

RESEARCH ARTICLE

Loci and natural alleles underlying robust roots and adaptive domestication of upland ecotype rice in aerobic conditions

Yan Zhao¹*, Hongliang Zhang¹*, Jianlong Xu^{2,3}*, Conghui Jiang¹, Zhigang Yin¹, Haiyan Xiong¹, Jianyin Xie¹, Xueqiang Wang¹, Xiaoyang Zhu¹, Yang Li¹, Weipeng Zhao¹, Muhammad Abdul Rehman Rashid^{1,4}, Jinjie Li¹, Wensheng Wang², Binying Fu², Guoyou Ye^{3,5}, Yan Guo⁶, Zhiqiang Hu², Zhikang Li^{2,3*}, Zichao Li^{1*}

1 Key Lab of Crop Heterosis and Utilization of Ministry of Education and Beijing Key Lab of Crop Genetic Improvement, China Agricultural University, Beijing, China, **2** Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, **3** Shenzhen Institute for Innovative Breeding, Chinese Academy of Agricultural Sciences, Shenzhen, China, **4** University of Agriculture Faisalabad, Sub-campus Burewala-Vehari, Pakistan, **5** International Rice Research Institute, Manila, Philippines, **6** State Key Laboratory of Plant Physiology and Biochemistry, College of Biological Sciences, China Agricultural University, Beijing, China

* These authors contributed equally to this work.

* zhkli1953@126.com (ZhL); lizichao@cau.edu.cn (ZL)



OPEN ACCESS

Citation: Zhao Y, Zhang H, Xu J, Jiang C, Yin Z, Xiong H, et al. (2018) Loci and natural alleles underlying robust roots and adaptive domestication of upland ecotype rice in aerobic conditions. *PLoS Genet* 14(8): e1007521. <https://doi.org/10.1371/journal.pgen.1007521>

Editor: Nathan M. Springer, University of Minnesota, UNITED STATES

Received: December 11, 2017

Accepted: June 27, 2018

Published: August 10, 2018

Copyright: © 2018 Zhao 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 raw RNAseq data are available from the NCBI SRA database (accession number SRP151096). All other relevant data are within the paper and its Supporting Information files.

Funding: The research was supported by grants from the Ministry of Science and Technology in China (2016YFD0100101-09, 2015DFG31900, 2015BAD02B01), the National Natural Science Foundation of China (31061140458), and Key Program of Hainan Department of Science and

Abstract

A robust (long and thick) root system is characteristic of upland *japonica* rice adapted to drought conditions. Using deep sequencing and large scale phenotyping data of 795 rice accessions and an integrated strategy combining results from high resolution mapping by GWAS and linkage mapping, comprehensive analyses of genomic, transcriptomic and haplotype data, we identified large numbers of QTLs affecting rice root length and thickness (RL and RT) and shortlisted relatively few candidate genes for many of the identified small-effect QTLs. Forty four and 97 QTL candidate genes for RL and RT were identified, and five of the RL QTL candidates were validated by T-DNA insertional mutation; all have diverse functions and are involved in root development. This work demonstrated a powerful strategy for highly efficient cloning of moderate- and small-effect QTLs that is difficult using the classical map-based cloning approach. Population analyses of the 795 accessions, 202 additional upland landraces, and 446 wild rice accessions based on random SNPs and SNPs within robust loci suggested that there could be much less diversity in robust-root candidate genes among upland *japonica* accessions than in other ecotypes. Further analysis of nucleotide diversity and allele frequency in the robust loci among different ecotypes and wild rice accessions showed that almost all alleles could be detected in wild rice, and pyramiding of robust-root alleles could be an important genetic characteristic of upland *japonica*. Given that geographical distribution of upland landraces, we suggest that during domestication of upland *japonica*, the strongest pyramiding of robust-root alleles makes it a unique ecotype adapted to aerobic conditions.

Technology (ZDYF2016217). 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.

Author summary

Asian cultivated rice is well-known for its rich-within-species diversity with two major subspecies, *indica* and *japonica* and subpopulation differentiation. A robust (long and thick) root system that is characteristic of upland *japonica* rice represents a predominant ecotype grown under aerobic and rain-fed conditions. In this study, we identified candidate genes for root length and root thickness, and validated five root length candidates by T-DNA insertional mutations. Further analyses of an Asian cultivated and wild rice population were performed based on random SNPs and SNPs within robust loci. The findings hold promise for application in improving drought resistance and also reveal the adaptive domestication history of upland rice as a unique Asian cultivated rice ecotype.

Introduction

Asian cultivated rice (*Oryza sativa* L., *O. sativa*) is a staple food for half the world's population. Grown in diverse environments worldwide, *O. sativa* is also well-known for its rich-within-species diversity with two major subspecies, *indica* (also referred to as *Xian*) and *japonica* (also referred to as *Geng*) and subpopulation differentiation [1–6]. Upland rice, particularly upland *japonica* rice, represents a predominant ecotype grown under aerobic and rain-fed conditions in mountainous areas of Southwest China, South and Southeast Asia, Africa and Latin America [7–11]. Most upland rice accessions have a robust root (long and thick) system [10,12,13] that confers the ability of upland to adapt to aerobic and drought-prone conditions of rain-fed, hilly environments by absorbing more water and nutrients from deeper soil zones [14–18], implying that robust roots could be a key physiological feature selected during adaptation to aerobic conditions of upland rice. Thus, exploration of loci and natural alleles underlying robust roots and adaptation analysis of upland rice using these loci are conducive to developing new rice varieties with improved drought resistance and water use efficiency, when plant breeders are facing an increasing challenge of water shortage caused by global climate change.

Breeding rice lines with robust roots based on phenotypic selection has been challenging because conventional breeding approaches based on phenotypic selection are ineffective for improving root traits [19–21]. Thus, tremendous efforts have been made in identifying QTLs/genes contributing to robust roots to facilitate marker- or genome-assisted breeding of drought resistance rice varieties by improving root traits. To date, many QTLs for root length (or depth) and thickness in rice have been identified in bi-parental crosses, such as *DRO1*, *DRO2*, *DRO3*, *SOR1*, *qSOR1*, *qRL6.1*, *qRL7* and *qRT9* [22–30]. *OsbHLH120* that controls root thickness was cloned by map-based cloning using introgression lines [23,31]. Additionally, *DRO1* affects deep rooting by controlling root angle and the mutation of *SOR1* leads to observed deficiency in root gravitropic response and soil-surface rooting [30,32]. Traditional methods using linkage mapping and map-based cloning for QTL identification and cloning are laborious and inefficient. Thus, many questions remain to be answered in understanding allelic diversity at QTL controlling robust roots in the primary gene pool of rice and their roles in the adaptation to aerobic conditions of upland rice.

Genome-wide association and transcriptome analyses have become feasible with recent advances in high-throughput sequencing technologies, which enable capture of genomic variation affecting complex traits. These include genome-wide association study (GWAS) [33–36], genome-wide screening of elite single-nucleotide polymorphism (SNP) alleles [12], determination of genome-wide expression profiles under different drought stress conditions [37–39], MutMap based on whole-genome sequencing of bulked DNA of F₂ segregants [40], Ho-

LAMap that joins GWAS and multiple bi-parental linkage analysis [41]. Using these methods, wide natural variation was explored at loci associated with root architecture, such as *Nall*, *OsJAZ1* and *Nced* [12,36]. Although each of these approaches has advantages and disadvantages in revealing specific aspects of the genetic and/or molecular mechanisms underlying complex traits, when taken together, they provide a powerful strategy for genetic and molecular dissection of complex traits.

Based on hydrological conditions ranging from fully aerobic, temporarily and fully anaerobic, four major rice ecosystems were categorized by International Rice Research Institute (IRRI), including upland, rainfed lowland, irrigated (also called lowland or paddy) and flood-prone (also called deepwater) [7,33,42,43]. From the 1970s many morphologic and genetic studies distinguished upland and lowland rice genotypes, and investigated the domestication of upland rice [4,13,18,44]. Among morphological traits for drought resistance, such as root characteristics [45], leaf rolling [46] and ratio of shoot and root [47], root length and root thickness have been positively related to field drought resistance [48,49], implying that robust roots can be used to distinguish upland and lowland rice. Further genetic differentiation between upland and lowland rice was explored by simple sequence repeats (SSRs) located in expressed sequence tags (ESTs) isolated from rice in response to drought treatment [44,50]. These previous studies suggested that there was considerable morphologic and genetic differentiation between upland and lowland rice. Recently, a phylogenetic analysis of 3,029,822 SNPs detected from 166 rice accessions (82 lowland accessions, 84 upland accessions and 25 wild rice accessions) showed that all upland *japonica* accessions clustered together, suggesting that upland *japonica* has a single origin [18]. Despite these advances, deep sequencing of a wider collection of *O. sativa* was needed to investigate the adaptation to aerobic conditions of upland rice. Moreover, given differences in phylogenetic trees using whole-genome sequencing data and domestication loci [2], genetic analysis using loci associated with robust roots may be a better way to gain insights into the adaptive domestication process of upland rice.

In this study, we demonstrate an integrated strategy that combined high resolution mapping by GWAS and linkage mapping, comprehensive bioinformatics analyses of genomic, transcriptomic and haplotype data from deep genomic and transcriptomic sequencing, and large scale phenotyping of two large sets of *O. sativa* accessions for genetic dissection and high efficiency cloning of small-effect QTLs underlying robust roots in rice. Our results revealed insights into the adaptive domestication history of upland rice as a unique *O. sativa* ecotype that is adapted to the aerobic conditions of highlands in tropical and subtropical environments.

Results

Population structure and phenotypic characterization of root traits of 795 *O. sativa* accessions

The 795 *O. sativa* accessions used were from 41 countries worldwide, including a 189 accession mini-core collection retaining 70.65% of the genotypic variation and 76.97% of the phenotypic variation of 4,310 Chinese *O. sativa* accessions [51], and 525 lines in the International Rice Molecular Breeding Network [52], including 40 upland accessions (S1 Table). By comparison with the RGAP 7 Nipponbare reference genome, we obtained 15,133,187 SNPs from the 795 genomes with an average sequencing depth of 15×, including 12,648,777 SNPs with missing rates ≤ 50% (S1A Fig). After removing SNPs with missing rates > 50% and minor allele frequencies < 5%, we identified 3,291,151 (3.3 million) SNPs representing high density. Using 154,516 SNPs with missing rates ≤ 50%, minor allele frequencies ≥ 5% and r^2 of linkage disequilibrium (LD) ≤ 0.3, we calculated varying levels of K means, the *Indica* and *Japonica*

varietal groups appeared clearly at $K = 2$ (S1B and S1C Fig), which was supported by principal component (PC) plotting, and a kinship matrix constructed using 3.3 million SNPs and neighbor-joining tree constructed using 54,853 evenly distributed SNPs (Fig 1A and 1B, S1D Fig). Referring to the recent research results of 3,010 rice accessions [1], we classified the 795 *O. sativa* accessions into two major subspecies, *indica* (506 accessions) and *japonica* (289 accessions) for further phenotypic analysis and GWAS in two subpopulations (Fig 1A and 1B, S1 Table).

The 795 *O. sativa* accessions showed significant amount of variations for the measured root traits, ranging from 8.1 to 19.4 cm for root length (RL) and from 0.46 to 0.95 mm for root thickness (RT) (S2 Fig, S2 Table). Upland *japonica* accessions had mean RL and RT of 14.5 ± 1.6 cm and 0.79 ± 0.07 mm, significantly higher than those of the other three ecotypes (upland *indica*, lowland *indica* and lowland *japonica*) (Fig 1C and 1D). No obvious differences in mean root weight were detected between the upland *japonica* and other ecotypes, although there was considerable variation in root weight among different accessions within each subpopulation (S3 Fig, S2 Table). Clearly, upland *japonica* accessions as a group showed extreme phenotypes for robust root systems, and our population showed sufficient variation in RL and RT to uncover the genetic architecture of robust roots.

Identification of robust root QTLs by GWAS

Using the root trait data and 3.3 million SNPs, we performed GWAS to identify important QTLs for RL and RT. To minimize false positives due to population structure [35,53,54], we compared the general linear model (GLM) (S4 and S5 Figs) and compressed mixed linear model (CMLM) (Fig 2) and determined that the CMLM model effectively minimized false positive rates in detecting RL and RT QTLs in our populations. A threshold of $-\log(P) = 4$ was determined using a conditional permutation test (S6 Fig). We defined a QTL as including ≥ 3 clustered significant SNPs within distances ≤ 170 kb between adjacent ones, given LD decay values of ~ 123 kb and ~ 167 kb in *indica* and *japonica* populations, respectively [53]. With this criterion we identified 21, 14 and 4 RL QTLs in the whole, *indica* and *japonica* populations, respectively (Fig 2). Detailed comparisons indicated that 1 of 4 QTLs in *japonica* and 7 of 14 QTLs in *indica* mapped to identical positions as RL QTLs identified in the whole population, while most other QTLs detected in both subpopulations were near those identified in the whole population, indicating that the slight differences in position of detected RL QTLs resulted from LD and population structure (Fig 2, S3 Table). We also identified 22, 28 and 46 RT QTLs in the whole, *indica* and *japonica* populations (Fig 2, S4 Table). Given that most of RL QTLs in both subpopulations could be identified in the whole population, the 21 RL QTLs in the whole population and 96 RT QTLs in all three populations were considered target QTLs for further exploration of functional genes.

Determination of candidate genes within the RL and RT QTLs

We took a four-step integrated strategy to determine the high-confidence candidate genes for the detected RL and RT QTLs (see Materials and Methods for detail).

