RESEARCH   **Open Access**

OGRO: The Overview of functionally characterized Genes in Rice online database

Eiji Yamamoto[1], Jun-ichi Yonemaru[1*], Toshio Yamamoto[1] and Masahiro Yano[2]

Abstract

Background: The high-quality sequence information and rich bioinformatics tools available for rice have contributed to remarkable advances in functional genomics. To facilitate the application of gene function information to the study of natural variation in rice, we comprehensively searched for articles related to rice functional genomics and extracted information on functionally characterized genes.

Results: As of 31 March 2012, 702 functionally characterized genes were annotated. This number represents about 1.6% of the predicted loci in the Rice Annotation Project Database. The compiled gene information is organized to facilitate direct comparisons with quantitative trait locus (QTL) information in the Q-TARO database. Comparison of genomic locations between functionally characterized genes and the QTLs revealed that QTL clusters were often co-localized with high-density gene regions, and that the genes associated with the QTLs in these clusters were different genes, suggesting that these QTL clusters are likely to be explained by tightly linked but distinct genes. Information on the functionally characterized genes compiled during this study is now available in the Overview of Functionally Characterized Genes in Rice Online database (OGRO) on the Q-TARO website (http://qtaro.abr.affrc.go.jp/ogro). The database has two interfaces: a table containing gene information, and a genome viewer that allows users to compare the locations of QTLs and functionally characterized genes.

Conclusions: OGRO on Q-TARO will facilitate a candidate-gene approach to identifying the genes responsible for QTLs. Because the QTL descriptions in Q-TARO contain information on agronomic traits, such comparisons will also facilitate the annotation of functionally characterized genes in terms of their effects on traits important for rice breeding. The increasing amount of information on rice gene function being generated from mutant panels and other types of studies will make the OGRO database even more valuable in the future.

Keywords: Rice (*Oryza sativa* L), Functionally characterized genes, QTL, Database

Background

Rice is a model plant species for which many genetic and genomic resources have been developed. These resources include high-quality genome sequence information (Goff et al. 2002; Yu et al. 2002; International Rice Genome Sequencing Project 2005), high-efficiency transformation systems (Hiei and Komari 2008), bioinformatics tools and databases (reviewed by Nagamura and Antonio 2010), mutant panels (Chern et al. 2007; Miyao et al. 2007), and publicly available populations for genetic analysis such as backcross inbred lines (BILs) and chromosome segment substitution lines (CSSLs)

(Fukuoka et al. 2010). These resources have contributed to remarkable advances in rice functional genomics during the last two decades, and many genes have been functionally characterized (Jiang et al. 2011). Because rice is an important food crop as well as a model plant, information derived from functional genomics research needs to be applied to rice breeding.

So far, most of the genomics research that has been applied to rice breeding has been related to quantitative trait locus (QTL) analysis, because, in many cases, agronomically useful alleles represent naturally occurring allelic variations that were identified as QTLs in cultivars, landraces, or wild species (Yamamoto et al. 2009; Xing and Zhang 2010; Miura et al. 2011). Information on rice QTLs from published articles has been compiled and is publicly available in the Gramene-QTL database (Ni

* Correspondence: yonemaru@affrc.go.jp
[1]National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan
Full list of author information is available at the end of the article

et al. 2009); http://www.gramene.org/qtl/) and the QTL Annotation Rice Online database (Q-TARO; Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/). Several of the genes responsible for QTLs have been cloned, but most have not yet been identified. Mapped QTL regions are often long enough to contain many genes, and introgression of such QTL regions may result in linkage drag, which results from the introgression of one or more unfavorable genes that are closely linked to the genes responsible for the target QTL. In cases where a QTL has been fine-mapped or the causal gene(s) have been identified, the problem of linkage drag can be overcome by means of marker-assisted selection of recombinants between the target gene or QTL and nearby unfavorable genes (Fukuoka et al. 2009).

With the exception of genes that have been identified as those responsible for QTLs, most of the functionally characterized genes in rice have not been analyzed for allelic variation and functional differences in natural populations. However, such information is useful for QTL cloning using the candidate gene approach and for candidate gene association studies (Ehrenreich et al. 2009; Emanuelli et al. 2010). For these approaches, it is necessary to make the list of candidate genes involved in the trait of interest readily available for individual experimental design. It is also important that the genomic locations of functionally characterized genes can be readily compared with the location of QTLs involved in the same trait. Rice databases such as Gramene (Youens-Clark et al. 2011) and Oryzabase (Kurata and Yamazaki 2006) include information on gene function from published research. However, it is necessary to rearrange the data provided by these databases for carrying out the abovementioned approaches. We also found that several functionally characterized genes are not included in those databases, probably because information on such genes was published in agronomy and breeding journals rather than in genetics, genomics, or molecular biology journals.

In this study, our goal was to facilitate the application of gene function information to the study of natural variation in rice. To accomplish this, we comprehensively searched for articles related to rice functional genomics and established a list of functionally characterized genes. Information on each gene was summarized to facilitate direct comparison with QTL information from Q-TARO (Yonemaru et al. 2010). We also compared the genomic locations of functionally characterized genes and QTLs. The information on functionally characterized genes obtained in this study was compiled in a new database, the Overview of Functionally Characterized Genes in Rice Online database (OGRO), which is located on the Q-TARO website (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/ogro).

Results and discussion
Extraction of information on functionally characterized genes in rice

To establish the list of functionally characterized genes in rice, we conducted a comprehensive search for articles related to rice functional genomics, and we extracted information on gene function by manually checking every article identified in the search. As of 31 March 2012, 702 functionally characterized genes were annotated based on the information from 707 articles. The categories of information extracted for each of the functionally characterized genes are listed in Table 1. The list of functionally characterized genes includes seven microRNAs (miRNAs) that have been associated with specific phenotypes (Xie et al. 2006; Zhu et al.

