

A Computational Systems Biology Study for Understanding Salt Tolerance Mechanism in Rice

Juexin Wang^{1,2}, Liang Chen¹, Yan Wang¹, Jingfen Zhang², Yanchun Liang^{1*}, Dong Xu^{1,2*}

1 College of Computer Science and Technology, Jilin University, Changchun, China, **2** Digital Biology Laboratory, Computer Science Department, and Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, Missouri, United States of America

Abstract

Salinity is one of the most common abiotic stresses in agriculture production. Salt tolerance of rice (*Oryza sativa*) is an important trait controlled by various genes. The mechanism of rice salt tolerance, currently with limited understanding, is of great interest to molecular breeding in improving grain yield. In this study, a gene regulatory network of rice salt tolerance is constructed using a systems biology approach with a number of novel computational methods. We developed an improved volcano plot method in conjunction with a new machine-learning method for gene selection based on gene expression data and applied the method to choose genes related to salt tolerance in rice. The results were then assessed by quantitative trait loci (QTL), co-expression and regulatory binding motif analysis. The selected genes were constructed into a number of network modules based on predicted protein interactions including modules of phosphorylation activity, ubiquity activity, and several proteinase activities such as peroxidase, aspartic proteinase, glucosyltransferase, and flavonol synthase. All of these discovered modules are related to the salt tolerance mechanism of signal transduction, ion pump, abscisic acid mediation, reactive oxygen species scavenging and ion sequestration. We also predicted the three-dimensional structures of some crucial proteins related to the salt tolerance QTL for understanding the roles of these proteins in the network. Our computational study sheds some new light on the mechanism of salt tolerance and provides a systems biology pipeline for studying plant traits in general.

Citation: Wang J, Chen L, Wang Y, Zhang J, Liang Y, et al. (2013) A Computational Systems Biology Study for Understanding Salt Tolerance Mechanism in Rice. PLoS ONE 8(6): e64929. doi:10.1371/journal.pone.0064929

Editor: Ying Xu, University of Georgia, United States of America

Received: January 9, 2013; **Accepted:** April 19, 2013; **Published:** June 7, 2013

Copyright: © 2013 Wang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors are grateful to the support of National Natural Science Foundation of China (grants 61073075 and 61272207), the Science-Technology Development Project from Jilin Province of China (grant 20120730), National Institutes of Health (R21/R33-GM078601 and R01-GM100701), National Science Foundation (#DBI-0421620), and Department of Energy (DE-SC0004898). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ycliang@jlu.edu.cn (YL); xudong@missouri.edu (DX)

Introduction

Salinity is one of agriculture's most crucial problems in large parts of the world [1]. Rice (*Oryza sativa* L.), which provides a major food source for about half of the global population, is considered as the most important cereal crop in agriculture, but it is salt susceptible [2]. Soil salinity is a major abiotic stress, which limits rice production in about 30% of the rice-growing area worldwide [3,4]. Under the heavy pressure of global population explosion and global climate change, studying rice salt tolerance is of high importance. Genetic improvements leading to salt tolerance of cereal crops in molecular breeding could help maintain stable global food supply [5]. Some traditional cultivars and landraces have been identified as tolerant to abiotic stresses, despite their undesirable agronomic traits such as tall plant stature, photosensitivity, poor grain quality and low yield. For example, Pokkali, an Indian landrace, can maintain high K⁺/Na⁺ ratio in shoot in a high salinity environment, and it could be a donor of salt-tolerance strains in breeding programs. FL478, an F2-derived F8, inherited the salt tolerance property in recombinant inbred lines from parents Pokkali and IR29. FL478 is also an improved indica cultivar used as a salt-susceptibility standard [6].

With years of continuous exploration, some general molecular mechanisms of salt tolerance in plants have been revealed. The

high-salinity environment mainly disrupts the ionic and osmotic equilibrium of cells, and as a result, genes in several pathways are activated in response to high sodium concentration. Pathways related to ion pumps [7], calcium [8], SOS pathway [9], ABA (abscisic acid) [10], mitogen-activated protein kinases [11], glycine betaine [12], proline [13], reactive oxygen species [14], and DEAD-box helicases [15] are of significance in high salinity environment. They play different roles in maintaining high K⁺/Na⁺ ratio, synthesizing and segregating ions, and controlling ion concentration [16]. The genes and transcription factors that encode or regulate these components often demonstrate irregular activities in a high salinity environment. At the cell level, the most significant activities in dealing with excessive salt in plants is pumping ions out of a cell to keep the ion equilibrium, while the vacuole located in the cell helps store some ions. In salt-resistant detoxifying mechanisms, especially sequestration by vacuole [17], many salt tolerance genes with high level of activities in a high salinity environment are related to vesicle, membrane and ion transport. For example, H⁺-ATPase as a proton pump on cytoplasmic vesicle maintains the ion equilibrium of the cell by pumping H⁺ to the vacuole to retain pH and transmembrane proton gradient [18]; Na⁺ transporter plays an important role in maintaining high Na⁺/K⁺ ratio in various tissues [19,20]. However, the global picture of salt tolerance mechanisms,

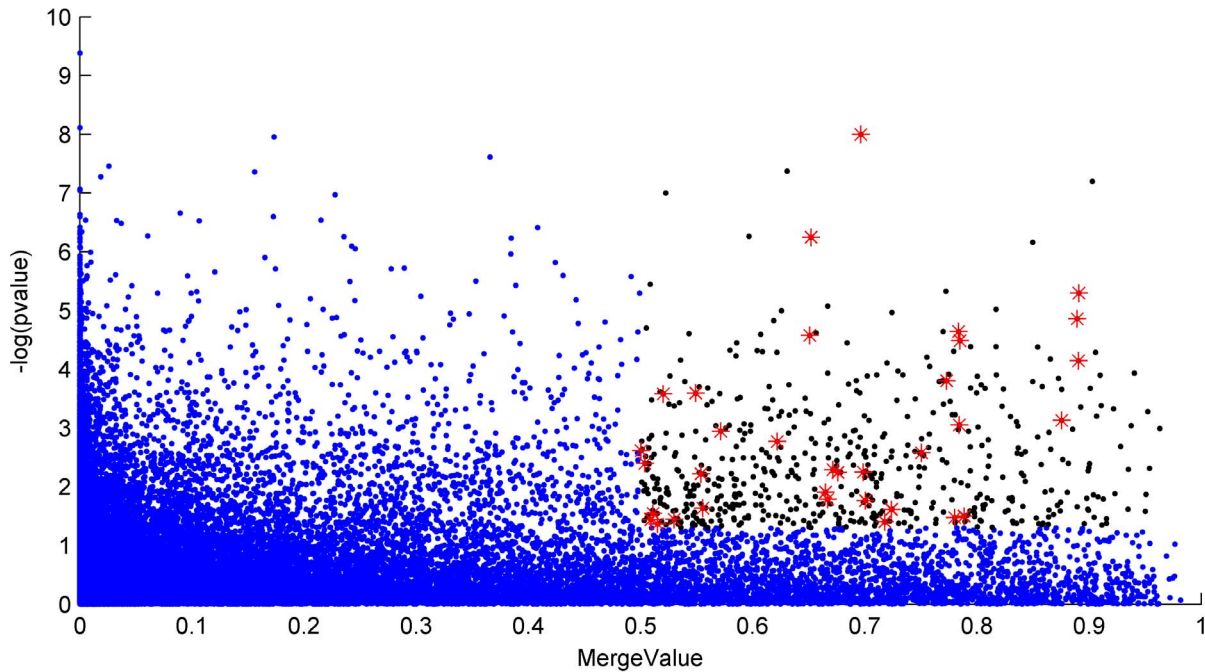


Figure 1. Improved volcano plot of GSE14403. The horizontal axis represents *MergeValue* obtained by bootstraps SVM-RFE. The vertical axis shows the $-\log(\text{p-value})$ from t-test. Black dots indicate selected probes with *MergeValue* threshold of 0.5 and t-test p-value threshold of 0.05. Red stars indicate selected probes mapped on QTL region. Blue dots indicate unselected probes. doi:10.1371/journal.pone.0064929.g001

especially rice-specific salt tolerance mechanisms is still unclear; for example, how ABA induces H_2O_2 control and how a plant transduces signals in response to salt tolerance are largely unknown.

