

RESEARCH ARTICLE

# Whole-Genome Analysis Revealed the Positively Selected Genes during the Differentiation of *indica* and Temperate *japonica* Rice

Xinli Sun<sup>1,2\*</sup>, Qi Jia<sup>1,2</sup>, Yuchun Guo<sup>1,2</sup>, Xiujuan Zheng<sup>1,2</sup>, Kangjing Liang<sup>1,2</sup>

**1** Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, Fujian Agriculture & Forestry University, Fuzhou, China, **2** College of Crop Science, Fujian Agriculture & Forestry University, Fuzhou, China

\* [xinlisun@hotmail.com](mailto:xinlisun@hotmail.com)



## Abstract

To investigate the selective pressures acting on the protein-coding genes during the differentiation of *indica* and *japonica*, all of the possible orthologous genes between the Nippon-bare and 93–11 genomes were identified and compared with each other. Among these genes, 8,530 pairs had identical sequences, and 27,384 pairs shared more than 90% sequence identity. Only 2,678 pairs of genes displaying a  $K_a/K_s$  ratio significantly greater than one were revealed, and most of these genes contained only nonsynonymous sites. The genes without synonymous site were further analyzed with the SNP data of 1529 *O. sativa* and *O. rufipogon* accessions, and 1068 genes were identified to be under positive selection during the differentiation of *indica* and temperate *japonica*. The positively selected genes (PSGs) are unevenly distributed on 12 chromosomes, and the proteins encoded by the PSGs are dominant with binding, transferase and hydrolase activities, and especially enriched in the plant responses to stimuli, biological regulations, and transport processes. Meanwhile, the most PSGs of the known function and/or expression were involved in the regulation of biotic/abiotic stresses. The evidence of pervasive positive selection suggested that many factors drove the differentiation of *indica* and *japonica*, which has already started in wild rice but is much lower than in cultivated rice. Lower differentiation and less PSGs revealed between the Or-It and Or-IIIa wild rice groups implied that artificial selection provides greater contribution on the differentiation than natural selection. In addition, the phylogenetic tree constructed with positively selected sites showed that the *japonica* varieties exhibited more diversity than *indica* on differentiation, and Or-III of *O. rufipogon* exhibited more than Or-I.

## OPEN ACCESS

**Citation:** Sun X, Jia Q, Guo Y, Zheng X, Liang K (2015) Whole-Genome Analysis Revealed the Positively Selected Genes during the Differentiation of *indica* and Temperate *japonica* Rice. PLoS ONE 10(3): e0119239. doi:10.1371/journal.pone.0119239

**Academic Editor:** Binying Fu, Institute of Crop Sciences, CHINA

**Received:** September 5, 2014

**Accepted:** January 11, 2015

**Published:** March 16, 2015

**Copyright:** © 2015 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors have no support or funding to report.

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

## Introduction

Asian cultivated rice (*O. sativa* L.) is one of the oldest and most important crop species. It is the primary source of food and livelihood for more than a third of Asia's population, accounting for 35–60% Asia's and ~20% the world's caloric intake respectively [1]. *O. sativa* has a broad geographic distribution across the world with a high phenotypic variability, an estimated 120,000 varieties [1]. Most varieties have been placed into two subspecies, *O. sativa* ssp. *indica* and *O. sativa* ssp. *japonica*, which differ in more than 40 characteristics, such as phenol reaction phenotype, KClO<sub>3</sub> resistance, cold sensitivity, drought tolerance, germination, seed shedding, length-width ratio of spikelet, apiculus hair length, awn length, digestion of endosperm in KOH solution, hardening of endosperm and first internode [2]. Some of these characteristics have been used to distinguish the *indica* and *japonica* varieties [2, 3]. Further analyses with ecological traits, isozymes and/or DNA markers confirmed and developed the above classification [2–8]. Studying 950 accessions with 4.1 million SNPs, Huang et al further divided the *japonica* subspecies into two sub-groups, temperate *japonica* and tropical *japonica*, and the *indica* subspecies into *indica* and aus sub-group [9].

The immediate progenitor of *O. sativa* is *O. rufipogon* [10]. Previous studies mostly focused on the domestication of wild rice, indicating that *O. sativa* was domesticated from *O. rufipogon* approximately 8200–13,500 years ago [2, 11]. However, these studies have provided two hypotheses about the origin(s) of two subspecies. One proposed that the domesticated rice originated from a single common wild ancestor, and differentiation of *indica* and *japonica* occurred after the domestication of cultivated species, which is supported by the analyses of well-characterized domestication genes and SNPs from 630 gene fragments in wild and cultivated rice accessions [11–16]. The other hypothesis suggested that two major rice types were domesticated separately from different populations of wild rice, supported by phylogenetic analyses that showed distinct clades in *O. sativa* for *indica* and *japonica* with different *O. rufipogon* accessions associated with each clade [17–22], as well as the whole-genome SNPs analyses [10, 23]. The SNPs analyses further indicated that *japonica* was first domesticated from a specific population of *O. rufipogon* around the middle area of the Pearl River in southern China, and that *indica* was subsequently developed from crosses between *japonica* and local wild rice as the initial cultivars spread into South East and South Asia [10].

Incomplete observations with one or several isozymes or domestication-related genes have only dropped small hints about the domestication processes, whereas the application and development of molecular markers can provide considerable information for the understanding of rice evolution. However, analyses with molecular markers such as RFLPs, SSRs and SNPs, which are usually caused by mutation, need to specify whether the mutation is neutral or not. A neutral mutation cannot change the gene function, and has no effect on fitness. Thus, such mutations provide less useful information for evolutionary analyses, and may interfere with the prediction. An advantageous mutation would have a positive effect on phenotype, and increase the fitness of the organism. These mutations will be accumulated in the gene pool. Conversely, deleterious mutations would decrease the fitness of the organism, and get typically eliminated from the gene pool by selection. Thus, positively selected genes (PSGs) carry much more information that is relevant to the evolutionary history of a species than negatively selected genes. Furthermore, the PSGs are, or have been, functionally important, and identification will facilitate the understanding of genetic variation that contributes to phenotypic diversity, and help to annotate the functional genome. Therefore, the emphasis of the domestication and differentiation analyses should be placed on the PSGs.

Differentiation of *indica* and *japonica* was driven by both artificial and natural selection, which directly acted on many characteristics [2, 24–27]. Selection results in a difference in gene

frequencies between populations. Various factors have been known to be as selective forces, e.g., temperature, light condition, day length, soil fertility, stress conditions like drought, submergence, salinity, pollution, and herbicide use [2]. Protein evolution is the outcome of interaction between mutational processes and selective forces; therefore, analyzing the coding region of a genome is fundamental to understand how selection influences evolution. As a model organism, *O. sativa* is a well-characterised species with a small genome (389 Mb). The *indica* variety 93–11 and the *japonica* variety Nipponbare have been fully sequenced [28], and more than 2000 accessions including wild rice have been partly sequenced (<http://ricevarmap.ncpgr.cn/django/home/>) [10]. These features afford unique opportunities to explore differentiation of *indica* and *japonica* via genomic approaches. In this study, the genes under positive selection were identified and analyzed systematically in order to provide information for further understanding of cultivated rice evolution.

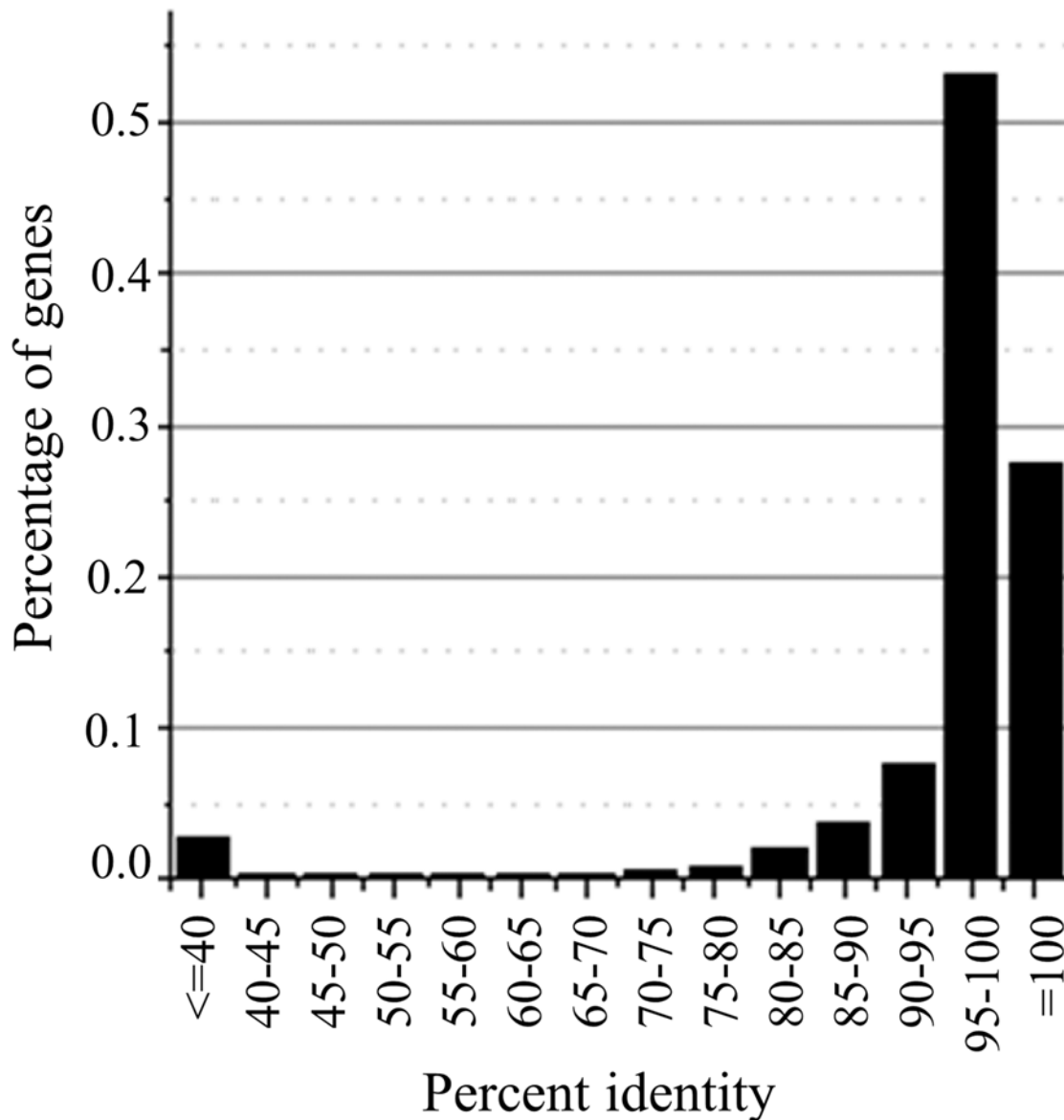
## Results

All of the gene annotations for Nipponbare and 93–11 were downloaded from online public databases. There were 40,354 and 67,393 annotations for Nipponbare from the Rice Annotation Project Database (RAP-DB) and Rice Genome Annotation Project (RGAP), respectively, and 40,745 annotations for 93–11 from Rice Information System (RISe). The annotations of RGAP contained more alternatively spliced genes, transposons and retrotransposon genes. The databases of Nipponbare protein sequences retrieved from RAP-DB and RGAP were queried with the 93–11 protein sequences to identify pairs of orthologs, respectively. Combining the two BLAST results, 30,995 pairs of orthologous genes were found. The identity values of these orthologs were re-calculated according to the ClustalW2 result, and the resulting distribution of percent identities was shown in Fig. 1A total of 8,530 gene pairs exhibited 100% identity, and 27,384 pairs had more than 90% identity. These pairs were used to evaluate positive selection between *indica* and *japonica*.

### Positively selected genes between 93–11 and Nipponbare

Positive selection is often evaluated by the ratio of nonsynonymous/synonymous substitution rates. This ratio,  $K_a/K_s$ , is expected to be greater than 1.0 in the case of positive selection [29]. The orthologous genes with high identity values (>90%) were used to detect instances of positive selection between 93–11 and Nipponbare by estimating their synonymous and nonsynonymous substitution rates. The  $K_a$  and  $K_s$  values of those genes were obtained with NG, gNG, YN, MYN, and maximum likelihood, respectively, and the results with maximum likelihood, gNG and MYN were shown in Fig. 2. The distributions of the  $K_a$  and  $K_s$  values were very narrow, with 99% of those genes displaying the  $K_a$  and  $K_s$  values below 0.3, and 18.5–24.5% of the gene pairs (more than 5000 pairs) showed  $K_a/K_s > 1$  (Fig. 2). Fisher's test was used to identify the  $K_a/K_s$  ratios that were significantly higher than one, suggesting more than 10% genes were positively selected during the differentiation of 93–11 and Nipponbare. In addition, the average percent identity of the PSGs was 99.29%.