Table 1 Information on functionally characterized genes extracted from each article

Gene information item	Remarks
Gene	Unabbreviated gene name
Gene symbol	Abbreviated gene name
Major category	Corresponds to the criteria used in Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/)
Category of objective character	Corresponds to the criteria used in Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/)
Chr	Chromosome number
Genome start	Corresponds to IRGSP pseudomolecules build 4 (http://rgp.dna.affrc.go.jp/E/IRGSP/Build4/build4.html)
Genome end	Corresponds to IRGSP pseudomolecules build 4 (http://rgp.dna.affrc.go.jp/E/IRGSP/Build4/build4.html)
Locus ID	RAP locus (Rice Annotation Project 2008; http://rapdb.dna.affrc.go.jp/), MSU *Osa1* rice locus (Yuan et al. 2005; http://rice.plantbiology.msu.edu/), osa-miRNA ID (Griffiths-Jones et al. 2008; http://www.mirbase.org/), or GenBank (http://www.ncbi.nlm.nih.gov/genbank/) accession number
Method of isolation	The term "natural variation" was used for genes functionally characterized by using cultivars, landraces, or wild relatives. The term "knockdown/overexpression" indicates that the genes were characterized using both knockdown and overexpression transgenic plants.
Objective character	Phenotypes described in each of the articles
Reference	Identified by the Digital Object Identifier (doi)

2009; Gao et al. 2010; Gao et al. 2011a). Figure 1A shows the genomic distribution of the 702 functionally characterized genes. Among these, four genes were absent from the reference genome sequence (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare): *qSW5/GW5* for grain size (Shomura et al. 2008; Weng et al. 2008), *Sub1A* for submergence tolerance (Xu et al. 2006), and *SK1* and *SK2* for internode elongation in floating rice (Hattori et al. 2009). Genomic regions with few to no functionally characterized genes generally corresponded to heterochromatic regions (Figure 1A; Cheng et al. 2001; Li et al. 2008).

There are 44 755 gene loci, excluding transposable elements (TEs) and ribosomal protein or tRNA loci, in RAP (Rice Annotation Project 2008; http://rapdb.dna.affrc.go.jp/), and 491 miRNA loci in release 18 miRbase (Griffiths-Jones et al. 2008; http://www.mirbase.org/). The functionally characterized genes compiled during this study represent only 1.6% of these loci. In Arabidopsis, a model dicot species, 5826 genes have been functionally characterized, accounting for more than 20% of the gene loci in this species (Lamesch et al. 2012). Considering both the number and the proportion of functionally characterized genes in Arabidopsis, it seems that the functional characterization of rice genes is far from complete.

For the gene information item "method of isolation" (Table 1), the genes identified by using cultivars, landraces, or wild relatives were described as "natural variation". Among the 702 functionally characterized genes, 11% (80 genes) had been identified through natural variation. Another 41% (286 genes) were identified by mutant analysis, and 48% (336) were identified by using transgenic plants (isolation method classified as "overexpression", "knockdown", "knockdown/overexpression", or "others"; Figure 1B). This breakdown indicates that both forward- and reverse-genetics approaches are valuable methods in rice functional genomics.
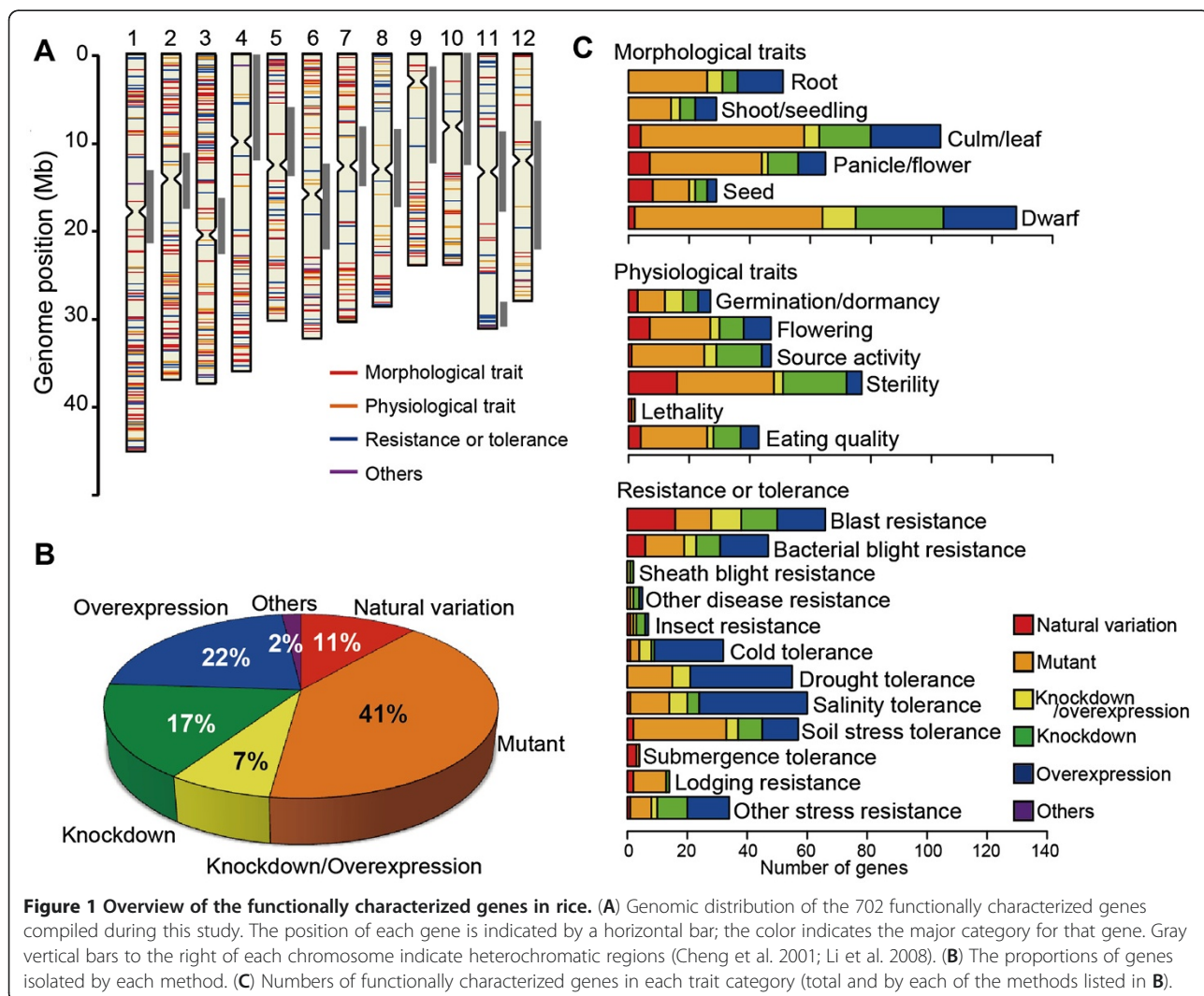

