Plant Biotechnology Journal

Plant Biotechnology Journal (2013) **11**, pp. 87–100 doi: 10.1111/pbi.12010

New allelic variants found in key rice salt-tolerance genes: an association study

Sónia Negrão[1,2], M. Cecília Almadanim[1,2], Inês S. Pires[1,2], Isabel A. Abreu[1,2], João Maroco[3], Brigitte Courtois[4], Glenn B. Gregorio[5], Kenneth L. McNally[5] and M. Margarida Oliveira[1,2,*]

[1]*Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal*

[2]*iBET, Oeiras, Portugal*

[3]*ISPA (Instituto Superior de Psicologia Aplicada), Grupo de Estatística e Matemática, Lisboa, Portugal*

[4]*CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Dévelopment), UMR AGAP, Montpellier, France*

[5]*IRRI (International Rice Research Institute), Metro Manila, The Philippines*

Received 18 July 2010;
revised 17 September 2012;
accepted 19 September 2012.
**Correspondence* (fax +351 214421161; email mmolive@itqb.unl.pt)

Keywords: *Oryza sativa*, salt-stress, association analysis, OsRMC, OsNHX1, OsCPK17, OsHKT1;5 (SKC1) (HKT8), salT.

Summary

Salt stress is a complex physiological trait affecting plants by limiting growth and productivity. Rice, one of the most important food crops, is rated as salt-sensitive. High-throughput screening methods are required to exploit novel sources of genetic variation in rice and further improve salinity tolerance in breeding programmes. To search for genotypic differences related to salt stress, we genotyped 392 rice accessions by EcoTILLING. We targeted five key salt-related genes involved in mechanisms such as Na^+/K^+ ratio equilibrium, signalling cascade and stress protection, and we found 40 new allelic variants in coding sequences. By performing association analyses using both general and mixed linear models, we identified 11 significant SNPs related to salinity. We further evaluated the putative consequences of these SNPs at the protein level using bioinformatic tools. Amongst the five nonsynonymous SNPs significantly associated with salt-stress traits, we found a T67K mutation that may cause the destabilization of one transmembrane domain in OsHKT1;5, and a P140A alteration that significantly increases the probability of OsHKT1;5 phosphorylation. The K24E mutation can putatively affect SalT interaction with other proteins thus impacting its function. Our results have uncovered allelic variants affecting salinity tolerance that may be important in breeding.

Introduction

Rice, *Oryza sativa* L., is one of the most important crop species and the major food crop for much of the world's population (Molina *et al.*, 2011). It has become clear in the past few years that abiotic stresses will be increasingly important due to climate change, land degradation and declining water quality (Wassmann *et al.*, 2009). Soil salinity is a major abiotic stress in plant agriculture worldwide (Zhu, 2001). To improve stress tolerance, breeders depend to a great extent on the ability to make use of existing or novel sources of genetic variation (McCouch and Kovach, 2008).

Amongst cereals, rice is the most sensitive to salt stress. Large genetic variability has been reported in salinity tolerance amongst rice varieties (reviewed by Negrão *et al.*, 2011). Several breeding programmes throughout the world have used a number of *indica* types as salt-tolerant donors. Amongst them, 'Pokkali', 'Cheriviruppu' and 'Nona Bokra' are the most famous direct or remote parents, by carrying diverse response mechanisms to salinity (Zeng, 2005). However, the distribution of salinity tolerance amongst different *O. sativa* varietal groups is still not well documented. Because the most common forms of genetic variation within natural populations are single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels), a further step can be achieved by investigating such variation in key-response genes from diverse germplasm sources (Raghavan *et al.*, 2007). A clear example of the importance of genotypic differences between varieties in a salt-stress-related gene is given by the allelic differences found in *OsHKT*1;5 gene. The six nucleotide substitutions in the coding region leading to four amino acid changes present in 'Nona Bokra' compared to 'Koshihikari' enhanced the overall Na^+ transport activity (Ren *et al.*, 2005), showing the importance of discovering superior alleles in salt-stress-related genes.

In rice, several projects aim to provide research community with extensive information on the genetic variation present in diverse germplasm sources. It is the case of OryzaSNP consortium (http://www.OryzaSNP.org) that analysed 100Mbp of nonrepeated sequences in 20 cultivated rice accessions (McNally *et al.*, 2009). Another research platform used a diversity panel of 400 *O. sativa* and 100 *O. rufipogon* accessions, to link sequence diversity with physiological, morphological and agronomic variations (Tung *et al.*, 2010; Zhao *et al.*, 2011). It is hoped that an outcome of these projects will be the easier identification and tracking of new allelic variations and the conversion of that information into cost-effective tools for applied plant improvement. In allele mining, recent developments such as 'next-generation sequencing' have made possible whole-genome and gene-targeted surveys by covering a number of different varieties. This exciting new age of genomics is providing the basis to discover superior alleles through 'mining' the gene(s) of interest, from diverse genetic resources. However, many research programmes have limited funding and need more cost-effective techniques.

© 2012 The Authors
Plant Biotechnology Journal © 2012 Society for Experimental Biology, Association of Applied Biologists and Blackwell Publishing Ltd

Previously, we reviewed the recent updates on salinity stress in rice summarizing some of the important and/or interesting candidate genes recognized as involved in tolerance mechanisms (Negrão *et al.*, 2011). In the present study, we selected five key genes involved in several aspects of salinity. All of these genes have been previously described and characterized as related to salt-tolerance enhancement in rice, through different mechanisms.

Signal perception and transduction pathways are the first components of plant adaptive response to stress, and therefore, we decided to analyse the allelic variants of *OsCPK*17 and *OsRMC*. Under stress, a transient increase in cytosolic Ca^{2+} is the first step in the response cascades, and several groups of Ca^{2+}-binding proteins play an important role in signal transduction. Amongst these proteins, OsCPK17 seems to be down-regulated by cold, drought and salt in certain stress tolerant cultivars (Lin *et al.*, 2007). Additionally, when listing QTLs linked to salt-stress responses and co-localizing them with candidate genes, we noticed that *OsCPK*17 occurs within 4 QTLs linked to ion homoeostasis (Negrão *et al.*, 2011). We also included another salt-related gene, *OsRMC*, which is possibly involved in signalling. OsRMC is a jasmonic acid–induced DUF26 protein involved in rice root development (Chong *et al.*, 2007) and is up-regulated at the transcript level by high salinity (Rabbani *et al.*, 2003).

Two other key genes (*OsHKT*1;5 and *OsNHX*1) involved in ion homoeostasis were included, due to their importance under salinity both by osmotic adjustment and by keeping Na^+ away from the cytosol (Zhang and Blumwald, 2001). *OsHKT*1;5 encodes a member of the HKT-type transporters that is preferentially expressed in root parenchyma cells surrounding xylem vessels (Ren *et al.*, 2005). OsHKT1;5 physiological activity is responsible for retrieving Na^+ from the xylem sap, resulting in less Na^+ load in shoots, and hence maintaining low Na^+/K^+ shoot ratios in rice plants submitted to salt stress (Mian *et al.*, 2011; Negrão *et al.*, 2011). *OsNHX*1 encodes a vacuolar $(Na^+, K^+)/H^+$ antiporter localized in the tonoplast, allowing efficient compartmentation of Na^+ in the vacuole (Fukuda *et al.*, 2004).

Finally, we also included *Sal*T (reviewed by Negrão *et al.*, 2011) that co-localizes with SalTol QTL (chromosome 1) and was first isolated and characterized from the roots of salt-treated rice plants. Evidence supports its regulation by ABA-dependent and ABA-independent pathways and also suggests correlation of *Sal*T expression with production of osmoprotectants, such as trehalose and proline.

Although previous studies have indicated the importance of allelic variants in stress response and their crucial role in breeding programmes, an integrative approach that correlates newly found genetic diversity with tolerance phenotypes is essential. In this research work, we analysed by EcoTILLING (Comai *et al.*, 2004) a rice population composed of 392 accessions belonging to the five variety groups *indica*, *temperate* and *tropical japonica*, *aromatic* and *aus*, selected to be representative of rice genetic diversity. Of these accessions, 373 spanned the range of diversity of Generation Challenge Programme composite rice collection of 2339 accessions genotyped by 45 simple sequence repeat (SSR) markers and analysed for diversity (K.L. McNally, R. Manzano, M. Macatangay, M. Redondo, V. Lacorte, M. Zaidem, J. Detras, M. Barile, S. Mercado-Quilloy, M.E.B. Naredo L. Benoit, R. Rivallan, B. Courtois, C. Billot, A. Garavito, M. Lorieux, C.P. Martinez, T. Borba, R.V. Brondani, C. Brondani, M. Cissoko, M.-N. Ndjiondjop, A. Famoso, S.R. McCouch, N.R. Sackville Hamilton, and J.-C. Glaszmann, unpublished). The other 19 included OryzaSNP entries (McNally *et al.*, 2009) and salt-tolerant and sensitive types not in the composite collection. Our work aims to find new allelic variants for five genes involved in salinity tolerance (*OsCPK*17, *OsRMC*, *OsNHX*1, *OsHKT*1;5 and *Sal*T) and to associate the haplotype variation with salt-tolerance phenotypes. We also addressed the putative consequences of SNP polymorphisms at protein level using bioinformatics tools and discussed their presumed functional influence.

Results

Allelic variants found in the rice working set

We genotyped a working set composed of 392 rice accessions representative of the diversity present in *O. sativa.* Haplotypes were assigned by comparing the CJE digestion patterns (Figure 1). For most of the genes, we found some areas difficult to amplify, especially at intronic regions. Therefore, only some regions could be successfully amplified and digested, as shown in Table 1.

In a following step, we sequenced the representative haplotypes. Using the sequence information of 'Nipponbare' as reference (Matsumoto *et al.*, 2005), we discovered several new allelic variants listed in Table 2. We present in Table S2 all the haplotypes, and in Table S3, all SNPs/indels found in the target genes in UTRs, introns and cds. Our work focused only on the target genes's cds. Nonsynonymous SNPs, with the corresponding amino acid changes, are indicated in Table 3.