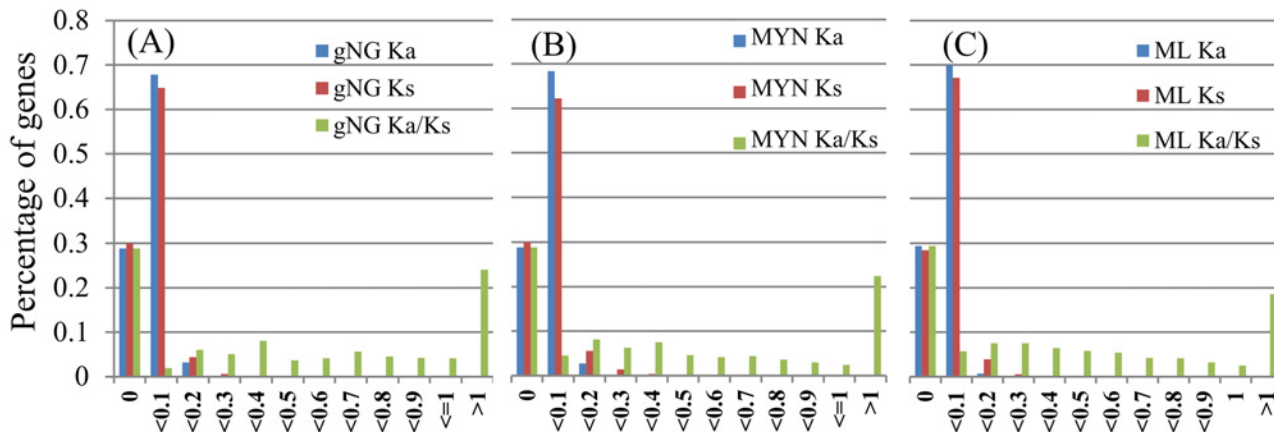
These PSGs were manually analyzed to remove annotation mistakes and ClustalW errors, followed by re-calculation. There were 2,977 PSGs detected with the gNG method and 2,799 PSGs with MYN. Among them, 2,664 PSGs were shared via both approaches. Interestingly, the synonymous substitution numbers of most these PSGs are zero (S1 Table). We denoted such type of genes as nonsynonymous substitution genes (NSSGs) including the genes whose  $K_a/K_s$  ratios were not significantly higher than one in Fisher's test (S1 Table). Most of the NSSGs exhibited one or two substitutions (Fig. 3). Only seven PSGs with synonymous substitution sites were detected by at least one method (S2 Table).



**Fig 1. The distribution of the percent identity between the possible orthologs.** The most similar proteins between 93–11 and Nipponbare were selected with BLAST, and 30995 pairs of proteins were obtained; each pair was analyzed via ClustalW 2 to obtain the percent identity.

doi:10.1371/journal.pone.0119239.g001

To further investigate the selective pressure acting on protein-coding genes, all of the genes whose Ka/Ks ratios is not significantly higher than one were analyzed with an alternative approach to calculate the Ka/Ks ratios on sliding windows of fixed size. The alignment slicing procedure with sliding windows of 100 codons and a window shift of 34 codons generated 108,358 windows with less than 50% gap. Only 14 PSGs were identified by at least one of above methods (MYN, YN, NG and ML) after removing the annotation mistakes and ClustalW errors (S3 Table). Because it is difficult to do further analysis with the SNP data for these genes, we focused on NSSGs in the following study.



**Fig 2. The distribution of Ka, Ks and Ka/Ks.** (A) Using the gNG method. (B) Using the MYN method. (C) Using the Maximum Likelihood method. Note: Ka/Ks were specified as zero if both Ka and Ks were zero (5247 genes).

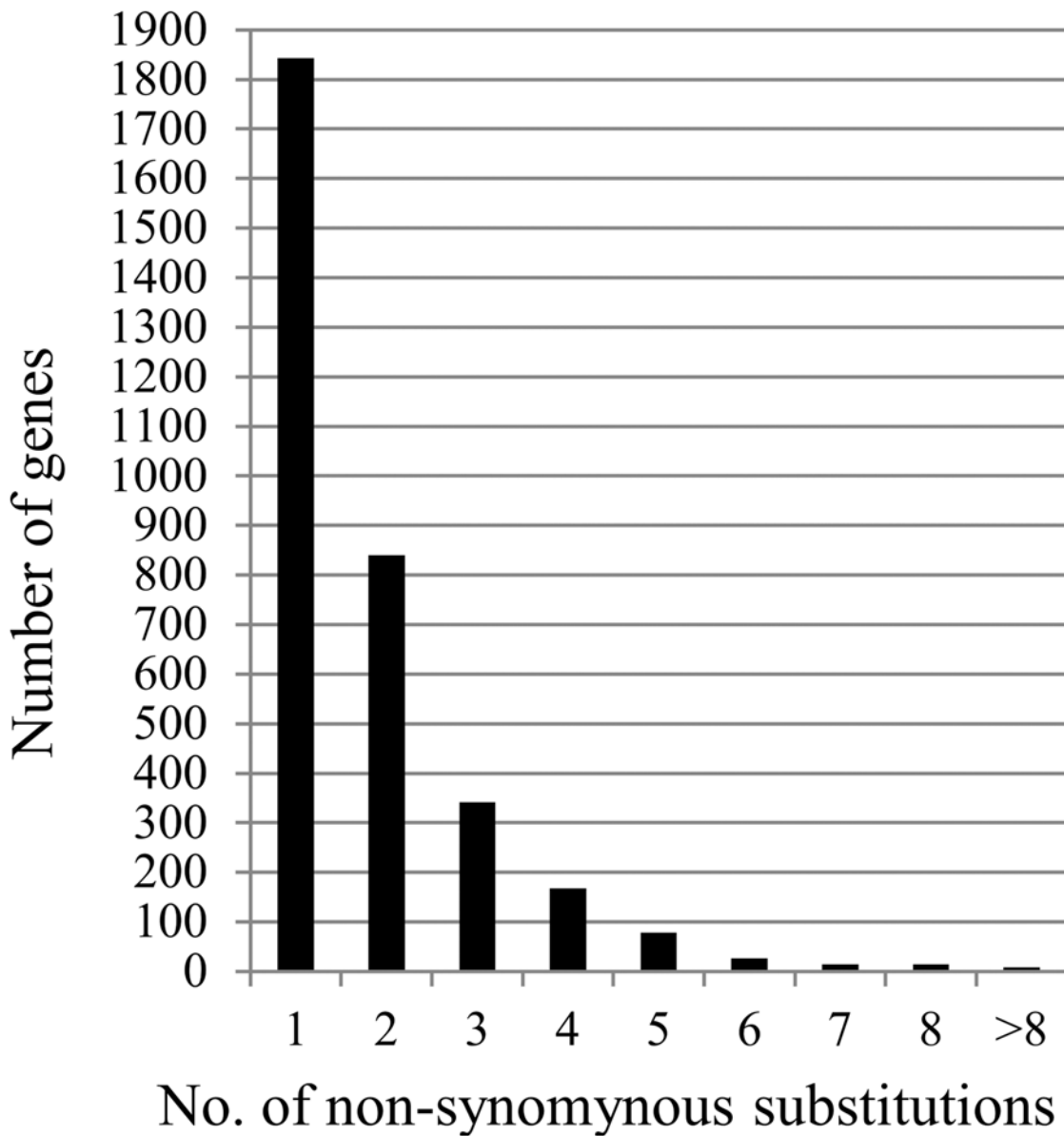
doi:10.1371/journal.pone.0119239.g002

### SNP data can be used to estimate the positively selected sites in NSSGs between *indica* and temperate *japonica*

To detect whether the above PSGs found between 93–11 and Nipponbare were also under selection among most of the other *indica* and *japonica* accessions, the data containing 520 *indica* (excluding aus rice) and 409 temperate *japonica* accessions with 4.1 million SNPs was downloaded [10] and analyzed, as 93–11 belonged to *indica* and Nipponbare to temperate *japonica*. The SNPs in the exons of all PSGs were exposed, and most of the positively selected sites (PSSs) in the PSGs were able to be found in these SNPs. In addition, some new sites that changed amino acids were also discovered. Taking chromosome 1 for example, we found 2847 SNPs in the exons of 497 NSSGs, which included 687 (77.6%) PSSs revealed between 93–11 and Nipponbare. There were 198 (22.4%) PSSs that did not contain in these SNPs, so we speculated that these PSSs were specific to the differences between 93–11 and Nipponbare or the SNP coverage was not enough.

We calculated the  $F_{st}$  value of each SNP site based on its frequency in *indica* and temperate *japonica*, which was thought to be a measure of population differentiation due to genetic structure [30–32]. Most of these SNP sites possess lower  $F_{st}$  values, with 70.6% less than 0.25 and 60.7% less than 0.1. These sites could less affect the early differentiation between *indica* and *japonica*, and thus we did not examine whether their alterations changed the amino acids. The frequencies of these SNPs were calculated in total, indicating that the minor base frequencies of 1687 (83.9%) SNP sites were less than 0.1. It showed that these SNPs merely affected a small amount of rice accessions when they changed the amino acids, explaining why so many SNPs could not be found in PSSs.

We found 115 new nonsynonymous sites and 80 synonymous sites in the SNP data (Table 1). The difference of 93–11 and Nipponbare could represent about 73.2% diversity among all the *indica* and temperate *japonica* accessions when the  $F_{st}$  values are not less than 0.25. This suggested that it was possible to find synonymous sites in NSSGs in the study of the differentiation of *indica* and temperate *japonica*. It is difficult to calculate the number of the synonymous sites in the unsequenced region, but the probability could be estimated in the sequenced region, and the two regions are supposed to have an identical probability. The average probability is about 11.0%, but the probability is much less for the NSSGs with higher  $F_{st}$



**Fig 3. The distribution of the numbers of non-synonymous substitutions in NSSGs.**

doi:10.1371/journal.pone.0119239.g003

values (Table 1). These results inferred that SNPs data could be used to estimate the positively selected sites in NSSGs between *indica* and temperate *japonica*.

### Positively selected genes between *indica* and temperate *japonica*

We discovered 313 genes that only contain nonsynonymous sites ( $F_{st} \geq 0.25$ ) in chromosome 1 (S4 Table). To detect whether these SNP loci were the signatures of adaptation during the genetic differentiation between populations, Lositan, an  $F_{st}$  related statistic method, was employed to identify the outliers that were positively influenced by selection, acting on either the locus itself or the closely linked locus. We identified 105 (19.9%) outliers within 99% confidence interval and additional 56 (10.5%) within 95% confidence interval. Almost all of the sites whose  $F_{st}$  values are more than 0.95 were outliers. However, the sites with the  $F_{st}$  value of one

**Table 1. Number of nonsynonymous sites and new synonymous sites of chromosome 1 between *indica* and temperate *japonica*.**

Fst	shared <sup>1</sup>	New Nonsyn. <sup>2</sup>	New syn. <sup>3</sup>	total	Probability <sup>4</sup>
> = 0.95–1	204	9	6	219	0.027
0.9–0.95	45	6	5	56	0.089
0.8–0.9	59	8	11	78	0.141
0.7–0.8	52	14	7	73	0.096
0.6–0.7	37	10	14	61	0.230
0.5–0.6	45	12	8	65	0.123
0.4–0.5	41	18	9	68	0.132
0.3–0.4	32	17	8	57	0.140
0.25–0.3	17	21	12	50	0.240
< 0.25	155				

<sup>1</sup>Number of nonsynonymous sites (PSSs) shared by sequence and SNP analyses.

<sup>2</sup>Number of new nonsynonymous sites found between *indica* and temperate *japonica*

<sup>3</sup>Number of new synonymous sites found between *indica* and temperate *japonica*

<sup>4</sup>The probability of synonymous site occurred in NSSGs.

doi:10.1371/journal.pone.0119239.t001

were not able to be detected as outliers because one is the biggest Fst value, and the program cannot distinguish which is bigger between the simulation and the sample Fst. These sites should be under selection during the differentiation of *indica* and *japonica*, and have already been completely fixed. We also detected some outliers with lower Fst values, which would infer the recent selection occurred (Table 2).

