Indian J. Genet., **66**(2): 137-139 (2006)

Short Communication

Rice-QTL-Marker Database — A breeder friendly resource tool for easy deployment of markers for selection of complex traits in rice (*Oryza sativa* L.)

Girish T. N. and Shailaja Hittalmani[1]

Department of Genetics and Plant Breeding, University of Agricultural Sciences, Bangalore 560 065

(Received: March 2006; Revised: July 2006; Accepted: July 2006)

Rice was chosen as a model organism for genome sequencing because of its economic and utility value, small genome size, and syntenic relationship with other cereal species. Many genes rather than a single locus often influence several economically important traits, that are known as quantitative trait loci (QTL). Quantitative trait locus is a polymorphic locus which contains alleles that differentially affect the expression of a continuously distributed phenotypic trait. QTL mapping is a useful approach for detection of complex traits in crops, evolution and medicine, molecular dissection of quantitative phenotypes, supplementing the principles of classical quantitative genetics. This has accelerated the progress in the manipulation of plant and animal genomes [1].

Although the concepts for detecting QTL were developed more than 75 years ago [2] the availability of DNA markers and powerful biometric methods has led to considerable progress in QTL mapping in plant science. The most obvious applications of QTL analysis have been in Marker Assisted Selection during breeding and pre-breeding and QTL cloning. The success in both endeavours primarily depends on the reliability and accuracy of the QTL analysis where information has been obtained. For these endeavours to be effectively and efficiently carried out on a large scale it is necessary that the outcome of QTL studies carried out extensively be documented and made available.

A database literally means a collection of information that has been systematically organized for easy access and analysis. Databases typically are computerized. A database may contain bibliographic, textual or numeric data and the data are usually structured so that they may be searched in a number of ways. There are a wide variety of databases [3-6], from simple tables stored in a single file to very large databases with many millions of records, stored in rooms full of disk drives or other peripheral electronic storage devices. Effort has been made to develop one such database on QTLs governing various traits exclusively in rice from the work carried out meticulously by various highly pricised scientific research.

The QTLs governing various traits in rice have been compiled from the published articles in the National and International Journals, Proceedings of the International Symposiums and News Letters. Meticulously carried out experiments by postgraduate researchers at the Genetics and Plant Breeding department, University of Agricultural Sciences, Bangalore, is also considered for this application unit. The data is considered authentic and is stored in high user-friendly software Excel with several options for quick search.

Features and application

The Database is simple, structured and exhaustive: Databases are generally classified by programming model associated with them. Rice QTL-Marker Database is a flat file database based on the flat (or table) model which consists of a single, two-dimensional array of data elements, where all members of a given column are assumed to be similar values, and all members of a row are assumed to be related to one another. In a flat file database all the information is stored in plain text files, one database record per line. Each record is divided into fields using delimiters or at fixed column positions. The information on QTLs in Rice QTL-Marker Database is presented in a ready to use and easy to understand Microsoft Excel format (Table 1). By being present in the excel format the database extends the access of information to even those who are just acquainted with the basics of computer programmes and program applications The database provides the following types of information on a QTL to the user:

Trait category: Each trait belongs to a parent trait

---

[1] Corresponding author: shailajah_maslab@rediffmail.com

Table 1. Example of yield and yield related traits in flat-file format of Rice-QTL-Marker-Database

Sl. No.	Trait	Chr#	Marker Interval	Authors	References	LOD	%Var/R^2
1	Number of grains per panicle	4	RM226 - C513	He *et al.*, 2001	Crop Science, 41:1240-1246	2.05	12.90
2	Grain weight	3	RM55 - RM203	Hua *et al.*, 2002	Genetics, 162: 18851895	4.60	4.40
3	1000 grain weight	3	RM55 - RM49	Hittalmani *et al.*, 2003	TAG, 107:679690	4.03	14.40
4	Grain density	7	RM234 - CDO405	Xu *et al.*, 2004	Plant Breeding, 123:43-50	3.71	6.00
5	Grain length	5	RM163 - RM161	Xu *et al.*, 2004	Plant Breeding, 123:43-50	8.79	10.00
6	Yield per plant	2	RM262 - RM183	Marri *et al.*, 2005	BMC Genetics, 6:33-33	4.35	14.20
7	Harvest index	2	RM183 - RM263	Marri *et al.*, 2005	BMC Genetics, 6:33-33	3.12	5.83
8	Panicle length	9	RM242 - RM205	Marri *et al.*, 2005	BMC Genetics, 6 :33-33	9.48	20.85
9	Number of panicles per plant	2	RM262 - RM183	Marri *et al.*, 2005	BMC Genetics, 6:33-33	2.50	6.80
10	Spikelet fertility	5	RM13 - RM405	Song *et al.*, 2005	TAG, 110: 205-2111	11.96	25.22