By analysing the coding sequence allelic diversity, we observed that *OsHKT*1;5 had the highest genetic variability with 29 SNPs and two indels, corresponding to 15 haplotypes and 14 different proteins, whilst *OsRMC* had the lowest variability with only four haplotypes and three different proteins (99.2% of which belonging to the same haplotype group). Regarding *Sal*T, we identified 13 SNPs and no indels in the cds, corresponding to six different haplotypes and six different proteins. In the *OsNHX*1

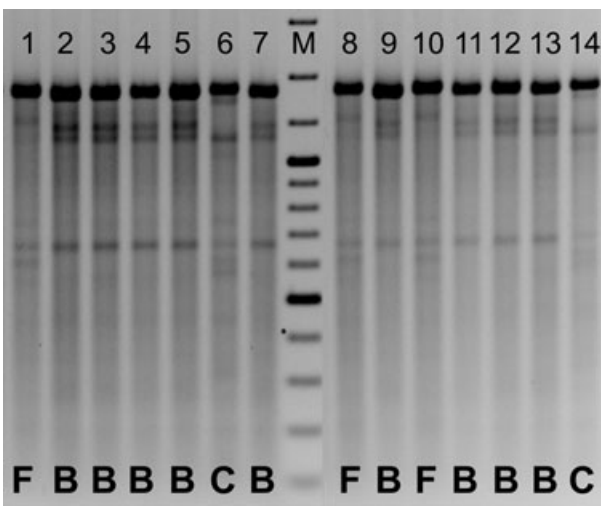

Figure 1 Analysis of the EcoTILLING digestion patterns in agarose gel for *OsHKT*1;5 (P5) in different rice varieties. The varieties shown were contrasted with 'IR64' and generated three haplotype groups (B, C and F) according to CJE-cleaved products. 1- PELITA JANGGUT; 2- POPOT; 3- KINANDANG PUTI; 4- DA1; 5- LAL AMAN; 6- PATNAI 23; 7- IRGC 31293; 8- KHAO PON; 9- MALLIGAI(KOTTAMALLI SAMBA); 10- TUNGHWANPEI; 11- DE ABRIL; 12- LAGEADO; 13- BALA; 14- SINNA SITHIRA KALI.

© 2012 The Authors

Table 1 Distribution frequency of successful digestion in EcoTILLING of each fragment efficiently amplified with the different primer pairs (see Table 5 for correlation). Percentage of successful digestion refers to the percentage of accessions digested within the 392 genotype panel

Target gene/locus	Primer ID	Start position*	End position*	Successful digestion (%)
*OsCPK*17	P1	74	1107	5.10
(*Os07g06740*)	P3	1437	2530	100
	P4	2348	3470	100
	P6	3924	4902	98.21
OsRMC	P1	1537	2758	100
(*Os04g56430*)				
*OsNHX*1	P1	−258	866	98.21
(*Os07g47100*)	P3	1727	2597	98.98
	P4	2493	3349	98.47
	P5	3197	4264	94.90
	P6	4165	5239	3.32
OsHKT1;5	P1	3803	4441	94.64
(*Os01g20160*)	P5	436	1813	97.19
	P6	−439 (−414)*	626	7.65
*Sal*T (*Os01g24710*)	P1	130	1193	96.17
	P2	1011	1812	65.56

*In reference to the respective *locus* (in Nipponbare sequence).

cds, we identified 10 SNPs and no indels, corresponding to six different haplotypes and four different proteins. In the coding region of *OsCPK*17, we found 14 SNPs and one indel, which corresponded to nine haplotypes and five different proteins. In general, accessions with the same haplotype belonged to the same varietal group. For instance, in *Sal*T, we observed that approximately 82% of the *japonica* accessions were haplotype A, whilst 72% of *indica* accessions were haplotype B.

In the case of *OsHKT*1;5, we identified 13 new haplotypes besides the two previously described by Ren *et al.* (2005) as 'NonaBokra' and 'Koshihikari' alleles. Interestingly, 'Pokkali', a well-known donor accession for salt tolerance, also displays the 'NonaBokra' haplotype (haplotype E). However, other salt-tolerant varieties such as 'FL478', 'IR52724-2B-6-2B-1-1' or 'Hasawi' carried different *OsHKT*1;5 haplotypes. We could not find distinct haplotypes shared by most tolerant accessions for any of the target genes.

Haplotype networks based on all the SNPs found in the coding regions for each gene were constructed to infer the relationships amongst the haplotypes (see Figure 2 and Figure S1). In Figure 2, two distinct examples of haplotype networks are shown for *OsRMC* (with a dominant haplotype and a few rare ones and low diversity) and *OsHKT*1;5 (with 15 haplotypes of similar frequency).

Association analysis between SNP/haplotypes and salt-stress traits

We phenotypically characterized 59 accessions, representative of all the allelic variants found at the coding sequence for each gene. Our results (Table S4) show a diverse response towards salt stress ranging from sensitive to tolerant phenotypes. These results include the characterization of salt-tolerant genotypes 'Hasawi', 'Kun Min Tsieh Hunan', 'Davao' and 'ARC 7229'. Generally, we observed that tolerant genotypes have a significantly lower shoot Na^+ concentration compared to sensitive genotypes, which is in accordance with previous results (Walia *et al.*, 2005; Senadheera *et al.*, 2009; Cotsaftis *et al.*, 2010). Additionally, our results support that tolerant genotypes are able to maintain the osmotic potential under salt stress. Moreover, the osmotic potential dropped significantly under salt stress, as previously reported (Siringam *et al.*, 2012).

For population structure, PCA analysis was performed on the 50 accessions that covered the identified haplotypes, which were phenotyped for salt tolerance. The coordinates in the first 10 PCA axes captured 68% of the total population variability (which included the five variety groups) and were kept for further analyses. When observing the first two axes of PCA in Figure 3, we can see two major clusters, corresponding to *indica* and *japonica* (temperate and tropical) genetic groups and two minor clusters corresponding to *aus* and *aromatic* type, indicating that haplotype variation at the five target genes was dispersed across all variety groups in our working set.

For association studies, we first had to eliminate rare alleles (frequency <5%) because of the strong unbalance amongst the genotypic classes. The final number of alleles in the cds of the target genes thus reduced from 71 to 32 SNPs, and the number of haplotypes reduced from 40 to 19. After removing rare alleles, *OsRMC* became monomorphic and was thus eliminated from further testing. Amongst the 392 accessions, several rare variants were observed such as haplotype F in OsCPK17 (present in only one accession) that has a deletion of three AA residues (EDV) compared to all other haplotypes (Table 2). This particular change can interfere with protein conformation and/or activity (Table 3), but was not evaluated due to the exclusion of rare variants.

We only performed association analysis for polymorphisms detected in the coding portions. We found significant associations between both nonsynonymous (amino acid residue alteration) and synonymous SNPs.

We found significant associations between SNP polymorphism or haplotypes and the studied traits with three tested statistical methods, GLM_Q, MLM_K and MLM_Q+K (Table 4). The significant associations detected with GLM_Q proved to be less stringent, because this model detected the highest number of significant associations. In fact, this was already expected, because this method does not account for all possible background effects putatively leading to false positives. For *OsCPK*17 gene, all three tested models supported associations with biomass traits analysed under salt stress (shoot fresh weight and root fresh and dry weights), thus increasing the confidence in the results obtained for *OsCPK*17.

A total of 11 SNPs of 32 from four candidate genes (*OsHKT*1;5, *Sal*T; *OsCPK*17 and *OsNHX*1) showed significant associations with one or more salt-related traits. From all the phenotypic traits considered, only shoot dry weight, root length, chlorophyll a/b ratio, K^+ content in shoots and salt injury were not significantly associated with any SNP or haplotype. We found several salt-tolerant accessions that did not have specific gene haplotypes with significant associations.

In the following paragraphs, we discuss in detail the influence of the AA residue changes in the overall protein conformation and/or activity. Nevertheless, we can estimate the putative influence of a particular AA residue in a particular trait, by its increased or reduced effect on the trait value. In SalT, the existence of the K24 residue (instead of E24) is estimated to increase by 0.57 mM the Na^+ concentration in shoots of salt-stress plants. Thus, we suppose that E24 residue presence significantly

© 2012 The Authors

Table 2 Combined information about the total number of SNPs and indels identified in the target genes. Data include the polymorphisms found within the coding sequence, intronic and UTR regions. Allelic information found at the coding sequence level and more thoroughly investigated by association analysis is highlighted in bold. The allelic frequency for each haplotype found in the coding sequence is also presented

Gene	Length (bp)	No. SNPs in cds (No. SNPs in the *locus*)	No indels	No haplotypes	Haplotype	Allelic frequency (%)
*OsHKT*1;5	4487	**29** (57)	**2** (9)	**15** (17)	A	11.67
					B	1.33
					C	22.00
					D	5.00
					E	6.67
					F	2.33
					G	25.67
					H	14.33
					I	0.33
					J	0.67
					K	7.67
					L	0.33
					M	0.67
					N	1.00
					O	0.33
*OsNHX*1	4923	**10** (54)	**0** (13)	**6** (13)	A	92.90
					B	6.01
					C	0.27
					D	0.27
					E	0.27
					F	0.27
SalT	1576	**13** (109)	**0** (18)	**6** (13)	A	30.77
					B	49.34
					C	7.43
					D	0.80
					E	0.53
					F	11.14
OsRMC	1070	**3** (3)	**0** (0)	**4** (4)	A	99.23
					B	0.26
					C	0.26
					D	0.26
*OsCPK*17	4814	**14** (38)	**1** (6)	**9** (13)	A	45.15
					B	48.72
					C	1.02
					D	0.26
					E	3.83
					F	0.26
					G	0.26
					H	0.26
					I	0.26

increases stress tolerance by leading to a reduction in Na^+ concentration in shoots. In addition, we estimate that for R184H mutation, the presence of R184 will correspond to a 0.08 mM increase in K^+ concentration in the roots of salt-stressed plants.

Nonsynonymous SNPs – Changes at protein level

Our results include SNPs and small indels in the five target genes not only at cds but also in the 3′ and 5′ UTR and in introns (whenever possible). The complete information is available in Table S3. As previously mentioned, we focused our association analysis at the coding region level and thus decided to further explore the amino acid residues changes (Table 3) and their putative influence at the protein structural level. The results obtained are merely indicative of the potential effects of each SNP, but constitute a good starting hypothesis to further study predicted effects.