Figure 1 Overview of the functionally characterized genes in rice. (**A**) Genomic distribution of the 702 functionally characterized genes compiled during this study. The position of each gene is indicated by a horizontal bar; the color indicates the major category for that gene. Gray vertical bars to the right of each chromosome indicate heterochromatic regions (Cheng et al. 2001; Li et al. 2008). (**B**) The proportions of genes isolated by each method. (**C**) Numbers of functionally characterized genes in each trait category (total and by each of the methods listed in **B**).

We annotated the functionally characterized genes based on the phenotypes described in each of the articles (Table 1). The phenotypes related to each gene were classified into "major category" and "category of objective character" (Table 1). These categories are identical to those used in Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/). Genes associated with multiple traits were counted within each relevant category.
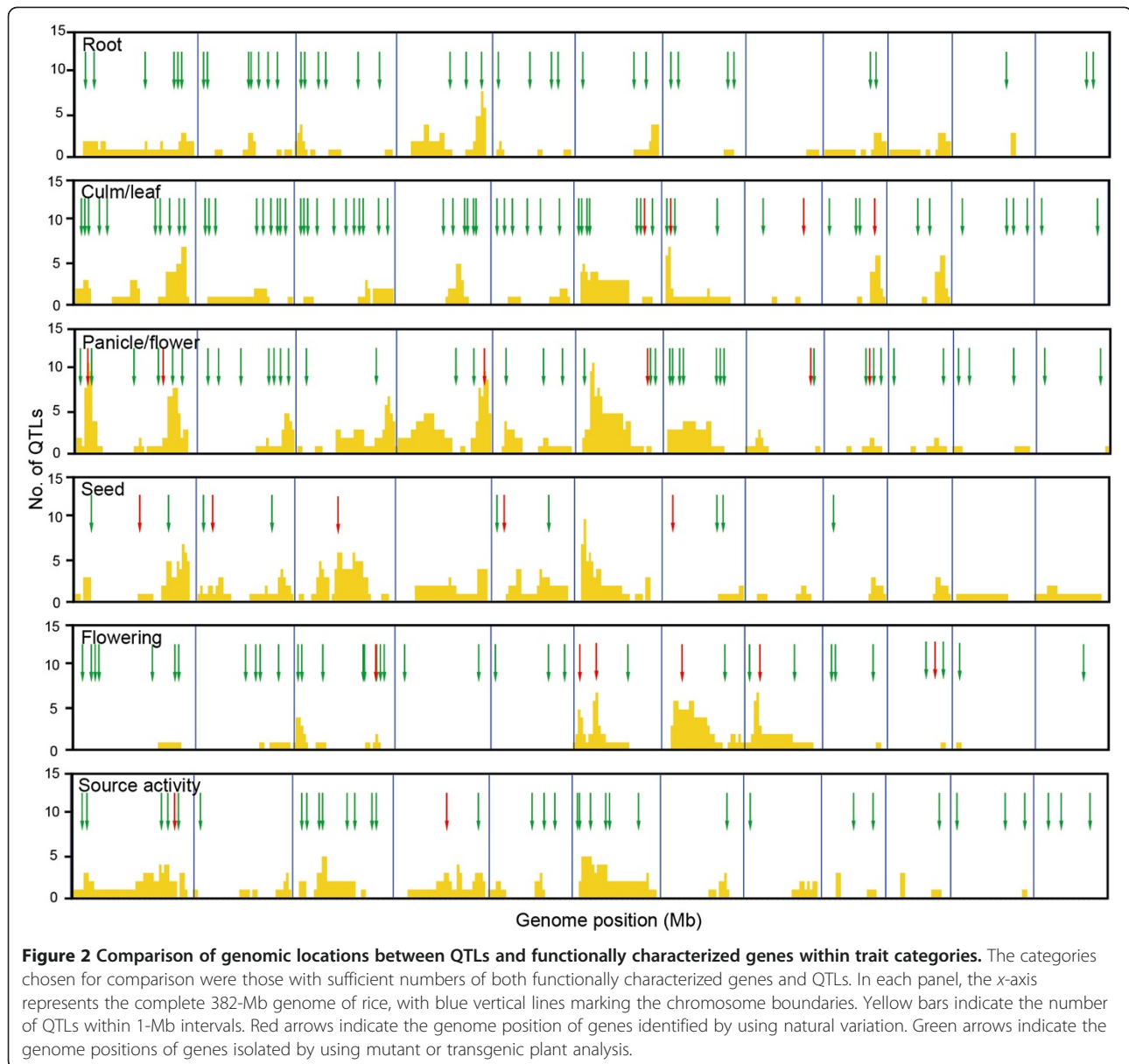
The number of functionally characterized genes within each category is shown in Figure 1C. The variability in the number of functionally characterized genes among the different categories (Figure 1C) probably reflects the agronomic importance of each trait and the interests of individual researchers rather than the actual number of genes involved in each trait. In the major category "resistance or tolerance", transgenic approaches ("overexpression", "knockdown", and "knockdown/overexpression") were used for functional analysis more frequently than for genes in the major categories "morphological trait" and "physiological trait" (Figure 1C). This difference might be due to the difficulty in screening mutant and natural populations for traits related to resistance or tolerance. Within the major category "resistance or tolerance", most of the genes in the categories "cold", "drought", and "salinity" were characterized by overexpression analysis (Figure 1C). The overexpressing plants often showed pleiotropic effects such as growth retardation (Abbasi et al. 2004; Ye et al. 2009; Nakashima et al. 2007), suggesting that complex mechanisms control these abiotic stress tolerances in rice.