First, we constructed two varietal pools for RL and RT from each of the *japonica* and *indica* subpopulations with each pool containing 20 accessions with extreme phenotypes, i.e. a long-root pool and a short-root pool for RL, or a thick-root pool and a thin-root pool for RT (S1 Table). We then performed pooled- χ^2 -tests to compare the allele frequencies of all significant SNPs at the target QTLs between the two pools for each root trait because varieties with extreme phenotypes were expected to possess more QTL alleles for increased trait value, whereas those of extremely low phenotypic values would have fewer. Within the 21 RL QTLs,

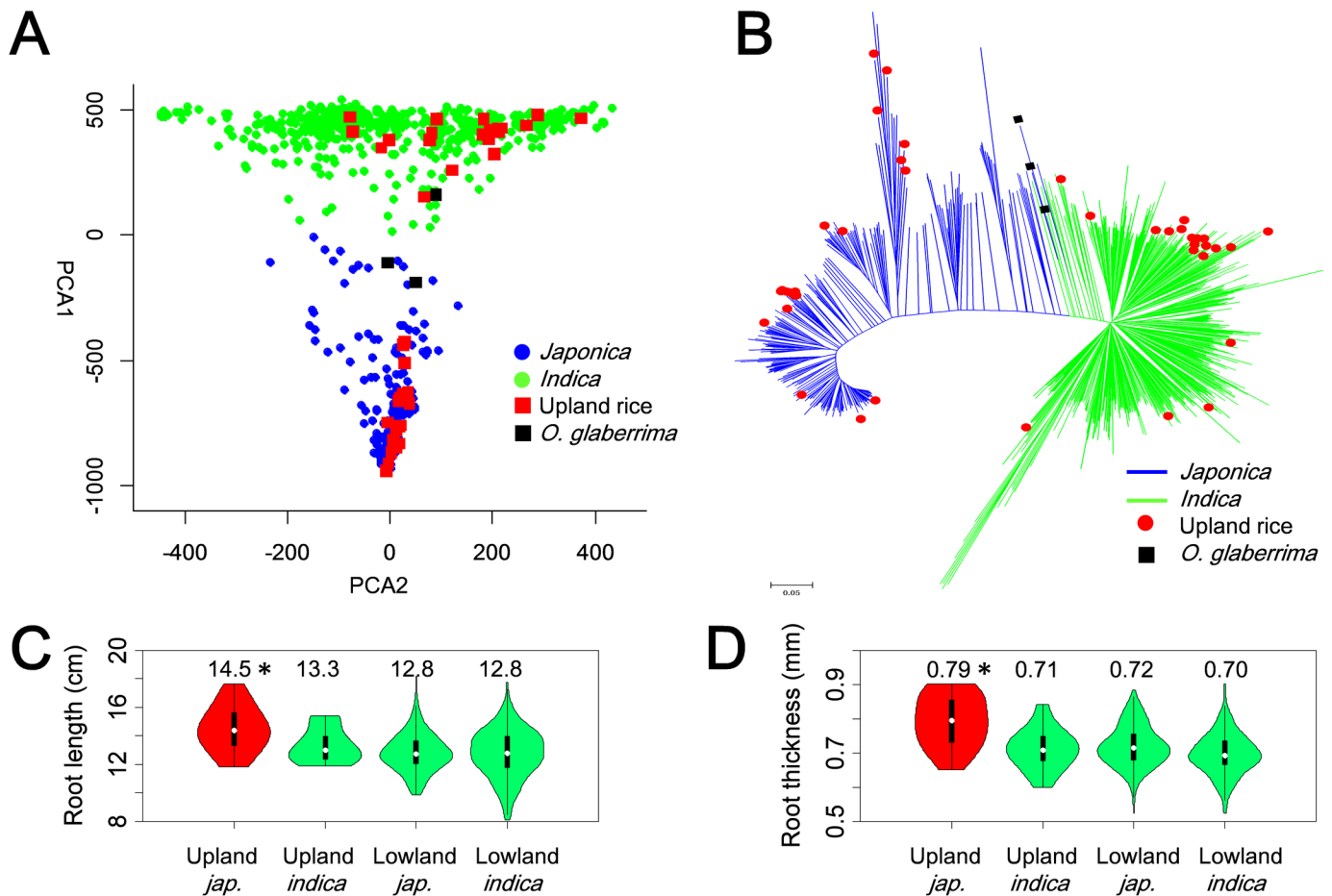


Fig 1. Population structure of 795 *O. sativa* (rice) accessions, phenotypic diversity in root traits. (A) Principal component (PC) plots of different ecotypes and subpopulations; and (B) neighbor-joining tree of all accessions and distribution of upland rice accessions in the N-J tree. PC analysis was performed using 3.3 million SNPs with missing data rates $\leq 50\%$ and minor allele frequency $\geq 5\%$. N-J tree was constructed using 54,853 SNPs evenly distributed throughout the genome. Distribution of (C) root length and (D) root thickness among different rice ecotypes. Y-axes show the median (white point), 95% of confidence intervals (black bars) and range (colored shapes) of the trait phenotypes. X-axes indicate different ecotypes, ordered by upland japonica, upland indica, lowland japonica and lowland indica. Numbers above violins are mean phenotypic values of different populations. Ecotypes colored red with * were significantly higher than those colored green ($P < 0.05$ detected by one-way ANOVA).

<https://doi.org/10.1371/journal.pgen.1007521.g001>

there were 254 annotated genes containing 857 significant SNPs (S3 Table). Among them, there were 174 genes with 282 large-effect SNPs in the open reading frames (ORFs) and 141 genes with 271 significant SNPs in their 2000 bp upstream promoter regions. The pooled- χ^2 -tests based on allelic frequency differences for SNPs around each detected QTL between the long-root and short-root pools allowed us to reduce the number of significant SNPs (candidate genes) from 857 to 230 (from 254 to 85) for 18 RL QTLs with QTL alleles for increased root length were much more abundant in the long-root varieties than in the short-root varieties at all 230 SNPs (Fig 3A, S5 and S6 Tables).

We then compared the QTLs identified by GWAS with QTLs detected previously by linkage mapping in two bi-parental crosses and found 10 [55–60] of the 21 RL QTLs and 23 [22,23,25,59,61] of the 96 RT QTLs were common (S7 Table). We screened all 380 SNPs significantly associated with RL in 10 common QTL regions containing 102 candidate genes and found that 119 of these SNPs located in 35 genes in 7 QTLs showed differences between the parents of the bi-parental populations (S8 Table). Using the results of pooled- χ^2 -tests, we

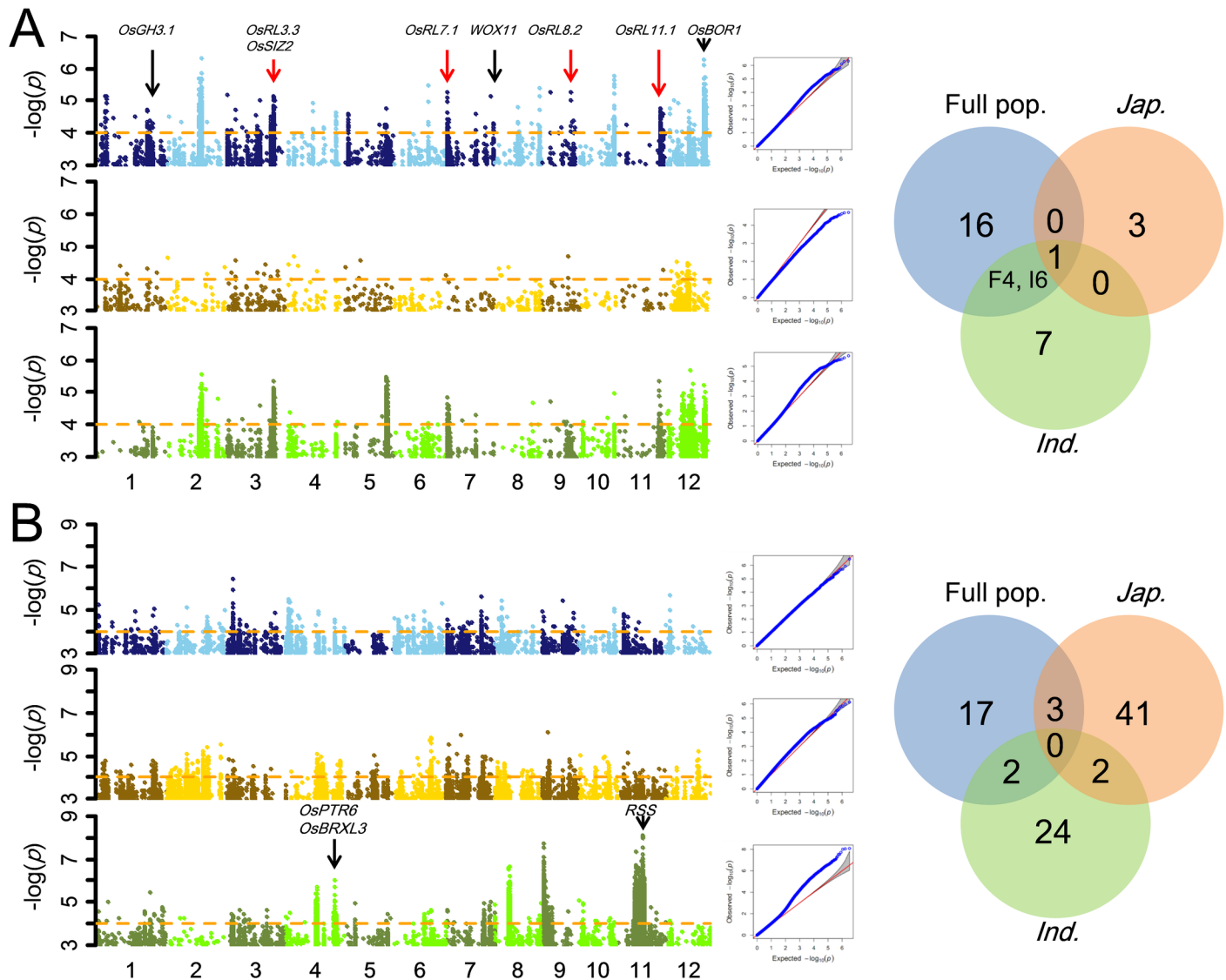


Fig 2. Identification of robust-root QTLs by GWAS. Manhattan plots and quantile-quantile (Q-Q) plots in the whole population (upper left, blue), *japonica* (middle left, yellow) and *indica* (lower left, green) and Venn diagram of QTL number detected in different populations (right) for (A) root length and (B) root thickness. Black and red arrows mark six known genes and five QTL genes identified/confirmed in this study in their corresponding QTLs, respectively. F4 and I6 in the Venn diagram showed that 4 QTLs in whole population overlapped with 6 QTLs in *indica*.

<https://doi.org/10.1371/journal.pgen.1007521.g002>

shortlisted the candidate genes from 254 to 81 within 17 RL QTLs (Fig 3B and 3C, S7 Fig). There were 146 significant SNPs in the 81 genes, including 82 SNPs in ORFs of 42 candidate genes, while 77 SNPs in the promoters of 46 candidate genes (S9 Table). These candidate genes included *OsGH3.1*, *WOX11* and *OsBOR1* with known functions in root development. Similar analyses were performed on the 96 RT QTLs and resulted in 466 genes as the most likely candidates for the 29 RT QTLs (S10 Table).

In steps 3 and 4, to further reduce the number of candidate genes for the identified root trait QTLs, we performed transcriptomic analyses on 6 robust-root *japonica* upland varieties and 6 non-robust-root lowland *japonica* lines. We expected that most real QTL genes in these two pools of varieties might indicate differences in three patterns: root-specific expression, differential expression between two pools, and no differential expression but having significant

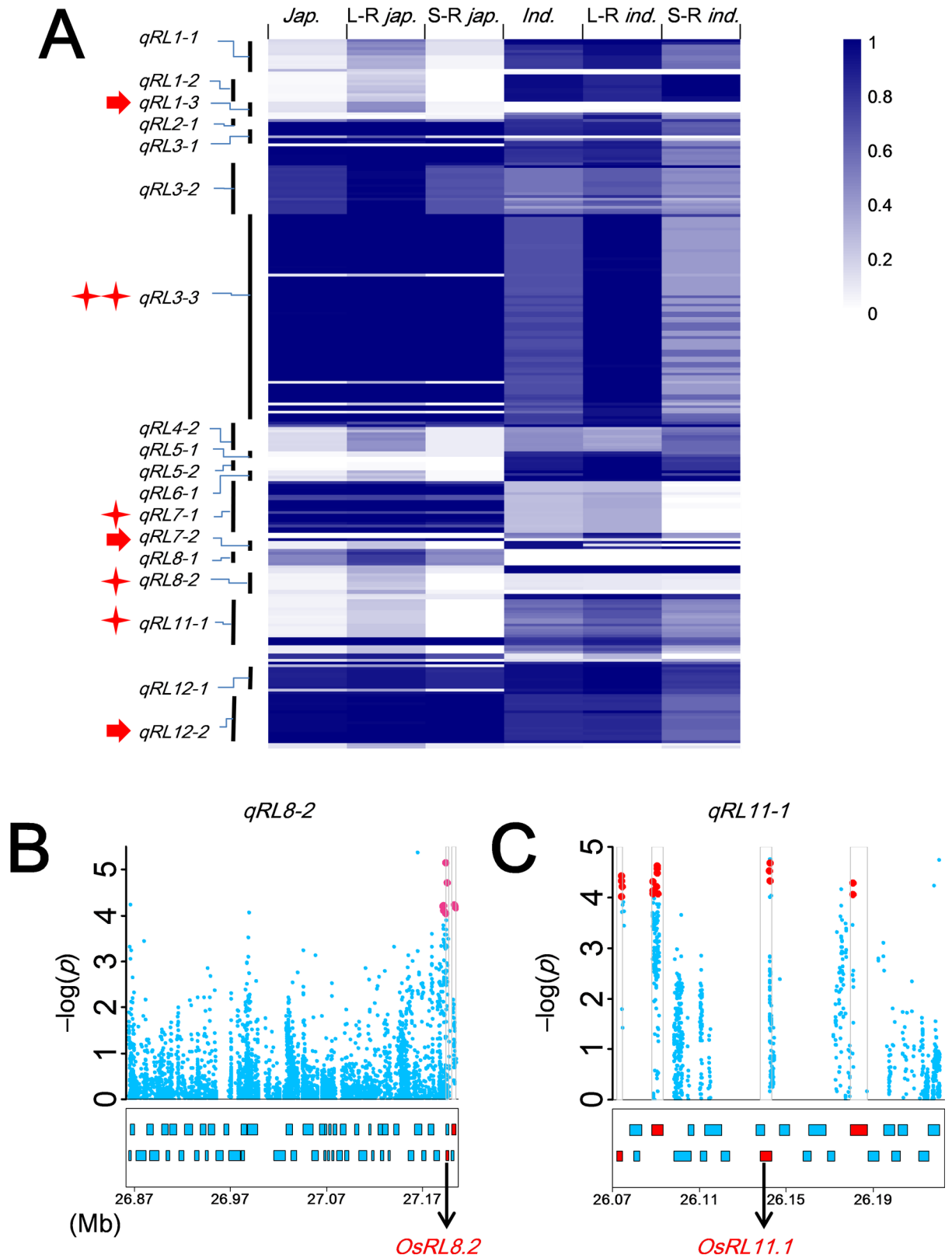


Fig 3. Determination of candidate genes for root-length QTLs by integrated analyses of GWAS, linkage mapping and pooled- χ^2 -tests. (A) Heat map showing significant differences in long-root allele frequencies at 230 SNPs in 85 genes in 18 GWAS detected root-length QTL