OsHKT1;5

In our association studies, we found three nonsynonymous (with altered amino acid) SNPs with significance (Table 4). We started by analysing the position of these three nonsynonymous SNPs within the putative structural domains of the OsHKT1;5 protein using InterProScan. The InterPro database integrates several domain-identification search engines and allows creation of a

© 2012 The Authors

Table 3 List of nonsynonymous amino acid changes originated by identified SNPs in all target genes. Amino acid numbering refers to reference haplotype (Nipponbare accession). Full information data about the intronic and coding region as well as UTR's are available in Table S3

	OsHKT1;5	OsNHX1	SalT	OsRMC	OsCPK17
Coding sequence	S2N	M3L	P8L	T33A	G40D
	F66L	I353L	K24E	L240F	Y41C
	T67K	S477N	G28S		Deletion of E149DV
	Deletion of S119		S72T		K317Q
	A123P		Q74H		T472N
	D128N		E114D		
	S134N				
	P140A				
	P144R				
	V162I				
	R184H				
	S197P				
	Insertion of T206				
	L231V				
	A233V				
	H332D				
	S338P				
	V349M				
	G364S				
	L395V				
	P508S				
	L531F				

unique, nonredundant characterization of a given protein family, domain or functional site (Zdobnov and Apweiler, 2001). This analysis identified a cation transport protein (TrkH)-type domain (previously described in active sodium pumps) that spans approximately two-third of the protein in its C-terminus. Furthermore, we have identified a TMHMM motif, consisting of nine transmembrane regions, also consistent with OsHKT1;5 function as a membrane Na^+ transporter. Additionally, a putative signal peptide was predicted for the protein N-terminal, but it could not be confirmed by other specific bioinformatic tools. From this analysis, we could also conclude that the identified SNPs are all at the N-terminal part of the protein, at outside the TrkH-domain and thus may be within a protein region with regulatory functions.

Additionally to analysing the consequences of these SNPs in the predicted secondary structure of the protein, we also analysed their influence in putative post-translational modifications (PTM), such as phosphorylation, ubiquitination and SUMOylation. Haplotype F is unique in carrying the T67K mutation. This alteration happens within the second predicted transmembrane region (within the TMHMM domain) of the protein and is expected to have a significant structural effect. Indeed, the analysis of haplotype F, using InterProScan, no longer predicts the existence of the second transmembrane helix, due to the destabilizing effect of T67K (Figure 4). Loss of the threonine phosphorylation or gain of the lysine SUMOylation or ubiquitination are unlikely in this case, because both residues are poor targets for these modifications, according to NetPhos 2.0, SUMOsp and UbPred

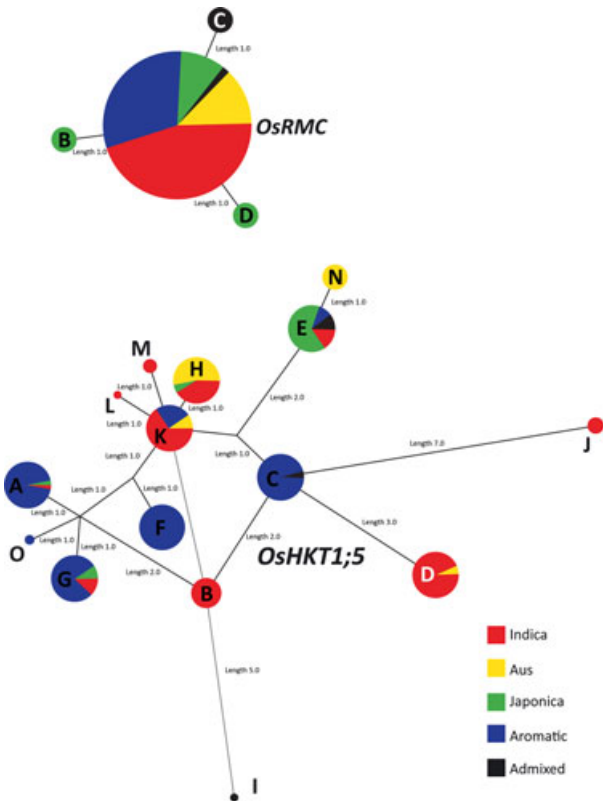

Figure 2 Distribution of the observed haplotypes in genes *OsRMC* and *OsHKT*1;5. The size of the circles is proportional to haplotype frequency, and the length of the connecting lines is proportional to the number of mutation steps between genotypes and their proximal states. Pie charts indicate the contribution of the rice varietal groups to a particular haplotype (Table S1).

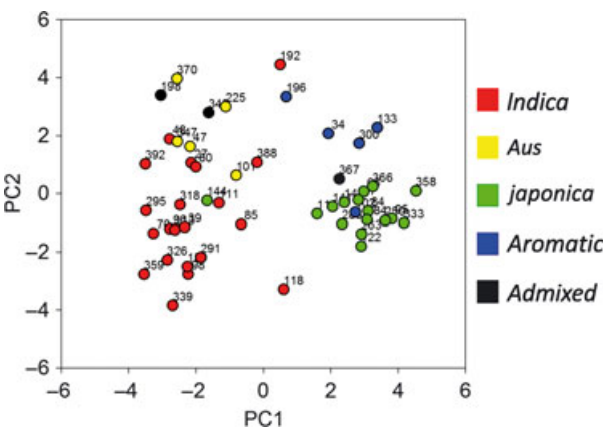

Figure 3 Projection of the 50 accessions on the plane of the two-first axes of the Principal Component Analysis indicative of the population structure existing in our rice working set.

predictions, respectively (Blom *et al.*, 1999; Ren *et al.*, 2009; Radivojac *et al.*, 2010). The P140A alteration that occurs in haplotypes B, C, D, E and N relative to the reference haplotype A does not seem to have a significant impact on the protein secondary structure, because it occurs at a region predicted to be mostly random coil. Instead, this modification occurs within a

© 2012 The Authors

Table 4 Significant associations ($\alpha < 0.01$) identified between traits related to salt stress and SNP positions/haplotypes, targeting salt-related genes

Target gene	SNP position	AA residue change	Influence at structural protein level	Shoot Length		Shoot Fresh Weight		Shoot Dry Weight		Root Length		Root Fresh Weight		Root Dry Weight		Chlorophyll A		Chlorophyll B		Chlorophyll Total		Chlorophyll A/B		Na^+ Shoot Content		K^+ Shoot Content		Na_+/K_+ Shoot Content		Na^+ Root Content		K^+ Root Content		Na^+/K^+ Root Content		Na^+/K^+ Shoot/Root Contert		Osmotic Potential		Salt Injury	
				NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S	NS	S
HKT1;5	367	T67K	Destabilization of one TM region																					c				c										c		b			
	585	P140A	Increases the phosphoryl. probability																																						a		
	718	R184H	Potential effect on transporter regulation																															b									
	1286	Syn. mutation	-	a												a		a		a												a				a								
	1319	Syn. mutation	-																															b									
SalT	790	K24E	From a basic to an acid residue. Can affect SALT interactions with other proteins																							c																	
CPK17	602	Syn. mutation	-																							a																	
	771	Syn. mutation	-			a	ab					b	abc	ab	ab																												
	1959	Syn. mutation	-			a	ab					b	abc	ab	ab																												
	-	Haplotype	-										a		a																					b							
NHX1	4223	Syn. mutation	-																																							b	
	4285	S477N	May represent a loss of phosphoryl. site	a																																							
	-	Haplotype	-	a																																							

Legend: a- GLM_Q model; b- MLM_K model; c- MLM_Q+K model; TM- Transmembrane; Syn.-Synonymous. NS- Non-stress conditions (EC=0 dS/m); S- salt-stress conditions (EC=12 dS/m).

© 2012 The Authors

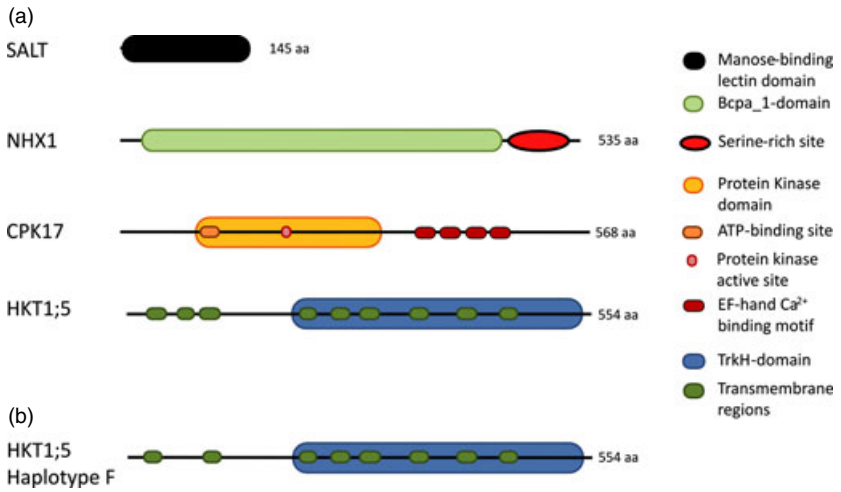

Figure 4 Schematic representation of the domains present in the studied proteins. All domains, except the serine-rich site, were identified using InterProScan. The serine-rich site was predicted based on phosphorylation probability data obtained with NetPhos 2.0. (a) All diagrams represent proteins from Nipponbare (reference haplotype); (b) Schematic representation of OsHKT1;5 for haplotype F, showing the putative destabilization of one transmembrane region.

cluster of potential phosphorylation sites (S134, T141, T142, S143 and T147). If this is indeed a site for OsHKT1;5 regulation by phosphorylation, the P to A alteration significantly increases the phosphorylation probability of T141 and T142, leaving the rest unchanged (Table S5), and thus could affect protein function. These calculations were obtained with NetPhos 2.0 (Blom *et al.*, 1999). Finally, the R184H modification observed in haplotypes E and N does not affect phosphorylation probability within a second putative cluster of phosphorylation (S194, S195, S196, S198, S200 and T204) sitting at the beginning of the TrkH-like domain (NetPhos 2.0 prediction).