Comparison of genomic locations between functionally characterized genes and QTLs

QTL analysis has been used extensively in rice to identify the chromosomal locations and phenotypic contributions of QTLs, and this information has been compiled in two databases, Gramene-QTL database (Ni et al. 2009; http://www.gramene.org/qtl/) and Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/). The gene function information gathered in the present study was arranged to enable direct comparison with the QTL information (trait classification and genomic location) in Q-TARO (Table 1). We performed a genome-wide comparison of the genomic locations of functionally characterized genes in several trait categories with those of the QTLs in Q-TARO (Figure 2). Because most of the genes functionally characterized by using natural variation were identified as genes responsible for QTLs, it was not unexpected to find that most of their genomic locations were included in QTL regions associated with those same traits (Figure 2). Moreover, the genomic locations of functionally characterized genes identified by mutant and transgenic plant analysis also showed good

correspondence with QTL locations (Figure 2), indicating that these genes are good candidates for the genes responsible for the QTLs. The increasing amount of information on rice gene function will make the candidate gene approach for identifying genes responsible for QTLs even more powerful.

Many QTLs tend to be co-localized in specific genomic regions (QTL clusters) even though they control different traits (Yonemaru et al. 2010; Zhao et al. 2011). To survey whether functionally characterized genes were also arranged in such clusters, we calculated the distribution of functionally characterized genes and compared it with the genomic locations of the QTL clusters (Figure 3). In this comparison, we also included the gene density of RAP loci (Rice Annotation Project et al. 2008; http://rapdb.dna.affrc.go.jp/). There was good correspondence between the genomic locations of functionally characterized genes and RAP locus gene density (Figure 3). Furthermore, functionally characterized genes and QTLs also showed high co-localization (Figure 3), indicating that QTLs tended be located in regions of high gene density. Regarding the genetic basis of the QTL clusters, two main possibilities are generally considered: the pleiotropic effects of one or a few genes, or the effects of multiple genes that are tightly linked to one another. Several genes responsible for QTLs have been reported to have pleiotropic effects; for example, *SCM2* is involved in panicle architecture, culm length, and culm mechanical strength (Ookawa et al. 2010), and *IPA/WFP* is involved in panicle architecture, panicle number, and culm mechanical strength (Jiao et al. 2010; Miura et al. 2010). However, when we examined the genomic location of QTL clusters and genes identified by using natural variation, we found that the QTL clusters often contained multiple genes identified by using natural variation (Figure 3). For example, on the long arm of chromosome 1, which contains the largest QTL cluster region, there were four genes that had been identified by using natural variation: *Pi37* for blast resistance (Lin et al. 2007), *qSH1* for seed shattering (Konishi et al. 2006), *qNPQ1-2* for photosynthetic capacity (Kasajima et al. 2011), and *sd1* for culm length (Sasaki et al. 2002). On the short arm of chromosome 6, the location of the second-largest QTL cluster region, there were eight genes that had been identified by using natural variation: *wx* (Wang et al. 1995) and *alk* (Gao et al. 2011b) for eating quality, *Hd3a* (Kojima et al. 2002) and *Hd1* (Yano et al. 2000) for heading date, *DPL2* (Mizuta et al. 2010) and *S5* (Chen et al. 2008) for sterility, and *Pi2/Pi9* (Zhou et al. 2006) and *Pi25/Pid3* (Qu et al. 2005) for blast resistance. Although the genes responsible for most QTLs are still unidentified, considering these examples along with the data showing co-localization of QTL clusters and high-density gene regions (Figure 3) suggests that many QTL clusters are caused by groups of distinct but tightly linked genes controlling different traits.
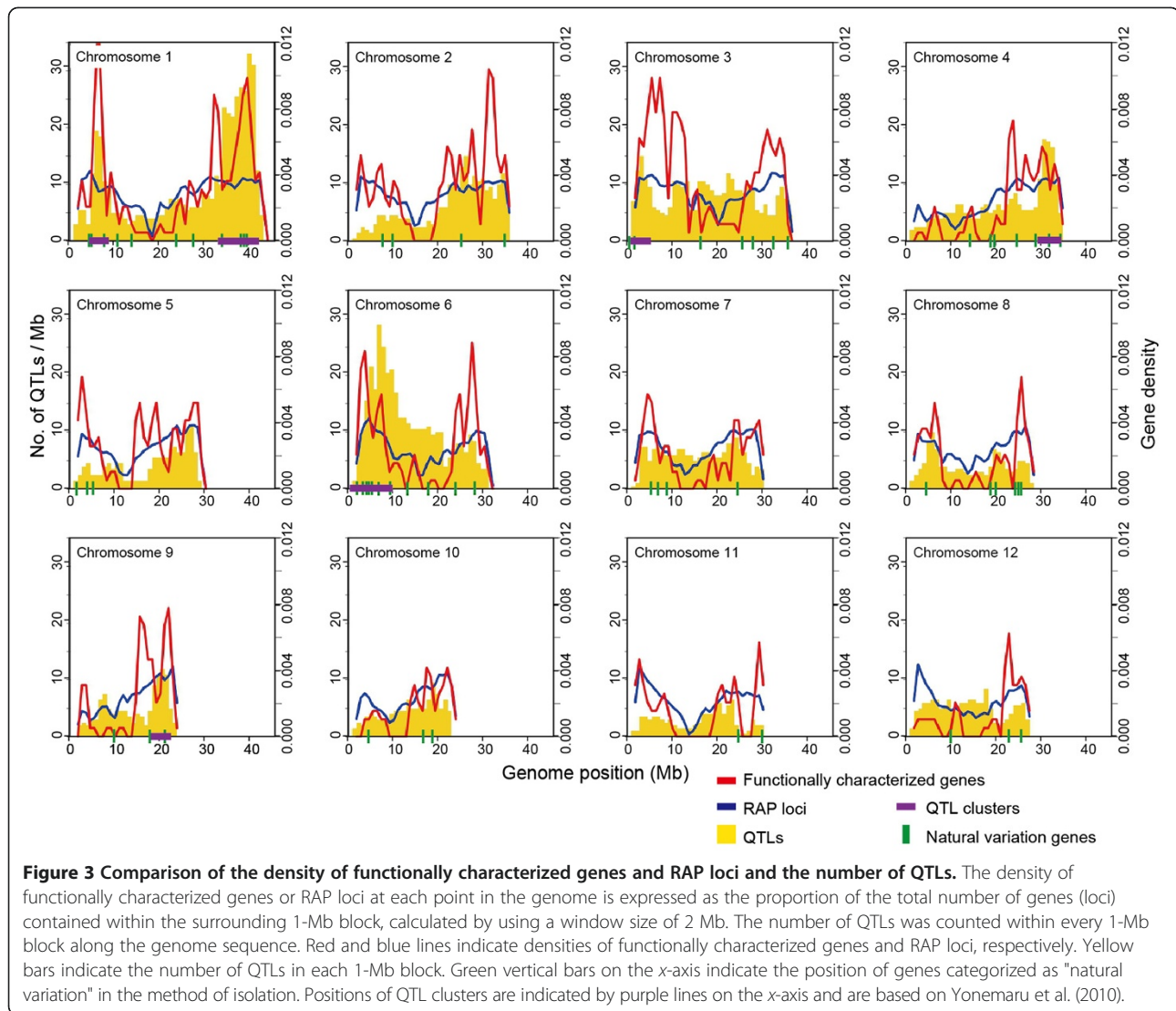
Figure 2 Comparison of genomic locations between QTLs and functionally characterized genes within trait categories. The categories chosen for comparison were those with sufficient numbers of both functionally characterized genes and QTLs. In each panel, the *x*-axis represents the complete 382-Mb genome of rice, with blue vertical lines marking the chromosome boundaries. Yellow bars indicate the number of QTLs within 1-Mb intervals. Red arrows indicate the genome position of genes identified by using natural variation. Green arrows indicate the genome positions of genes isolated by using mutant or transgenic plant analysis.