We detected 173 (55.3%) genes (including the genes with the site whose Fst is one) with positively selected outliers and only nonsynonymous sites (S4 Table). Natural and/or artificial selection could directly act on these genes. Some genes with more than one nonsynonymous site

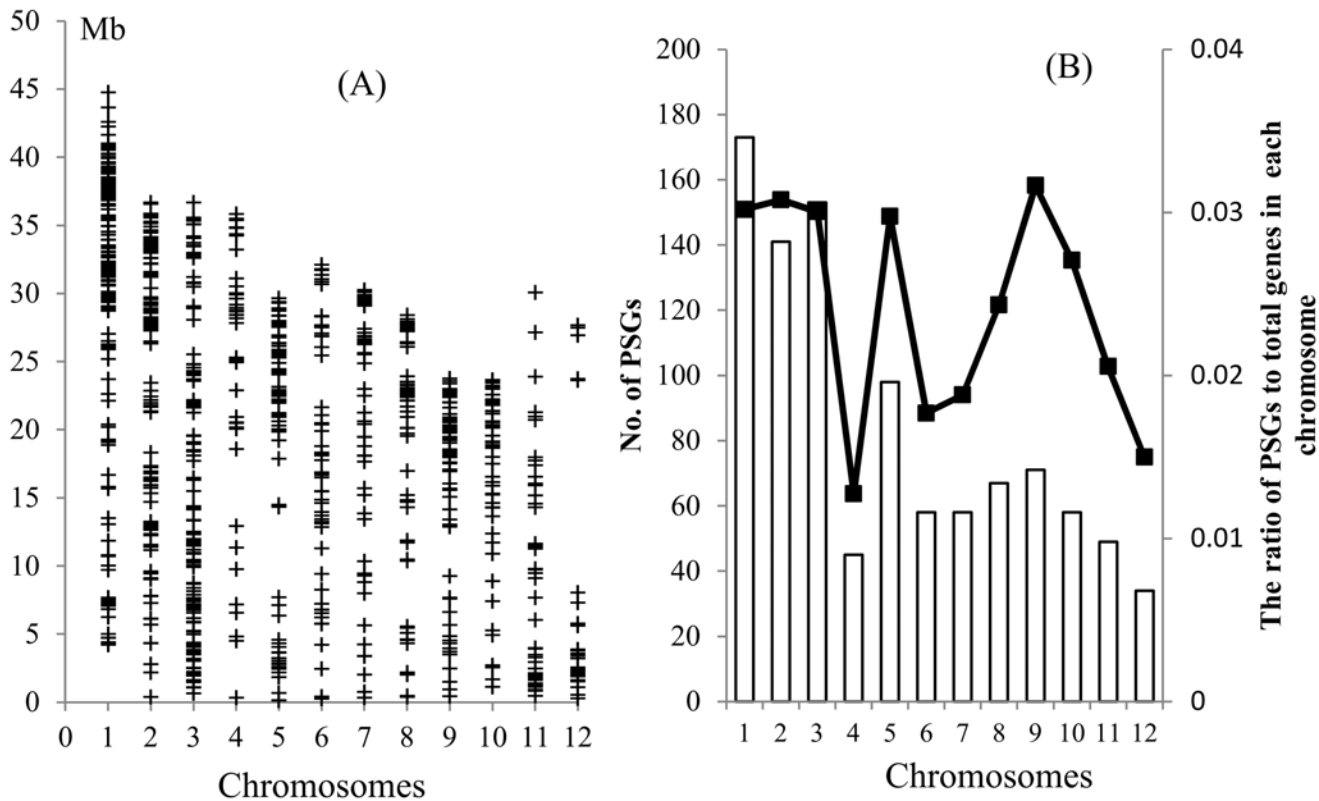
**Table 2. The distribution of NSSGs and outliers along Fst values between *indica* and temperate *japonica*, and the distribution of outliers along Fst values between Or-I and Or-III.**

Fst values	No. of NSSGs between <i>indica</i> and temperate <i>japonica</i> <sup>1</sup>		No. of outliers between <i>indica</i> and temperate <i>japonica</i>		No. of outliers between Or-I and Or-III <sup>2</sup>
	Chr.1	Chr.2–12	Chr. 1	Chr. 2–12	
1	76	224	99	252	30
0.95–<1	68	310	93	429	66
0.9–0.95	19	162	34	240	64
0.8–0.9	34	68	13	85	22
0.7–0.8	34	4			
0.6–0.7	21	7	1	3	
0.5–0.6	27	11	9	18	2
0.4–0.5	25	22	5	44	8
0.25–0.4	9	24	6	56	28
Total	313	832	161	1127	220

<sup>1</sup>Only the PSS with highest Fst was considered for some PSGs with more than one PSS.

<sup>2</sup>The number of the PSSs used to analyze *O. rufipogon* is 1354, including some of the PSSs from the PSGs with a synonymous site.

doi:10.1371/journal.pone.0119239.t002



**Fig 4. The distribution of the PSGs along rice chromosomes.** (A) The distribution of the PSGs along the chromosomes. '+' indicates the positions of the genes on the chromosomes. (B) The numbers of the PSGs (bars) and the ratios of PSGs (lines) to total genes in each chromosome.

doi:10.1371/journal.pone.0119239.g004

were measured again after combining the SNPs to get the haplotypes of the genes, but no additional genes were found to be under positive selection (data not shown).

We found some NSSGs with synonymous sites between *indica* and temperate *japonica*, in which some synonymous sites were detected as outliers, and these may be hitchhiking sites (S5 Table). The genes with higher  $F_{st}$  values nonsynonymous sites and lower synonymous sites were further analyzed, and those with nonsynonymous outliers should be under positive selection (S5 Table).

All NSSGs on the other chromosomes were analyzed between *indica* and temperate *japonica* with Lositan. The outlier sites corresponding to the amino residues and their positions on the proteins were shown in S6 Table. Additional 832 genes (1005 genes in total including chromosome 1) containing only nonsynonymous outlier sites were revealed (Table 2 and S6), and 83.6% of these genes have at least one high- $F_{st}$ -value site ( $F_{st} > 0.9$ ).

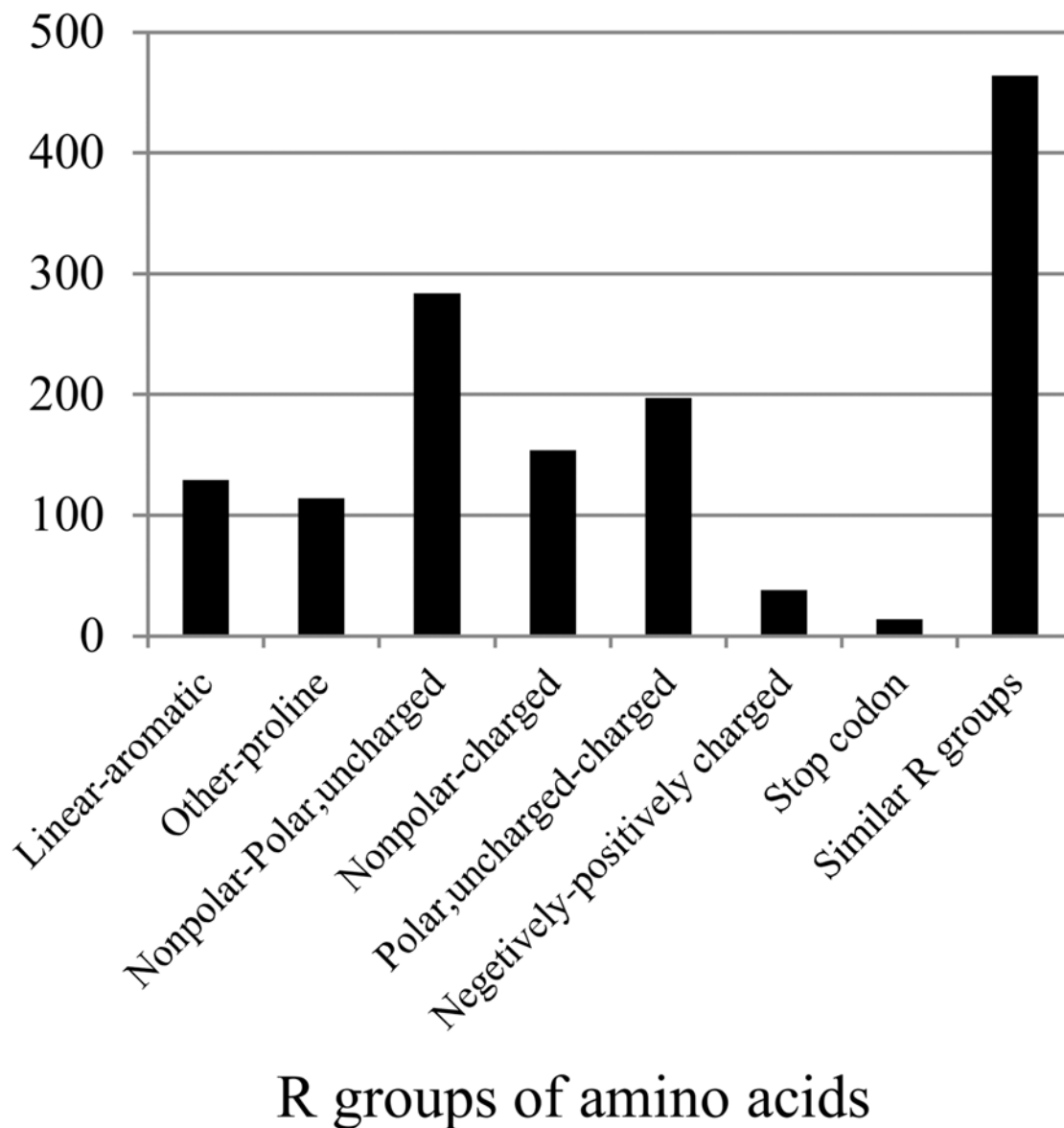
The PSGs were unevenly distributed among and along the chromosomes (Fig. 4). The numbers of the PSGs in the chromosome 1, 2, 3 are about 2 times more than those in the chromosome 4, 6, 7, 10, 11 and 12 (Fig. 4B). The gene density was usually higher near the ends of the chromosomes, but the distribution patterns on different chromosomes were distinct. For example, the PSGs were preferentially located near the ends of the short arms of chromosome 11 and 12, and the ends of the long arms of chromosomes 1, 2, 4, 7, 8, 9 and 10 (Fig. 4A). The ratios of the PSGs to total genes in each million base pairs were calculated, and the pattern of these ratios along the chromosomes was similar but not identical to that of the numbers of the PSGs (S1 Fig). In addition, the distribution of the ratios of PSGs to total genes in each chromosome was similar with that of the numbers of the PSGs except chromosome 9, in which the



gene density is much higher with respect to the number of PSGs (Fig. 4). The result indicated that the uneven distribution was independent on the gene density in a chromosome.

We found 1393 nonsynonymous outlier sites including one site whose change altered intron 3' splice site and 14 sites whose changes resulted in stop codon. We found that 33.3% sites were replaced by the amino acids with similar R group (side chain), and the rest by the amino acids with different property (Fig. 5). Some substitutes were able to severely change the structures of the proteins, for example, the substitution of proline (Fig. 5).

We uncovered and selected additional 63 genes that contained only one synonymous site and some nonsynonymous outliers, in which the  $F_{st}$  value of the synonymous site was not higher than at least one nonsynonymous site (S7 Table). Some of these genes contained a



**Fig 5. The number of each type of substitutions in the proteins encoded by the PSGs.**

doi:10.1371/journal.pone.0119239.g005

nonsynonymous site, which altered the codon to stop codon, before the synonymous site on the gene. These genes were included in [S7 Table](#) as well, and could be under positive selection.

## Functional classification of PSGs

Based on the gene annotations, we summarized the possible functions of the PSGs ([S8 Table](#)). However, the function annotations of at least 223 PSGs are hypothetical or unknown, and some other PSGs contained two or more domains or motifs. Many proteins encoded by PSGs possessed zinc finger domain, binding domain, F-box domain or ankyrin repeat etc., or belonged to the protein families of transcription factor, transferase, protein kinase, peptidase, synthase, transporter or hydrolase etc. ([S8 Table](#)).