Multiple sources of data can enhance the understanding of salt tolerance. The genetic variations of different rice responses to salt stress may shed some light on the roles of various genes in salt tolerance. The availability of rice genome sequencing [21,22] further paved the way for in-depth study of rice salt tolerance. *Oryza sativa* microarray gene expression data have provided information on regulatory networks of salinity response. Kawasaki et al. analyzed the initial phase of salt stress in rice based on gene expression profiles [23]. Huang et al. identified a zinc finger protein named DST that regulates drought and salt tolerance in rice [24]. Zhang et al. studied OsGAPC3 over-expression in rice tolerance [25]. Mito et al. found that expression of DREB- and ZAT- related genes might be involved in the salt tolerance of the AtMYB102 chimeric repressor line [26]. Schmidt et al. examined transcription factors like heat shock factors (HSFs) in response to salinity environment and they characterized OsHsfC1b as playing a role in ABA-mediated salt stress tolerance in rice [27]. Nevertheless, these studies were mainly focused on a single gene or some isolated genes, and they lack systems-level understanding of the global molecular mechanism of salt tolerance given that salt resistance reacts in a coordinated and effective manner. In view of these findings, we conducted a systems-level study to fill the gap between isolated genes and the global mechanism of salt tolerance.

Among tens of thousands of genes in microarray data, it is challenging to choose the set of genes that are most relevant to salt tolerance [28,29]. Biologists often use a volcano plot method, which reflects both fold of change and its statistical significance at the same time in a heuristic fashion [30]. However, such a method may not be sufficient to discover some complex relationships between genes and a certain phenotype, trait, or condition [31].

Some statistical methods for clustering and classification are extensively used to deal with this problem [32,33]. Several machine-learning methods, such as random forest [34] and SVM-RFE (support vector machine recursive feature elimination) [35] were also developed for this purpose. RFE is commonly used for feature selection, and it eliminates features iteratively until getting the minimum subset of features. By combining SVM with the RFE procedure, SVM-RFE becomes an effective method in selecting and ranking genes in microarray data analyses [36,37,38]. In this study, we improved the volcano plot method using bootstrapping SVM-RFE to select informative genes related to salt tolerance from microarray datasets.

There are both challenges and advantages in analyzing rice data. For crops, there are typically limited experimental data available, and little bioinformatics work has been done on these data. On the other hand, crops, especially rice, have rich data related to traits, such as Quantitative Trait Locus (QTLs). QTLs are stretches of DNA containing or linked to genes that underlie a quantitative trait. It is a classical and widely used breeding method in identifying the actual genes underlying a trait in breeding experiments. With the availability of rice genome sequence, QTLs provide useful relationships between genes in a QTL region and its corresponding trait [39,40]. In this work, we used QTLs in validating selected informative gene sets. We also used predicted protein-protein interactions to build a protein network of informative genes. Some crucial genes in the network with QTL evidence were studied by protein structure prediction, co-expression and gene regulatory motif analysis. Our computational study provides useful hypotheses for studying salt tolerance and may help improve molecular breeding of rice in salinity.

Table 1. GO term enrichment analysis on gene selected from microarray by AgriGO.

GO term	Ontology	Description	Number in selected gene set	Number in Background	p-value	FDR
GO: 0034641	P	cellular nitrogen compound metabolic process	17	459	2.40E-05	0.011
GO: 0005506	F	iron ion binding	19	432	7.10E-07	0.0002
GO: 0016491	F	oxidoreductase activity	32	1141	4.30E-06	0.0006
GO: 0043169	F	cation binding	51	2582	0.00014	0.01
GO: 0043167	F	ion binding	51	2584	0.00015	0.01
GO: 0020037	F	heme binding	9	205	0.00055	0.031
GO: 0046906	F	tetrapyrrole binding	9	217	0.00082	0.039
GO: 0031982	C	Vesicle	131	7454	3.40E-06	0.0001
GO: 0016023	C	cytoplasmic membrane-bounded vesicle	131	7445	3.20E-06	0.0001
GO: 0031988	C	membrane-bounded vesicle	131	7445	3.20E-06	0.0001
GO: 0031410	C	cytoplasmic vesicle	131	7454	3.40E-06	0.0001
GO: 0031224	C	intrinsic to membrane	23	820	8.10E-05	0.002
GO: 0016021	C	integral to membrane	21	804	0.0004	0.0083

Ontology "P" indicates Biological Process, Ontology "F" indicates Molecular Function, and Ontology "C" indicates Cellular Component. "Number in selected gene set" is the number of genes in the query gene list. "Number in background" is the number of genes in the proteome. P-value represents the statistical significance of the gene enrichment test. FDR means False Discovery Rate.

doi:10.1371/journal.pone.0064929.t001

Results

Using the microarray data, GSE14403, to compare salinity susceptible and resistant rice genotypes with the Gene Expression Omnibus (GEO), we chose the threshold as 0.05 in t-test p-value and 0.5 in *Merge Value* (as described in the Method section) to obtain 556 probes in the improved volcano plot in Figure 1. We assume that many of these 556 genes are related to salt tolerance, and they are listed in Table S1. Table 1 lists the gene enrichment result using AgriGO [41], a plant-specific GO term enrichment analysis. In molecular functions, the chosen genes are over represented in the categories of iron binding, cation binding, ion binding and heme binding—all of which may be active due to the high ion concentration in salinity. The significant behavior of these genes in oxidoreductase activity may be related to electron transport in complex chemical reactions that balances the charges during ion transport. The oxidoreductase activity may also be related to reactive oxygen intermediates (ROI) that are produced in response to oxidative stress due to a water deficit during salinity stress [14]. ROI can seriously disrupt normal metabolism through oxidative damage on lipids [42,43], proteins and nucleic acids [42,44]. The increased oxidoreductase activity is consistent with known activation of the antioxidative enzymes such as catalase (CAT), ascorbate peroxidase (APX), guaiacol peroxidase (POD),

glutathione reductase (GR), and superoxide dismutase (SOD) under salt stress in plants [16]. In biological processes, the cellular nitrogen compound metabolic process is over represented. As proline and glycine betaine accumulate under stress, they are correlated with osmotic adjustment to improve plant salinity tolerance [45]. Proline is also involved in scavenging free radicals, stabilizing subcellular structures and buffering cellular redox potential under stresses. Polyamines can be synthesized under salt tolerance [46]. In this sense, we speculate these nitrogen-containing compounds may be synthesized in "the cellular nitrogen compound metabolic process." In cellular components, according to the gene enrichment analysis, most of the chosen genes are related to vesicle and membrane, which is consistent with the detoxifying mechanism of salt resistant genotypes, especially in sequestration by vacuole [17]. It is plausible to infer that some of these chosen proteins on membranes act as transporting ions outside the cell or to the vacuole to maintain pH, transmembrane proton gradient [18], and high K⁺/Na⁺ ratio [19].

In order to assess the performance of improved volcano plot, we developed a Microarray-QTL test by using the QTL information as a criterion to evaluate the reliability of chosen genes. As shown in Table 2, our chosen genes are compared with the QTL regions mapped in the whole genome and these genes show high statistical

Table 2. Evaluation of choosing salt-tolerance genes by QTLs.

Methods	Sample Size	Original QTL region		Flanking region with length of the QTL region			
		Number of Hits	Percentage of Population	Microarray-QTL test p-value	Number of Hits	Percentage of sample	Microarray-QTL test p-value
Improved Volcano Plot	556 selected	34	6.12%	4.21e-7	94	16.91%	7.44e-18

The Number of hits and the percentage of samples depict the number and percentage of the chosen genes that can be mapped into the (extended) QTLs regions. The Microarray-QTL test p-value means the statistical significance of these chosen genes according to the QTLs.