Understanding the cause of these QTL clusters is important for designing breeding strategies. If QTL clusters contain tightly linked but distinct genes, as in these examples, the problem of linkage drag can be overcome by repeated crossing and careful marker-assisted selection to break the linkage between desirable and undesirable genes and to select the ideal combination of alleles.

Public database of functionally characterized genes in rice

OGRO is available on the Q-TARO website (Figure 4; Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/ogro). Figure 4A shows a screen shot of an information table

from the database. The desired information is displayed by selecting trait categories and chromosome locations from the drop-down boxes, or by specifying search text (Figure 4A, top). To view the genomic location of a target gene or to compare the genomic locations of genes and QTLs, the user can either specify the genomic location in the genome viewer or zoom in graphically (Figure 4B). By default, the loci are grouped by trait category, allowing the locations of genes and QTLs for each trait to be easily compared. Dragging and dropping of the horizontal information bars at the top of the window facilitates the comparison of genes and QTLs in different trait categories.

Although recent advances in next-generation sequencing technologies have enabled re-sequencing of a large

Figure 3 Comparison of the density of functionally characterized genes and RAP loci and the number of QTLs. The density of functionally characterized genes or RAP loci at each point in the genome is expressed as the proportion of the total number of genes (loci) contained within the surrounding 1-Mb block, calculated by using a window size of 2 Mb. The number of QTLs was counted within every 1-Mb block along the genome sequence. Red and blue lines indicate densities of functionally characterized genes and RAP loci, respectively. Yellow bars indicate the number of QTLs in each 1-Mb block. Green vertical bars on the *x*-axis indicate the position of genes categorized as "natural variation" in the method of isolation. Positions of QTL clusters are indicated by purple lines on the *x*-axis and are based on Yonemaru et al. (2010).

number of rice genomes (Xu et al. 2011) as well as high-throughput genotyping and large-scale genetic variation surveys (McNally et al. 2009; Ebana et al. 2010; McCouch et al. 2010; Nagasaki et al. 2010; Yamamoto et al. 2010), analysis of gene function is still indispensable both for understanding fundamental phenomena and for genomics-based breeding. Increasing numbers of mutant panels have been developed in rice, and their comprehensive analysis is ongoing (Chern et al. 2007). These experiments will provide additional information on gene function, which will be added to the database as it becomes available.

Conclusion

In this study, we comprehensively searched for articles related to rice functional genomics and extracted information on 702 functionally characterized genes (Figure 1). The information on each gene was organized