Blast2GO, RAP-DB and MSU Rice Genome Annotation Project were used to reveal the GOs of these PSGs, and cellular component, molecular function and biological process annotations were found in only 492, 626 and 499 PSGs, respectively. These genes were involved in 218 molecular functions and 213 biological processes, respectively ([S9 Table](#) and [S10 Table](#)). More attentions should be paid for 402 (64.2%) proteins (including sequence-specific DNA binding transcription factor activity) encoded by the PSGs with binding activity, 116 (18.5%) with hydrolase activity, 93 (14.9%) with transferase activity. There were 118 (23.6%) PSGs involved in macromolecular metabolic processes, 75 (15.0%) in biological regulation, 64 (12.8%) in transport (transmembrane transport and vesicle-mediated transport) and 63 (12.6%) in response to stimulus (including cellular response to stimulus). Some proteins could contain more than one activity (462 proteins) or be involved in more than one process (316 proteins), or vice versa.

To reveal whether the PSGs contained some genes of known function, we searched NCBI, QTARO (<http://qtaro.abr.affrc.go.jp/ogro/table>) and Google Scholar with the IDs of RAP-DB and RGAP. We found 29 genes have been characterized, and 47 genes' expression patterns have been revealed ([Table 3](#) and [S11](#)). More than half of these genes are involved in the regulation of biotic and/or abiotic stresses, and seven of these genes regulate the germ cell development ([Table 3](#)), especially *S5* gene (Os06t0213100 or LOC\_Os06g11010), which regulates the hybrid sterility between *indica* and *japonica* variety [[33](#)] and was positively selected during differentiation of *indica* and *japonica* ([S6 Table](#)). The expressions of most genes can be induced or repressed by biotic and/or abiotic stresses ([S11 Table](#)). These results implied that many artificial/environmental factors could directly act on the genes and accelerated the variety differentiation. In particular, some known PSGs (*Dpl2* and *S5* genes) are involved in reproductive isolation ([Table 3](#)).

## Differentiation and positive selection among the *O. rufipogon* accessions

*O. sativa* have been classified into five groups—*indica*, aus, temperate *japonica*, tropical *japonica* and intermediate, while *O. rufipogon* into three groups—Or-I, Or-II and Or-III [[9](#), [10](#)]. To investigate the population differentiation on *indica* and *japonica*, we constructed a neighbour-joining tree with 446 *O. rufipogon* accessions and 1,083 *O. sativa* varieties based on the PSSs revealed between *indica* and temperate *japonica*, in which some PSSs were deleted for lack of data in some groups. The results showed that *indica* and temperate *japonica* were separately located on the two sides of the phylogenetic tree with the largest difference as prediction. The temperate and tropical *japonica* varieties distributed over a large range on the tree comparing with *indica* and aus which were clustered together. Most *indica* including the aus accessions seemed to generate from one progenitor. *O. rufipogon* were located between *indica* and *japonica*, meanwhile a small number of the Or-I accessions were in *indica* or aus group. Most of the

**Table 3. The PSGs of known function.**

Locus ID	Gene	Isolation or expression	Characters	Functions	Ref*
Os01t0678600	<i>asl1</i>	Mutant	Albino seedling lethality	Chloroplast development.	[34]
Os01t0695900	<i>OsMYB4</i>	Overexpression	Chilling and freezing tolerance	Cold tolerance	[35, 36]
Os01t0756700	<i>OsKAT1</i>	Overexpression	Salinity tolerance	Salinity tolerance in protoplast. Maintenance of cytosolic cation homeostasis.	[37]
Os01t0816100	<i>OsNAC4</i>	Knockdown	Blast resistance	Blast resistance. HR cell death.	[38–40]
Os01t0831000	<i>lax</i>	Mutant	Culm leaf, rachis-branches, lateral spikelet	Lateral organ development. Axillary meristem formation.	[41–44]
Os01t0867300	<i>Osabf1</i>	Mutant	Sensitive to drought and salinity treatment.	Drought and salinity tolerance.	[45]
Os01t0872800	<i>OsPdk1</i>	Mutant and overexpression	Overexpression of ospdk1 enhanced basal resistance against bacterial blight resistance and blast resistance	Ospdk1 participates in signal transduction through pathogen recognition	[46]
Os01t0929600	<i>rtS</i>	Knockdown	Sterility	Pollen development. Anther development.	[47]
Os02t0664000	<i>OsGPX3</i>	Knockdown	Dwarf and shorter root. Accumulation of H <sub>2</sub> O <sub>2</sub>	Root development. Dwarfism. H <sub>2</sub> O <sub>2</sub> homeostasis.	[48]
Os02t0766700	<i>OsZIP23</i>	Mutant and overexpression	Decreased sensitivity to ABA and tolerance to salinity and drought stress.	Drought and salinity tolerance. ABA sensitivity.	[49]
Os03t0119966	<i>rim1</i>	Mutant	Rice dwarf virus resistance.	Rice dwarf virus resistance.	[50]
Os03t0285800	<i>OsMAPK5</i> or <i>OsMPK3</i>	Knockdown and Overexpression	Bacterial blight and blast resistance; cold, drought and salinity tolerance	Positively regulate response to biotic and abiotic stress, and JA pathway	[51–55]
Os03t0821300	<i>xb15</i>	Mutant	Bacterial blight resistance	Resistance to Xoo. Regulation of cell death.	[56]
Os05t0420300	<i>serf1</i>	Mutant	Sensitive to salt stress	Salinity tolerance.	[57]
Os06t0184100	<i>DPL2</i>	Natural variation	Sterility	Hybrid sterility. Pollen germination. Interaction with DPL1	[58]
Os06t0213100	<i>S5</i>	Natural variation	Sterility	Single locus hybrid sterility.	[33]
Os06t0354700	<i>nyc3</i>	Mutant	Stay green	Leaf senescence. Chlorophyll degradation.	[59]
Os06t0665400	<i>apo1</i> or <i>SCM2</i>	Mutant and natural variation	Grain number. Lodging resistance	Floral organ identity; panicle branching; culm strength	[60–63]
Os06t0712700	<i>spw1</i> or <i>OsMADS16</i>	Mutant and overexpression	Alter floral organ	Floral organ formation.	[64, 65]
Os06t0724900	<i>ila1</i>	Mutant	Increase leaf angle	Abnormal vascular bundle formation and cell wall composition in the leaf lamina joint.	[66]
Os07t0687700	<i>rTGA2.1</i>	Knockdown	Bacterial blight resistance and reduced plant stature	Resistance to <i>Xanthomonas oryzae</i> pv. <i>Oryzae</i> . Growth retardation.	[67]
Os08t0139000	<i>OsDEG10</i>	Knockdown	Sensitive to high light and cold stresses	High-light and cold tolerance.	[68]
Os08t0522400	<i>OsAPx-R</i>	Knockdown	Dwarf	Delay development and disturb steady state of the antioxidant	[69]
Os09t0439200	<i>OsJAZ8</i>	Overexpression	Bacterial blight resistance	JA induced resistance to <i>Xanthomonas oryzae</i> pv. <i>Oryzae</i> .	[70]
Os09t0441900	<i>DEP1</i>	Natural variation	Dense and erect panicle.	Enhance meristematic activity. Conferring cadmium tolerance	[71, 72]
Os09t0507200	<i>OsMADS8</i>	Knockdown	Panicle flower	Floral organ formation.	
Os09t0522000	<i>OsDREB1B</i>	Overexpression	Cold, drought and salinity tolerance	Regulators of the abiotic stress	[73–76]

(Continued)

Table 3. (Continued)

Locus ID	Gene	Isolation or expression	Characters	Functions	Ref*
Os09t0537700	<i>OsRNS4</i>	Overexpression	Salinity tolerance	Salinity tolerance. Positive regulation in ABA response	[77]
Os12t0572800	<i>mel2</i>	Mutant	Developmental aberration of germline and nursery cells	Regulate the premeiotic G1/S-phase transition of male and female germ cells,	[78]

\*reference

doi:10.1371/journal.pone.0119239.t003

intermediate varieties were located between tropical *japonica* and Or-III, Or-II and Or-III, or Or-I and Or-II. The aromatic varieties in the collection were put together with some intermediate varieties (S2A Fig). We also constructed a Minimum Evolution tree and obtained a similar result (data not shown). To explore the phylogenetic relationships of the wild rice further, we constructed a tree only with the *O. rufipogon* accessions. Most of Or-I and Or-II were concentrated together, whereas Or-III distributed over a large range. Unexpectedly, some of Or-III seemed more close to the Or-II group (S2B Fig). We selected some *O. rufipogon* accessions into three new groups—Or-It, Or-IIIIt and Or-IIIIt for the next analysis, which were concentrated on the tree separately. Almost all of the accessions in Or-IIIIt were from China.

The phylogenetic tree exhibited that the degree of the *indica-japonica* differentiation of the wild rice was between *indica* and *japonica*, and Or-IIIIt was between Or-It and Or-IIIIt. We then measured the *Fst* values according to the frequencies of the PSSs, as shown in Table 4. The biggest average *Fst* value was between *indica* and temperate *japonica* as we expected, followed by the one between temperate *japonica* and Or-It. However, the *Fst* value between *indica* and Or-IIIIt was much lower than that between temperate *japonica* and Or-It. The *indica-japonica* differentiation of the wild rice is mainly between Or-It and Or-IIIIt, which exhibited a relatively high level of population differentiation (Table 4).

The PSSs were also revealed to check whether they were outliers during population differentiation in the wild rice. It showed that 23.2% PSSs were found to be under positive selection between Or-It and Or-IIIIt, but 29.7% outliers were associated with the *Fst* values lower than 0.25 (S12 Table). The GOs of these PSGs with the *Fst* values over 0.25 were shown in S12 Table, and no significant difference was found with the GOs' distribution of *indica*~temperate *japonica*. The distribution of the outliers along *Fst* values were shown in Table 2. Most of the outliers with a high *Fst* value between Or-It and Or-IIIIt also exhibited a high *Fst* value between *indica* and temperate *japonica* (S13 Table). These results inferred that the differentiation of *indica*

Table 4. The *Fst* values between the rice groups.

	Or-It	Or-IIIIt	Or-IIIIt	TeJ*	TrJ*	<i>Indica</i>
Or-IIIIt	0.192					
Or-IIIIt	0.476	0.247				
TeJ	0.802	0.561	0.230			
TrJ	0.658	0.452	0.198	0.076		
<i>Indica</i>	0.064	0.299	0.592	0.919	0.771	
aus	0.040	0.196	0.443	0.750	0.610	0.102

\*TeJ: temperate *japonica*; TrJ: tropical *japonica*

doi:10.1371/journal.pone.0119239.t004

and *japonica* should have started before domestication, and the differentiation was becoming stronger during and after domestication.

## Discussion

### The basis for understanding the differentiation of *indica* and *japonica*

The *indica* and *japonica* types exist as natural varieties that differ in their adaptation to distinct climatic, ecogeographic and cultural conditions [79]. The rice cultivars in the temperate countries such as Japan, Korea and northern China are exclusively *japonicas*, and those grown in the tropical and subtropical regions such as Thailand, Burma, India and southern China are usually *indicas*. In addition, some *japonicas* are also distributed in high altitude areas of the tropics [2]. Both natural and artificial selection have affected the distribution of *indica* and *japonica* rice, and brought about many different morphological and physiological traits between the two groups. Our study discovered that these PSGs were involved in 213 biological processes, and had 218 molecular functions except the genes without functional annotation. Many of these proteins encoded by the PSGs had binding activity, and were involved in response to stimulus and in biological regulation (S9 Table and S10 Table). More than half of the PSGs with known function and/or expression were involved in the responses to biotic/abiotic stresses, and some of them (*Dpl2* and *S5* genes) are involved in reproductive isolation (Table 3 and S11). These results implied that selection played an important role in the differentiation of *indica* and *japonica*, and these PSGs might directly or indirectly regulate and control these different traits. Further studies on the functions of these genes are essential to reveal the mechanism underlying the differentiation between varieties. Our results provided the basis for a comprehensive and systematic understanding of the differentiation of *indica* and *japonica*, and would help explain some important inter-subspecies differences. In addition, each target of positive selection has a story to tell about the historical forces and events that have shaped the history of a population.