category. These are assigned by the agronomic importance.

Abiotic stress: Traits related to tolerance to stresses from abiotic environment, e.g., light, temperature, water, soil or chemicals.

Growth and development: Traits directly measuring plant parts such as root, stem or leaf and those related to its development.

Biotic stress: Traits related to tolerance/resistance against stresses from pathogens and pests.

Grain quality: Traits of economic importance that may affect product quality of rice.

Yield traits: Traits contributing directly to yield based on economic value.

Seed and seedling traits: Traits related to seed characteristics like seed dormancy and seedling vigor like mesocotyl length, α-Amylase activity etc.

Miscellaneous: Other traits like floating ability, lodging resistance etc.

Trait name: Name of the trait for which the QTL was detected.

Linkage group: The linkage group or the chromosome number on which the QTL was mapped.

Marker interval: The markers flanking the QTL region.

Reference: The Authors of the article from which the QTLs have been compiled and its source.

LOD/ LR/prob: The statistical entity which determines the probability of occurrence of the QTL for the corresponding trait on a chromosome.

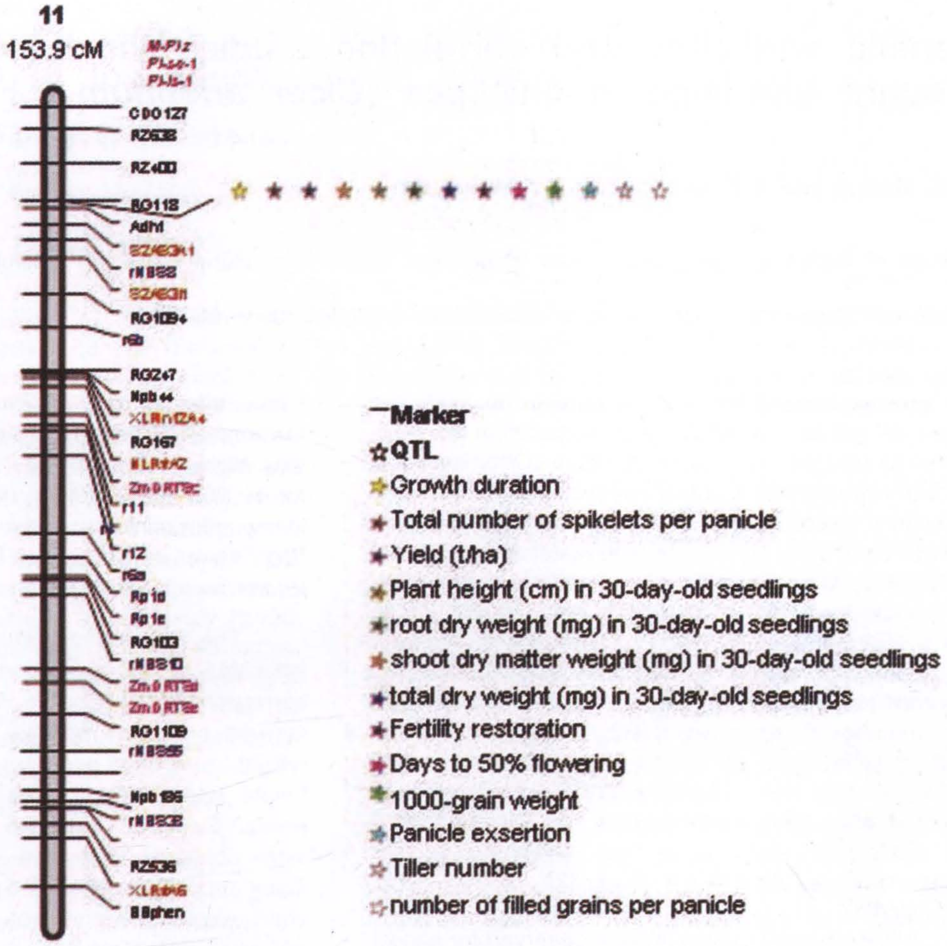
%Var/R2: The contribution of a QTL to the variation in trait expressed as per cent which helps the user to decide whether the QTL is useful, if deployed.

