance, Na+ exclusion and allele mining of HKT1;5 in *Oryza sativa* and *O. glaberrima*: many sources, many genes, one mechanism? *BMC plant biology,* 13**,** 32.

Ren, Z. H., Gao, J. P., Li, L. G., Cai, X. L., Huang, W., Chao, D. Y., ..., Zhu, M. Z. (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nature genetics,* 37**,** 1141-1146.

Rosas-Santiago, P., Lagunas-Gómez, D., Barkla, B. J., Vera-Estrella, R., Lalonde, S., Jones, A., ..., Frommer, W. B. (2015) Identification of rice cornichon as a possible cargo receptor for the Golgi-localized sodium transporter OsHKT1;3. *Journal of experimental botany,* 66**,** 2733-2748.

Rubio, F., Schwarz, M., Gassmann, W. & Schroeder, J. I. (1999) Genetic selection of mutations in the high affinity K+ transporter HKT1 that define functions of a loop site for reduced Na+ permeability and increased Na+ tolerance. *Journal of Biological Chemistry,* 274**,** 6839-6847.

Sabouri, H., and A. Sabouri. 2008. 'New evidence of QTLs attributed to salinity tolerance in rice', *African Journal of Biotechnology*, 7: 4376-83.

Sas Institute, I. (2011) SAS 9.3. Software. SAS Institute, Inc. Cary, NC.

Senadheera, P., Singh, R. & Maathuis, F. J. (2009) Differentially expressed membrane transporters in rice roots may contribute to cultivar dependent salt tolerance. *Journal of experimental botany,* 60**,** 2553-2563.

Si, L., Chen, J., Huang, X., Gong, H., Luo, J., Hou, Q., ..., Zhou, T. (2016) OsSPL13 controls grain size in cultivated rice. *Nature genetics,* 48**,** 447-456.

Tester, M. & Davenport, R. (2003) Na+ tolerance and Na+ transport in higher plants. *Annals of botany,* 91**,** 503-527.

Tiwari, Sushma, Krishnamurthy Sl, Vinod Kumar, Balwant Singh, A. R. Rao, Amitha Mithra Sv, Vandna Rai, Ashok K. Singh, and Nagendra K. Singh. 2016. 'Mapping QTLs for Salt Tolerance in Rice (Oryza sativa L.) by Bulked Segregant Analysis of Recombinant Inbred Lines Using 50K SNP Chip', *PLOS ONE*, 11: e0153610.

Ueda, Y., Frimpong, F., Qi, Y., Matthus, E., Wu, L., Höller, S., ..., Kraska, T. (2015) Genetic dissection of ozone tolerance in rice (Oryza sativa L.) by a genome-wide association study. *Journal of experimental botany,* 66**,** 293-306.

Véry, A. A., Nieves-Cordones, M., Daly, M., Khan, I., Fizames, C. & Sentenac, H. (2014) Molecular biology of K+ transport across the plant cell membrane: What do we learn from comparison between plant species? *Journal of Plant Physiology,* 171**,** 748-769.

Vij, S., Giri, J., Dansana, P. K., Kapoor, S. & Tyagi, A. K. (2008) The receptor-like cytoplasmic kinase (OsRLCK) gene family in rice: organization, phylogenetic relationship, and expression during development and stress. *Molecular Plant,* 1**,** 732-750.

Wang, H., Xu, X., Vieira, F. G., Xiao, Y., Li, Z., Wang, J., ..., Nielsen, R (2016) The power of inbreeding: NGS based GWAS of rice reveals convergent evolution during rice domestication. *Molecular Plant, 9, 975,895*.

Wang, R., Jing, W., Xiao, L., Jin, Y., Shen, L. & Zhang, W. (2015) The Rice High-Affinity Potassium Transporter1;1 Is Involved in Salt Tolerance and Regulated by an MYB-Type Transcription Factor. *Plant Physiology,* 168**,** 1076-1090.

Yeo, A., Izard, P., Boursier, P. & Flowers, T. (1991) Short-and long-term effects of salinity on leaf growth in rice (Oryza sativa L.). *Journal of Experimental Botany,* 42**,** 881-889.

Yoshida, S., Forno, D., Cock, J. & Gomez, K. (1976) Laboratory manual for physiological studies of rice. Manila (Philippines). International Rice Research Institute.

Yuan, F., Yang, H., Xue, Y., Kong, D., Ye, R., Li, C., ..., Zhang, J. (2014) OSCA1 mediates osmotic-stress-evoked Ca2+ increases vital for osmosensing in Arabidopsis. *Nature*, 514, 367-371.

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., ..., Norton, G. J. (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nature communications,* 2**,** 467.

Zhao, X. Q., Zhao, S. P. & Shi, W. M. (2008) Enhancement of NH4+ Uptake by NO3- in Relation to Expression of Nitrate-Induced Genes in Rice (Oryza sativa) Roots. *Pedosphere,* 18**,** 86-91.