### Whole-genome selection screening is necessary to study the differentiation of *indica* and *japonica*

By 4000 years ago, human societies worldwide had completed the domestication of all major crop species [80]. In the past ten or more years, researchers have identified the several specific genes that control some of the most important morphological changes associated with domestication. These genes include *tb1* [81] and *tga1* in maize [82]; *qSH1* [83], *sh4* [16], *Prog1* [12, 13] and *Rc* [84] in rice; *fw2.2* in tomato [85]; and the *Q* gene in wheat [86]. Although only a few domestication genes have been well documented, these analyses provided a great deal of information important to the understanding of how domestication modified plant development to produce today's crops. Nevertheless, these data have not been sufficient to reveal the mechanism of domestication yet. Even fewer genes have been identified as involved in the differentiation of *indica* and *japonica*. Given this background, a whole-genome selection screen is a useful strategy for understanding the domestication and differentiation of *indica* and *japonica*. In this study, we revealed 1068 genes throughout the genome that underwent positive selection during differentiation, but found only 29 genes of known function. There were 15 genes involved in the regulation of biotic and/or abiotic responses; seven genes regulate the germ cell development. All of these genes except *S5* and *Dpl2* have not been reported to be involved in differentiation of *indica* and *japonica* in previous work. In addition, we found other 47 PSGs in response to various environment factors (Table 3 and S11). Our study laid the foundation for further research on evolution of cultivated rice.

## The large differences between the Nipponbare and 93–11 proteomes due to the differences of gene annotations

When the genes in Nipponbare and 93–11 were compared, orthologs could not be found for more than 10,000 genes. This suggested a major difference between the Nipponbare and 93–11 proteomes. However, for most of these genes, highly similar sequences were found in the Nipponbare or 93–11 genome when used as queries to search the other genome. This result implied orthologs could not be found for these genes mainly due to the differences in the annotations of Nipponbare and 93–11. Further evidence to support this view came from the two different systems used to annotate the Nipponbare genes, RAP-DB and RGAP. We used the 93–11 protein sequences as queries to search the RAP-DB and RGAP databases, and obtained 21,884 and 25,538 orthologs, respectively. When these two sets of results were combined, 30,995 pairs of orthologs were discovered. In addition, previous study showed that approximately one-third of the automated annotations contained errors in the NBS-LRR encoding genes in Arabidopsis, and more than one-third in LRR-kinase genes in rice [85, 87]. The results suggested that inadequate gene annotation was the main impediment to finding the orthologous relationships between the Nipponbare and 93–11 genes. In this study, we selected a lower standard to reveal all of the possible orthologs, and then found the PSGs with a higher standard and manually corrected the annotation and ClustalW mistakes. These greatly reduced error rate and workload.

## More genes than those detected were under positive selection during differentiation of *indica* and *japonica*

Several considerations make us to suppose that the actual number of the PSGs during differentiation of *indica* and *japonica* would be far more than that detected in this study. Firstly, some annotation errors brought about that no orthologs were found between 93–11 and Nipponbare. Secondly, some genes are pseudogenes in 93–11 or Nipponbare, but are functional genes in another. For example, the *phr1* gene lost its function due to an 18 nucleotide deletion in the *japonica* lines, but it remained functional in the *indica* lines [25]. The *Phr1* gene encoding a polyphenol oxidase controls the phenol reaction, which is an important trait for distinguishing *indica* and *japonica*. The grains of the *indica* cultivars turn brown after being soaked in phenol solution, whereas those of the *japonica* cultivars do not [2]. The genetic test revealed positive selection for the 18 bp deletion [25]. Unfortunately, our study failed to detect this selection on the *Phr1* gene because the method used in this research was not suited for analyzing deletion. Thirdly, the method based on the average  $K_a/K_s$  ratio over all the sites in a sequence is low powerful to detect positive selection comparing with PAML-codeml because adaptive evolution occurs at only a few sites, as most amino acids in a protein are under structural and functional constraints [88, 89]. That is why only a few PSGs with synonymous sites were revealed and so many NSSGs were discovered in this study. However, the recent data are not fit to PAML-codeml. Fourthly, we adopted a relatively stringent condition, which required that the PSGs should be detected by both of methods at the same time. It led to the results that some supposed PSGs would be excluded because of the stringent condition.

## Artificial selection accelerate *indica-japonica* differentiation

Our results indicated that the differentiation has already started in wild rice, but this differentiation is very low. Only 16.4% PSSs in *indica-japonica* were also detected in Or-It~ Or-IIIIt, and the average  $F_{st}$  value between Or-It and Or-IIIIt is 0.476 comparing with 0.919 between *indica* and temperate *japonica*. In addition, the populations of Or-It and Or-IIIIt only constitute a small part of wild rice. If the differentiation in wild rice was supposed to be driven by natural

selection, the *indica-japonica* differentiation in cultivated rice could be driven by natural and artificial selection. Moreover, artificial selection is much more powerful than natural selection on the differentiation.

## Differentiation of *indica* and *japonica* is one of the most important evolution directions

We used the PSSs to reconstruct the phylogenetic tree, which showed that temperate *japonica* is far from all of wild rice, but *indica* and Or-I were almost clustered together (S2A Fig). This result looks like that in the principal component analysis (PCA) plot of 1529 accessions with ~8 million SNP sites in the Huang's paper (Supplementary Figure 13), in which the *japonica* varieties clearly segregate from the other groups, and some Or-I accessions mixed with the *indica* varieties [10]. This implied that *indica* came from Or-I, whereas *japonica* maybe derived from another wild rice group that is similar to Or-III and not included in the collection. However, Huang et al suggested that *japonica* was first domesticated from Or-III in southern China, and was subsequently crossed to Or-I wild rice in South East Asia and South Asia, thus generating *indica* after many cross-differentiation-selection cycles according to the analysis of domestication loci [10]. That seemed inconsistent with the results from all the SNP data [10] and our tree.

The first component in the PCA plot separated *indica* and *japonica*, and the second component separated *O. sativa* and *O. rufipogon* [10]. The result implied that differentiations of *indica* and *japonica*, and wild and cultivated rice are two main evolution directions. The differentiation of *indica* and *japonica* started in wild rice. The accessions of *japonica* and Or-III distributed over a large range in our tree, whereas that of *indica* and Or-I concentrated together. This inferred that the *japonica* varieties exhibited more diversity than *indica* on differentiation, Or-III than Or-I (S2 Fig). Thus, the study on the origin and evolution of *indica* and *japonica* should consider the power that acted on the differentiation. However, Huang's model neglected the *indica-japonica* evolution direction [10], and thus could not explain the evolution of *indica* and *japonica* well.

## Materials and Methods

### The sequences and SNP data used in the study

The annotations of the genes of Nipponbare (temperate *japonica*) were downloaded from RAP-DB (<http://rapdb.dna.affrc.go.jp/download/irgsp1.html>) and Rice Genome Annotation Project ([ftp://ftp.plantbiology.msu.edu/pub/data/Eukaryotic\\_Projects/o\\_sativa/annotation\\_dbs/pseudomolecules/version\\_6.1/all.dir/](ftp://ftp.plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.1/all.dir/)), respectively. The gene annotations for 93–11 (*indica*) were downloaded from the RISe database ([ftp://ftp.genomics.org.cn/pub/ricedb/rice\\_update\\_data/GLEAN\\_genes/Beijing\\_indica/GLEAN\\_genes/](ftp://ftp.genomics.org.cn/pub/ricedb/rice_update_data/GLEAN_genes/Beijing_indica/GLEAN_genes/)). The genome DNA sequences of Nipponbare and 93–11 were obtained from the IRGSP/RAP build 5 dataset (<http://rapdblegacy.dna.affrc.go.jp/download/index.html>) and ricedb/RGPVs9311/9311 ([ftp://ftp.genomics.org.cn/pub/ricedb/rice\\_update\\_data/genome/9311/](ftp://ftp.genomics.org.cn/pub/ricedb/rice_update_data/genome/9311/)), respectively. All of the SNP data were download from Rice Haplotype Map Project Database (<http://www.ncgr.ac.cn/RiceHap2/index.html>).

### Sequence alignment and the discovery of the orthologous genes between *indica* and *japonica*

The stand-alone BLAST programs (ncbi-blast-2.2.24+.exe; <ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/>) were used to search each database with default parameters. The BLAST results were parsed with the Bio::SearchIO module in perl (<http://search.cpan.org/~>

[cjfields/BioPerl-1.6.901/Bio/SearchIO.pm](#)). The results were manually treated with EXCEL to find the most similar genes on the collinear regions of the 93–11 and Nipponbare chromosomes, which were considered orthologous. The pairs of orthologs were aligned with ClustalW2 [90] using the Bio::Tools::Run::Alignment::Clustalw module (<http://search.cpan.org/~cjfields/BioPerl-Run-1.006900/lib/Bio/Tools/Run/Alignment/Clustalw.pm>). The percent identity values were calculated according to the ClustalW2 results. To convert a multiple sequence alignment of proteins to a codon alignment of the corresponding DNA sequences, PAL2NAL [91] was implemented using the Bio::Tools::Run::Alignment::Pal2Nal module (<http://search.cpan.org/~cjfields/BioPerl-Run-1.006900/lib/Bio/Tools/Run/Alignment/Pal2Nal.pm>).

## Analysis of selective pressure

Yang and Nielsen (YN) [92], maximum likelihood (ML) [93], Nei & Gojobori (NG) [94], MYN [95], gNG [96], and gMYN [96] methods were adopted to estimate the synonymous and nonsynonymous substitution rates in pairwise comparisons of protein-coding DNA sequences. The program PAML [97], Bio::Tools::Run::Phylo::PAML::Codeml (<http://search.cpan.org/~cjfields/BioPerl-Run-1.006900/lib/Bio/Tools/Run/Phylo/PAML/Codeml.pm>) and Bio::Tools::Run::Phylo::PAML::Yn00 modules (<http://search.cpan.org/~cjfields/BioPerl-Run-1.006900/lib/Bio/Tools/Run/Phylo/PAML/Yn00.pm>) were used to implement the YN, ML and NG methods; KaKs\_calculator 2.0 [29] was used to execute the gNG, MYN and gMYN methods. The outcomes of these methods were compared, and only the results of gNG and MYN were further analyzed. The outputs were ordered by the Ka/Ks value and by the significance level of the associated Fisher's test, which indicates whether the Ka/Ks ratio is significantly different from one.

The positions of the exons of the PSGs were revealed with Blast, and the SNPs in exons were selected, and their Fsts were calculated with EXCEL. All SNPs with more than 0.25 Fst values were manually checked to see if they changed the amino acids with the help of Sequencher (<http://www.genecodes.com>) and DNAsis Max Trial 1.0 (<http://www.miraibio.com/download/>).

Lositan, a Fst related statistic method, was employed to identify outliers for selection detection [98]. The phylogenetic tree was constructed with MEGA6 (<http://megasoftware.net/>).

## GO analysis

The GO annotations of the PSGs were retrieved from Rice Genome Annotation Project (<http://rice.plantbiology.msu.edu/>) and RAP-DB, or searched with Blast2GO using the default threshold [99, 100]. The GO annotations for biological processes, molecular functions and cellular component categories were classified into different groups with OBO-edit version 2.1.1 (<http://sourceforge.net/projects/geneontology/files/>). The chromosomal positions of the PSGs were recovered by searching the Nipponbare genome sequences, parsing the sequences with the Bio::SearchIO module and then mapping the resulting data with EXCEL.

## Supporting Information

**S1 Table. List of the genes whose Kss are zero and Kas are above zero.**

(XLSX)

**S2 Table. The positively selected genes detected with KaKs\_calculator, whose Kss are more than zero.**

(XLSX)



**S3 Table. The positively selected genes detected with KaKs\_calculator on sliding windows of fixed size.**

(XLSX)

**S4 Table. The NSSGs with outliers in chromosome 1.**

(XLSX)

**S5 Table. The NSSGs with outliers and synonymous sites between *indica* and temperate *japonica* in chromosome 1.**

(XLSX)

**S6 Table. The NSSGs with outliers in rice genome except chromosome 1.**

(XLSX)

**S7 Table. The genes that were selected as PSGs contained one synonymous site or nonsense mutant site.**

(XLSX)

**S8 Table. The possible functions of the PSGs.**

(DOCX)

**S9 Table. The molecular functions of the positively selected genes according to GO analysis.**

(XLSX)

**S10 Table. The biological processes of the positively selected genes according to GO analysis.**

(XLSX)

**S11 Table. The PSGs with the expression pattern found in the NCBI and Google scholar database.**

(DOCX)

**S12 Table. The comparison of the outliers revealed between Or-It and Or-IIIIt, and between *indica* and temperate *japonica*.**

(XLSX)

**S13 Table. The Molecular functions and biological processes of the positively selected genes between Or-It and Or-IIIIt according to GO analysis.**

(XLSX)

**S1 Fig. The ratios of PSGs to total genes in each Mb along the chromosomes.**

(TIF)

**S2 Fig. The population differentiation in 1529 rice accessions.** (A) Neighbor-joining tree of 446 *O. rufipogon* accessions and 1,083 *O. sativa* varieties constructed with the PSSs. The five divergent groups, *indica*, *aus*, temperate *japonica*, tropical *japonica* and intermediate were indicated with different colors. The scale bar indicates the simple matching distance. (B) Neighbor-joining tree of 446 *O. rufipogon* accessions constructed with the PSSs.