regions of the two extreme phenotypic pools (long-root vs short-root varietal pools) within rice populations (**S1 Table**), in which different colors show different long-root allele frequencies; rows of the heat map are SNP sites and QTLs; columns are long-root allele frequencies in different populations (*japonica*, long-root *japonica*, short-root *japonica*, *indica*, long-root *indica* and short-root *indica*); red arrows and stars show QTLs with known genes and candidate genes identified and confirmed for root length in this study. Identification of candidate genes in regions of (B) *qRL8-2* and (C) *qRL11-1*. The top of each panel shows the entire QTL region identified by GWAS using the whole population and CMLM, and X-axis indicates the physical position (Mb) of the QTL regions in the rice genome. Negative log₁₀-transformed *P* values are plotted on the vertical axis, in which dots show positions and $-\log(P)$ values of all significant SNPs in the QTL region and the red points indicate SNPs with significant differences in allele frequency of GWAS detected QTLs between the root-length pools and differential alleles between parents in previously reported bi-parental mapping populations; pink points represent SNPs with significant differences in allele frequency between pools, when the QTL was mapped only in our GWAS; and skyblue points show others. At the bottom of each panel, annotated genes are indicated by skyblue boxes; candidate genes are colored in red. Functional genes verified are red-labeled.

<https://doi.org/10.1371/journal.pgen.1007521.g003>

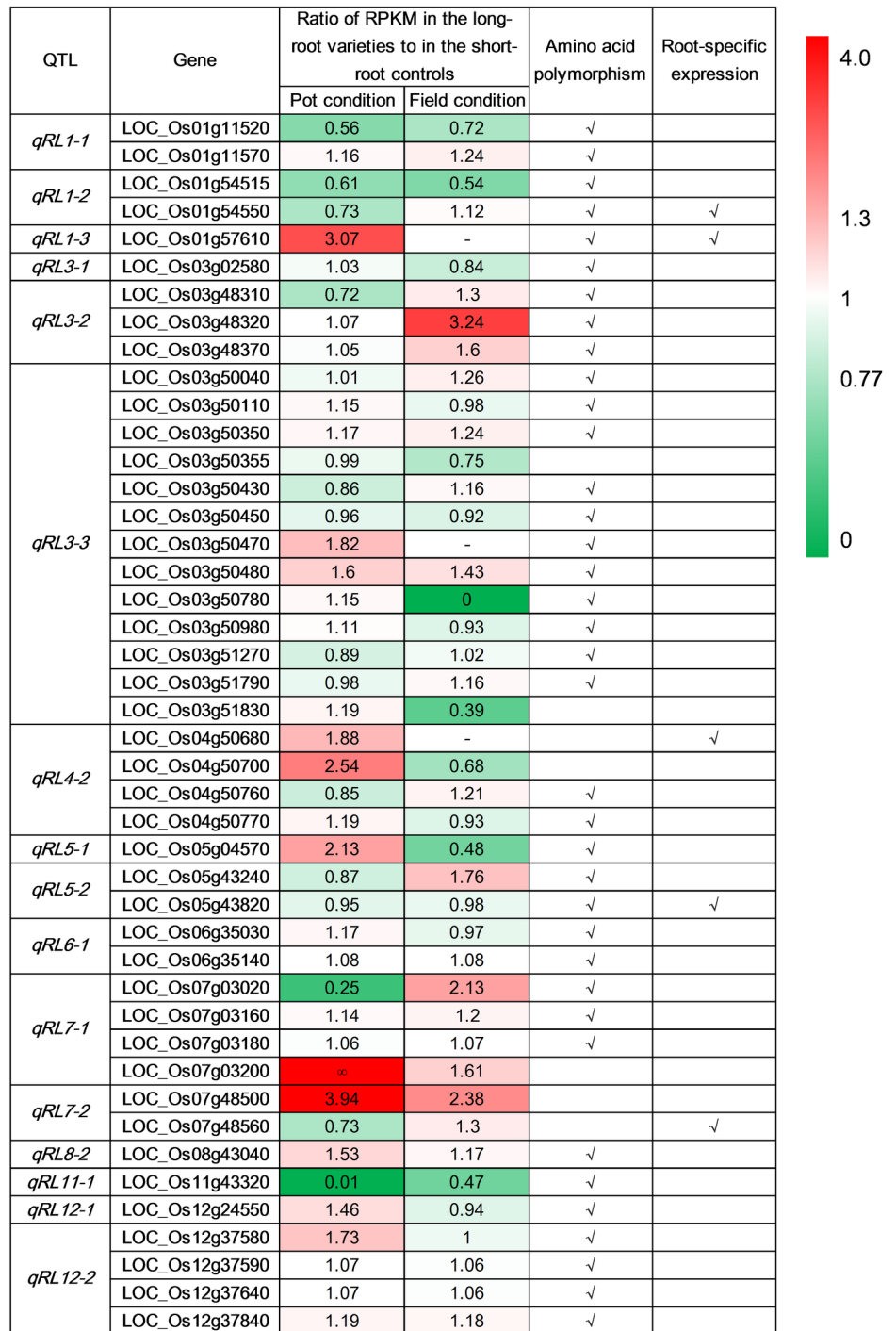
phenotypic differences between alleles at non-synonymous SNP. Based on thresholds of (RPKM, Reads Per Kilobase per Million mapped reads) $\text{RPKM}_{\text{Robust-roots}}/\text{RPKM}_{\text{Controls}} > 1.3$ or < 0.77 between the robust-root varieties and non-robust controls in our transcriptomic data, we identified 44 candidate genes for 16 RL QTLs, including 23 genes showing root-specific differential expression between the long-root and short-root pools, 37 genes containing many large-effect non-synonymous SNPs or InDels between the long-root and short-root groups, and 5 specifically expressed root genes (**Fig 4A, S11–S13 Tables**). Applying the same strategy, 97 candidate genes were shortlisted for 20 of the RT QTLs (**S14 and S15 Tables**). Of all these candidate genes, six (*OsGH3.1* for *qRL1-3*, *WOX11* for *qRL7-2*, *OsBOR1* for *qRL12-2*, *OsPTR6* for *qRT4-4*, *OsBRXL3* for *qRT4-4*, and *RSS3* for *qRT11-4*) appeared to be the most likely QTL candidate genes based on the combined genome mapping and transcriptomic evidence (**Figs 2 and 4B**). *OsGH3.1* (LOC_Os01g57610) is an indole-3-acetic acid (IAA) amido synthetase, a member of the GH3 family in which several genes are known to control root architecture and drought resistance [62–64]. *WOX11* (LOC_Os07g48560) was reported to be a key regulator specifically expressed in rice crown roots [65,66]. *OsBOR1* (LOC_Os12g37840) was reported to regulate root development in rice grown under limited boron availability [67]. Of the three genes for RT, over-expression of *OsPTR6* (LOC_Os04g50950) increases root growth in rice [68], whereas *OsBRXL3* (LOC_Os04g51172) is a member of the BREVIS RADIX family, which is known to influence cell proliferation and root elongation in Arabidopsis and rice [69]. *RSS3* (LOC_Os11g25920) regulates root cell elongation under salinity stress [70].

Identification of important QTL alleles and validation of the QTL candidate genes

Fig 5, S8 and S9 Figs show results from haplotype analyses using all non-synonymous SNPs and InDels within the ORFs and promoters of all candidate genes for the identified RL and RT QTLs. Significant phenotypic differences were detected between or among different haplotypes at 9 loci for 7 RL QTLs and 14 candidate genes for 7 RT QTLs with 2–5 major haplotypes at each of them. Gene ontology analysis indicated that 16 of the 23 genes have diverse biological functions and the remaining 7 genes have no known functions (**Table 1**).

For RL candidate genes, we detected 3 haplotypes at LOC_Os03g50980 (*OsSIZ2*) with Hap1 present only in some *indica* accessions. Hap2 was the predominant one and was associated with longer RL in both *indica* and *japonica* populations, while Hap3 had a low frequency and was associated with shorter roots in both populations (**Fig 5D**). Similarly, three major haplotypes were detected at LOC_Os11g43320 (*OsRL11.1*) with Hap1 associated with longer roots. Its frequency was high in *indica* and low in *japonica* (**Fig 5E**). The remaining 7 RL loci showed greater *indica-japonica* differentiation, with 1–3 of the haplotypes absent in either *indica* or *japonica* accessions (**Fig 5A–5C, S8 Fig**). There were *japonica*-specific allele variations for RT

A



B

	GWAS+PT	GWAS+LM	GWAS+PT+LM	GWAS+PT+LM+TN
Number of QTL	21-18	10-7	21-17	21-16
Number of Sig. SNP	857-230	380-119	857-146	-
Number of genes	254-85	102-35	254-81	254-44

Fig 4. Determination of candidate genes for root-length QTLs by integrated genomic and transcriptomic analyses. (A) Heat map of the RPKM ratio in long-root varieties to short-root controls. Different colors show ratios and blanks

indicate no expression. Rows of the heat map correspond to the 44 candidate genes for root length QTLs listed on the left of the table. The columns indicate the phenotypic differences between the long-root and short-root pools in the pot and field conditions. \checkmark or \times in column 'Amino acid polymorphism' indicates the presence or absence of large-effect SNPs/InDels and the significant phenotypic differences between the varietal pools of extreme phenotypes, parents or populations; column 'Root-specific' shows the root-specific expression of the listed genes. Detailed expression and large-effect SNPs/InDels with amino acid polymorphisms are listed in **S11–S13 Tables**. (B) Screening effectiveness of each step in the integrated strategy. 'PT', 'LM' and 'DE' indicate 'pooled- χ^2 -tests', 'linkage mapping' and 'transcriptomic and non-synonymous SNPs analysis'. Numbers in the block represent the number of QTLs, significant SNPs and candidate genes before and after screening.

<https://doi.org/10.1371/journal.pgen.1007521.g004>

on LOC_Os01g57610 (*OsGH3.1*) and LOC_Os08g43040 (*OsRL8.2*) since their haplotype differences were present only in *japonica* accessions, whereas LOC_Os03g50470 (*OsRL3.2*) showed *indica*-specific allele variations for RT.

We detected 2 to 5 haplotypes and observed a similar *indica-japonica* differentiation at the 14 candidate loci for RT QTLs (**S9 Fig**). Of these, LOC_Os02g13400 (*OsRT2.1*), LOC_Os04g02780 (*OsRT4.1*), LOC_Os04g31000 (*OsRT4.4*), and LOC_Os06g37840 (*OsRT6.2*) showed *japonica*-specific allele variations for RT since significant phenotypic differences were present only between or among *japonica* haplotypes. Similarly, LOC_Os04g30920 (*OsRT4.3*), LOC_Os04g31010 (*OsRT4.5*), LOC_Os04g31020 (*OsRT4.6*), LOC_Os04g31120 (*OsRT4.7*), LOC_Os04g51060 (*OsRT4.8*), LOC_Os04g51080 (*OsRT4.9*), and LOC_Os08g15050 (*OsRT8.2*) possessed *indica*-specific allele variation. Only three loci, LOC_Os04g30490 (*OsRT4.2*), LOC_Os06g36810 (*OsRT6.2*), LOC_Os08g14660 (*OsRT8.1*) showed significant phenotypic differences between haplotypes in both *indica* and *japonica* populations.

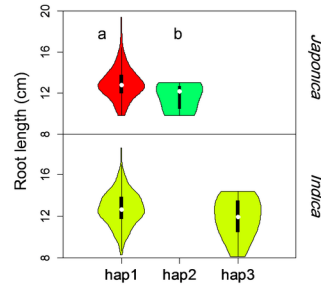
To validate the candidate genes related to robust roots, five mutant lines in Dongjin or Hwayoung genetic backgrounds (*Ti-OsRL3.3*, *Ti-OsSIZ2*, *Ti-OsRL7.1*, *Ti-OsRL8.2* and *Ti-OsRL11.1*) each containing a T-DNA insertion in one of the five RL candidate genes (LOC_Os03g50350, LOC_Os03g50980, LOC_Os07g03160, LOC_Os08g43040 and LOC_Os11g43320) (**S10 Fig, S16 Table**) for the 4 RL QTLs were evaluated on MS medium for RL, root number and shoot length, together with the Dongjin or Hwayoung wild types and controls (mutant lines in *HsfA4a*, LOC_Os01g54550). The homozygous and heterozygous mutant plants at all loci, except for *Ti-OsSIZ2* showed highly significant reductions in root and shoot growth compared with the wild types and controls (**Fig 6A and 6D**). *Ti-OsRL3.3* mutant (LOC_Os03g50350) plants in particular had no primary root development 5 days after germination, even though the crown roots initiated in the same way as the wild type (**Fig 6B and 6C, S11 Fig**). Curiously, mutant *Ti-OsSIZ2* (LOC_Os03g50980) showed a slight but non-significant reduction in RL, but a highly significant shoot length reduction when compared with the wild type (**Fig 6D**). There was a highly significant difference between *Ti-OsRL3.3* and the wild type in root number (**Fig 6D**). *OsSIZ2* was previously shown to complement a dwarf phenotype in Arabidopsis [71], and its homolog (*OsSIZ1*) affected primary root length, crown root number and plant height by regulating auxin accumulation [72,73]. Thus, the detected RL QTL, *qRL3-3*, could be due to either of the two non-synonymous SNPs in LOC_Os03g50350, whereas LOC_Os03g50980 was located in a different LD block and showed no interaction with LOC_Os03g50350 (**S12 Fig**).