to enable direct comparison with the QTL information in Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc. go.jp/), which will facilitate a candidate-gene approach to identifying the genes responsible for QTLs (Figure 2). Because the QTL descriptions in Q-TARO contain information on agronomic traits, such comparisons will also facilitate the annotation of functionally characterized genes in terms of their effects on traits important for rice breeding. We found that the genes responsible for QTLs in QTL clusters were identified as different genes (Figure 3). Considering this evidence along with the data showing co-localization of QTL clusters and high-density gene regions (Figure 3), our results suggest that many QTL clusters are caused by distinct but tightly linked genes. Information on the functionally characterized genes compiled in this study is now available in OGRO on the Q-TARO Web site (Figure 4; http:// qtaro.abr.affrc.go.jp/ogro). The increasing amount of
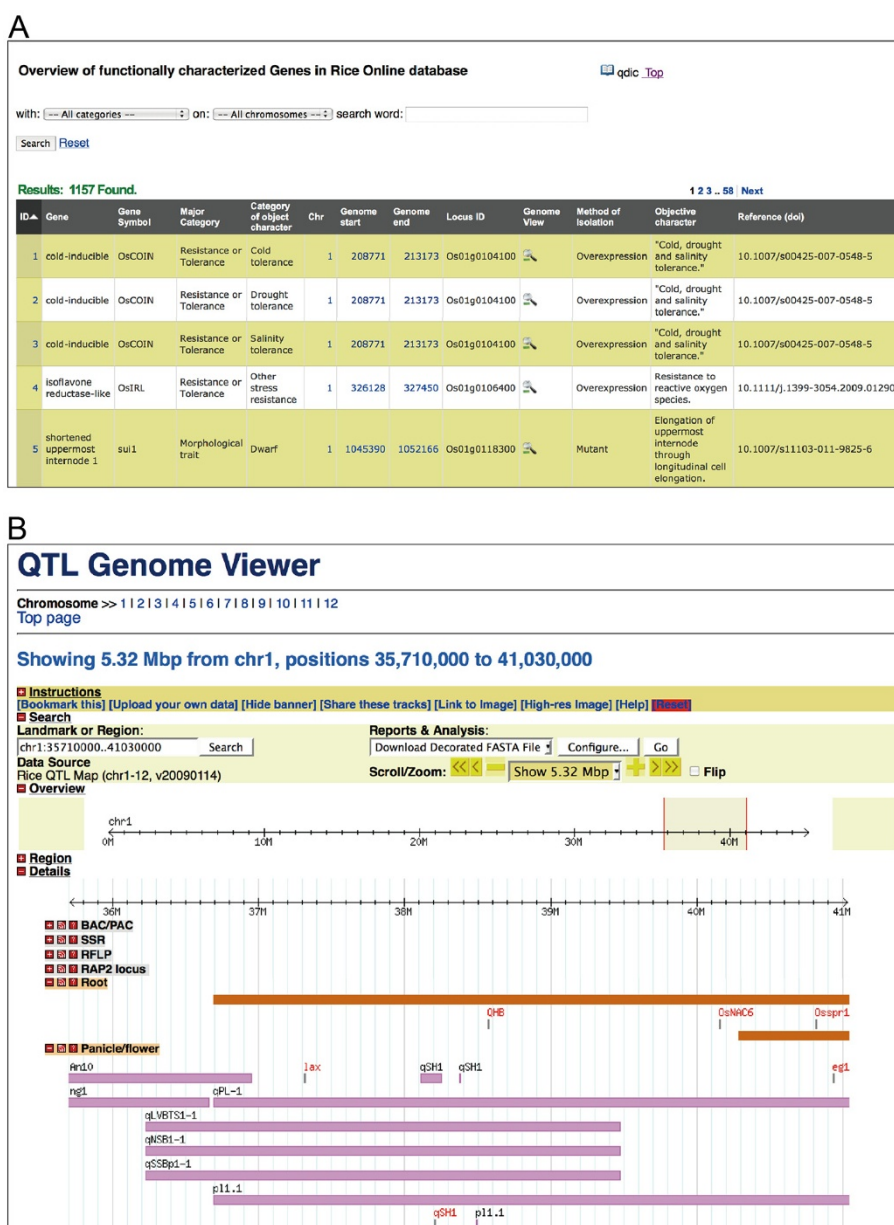
Figure 4 Screen shots of the Overview of Functionally Characterized Genes in Rice Online database (OGRO) (http://qtaro.abr.affrc.go. jp/ogro). (**A**) Gene information table. All displayed information can be exported as comma-separated values (CSV format). (**B**) OGRO genome viewer. This viewer can be used to compare the locations of QTLs with those of functionally characterized genes.

information on rice gene function being generated from mutant panels and other types of studies will make the OGRO database even more valuable in the future.

Methods
Extraction of gene information from published articles
Functional genomics studies have been done using many different approaches, and the degree of functional characterization differs substantially among genes. To avoid ambiguity, we established two main criteria for functionally characterized genes in rice. The first was verification of function: gene function had to be demonstrated in rice through direct evidence based on complementation tests, mutant analysis, or transgenic plant analysis. The second was verification of the phenotype: there had to be evidence that the function of the gene affected the phenotype of the rice plant. Functional analysis using other organisms such as yeast and

Arabidopsis was not counted as meeting this criterion because such experiments do not necessarily indicate that the gene has a biological role in rice.

Articles related to rice functional genomics were identified by searching the Web of Science database (http://apps.webofknowledge.com/) with the search terms "rice" and "Oryza sativa". Because rice studies span a broad range of research fields, the following categories were surveyed: AGRICULTURE MULTIDISCIPLINARY, AGRONOMY, BIOTECHNOLOGY & APPLIED MICROBIOLOGY, CELL BIOLOGY, GENETICS & HEREDITY, MULTIDISCIPLINARY SCIENCES, and PLANT SCIENCES. To make this search comprehensive, the time span was set to "All" (i.e., all publications since 1899). As of 31 March 2012, we identified a total of 14 102 articles using these search conditions. All of the articles were then manually checked, and articles containing information on gene function that met our criteria for functionally characterized genes were selected. The result was a total of 707 articles. For each gene meeting the criteria for a functionally characterized gene, we extracted information including the gene locus ID, genome position, method of isolation, related traits, and reference information (doi) (Table 1). Whenever possible, the RAP ID number (Rice Annotation Project 2008; http://rapdb.dna.affrc.go.jp/) was used as the gene locus ID number. If there was no corresponding ID in RAP, the Michigan State University (MSU) locus number (Yuan et al. 2005; http://rice.plantbiology.msu.edu/) or GenBank (http://www.ncbi.nlm.nih.gov/genbank/) accession number was used. Information on genome position (start and end) was based on International Rice Genome Sequencing Project (IRGSP) Pseudomolecules build 4.0 (http://rgp.dna.affrc.go.jp/E/IRGSP/Build4/build4.html).

The genome positions of genes not found in the reference genome (Oryza sativa L. ssp. japonica cv. Nipponbare) were indicated by using either a position adjacent to the deleted sequence or the positions of the flanking markers used for positional cloning. Under the method of isolation, "knockdown/overexpression" indicates that the genes were characterized by using both knockdown and overexpression transgenic plants.

Comparison of genomic locations and densities between functionally characterized genes and QTLs

We compared the relative genome positions and distributions of functionally characterized genes and QTLs within each of the trait categories. The genome position of each functionally characterized gene was represented by the midpoint between the genome start and genome end positions (Table 1). The QTL information was extracted from Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/).