(TIF)

## Acknowledgments

We wish to thank Miss Lihua Zhang, a PhD student at the South Dakota University, for her assistance with Perl programming, and thank Dr. Bin Han's group for their SNP data of 1529 *O. sativa* and *O. rufipogon* accessions.

## Author Contributions

Conceived and designed the experiments: XS. Performed the experiments: XS. Analyzed the data: XS YG XZ QJ KL. Contributed reagents/materials/analysis tools: XS. Wrote the paper: XS QJ. XS.

## References

1. Khush GS. Origin, dispersal, cultivation and variation of rice. *Plant molecular biology*. 1997; 35(1–2):25–34. Epub 1997/09/18.
2. Oka HI, editor. Origin of cultivated rice. Tokyo: Japan Scientific Societies Press; 1988.
3. Morishima H, Oka H. Phylogenetic differentiation of cultivated rice. XXII. Numerical evaluation of the Indica—Japonica differentiation. *Jpn J Breed*. 1981; 31:402–13.
4. Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S. Genetic structure and diversity in *Oryza sativa* L. *Genetics*. 2005; 169(3):1631–8. Epub 2005/01/18. [genetics.104.035642](https://doi.org/10.1534/genetics.104.035642) [pii] doi: [10.1534/genetics.104.035642](https://doi.org/10.1534/genetics.104.035642) PMID: [15654106](https://pubmed.ncbi.nlm.nih.gov/15654106/)
5. Glaszmann JC. Isozymes and classification of Asian rice varieties. *Theor Appl Genet*. 1987; 74:21–30. doi: [10.1007/BF00290078](https://doi.org/10.1007/BF00290078) PMID: [24241451](https://pubmed.ncbi.nlm.nih.gov/24241451/)
6. Second G. Evolutionary relationships in the sativa group of *Oryza* based on isozyme data. *Genet Sel Evol*. 1985; 17:89–114. doi: [10.1186/1297-9686-17-1-89](https://doi.org/10.1186/1297-9686-17-1-89) PMID: [22879188](https://pubmed.ncbi.nlm.nih.gov/22879188/)
7. Wang ZY, Tanksley SD. Restriction fragment length polymorphism in *Oryza sativa* L. *Genome*. 1989; 32(6):1113–8.
8. Zhang P, Li J, Li X, Liu X, Zhao X, Lu Y. Population structure and genetic diversity in a rice core collection (*Oryza sativa* L.) investigated with SSR markers. *PloS one*. 2011; 6(12):e27565. Epub 2011/12/14 doi: [10.1371/journal.pone.0027565](https://doi.org/10.1371/journal.pone.0027565) PMID: [22164211](https://pubmed.ncbi.nlm.nih.gov/22164211/)
9. Huang X, Zhao Y, Wei X, Li C, Wang A, Zhao Q, et al. Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature genetics*. 2012; 44(1):32–9. doi: [10.1038/ng.1018](https://doi.org/10.1038/ng.1018)
10. Huang X, Kurata N, Wei X, Wang ZX, Wang A, Zhao Q, et al. A map of rice genome variation reveals the origin of cultivated rice. *Nature*. 2012; 490(7421):497–501. doi: [10.1038/nature11532](https://doi.org/10.1038/nature11532) PMID: [23034647](https://pubmed.ncbi.nlm.nih.gov/23034647/)
11. Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, et al. Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(20):8351–6. Epub 2011/05/04. 1104686108 [pii] doi: [10.1073/pnas.1104686108](https://doi.org/10.1073/pnas.1104686108) PMID: [21536870](https://pubmed.ncbi.nlm.nih.gov/21536870/)
12. Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, et al. Control of a key transition from prostrate to erect growth in rice domestication. *Nature genetics*. 2008; 40(11):1360–4. Epub 2008/09/30. ng.197 [pii] doi: [10.1038/ng.197](https://doi.org/10.1038/ng.197) PMID: [18820699](https://pubmed.ncbi.nlm.nih.gov/18820699/)
13. Jin J, Huang W, Gao JP, Yang J, Shi M, Zhu MZ, et al. Genetic control of rice plant architecture under domestication. *Nature genetics*. 2008; 40(11):1365–9. Epub 2008/09/30. ng.247 [pii] doi: [10.1038/ng.247](https://doi.org/10.1038/ng.247) PMID: [18820696](https://pubmed.ncbi.nlm.nih.gov/18820696/)
14. Zhu BF, Si L, Wang Z, Zhou Y, Zhu J, Shangguan Y, et al. Genetic control of a transition from black to straw-white seed hull in rice domestication. *Plant physiology*. 2011; 155(3):1301–11. doi: [10.1104/pp.110.168500](https://doi.org/10.1104/pp.110.168500) PMID: [21263038](https://pubmed.ncbi.nlm.nih.gov/21263038/)
15. Zhang LB, Zhu Q, Wu ZQ, Ross-Ibarra J, Gaut BS, Ge S, et al. Selection on grain shattering genes and rates of rice domestication. *The New phytologist*. 2009; 184(3):708–20. Epub 2009/08/14. NPH2984 [pii] doi: [10.1111/j.1469-8137.2009.02984.x](https://doi.org/10.1111/j.1469-8137.2009.02984.x) PMID: [19674325](https://pubmed.ncbi.nlm.nih.gov/19674325/)
16. Li C, Zhou A, Sang T. Rice domestication by reducing shattering. *Science*. 2006; 311(5769):1936–9. Epub 2006/03/11. 1123604 [pii] doi: [10.1126/science.1123604](https://doi.org/10.1126/science.1123604) PMID: [16527928](https://pubmed.ncbi.nlm.nih.gov/16527928/)
17. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fedel-Alon A, York TL, et al. Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS genetics*. 2007; 3(9):1745–56. Epub 2007/10/03. 07-PLGE-RA-0120 [pii] doi: [10.1371/journal.pgen.0030163](https://doi.org/10.1371/journal.pgen.0030163) PMID: [17907810](https://pubmed.ncbi.nlm.nih.gov/17907810/)
18. Cheng C, Motohashi R, Tsuchimoto S, Fukuta Y, Ohtsubo H, Ohtsubo E. Polyphyletic origin of cultivated rice: based on the interspersed pattern of SINEs. *Mol Biol Evol*. 2003; 20(1):67–75. Epub 2003/01/10 PMID: [12519908](https://pubmed.ncbi.nlm.nih.gov/12519908/)
19. Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(25):9578–83. Epub 2006/06/13. 0603152103 [pii] doi: [10.1073/pnas.0603152103](https://doi.org/10.1073/pnas.0603152103) PMID: [16766658](https://pubmed.ncbi.nlm.nih.gov/16766658/)

20. Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A, et al. Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. *Theor Appl Genet.* 2007; 114(4):731–43. Epub 2007/01/16 doi: [10.1007/s00122-006-0473-1](https://doi.org/10.1007/s00122-006-0473-1) PMID: [17219210](https://pubmed.ncbi.nlm.nih.gov/17219210/)
21. Tang T, Lu J, Huang J, He J, McCouch SR, Shen Y, et al. Genomic variation in rice: genesis of highly polymorphic linkage blocks during domestication. *PLoS genetics.* 2006; 2(11):e199. Epub 2006/11/23. 06-PLGE-RA-0344R2 [pii] doi: [10.1371/journal.pgen.0020199](https://doi.org/10.1371/journal.pgen.0020199) PMID: [17112320](https://pubmed.ncbi.nlm.nih.gov/17112320/)
22. Zhu Q, Ge S. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *The New phytologist.* 2005; 167(1):249–65. Epub 2005/06/14. NPH1406 [pii] doi: [10.1111/j.1469-8137.2005.01406.x](https://doi.org/10.1111/j.1469-8137.2005.01406.x) PMID: [15948847](https://pubmed.ncbi.nlm.nih.gov/15948847/)
23. He Z, Zhai W, Wen H, Tang T, Wang Y, Lu X, et al. Two evolutionary histories in the genome of rice: the roles of domestication genes. *PLoS genetics.* 2011; 7(6):e1002100. Epub 2011/06/23 doi: [10.1371/journal.pgen.1002100](https://doi.org/10.1371/journal.pgen.1002100) PGENETICS-D-10-00553 [pii] PMID: [21695282](https://pubmed.ncbi.nlm.nih.gov/21695282/)
24. Huang CL, Hung CY, Chiang YC, Hwang CC, Hsu TW, Huang CC, et al. Footprints of natural and artificial selection for photoperiod pathway genes in *Oryza*. *The Plant journal: for cell and molecular biology.* 2012; 70(5):769–82. doi: [10.1111/j.1365-313X.2012.04915.x](https://doi.org/10.1111/j.1365-313X.2012.04915.x)
25. Yu Y, Tang T, Qian Q, Wang Y, Yan M, Zeng D, et al. Independent losses of function in a polyphenol oxidase in rice: differentiation in grain discoloration between subspecies and the role of positive selection under domestication. *The Plant cell.* 2008; 20(11):2946–59. Epub 2008/11/27. tpc.108.060426 [pii] doi: [10.1105/tpc.108.060426](https://doi.org/10.1105/tpc.108.060426) PMID: [19033526](https://pubmed.ncbi.nlm.nih.gov/19033526/)
26. Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD. Selection under domestication: evidence for a sweep in the rice waxy genomic region. *Genetics.* 2006; 173(2):975–83. doi: [10.1534/genetics.106.056473](https://doi.org/10.1534/genetics.106.056473) PMID: [16547098](https://pubmed.ncbi.nlm.nih.gov/16547098/)
27. Takano-Kai N, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, et al. Evolutionary history of GS3, a gene conferring grain length in rice. *Genetics.* 2009; 182(4):1323–34. doi: [10.1534/genetics.109.103002](https://doi.org/10.1534/genetics.109.103002) PMID: [19506305](https://pubmed.ncbi.nlm.nih.gov/19506305/)
28. Project IRGS. The map-based sequence of the rice genome. *Nature.* 2005; 436(7052):793–800. Epub 2005/08/16 doi: [10.1038/nature03895](https://doi.org/10.1038/nature03895) PMID: [16100779](https://pubmed.ncbi.nlm.nih.gov/16100779/)
29. Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J. KaKs\_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics.* 2006; 4(4):259–63. Epub 2007/05/29. S1672-0229(07)60007-2 [pii] doi: [10.1016/S1672-0229\(07\)60007-2](https://doi.org/10.1016/S1672-0229(07)60007-2) PMID: [17531802](https://pubmed.ncbi.nlm.nih.gov/17531802/)
30. Wright S. The genetical structure of populations. *Annals of Eugenics.* 1951; 15:323–54. PMID: [24540312](https://pubmed.ncbi.nlm.nih.gov/24540312/)
31. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution.* 1965; 19:395–420.
32. Wright S. *Variability Within and Among Natural Populations*: Univ. of Chicago Press, Chicago; 1978.
33. Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, et al. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. *Proceedings of the National Academy of Sciences of the United States of America.* 2008; 105(32):11436–41. doi: [10.1073/pnas.0804761105](https://doi.org/10.1073/pnas.0804761105) PMID: [18678896](https://pubmed.ncbi.nlm.nih.gov/18678896/)
34. Gong X, Jiang Q, Xu J, Zhang J, Teng S, Lin D, et al. Disruption of the rice plastid ribosomal protein s20 leads to chloroplast developmental defects and seedling lethality. *G3.* 2013; 3(10):1769–77. doi: [10.1534/g3.113.007856](https://doi.org/10.1534/g3.113.007856) PMID: [23979931](https://pubmed.ncbi.nlm.nih.gov/23979931/)
35. Park MR, Yun KY, Mohanty B, Herath V, Xu F, Wijaya E, et al. Supra-optimal expression of the cold-regulated OsMyb4 transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. *Plant, cell & environment.* 2010; 33(12):2209–30. doi: [10.1111/j.1365-3040.2010.02221.x](https://doi.org/10.1111/j.1365-3040.2010.02221.x)
36. Vannini C, Locatelli F, Bracale M, Magnani E, Marsoni M, Osnato M, et al. Overexpression of the rice Osmyb4 gene increases chilling and freezing tolerance of *Arabidopsis thaliana* plants. *The Plant journal: for cell and molecular biology.* 2004; 37(1):115–27. PMID: [14675437](https://pubmed.ncbi.nlm.nih.gov/14675437/)
37. Obata T, Kitamoto HK, Nakamura A, Fukuda A, Tanaka Y. Rice shaker potassium channel OsKAT1 confers tolerance to salinity stress on yeast and rice cells. *Plant physiology.* 2007; 144(4):1978–85. doi: [10.1104/pp.107.101154](https://doi.org/10.1104/pp.107.101154) PMID: [17586689](https://pubmed.ncbi.nlm.nih.gov/17586689/)
38. Kaneda T, Taga Y, Takai R, Iwano M, Matsui H, Takayama S, et al. The transcription factor OsNAC4 is a key positive regulator of plant hypersensitive cell death. *The EMBO journal.* 2009; 28(7):926–36. doi: [10.1038/emboj.2009.39](https://doi.org/10.1038/emboj.2009.39) PMID: [19229294](https://pubmed.ncbi.nlm.nih.gov/19229294/)