Pyramiding of robust-root alleles within upland *japonica* in adapting to aerobic conditions

Initially, we constructed a phylogenetic tree of 997 *O. sativa* accessions (with 242 upland rice landraces from 21 countries) and 446 wild rice (*O. rufipogon*) accessions [2] based on the evenly distributed 90,838 SNPs (**S13 Fig**). The 242 upland rice landrace accessions were

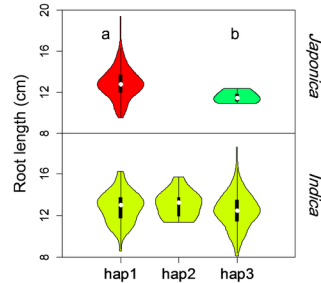
A *LOC_Os03g50350 (qRL3-3)*

Hap.	CDS		<i>Jap.</i>		<i>Ind.</i>	
	28731233	28732204	Num.	Root length (cm)	Num.	Root length (cm)
hap1	C	A	253	12.95	340	12.74
hap2	C	G	5	11.64	0	-
hap3	G	G	0	-	14	11.87



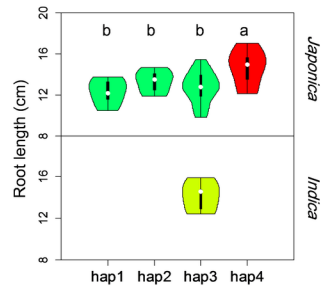
B *LOC_Os07g03160 (qRL7-1)*

Hap.	CDS		<i>Jap.</i>		<i>Ind.</i>	
	1229401	1233481	Num.	Root length (cm)	Num.	Root length (cm)
hap1	A	C	225	12.88	76	12.82
hap2	T	C	0	-	13	13.1
hap3	T	G	4	11.54	298	12.48



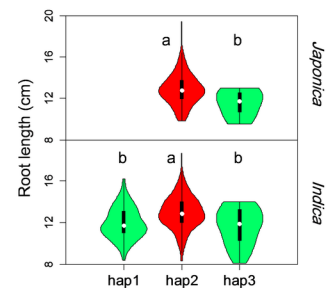
C *LOC_Os08g43040 (qRL8-2)*

Hap.	Promoter		<i>Jap.</i>		<i>Ind.</i>	
	27201417	27201966	Num.	Root length (cm)	Num.	Root length (cm)
hap1	CATAT	C	7	12.3	0	-
hap2	CATAT	CAG	7	13.31	0	-
hap3	C	CAG	32	12.74	6	13.85
hap4	CAT	CAG	6	14.99	0	-



D *LOC_Os03g50980 (qRL3-3)*

Hap.	CDS						<i>Jap.</i>		<i>Ind.</i>		
	29129116	29130305	29132728	29133285	29133685	29133749	29133985	Num.	Root length (cm)	Num.	Root length (cm)
hap1	G	C	C	G	A	A	A	0	-	92	11.95
hap2	T	C	C	A	G	A	A	236	12.88	259	12.97
hap3	T	T	T	A	G	G	C	6	11.54	32	11.61



E *LOC_Os11g43320 (qRL11-1)*

Hap.	CDS						<i>Jap.</i>		<i>Ind.</i>		
	26142352	26142458	26142464	26142674	26142706	26142815	26142889	Num.	Root length (cm)	Num.	Root length (cm)
hap1	A	A	T	A	G	C	T	10	13.22	140	12.89
hap2	A	G	T	G	G	T	C	201	12.81	49	12.3
hap3	G	G	G	G	A	T	C	4	12.61	60	11.98

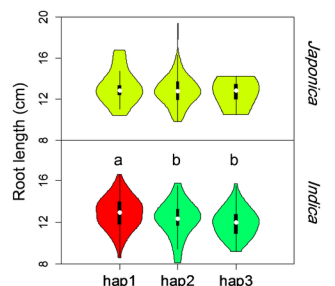


Fig 5. Haplotype analyses of five QTL genes for root length. Gene structures (left) and root lengths of different haplotypes (right) of (A) *OsRL3.3* (LOC_Os03g50350) for *qRL3-3*, (B) *OsSIZ2* (LOC_Os03g50980) for *qRL3-3*, (C) *OsRL7.1* (LOC_Os07g03160) for *qRL7-1*, (D) *OsRL11.1* (LOC_Os11g43320) for *qRL11-1*, (E) *OsRL8.2* (LOC_Os08g43040) for *qRL8-2*, where red colored numbers show the key positions where amino acid substitutions occurred among major haplotypes, that resulted in significant differences in root length; among violin maps of different haplotypes in subpopulations *japonica* (upper right) and *indica* (lower right), different letters above the violins indicate significant differences ($P < 0.05$) when analyzed by Duncan's test or independent-sample t-tests.

<https://doi.org/10.1371/journal.pgen.1007521.g005>

divided into 90 tropical *japonicas*, 41 temperate *japonicas*, 3 *japonicas* (*japonica* but not clearly into tropical or temperate ecotypes), 73 *indicas* and 35 intermediate-types (cultivated rice but not clearly into *japonica* or *indica*). We then constructed a phylogenetic tree of 997 *O. sativa* and 446 wild rice accessions (Fig 7A, S14 Fig) using 5,779 SNPs in the robust-root candidate genes. The upland rice landraces in *japonica* and *indica* were clearly separated at two nodes after differentiation of *japonica* and *indica*. By further comparing their phenotypic and genotypic differences, we found most upland landraces with robust roots and very high frequencies of long-root and thick-root alleles were clustered in two sub-branches within subspecies *japonica* and *indica*, which were designated L-T-root ecotypes (Fig 7A and 7B). *Japonica* L-T-root ecotypes included 141 upland accessions (89 tropical *japonica*, 36 temperate *japonica*, 1 *japonica* and 15 intermediate), whereas the *indica* L-T-root ecotypes included 59 upland accessions (51 *indica* and 8 intermediate) (S14 Fig, S1 Table). Close examination of the geographical distribution of the upland landraces indicated that the majority of *japonica* L-T-root accessions were from hilly areas with rain-fed farming practices [7,74–78] in South China, Southeast Asia and Africa, as were most *indica* L-T-root landraces from the rain-fed areas of South China, Southeast Asia and India (S1 Table). Clearly, those upland landraces became adapted to the rain-fed upland aerobic conditions of the respective hilly areas. These results suggested that the lower diversity in robust-root candidate genes was a distinctive characteristic of upland rice in adapting to aerobic conditions (Fig 7A and 7B).

We then examined the ancestral states of 59 and 136 non-synonymous SNPs associated with RL and RT; 53 of 59 RL alleles and 108 of 136 RT alleles were detected in the 446 wild accessions. In particular, the frequencies of the long-root and thick-root alleles were lowest in the three wild rice populations, Or-I (49% and 43%, respectively), Or-II (53% and 40%) and Or-III (55% and 38%) [2], followed by *indica* (57% and 54%), lowland temperate *japonica* (79% and 45%), lowland tropical *japonica* (77% and 58%), and upland temperate *japonica* (85% and 57%), whereas upland tropical *japonica* had the highest long-root and thick-root allele frequencies at 86% and 64% (Fig 7C). The results suggested that almost all alleles were present in wild rice, and there was obvious pyramiding of robust-root alleles in cultivated rice, especially in upland *japonica*.

Given that a reduction in nucleotide diversity was an important selective signature [2], we computed the ratios of genetic diversity between different branches ($\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}}$) in *japonica* and *indica*. Twenty-six of the 44 RL candidates and 56 of the 97 RT candidates were under selection with greatly reduced diversity in *japonica* L-T-root accessions ($\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}} > 4$), whereas only 3 of the 44 RL candidates and 17 of the 97 RT candidates were under selection in *indica* L-T-root accessions (Fig 7D, S17 Table). When comparing the genetic diversity in the wild rice accessions with the *japonica* and *indica* subspecies, (π_w/π_j and π_w/π_i), we found 16 of the 44 RL and 9 of 97 RT candidate genes showing strong selective signals in subspecies *japonica*, in contrast to only 1 of 97 RT candidates in *indica* (Fig 7D, S17 Table). The mean value of $\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}}$ in *japonica* for all candidate genes was 14.6, which is much higher than those of $\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}}$ in *indica* (2.1), π_w/π_j (3.7) and π_w/π_i (1.2) (S17 Table). Eleven RL candidate genes and 16 RT candidate genes showed very strong

selective signals with much higher $\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}}$ value in *japonica* (> 20), and should be key genes for aerobic adaptation of upland *japonica* (S17 Table). Among them, genes *OsRL3.3* and *OsSIZ2* with $\pi_{\text{non-L-T-root}}/\pi_{\text{L-T-root}}$ values in *japonica* of 102.1 and 29.6 were shown to be associated with RL by T-DNA insertion mutant lines (Table 1). Further Tajima's *D* Test was performed for each of the robust-root candidate genes in wild rice, and the four ecotypes of cultivated rice. Tajima's *D* values of 5, 3, 48, 56 and 92 genes were negative in wild rice, non-L-T-root in *indica*, L-T-root in *indica*, non-L-T-root in *japonica* and L-T-root in *japonica*, respectively. Only one, 0, 7, 4 and 66 gene showed strong selection with Tajima's *D* < -1 in above separated groups (S17 Table). These observations indicated that significantly more robust-root candidate genes had undergone natural selection during adaptation to aerobic conditions of the upland *japonica* accessions than in other ecotypes. Therefore, we concluded that almost all robust-root alleles were inherited from wild rice, and they gradually accumulated in 4 *japonica* ecotypes during and after domestication of tropical and temperate *japonica* subpopulations, and more robust-root alleles were accumulated by natural and artificial selection in the upland *japonica* ecotype than in other ecotypes in adapting to aerobic conditions (Fig 7E).

Discussion

Efficient discovery of moderate- and small-effect alleles for robust roots

Considerable effort has been made to understand the genes underlying variation of complex traits by cloning QTLs. However, QTL cloning by using classical map-based approach is time-

Table 1. Description of 9 and 14 genes for root length and root thickness identified by an integrated analyses.

Trait	QTL	Name	ID.	Number of haplotypes	Number of non-synonymous SNP and key InDel	Subpopulations with superior haplotypes	$\pi_{\text{non-L-T-root}} / \pi_{\text{L-T-root}}$ in <i>japonica</i>	$\pi_{\text{non-L-T-root}} / \pi_{\text{L-T-root}}$ in <i>indica</i>	π_w / π_j	π_w / π_i	Description
RL	<i>qMRL1-1</i>	<i>OsRL1.1</i>	LOC_Os01g11520	4	2	All	0.88	1.27	0.75	1.25	RING-H2 finger protein, putative, expressed
	<i>qMRL1-3</i>	<i>OsGH3.1</i>	LOC_Os01g57610	4	3	<i>Japonica</i>	3.71	1.03	1.04	0.83	<i>OsGH3.1</i> —Probable indole-3-acetic acid-amido synthetase, expressed
	<i>qMRL3-2</i>	<i>OsRL3.1</i>	LOC_Os03g48370	4	3	All	15.92	1.51	3.86	0.79	Disease resistance RPP13-like protein 1, putative, expressed
	<i>qMRL3-3</i>	<i>OsRL3.3*</i>	LOC_Os03g50350	3	2	All	102.07	2.84	16.96	2.56	Coatomer alpha subunit, putative, expressed
		<i>OsSIZ2*</i>	LOC_Os03g50980	3	7	All	29.64	1.12	5.18	1.05	ATSIZ1/SIZ1, putative, expressed
		<i>OsRL3.2</i>	LOC_Os03g50470	2	2	All	15.02	2.24	5.79	1.01	Expressed protein
	<i>qMRL7-1</i>	<i>OsRL7.1*</i>	LOC_Os07g03160	3	2	All	3.46	1.93	10.95	2.82	WD domain, G-beta repeat domain containing protein, expressed
	<i>qMRL8-2</i>	<i>OsRL8.2*</i>	LOC_Os08g43040	4	2	Upland <i>japonica</i>	1.69	1.23	1.79	2.10	Transferase family protein, putative, expressed
	<i>qMRL11-1</i>	<i>OsRL11.1*</i>	LOC_Os11g43320	3	7	All	1.05	0.85	0.89	0.54	NBS-LRR type disease resistance protein, putative, expressed

(Continued)

Table 1. (Continued)

Trait	QTL	Name	ID.	Number of haplotypes	Number of non-synonymous SNP and key InDel	Subpopulations with superior haplotypes	$\pi_{\text{non-L-T-root}} / \pi_{\text{L-T-root}}$ in japonica	$\pi_{\text{non-L-T-root}} / \pi_{\text{L-T-root}}$ in indica	π_w / π_j	π_w / π_i	Description
RT	<i>qBRT2-2</i> (J)	<i>OsRT2.1</i>	LOC_Os02g13400	3	4	Japonica	0.59	2.39	0.88	4.81	Expressed protein
	<i>qBRT6-2</i> (J)	<i>OsRT6.1</i>	LOC_Os06g36810	2	2	All	1.48	0.80	1.70	0.96	Expressed protein
	<i>qBRT6-3</i> (J)	<i>OsRT6.2</i>	LOC_Os06g37840	5	5	Japonica	1.53	1.27	0.86	0.95	Resistance protein, putative, expressed
	<i>qBRT4-1</i> (F)	<i>OsRT4.1</i>	LOC_Os04g02780	3	2	All	1.41	0.99	0.67	0.80	Amidase family protein, putative, expressed
	<i>qBRT4-1</i> (I)	<i>OsRT4.2</i>	LOC_Os04g30490	3	2	All	0.61	2.50	0.89	1.04	MATE efflux family protein, putative, expressed
		<i>OsRT4.3</i>	LOC_Os04g30920	3	10	Indica	12.85	4.43	3.11	0.53	Expressed protein
		<i>OsRT4.4</i>	LOC_Os04g31000	3	6	All	3.65	4.65	2.04	0.68	Methyltransferase domain containing protein, expressed
		<i>OsRT4.5</i>	LOC_Os04g31010	3	2	Indica	10.85	3.65	4.35	0.81	Expressed protein
		<i>OsRT4.6</i>	LOC_Os04g31020	3	7	Indica	2.80	3.34	3.29	0.91	Expressed protein
		<i>OsRT4.7</i>	LOC_Os04g31120	3	5	Indica	5.60	2.30	4.36	0.83	<i>OsFBK14</i> —F-box domain and kelch repeat containing protein, expressed
	<i>qBRT4-4</i> (I)	<i>OsRT4.8</i>	LOC_Os04g51060	3	9	Indica	2.93	2.25	3.31	0.60	Expressed protein
		<i>OsRT4.9</i>	LOC_Os04g51080	2	2	Indica	1.87	1.66	2.14	0.67	Scramblase, putative, expressed
	<i>qBRT8-2</i> (I)	<i>OsRT8.1</i>	LOC_Os08g14660	5	12	All	1.12	0.99	1.34	0.87	SET domain containing protein, expressed
		<i>OsRT8.2</i>	LOC_Os08g15050	3	2	Indica	0.64	1.46	0.91	0.97	CCT/B-box zinc finger protein, putative, expressed

*, Root length candidate gene is validated by T-DNA insertional mutation.