SalT

Our association studies show that one of the 6 nonsynonymous SNPs found within SalT has an impact on Na^+ shoot content in salt-stress conditions, namely the K24E alteration, found only in haplotype F. Because this residue occurs within the first 30 AA residues of the protein, although no targeting sequence or signal peptide should be expected, we searched for these sequences, using InterProScan and Predotar 1.03 (http://urgi.versailles.inra.fr/predotar/predotar.html). No targeting sequences were detected, confirming that SalT should be accumulating in the cytoplasm. We have also analysed whether an alteration in a putative PTM could somehow account for the association result. K residues can be ubiquitinated, or SUMOylated, whilst E residues can be methylated. The K residue in haplotypes A, B, C and D is prone neither to ubiquitination nor to SUMOylation, according to predictions from UbPred and SUMOsp, respectively (Radivojac *et al.*, 2010; Ren *et al.*, 2009). Methylation of E in haplotype F was not assessed, due to the lack of bioinformatics tools. The K24E influence on the phosphorylation of neighbouring AA residues was ruled out due to the absence of potential phosphorylation sites at SalT N-terminus. Given these predictions, our hypothesis is that, because both K or E residues are predicted to be solvent exposed [according to NetSurfP 1.1 calculations (Petersen *et al.*, 2009)], a change in a positive to negative local charge would affect SalT interactions with other proteins and thus modify its function as a lectin.

OsNHX1

Amongst the six haplotypes identified at the coding sequence level, we found two nonsynonymous SNPs. Within these, only S477N, detected in haplotype B, was found to correlate with shoot length in our association studies. The loss of a S residue can imply the loss of a putative phosphorylation site, so we calculated the potential for phosphorylation sites at the S, T and Y residues, of the protein using NetPhos 2.0 (Blom *et al.*, 1999). S477 sits in a cluster of S residues with high phosphorylation probability and is itself a potential target for this PTM (Table S5). Moreover, calculations using NetSurfP 1.1 predict that it sits at a transition zone between an α-helix and the remaining C-terminus of the protein that appears to be fairly unstructured in a randomly coiled conformation. Together, these two observations point to the existence of a regulatory C-terminus domain of OsNHX1 that can be strongly affected by phosphorylation and may account for the correlation we observed between S477N and 'shoot length'. Further supporting this hypothesis, InterProScan identified a putative B_cpa1 domain, normally found in plant proteins involved in salt tolerance due to Na^+ uptake into vacuoles that spans from the N-terminus of the protein until approximately 15–25 AA residues from the hypothetical site regulated by phosphorylation (Figure 4).

OsCPK17

OsCPK17 belongs to a large multigene family and has been annotated in databases as a calcium-dependent protein kinase. Indeed, InterProScan analysis identifies the typical domains to support this. It contains a protein kinase domain at its N-terminal that includes an ATP-binding site and the serine/threonine active site. This is followed by an EF-hand calcium-binding site, comprising four calcium-binding EF-hand motifs. A potential signal peptide at the N-terminal was also identified, but could not be confirmed by other specific bioinformatics tools (Predotar or SignalP). Using NetPhos 2.0, we also identified a serine-rich motif at the N-terminal with a high phosphorylation probability (Table S5). We have identified five different AA changes for OsCPK17, but none of the nonsynonymous SNPs or small indels

© 2012 The Authors

was found to be correlated with the studied phenotypes in our association studies.

Discussion

The power of EcoTILLING as a genotyping technique

EcoTILLING technique employs a mismatch-specific nuclease that cleaves amplified PCR fragments at the site of a nucleotide polymorphism (Comai *et al.*, 2004). It was first developed to discover natural variants in *Arabidopsis* ecotypes (Comai *et al.*, 2004), and since then, it has been successfully employed in discovering and evaluating nucleotide diversity in humans, wheat, poplar, banana, Brassica and other organisms (Gilchrist *et al.*, 2006; Till *et al.*, 2006; Li *et al.*, 2010; Till *et al.*, 2010; Wang *et al.*, 2010).

In our study, we used EcoTILLING with an agarose-based detection system to reduce genotyping costs (Raghavan *et al.*, 2007). Although some primer pairs have not functioned, making it impossible to cover the whole target genes sequence, this system proved to be an effective genotyping technique for finding new natural allelic variants in rice. With EcoTILLING as genotyping strategy, we successfully defined haplotypes for a total of 392 rice accessions representative of *O. sativa* diversity. We identified a total of 301 novel SNPs and indels (per whole aligned DNA sequence) from five key salt-related rice genes, namely *OsCPK*17, *OsRMC*, *OsNHX*1, *OsHKT*1;5 and *Sal*T. Collectively, we found 40 new allelic variants in the target coding sequences: 15 haplotypes in *OsHKT*1;5, four in *OsRMC*, six in both *Sal*T and *OsNHX*1 and nine different haplotypes in *OsCPK*17. Expanding the allelic diversity to the whole gene level (including cds, intronic and both 5′ and 3′ UTR's regions), EcoTILLING enabled the identification of 60 new haplotypes in 392 accessions. It deserves mentioning that, excluding the reference haplotype ('Nipponbare'), only two allelic variants of 60 were previously described for *OsHKT*1;5 (Ren *et al.*, 2005; Cotsaftis *et al.*, 2010). Another interesting result from our study was that we could not cluster all the major salt-tolerant accessions into one haplotype. This result supports the fact that different genotypes have different mechanisms to cope with salt stress and that no accession carries all the favourable alleles for all the target loci. This fact was previously reported for salinity tolerance using another rice diversity set (European Rice Core Collection) (Ahmadi *et al.*, 2011).

In recent years, the introduction of instruments capable of producing millions of DNA sequence reads in a single run has changed genetic investigation by providing whole-genome and gene-targeted surveys at an incredible speed. Nevertheless, this state-of-the art 'next-generation' DNA sequencing requires appropriate funding and suitable computational algorithms to provide reliable data, whilst often more cost-effective techniques are required. The agarose gel detection system that we used for EcoTILLING is one such strategy, offering researchers with limited resources and the possibility to discover new allelic variants relevant, for instance, to breeding purposes.

Association analysis between haplotypes and salt tolerance in rice

There is growing interest on association mapping to link specific genes with certain phenotypic traits (Hall *et al.*, 2010). In natural *Arabidopsis* populations, a genome-wide association study showed a strong association between *AtHKT*1;1 (orthologue for *OsHKT*1;5) and local adaptation to saline environments (Baxter *et al.*, 2010).

In this work, we targeted five key genes previously described and characterized as related to salt-tolerance enhancement in rice, through different mechanisms such as Na^+/K^+ ratio equilibrium, signalling cascade and stress protection. In our work, 21 rare alleles were excluded from the association tests, which limited the number of inferences we could establish between specific SNPs and their relevance in salt tolerance. The *OsRMC* gene was found to be very conserved with only four haplotypes. Because three of those were rare alleles, their exclusion made it impossible to evaluate the gene. *OsRMC* extreme conservation may be due its close proximity to *sh*4 (locus responsible for reducing grain shattering along rice domestication, located 34 020 Kb apart) (He *et al.*, 2011). Another 18 SNPs, identified as rare in the target genes, could not be analysed in our association panel. The impact of this limitation suggests the importance of further exploring the putative functional importance of certain rare alleles. The use of bi-parental mapping populations is one of the major methods to confirm the association of rare alleles with target traits (Rafalski, 2002).

We analysed our results using three different statistical models, representing increasing degrees of control of relatedness from Q to K up to Q+K, which is generally the most efficient, notably in situation of complex population structure (Yu *et al.*, 2006). As often observed, the Q model was the less stringent, followed by K and then Q+K. However, because in plants, the best model can sometime differ depending on the trait (Yu *et al.*, 2006), we decided to evaluate all SNPs considered to be significant by any of the three models. Using a ρ value ≤ 0.01, we found significant associations at 11 SNP positions (Table 4). However, there is a risk that some of the associations identified only by the Q or by the K models are false positives. One way to overcome this concern is to validate the association by replicating and verifying the associations (Ingvarsson and Street, 2011). Thus, to separate true from false positives, it will be necessary to validate the five nonsynonymous SNPs and their significant influence by functional analysis of the salt-tolerance mechanism. It may also be necessary to use a larger population size because the power to detect an association is directly linked to this parameter (Zhu *et al.*, 2008).

Additionally to the nonsynonymous SNPs found relevant in our association studies, we found six synonymous SNPs (Table 4). Two hypotheses can explain their significance, the most likely one being that these SNPs do not correspond to the functional mutation but are in linkage disequilibrium (LD) with it. The resolution in association mapping is determined by the extent of LD. All markers in strong LD with a functional mutation will appear as significant in association tests; this is one of the method limitations. In rice, the extent of LD depends on the population and genomic region considered, but is often over 45 kb (Mather *et al.*, 2007). Another possible explanation is that these synonymous SNPs do not lead to AA residue changes at protein level; however, they are documented as leading to changes in mRNA structure, stability and splicing or even in delays or acceleration of protein folding that can result in different final protein conformations and thus functional alterations [for an extended review see (Hunt *et al.*, 2009)]. All of the six identified synonymous SNPs are within exonic regions; thus, they are expected to be conserved in the mature mRNA and will not interfere with mRNA splicing processes. Most of the synonymous SNPs affecting mRNA stability are located at the 3′-UTR, but some examples exist for their presence within coding regions, and so this hypothesis cannot be ruled out. Nevertheless, more likely, the identified SNPs may affect mRNA structure and consequently protein translation

© 2012 The Authors

and/or protein folding. These are interesting hypotheses that deserve being further explored.

Putative importance of the new allelic variants identified

Assessing the genetic diversity contained in large germplasm collections is a major challenge (Glaszmann *et al.*, 2010). Moreover, researchers should be capable of combining the newly obtained genetic information with its corresponding phenotypic effect. In our studies, we identified 11 significant SNPs related to salt stress, and using bioinformatic tools, we could assess the putative influence of nonsynonymous SNPs at protein level.