doi:10.1371/journal.pone.0064929.t002

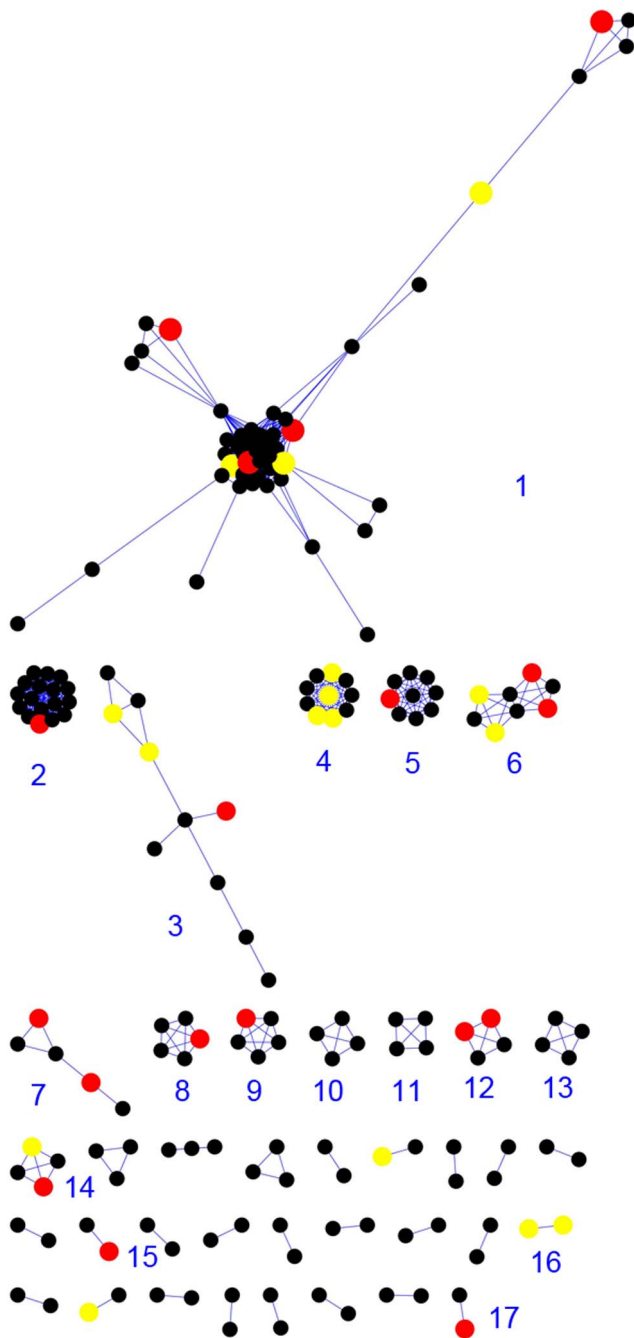


Figure 2. Salt-tolerance protein interaction modules that are related to QTLs. Red nodes represent the proteins located in QTL regions. Yellow nodes represent the proteins located in flanked regions with the lengths of the QTL regions.
doi:10.1371/journal.pone.0064929.g002

significance in related QTL regions by our Microarray-QTL test. We also compared the performance of other feature selection methods in Document S1.

We constructed a rice salt tolerance protein interaction network using the 556 genes selected by improved volcano plot methods as the nodes and protein-protein interaction data from DIPOS as the edges [47]. By merging the isoforms, the network contains 189 nodes and 705 edges, as visualized by Cytoscape [48] in Figure 2.

By analyzing the constructed network, we identified 17 modules with nodes located in QTL and flanked QTL region. Including

flanked regions could help give some tolerance to the errors in QTL mapping [49]. Table 3 shows the functions of these identified modules related to known salt tolerance mechanisms.

The largest module in this network contains 51 genes of 47 merged nodes and 372 edges. Figure 3 depicts the radical layout of this module and Table S2 shows the node annotation in detail. The GO enrichment analysis reveals related salt tolerance activities in Table 4, and protein phosphorylation represents the most significant function. It is known that protein phosphorylation plays a vital role in ion homeostasis under salinity stress in Arabidopsis [50,51]. Under the salinity stress, phosphorylation often becomes active in signaling pathways; for example, MAPK transduces salt and other abiotic stress signals. In rice, Os-MAPK5 as a kinase can be triggered by salt, drought, wounding, cold, and ABA, resulting in an increase in tolerance to these abiotic stresses [10]. The Na⁺/H⁺ antiporter SOS1 mediates Na⁺ efflux. SOS2, a Ser/Thr protein kinase with N-terminal kinase catalytic domain regulates the activity of SOS1. SOS3, which senses salt stress-induced Ca²⁺ signature, thereby activating SOS2 to transduce the salt stress signal. Abscisic acid-mediated phosphorylation also plays a significant role in many activities within cytoplasmic proteins in rice under salinity stress [52]. The largest module that we identified includes these genes and others related to phosphorylation in nucleotide binding and kinase activities.

At the transcription level, we also checked whether the 51 genes in the largest module are transcriptionally co-regulated by examining if the promoter regions of these genes share conserved motifs as the regulatory elements. Three candidate motifs are predicted, and only one motif is validated by sequence comparison with known *cis* regulatory motifs in the PLACE database [53] as shown in Figure 4 and Table S3. This conserved motif is detected as “TCTCTCTCT”, the CTRMCAMV35S motif, which is a CT-rich motif found in a 60-nt region downstream of the transcription start site of the CaMV 35S RNA, and can enhance gene expression [54]. We also formed arbitrary reference gene sets selected randomly in whole genome scale and Chi-Square test on the validated motif showed significant statistical significance at p-value of 4.72e-13 (Table S4). We also checked the co-expression among the genes in this module. The averaged Pearson correlation coefficient of expression profiles inside the module is 0.402, comparing to 0.241 between genes in the module and randomly selected 51 genes in whole genome (Document S2). The motif analysis and co-expression analysis provide some support that this module is likely co-regulated.

At the whole genome level, we mapped all 556 selected genes, QTL regions and extended QTL regions by their individual positions on the rice genome of MSU Rice Genome Annotation (Osa1) Release 6 [55] in Figure 5 using Circos [56]. Some selected genes are located in the QTL and extended QTL regions. We also mapped the 51 genes in the largest module on the genome, together with their interactions.

We mapped the 51 genes of the largest module to the KEGG pathway (<http://www.genome.jp/kegg/pathway.html>). Three genes can be mapped to their Arabidopsis orthologs in the Plant-Pathogen Interaction Pathway [57,58], as depicted in Figure 6. In this environmental adaptation pathway, CNGCs (Os02g0789100) is a cyclic nucleotide gated channel, CDPK (Os01g0622600) is a calcium-dependent protein kinase, and CaMCM1 (Os01g0505600) is the calcium-binding protein CML. All three of these proteins interact with each other and cooperate together in response to calcium ion signaling.

One of the most interesting genes is LOC_Os01g52640.3, which is a hub gene in the largest module and overlaps with a QTL region. This gene corresponds to a hypothetical protein

Table 3. Modules of salt-tolerance network and their potential roles in salt tolerance.

NO	Function	Nodes	Edges	# of QTL hits	# of Extended QTL hits	Salt Tolerance Mechanism	QTL hits	Extended QTL hits
1	phosphorylation	47	372	4	7	Signal Transduction/ion homeostasis/ABA	LOC_Os01g611070.1, LOC_Os07g09420.1, LOC_Os07g12240.1, LOC_Os01g59580.1	LOC_Os01g70840.1/LOC_Os01g70840.2, LOC_Os09g29510.1, LOC_Os01g52640.3
2	cytochrome P450	17	136	1	1	ABA	LOC_Os09g26960.1	
3	NA	10	11	1	3	NA	LOC_Os07g08840.1	LOC_Os09g20220.1, LOC_Os01g70990.1
4	peroxidase	10	45	0	4	Antioxidate		LOC_Os06g33080.1, LOC_Os06g35480.1, LOC_Os06g33100.1, LOC_Os06g32990.1
5	aspartic proteinase nepenthesin	9	36	1	1	Agreement	LOC_Os01g64840.1	
6	Ubiquitin	8	19	2	4	ABA/degradation pathway	LOC_Os01g64620.1, LOC_Os01g60730.1/ LOC_Os01g60730.2	LOC_Os09g31031.2, LOC_Os09g27930.1
7	starch	5	5	2	2	Agreement	LOC_Os04g33920.1, LOC_Os01g63810.1	
8	glucosyltransferase	5	10	1	1	Agreement	LOC_Os07g10190.1	
9	flavonol synthase	5	10	1	1	ROS-scavenging	LOC_Os01g61610.1/LOC_Os01g61610.2/ LOC_Os01g61610.3	
10	glycosyl hydrolases family	4	6	0	0	Agreement		
11	MYB family transcription factor	4	6	0	0	Agreement		
12	gibberellin receptor GID1L2	4	6	2	2	Agreement	LOC_Os07g06840.1, LOC_Os07g06860.1	
13	O-methyltransferase	4	6	0	0	Sodium sequestration		
14	Phosphatidylethanolamine-binding	4	6	1	2	Cell signalling	LOC_Os04g33570.1, LOC_Os04g41130.1	
15	aldo/keto reductase family	2	1	1	1	Antioxidate	LOC_Os04g37490.1	
16	cytokinin dehydrogenase	2	1	0	2	ABA		LOC_Os01g56810.1, LOC_Os06g35650.1
17	cysteine synthase/indole-3-glycerol phosphate lyase	2	1	1	1	Agreement	LOC_Os06g36850.1	

The number of members and interactions of the module are shown in column of "Nodes and Edges." "# of QTL hits" and "# of Extended QTL hits" depict the number of module members that can be mapped into the (extended) QTLs regions. "Salt tolerance mechanism" shows the process or essential component that could participate in salt tolerance response. "Agreement" means the component is only consistent with suggestions in other studies but there is no direct evidence to link it to known salt tolerance mechanism. "QTL hits" and "Extended QTL hits" illustrate the member of module mapped into the (extended) QTLs regions.
doi:10.1371/journal.pone.0064929.t003