"F", "J" and "I" in bracket showed that the GWAS QTL was detected in whole population, japonica and indica, respectively.

<https://doi.org/10.1371/journal.pgen.1007521.t001>

consuming and effective only for large-effect QTLs. Cloning QTLs of moderate to small effect is highly challenging owing to the difficulty of phenotyping large populations, particularly for plant root traits such as RL and RT that are difficult to measure under field conditions. In this study we adopted a modified hydroponics system (see [Materials and Methods](#)) to measure root traits in rice. This allowed genetic potential to be assessed in the absence of stress [33,79]. Genetically, it remains an unanswered question as to whether large-effect QTLs are primary sources of genetic variation for complex quantitative traits, even though most cloned QTLs in rice and other plants are regulatory genes with large and pleiotropic effects on multiple traits [80,81]. Thus, identification of large numbers of RL and RT QTLs and cloning five of the RL QTLs of small effect in this study should be considered highly successful. This improvement was attributed to our integrated strategy that combined results from high resolution mapping by GWAS and linkage mapping, and comprehensive bioinformatic analyses of genomic, transcriptomic and haplotype data to shortlist QTL candidate genes plus validation of key candidate genes by insertional mutants. With the increasing availability of various kinds of -omics

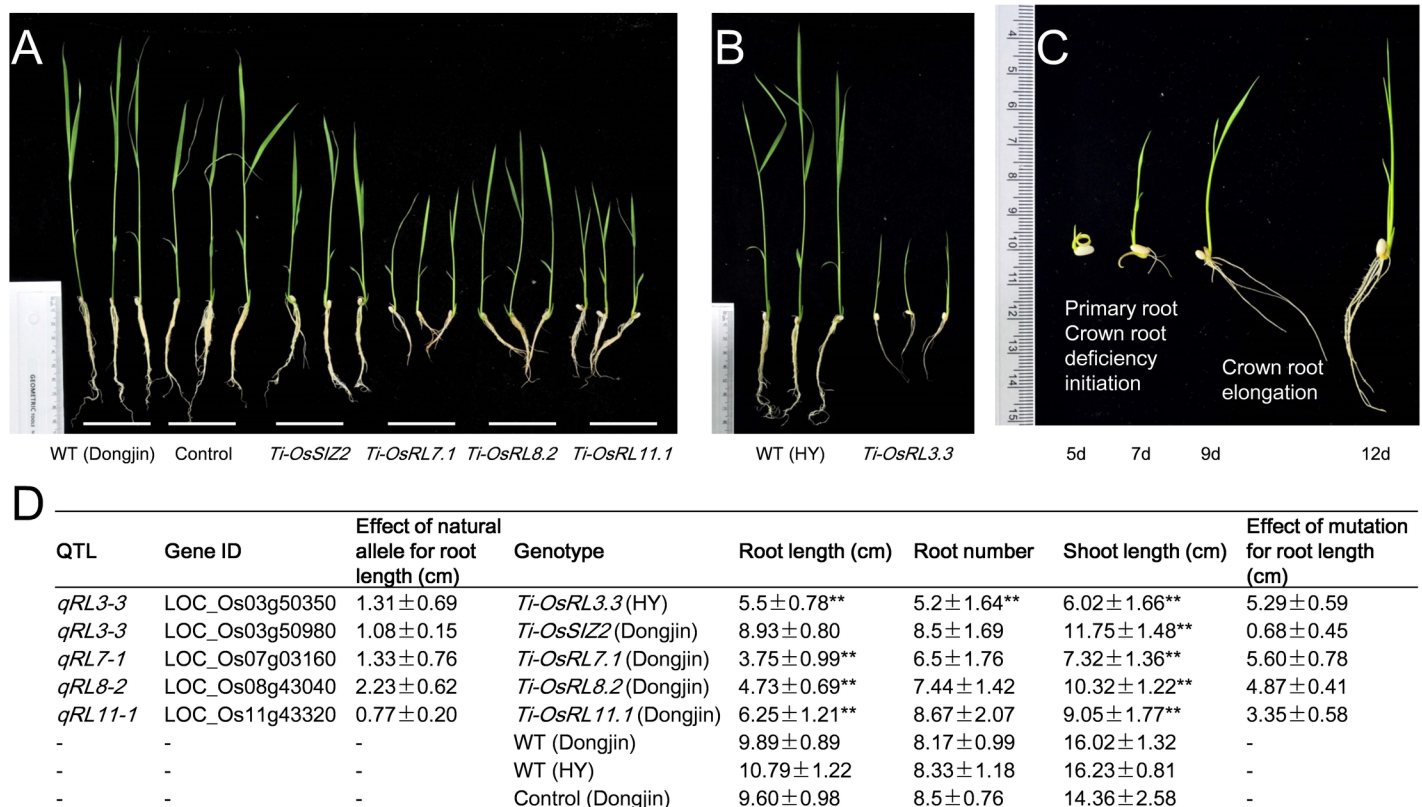


Fig 6. Validation of five QTL genes for root length by T-DNA insertion mutants. (A) Comparisons of root lengths of 12-day-old seedlings among the wild type (WT, Dongjin), control (mutant with a T-DNA insertion in *HsfA4a*) and T-DNA insertion mutant lines *Ti-OsSIZ2*, *Ti-OsRL7.1*, *Ti-OsRL8.2* and *Ti-OsRL11.1*. (B) Comparisons of root lengths of 12-day-old seedlings among the wild type (WT, Hwayoung) and T-DNA insertion mutant line *Ti-OsRL3.3*. (C) The root morphologies of line *Ti-OsRL3.3* at different growth stages. (D) Comparison of root lengths, root numbers and shoot lengths among the wild type, control and different T-DNA insertion lines at 12 days after germination. **, significant ($P < 0.01$) difference between the mutant line compared with WT and control.

<https://doi.org/10.1371/journal.pgen.1007521.g006>

data and genetic stocks such as genome-wide insertional mutants and introgression lines [82–85], it is expected that the integrated strategy demonstrated here will be widely applied in large scale molecular dissection of genes underlying complex traits by efficient cloning and characterization of small- and moderate-effect QTLs. In this respect, our results revealed several unique properties of genes underlying variation of complex traits and shed light on how natural QTL alleles contribute to robust-root systems and adaptation of rice to the aerobic conditions.

Our results indicated that major haplotypes, consisting of non-synonymous SNPs in the coding sequence (CDS) and/or promoter regions within single loci, represent the most important allelic diversity of moderate- and small-effect QTLs underlying robust-root variation in the rice populations. This conclusion was supported by at least five pieces of evidence. First, when examining large numbers of candidate genes in >100 RL and RT QTL regions, we detected 2 to 5 predominant haplotypes (alleles) associated with significant differences in RL or RT, each comprising 2 to 23 non-synonymous SNPs in the CDS or promoter region of each RL and RT QTL candidate locus, including the 5 validated RL QTL genes (Fig 6, S12 Fig). Second, compared with random SNPs, specific haplotypes at single loci showed much greater differentiation in gene frequency among populations, indicating they were important targets of selection during adaptation to aerobic conditions. Third, phenotypic differences between different haplotypes at each of the cloned RL gene loci, though statistically significant, were quite small and did not appear to have pleiotropic effects when compared with lines having

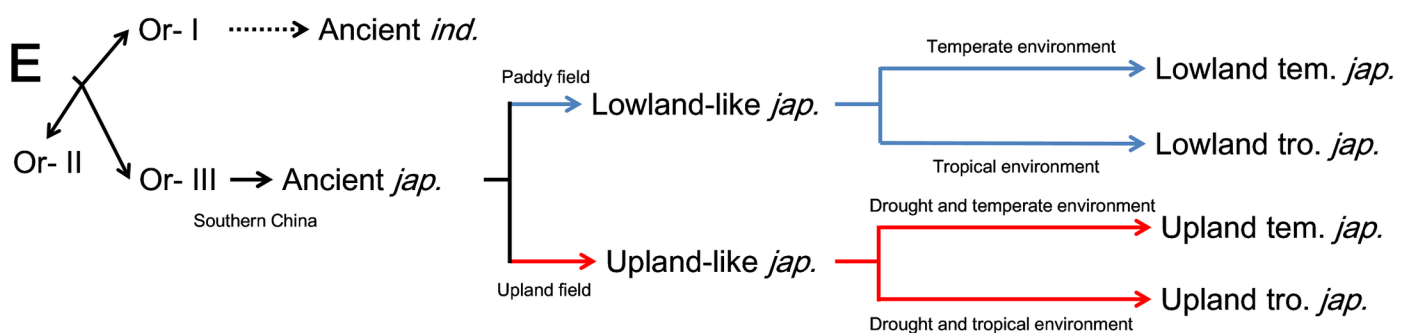
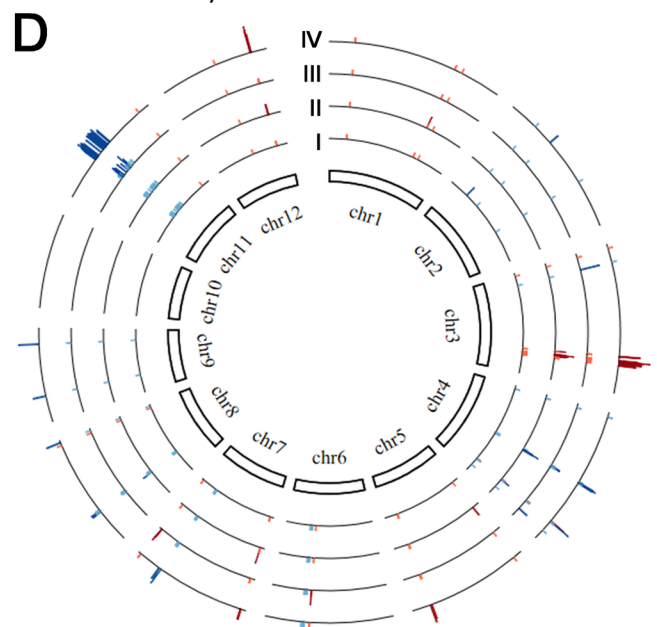
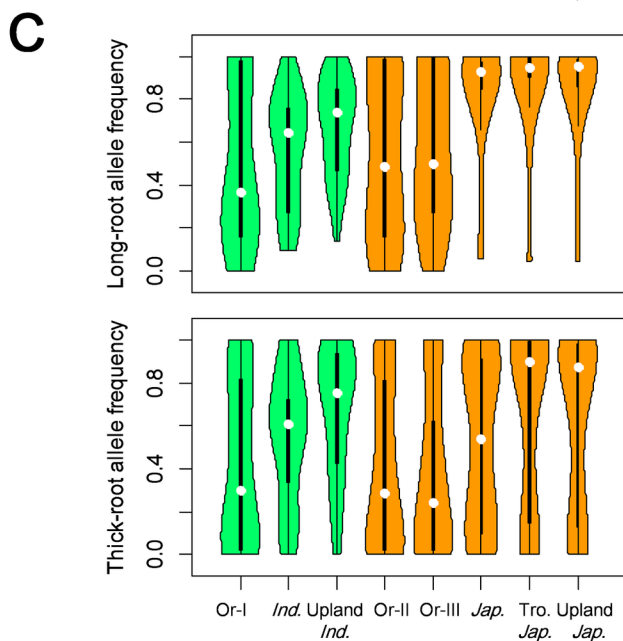
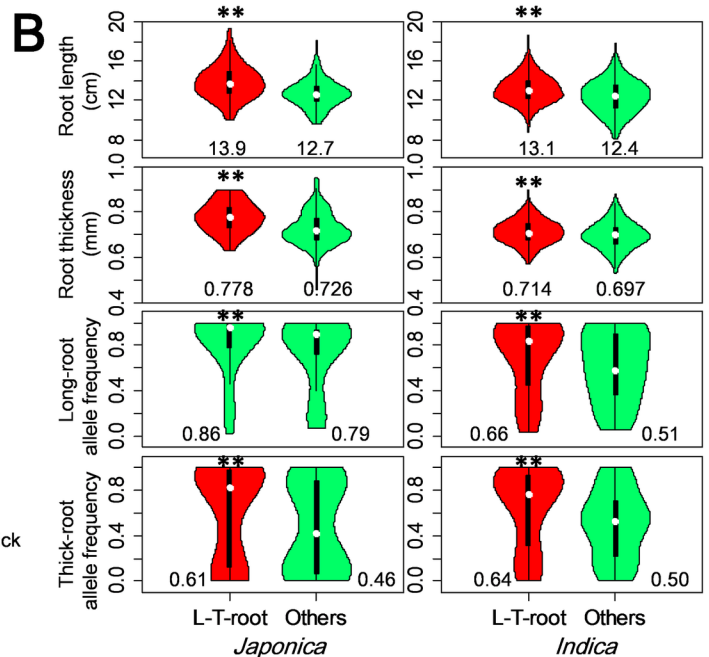
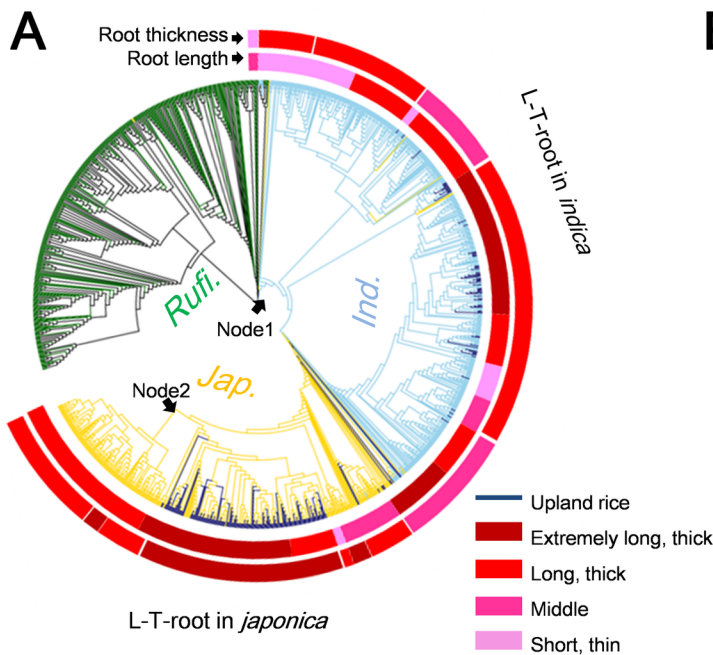