Zhu, D., Kang, H., Li, Z., Liu, M., Zhu, X., Wang, Y., ..., Wang, D. (2016) A Genome-Wide Association Study of Field Resistance to Magnaporthe Oryzae in Rice. *Rice,* 9**,** 44.

**FIGURE LEGENDS**

**Figure 1**. RGR of 306 rice accessions. **(A)** Plants exposed to control hydroponic medium for 30 days. **(B)** Plants exposed to 50 mM NaCl for 7 days. **(C)** As (B) but exposure for 30 days. Graphs show the means ±SE of three biological replicates.

**Figure 2.** RGR-reduction (treatment/control) of 306 rice accessions. **(A)** Plants exposed to 50 mM NaCl for 7 days. **(B)** As (A) but exposure for 30 days. Graphs show the means ±SE of three biological replicates.

**Figure 3.** Pot grown accessions that were selected on the basis of tolerance and sensitivity during hydroponics experiments largely retain their phenotype. Tolerance is expressed as a percentage reduction in shoot growth when grown in salinised pots watered with 50 mM NaCl) compared to controls (water only). Data show average of three plants ± SE.

**Figure 4.** Correlation between tissue ion levels and growth. **(A)** Correlation between root K+ and RGR. **(B)** Correlation between shoot Na+ and RGR. Tissue cation and growth data were from long term trial. \*\*\*\*Correlation is significant at the p<0.0001 level.

**Figure 5.** Manhattan plot showing association signal. GWAS analysis was carried out using shoot Na+ of the 306 accessions as phenotypic data. Plants were exposed to medium term salt stress (50 mM NaCl). X-axis indicates the SNP positions across the 12 rice chromosomes. The y-axis shows the *P*-value for association test at each locus on a log scale. The horizontal dashed line represents the significance threshold.

**Figure 6.** Ontology analysis of candidate genes. **(A)** Several functional categories were over represented especially those related to lipid binding, nuclease/hydrolase activity and transport. **(B)** Cellular compartments such as the membrane and cell wall showed a disproportionally high number of candidate genes.

**Figure 7.** Correlation between GWAS and QTL studies. The adapted QTL cartographer shows the 12 rice chromosomes and indicates where the main association signals from this study are localised (red arrowheads on the left hand side). On the other side of each chromosome are major QTLs identified from multiple studies (see legend) where rice was exposed to saline conditions.

**Table 1.** Summary of K+ and Na+ concentrations of accessions exposed to control and saline conditions. The table summarises the range, mean values and standard error for K+ and Na+ concentrations in root and shoot tissues of the 306 accessions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Phenotype | Control | Short term | Medium term | Long term |
| Root K+ | 240 - 888; 440 ± 5.7 | 346 - 793; 560 ± 4.6 | 249 - 1170; 522 ± 10.5 | 73 - 593; 314 ± 5.7 |
| Shoot K+ | 289 - 1010; 529 ± 9.1 | 527 - 927; 718 ± 3.6 | 250 - 1093; 657 ± 8.9 | 203 - 1208; 586 ± 13.9 |
| Root Na+ | - | 149 - 546; 334 ± 3.4 | 235 - 894; 502 ± 8.5 | 244 - 720; 457 ± 4.3 |
| Shoot Na+ | - | 13 - 53; 25 ± 0.3 | 10 - 907; 312 ± 11.1 | 62 - 1939; 692 ± 20.8 |

**Table 2:** Curated list of candidate genes from short, medium and long term NaCl treatment

**================================================================**

**Locus Gene**

**=================================================================**

**Abiotic stress related**

LOC\_Os10g42820 early-responsive to dehydration protein-related, putative, expressed