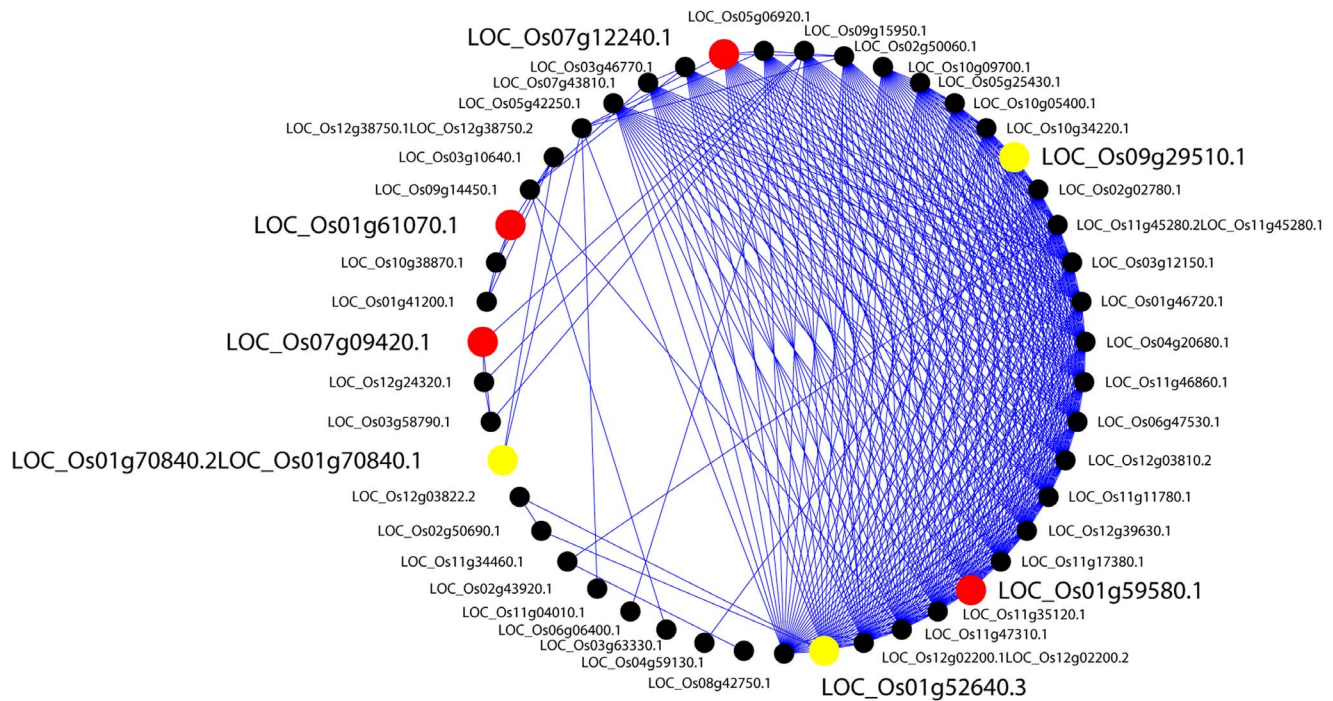


Figure 3. The largest module in the salt tolerance protein interaction network. Black nodes indicate genes covered by QTLs. Yellow nodes indicate genes covered by extended QTLs. doi:10.1371/journal.pone.0064929.g003

Os01g0725800, which interacts with 32 of the 51 proteins in the module. It contains four InterPro domains, namely, IPR000719, IPR001680, IPR011046, and IPR011009. IPR011009 domains can also be found in RIO kinase (IPR018935), a SPA1-related, serine/threonine-specific and tyrosine-specific protein kinase. This protein also has an ortholog in *Arabidopsis thaliana* as SPA4 (SPA1-RELATED 4), which is a binding protein and a signal transducer. We applied MUFOLD [59,60] to predict the structure for LOC_Os01g52640.3. Using the identified templates of 2GNQ, 3EMH, and 3DM0 in PDB, we constructed the model for the protein region of 196–627 for the protein with the length of 432, as shown in Figure 7. The protein structural model contains the WD40 structure motif repeats, each with a typtophan-aspartic (W-D) dipeptide termination. As WD40 proteins often play important roles in signal transduction and transcription regulation [61], the structure prediction suggests that this protein may be related to signal transduction in the salt resistance process.

We also applied MUFOLD for structure predictions of hub proteins LOC_Os01g59580.1 and LOC_Os01g46720.1, each of which has a node degree of 30. According to the remote homology detection, LOC_Os01g59580.1 has a template 2QKW, which is in the process of phosphorylation (GO: 0006468), and plays a role as ATP binding (GO: 0005524), protein binding (GO: 0005515) and protein serine/threonine kinase activity (GO: 0004674) [62]. LOC_Os01g59580.1 may have these activities as well. LOC_Os01g46720.1 has templates 1PKD, 3EZR, 2W06 and 2W17, which are involved in anaphase-promoting complex-dependent, proteasomal ubiquitin-dependent protein catabolic process (GO: 0031145), and cyclin-dependent protein kinase activity (GO: 0004693). The details of predicted structural models are described in Document S3.

Besides exploring the largest module related to phosphorylation activity in a systems biology point of view, we also explored the other 16 identified modules in a similar fashion. Among these

modules, Module 3 has divergent functions, but other modules converge to consistent functions related to salt tolerance. Module 2 of cytochrome P450 [63], Module 6 of ubiquitin activity [64,65], and Module 16 of cytokinin dehydrogenase [66,67] are related to ABA mediated in salt tolerance. Module 4 of peroxidase precursor [68], Module 9 of flavonol synthase [69,70], and Module 15 of the aldo/keto reductase (AKR) family [71] are all related to ROS-scavenging in salt tolerance. Module 13 of O-methyltransferase [72] is associated with sodium sequestration. Some previous experimental studies on the abiotic stress of plants based on gene expression patterns also linked salt tolerance to aspartic proteinase nepenthesin [73] in Module 5, starch [74] in Module 7, glucosyltransferase [73] in Module 8, glycosyl hydrolases [75] in Module 10, MYB family [73] in Module 11, gibberellin receptor GID1L2 [76] in Module 12, phosphatidylethanolamine-binding [77] in Module 14, and cysteine synthase [78] in Module 17. A detailed analysis on each of these 16 modules can be found in Document S4.

Discussion

In this paper, we focus on the salt tolerance mechanism of root tissue in rice. As there are commonly three samples in just one genotype of one condition on microarray experiments in contrast to tens of thousands of probe sets, it is a great challenge to determine feature selection on this small-sample, but high-dimension data. Classical statistical feature selection methods, such as t-test assume the samples follow some specific distribution as its hypothesis; however, the limited number of samples narrows the usage of these statistical methods. From the feature selection prospective, the volcano plot method uses two dimensions of fold change and t-test p-value to select genes in microarray analysis. It is a fast, simple, and widely used method. However, the fold change of each differential expression gene does not necessarily

Table 4. AgriGO term enrichment analysis on the 51 genes in the largest module.

GO term	Ontology	Description	Number of genes	Number in Background	p-value	FDR
GO: 0016310	P	Phosphorylation	9	1080	4.9e-06	0.00033
GO: 0006468	P	protein amino acid phosphorylation	8	887	8.8e-06	0.00033
GO: 0006796	P	phosphate metabolic process	9	1206	1.2e-05	0.00033
GO: 0006793	P	phosphorus metabolic process	9	1206	1.2e-05	0.00033
GO: 0043687	P	post-translational protein modification	8	1236	8.8e-05	0.002
GO: 0034641	P	cellular nitrogen compound metabolic process	5	459	0.00016	0.0027
GO: 0006464	P	protein modification process	8	1359	0.00017	0.0027
GO: 0043412	P	macromolecule modification	8	1406	0.00021	0.003
GO: 0009069	P	serine family amino acid metabolic process	5	538	0.00034	0.0043
GO: 0044267	P	cellular protein metabolic process	8	2166	0.0031	0.033
GO: 0006520	P	cellular amino acid metabolic process	5	918	0.0034	0.033
GO: 0044106	P	cellular amine metabolic process	5	918	0.0034	0.033
GO: 0019538	P	protein metabolic process	9	2770	0.0041	0.036
GO: 0009308	P	amine metabolic process	5	1002	0.0049	0.041
GO: 0001883	F	purine nucleoside binding	9	1171	9.3e-06	0.00012
GO: 0001882	F	nucleoside binding	9	1171	9.3e-06	0.00012
GO: 0000166	F	nucleotide binding	11	1686	5.2e-06	0.00012
GO: 0005524	F	ATP binding	9	1071	4.6e-06	0.00012
GO: 0030554	F	adenyl nucleotide binding	9	1171	9.3e-06	0.00012
GO: 0032559	F	adenyl ribonucleotide binding	9	1074	4.7e-06	0.00012
GO: 0032555	F	purine ribonucleotide binding	9	1218	1.3e-05	0.00012
GO: 0032553	F	ribonucleotide binding	9	1218	1.3e-05	0.00012
GO: 0017076	F	purine nucleotide binding	9	1317	2.3e-05	0.00019
GO: 0016301	F	kinase activity	9	1464	5.1e-05	0.00039
GO: 0004672	F	protein kinase activity	7	1102	0.00026	0.0018
GO: 0016773	F	phosphotransferase activity, alcohol group as acceptor	7	1238	0.00051	0.0032
GO: 0004674	F	protein serine/threonine kinase activity	6	949	0.00069	0.004
GO: 0016772	F	transferase activity, transferring phosphorus-containing groups	9	2197	0.00091	0.005
GO: 0005488	F	Binding	21	8681	0.0015	0.0076