Within the analysed *OsHKT*1;5 haplotypes, we found three nonsynonymous SNPs with significant associations (Table 4). We observed two residue differences between 'Nipponbare' (haplotype A) and IR29 (haplotype B, which also includes IR64), specifically D129N and P140A. Ren *et al.* (2005) described six SNPs in the *OsHKT*1;5 coding region that lead to four AA changes, namely P140A, R184H, H332D and L395V. Our work confirmed the existence of these AA residue changes in haplotype E (which includes 'NonaBokra' and 'Pokkali'). Recently, another research group fully sequenced the 'Pokkali' allele and confirmed this result (Cotsaftis *et al.*, 2010). From the four previously described nonsynonymous changes, we could find significant associations for two mutations, namely P140A and R184H. The other two mutations, although identified in the assessed population, were not significantly associated with any trait. The putative importance of L395V substitution was previously discussed in the 3D model of both Nipponbare and Pokkali HKT1;5 (Cotsaftis *et al.*, 2012). The 3D model indicated that L395V change is positioned in close proximity of Gly391 (near the pore entrance in both transporters), being hypothesized that L395V mutation could directly influence pore rigidity thus slowing down Na^+ transport rates (Cotsaftis *et al.*, 2012). From our results, no significant association was found for this replacement, hence reinforcing the need to confirm individually each of these point mutations physiologically. For OsHKT1;5, the new allelic variant T67K mutation (significantly associated) seems particularly interesting because it could lead to the destabilization of a transmembrane helix due to the introduction of a positive charge in the membrane interior, thus interfering with protein structure. We found that OsHKT1;5 has nine putative transmembrane regions. This result suggests that the transporter architecture is similar to its homologues from wheat (TaHKT1;5) and *Arabidopsis* (AtHKT1;1) (Schachtman and Schroeder, 1994; Wang *et al.*, 1998; Uozumi *et al.*, 2000). These proteins belong to a family of K^+ transporters, comprising four sequential MPM (membrane–pore–membrane) motifs that probably have arisen from gene duplication and fusion events of a common bacterial K^+ channel ancestor, similar to *Streptomyces lividans* KscA (Doyle *et al.*, 1998; Durell *et al.*, 1999). Thus, a systematic investigation to reveal OsHKT1;5 real topological structure for both haplotypes to confirm T67K importance is essential. Haplotype E allelic variants carry both P140A and R184R alterations. According to recent molecular models of OsHKT1;5, both alterations sit in a cytoplasm-exposed loop and are not expected to have direct influence in Na^+ transport (Cotsaftis *et al.*, 2012). However, the influence of these residues on the transporter's function could be at regulatory level, eventually affecting a putative phosphorylation of OsHKT1;5. The P140A alteration, for instance, could affect the phosphorylation probability of T141 and T142. Nevertheless, if this is indeed a regulation mechanism for OsHKT1;5 in rice, it does not seem to be conserved in wheat because not only the segment of AA residues that contain P140 in rice is not present in TaHKT1;5, but also T141/T142 residues are not conserved. More farfetched but still worthy of mention is that R184 is part of one of two consensus sequences R-X-X-S/T found in OsHKT1;5. These consensus sequences were found to be recognized by SnRK2 protein kinases for phosphorylation of serine/threonine residues (Furihata *et al.*, 2006; Vlad *et al.*, 2008). Although evidence that HKT-family members are regulated by phosphorylation is still lacking, this is an interesting possibility to consider. Indeed, it is known in *Arabidopsis* that SnRK2.6 is able to directly phosphorylate a voltage-dependent K^+ channel (KAT1), negatively impacting K^+ transport (Sato *et al.*, 2009).

Two rice lectins responsive to several stresses including salt were previously described; the salt-induced gene *Sal*T (GenBank accession Z25811) found in *indica* rice (Claes *et al.*, 1990), and the mannose-binding rice lectin (MRL; GenBank accession AB012605) found in *japonica* (Teraoka *et al.*, 1990), with only two AA residue (P8L and Q74H) differences between the two proteins alignment. From our study, we now assume that the earlier described genes are in fact two allelic variants of *Sal*T, with MRL corresponding to haplotype A and *Sal*T being in fact haplotype B. Our association studies found a new *Sal*T allelic variant, K24E only present in haplotype F, which was significantly associated with Na^+ concentration in shoots under salt. This residue mutation, causing a positive to negative charge change, can affect SalT interactions with other proteins and thus impacting SalT function as a lectin, because it acts as a potent agglutinin (Branco *et al.*, 2004). Previous studies have demonstrated that SalT mediates essential protein–carbohydrate interactions in plants and binds to mannose, methyl α-mannopyranoside and trehalose (Zhang *et al.*, 2000). However, the mechanism by which SalT confers tolerance is still not clear, and further studies are required to fully assess its role in salinity.

In *Arabidopsis*, AtNHX1 has been shown to be regulated by the SOS2 kinase (Qiu *et al.*, 2004). However, Qiu *et al.* (2004) could not observe *in vitro* phosphorylation of NHX1 by SOS2 and suggested that either it was only a technique limitation or a membrane-associated intermediate exists between SOS2 and NHX1. In our work, we found a significant nonsynonymous mutation in OsNHX1, S477N present in haplotype B (which includes IR29 and tolerant accession FL478) that was associated with shoot length. S477 sits within a group of S residues with high phosphorylation probability and is putatively phosphorylated. This residue is located in the C-terminal of the protein that by homology with AtNHX1 should be in the cytoplasm (Sato and Sakaguchi, 2005) and therefore may affect the regulation of OsNHX1, interfering with its function. To confirm this hypothesis, further studies are required. In fact, the presence of two distinct salt-response accessions in haplotype B points to the need of an integrative study regarding favourable alleles.

In summary, our study provides detailed information about natural allelic variants existing in rice accessions worldwide. The molecular diversity of five target salt-related key genes was associated with salt tolerance or sensitivity, allowing the identification of favourable haplotypes. Our results open new lines of research to functionally validate the polymorphisms found, aiming to relate an allelic variant to a molecular or physiological phenotype. Desirably, the nonsynonymous mutations should be tested for their impact in protein function, such as in heterologous expression systems. Furthermore, the overall impact of these alterations should be tested in physiological assays in rice plants,

© 2012 The Authors

preferentially in knockout target genes backgrounds. After further confirmation of the relevance of these new variants in progeny testing, these results may be used by rice breeders, to identify accessions that can be used to improve salt tolerance. However, we should not underestimate the complexity of salt-tolerance traits that require the involvement of several key genes. Although we have found correlations between some point mutations in the target genes and the studied phenotypic traits, this is not enough to assign a certain tolerant phenotype to a single SNP. On the contrary, we believe that our study highlights the fact that a favourable allelic variant in one key gene is not enough to provide salt tolerance, and a breeding strategy combining several favourable allelic variants at key gene loci should be used.

Experimental procedures

Rice accessions

We used a working set composed of 392 *O. sativa* accessions listed in Supporting Information Table S1. All accessions were provided by International Rice Gene Bank held at IRRI, Philippines. This set is comprised of representative genotypes of the five variety groups commonly distinguished in *O. sativa* (*indica*, *temperate* and *tropical japonica*, *aromatic* and *aus*). Of these accessions, 373 were selected from the composite collection of 2339 accessions that was designed to cover the diversity and geographic origins of traditional and improved rice types, and genotyped using 45 SSR markers within the 'Generation Challenge Programme' (GCP). The 373 accessions from GCP composite collection contained 82% of the alleles and 95% of the diversity (Shannon–Weaver index ratio) of the composite collection. Both GCP composite collection and the working set showed population structures similar to that found by Garris *et al.* (2005) on their own collection (K.L. McNally, R. Manzano, M. Macatangay, M. Redondo, V. Lacorte, M. Zaidem, J. Detras, M. Barile, S. Mercado-Quilloy, M.E.B. Naredo L. Benoit, R. Rivallan, B. Courtois, C. Billot, A. Garavito, M. Lorieux, C.P. Martinez, T. Borba, R.V. Brondani, C. Brondani, M. Cissoko, M.-N. Ndjiondjop, A. Famoso, S.R. McCouch, N.R. Sackville Hamilton, and J.-C. Glaszmann, unpublished). These 373 accessions were complemented by 19 accessions chosen from OryzaSNP panel (McNally *et al.*, 2009) and also salt-tolerant and sensitive genotypes (not available from GCP composite collection).

Primer design

The complete genomic sequences (based on 'Nipponbare' sequence) of the five target genes (*OsCPK*17, *OsRMC*, *OsHKT*1;5, *OsNHX*1 and *Sal*T) were retrieved from OryGenesDB (http://orygenesdb.cirad.fr/). Primer suite software (http://www-fgg.eur.nl/kgen/primer/) was used to design primer pairs for ~1Kb amplicons spanning the whole gene sequence (including cds, intronic and UTR regions) with different amplicons overlapping by about 200 bp. On testing, several primer pairs were discarded due to low specificity in regions of high variability (such as spanning an intron). All primers are listed in Table 5.