39. Taga Y, Takai R, Kaneda T, Matsui H, Isogai A, Che FS. Role of OsHSP90 and IREN, Ca<sup>2+</sup>-dependent nuclease, in plant hypersensitive cell death induced by transcription factor OsNAC4. *Plant signaling & behavior*. 2009; 4(8):740–2. doi: [10.1038/emboj.2009.39](https://doi.org/10.1038/emboj.2009.39)
40. Kawahara Y, Oono Y, Kanamori H, Matsumoto T, Itoh T, Minami E. Simultaneous RNA-seq analysis of a mixed transcriptome of rice and blast fungus interaction. *PloS one*. 2012; 7(11):e49423. doi: [10.1371/journal.pone.0049423](https://doi.org/10.1371/journal.pone.0049423) PMID: [23139845](https://pubmed.ncbi.nlm.nih.gov/23139845/)
41. Komatsu M, Maekawa M, Shimamoto K, Kyojuka J. The LAX1 and FRIZZY PANICLE 2 genes determine the inflorescence architecture of rice by controlling rachis-branch and spikelet development. *Developmental biology*. 2001; 231(2):364–73. doi: [10.1006/dbio.2000.9988](https://doi.org/10.1006/dbio.2000.9988) PMID: [11237465](https://pubmed.ncbi.nlm.nih.gov/11237465/)
42. Yang F, Wang Q, Schmitz G, Muller D, Theres K. The bHLH protein ROX acts in concert with RAX1 and LAS to modulate axillary meristem formation in Arabidopsis. *The Plant journal: for cell and molecular biology*. 2012; 71(1):61–70. doi: [10.1111/j.1365-3113X.2012.04970.x](https://doi.org/10.1111/j.1365-3113X.2012.04970.x) PMID: [22372440](https://pubmed.ncbi.nlm.nih.gov/22372440/)
43. Mach J. Rice axillary meristem formation requires directional movement of LAX PANICLE1 protein. *The Plant cell*. 2009; 21(4):1027. doi: [10.1105/tpc.109.210410](https://doi.org/10.1105/tpc.109.210410) PMID: [19346463](https://pubmed.ncbi.nlm.nih.gov/19346463/)
44. Oikawa T, Kyojuka J. Two-Step Regulation of LAX PANICLE1 Protein Accumulation in Axillary Meristem Formation in Rice. *The Plant cell*. 2009; 21(4):1095–108. doi: [10.1105/tpc.108.065425](https://doi.org/10.1105/tpc.108.065425) PMID: [19346465](https://pubmed.ncbi.nlm.nih.gov/19346465/)
45. Amir Hossain M, Lee Y, Cho JI, Ahn CH, Lee SK, Jeon JS, et al. The bZIP transcription factor OsABF1 is an ABA responsive element binding factor that enhances abiotic stress signaling in rice. *Plant molecular biology*. 2010; 72(4–5):557–66. doi: [10.1007/s11103-009-9592-9](https://doi.org/10.1007/s11103-009-9592-9) PMID: [20087755](https://pubmed.ncbi.nlm.nih.gov/20087755/)
46. Matsui H, Miyao A, Takahashi A, Hirochika H. Pdk1 kinase regulates basal disease resistance through the OsOxi1-OsPti1a phosphorylation cascade in rice. *Plant & cell physiology*. 2010; 51(12):2082–91. doi: [10.1093/pcp/pcq167](https://doi.org/10.1093/pcp/pcq167)
47. Luo H, Lee JY, Hu Q, Nelson-Vasilchik K, Eitas TK, Lickwar C, et al. RTS, a rice anther-specific gene is required for male fertility and its promoter sequence directs tissue-specific gene expression in different plant species. *Plant molecular biology*. 2006; 62(3):397–408. doi: [10.1007/s11103-006-9031-0](https://doi.org/10.1007/s11103-006-9031-0) PMID: [16897470](https://pubmed.ncbi.nlm.nih.gov/16897470/)
48. Passaia G, Fonini LS, Caverzan A, Jardim-Messeder D, Christoff AP, Gaeta ML, et al. The mitochondrial glutathione peroxidase GPX3 is essential for H<sub>2</sub>O<sub>2</sub> homeostasis and root and shoot development in rice. *Plant science: an international journal of experimental plant biology*. 2013; 208:93–101. doi: [10.1016/j.plantsci.2013.03.017](https://doi.org/10.1016/j.plantsci.2013.03.017)
49. Xiang Y, Tang N, Du H, Ye H, Xiong L. Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. *Plant physiology*. 2008; 148(4):1938–52. doi: [10.1104/pp.108.128199](https://doi.org/10.1104/pp.108.128199) PMID: [18931143](https://pubmed.ncbi.nlm.nih.gov/18931143/)
50. Yoshii M, Shimizu T, Yamazaki M, Higashi T, Miyao A, Hirochika H, et al. Disruption of a novel gene for a NAC-domain protein in rice confers resistance to Rice dwarf virus. *The Plant journal: for cell and molecular biology*. 2009; 57(4):615–25. doi: [10.1111/j.1365-3113X.2008.03712.x](https://doi.org/10.1111/j.1365-3113X.2008.03712.x) PMID: [18980655](https://pubmed.ncbi.nlm.nih.gov/18980655/)
51. Wang Q, Li J, Hu L, Zhang T, Zhang G, Lou Y. OsMPK3 positively regulates the JA signaling pathway and plant resistance to a chewing herbivore in rice. *Plant cell reports*. 2013; 32(7):1075–84. doi: [10.1007/s00299-013-1389-2](https://doi.org/10.1007/s00299-013-1389-2) PMID: [23344857](https://pubmed.ncbi.nlm.nih.gov/23344857/)
52. Xie G, Kato H, Sasaki K, Imai R. A cold-induced thioredoxin h of rice, OsTrx23, negatively regulates kinase activities of OsMPK3 and OsMPK6 in vitro. *FEBS letters*. 2009; 583(17):2734–8. doi: [10.1016/j.febslet.2009.07.057](https://doi.org/10.1016/j.febslet.2009.07.057) PMID: [19665023](https://pubmed.ncbi.nlm.nih.gov/19665023/)
53. Xie G, Kato H, Imai R. Biochemical identification of the OsMKK6-OsMPK3 signalling pathway for chilling stress tolerance in rice. *The Biochemical journal*. 2012; 443(1):95–102. doi: [10.1042/BJ20111792](https://doi.org/10.1042/BJ20111792) PMID: [22248149](https://pubmed.ncbi.nlm.nih.gov/22248149/)
54. Rao KP, Vani G, Kumar K, Wankhede DP, Misra M, Gupta M, et al. Arsenic stress activates MAP kinase in rice roots and leaves. *Archives of biochemistry and biophysics*. 2011; 506(1):73–82. doi: [10.1016/j.abb.2010.11.006](https://doi.org/10.1016/j.abb.2010.11.006) PMID: [21081102](https://pubmed.ncbi.nlm.nih.gov/21081102/)
55. Kishi-Kaboshi M, Okada K, Kurimoto L, Murakami S, Umezawa T, Shibuya N, et al. A rice fungal MAMP-responsive MAPK cascade regulates metabolic flow to antimicrobial metabolite synthesis. *The Plant journal: for cell and molecular biology*. 2010; 63(4):599–612. doi: [10.1111/j.1365-3113X.2010.04264.x](https://doi.org/10.1111/j.1365-3113X.2010.04264.x) PMID: [20525005](https://pubmed.ncbi.nlm.nih.gov/20525005/)
56. Park CJ, Peng Y, Chen X, Dardick C, Ruan D, Bart R, et al. Rice XB15, a protein phosphatase 2C, negatively regulates cell death and XA21-mediated innate immunity. *PLoS biology*. 2008; 6(9):e231. doi: [10.1371/journal.pbio.0060231](https://doi.org/10.1371/journal.pbio.0060231) PMID: [18817453](https://pubmed.ncbi.nlm.nih.gov/18817453/)
57. Schmidt R, Mieulet D, Hubberten HM, Obata T, Hoefgen R, Fernie AR, et al. Salt-responsive ERF1 regulates reactive oxygen species-dependent signaling during the initial response to salt stress in rice. *The Plant cell*. 2013; 25(6):2115–31. doi: [10.1105/tpc.113.113068](https://doi.org/10.1105/tpc.113.113068) PMID: [23800963](https://pubmed.ncbi.nlm.nih.gov/23800963/)