LOC\_Os04g48030 heat stress transcription factor B-1, putative

LOC\_Os03g06630 heat stress transcription factor, putative

LOC\_Os03g06360 late embryogenesis abundant protein D-34, putative, expressed

LOC\_Os02g42210 membrane-associated salt-inducible protein like, putative

**Anion transport**

LOC\_Os02g48880 CLCe/f, chloride transporter, chloride channel family

LOC\_Os04g47930 aluminium-activated malate transporter

LOC\_Os03g06520 SULTR3;2, sulphate transporter

**Kinases**

LOC\_Os11g35274 protein kinase domain containing protein, expressed

LOC\_Os11g44560 RLCK348, protein kinase domain containing protein

LOC\_Os04g52590 protein kinase domain containing protein, expressed

LOC\_Os02g02780 protein kinase family protein

LOC\_Os02g02040 protein kinase, putative

LOC\_Os06g03970 receptor-like protein kinase 5 precursor, putative, expressed

LOC\_Os08g25380 serine/threonine-protein kinase BRI1-like 1 precursor

LOC\_Os02g42110 wall-associated receptor kinase-like 22 precursor

LOC\_Os02g42150 OsWAK14 - OsWAK receptor-like protein kinase, expressed

**Phosphatases**

LOC\_Os12g05540 Ser/Thr protein phosphatase family protein

LOC\_Os02g48840 dual specificity protein phosphatase, putative, expressed

**ABC transporters**

LOC\_Os06g06440 ABCC15, ABC transporter, ATP-binding protein

LOC\_Os08g43120 ABCG45, PDR ABC transporter

**Ca2+ binding and signalling**

LOC\_Os01g11414 CAX1, proton/calcium exchanger protein, putative, expressed

LOC\_Os01g37690 CAX1a, sodium/calcium exchanger protein

LOC\_Os09g38580 CNGC9, cyclic nucleotide-gated ion channel, putative, expressed

LOC\_Os02g02540 glutamate receptor, putative, expressed

LOC\_Os06g06130 glutamate receptor, putative, expressed

LOC\_Os01g48680 TPC1, two pore calcium channel protein 1, putative, expressed

LOC\_Os10g42820 OSCA1;4, hyperosmolality-gated calcium-permeable channel

LOC\_Os11g44680 calmodulin binding protein, putative, expressed

LOC\_Os12g05420 calmodulin-binding protein, putative, expressed

LOC\_Os04g48160 IQ calmodulin-binding motif family protein

LOC\_Os02g08140 CAMK\_KIN1/SNF1/Nim1, calcium/calmodulin dependent kinase

LOC\_Os11g02240 CIPK15, CAMK\_KIN1/SNF1/Nim1 calcium/calmodulin dependent protein kinases, expressed

**Peptide/amino acid transport**

LOC\_Os02g49060 AAP10, amino acid transporter, putative, expressed

LOC\_Os01g66010 AAP8, amino acid transporter

LOC\_Os01g11160 CAT1, amino acid permease family protein

LOC\_Os01g37590 PTR2, peptide transporter PTR2, putative, expressed

**Sugar transport**

LOC\_Os10g21590 Sugar transporter family protein, putative, expressed

LOC\_Os06g30950 Sugar transporter-related, putative

LOC\_Os10g41190 Monosaccharide transporter family protein

LOC\_Os10g42830 Vacuolar hexose transporter

**Transcription factors**

LOC\_Os08g38210 transcription factor BIM2, putative, expressed

LOC\_Os08g43160 TCP family transcription factor, putative, expressed

LOC\_Os01g11350 bZIP transcription factor domain containing protein, expressed

LOC\_Os08g26880 bZIP transcription factor domain containing protein, expressed

LOC\_Os01g19330 MYB family transcription factor, putative, expressed

LOC\_Os10g41200 MYB family transcription factor, putative, expressed

LOC\_Os10g41260 MYB family transcription factor, putative, expressed

LOC\_Os02g40530 MYB family transcription factor, putative, expressed

LOC\_Os01g62410 MYB family transcription factor, putative, expressed

LOC\_Os06g06360 WRKY113, expressed

LOC\_Os11g02530 WRKY40, expressed

LOC\_Os11g02480 WRKY46, expressed

LOC\_Os05g39720 WRKY70, expressed

**Other transporter**

LOC\_Os04g47220 PIP1;2, aquaporin protein, putative, expressed

**Phospholipases**

LOC\_Os04g35100 phospholipase C, putative, expressed

**Cation transport**

LOC\_Os04g52310 ZIP3, metal cation transporter

LOC\_Os05g39560 ZIP5, metal cation transporter

LOC\_Os05g39540 ZIP9, metal cation transporter

LOC\_Os03g53780 ammonium transporter 2

LOC\_Os02g40710 AMT1;2, ammonium transporter protein, putative, expressed

LOC\_Os02g40730 AMT1;3, ammonium transporter protein, putative, expressed

LOC\_Os01g11250 potassium channel KAT1

LOC\_Os05g35410 potassium channel AKT2/3, putative, expressed

LOC\_Os07g48130 HAK9, potassium transporter

LOC\_Os04g52390 HAK11, potassium transporter, putative, expressed

LOC\_Os02g49760 HAK25, potassium transporter, putative, expressed

LOC\_Os02g07830 OsHKT1;3 - Na+ transporter

LOC\_Os05g39600 ATCHX15, putative, expressed

LOC\_Os02g08018 calcium-transporting ATPase 9

LOC\_Os01g11414 CAX1, proton/calcium exchanger protein, putative, expressed

LOC\_Os01g37690 CAX1a, sodium/calcium exchanger protein

LOC\_Os11g43860 sodium/calcium exchanger protein

LOC\_Os02g02540 glutamate receptor, putative, expressed

LOC\_Os06g06130 glutamate receptor, putative, expressed

LOC\_Os09g38580 CNGC9, cyclic nucleotide-gated ion channel, putative, expressed

LOC\_Os01g48680 TPC1, two pore calcium channel protein 1, putative, expressed

LOC\_Os10g42820 OSCA1;4, hyperosmolality-gated calcium-permeable channel