One gene can be in multiple GO families. In "Ontology", "P" indicates biological process and "F" indicates molecular function.
doi:10.1371/journal.pone.0064929.t004

reveal the nature of biological meaning. In this case, using machine-learning methods could be a good alternative in microarray analysis. The improved volcano plot method using some specific criteria like *MergeValue* based on an SVM-RFE

procedure could improve the performance. The improved method used all three salt resistant genotypes as a whole to mine the common pattern of salt tolerance, which helps overcome the disadvantage of a limited number of samples. The improved

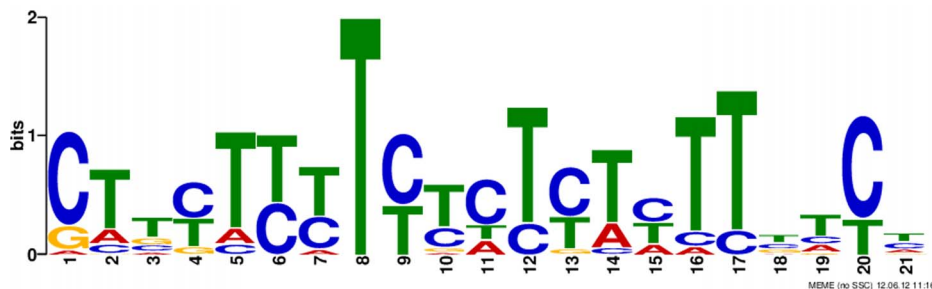


Figure 4. Detected motif in the upstream sequences of the largest module.
doi:10.1371/journal.pone.0064929.g004

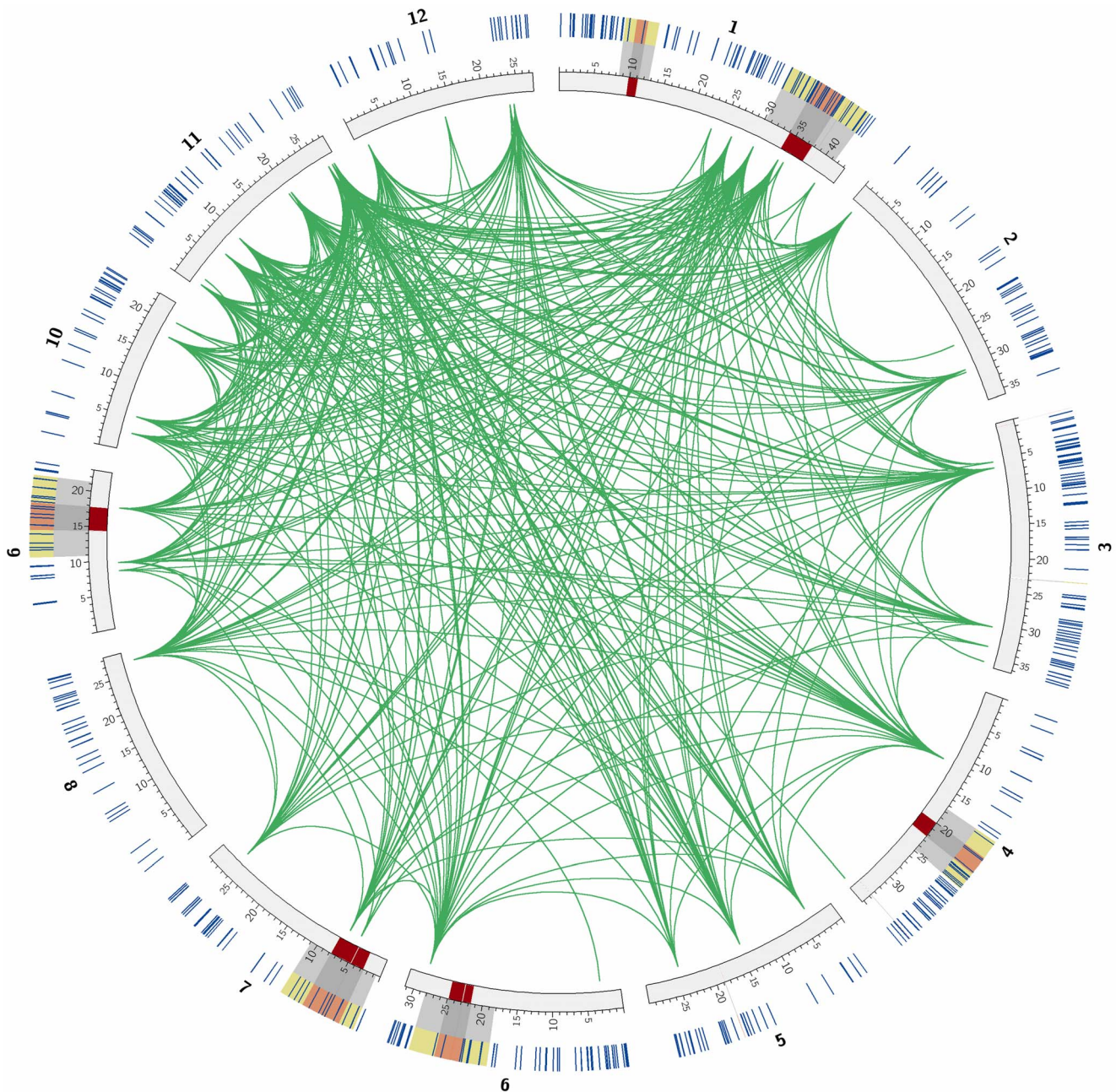


Figure 5. The whole genome mapping of selected genes and their interactions. Blue lines in the outer circle represent all the 556 selected genes. Red regions on the chromosome are the QTLs while the grey and yellow regions are the extended QTLs with one QTL length at each side of the flanking region. Inside the circle, green links show the protein-protein interactions among 51 genes in the largest module.
doi:10.1371/journal.pone.0064929.g005

method also used a bootstraps approach to make the feature selection more robust.

We incorporated the QTL information with transcription profiles to identify genes related to drought response. For a given QTL, there may be 25–30 genes per cM (~270 kbp in rice) [79]. Given so many possible genes in a QTL region associated to a phenotype, the proposed Microarray-QTL test, which used the same mechanism as GO term enrichment analysis, could help evaluate the relative relevance of these QTL genes to the phenotype quantitatively. Furthermore, we also combined predicted protein-protein interactions, protein structure prediction and gene regulatory motif analysis in studying potential genes

related to salt tolerance. Such a systems approach is powerful in providing high-confidence predictions of salt-tolerant genes. Our study may provide richer and more concise predictions than a study done by Cotsaftis et al. [80], which only used the expression level of the gene probes in transcript profiling to predict salt-tolerant genes.

Walia et al. [5] summarized the following components in the salinity response based on their microarray study: 1) adaptive response, 2) non-adaptive response, 3) response to salt injury, and 4) heritable responses conferring tolerance. While these general categorizations are important, they do not provide a detailed mechanism, especially in terms of genes involved and these

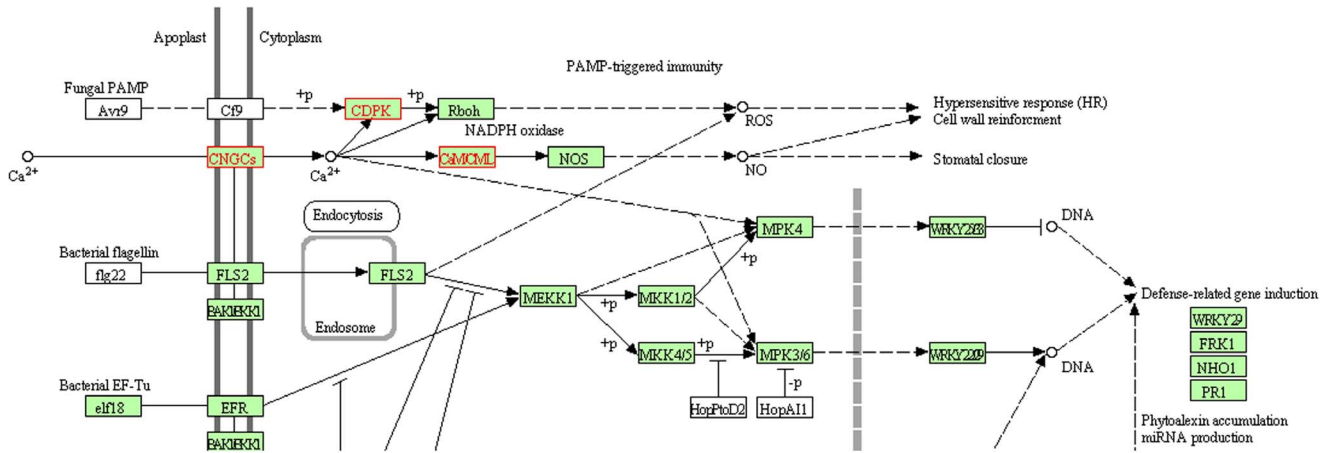


Figure 6. Part of the Arabidopsis Plant-Pathogen Interaction pathway in KEGG (http://www.genome.jp/kegg-bin/show_pathway?ath04626+AT1G18210+AT3G51850+AT1G01340), where white boxes indicate that no genes have been assigned, green boxes have known genes in Arabidopsis, and boxes highlighted in red show the three mapped genes in the largest module. doi:10.1371/journal.pone.0064929.g006