EcoTILLING

We extracted DNA from leaf tissue according to Fulton *et al.* (1995) and adjusted concentration to 0.5 ng/μL. For EcoTILLING, DNA from each genotype was contrasted with either 'IR64' (for *indica*, *aus* and admixed accessions) or 'Nipponbare' (for *japonica*, *aromatic* and admixed accessions) separately in a 1:1 ratio. We amplified circa 1Kb overlapping segments of each gene using specific primers (Table 5). Polymerase chain reaction (PCR) was performed in 14-μL final volume, using 3.5 ng of total DNA as template, 0.4 U/reaction of Taq DNA polymerase (Promega, Madison, WI, USA). PCR cycling conditions were set at 95 °C for 3 min, followed by 35 cycles of 95 °C for 20 s, 57–64 °C

Table 5 Primer sequences used in EcoTILLING for PCR amplification of the five target genes. Primers were designed to amplify overlapping fragments covering the whole gene sequence

Target gene/locus	Primer ID	Product size (bp)	Forward primer sequence (5'T3')	Reverse primer sequence (5'T3')	Ann. Temp. (°C)
OsHKT1; 5 (*Os01g20160*)	P1	1096	GCATGTCCTTTGTGGTCTATATGGTG	AGAGAAAGTGAGGAAAATCCACATC	60
	P2	1031	CATGAGAGTGAGCTTCCCTTGTTTG	CCACCTTACACCACTTTCGTACCTG	Discard
	P3	1181	CCACCACTATGCGTACCTGATAACC	GATTGACGCTTTCAGTTTGTTGGAG	Discard
	P4	999	AAGGGAGTATGTGCTTGCAGTTGAC	GTGCTCTACGTGGTGATGATGTATG	Discard
	P5	1160	ATTTAGATGCCTCGAGACCCTGCAC	CTACTTCACCAACGCCAAGTACTCC	62
	P6	1090	TCCGACGTCCTAACCCTAGCTCTAC	TGGTGATGATTTACGCACATGAGAC	65
SalT (*Os01g24710*)	P1	1087	TTCCAGACCTTCCAAAGAATCCAA	ACAAGGAAATTTAAGCGACCACGA	62
	P2	825	ACCACTCAACACCGGTAGGACACT	GCAGATTAAACTGGGCTCCTCTGA	64
OsRMC (*Os04g56430*)	P1	1240	ATGCGGCCGCCGAGTGCGACGAATGTA	CTACTCACGCAGCACCACC	68
OsCPK17 (*Os07g06740*)	P1	1034	CGCAATGCACCACCATTACC	CCGATCCAAATCCACCAACC	64
	P2	1171	CGGCCATAATGCTTCATCAG	GTACCAGTGCGTGGAGAAGG	Discard
	P3	1094	CTGGAAATAGAAGGGGCATAGTC	CTTAAAATGCCAAGGGAATGG	57
	P4	1123	CCACTAGTAGCAATGAGCATATGAAG	GATGAAATTGCGGGATTGAG	61
	P5	1088	CACTCCATTCTACATTCAGGGGTAG	GCAAGGCCAGCTAAGTTTTG	Discard
	P6	979	AATTGGAGGTTGGGCCATAG	TGTGAGGTGGAAGAAGCAAAC	61
OsNHX1 (*Os07g47100*)	P1	1173	AGCATGACACAGGCCCTCTTAAAC	ACAGCTAGGCCCAGAAGAAACAAAC	67
	P2	1132	CTGTGGTTAGATTTCGGCTTCAGTTC	GAAAATCTCCTACATCCAGCGTTCC	Discard
	P3	895	TCTTCCGTGAGTTGGTAGTGCTTTC	ATACTGCATCAAAGCAGAGGCACAC	63
	P4	881	GCTCAGTGTTAACTTGGATGTTGGTG	ATAGTCTTGCACTGACGGACTGACC	65
	P5	1092	CTGGAGACAGCAAGTGAGTATCTGG	TGTACTCTCGAGGTCAGAACCTTGC	66
	P6	1099	AAAGCCATTGATCAGGCTGCTG	AGGGGTAAGACTTCGAAACCAGGTC	67

© 2012 The Authors

(depending on primer specificity, see Table 5) for 30 s, 72 °C for 30 s and a final extension of 7 min at 72°C. PCR products were denatured at 99 °C for 10min and renatured initially at 70 °C for 20 s followed by 69 cycles with a temperature decrease by 0.3 ° C per cycle. Celery juice extract (CJE) was produced using Till *et al.* (2004) technique. Whenever the amplicons differed in sequence content between reference and target germplasm, heteroduplex mismatch molecules formed upon re-annealing. Mismatch cleavage (CJE digestion) and EcoTILLING analysis in agarose gel were performed according to Raghavan *et al.* (2007).

Identification of haplotypes

The EcoTILLING profiles enabled the assembly of accessions into haplotype groups by comparing digestion patterns. To confirm the polymorphisms identified by EcoTILLING, accessions representative of the different haplotypes were randomly selected, and the different fragments obtained for each target gene were sequenced. PCR products were purified using Roche Diagnostics kit (Germany) and sequenced by Beckman Coulter Genomics (http://www.cogenicsonline.com/) for both strands. To minimize sequencing errors (by confirming sequence chromatograms for each sample), each sequencing reaction was performed twice. The number, position and type of SNPs in the haplotypes are listed in Table S3. To further improve analysis of the genotyping data, we performed a network analysis of different haplotype groups using Haplophyle software (http://haplophyle.cirad.fr/Haplophyle/) with default parameters.

Phenotyping for salt-stress tolerance

For association analysis with salt stress, we assembled a set of accessions representing all the haplotypes existing at cds level for each gene. Whenever possible, we included three accessions per haplotype in the phenotyping assay. The final list included 59 rice accessions representative of all the allelic variants found (Table S4).

The selected accessions were evaluated for salinity tolerance at seedling stage as previously described (Gregorio *et al.*, 1997). The phenotyping experiment included a control and a salt-stress treatment, with the presence of two checks: susceptible 'IR29' and tolerant 'FL478' (IR66946-3R-178-1-1) in all trays. The experimental design was a split plot design with 11 genotypes per tray and 16 plants per genotype. Salt-stress imposition was performed 10 days after sowing, by supplementing Yoshida solution (Yoshida *et al.*, 1976) with 6 g/L of NaCl (EC = 12 dS/m). In the control trays, no salt was added, and plants were maintained at 0 dS/m of electrical conductivity (EC). The experiment was conducted in a glasshouse maintained at approximately 29 °C/22 °C day/night with 70% relative humidity.

Test entries were visually rated for injury symptoms at 10 days after initial salinization using a 1–9 scale according to Gregorio *et al.* (1997). After 10 days of salt imposition, we evaluated several parameters such as plant height, root and shoot fresh and dry weight. Chlorophyll a and b were determined by extracting freeze-dried material (4th fully expanded leaf) in 80% acetone overnight, and readings were carried out using UV Spectrophotometer (UV-1800; Shimadzu, Kyoto, Japan). The determination of Na^+ and K^+ content in shoots (3rd fully expanded leaf) and roots was performed by extracting both ions in acetic acid (0.1 N) overnight at 80 °C. The concentration of Na^+ and K^+ in the extract was determined using a flame photometer (model 420; Sherwood Scientific, Cambridge, UK). Leaf osmotic potential was estimated in 10 μL of leaf juice extract using a Vapro Osmometer 5520 (Wescor Inc., Logan, UT).

For biomass measurements, we used five plants per accession, whilst for all other parameters, we used three plants per accession, for both control and salt conditions. Special care was taken for Na^+ and K^+ determination, with all samples washed thoroughly using nanopure water.

Population structure and kinship determination

Nine of the accessions listed in Table S4 were only genotyped as checks for salt tolerance. The association study was conducted on the 50 remaining genotypes. Population structure and kinship amongst the panel accessions are factors known to induce false positives in the association tests (Yu *et al.*, 2006). These factors need to be controlled in the analyses. To establish population structure, we used previously obtained genotypic data of the 50 accessions, at 36 independent SSR loci (K.L. McNally, R. Manzano, M. Macatangay, M. Redondo, V. Lacorte, M. Zaidem, J. Detras, M. Barile, S. Mercado-Quilloy, M.E.B. Naredo L. Benoit, R. Rivallan, B. Courtois, C. Billot, A. Garavito, M. Lorieux, C.P. Martinez, T. Borba, R.V. Brondani, C. Brondani, M. Cissoko, M.-N. Ndjiondjop, A. Famoso, S.R. McCouch, N.R. Sackville Hamilton, and J.-C. Glaszmann, unpublished). A Principal Component Analysis (PCA) was run on the 50 accessions x 36 SSR matrix using TASSEL software (http://www.maizegenetics.net/tassel). By observing the cumulative proportion of the explained variability as recommended by TASSEL, we decided to choose 10 PCA axes. Kinship amongst these accessions was estimated using the 'Kin' function of TASSEL.

Genotype–phenotype association analysis

Association analysis between target genes and the investigated traits was performed using TASSEL. For each gene, the analyses were run both on individual SNPs and on haplotypes. We excluded from the analysis all the haplotypes and SNPs with an allele frequency <5%. We tested associations between SNPs/haplotypes and phenotypes separately for control and salt-stress conditions.

To limit spurious associations due to multiple levels of relatedness between accessions, we used three statistical models: (i) general linear model (GLM_Q), where the SNP/Haplotype is considered as a fixed effect and the coordinates of each accession in the 10 PCA axes (*Q* matrix) used as population structure factors; (ii) mixed linear model (MLM_K), where the SNP/Haplotype is considered as fixed effects, and the kinship matrix (*K*) is incorporated as a random effect; (iii) MLM_K+Q model, in which both the Q and the K matrices are considered. To correct for multiple tests, we used Bonferonni correction. Because we had four independent target genes within which the polymorphisms were in high LD, a significance threshold of 0.01 was used for all tested models.

Evaluation of putative changes at protein level

At protein level, we investigated putative consequences of the significant associations found in SNPs (Table 4). For this purpose, we used several bioinformatics tools available at ExPASy Bioinformatics Resource Portal (http://expasy.org/tools/). We used InterProScan to identify protein domains. To further analyse the SNP influence in putative post-translational modifications (PTM), such as phosphorylation, ubiquitination and SUMOylation, we used the following software predictions, respectively: NetPhos 2.0, UbPred and SUMOsp.

© 2012 The Authors

Acknowledgements

Fundação para a Ciência e a Tecnologia (FCT) through National Funds: PTDC/AGR-GPL/70920/2006 and # PEst-OE/EQB/LA0004/2011. Sónia Negrão and Isabel A. Abreu gratefully acknowledge FCT for grants SFRH/BPD/34593/2007 and SFRH/BPD/78314/2011. All plant material was kindly provided by the International Rice Genebank (IRRI, Philippines). We also thank Cláudio Soares for critical reading, M. Elizabeth Naredo (IRRI) for laboratory support and Naireen Vispo, Junrey Amas and Mona Lisa Jubay (IRRI) for phenotyping assistance.

References

Ahmadi, N., Negrão, S., Katsantonis, D., Frouin, J., Ploux, J., Letourmy, P., Droc, G., Babo, P., Trindade, H., Bruschi, G., Greco, R., Oliveira, M., Piffanelli, P. and Courtois, B. (2011) Targeted association analysis identified *japonica* rice varieties achieving Na^+/K^+ homeostasis without the allelic make-up of the salt tolerant *indica* variety Nona Bokra. *Theor. Appl. Genet.* **123**, 881–895.