Fig 7. A hypothetical adaptive domestication history of upland rice accessions with robust roots. (A) The phylogenetic tree of the 446 *O. rufipogon* and 997 *O. sativa* accessions based on 5,779 SNPs within the candidate genes for robust roots. (B) Comparisons of robust-root phenotypes and allele frequencies between accessions in the long-thick-root (L-T-root) region and other varieties in the corresponding subpopulations; in which the X-axis indicates different groups, and the Y-axis indicates the mean root length and thickness values plotted in the top panel and the long-root and thick-root allele frequencies plotted in the bottom of panel. There are significant differences between red violin and green violin distributions in each plot, and the number represents average value of the adjacent violin. (C) Comparisons of frequencies of the long-root alleles and thick-root alleles at candidate genes in *O. rufipogon* and *O. sativa* subpopulations and ecotypes. Y-axes show allele frequencies for long-roots (upper) and thick-roots (lower), and X-axis represents different groups. Or-I is *indica*-like *O. rufipogon* accessions colored blue as *indica*, and Or-II and Or-III are intermediate and *japonica*-like *O. rufipogon* accessions colored yellow as *japonica*. (D) Ratios of genetic diversity among different groups. The log₂-transformed ratios of π_w/π_l , π_w/π_t , $\pi_{non-L-T-root}/\pi_{L-T-root}$ in *indica* and $\pi_{non-L-T-root}/\pi_{L-T-root}$ in *japonica* are plotted against the position on each chromosome, marked by Roman numerals I, II, III and IV. Red and blue represent root length and root thickness candidate genes and the bar lengths indicate log₂-transformed ratios. Details are listed in S17 Table. (E) Adaptive domestication schematics of upland rice in aerobic conditions. Black lines and dotted line show known and unclear domestication history. Blue and red lines indicate adaptive domestication history of lowland and upland *japonica* ecotypes in tropical and subtropical environments and temperate environment, respectively.

<https://doi.org/10.1371/journal.pgen.1007521.g007>

insertional mutations in the same genes. Fourth, phenotypic differences between haplotypes at many of these QTL gene loci could be, at least partially, due to differences at the transcriptomic level. This suggests that fine tuning of complex traits in crop improvement can be achieved by careful manipulation of cloned QTL genes at the transcriptional level using various molecular technologies. Fifth, QTL genes affecting the same RL or RT phenotype appeared to have diverse molecular functions, and few of them were regulatory genes (Table 1). These properties distinguish the QTL genes identified in this study from most cloned large-effect QTL genes reported previously. Nevertheless, it remains a challenge to understand how different alleles at moderate- to small-effect QTL loci control complex traits at the molecular level and how they interact with the environment.

Adaptive domestication of the upland *japonica* ecotype

O. sativa is well known for its rich diversity and wide subspecies and population differentiation [2,3]. The domestication analysis using 446 *O. rufipogon* and 1,083 *O. sativa* accessions indicated that *japonica* was first domesticated from a specific population of *O. rufipogon* in the Pearl River Valley in southern China [2]. However, adaptation process to aerobic conditions of upland *japonica* rice remains unclear.

Our previous survey of 50,265 rice landraces (33,665 *indica* and 16,784 *japonica*) indicated that upland *japonica* landraces were distributed primarily in hilly areas of four southern and southeastern China provinces (41% in Yunnan, 10% in Guangxi, 16% in Guizhou and 21% in Hainan, respectively) [11] (S15 Fig). The genetic structure of *O. sativa* indicated that *japonica* was clearly differentiated between soil water regime ecotypes (uplands and lowlands) in Yunnan [9], Guizhou [86] and many parts of China [4–6], caused by different environments and cropping systems. Based on previous studies of geographical distribution and population structure, we have reasons to believe that upland *japonica* is a unique ecotype with specific morphologies and genotypes associated with drought resistance.

Consistent with previous research, our study verified that robust roots are characteristics of upland rice [10,12,13]. Therefore, we analyzed adaptation process to aerobic conditions of upland *japonica* rice based on 5,779 SNPs in the robust-root candidate genes. The phylogenetic tree showed that there was lower diversity and higher population structure consistency in robust-root candidate genes in the L-T root ecotype in *japonica* (89 tropical *japonica*, 36 temperate *japonica*, 1 *japonica* and 15 intermediate), suggesting that the genotype of robust roots in upland rice was different from that in other ecotypes. By further analysis of nucleotide diversity and allele frequency at robust loci, our study provided new insights into adaptive domestication of the upland rice ecotype. In summary, upland rice accumulated more robust-root alleles in adapting to aerobic conditions during domestication than present in other ecotypes. Additionally, we suggest that the upland *japonica* ecotype is a typical upland rice ecotype

with the most robust roots, highest number of robust-root alleles, and strongest selective signals among all tested ecotypes.

Materials and methods

Materials

A total of 795 *O. sativa* accessions from the 3000 Rice Genome Project (3KRGP) [83,84] were used for identification of RL and RT QTLs, including mini-core collections constructed by Zichao Li in China Agricultural University and 525 lines in the International Rice Molecular Breeding Network constructed by Zhikang Li in Chinese Academy of Agricultural Science [51,52]. All information of cultural type for these rice accessions were from The Catalog of Rice Germplasm Resources in China and IRRI, respectively (S1 Table) [11]. For an adaptive domestication analysis, we added 202 additional upland landrace accessions from 19 countries from 3KRGP and 446 *O. rufipogon* accessions from Asia and Oceania [2]. The background information of cultural type for the additional 202 upland landrace accessions was from IRRI.

Phenotyping of root length and root thickness

All 795 *O. sativa* accessions were used for phenotyping RL and RT. Two robust-root varieties, HGL and IRAT109, and two lowland temperate *japonica* varieties, Nipponbare and Yuefu, were used as L-T-root and non-L-T-root checks. Four additional long-root varieties (CH1052, CH1086, CH1198 and CH1183) and four non-long-root varieties (CH1179, CH1008, CH1283 and CH1122) were also used as checks in transcriptomic analysis (S1 Table). The hydroponic culture experiment was conducted at China Agricultural University in 2014. Seed of each accession were washed with distilled-water and germinated at 32°C for 64 hours. Five uniformly germinated seeds of each variety were placed into five wells on a plastic foam frame with gauze bottom (30 cm × 42 cm) containing 130 (10 wells × 13 rows) wells, with 110 plants of 22 varieties and 20 plants of two bordering rows (S16 Fig). Two frames were floated in a plastic box (60 cm × 42 cm × 18 cm) containing Yoshida nutrient solution [87]. The pH (5.5) and concentration of the nutrient solution were adjusted twice daily with NaOH and distilled water, and the solution was replaced weekly. Plants were grown under natural conditions but protected with tarpaulins on rainy days. The experiment was conducted twice as two replications from May 24 to June 15 and from June 23 to July 15 in 2014, during which the average daily temperatures fluctuated from 19.9°C to 31.7°C and from 23.9°C to 32.9°C, respectively. Seedlings were allowed to grow for 23 days and then five plants per variety were sampled and measured for RL, RT and root weight using a method described previously [23]. The mean values of all five plants of each accession were used as the input data in the data analyses.

Genome sequencing

Sequencing data of the 795 and additional upland landraces were obtained from the 3KRGP, and had an average sequencing depth of 15× and generated > 15 million SNPs and > 2 million small InDels when compared with the Nipponbare reference genome [84]. The sequencing data of 446 *O. rufipogon* accessions were downloaded from the public database [2] (S1 File).

Transcriptome sequencing

Six robust-root cultivars and 6 non-robust-root controls were used for transcriptomic analyses. The 12 varieties were planted in pot and field experiments. At the four-leaf stage, roots and shoots of each variety were sampled for extraction of total RNA using TRIzol reagent. High-quality RNA was used for construction of RNA-seq libraries with three biological replicates for

each sampled variety. Upon completion of sequencing libraries using the TruSeq RNA sample preparation kit (Illumina), sequencing was performed on an Illumina HiSeq 2500. More than 12.8 million 101-bp single-ends were generated in each sample. Mapping of RNAseq reads and transcript abundance RPKM was performed by the Cufflinks package version 2.0, based on the rice reference genome and gene model annotation file (GFF transformed GTF file) from the Ensembl database (<http://plants.ensembl.org/index.html>). The basic information of the transcriptomic data is provided in [S18 Table](#).

Population structure

Based on 3.3 million un-imputed SNPs with missing rates $\leq 50\%$ and minor allele frequencies $\geq 5\%$, we further extracted 154,516 SNPs with missing rates $\leq 50\%$, minor allele frequencies $\geq 5\%$ and r^2 of LD ≤ 0.3 using PLINK [88]. The software ADMIXTURE 1.3 was used to calculate varying levels of K ($K = 1-10$), and $K = 8$ was a suggesting modeling choice ([S1B and S1C Fig](#)) [89].

GWAS and conditional permutation test

Based on 3.3 million un-imputed SNPs with missing rates $\leq 50\%$ and minor allele frequencies $\geq 5\%$, principal component (PC) and kinship analyses were computed to verify population structure and relationship among the 795 cultivated rice accessions in software GAPIT. The first three PCs were used to construct the PC matrix. We performed GWAS with a Compressed Mixed Linear Model (CMLM) with PC and kinship and GLM with PC using default settings of software GAPIT. A conditional permutation was used to determine the significance threshold. We divided the phenotype (Y) of each accession into the original genotypic effect (G) and the fixed effect of population structure (P). P was estimated by the average effect of each PC in the PC matrix on each individual through regression analysis of each PC on Y ; the remainder was G after excluding P from Y . G was randomly reshuffled as G_r , and the new phenotype of each accession was reconstructed as $P + G_r$. The conditional permutation test was executed using CMLM with the same parameters and PC matrix. A total of 1,000 sets of $P + G_r$ were performed for root length and root thickness. SNPs with $-\log(P) \geq 2$ in the GWAS using original phenotypes were extracted to improve the computational efficiency. Finally a threshold of $-\log(P) = 4$ was determined, which was higher than the 95th percentile of 1,000 conditional permutation tests. LD heatmaps of two candidate genes (*OsRL3.3* and *OsSIZ2*) were generated using the R package “LD heatmaps”.

Pooled- χ^2 -tests

For screening SNPs associated with target traits, we constructed two pools for RL and RT by selecting accessions with extreme phenotypes from typical *indica* and *japonica* populations, then performed a chi-square test on the allele frequency of each SNP executed by in-house Perl scripts. To reduce genetic differences in unrelated target trait between polar pools, we took no account of varieties at PC1 values ranging from -400 to 300. Varieties with PC1 < -400 were defined as typical *japonica* and varieties with PC1 > 300 were considered to be typical *indica* ([Fig 1A](#)). Each pool included 20 varieties with extreme phenotypes in typical *indica* and *japonica*, respectively ([S1 Table](#)).

Analysis of T-DNA insertional mutants

After haplotype analysis of candidate genes for robust roots, we obtained T-DNA insertion mutants for 9 RL candidate genes. Six mutant plants containing T-DNA insertions (*Ti-OsRL3.3*, *Ti-OsSIZ2*, *Ti-OsRL7.1*, *Ti-Os8.2*, *Ti-Os11.1* and *HsfA4a*) in ‘Dongjin’ or ‘Hwayoung’ background were from the POSTECH Biotech Center, Republic of Korea. Seeds of the six

mutant lines were surface-sterilized with 75% ethanol for 2 min and in 20% NaClO for 30 min, and thoroughly washed with sterile water. Medium cultivation experiments were conducted in a glasshouse with a 14h light/10h darkness and temperature 28°C using MS Medium, pH 5.8 (S17 Fig). After 12 days, RL, RT and root weight were measured with a ruler and an electronic scales. DNA was extracted from leaves and was used for identification of homozygous and heterozygous mutants by PCR using gene-specific primers (LP and RP) coupled with a T-DNA-specific primer (RB or LB). PCR was conducted with an initial step of incubation at 95°C for 5 min, followed by a second step of 35 cycles of 95°C for 40s, 58°C (55°C) for 40s, and 72°C for 1 min 20s. The PCR products were genotyped by a sequencing company.

Phylogenetic analysis

We downloaded the SNP set (7,970,357) of 1,529 rice lines from a recent study and selected SNP data for the 446 wild rice accessions that were simply classified into three types Or-I, Or-II and Or-III in this study [2]. Since this dataset used IRGSP4 Nipponbare genome as reference, we extracted the 200 bp flanking regions from each side of the target SNP (200+1+200 = 401 bp) from the IRGSP4 reference and mapped 401 bp sequences to the RGAP 7 reference by BWA software [90]. We converted the SNP coordinate of IRGSP4 to RGAP 7 based on the mapping result and then integrated the genetic information for 997 cultivated rice accessions and 446 wild rice accessions following extraction of 5,779 SNPs within robust-root candidate genes (44 genes for RL and 97 genes for RT) from 1,443 accessions (S1 File). A neighbor-joining tree based on the polymorphisms was generated using Tassel 5 and Mega 6, and the nucleotide diversity (π) [91] was calculated using an in-house Perl script (S1 File). We also extracted 90,838 SNPs evenly distributed throughout the genomes of 1,443 accessions, and a constructed neighbor-joining tree as control.