processes. Our study serves as an attempt to fill this gap. According to our constructed network, one kernel module of phosphorylation activity is detected. The role of phosphorylation in abiotic stress has been actively studied in recent years in addition to its relationship to salinity stress [81], cold stress [82] and heat stress [83]. Our result shows the central role of phosphorylation and redox action in salt tolerance, with implication of activity in signal transduction and oxidation. As our protein interaction networks were constructed from predicted interaction, our predictions of network modules may have significant false positives, which need further biological experiments to validate.

Our study may provide some useful hypotheses for researchers to design experiments for studying salt resistance and some

guidance for molecular breeders to improve traits. Since some key proteins have been predicted and mapped to QTLs in our study, which means researchers could conduct experiments to clone and validate these genes. We plan to apply our computational pipeline to study other traits and other species.

Materials and Methods

Data Source

We obtained rice microarray data from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>), which contains 14 datasets, 182 platforms, 5,210 samples and 374 series

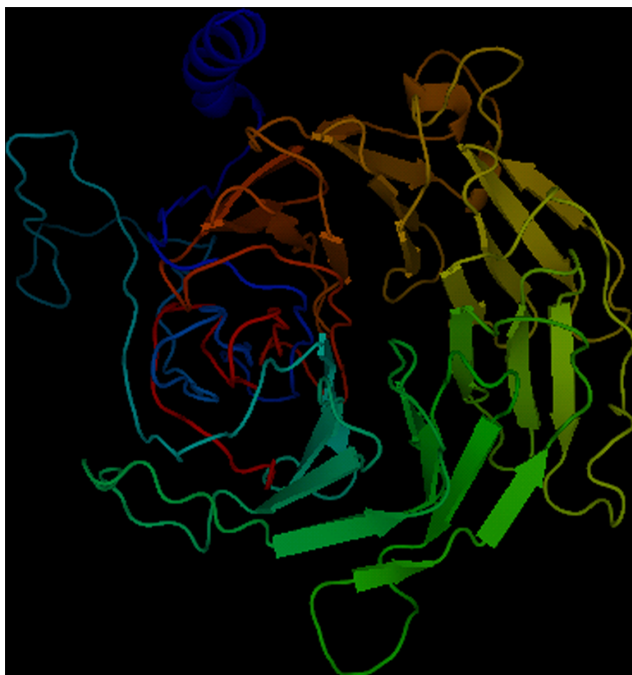


Figure 7. Predicted structural model of protein Os01g0725800. doi:10.1371/journal.pone.0064929.g007

Table 5. GSE14403 resistant samples description.

No	Sample	Genotype	Repeats	Stress condition	Class Label
1	GSM359902	FL478	1	Control	+1
2	GSM359903	FL478	2	Control	+1
3	GSM359904	FL478	3	Control	+1
4	GSM359905	FL478	1	Salt-treated	-1
5	GSM359906	FL478	2	Salt-treated	-1
6	GSM359907	FL478	3	Salt-treated	-1
7	GSM359913	IR63731	1	Control	+1
8	GSM359914	IR63731	2	Control	+1
9	GSM359915	IR63731	3	Control	+1
10	GSM359916	IR63731	1	Salt-treated	-1
11	GSM359917	IR63731	2	Salt-treated	-1
12	GSM359918	IR63731	3	Salt-treated	-1
13	GSM359919	Pokkali	1	Control	+1
14	GSM359920	Pokkali	2	Control	+1
15	GSM359921	Pokkali	3	Control	+1
16	GSM359922	Pokkali	1	Salt-treated	-1
17	GSM359923	Pokkali	2	Salt-treated	-1
18	GSM359924	Pokkali	3	Salt-treated	-1

doi:10.1371/journal.pone.0064929.t005

Table 6. QTLs related to salt tolerance in rice from the Gramene QTL library.

No.	QTL name	Chr	Mapped positions	Note
1	AQEM001	1	33,956,950–37,713,775	
2	AQEM006	1	9,820,009–11,232,822	
3	AQEM008	1	33,956,950–37,713,775	Same as 1
4	AQGR001	1	38,530,957–38,531,467	
5	AQGR002	3	22,798,284–22,830,744	
6	AQCL001	3	484,860–485,333	
7	AQCL002	4	33,663,984–33,664,487	
8	AQEM009	4	19,928,370–22,355,854	
9	AQCL003	5	18,874,932–18,875,558	
10	AQEM002	6	21,605,889–24,919,236	
11	AQCL004	6	22,862,400–22,862,821	
12	AQEM003	7	4,573,316–7,739,951	
13	AQEM004	7	2,633,784–4,575,215	
14	AQEM010	7	2,633,784–4,575,215	Same as 13
15	AQEM005	7	2,633,784–4,575,215	Same as 13
16	AQEM011	7	2,633,784–4,575,215	Same as 13
17	AQEM007	9	14,362,062–17,837,010	

doi:10.1371/journal.pone.0064929.t006

of *Oryza sativa*. Among these datasets, we used GSE14403 submitted by Ute Baumann January 13, 2009 and last updated August 2, 2012 to analyze salt tolerance. Unlike other datasets such as GSE3053 and GSE13735, this dataset contains the largest size of samples ever gathered related to salt tolerance of roots. The data were generated from Affymetrix Rice Genome Array (GPL 2025 in GEO), which contains 57,381 probes and each probe corresponded to an individual gene.

In the dataset GSE14403, we used salt resistant genotypes FL478, Pokkali and IR63731, and salt susceptible genotype IR29 under control and salinity-stressed conditions during vegetative growth, which ranged from GSM359902 to GSM359924. We merged these three salt-tolerant plants together as the salt-tolerant group. In the microarray experiments that collected the data [80], seedlings were cultured in sand and irrigated with a nutrient solution for 22 days (salt-treated) and 30 days (control) after germination, respectively. Salinity treatment was applied by adding NaCl and CaCl₂ (5:1 molar concentration) in two steps over a period of 3 days (final electrical conductivity: 7.4 dS m⁻¹) to prevent osmotic shock. All plants were harvested on day 30. All the data were collected from the root tissue of the plants. Table 5 gives the specific description of the data used, and the class label (−1/1) is according to the different stress conditions. Table 6 shows 17 QTLs (13 unique QTLs) detected with salt tolerance in Gramene (<http://www.gramene.org/qtl/>). We mapped these QTLs regions using Gramene annotation of the rice genome of MSU Rice Genome Annotation (Osa1) Release 6 [55].

We obtained predicted protein-protein interaction data from Database of Interacting Proteins in *Oryza Sativa* (DIPOS) (<http://csb.shu.edu.cn/dipos/>) [47]. This database used two different but complementary methods, i.e., interologs and domain interactions based methods to predict protein interactions for rice. DIPOS contains 14,614,067 pairwise interactions among 27,746 proteins, covering about 41% of the whole *Oryza sativa* proteome.

Data Preprocessing

The preprocessing step intended to overcome the noises, including missing probes and mislabeled probes. We conducted the analysis using Bioconductor [84]. We used Bioconductor's affy package for estimation of expression values by Robust Multi-chip Average (RMA) [85]. The RMA procedure consists of three steps: a background adjustment, quantile normalization and summarization. As there are 123 probe sets designed for control in this microarray, the dataset excluded these probe sets and 57,258 valid probe sets were obtained for further analysis.

Choice of Gene List using Improved Volcano Plot

We proposed an improved volcano plot method to choose genes in this dataset. The improved method has a new measure *MergeValue* for selecting genes. The detail of this method is described in Document S5.