Baxter, I., Brazelton, J.N., Yu, D.N., Huang, Y.S., Lahner, B., Yakubova, E., Li, Y., Bergelson, J., Borevitz, J.O., Nordborg, M., Vitek, O. and Salt, D.E. (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.

Blom, N., Gammeltoft, S. and Brunak, S. (1999) Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* **294**, 1351–1362.

Branco, A.T., Bernabe, R.B., dos Santos Ferreira, B., de Oliveira, M.V., Garcia, A.B. and de Souza Filho, G.A. (2004) Expression and purification of the recombinant SALT lectin from rice (*Oryza sativa* L.). *Protein Expr. Purif.* **33**, 34–38.

Chong, K., Jiang, J.F., Li, J.H., Xu, Y.Y., Han, Y., Bai, Y., Zhou, G.X., Lou, Y.G. and Xu, Z.H. (2007) RNAi knockdown of *Oryza sativa* root meander curling gene led to altered root development and coiling which were mediated by jasmonic acid signalling in rice. *Plant, Cell Environ.* **30**, 690–699.

Claes, B., Dekeyser, R., Villarroel, R., Vandenbulcke, M., Bauw, G., Vanmontagu, M. and Caplan, A. (1990) Characterization of a rice gene showing organ-specific expression in response to salt stress and drought. *Plant Cell*, **2**, 19–27.

Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R. and Henikoff, S. (2004) Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *Plant J.* **37**, 778–786.

Cotsaftis, O., Plett, D., Johnson, A.A., Walia, H., Wilson, C., Ismail, A.M., Close, T.J., Tester, M. and Baumann, U. (2010) Root-specific transcript profiling of contrasting rice genotypes in response to salinity stress. *Mol. Plant*, **4**, 25–41.

Cotsaftis, O., Plett, D., Shirley, N., Tester, M. and Hrmova, M. (2012) A two-staged model of Na^+ exclusion in rice explained by 3D modeling of HKT transporters and alternative splicing. *PLoS ONE*, **7**, e39865.

Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A.L., Gulbis, J.M., Cohen, S.L., Chait, B.T. and MacKinnon, R. (1998) The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science*, New York, NY, **280**, 69–77.

Durell, S.R., Hao, Y.L., Nakamura, T., Bakker, E.P. and Guy, H.R. (1999) Evolutionary relationship between K(+) channels and symporters. *Biophys. J.* **77**, 775–788.

Fukuda, A., Nakamura, A., Tagiri, A., Tanaka, H., Miyao, A., Hirochika, H. and Tanaka, Y. (2004) Function, intracellular localization and the importance in salt tolerance of a vacuolar Na^+/H^+ antiporter from rice. *Plant Cell Physiol.* **45**, 146–159.

Fulton, T.M., Chunwongse, J. and Tanksley, S.D. (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol. Biol. Rep.* **13**, 207–209.

Furihata, T., Maruyama, K., Fujita, Y., Umezawa, T., Yoshida, R., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2006) Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. *Proc. Natl Acad. Sci. USA*, **103**, 1988–1993.

Garris, A.J., Tai, T.H., Coburn, J., Kresovich, S. and McCouch, S. (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics*, **169**, 1631–1638.

Gilchrist, E.J., Haughn, G.W., Ying, C.C., Otto, S.P., Zhuang, J., Cheung, D., Hamberger, B., Aboutorabi, F., Kalynyak, T., Johnson, L., Bohlmann, J., Ellis, B.E., Douglas, C.J. and Cronk, Q.C. (2006) Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Mol. Ecol.* **15**, 1367–1378.

Glaszmann, J.C., Kilian, B., Upadhyaya, H.D. and Varshney, R.K. (2010) Accessing genetic diversity for crop improvement. *Curr. Opin. Plant Biol.* **13**, 167–173.

Gregorio, G.B., Senadhira, D. and Mendoza, R.D. (1997) *Screening Rice for Salinity Tolerance*. IRRI discussion papers No 22. Manila (Philippines): International Rice Research Institute.

Hall, D., Tegstrom, C. and Ingvarsson, P.K. (2010) Using association mapping to dissect the genetic basis of complex traits in plants. *Brief. Funct. Genomics*, **9**, 157–165.

He, Z., Zhai, W., Wen, H., Tang, T., Wang, Y., Lu, X., Greenberg, A.J., Hudson, R.R., Wu, C.-I. and Shi, S. (2011) Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLoS Genet.* **7**, e1002100.

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

Ingvarsson, P.K. and Street, N.R. (2011) Association genetics of complex traits in plants. *New Phytol.* **189**, 909–922.

Li, X.P., Zhao, X.Q., He, X., Zhao, G.Y., Li, B., Liu, D.C., Zhang, A.M., Zhang, X. Y., Tong, Y.P. and Li, Z.S. (2010) Haplotype analysis of the genes encoding glutamine synthetase plastic isoforms and their association with nitrogen-use- and yield-related traits in bread wheat. *New Phytol.* **189**, 449–458.

Lin, Y.J., Wan, B.L. and Mou, T.M. (2007) Expression of rice Ca^{2+}-dependent protein kinases (CDPKs) genes under different environmental stresses. *FEBS Lett.* **581**, 1179–1189.

Mather, K.A., Caicedo, A.L., Polato, N.R., Olsen, K.M., McCouch, S. and Purugganan, M.D. (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics*, **177**, 2223–2232.

Matsumoto, T., Wu, J.Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., Mizuno, H., Yamamoto, K., Antonio, B.A., Baba, T., Sakata, K., Nagamura, Y., Aoki, H., Arikawa, K., Arita, K., Bito, T., Chiden, Y., Fujitsuka, N., Fukunaka, R., Hamada, M., Harada, C., Hayashi, A., Hijishita, S., Honda, M., Hosokawa, S., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, K., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Kurita, K., Katagiri, S., Kikuta, A., Kobayashi, H., Kobayashi, N., Machita, K., Maehara, T., Masukawa, M., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nagata, Y., Naito, S., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Meguro, A., Negishi, M., Ohta, I., Ohta, T., Okamoto, M., Ono, N., Saji, S., Sakaguchi, M., Sakai, K., Shibata, M., Shimokawa, T., Song, J.Y., Takazaki, Y., Terasawa, K., Tsugane, M., Tsuji, K., Ueda, S., Waki, K., Yamagata, H., Yamamoto, M., Yamamoto, S., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H.S., Yano, M., Sasaki, T., Yuan, Q.P., Shu, O.T., Liu, J., Jones, K.M., Gansberger, K., Moffat, K., Hill, J., Bera, J., Fadrosh, D., Jin, S.H., Johri, S., Kim, M., Overton, L., Reardon, M., Tsitrin, T., Vuong, H., Weaver, B., Ciecko, A., Tallon, L., Jackson, J., Pai, G., Van Aken, S., Utterback, T., Reidmuller, S., Feldblyum, T., Hsiao, J., Zismann, V., Iobst, S., de Vazeille, A.R., Buell, C.R., Ying, K., Li, Y., Lu, T.T., Huang, Y.C., Zhao, Q., Feng, Q., Zhang, L., Zhu, J.J., Weng, Q.J., Mu, J., Lu, Y.Q., Fan, D.L., Liu, Y.L., Guan, J.P., Zhang, Y.J., Yu, S.L., Liu, X.H., Zhang, Y., Hong, G.F., Han, B., Choisne, N., Demange, N., Orjeda, G.,Samain, S., Cattolico, L., Pelletier, E., Couloux, A., Segurens, B., Wincker, P., D'Hont, A., Scarpelli, C., Weissenbach, J., Salanoubat, M., Quetier, F., Yu, Y., Kim, H.R., Rambo, T., Currie, J., Collura, K., Luo, M.Z., Yang, T.J., Ammiraju, J.S.S., Engler, F., Soderlund, C., Wing, R.A., Palmer, L.E., de la Bastide, M., Spiegel, L., Nascimento, L., Zutavern, T., O'Shaughnessy, A., Dike, S., Dedhia, N., Preston, R., Balija, V., McCombie, W.R., Chow, T.Y., Chen, H.H., Chung, M. C., Chen, C.S., Shaw, J.F., Wu, H.P., Hsiao, K.J., Chao, Y.T., Chu, M.K., Cheng, C.H., Hour, A.L., Lee, P.F., Lin, S.J., Lin, Y.C., Liou, J.Y., Liu, S.M., Hsing, Y.I., Raghuvanshi, S., Mohanty, A., Bharti, A.K., Gaur, A., Gupta, V., Kumar, D., Ravi, V., Vij, S., Kapur, A., Khurana, P., Khurana, P., Khurana, J.P., Tyagi, A.K., Gaikwad, K., Singh, A., Dalal, V., Srivastava, S., Dixit, A., Pal, A.

© 2012 The Authors

K., Ghazi, I.A., Yadav, M., Pandit, A., Bhargava, A., Sureshbabu, K., Batra, K., Sharma, T.R., Mohapatra, T., Singh, N.K., Messing, J., Nelson, A.B., Fuks, G., Kavchok, S., Keizer, G., Llaca, E.L.V., Song, R.T., Tanyolac, B., Young, S., Il, K. H., Hahn, J.H., Sangsakoo, G., Vanavichit, A., de Mattos, L.A.T., Zimmer, P. D., Malone, G., Dellagostin, O., de Oliveira, A.C., Bevan, M., Bancroft, I., Minx, P., Cordum, H., Wilson, R., Cheng, Z.K., Jin, W.W., Jiang, J.M., Leong, S.A., Iwama, H., Gojobori, T., Itoh, T., Niimura, Y., Fujii, Y., Habara, T., Sakai, H., Sato, Y., Wilson, G., Kumar, K., McCouch, S., Juretic, N., Hoen, D., Wright, S., Bruskiewich, R., Bureau, T., Miyao, A., Hirochika, H., Nishikawa, T., Kadowaki, K. and Sugiura, M. (2005) The map-based sequence of the rice genome. *Nature*, **436**, 793–800.

McCouch, S.R. and Kovach, M.J. (2008) Leveraging natural diversity: back through the bottleneck. *Curr. Opin. Plant Biol.* **11**, 193–200.