Supporting information

S1 Fig. Quality test of genome sequencing, and the population structure and kinship of 795 *O. sativa* (rice) accessions.

(DOCX)

S2 Fig. Histograms of phenotypic diversity in root traits of the whole population and different ecotypes of rice.

(DOCX)

S3 Fig. Distribution of root weight among different rice ecotypes.

(DOCX)

S4 Fig. Genome-wide association study of root length under GLM.

(DOCX)

S5 Fig. Genome-wide association study of root thickness under GLM.

(DOCX)

S6 Fig. The conditional permutation test for determining the threshold in the whole population using the CMLM model.

(DOCX)

S7 Fig. Identification of putative functional genes in 19 QTL regions detected by GWAS.

(DOCX)

S8 Fig. Haplotype analyses of 4 key root length candidate genes.

(DOCX)

S9 Fig. Haplotype analyses of 14 key root thickness candidate genes.

(DOCX)

S10 Fig. Genotypic analysis of four mutant lines and control plants.

(DOCX)

S11 Fig. Phenotypes of primary roots in wild type Dongjin and T-DNA insertion mutant *Ti-OsRL3.3* lines.

(DOCX)

S12 Fig. LD heatmap for 9 non-synonymous SNPs in *OsRL3.3* and *OsSIZ2* for *qRL3-3*.

(DOCX)

S13 Fig. The neighbor-joining tree of 997 *O. sativa* and 446 wild *O. rufipogon* accessions based on evenly distributed 90,838 SNPs across the genome.

(DOCX)

S14 Fig. Phylogenetic relationships among the sampled population based on SNPs within the robust-root candidate genes.

(DOCX)

S15 Fig. Geographical distributions of 3,027 upland *japonica* landraces from 50,265 rice landraces in China.

(DOCX)

S16 Fig. The phenotyping system for evaluating rice root traits under the hydroponic conditions.

(DOCX)

S17 Fig. Root phenotypes of wild type Dongjin, control (mutant with T-DNA insertion in known gene *HsfA4a*), and T-DNA insertion mutant lines *Ti-OsSIZ2*, *Ti-OsRL7.1*, *Ti-OsRL8.2* and *Ti-OsRL11.1*.

(DOCX)

S1 Table. Description of 795 rice accessions, 202 additional landraces and 446 wild rice accessions.

(XLSX)

S2 Table. Descriptive statistics of 3 root traits among different ecotypes.

(XLSX)

S3 Table. Summary of QTLs and SNPs associated with root length detected by GWAS.

(XLSX)

S4 Table. Summary of QTLs and SNPs associated with root thickness detected by GWAS.

(XLSX)

S5 Table. Distribution of different allele types at 857 SNP sites in pools of long and short roots and corresponding subpopulations.

(XLSX)

S6 Table. Long-root allele frequencies of 230 SNP sites with significant difference between pools of long and short roots.

(XLSX)

S7 Table. Description of common root length and root thickness QTLs from GWAS and linkage mapping in rice.

(XLSX)

S8 Table. Summary of 119 SNPs from 7 root length common QTL identified by GWAS and linkage mapping.

(XLSX)

S9 Table. Summary of 146 SNPs in 81 genes identified by GWAS, linkage mapping and pooled- χ^2 -tests.

(XLSX)

S10 Table. Screening of SNPs related to root thickness in GWAS-QTLs by pool- χ^2 -test and linkage mapping.

(XLSX)

S11 Table. RPKM of 44 root length candidate genes in long-root varieties and controls.

(XLSX)

S12 Table. SNPs leading amino acid polymorphism for 44 root length candidate genes in pools, parents and populations.

(XLSX)

S13 Table. InDels leading amino acid polymorphism for 20 root length candidate genes in pools and populations.

(XLSX)

S14 Table. Analysis of expression and amino acid variation for 97 root thickness key candidate genes.

(XLSX)

S15 Table. SNPs leading amino acid polymorphisms for root thickness candidate genes in pools, parents and populations.

(XLSX)

S16 Table. Primers used in this study.

(XLSX)

S17 Table. Ratio of genetic diversity and Tajima's *D* test of root length and root thickness candidate genes among different populations.

(XLSX)

S18 Table. RPKM of all annotated gene in long-root varieties and controls.

(XLSX)

S1 File. Main Perl scripts used in our study.

(ZIP)

Acknowledgments

We thank Robert McIntosh (University of Sydney) for critical reading and suggested revisions to the manuscript.

Author Contributions

Conceptualization: Xueqiang Wang, Zhikang Li, Zichao Li.

Data curation: Yan Zhao, Jianlong Xu, Zhigang Yin, Yang Li, Weipeng Zhao, Zhikang Li, Zichao Li.

Formal analysis: Yan Zhao, Conghui Jiang, Haiyan Xiong, Jianyin Xie, Xiaoyang Zhu, Zhiqiang Hu.

Funding acquisition: Zichao Li.

Investigation: Yan Zhao.

Methodology: Yan Zhao, Hongliang Zhang, Jianlong Xu, Zhikang Li.

Project administration: Zichao Li.

Resources: Zhikang Li.

Supervision: Jianlong Xu, Jinjie Li, Wensheng Wang, Binying Fu, Guoyou Ye, Yan Guo.

Validation: Yan Zhao, Conghui Jiang.

Visualization: Xiaoyang Zhu.

Writing – original draft: Yan Zhao, Zhikang Li.

Writing – review & editing: Muhammad Abdul Rehman Rashid, Zhikang Li, Zichao Li.

References

1. Wang W, Mauleon R, Hu Z, Chebotarov D, Tai S, et al. (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557: 43–49. <https://doi.org/10.1038/s41586-018-0063-9> PMID: 29695866
2. Huang X, Kurata N, Wei X, Wang Z-X, Wang A, et al. (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490: 497–501. <https://doi.org/10.1038/nature11532> PMID: 23034647
3. Civián P, Craig H, Cox CJ, Brown TA (2015) Three geographically separate domestications of Asian rice. *Nature Plants* 1: 15164. <https://doi.org/10.1038/nplants.2015.164> PMID: 27251535
4. Zhang D, Zhang H, Wang M, Sun J, Qi Y, et al. (2009) Genetic structure and differentiation of *Oryza sativa* L. in China revealed by microsatellites. *Theoretical and Applied Genetics* 119: 1105–1117. <https://doi.org/10.1007/s00122-009-1112-4> PMID: 19649611
5. Zhang D, Wang M, Qi Y, Sun J, Wang F, et al. (2012) Genetic structure and eco-geographical differentiation of cultivated *keng* rice (*Oryza sativa* L. subsp. *japonica*) in China revealed by microsatellites. *Journal of Integrative Agriculture* 11: 1755–1766.
6. Zhang D, Zhang H, Qi Y, Wang M, Sun J, et al. (2013) Genetic structure and eco-geographical differentiation of cultivated *Hsien* rice (*Oryza sativa* L. subsp. *indica*) in China revealed by microsatellites. *Chinese Science Bulletin* 11: 344–352.
7. Bridhikitti A, Overcamp TJ (2012) Estimation of Southeast Asian rice paddy areas with different ecosystems from moderate-resolution satellite imagery. *Agriculture Ecosystems & Environment* 146: 113–120.
8. Zhang L, Cao G, Han L (2010) Analysis of genetic similarity for upland landrace rices from different geographical origins in China. *Scientia Agricultura Sinica* 43: 3481–3488.
9. Zhang H, Sun J, Wang M, Liao D, Zeng Y, et al. (2007) Genetic structure and phylogeography of rice landraces in Yunnan, China, revealed by SSR. *Genome* 50: 72–83. <https://doi.org/10.1139/g06-130> PMID: 17546073
10. IRRI (1970) Upland and lowland conditions for rice: International Rice Research Institute Annual Report.
11. CAAS (1992) The catalog of rice germplasm resources in China. Beijing: China Agriculture Science Press.
12. Lyu J, Zhang S, Dong Y, He W, Zhang J, et al. (2013) Analysis of elite variety tag SNPs reveals an important allele in upland rice. *Nature Communications* 4: 2138. <https://doi.org/10.1038/ncomms3138> PMID: 23828614
13. Gupta PC, O'Toole JC (1986) Upland rice: a global perspective: International Rice Research Institute.
14. Gowda VRP, Henry A, Yamauchi A, Shashidhar HE, Serraj R (2011) Root biology and genetic improvement for drought avoidance in rice. *Field Crops Research* 122: 1–13.

15. Yoshida S, Hasegawa S (1982) The rice root system: its development and function: International Rice Research Institute. 97–114 p.
16. Comas LH, Becker SR, Cruz VMV, Byrne PF, Dierig DA (2013) Root traits contributing to plant productivity under drought. *Frontiers in Plant Science* 4: 442. <https://doi.org/10.3389/fpls.2013.00442> PMID: 24204374
17. Lafitte HR, Champoux MC, McLaren G, O'Toole JC (2001) Rice root morphological traits are related to isozyme group and adaptation. *Field Crops Research* 71: 57–70.
18. Lyu J, Li B, He W, Zhang S, Gou Z, et al. (2014) A genomic perspective on the important genetic mechanisms of upland adaptation of rice. *BMC Plant Biology* 14: 160. <https://doi.org/10.1186/1471-2229-14-160> PMID: 24920279
19. Coudert Y, Périn C, Courtois B, Khong NG, Gantet P (2010) Genetic control of root development in rice, the model cereal. *Trends in Plant Science* 15: 219–226. <https://doi.org/10.1016/j.tplants.2010.01.008> PMID: 20153971
20. Richards RA (2008) Genetic opportunities to improve cereal root systems for dryland agriculture. *Plant Production Science* 11: 12–16.
21. de Dorlodot S, Forster B, Pagès L, Price A, Tuberosa R, et al. (2007) Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends in Plant Science* 12: 474–481. <https://doi.org/10.1016/j.tplants.2007.08.012> PMID: 17822944
22. Qu Y, Mu P, Zhang H, Chen CY, Gao Y, et al. (2008) Mapping QTLs of root morphological traits at different growth stages in rice. *Genetica* 133: 187–200. <https://doi.org/10.1007/s10709-007-9199-5> PMID: 17823843
23. Li J, Wang D, Xie Y, Zhang H, Hu G, et al. (2011) Development of upland rice introgression lines and identification of QTLs for basal root thickness under different water regimes. *Journal of Genetics and Genomics* 38: 547–556. <https://doi.org/10.1016/j.jgg.2011.08.005> PMID: 22133686
24. Obara M, Tamura W, Ebitani T, Yano M, Sato T, et al. (2010) Fine-mapping of qRL6.1, a major QTL for root length of rice seedlings grown under a wide range of NH₄⁺ concentrations in hydroponic conditions. *Theoretical and Applied Genetics* 121: 535–547. <https://doi.org/10.1007/s00122-010-1328-3> PMID: 20390245
25. Li Z, Mu P, Li C, Zhang H, Li Z, et al. (2005) QTL mapping of root traits in a doubled haploid population from a cross between upland and lowland *japonica* rice in three environments. *Theoretical and Applied Genetics* 110: 1244–1252. <https://doi.org/10.1007/s00122-005-1958-z> PMID: 15765223
26. Wang H, Xu X, Zhan X, Zhai R, Wu W, et al. (2013) Identification of qRL7, a major quantitative trait locus associated with rice root length in hydroponic conditions. *Breeding Science* 63: 267–274. <https://doi.org/10.1270/jsbbs.63.267> PMID: 24273421
27. Uga Y, Okuno K, Yano M (2011) Dro1, a major QTL involved in deep rooting of rice under upland field conditions. *Journal of Experimental Botany* 62: 2485–2494. <https://doi.org/10.1093/jxb/erq429> PMID: 21212298
28. Uga Y, Yamamoto E, Kanno N, Kawai S, Mizubayashi T, et al. (2013) A major QTL controlling deep rooting on rice chromosome 4. *Scientific Reports* 3: 3040. <https://doi.org/10.1038/srep03040> PMID: 24154623
29. Uga Y, Hanzawa E, Nagai S, Sasaki K, Yano M, et al. (2012) Identification of qSOR1, a major rice QTL involved in soil-surface rooting in paddy fields. *Theoretical and Applied Genetics* 124: 75–86. <https://doi.org/10.1007/s00122-011-1688-3> PMID: 21894467
30. Hanzawa E, Sasaki K, Nagai S, Obara M, Fukuta Y, et al. (2013) Isolation of a novel mutant gene for soil-surface rooting in rice (*Oryza sativa* L.). *Rice* 6: 30. <https://doi.org/10.1186/1939-8433-6-30> PMID: 24280269
31. Li J, Han Y, Liu L, Chen Y, Du Y, et al. (2015) qRT9, a quantitative trait locus controlling root thickness and root length in upland rice. *Journal of Experimental Botany* 66: 2723–2732. <https://doi.org/10.1093/jxb/erv076> PMID: 25769309
32. Uga Y, Sugimoto K, Ogawa S, Rane J, Ishitani M, et al. (2013) Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nature Genetics* 45: 1097–1102. <https://doi.org/10.1038/ng.2725> PMID: 23913002
33. Courtois B, Audebert A, Dardou A, Roques S, Ghneim-Herrera T, et al. (2013) Genome-wide association mapping of root traits in a *japonica* rice panel. *Plos One* 8: e78037. <https://doi.org/10.1371/journal.pone.0078037> PMID: 24223758
34. Si L, Chen J, Huang X, Hao G, Luo J, et al. (2016) OsSPL13 controls grain size in cultivated rice. *Nature Genetics* 48: 447–458. <https://doi.org/10.1038/ng.3518> PMID: 26950093
35. Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P-c, et al. (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nature Genetics* 48: 927–934. <https://doi.org/10.1038/ng.3596> PMID: 27322545