58. Mizuta Y, Harushima Y, Kurata N. Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(47):20417–22. doi: [10.1073/pnas.1003124107](https://doi.org/10.1073/pnas.1003124107) PMID: [21048083](https://pubmed.ncbi.nlm.nih.gov/21048083/)
59. Morita R, Sato Y, Masuda Y, Nishimura M, Kusaba M. Defect in non-yellow coloring 3, an alpha/beta hydrolase-fold family protein, causes a stay-green phenotype during leaf senescence in rice. *The Plant journal: for cell and molecular biology*. 2009; 59(6):940–52. doi: [10.1111/j.1365-313X.2009.03919.x](https://doi.org/10.1111/j.1365-313X.2009.03919.x)
60. Ookawa T, Hobo T, Yano M, Murata K, Ando T, Miura H, et al. New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. *Nature communications*. 2010; 1:132. doi: [10.1038/ncomms1132](https://doi.org/10.1038/ncomms1132) PMID: [21119645](https://pubmed.ncbi.nlm.nih.gov/21119645/)
61. Ikeda-Kawakatsu K, Yasuno N, Oikawa T, Iida S, Nagato Y, Maekawa M, et al. Expression level of ABERRANT PANICLE ORGANIZATION1 determines rice inflorescence form through control of cell proliferation in the meristem. *Plant physiology*. 2009; 150(2):736–47. doi: [10.1104/pp.109.136739](https://doi.org/10.1104/pp.109.136739) PMID: [19386809](https://pubmed.ncbi.nlm.nih.gov/19386809/)
62. Ikeda K, Nagasawa N, Nagato Y. ABERRANT PANICLE ORGANIZATION 1 temporally regulates meristem identity in rice. *Developmental biology*. 2005; 282(2):349–60. doi: [10.1016/j.ydbio.2005.03.016](https://doi.org/10.1016/j.ydbio.2005.03.016) PMID: [15950602](https://pubmed.ncbi.nlm.nih.gov/15950602/)
63. Ikeda K, Ito M, Nagasawa N, Kyozuka J, Nagato Y. Rice ABERRANT PANICLE ORGANIZATION 1, encoding an F-box protein, regulates meristem fate. *The Plant journal: for cell and molecular biology*. 2007; 51(6):1030–40. doi: [10.1111/j.1365-313X.2007.03200.x](https://doi.org/10.1111/j.1365-313X.2007.03200.x) PMID: [17666027](https://pubmed.ncbi.nlm.nih.gov/17666027/)
64. Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, et al. SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice. *Development*. 2003; 130(4):705–18. PMID: [12506001](https://pubmed.ncbi.nlm.nih.gov/12506001/)
65. Lee S, Jeon JS, An K, Moon YH, Lee S, Chung YY, et al. Alteration of floral organ identity in rice through ectopic expression of OsMADS16. *Planta*. 2003; 217(6):904–11. doi: [10.1007/s00425-003-1066-8](https://doi.org/10.1007/s00425-003-1066-8) PMID: [12905025](https://pubmed.ncbi.nlm.nih.gov/12905025/)
66. Ning J, Zhang B, Wang N, Zhou Y, Xiong L. Increased leaf angle1, a Raf-like MAPKKK that interacts with a nuclear protein family, regulates mechanical tissue formation in the Lamina joint of rice. *The Plant cell*. 2011; 23(12):4334–47. doi: [10.1105/tpc.111.093419](https://doi.org/10.1105/tpc.111.093419) PMID: [22207574](https://pubmed.ncbi.nlm.nih.gov/22207574/)
67. Fitzgerald HA, Canlas PE, Chern MS, Ronald PC. Alteration of TGA factor activity in rice results in enhanced tolerance to *Xanthomonas oryzae* pv. *oryzae*. *The Plant journal: for cell and molecular biology*. 2005; 43(3):335–47. doi: [10.1111/j.1365-313X.2005.02457.x](https://doi.org/10.1111/j.1365-313X.2005.02457.x) PMID: [16045470](https://pubmed.ncbi.nlm.nih.gov/16045470/)
68. Park HY, Kang IS, Han JS, Lee CH, An G, Moon YH. OsDEG10 encoding a small RNA-binding protein is involved in abiotic stress signaling. *Biochemical and biophysical research communications*. 2009; 380(3):597–602. doi: [10.1016/j.bbrc.2009.01.131](https://doi.org/10.1016/j.bbrc.2009.01.131) PMID: [19285007](https://pubmed.ncbi.nlm.nih.gov/19285007/)
69. Lazzarotto F, Teixeira FK, Rosa SB, Dunand C, Fernandes CL, Fontenele Ade V, et al. Ascorbate peroxidase-related (APx-R) is a new heme-containing protein functionally associated with ascorbate peroxidase but evolutionarily divergent. *The New phytologist*. 2011; 191(1):234–50. doi: [10.1111/j.1469-8137.2011.03659.x](https://doi.org/10.1111/j.1469-8137.2011.03659.x) PMID: [21352234](https://pubmed.ncbi.nlm.nih.gov/21352234/)
70. Yamada S, Kano A, Tamaoki D, Miyamoto A, Shishido H, Miyoshi S, et al. Involvement of OsJAZ8 in jasmonate-induced resistance to bacterial blight in rice. *Plant & cell physiology*. 2012; 53(12):2060–72. doi: [10.1093/pcp/pcs145](https://doi.org/10.1093/pcp/pcs145)
71. Kunihiro S, Saito T, Matsuda T, Inoue M, Kuramata M, Taguchi-Shiobara F, et al. Rice DEP1, encoding a highly cysteine-rich G protein gamma subunit, confers cadmium tolerance on yeast cells and plants. *Journal of experimental botany*. 2013; 64(14):4517–27. doi: [10.1093/jxb/ert267](https://doi.org/10.1093/jxb/ert267) PMID: [24163402](https://pubmed.ncbi.nlm.nih.gov/24163402/)
72. Huang X, Qian Q, Liu Z, Sun H, He S, Luo D, et al. Natural variation at the DEP1 locus enhances grain yield in rice. *Nature genetics*. 2009; 41(4):494–7. doi: [10.1038/ng.352](https://doi.org/10.1038/ng.352) PMID: [19305410](https://pubmed.ncbi.nlm.nih.gov/19305410/)
73. Zhang T, Zhao X, Wang W, Pan Y, Huang L, Liu X, et al. Comparative transcriptome profiling of chilling stress responsiveness in two contrasting rice genotypes. *PloS one*. 2012; 7(8):e43274. doi: [10.1371/journal.pone.0043274](https://doi.org/10.1371/journal.pone.0043274) PMID: [22912843](https://pubmed.ncbi.nlm.nih.gov/22912843/)
74. Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, et al. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *The Plant journal: for cell and molecular biology*. 2003; 33(4):751–63. PMID: [12609047](https://pubmed.ncbi.nlm.nih.gov/12609047/)
75. Figueiredo DD, Barros PM, Cordeiro AM, Serra TS, Lourenco T, Chander S, et al. Seven zinc-finger transcription factors are novel regulators of the stress responsive gene OsDREB1B. *Journal of experimental botany*. 2012; 63(10):3643–56. doi: [10.1093/jxb/ers035](https://doi.org/10.1093/jxb/ers035) PMID: [22412187](https://pubmed.ncbi.nlm.nih.gov/22412187/)

76. Gutha LR, Reddy AR. Rice DREB1B promoter shows distinct stress-specific responses, and the over-expression of cDNA in tobacco confers improved abiotic and biotic stress tolerance. *Plant molecular biology*. 2008; 68(6):533–55. doi: [10.1007/s11103-008-9391-8](https://doi.org/10.1007/s11103-008-9391-8) PMID: [18754079](https://pubmed.ncbi.nlm.nih.gov/18754079/)
77. Zheng J, Wang Y, He Y, Zhou J, Li Y, Liu Q, et al. Overexpression of an S-like ribonuclease gene, OsRNS4, confers enhanced tolerance to high salinity and hyposensitivity to phytochrome-mediated light signals in rice. *Plant science: an international journal of experimental plant biology*. 2014; 214:99–105. doi: [10.1016/j.plantsci.2013.10.003](https://doi.org/10.1016/j.plantsci.2013.10.003)
78. Nonomura K, Eiguchi M, Nakano M, Takashima K, Komeda N, Fukuchi S, et al. A novel RNA-recognition-motif protein is required for premeiotic G1/S-phase transition in rice (*Oryza sativa* L.). *PLoS genetics*. 2011; 7(1):e1001265. doi: [10.1371/journal.pgen.1001265](https://doi.org/10.1371/journal.pgen.1001265) PMID: [21253568](https://pubmed.ncbi.nlm.nih.gov/21253568/)
79. Chang T-T, editor. *Origin, domestication, and diversification*. New Jersey USA: J. Wiley and Sons, Inc.; 2003.
80. Doebley JF, Gaut BS, Smith BD. The molecular genetics of crop domestication. *Cell*. 2006; 127(7):1309–21. Epub 2006/12/28. S0092-8674(06)01592-3 [pii] doi: [10.1016/j.cell.2006.12.006](https://doi.org/10.1016/j.cell.2006.12.006) PMID: [17190597](https://pubmed.ncbi.nlm.nih.gov/17190597/)
81. Doebley J, Stec A, Hubbard L. The evolution of apical dominance in maize. *Nature*. 1997; 386(6624):485–8. Epub 1997/04/03 doi: [10.1038/386485a0](https://doi.org/10.1038/386485a0) PMID: [9087405](https://pubmed.ncbi.nlm.nih.gov/9087405/)
82. Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, et al. The origin of the naked grains of maize. *Nature*. 2005; 436(7051):714–9. Epub 2005/08/05. nature03863 [pii] doi: [10.1038/nature03863](https://doi.org/10.1038/nature03863) PMID: [16079849](https://pubmed.ncbi.nlm.nih.gov/16079849/)
83. Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, et al. An SNP caused loss of seed shattering during rice domestication. *Science*. 2006; 312(5778):1392–6. Epub 2006/04/15. 1126410 [pii] doi: [10.1126/science.1126410](https://doi.org/10.1126/science.1126410) PMID: [16614172](https://pubmed.ncbi.nlm.nih.gov/16614172/)
84. Sweeney MT, Thomson MJ, Pfeil BE, McCouch S. Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *The Plant cell*. 2006; 18(2):283–94. Epub 2006/01/10. tpc.105.038430 [pii] doi: [10.1105/tpc.105.038430](https://doi.org/10.1105/tpc.105.038430) PMID: [16399804](https://pubmed.ncbi.nlm.nih.gov/16399804/)
85. Frary A, Nesbitt TC, Grandillo S, Knaap E, Cong B, Liu J, et al. fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. *Science*. 2000; 289(5476):85–8. Epub 2000/07/07. 8651 [pii] PMID: [10884229](https://pubmed.ncbi.nlm.nih.gov/10884229/)
86. Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai YS, Gill BS, et al. Molecular characterization of the major wheat domestication gene Q. *Genetics*. 2006; 172(1):547–55. Epub 2005/09/21. genetics.105.044727 [pii] doi: [10.1534/genetics.105.044727](https://doi.org/10.1534/genetics.105.044727) PMID: [16172507](https://pubmed.ncbi.nlm.nih.gov/16172507/)
87. Sun X, Wang GL. Genome-wide identification, characterization and phylogenetic analysis of the rice LRR-kinases. *PloS one*. 2011; 6(3):e16079. doi: [10.1371/journal.pone.0016079](https://doi.org/10.1371/journal.pone.0016079) PMID: [21408199](https://pubmed.ncbi.nlm.nih.gov/21408199/)
88. Anisimova M, Bielawski JP, Yang Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. *Mol Biol Evol*. 2001; 18(8):1585–92. PMID: [11470850](https://pubmed.ncbi.nlm.nih.gov/11470850/)
89. Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*. 1997; 13(5):555–6. Epub 1997/11/21 PMID: [9367129](https://pubmed.ncbi.nlm.nih.gov/9367129/)
90. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007; 23(21):2947–8. Epub 2007/09/12. btm404 [pii] doi: [10.1093/bioinformatics/btm404](https://doi.org/10.1093/bioinformatics/btm404) PMID: [17846036](https://pubmed.ncbi.nlm.nih.gov/17846036/)
91. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 2006; 34(Web Server issue):W609–12. Epub 2006/07/18. 34/suppl\_2/W609 [pii] doi: [10.1093/nar/gkl315](https://doi.org/10.1093/nar/gkl315)
92. Yang Z, Nielsen R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol*. 2000; 17(1):32–43. Epub 2000/02/10 PMID: [10666704](https://pubmed.ncbi.nlm.nih.gov/10666704/)
93. Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol*. 1998; 15(5):568–73. Epub 1998/05/15 PMID: [9580986](https://pubmed.ncbi.nlm.nih.gov/9580986/)
94. Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 1986; 3(5):418–26. Epub 1986/09/01 PMID: [3444411](https://pubmed.ncbi.nlm.nih.gov/3444411/)
95. Zhang Z, Li J, Yu J. Computing Ka and Ks with a consideration of unequal transitional substitutions. *BMC Evol Biol*. 2006; 6:44. Epub 2006/06/03. 1471–2148–6–44 [pii] doi: [10.1186/1471-2148-6-44](https://doi.org/10.1186/1471-2148-6-44) PMID: [16740169](https://pubmed.ncbi.nlm.nih.gov/16740169/)
96. Wang DP, Wan HL, Zhang S, Yu J. Gamma-MYN: a new algorithm for estimating Ka and Ks with consideration of variable substitution rates. *Biol Direct*. 2009; 4:20. Epub 2009/06/18. 1745–6150–4–20 [pii] doi: [10.1186/1745-6150-4-20](https://doi.org/10.1186/1745-6150-4-20) PMID: [19531225](https://pubmed.ncbi.nlm.nih.gov/19531225/)
97. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007; 24(8):1586–91. Epub 2007/05/08. msm088 [pii] doi: [10.1093/molbev/msm088](https://doi.org/10.1093/molbev/msm088) PMID: [17483113](https://pubmed.ncbi.nlm.nih.gov/17483113/)

98. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*. 2008; 9:323. doi: [10.1186/1471-2105-9-323](https://doi.org/10.1186/1471-2105-9-323) PMID: [18662398](https://pubmed.ncbi.nlm.nih.gov/18662398/)
99. Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 2005; 21(18):3674–6. Epub 2005/08/06 bti610 [pii] doi: [10.1093/bioinformatics/bti610](https://doi.org/10.1093/bioinformatics/bti610) PMID: [16081474](https://pubmed.ncbi.nlm.nih.gov/16081474/)
100. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res*. 2008; 36(10):3420–35. Epub 2008/05/01. gkn176 [pii] doi: [10.1093/nar/gkn176](https://doi.org/10.1093/nar/gkn176) PMID: [18445632](https://pubmed.ncbi.nlm.nih.gov/18445632/)