Validate Gene List from QTL

Firstly, we identified QTLs zone in the genome using Gramene QTL database (<http://www.gramene.org/qtl/>), where there are 17 QTLs ranging from chromosomes 1,3,4,5,6,7,9. Secondly, we mapped the candidate genes detected by the microarray probe sets to the genome. Thirdly, we defined a criterion named Microarray-QTLs (Eq.(1)) to evaluate the statistical significance of chosen genes related with the specific trait.

$$P(X=k) = \frac{\binom{m}{k} \binom{N-m}{n-k}}{\binom{N}{n}} \quad (1)$$

where N is the total number of valid genes in the microarray, m is number of valid genes covered by QTLs, n is the number of chosen genes by microarray feature selection method, and k is the number of chosen genes that are covered by QTLs. The Microarray-QTL test follows a Hypergeometric distribution, and the p-value reveals significance of chosen genes related with specific QTLs.

Motif Analysis on Promoters

To determine whether chosen genes are transcriptionally co-regulated, a motif analysis could help to find conserved motif regulatory elements in their promoters. We first used MEME [86] to predict motifs on the upstream region of 1,000 bps from the translation start site of the chosen genes. Then we used FIMO [87] to conduct a Chi-squared test on the significance of these motifs in comparison to randomly selected genes in the genome. We compared the identified motifs with known plant motifs from the PLACE database [53] using CompariMotif [88]. For each pair of compared motifs, if their similarity score is more than 4 and the percentage of their matched positions is more than 80%, these two motifs were considered identical.

Supporting Information

Table S1 556 genes selected by Improved Volcano Plot. (XLS)

Table S2 51 genes in the largest module. (XLS)

Table S3 Motif of the largest module discovered by PLACE database. (XLS)

Table S4 Chi-Square test on the validated motif of the largest module.

(XLS)

Document S1 Feature selection comparison of improved volcano Plot.

(DOC)

Document S2 Supplementary of Co-expression analysis.

(DOC)

Document S3 Supplementary of Protein Prediction by MUFOLD.

(DOC)

Document S4 Annotation for modules identified from the constructed network.

(DOC)

Document S5 Choice of gene list using improved volcano plot.

(DOC)

Author Contributions

Conceived and designed the experiments: JW YL DX. Analyzed the data: JW LC YW JZ. Wrote the paper: JW DX.

References

- Munns R (2002) Comparative physiology of salt and water stress. *Plant Cell Environ* 25: 239–250.
- Maas EV, Hoffman GJ (1977) Crop salt tolerance-current assessment. *Journal of Irrigation and Drainage Division, Proceedings of the American Society of Civil Engineers* 103: 115–134.
- Tanji KK (1990) Nature and extent of agricultural salinity. *Agricultural salinity assessment and management* 71: 1–17.
- Wu R, Garg A (2003) Engineering rice plants with trehalose-producing genes improves tolerance to drought, salt and low temperature. ISBN News Report <http://www.isb.vt.edu>.
- Walia H, Wilson C, Condamine P, Liu X, Ismail A et al. (2005) Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during the vegetative growth stage. *Plant Physiol* 139(2): 822–835.
- Bonilla P, Dvorak J, Mackill D, Deal K, Gregorio G (2002) RFLP and SSLP mapping of salinity tolerance genes in chromosome 1 of rice (*Oryza sativa* L.) using recombinant inbred lines. *Philippine Agricultural Scientist* 85: 68–76.
- Pons R, Cornejo M, Sanz A (2011) Differential salinity-induced variations in the activity of H⁺-pumps and Na⁺/H⁺ antiporters that are involved in cytoplasmic ion homeostasis as a function of genotype and tolerance level in rice cell lines. *Plant Physiology and Biochemistry* 49(12): 1399–1409.
- White PJ, Broadley MR (2003) Calcium in plants. *Annals of botany* 92(4): 487–511.
- Mahajan S, Tuteja N (2005) Cold, salinity and drought stresses: An overview. *Arch. Biochem. Biophys.* 444, 139–158.
- Teige M, Scheikl E, Eulgem T, Doczi R, Ichimura K, et al. (2004) The MKK2 pathway mediates cold and salt stress signaling in Arabidopsis. *Mol. Cell* 15(1), 141–52.
- Zhang T, Liu Y, Yang T, Zhang L, Xu S, et al. (2006). Diverse signals converge at MAPK cascades in plant. *Plant Physiol. Biochem.* 44, 274–283.
- Rhodes D, Hanson AD (1993). Quaternary ammonium and tertiary sulfonium compounds in higher-plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 44, 357–384.
- Thiery L, Leprince A, Lefebvre D, Ghars MA, Debabieux E, et al. (2004) Phospholipase D is a negative regulator of proline biosynthesis in Arabidopsis thaliana. *J. Biol. Chem.* 279, 14812–14818.
- Miller G, Shulaev V, Mittler R (2008) Reactive oxygen signaling and abiotic stress. *Physiologia Plantarum* 133(3): 481–489.
- Vashisht AA, Tuteja N (2006). Stress responsive DEAD-box helicases: A new pathway to engineer plant stress tolerance. *J. Phytochem. Photobiol.* 84, 150–160.
- Narendra Tuteja (2007) Mechanisms of High Salinity Tolerance in Plants. *Methods in Enzymology* 428: 419–438.
- Apse M, Aharon G, Snedden W, Blumwald E (1999) Salt tolerance conferred by overexpression of a vacuolar Na⁺/H⁺ antiporter in Arabidopsis. *Science* 285(5431): 1256–1258.
- Niu X, Narasimhan M, Salzman R, Bressan R, Hasegawa P, et al. (1996) NaCl Regulation of Plasma Membrane H⁺-ATPase Gene Expression in a Glycophyte and a Halophyte. *Plant Physiol* 111: 679–718.
- Ren Z, Gao J, Li L, Cai X, Huang W, Chao D et al. (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nature Genetics* 37(10): 1141–1146.
- Cotsafis O, Plett D, Shirley N, Tester M, Hrmova M (2012) A Two-Stage Model of Na⁺ Exclusion in Rice Explained by 3D Modeling of HKT Transporters and Alternative Splicing. *PLoS One* 7(7): e39865.
- Yu J, Hu S, Wang J, Wong G, Li S, et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp indica). *Science* 296(5565): 79–82.
- Yu J, Wang J, Lin W, Li S, Li H, et al. (2005) The Genomes of *Oryza sativa*: A history of duplications. *PLoS Biology* 3(2): 266–281.
- Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, et al. (2001) Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell* 13: 889–905.
- Huang X, Chao D, Gao J, Zhu M, Shi M et al. (2009) A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. *Genes & Development*. 23: 1805–1817.
- Zhang X, Rao X, Shi H, Li R, Lu Y (2011) Overexpression of a cytosolic glyceraldehyde-3-phosphate dehydrogenase gene OsGAPC3 confers salt tolerance in rice. *Plant Cell Tiss Organ Cult* 107: 1–11.
- Mito T, Seki M, Shinozaki K, Ohme-Takagi M, Matsui K (2011) Generation of chimeric repressors that confer salt tolerance in Arabidopsis and rice. *Plant Biotechnology Journal* 9: 736–746.
- Schmidt R, Schippers JH, Welker A, Mieulet D, Guiderdoni E et al. (2012) Transcription factor OsHsC1b regulates salt tolerance and development in *Oryza sativa* ssp. japonica. *AoB Plants*. pls011.
- Blum A, Langley P (1997) Selection of relevant feature and examples in machine learning. *Artificial Intelligence* 97(1–2): 245–271.
- Guoyon I, Elisseeff A (2003) An introduction to variable and feature selection. *Journal of Machine Learning Research* 3: 1157–1182.
- Cui X, Churchill G (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology* 4: 210.
- Liang Y, Zhang F, Wang J, Joshi T, Wang Y, et al. (2011) Prediction of Drought-Resistant Genes in Arabidopsis thaliana Using SVM-RFE. *PLoS ONE* 6(7): e21750.
- Eisen M, Spellman P, Brown P, Bostein D (1998) Cluster analysis and display for genome-wide expression patterns. *Proc Natl Acad Sci USA* 95: 14863–14868.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, et al. (2001) Gene expression profiles in hereditary breast cancer. *The New England Journal of Medicine* 344: 539–558.
- Lai H, Han B, Li L, Chen Y, Zhu L (2010) An Integrated Semi-Random Forests Based Approach to Gene Selection for Glioma Classification. *Acta Biophys Sin* 26(9): 833–845.
- Guyon I, Weston J, Barnhill S, Vapnik V (2002) Gene selection for cancer classification using support vector machines. *Machine Learning* 46: 389–422.
- Furlanello C, Maria S, Serler M, Giuseppe J (2003) An accelerated procedure for recursive feature ranking on Microarray Data. *Neural Networks* 16: 641–648.
- Duan K, Rajapakse J (2004) A variant of SVM-RFE for gene selection in cancer classification with expression data. *Computational Intelligence in Bioinformatics and Computational Biology. CIBCB '04.* 7–8: 49–55.
- Ding Y, Wilkins D (2006) Improving the performance of SVM-RFE to select genes in microarray data. *BMC Bioinformatics* 7 (Suppl 2): S12.
- Koyama M, Levesley A, Koebner R, Flowers T, Yeo A (2001) Quantitative Trait Loci for Component Physiological Traits Determining Salt Tolerance in Rice. *Plant Physiology* 125: 406–422.
- Seaton G, Haley C, Knott S, Kearsey M, Visscher P (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* 18 (2): 339–340.
- Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucl Acids Res.* 38: W64–W70.
- Fridovich I (1986). Biological effects of the superoxide radical. *Arch. Biochem. Biophys* 247: 1–11.
- Wise R, Naylor A (1987) Chilling-enhanced photooxidation: evidence for the role of singlet oxygen and superoxide in the breakdown of pigments and endogenous antioxidants. *Plant Physiol* 83: 278–282.
- Imlay J, Linn S (1988). DNA damage and oxygen radical toxicity. *Science* 240: 1302–1309.
- Vinocur B, Altman A (2005). Recent advances in engineering plant tolerance to abiotic stress: Achievements and limitations. *Curr. Opin. Biotech.* 16, 123–132.
- Wang Y, Nil H (2000) Changes in chlorophyll, ribulose biphosphate carboxylase-oxygenase, glycine betaine content, photosynthesis and transpiration in *Amaranthus tricolor* leaves during salt stress. *J. Hortic. Sci. Biotechnol.*, 75: 623–627.
- Sapkota A, Liu X, Zhao X, Cao Y, Liu J, et al. (2011) DIPOS: database of interacting proteins in *Oryza sativa*. *Mol. BioSyst.* 7: 2615–2621.
- Smoot M, Ono K, Ruscheinski J, Wang P, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27(3): 431–432.
- Jannink J, Bink Mc, Jansen Rc (2001) Using complex plant pedigrees to map valuable genes. *Trends in plant science* 6 (8): 337–42.