We also performed comparisons across all of the trait categories between the density of functionally characterized genes, the density of RAP loci and the number of QTLs. The density of functionally characterized genes or RAP loci at each point in the genome was expressed as the proportion of the total number of genes (loci) contained within the surrounding 1-Mb block, calculated by using a window size of 2 Mb. The number of QTLs was counted within every 1-Mb block along the genome sequence.

Database construction

All data on the functionally characterized genes annotated in this study were compiled in OGRO (http://qtaro.abr.affrc.go.jp/ogro). Like Q-TARO (Yonemaru et al. 2010; http://qtaro.abr.affrc.go.jp/), OGRO consists of two Web applications: a gene information table and a genome viewer. The Web applications were implemented as Perl scripts and CGI modules. The database was constructed using MySQL, a relational database management system. We used the GBrowse viewer (http://gmod.org/wiki/Main_Page), which was configured to access OGRO from within the Q-TARO genome viewer.

Author details
[1]National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan. [2]National Institute of Agrobiological Sciences, 1-2 Ohwashi, Tsukuba, Ibaraki 305-8634, Japan.

References
Abbasi F, Onodera H, Toki S, Tanaka H, Komatsu S (2004) OsCDPK13, a calcium-dependent protein kinase gene from rice, is induced by cold and gibberellin in rice leaf sheath. Plant Mol Biol 55:541–552

Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, Zhao J, Qiu S, Zhang X, Yao J, Liu K, Wang L, Xu C, Li X, Xue Y, Xia M, Ji Q, Lu J, Xu M, Zhang Q (2008) A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica-japonica hybrids in rice. Proc Natl Acad Sci USA 105:11436–11441

Cheng Z, Buell CR, Wing RA, Gu M, Jiang J (2001) Toward a cytological characterization of the rice genome. Genome Res 11:2133–2141

Chern CG, Fan MJ, Yu SM, Hour AL, Lu PC, Lin YC, Wei FJ, Huang SC, Chen S, Lai MH, Tseng CS, Yen HM, Jwo WS, Wu CC, Yang TL, Li LS, Kuo YC, Li SM, Li CP, Wey CK, Trisiriroj A, Lee HF, Hsing YI (2007) A rice phenomics study-

phenotype scoring and seed propagation of a T-DNA insertion-induced rice mutant population. Plant Mol Biol 65:427–438

Ebana K, Yonemaru JI, Fukuoka S, Iwata H, Kanamori H, Namiki N, Nagasaki H, Yano M (2010) Genetic structure revealed by a whole-genome single-nucleotide polymorphism survey of diverse accessions of cultivated Asian rice (Oryza sativa L.). Breed Sci 60:390–397

Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD (2009) Candidate gene association mapping of Arabidopsis flowering time. Genetics 183:325–335

Emanuelli F, Battilana J, Costantini L, Le Cunff L, Boursiquot JM, This P, Grando MS (2010) A candidate gene association study on muscat flavor in grapevine (Vitis vinifera L.). BMC Plant Biol 10:241

Fukuoka S, Nonoue Y, Yano M (2010) Germplasm enhancement by developing advanced plant materials from diverse rice accessions. Breed Sci 60:509–517

Fukuoka S, Saka N, Koga H, Ono K, Shimizu T, Ebana K, Hayashi N, Takahashi A, Hirochika H, Okuno K, Yano M (2009) Loss of function of a proline-containing protein confers durable disease resistance in rice. Science 325:998–1001

Gao P, Bai X, Yang L, Lv D, Li Y, Cai H, Ji W, Guo D, Zhu Y (2010) Over-expression of osa-MIR396c decreases salt and alkali stress tolerance. Planta 231: 991–1001

Gao P, Bai X, Yang L, Lv D, Pan X, Li Y, Cai H, Ji W, Chen Q, Zhu Y (2011a) osa-MIR393: a salinity- and alkaline stress-related microRNA gene. Mol Biol Rep 38:237–242

Gao Z, Zeng D, Cheng F, Tian Z, Guo L, Su Y, Yan M, Jiang H, Dong G, Huang Y, Han B, Li J, Qian Q (2011b) ALK, the key gene for gelatinization temperature, is a modifier gene for gel consistency in rice. J Integr Plant Biol 53:756–765

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P et al (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296:92–100

Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36(Database issue):D154–D158

Hattori Y, Nagai K, Furukawa S, Song XJ, Kawano R, Sakakibara H, Wu J, Matsumoto T, Yoshimura A, Kitano H, Matsuoka M, Mori H, Ashikari M (2009) The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. Nature 460:1026–1030

Hiei Y, Komari T (2008) Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. Nat Protoc 3:824–834

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800

Jiang Y, Cai Z, Xie W, Long T, Yu H, Zhang Q (2011) Rice functional genomics research: Progress and implications for crop genetic improvement. Biotechnol Adv. doi:10.1016/j.biotechadv.2011.08.013

Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, Qian Q, Li J (2010) Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nat Genet 42:541–544

Kasajima I, Ebana K, Yamamoto T, Takahara K, Yano M, Kawai-Yamada M, Uchimiya H (2011) Molecular distinction in genetic regulation of nonphotochemical quenching in rice. Proc Natl Acad Sci USA 108:13835–13840

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. Plant Cell Physiol 43:1096–1105

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312: 1392–1396

Kurata N, Yamazaki Y (2006) Oryzabase. An integrated biological and genome information database for rice. Plant Physiol 140:12–17

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 40(Database issue):D1202–1210

Li X, Wang X, He K, Ma Y, Su N, He H, Stolc V, Tongprasit W, Jin W, Jiang J, Terzaghi W, Li S, Deng XW (2008) High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. Plant Cell 20:259–276

Lin F, Chen S, Que Z, Wang L, Liu X, Pan Q (2007) The blast resistance gene Pi37 encodes a nucleotide binding site leucine-rich repeat protein and is a member of a resistance gene cluster on rice chromosome 1. Genetics 177:1871–1880

McCouch SR, Zhao K, Wright M, Tung CW, Ebana K, Thomson M, Reynolds A, Wang D, DeClerck G, Ali Md L, McClung A, Eizenga G, Bustamante C (2010) Development of genome-wide SNP assays for rice. Breed Sci 60:524–535

McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, Stokowski R, Ballinger DG, Frazer KA, Cox DR, Padhukasahasram B, Bustamante CD, Weigel D, Mackill DJ, Bruskiewich RM, Rätsch G, Buell CR, Leung H, Leach JE (2009) Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc Natl Acad Sci USA 106:12273–12278

Miura K, Ashikari M, Matsuoka M (2011) The role of QTLs in the breeding of high-yielding rice. Trends Plant Sci 16:319–326

Miura K, Ikeda M, Matsubara A, Song XJ, Ito M, Asano K, Matsuoka M, Kitano H, Ashikari M (2010) OsSPL14 promotes panicle branching and higher grain productivity in rice. Nat Genet 42:545–549

Miyao A, Iwasaki Y, Kitano H, Itoh J, Maekawa M, Murata K, Yatou O, Nagato Y, Hirochika H (2007) A large-scale collection of phenotypic data describing an insertional mutant population to facilitate functional analysis of rice genes. Plant Mol Biol 63:625–635

Mizuta Y, Harushima Y, Kurata N (2010) Rice pollen hybrid incompatibility caused by reciprocal gene loss of duplicated genes. Proc Natl Acad Sci USA 107:20417–20422

Nagamura Y, Antonio BA (2010) Current status of rice informatics resources and breeding applications. Breed Sci 60:549–555

Nagasaki H, Ebana K, Shibaya T, Yonemaru JI, Yano M (2010) Core single-nucleotide polymorphisms—a tool for genetic analysis of the Japanese rice population. Breed Sci 60:648–655

Nakashima K, Tran LS, Van Nguyen D, Fujita M, Maruyama K, Todaka D, Ito Y, Hayashi N, Shinozaki K, Yamaguchi-Shinozaki K (2007) Functional analysis of a NAC-type transcription factor OsNAC6 involved in abiotic and biotic stress-responsive gene expression in rice. Plant J 51:617–630

Ni J, Pujar A, Youens-Clark K, Yap I, Jaiswal P, Tecle I, Tung CW, Ren L, Spooner W, Wei X, Avraham S, Ware D, Stein L, McCouch S (2009) Gramene QTL database: development, content and applications. Database 2009:bap005

Ookawa T, Hobo T, Yano M, Murata K, Ando T, Miura H, Asano K, Ochiai Y, Ikeda M, Nishitani R, Ebitani T, Ozaki H, Angeles ER, Hirasawa T, Matsuoka M (2010) New approach for rice improvement using a pleiotropic QTL gene for lodging resistance and yield. Nat Commun 1:132. doi:10.1038/ncomms1132

Qu S, Liu G, Zhou B, Bellizzi M, Zeng L, Dai L, Han B, Wang GL (2005) The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site–leucine-rich repeat protein and is a member of a multigene family in rice. Genetics 172:1901–1914

Rice Annotation Project (2008) The Rice Annotation Project Database (RAP-DB): 2008 update. Nucleic Acids Res 36(Database issue):D1028–1033

Sasaki A, Ashikari M, Ueguchi-Tanaka M, Itoh H, Nishimura A, Swapan D, Ishiyama K, Saito T, Kobayashi M, Khush GS, Kitano H, Matsuoka M (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. Nature 416:701–702

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40:1023–1028

Wang ZY, Zheng FQ, Shen GZ, Gao JP, Snustad DP, Li MG, Zhang JL, Hong MM (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. Plant J 7:613–622

Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X, Wang J, Jiang L, Zhai H, Wan J (2008) Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. Cell Res 18:1199–1209

Xie K, Wu C, Xiong L (2006) Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. Plant Physiol 142:280–293

Xing Y, Zhang Q (2010) Genetic and molecular bases of rice yield. Annu Rev Plant Biol 61:421–442

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705–708

Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, Li J, He W, Zhang G, Zheng X, Zhang F, Li Y, Yu C, Kristiansen K, Zhang X, Wang J, Wright M, McCouch S, Nielsen R, Wang J, Wang W (2011) Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat Biotechnol 30:105–111

Yamamoto T, Nagasaki H, Yonemaru J, Ebana K, Nakajima M, Shibaya T, Yano M (2010) Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms. BMC Genomics 11:267

Yamamoto T, Yonemaru J, Yano M (2009) Towards the understanding of complex traits in rice: substantially or superficially? DNA Res 16:141–154

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene *CONSTANS*. Plant Cell 12:2473–2484

Ye H, Du H, Tang N, Li X, Xiong L (2009) Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. Plant Mol Biol 71:291–305

Yonemaru JI, Yamamoto T, Fukuoka S, Uga Y, Hori K, Yano M (2010) Q-TARO: QTL Annotation Rice Online Database. Rice 3:194–203

Youens-Clark K, Buckler E, Casstevens T, Chen C, Declerck G, Derwent P, Dharmawardhana P, Jaiswal P, Kersey P, Karthikeyan AS, Lu J, McCouch SR, Ren L, Spooner W, Stein JC, Thomason J, Wei S, Ware D (2011) Gramene database in 2010: updates and extensions. Nucleic Acids Res 39:D1085–1094

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92

Yuan Q, Ouyang S, Wang A, Zhu W, Maiti R, Lin H, Hamilton J, Haas B, Sultana R, Cheung F, Wortman J, Buell CR (2005) The Institute for Genomic Research Osa1 rice genome annotation database. Plant Physiol 138:18–26

Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, Bustamante CD, McCouch SR (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nat Commun 2:467. doi:10.1038/ncomms1467

Zhou B, Qu S, Liu G, Dolan M, Sakai H, Lu G, Bellizzi M, Wang GL (2006) The eight amino-acid differences within three leucine-rich repeats between Pi2 and Piz-t resistance proteins determine the resistance specificity to *Magnaporthe grisea*. Mol Plant Microbe Interact 19:1216–1228

Zhu QH, Upadhyaya NM, Gubler F, Helliwell CA (2009) Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*). BMC Plant Biol 9:149