36. Li X, Guo Z, Lv Y, Cen X, Ding X, et al. (2017) Genetic control of the root system in rice under normal and drought stress conditions by genome-wide association study. *Plos Genetics* 13: e1006889. <https://doi.org/10.1371/journal.pgen.1006889> PMID: 28686596
37. Huang L, Zhang F, Zhang F, Wang W, Zhou Y, et al. (2014) Comparative transcriptome sequencing of tolerant rice introgression line and its parents in response to drought stress. *BMC Genomics* 15: 1026. <https://doi.org/10.1186/1471-2164-15-1026> PMID: 25428615
38. Moumeni A, Satoh K, Kondoh H, Asano T, Hosaka A, et al. (2011) Comparative analysis of root transcriptome profiles of two pairs of drought-tolerant and susceptible rice near-isogenic lines under different drought stress. *BMC Plant Biology* 11: 174. <https://doi.org/10.1186/1471-2229-11-174> PMID: 22136218
39. Ambavaram MMR, Basu S, Krishnan A, Ramegowda V, Batlang U, et al. (2014) Coordinated regulation of photosynthesis in rice increases yield and tolerance to environmental stress. *Nature Communications* 5: 5302. <https://doi.org/10.1038/ncomms6302> PMID: 25358745
40. Takagi H, Tamiru M, Abe A, Yoshida K, Uemura A, et al. (2015) MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nature Biotechnology* 33: 445–449. <https://doi.org/10.1038/nbt.3188> PMID: 25798936
41. Yu J, Xiong H, Zhu X, Zhang H, Li H, et al. (2017) OsLG3 contributing to rice grain length and yield was mined by Ho-LAMap. *BMC Biology* 15: 28. <https://doi.org/10.1186/s12915-017-0365-7> PMID: 28385155
42. Khush GS (1997) Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology* 35: 25–34. PMID: 9291957
43. Colmer TD (2003) Aerenchyma and an inducible barrier to radial oxygen loss facilitate root aeration in upland, paddy and deep water Rice (*Oryza sativa* L.). *Annals of Botany* 91: 301–309. <https://doi.org/10.1093/aob/mcf114> PMID: 12509350
44. Xia H, Zheng X, Chen L, Gao H, Yang H, et al. (2014) Genetic differentiation revealed by selective loci of drought-responding EST-SSRs between upland and lowland rice in China. *Plos One* 9: e106352. <https://doi.org/10.1371/journal.pone.0106352> PMID: 25286109
45. Kato Y, Kamoshita A, Yamagishi J (2007) Evaluating the resistance of six rice cultivars to drought: restriction of deep rooting and the use of raised beds. *Plant & Soil* 300: 149–161.
46. Ganapathy S, Ganesh SK (2008) Heterosis analysis for physio-morphological traits in relation to drought tolerance in rice (*Oryza sativa* L.). *World Journal of Agricultural Sciences* 4: 623–629.
47. Selvaraj R, Robin S, Muthurajan R (2014) Correlation analysis for physio-morphological traits of drought resistance in advanced backcross population of rice (*Oryza sativa* L.). *Indian Journal of Animal Nutrition* 7: 2117–2122.
48. Ekanayake IJ, Garrity DP, Masajo TM, O'Toole JC (1985) Root pulling resistance in rice: inheritance and association with drought tolerance. *Euphytica* 34: 905–913.
49. Ekanayake IJ, O'Toole JC, Garrity DP, Masajo TM (1985) Inheritance of root characters and their relations to drought resistance in rice. *Crop Science* 25: 927–933.
50. Xia H, Xiong J, Tao T, Zheng X, Huang W, et al. (2015) Distinguishing upland and lowland rice ecotypes by selective SSRs and their applications in molecular-assisted selection of rice drought resistance. *Euphytica* 206: 11–20.
51. Zhang H, Zhang D, Wang M, Sun J, Qi Y, et al. (2011) A core collection and mini core collection of *Oryza sativa* L. in China. *Theoretical and Applied Genetics* 122: 49–61. <https://doi.org/10.1007/s00122-010-1421-7> PMID: 20717799
52. Yu SB, Xu WJ, Vijayakumar CHM, Ali J, Fu BY, et al. (2003) Molecular diversity and multilocus organization of the parental lines used in the International Rice Molecular Breeding Program. *Theoretical and Applied Genetics* 108: 131–140. <https://doi.org/10.1007/s00122-003-1400-3> PMID: 13679990
53. Huang X, Wei X, Sang T, Zhao Q, Feng Q, et al. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics* 42: 961–967. <https://doi.org/10.1038/ng.695> PMID: 20972439
54. Huang X, Zhao Y, Wei X, Li C, Wang A, et al. (2012) Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics* 44: 32–39.
55. Horii H, Nemoto K, Miyamoto N, Harada J (2006) Quantitative trait loci for adventitious and lateral roots in rice. *Plant Breeding* 125: 198–200.
56. Price AH, Steele KA, Moore BJ, Jones RGW (2002) Upland rice grown in soil-filled chambers and exposed to contrasting water-deficit regimes: II. Mapping quantitative trait loci for root morphology and distribution. *Field Crops Research* 76: 25–43.
57. Yadav R, Courtois B, Huang N, McLaren G (1997) Mapping genes controlling root morphology and root distribution in a doubled-haploid population of rice. *Theoretical and Applied Genetics* 94: 619–632.

58. MacMillan K, Emrich K, Piepho HP, Mullins CE, Price AH (2006) Assessing the importance of genotype × environment interaction for root traits in rice using a mapping population II: conventional QTL analysis. *Theoretical and Applied Genetics* 113: 953–964. <https://doi.org/10.1007/s00122-006-0357-4> PMID: 16896715
59. Hemamalini GS, Shashidhar HE, Hittalmani S (2000) Molecular marker assisted tagging of morphological and physiological traits under two contrasting moisture regimes at peak vegetative stage in rice (*Oryza sativa* L.). *Euphytica* 112: 69–78.
60. Yue B, Xue W, Xiong L, Yu X, Luo L, et al. (2006) Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. *Genetics* 172: 1213–1228. <https://doi.org/10.1534/genetics.105.045062> PMID: 16272419
61. Liu L, Mu P, Li X, Qu Y, Wang Y, et al. (2008) Localization of QTL for basal root thickness in japonica rice and effect of marker-assisted selection for a major QTL. *Euphytica* 164: 729–737.
62. Domingo C, Andres F, Tharreau D, Iglesias DJ, Talon M (2009) Constitutive expression of OsGH3.1 reduces auxin content and enhances defense response and resistance to a fungal pathogen in rice. *Molecular Plant Microbe Interactions* 22: 201–210. <https://doi.org/10.1094/MPMI-22-2-0201> PMID: 19132872
63. Zhang S, Li C, Cao J, Zhang Y, Zhang S, et al. (2009) Altered architecture and enhanced drought tolerance in rice via the down-regulation of indole-3-acetic acid by TLD1/OsGH3.13 activation. *Plant Physiology* 151: 1889–1901. <https://doi.org/10.1104/pp.109.146803> PMID: 19776160
64. Zhao S, Xiang J, Xue H (2013) Studies on the rice LEAF INCLINATION1 (LC1), an IAA-amido synthetase, reveal the effects of auxin in leaf inclination control. *Molecular Plant* 6: 174–187. <https://doi.org/10.1093/mp/sss064> PMID: 22888153
65. Zhao Y, Hu Y, Dai M, Huang L, Zhou D (2009) The WUSCHEL-related homeobox gene WOX11 is required to activate shoot-borne crown root development in rice. *Plant Cell* 21: 736–748. <https://doi.org/10.1105/tpc.108.061655> PMID: 19258439
66. Zhao Y, Cheng S, Song Y, Huang Y, Zhou S, et al. (2015) The interaction between rice ERF3 and WOX11 promotes crown root development by regulating gene expression involved in cytokinin signaling. *Plant Cell* 27: 2469–2483. <https://doi.org/10.1105/tpc.15.00227> PMID: 26307379
67. Nakagawa Y, Hanaoka H, Kobayashi M, Miyoshi K, Miwa K, et al. (2007) Cell-type specificity of the expression of OsBOR1, a rice efflux boron transporter gene, is regulated in response to boron availability for efficient boron uptake and xylem loading. *Plant Cell* 19: 2624–2635. <https://doi.org/10.1105/tpc.106.049015> PMID: 17675406
68. Fan X, Dan X, Chen J, Lu H, Xu Y, et al. (2014) Over-expression of OsPTR6 in rice increased plant growth at different nitrogen supplies but decreased nitrogen use efficiency at high ammonium supply. *Plant Science* 227: 1–11. <https://doi.org/10.1016/j.plantsci.2014.05.013> PMID: 25219300
69. Liu J, Liang D, Song Y, Xiong L (2010) Systematic identification and expression analysis of BREVIS RADIX-like homologous genes in rice. *Plant Science* 178: 183–191.
70. Toda Y, Tanaka M, Ogawa D, Kurata K, Kurotani K-i, et al. (2013) RICE SALT SENSITIVE3 forms a ternary complex with JAZ and class-C bHLH factors and regulates jasmonate-induced gene expression and root cell elongation. *Plant Cell* 25: 1709–1725. <https://doi.org/10.1105/tpc.113.112052> PMID: 23715469
71. Park HC, Kim H, Koo SC, Park HJ, Cheong MS, et al. (2010) Functional characterization of the SIZ/PIAS-type SUMO E3 ligases, OsSIZ1 and OsSIZ2 in rice. *Plant Cell Environment* 33: 1923–1934.
72. Wang H, Makeen K, Yan Y, Cao Y, Sun S, et al. (2011) OsSIZ1 regulates the vegetative growth and reproductive development in rice. *Plant Molecular Biology Reporter* 29: 411–417.
73. Miura K, Lee J, Gong Q, Ma S, Jin JB, et al. (2011) SIZ1 regulation of phosphate starvation-induced root architecture remodeling involves the control of auxin accumulation. *Plant Physiology* 155: 1000–1012. <https://doi.org/10.1104/pp.110.165191> PMID: 21156857
74. Atsushi M, Yusuke H, OS E., Godfrey A, Tatsushi T, et al. (2014) Rice green revolution and climatic change in East Africa: an approach from the technical efficiency of rainfed rice farmers in Uganda. *Agricultural Sciences* 5: 330–341.
75. Haneishi Y, Okello SE, Asea G, Tsuboi T, Maruyama A, et al. (2013) Exploration of rainfed rice farming in Uganda based on a nationwide survey: evolution, regionality, farmers and land. *African Journal of Agricultural Research* 8: 3318–3329.
76. Liu HB, Gou Y, Wang HY, Li HM, Wu W (2014) Temporal changes in climatic variables and their impact on crop yields in southwestern China. *International Journal of Biometeorology* 58: 1021–1030. <https://doi.org/10.1007/s00484-013-0686-3> PMID: 23736776
77. Pandey S, Bhandari H, Ding S, Prapertchob P, Sharan R, et al. (2007) Coping with drought in rice farming in Asia: insights from a cross-country comparative study. *Agricultural Economics* 37: 213–224.

78. Atlin GN, Lafitte HR, Tao D, Laza A, Amante A, et al. (2006) Developing rice cultivars for high-fertility upland systems in the Asian tropics. *Field Crops Research* 97: 43–52.
79. Courtois B, Lafitte HR (1999) Improving rice for drought-prone upland environments. *Genetic Improvement of Rice for Water-limited Environments*: 35–56.
80. Hu H, Xiong L (2014) Genetic engineering and breeding of drought-resistant crops. *Annual Review of Plant Biology* 65: 715–741. <https://doi.org/10.1146/annurev-arplant-050213-040000> PMID: [24313844](https://pubmed.ncbi.nlm.nih.gov/24313844/)
81. Gardner KM, Latta RG (2007) Shared quantitative trait loci underlying the genetic correlation between continuous traits. *Molecular Ecology* 16: 4195–4209. <https://doi.org/10.1111/j.1365-294X.2007.03499.x> PMID: [17850272](https://pubmed.ncbi.nlm.nih.gov/17850272/)
82. Li Z, Fu B, Gao Y, Xu J, Ali J, et al. (2005) Genome-wide introgression lines and their use in genetic and molecular dissection of complex phenotypes in rice (*Oryza sativa* L.). *Plant Molecular Biology* 59: 33–52. <https://doi.org/10.1007/s11103-005-8519-3> PMID: [16217600](https://pubmed.ncbi.nlm.nih.gov/16217600/)
83. Li JY, Wang J, Zeigler RS (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience* 3: 8. <https://doi.org/10.1186/2047-217X-3-8> PMID: [24872878](https://pubmed.ncbi.nlm.nih.gov/24872878/)
84. Li Z, Fu B, Gao Y, Wang W, Xu J, et al. (2014) The 3,000 rice genomes project. *Gigascience* 3: 7. <https://doi.org/10.1186/2047-217X-3-7> PMID: [24872877](https://pubmed.ncbi.nlm.nih.gov/24872877/)
85. Zhang Q, Li J, Xue Y, Han B, Deng XW (2008) Rice 2020: a call for an international coordinated effort in rice functional genomics. *Molecular Plant* 1: 1–5.
86. Zhang D, Zhang H, Wei X, Qi Y, Wang M, et al. (2007) Genetic structure and diversity of *Oryza sativa* L. in Guizhou, China. *Chinese Science Bulletin* 52: 343–351.
87. Yoshida S, Forno DA, Cock JH, Gomez KA (1976) Laboratory manual for physiological studies of rice: The International Rice Research Institute.
88. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: a Tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559–575. <https://doi.org/10.1086/519795> PMID: [17701901](https://pubmed.ncbi.nlm.nih.gov/17701901/)
89. Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* 19: 1655–1664. <https://doi.org/10.1101/gr.094052.109> PMID: [19648217](https://pubmed.ncbi.nlm.nih.gov/19648217/)
90. Li H, Coghlan A, Ruan J, Coin LJ, Hériché JK, et al. (2006) TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Research* 34: D572. <https://doi.org/10.1093/nar/gkj118> PMID: [16381935](https://pubmed.ncbi.nlm.nih.gov/16381935/)
91. Nei M (1987) *Molecular evolutionary genetics*: Columbia University Press.