50. Liu J, Zhu J (1998) A calcium sensor homolog required for plant salt tolerance. *Science* 280 (5371): 1943–5.
51. Zhu J (2003) Regulation of ion homeostasis under salt stress. *Curr. Opin. Plant Biol.* 6(5), 441–5.
52. Gupta K, Gupta B, Ghosh B, Sengupta DN (2012) Spermidine and abscisic acid-mediated phosphorylation of a cytoplasmic protein from rice root in response to salinity stress. *Acta Physiologica Plantarum* 34(1): 29–40.
53. Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic acids research* 27(1): 297–300.
54. Pauli S, Rothnie HM, Chen G, He X, Hohn T (2004). The cauliflower mosaic virus 35S promoter extends into the transcribed region. *J Virol.* 78: 12120–12128.
55. Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, et al. (2007) The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Research* 35: D883–D887.
56. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: an Information Aesthetic for Comparative Genomics. *Genome Res* 19: 1639–1645.
57. de Wit PJ (2007) How plants recognize pathogens and defend themselves. *Cell Mol Life Sci* 64: 2726–32.
58. Jones JD, Dang JL (2006) The plant immune system. *Nature* 444: 323–9.
59. Zhang J, Wang Q, Vantasin K, Zhang J, He Z, et al. (2011) A multi-layer evaluation approach for protein structure prediction and model quality assessment. *Proteins Volume 79, Issue Supplement S10: 172–184.*
60. Zhang J, Wang Q, Barz B, He Z, Kosztin I, et al. (2010) MUFOLD: A new solution for protein 3D structure prediction. *Proteins* 78(5): 1137–1152.
61. Neer EJ, Schmidt CJ, Nambudripad R, Smith TF (1994) The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371(6495): 297–300.
62. Xing W, Zou Y, Liu Q, Liu J, Luo X, et al. (2007) The structural basis for activation of plant immunity by bacterial effector protein AvrPto. *Nature* 449(7159): 243–7.
63. Kitahata N, Saito S, Miyazawa Y, Umezawa T, Shimada Y, et al. (2005) Chemical regulation of abscisic acid catabolism in plants by cytochromeP450 inhibitors. *Bioorganic & Medicinal Chemistry* 13(14): 4491–4498.
64. Ko J, Yang S, Han K (2006) Upregulation of an Arabidopsis RING-H2 gene, XERICO, confers drought tolerance through increased abscisic acid biosynthesis. *Plant Journal* 47: 343–355.
65. Dreher K, Callis J (2007) Ubiquitin hormones and biotic stress in plants. *Ann. Bot. (Lond.)* 99: 787–822.
66. Wang R, Pandey S, Li S, Gookin T, Zhao Z, et al. (2011) Common and unique elements of the ABA-regulated transcriptome of Arabidopsis guard cells. *BMC Genomics* 12: 216.
67. Sreenivasulu N, Harshvardhan V, Govind G, Seiler C, Kohli A (2012) Contrapuntal role of ABA: Does it mediate stress tolerance or plant growth retardation under long-term drought stress?. *Gene* 506(2): 265–273.
68. Dionisio-Sese M, Tobita S (1998) Antioxidant responses of rice seedlings to salinity stress. *Plant Science* 135(1): 1–9.
69. Fini A, Brunetti C, Ferdinando M, Ferrini F, Tattini M (2011) Stress-induced flavonoid biosynthesis and the antioxidant machinery of plants. *Plant Signal Behav.* 6(5): 709–711.
70. Agatia G, Biricolti S, Guidic L, Ferrini F, Fini A, et al. (2011) The biosynthesis of flavonoids is enhanced similarly by UV radiation and root zone salinity in *L. vulgare* leaves. *Journal of Plant Physiology* 168(3): 204–212.
71. Turóczy Z, Kis P, Török K, Cserhádi M, Lendvai Á, et al. (2011) Overproduction of a rice aldo-keto reductase increases oxidative and heat stress tolerance by malondialdehyde and methylglyoxal detoxification. *Plant Molecular Biology* 75(4–5): 399–412.
72. Mizuno H, Kawahigashi H, Kawahara Y, Kanamori H, Ogata J, et al. (2012) Global transcriptome analysis reveals distinct expression among duplicated genes during sorghum-Bipolaris sorghicola interaction. *BMC Plant Biology* 12: 121.
73. Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M (2002) Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *The Plant Journal* 31(3): 279–292.
74. Balibera M, Amico J, Bolarin M, Perez-Alfócea F (2000) Carbon partitioning and sucrose metabolism in tomato plants growing under salinity. *Physiol Plant* 110(4): 503–511.
75. Zhou J, Wang X, Jiao Y, Qin Y, Liu X, et al. (2007) Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. *Plant Mol Biol* 63: 591–608.
76. Ogawa I, Nakanishi H, Mori S, Nishizawa N (2009) Time course analysis of gene regulation under cadmium stress in rice. *Plant and soil* 325: (1–2): 97–108.
77. Caesar R, Blomberg A (2004) The stress-induced Tflp requires NatB-mediated acetylation to inhibit carboxypeptidase Y and to regulate the protein kinase A pathway. *Journal of Biological Chemistry* 279(37): 38532–38543.
78. Jacoby R, Millar A, Taylor N (2010) Wheat mitochondrial proteomes provide new links between antioxidant defense and plant salinity tolerance. *Journal of Proteome Research* 9(12): 6595–604.
79. Khurana P, Gaikwad K (2005) The map based sequence of the rice genome. *Nature* 436: 793–800.
80. Cotsaftis O, Plett D, Johnson A, Walia H, Wilson C, et al. (2011) Root-Specific Transcript Profiling of Contrasting Rice Genotypes in Response to Salinity Stress. *Molecular Plant* 4 (1): 25–41.
81. Chitteti BR, Peng Z (2007) Proteome and phosphoproteome differential expression under salinity stress in rice (*Oryza sativa* L.) roots. *J Proteome Res* 6(5): 1718–1727.
82. Komatsu S, Karibe H, Hamada T, Rakwal R (1999) Phosphorylation upon cold stress in rice (*Oryza sativa* L.) seedlings. *Theor Appl Genet* 98: 1304–1310.
83. Chen X, Zhang W, Zhang B, Zhou J, Wang Y, et al. (2011) Phosphoproteins regulated by heat stress in rice leaves. *Proteome Science* 9: 37.
84. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* 5(10): R80.
85. Bolstad B, Irizarry R, Astrand M, Speed T (2003) A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinformatics* 19(2): 185–193.
86. Bailey T, Bodén M, Buske F, Frith M, Grant C (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37: W202–W208.
87. Grant C, Bailey T, Noble W (2011) FIMO: Scanning for occurrences of a given motif. *Bioinformatics* 27(7): 1017–1018.
88. Davey N, Edwards R, Shields D (2007) The SLiMDisc server: short, linear motif discovery in proteins. *Nucleic Acids Res.* 35(Web Server issue): W455–9.