McNally, K.L., Childs, K.L., Bohnert, R., Davidson, R.M., Zhao, K., Ulat, V.J., Zeller, G., Clark, R.M., Hoen, D.R., Bureau, T.E., Stokowski, R., Ballinger, D. G., Frazer, K.A., Cox, D.R., Padhukasahasram, B., Bustamante, C.D., Weigel, D., Mackill, D.J., Bruskiewich, R.M., Ratsch, G., Buell, C.R., Leung, H. and Leach, J.E. (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. *Proc. Natl Acad. Sci. USA*, **106**, 12273–12278.

Mian, A.A., Prasad, S. and Maathuis, F.J.M. (2011) Improving crop salt tolerance: anion and cation transporters as genetic engineering targets. In: *Plant Nutrition and Abiotic Stress Tolerance III Plant Stress 5* (Anjum, N.A. and Lopez-Lauri, F., eds), pp. 64–72. UK: Global Science Books.

Molina, J., Sikora, M., Garud, N., Flowers, J.M., Rubinstein, S., Reynolds, A., Huang, P., Jackson, S., Schaal, B.A., Bustamante, C.D., Boyko, A.R. and Purugganan, M.D. (2011) Molecular evidence for a single evolutionary origin of domesticated rice. *Proc. Natl Acad. Sci. USA*, **108**, 8351–8356.

Negrão, S., Courtois, B., Ahmadi, N., Abreu, I., Saibo, N. and Oliveira, M.M. (2011) Recent updates on salinity stress in rice: from physiological to molecular responses. *Crit. Rev. Plant Sci.* **30**, 329–377.

Petersen, B., Petersen, T.N., Andersen, P., Nielsen, M. and Lundegaard, C. (2009) A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Struct. Biol.* **9**, 51.

Qiu, Q.S., Guo, Y., Quintero, F.J., Pardo, J.M., Schumaker, K.S. and Zhu, J.K. (2004) Regulation of vacuolar Na+/H+ exchange in *Arabidopsis thaliana* by the salt-overly-sensitive (SOS) pathway. *J. Biol. Chem.* **279**, 207–215.

Rabbani, M.A., Maruyama, K., Abe, H., Khan, M.A., Katsura, K., Ito, Y., Yoshiwara, K., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA get-blot analyses. *Plant Physiol.* **133**, 1755–1767.

Radivojac, P., Vacic, V., Haynes, C., Cocklin, R.R., Mohan, A., Heyen, J.W., Goebl, M.G. and Iakoucheva, L.M. (2010) Identification, analysis, and prediction of protein ubiquitination sites. *Proteins*,**78**, 365–380.

Rafalski, A. (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**, 94–100.

Raghavan, C., Naredo, M.E.B., Wang, H.H., Atienza, G., Liu, B., Qiu, F.L., McNally, K.L. and Leung, H. (2007) Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. *Mol. Breeding*, **19**, 87–101.

Ren, Z.H., Gao, J.P., Li, L.G., Cai, X.L., Huang, W., Chao, D.Y., Zhu, M.Z., Wang, Z.Y., Luan, S. and Lin, H.X. (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* **37**, 1141–1146.

Ren, J., Gao, X., Jin, C., Zhu, M., Wang, X., Shaw, A., Wen, L., Yao, X. and Xue, Y. (2009) Systematic study of protein sumoylation: development of a site-specific predictor of SUMOsp 2.0. *Proteomics*, **9**, 3409–3412.

Sato, Y. and Sakaguchi, S.M. (2005) Topogenic properties of transmembrane segments of *Arabidopsis thaliana* NHX1 reveal a common topology model of the Na^+/H^+ exchanger family. *J. Biochem.* **138**, 425–431.

Sato, A., Sato, Y., Fukao, Y., Fujiwara, M., Umezawa, T., Shinozaki, K., Hibi, T., Taniguchi, M., Miyake, H., Goto, D.B. and Uozumi, N. (2009) Threonine at position 306 of the KAT1 potassium channel is essential for channel activity and is a target site for ABA-activated SnRK2/OST1/SnRK2.6 protein kinase. *Biochem. J.* **424**, 439–448.

Schachtman, D.P. and Schroeder, J.I. (1994) Structure and transport mechanism of a high-affinity Potassium uptake transporter from higher-plants. *Nature*, **370**, 655–658.

Senadheera, P., Singh, R.K. and Maathuis, F.J.M. (2009) Differentially expressed membrane transporters in rice roots may contribute to cultivar dependent salt tolerance. *J. Exp. Bot.* **60**, 2553–2563.

Siringam, K., Juntawong, N., Chan-um, S., Boribookaset, T. and Kirdmanee, C. (2012) Salt tolerance enhancement in *indica* rica (*Oryza sativa* L.) seedlings using exogenous sucrose supplementation. *Plant Omics J.* **5**, 52–59.

Teraoka, T., Sakakibara, T., Den, E., Hosokawa, D. and Watanabe, M. (1990) A novel rice lectin specific to mannoside glucoside residues in rice seedlings. *Agric. Biol. Chem.* **54**, 3053–3056.

Till, B.J., Burtner, C., Comai, L. and Henikoff, S. (2004) Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res.* **32**, 2632–2641.

Till, B.J., Zerr, T., Bowers, E., Greene, E.A., Comai, L. and Henikoff, S. (2006) High-throughput discovery of rare human nucleotide polymorphisms by EcoTILLING. *Nucleic Acids Res.* **34**, e99.

Till, B.J., Jankowicz-Cieslak, J., Sagi, L., Huynh, O.A., Utsushi, H., Swennen, R., Terauchi, R. and Mba, C. (2010) Discovery of nucleotide polymorphisms in the Musa gene pool by EcoTILLING. *Theor. Appl. Genet.* **121**, 1381–1389.

Tung, C.-W., Zhao, K., Wright, M.H., Ali, M.L., Jung, J., Kimball, J., Tyagi, W., Thomson, M.J., McNally, K., Leung, H., Kim, H., Ahn, S.-N., Reynolds, A., Scheffler, B., Eizenga, G., McClung, A., Bustamante, C. and McCouch, S.R. (2010) Development of a research platform for dissecting Phenotype-Genotype associations in rice (*Oryza* spp.). *Rice*, **3**, 205–217.

Uozumi, N., Kim, E.J., Rubio, F., Yamaguchi, T., Muto, S., Tsuboi, A., Bakker, E. P., Nakamura, T. and Schroeder, J.I. (2000) The Arabidopsis HKT1 gene homolog mediates inward Na^+ currents in *Xenopus laevis* oocytes and Na^+ uptake in *Saccharomyces cerevisiae*. *Plant Physiol.* **122**, 1249–1259.

Vlad, F., Turk, B.E., Peynot, P., Leung, J. and Merlot, S. (2008) A versatile strategy to define the phosphorylation preferences of plant protein kinases and screen for putative substrates. *Plant J.* **55**, 104–117.

Walia, H., Wilson, C., Condamine, P., Liu, X., Ismail, A.M., Zeng, L.H., Wanamaker, S.I., Mandal, J., Xu, J., Cui, X.P. and Close, T.J. (2005) Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. *Plant Physiol.* **139**, 822–835.

Wang, T.B., Gassmann, W., Rubio, F., Schroeder, J.I. and Glass, A.D.M. (1998) Rapid up-regulation of HKT1, a high-affinity potassium transporter gene, in roots of barley and wheat following withdrawal of potassium. *Plant Physiol.* **118**, 651–659.

Wang, N., Shi, L., Tian, F., Ning, H., Wu, X., Long, Y. and Meng, J. (2010) Assessment of FAE1 polymorphisms in three Brassica species using EcoTILLING and their association with differences in seed erucic acid contents. *BMC Plant Biol.* **10**, 137.

Wassmann, R., Jagadish, S.V.K., Sumfleth, K., Pathak, H., Howell, G., Ismail, A., Serraj, R., Redona, E., Singh, R.K. and Heuer, S. (2009) Regional Vulnerability of Climate Change Impacts on Asian Rice Production and Scope for Adaptation. *Adv. Agron.* **102**, 91.

Yoshida, S., Forno, D., Cock, J. and Gomez, K. (1976) *Laboratory Manual for Physiological Studies of Rice*. Manila (Philippines): International Rice Research Institute.

Yu, J.M., Pressoir, G., Briggs, W.H., Bi, I.V., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S. and Buckler, E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203–208.

Zdobnov, E.M. and Apweiler, R. (2001) InterProScan–an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*, **17**, 847–848.

Zeng, L.H. (2005) Exploration of relationships between physiological parameters and growth performance of rice (*Oryza sativa* L.) seedlings under salinity stress using multivariate analysis. *Plant Soil*, **268**, 51–59.

Zhang, H.X. and Blumwald, E. (2001) Transgenic salt-tolerant tomato plants accumulate salt in foliage but not in fruit. *Nat. Biotechnol.* **19**, 765–768.

Zhang, W., Peumans, W.J., Barre, A., Astoul, C.H., Rovira, P., Rouge, P., Proost, P., Truffa-Bachi, P., Jalali, A.A. and Van Damme, E.J. (2000) Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (*Oryza sativa*) plants. *Planta*, **210**, 970–978.

Zhao, K., Tung, C.W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., McClung, A.M., Bustamante, C.D. and McCouch, S.R. (2011) Genome-wide association

© 2012 The Authors

mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2**, 467.
Zhu, J.K. (2001) Plant salt tolerance. *Trends Plant Sci.* **6**, 66–71.
Zhu, C., Gore, M., Buckler, E.S. and Yu, J. (2008) Status and prospects of association mapping in plants. *Plant Genome*, **1**, 5–20.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Distribution of the observed haplotypes in genes *Sal*T, *OsCPK*17 and *OsNHX1*.
Table S1 *Oryza sativa* germplasm accessions genotyped by EcoTILLING (392).
Table S2 Haplotype attribution (at the coding sequence level) of the 392 accessions for all target genes.
Table S3 Detailed information on the polymorphisms found (SNPs and indels) within the whole genomic region of the five target genes.
Table S4 Complete phenotypic data for the 59 accessions evaluated under control (EC = 0 dS/m) and salt-stress conditions (EC = 12 dS/m).
Table S5 Phosphorylation probability for several AA residues in the examined genes.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2012 The Authors