The database presently is compiled with a total of 3371 rice QTLs under 7 broad categories *viz.*, Seed and seedling related traits (101), Growth and Development (842), Yield traits (1206), Grain quality (140), Biotic stress (377), Abiotic stress (526) and Miscellaneous (179). The database is and will be continuously upgraded with information as when data on QTLs will be made available on public domain or any valid and reliable source.

Multiple QTL search/browse option: Each user for the QTLs would be interested in information on QTLs from different perspective. For instance a breeder may be interested in knowing whether QTLs governing a given trait are present on a particular chromosome, while the others may be interested in knowing the QTLs present between a given marker interval. Using the option Find/"Search", the user can browse through the QTL database by trait categories and traits or search chromosome or marker wise. Unlike in other databases comprehensive information on all these above aspects can be obtained in a single display in QTL-Marker database on a single page.

Highly informative: Methods related to QTL interval mapping involve the use of maximum likelihood or non-linear models to simultaneously estimate the position and effect of a QTL within a given interval. In interval mapping adjacent pairs of markers along a chromosome are scanned and likelihood profile of a QTL drawn. The maximum likelihood of a QTL is estimated for each interval. In other words, the log of odds ratio (LOD) of there being a QTL Vs no QTL at a particular point is determined. Those maxima in the profile which exceeds a specified significance level (LOD) indicates the likely sites of QTL. The QTL-Marker Database is highly informative in that in addition to the information the other databases offer, it provides the LOD value at which the QTLs were detected and the per cent contribution of a QTL to the phenotypic variation. From the LOD and %- variation values the breeder could decide upon the QTL to be selected for introgression in to the desired cultivars so as to improve their performance. The values would also help the molecular breeders/scientists in arriving at the QTLs whose fine mapping and cloning would help in a better understanding of their mode of action in terms of expressive units i.e., proteins.

Highly interpretive and analysis permissible: The QTL-Marker Database provides scope for various kinds of analysis being performed on it leading to many interesting revelations. A phenomenon where a single gene or chromosome region exerts influence over the expression of more than one trait is called Pleiotropy. The analysis of database facilitates the identification of pleiotropic chromosomal regions/ QTLs, the chromosome region with a vast pleiotropic effect (Fig 1). One could find clusters of pleiotropic regions on the chromosomes, select those contributing substantially to the traits they are influencing, clone, sequence and characterize them interms of their chemical nature or proteins. It is also possible to find out a trait that is associated with several of the markers or a single marker associated with several traits allowing for multiple trait selection simultaneously. Are QTLs of any of the traits specifically restricted or concentrated on one or few chromosomes or are they distributed randomly across chromosomes? Are any QTLs governing two different traits always associated together (linked) wherever they may be distributed on the chromosomes? An answer to such interesting queries can be obtained by subjecting the database to various kinds of analysis. There are many big databases available across genomes and for many crops like in www.gramene.org, but the one we are presenting is exclusively on rice and it is simple and ready to use use in a single page format.



This database is an effort to present voluminous data on QTLs in a meaningful and simple way so that the knowledge on QTLs is transformed into application. The database is created from nearly 70% of the available literature till date. The remaining and the additional large volume of data that is likely to come out in the very next months/years as QTL mapping in rice is at its helm of-late will be updated accordingly. An effort is in progress to include the traits that might have epistatic effect of the markers pair wise as and when the information will be available. Further the database would be automated thus enhancing its search potential as a source for varied analysis and launched on the Internet for public use. Until then the interested users will be provided the information upon contacting the authors: Shailajah_maslab @rediffmail.com.

Acknowledgement

The authors are thankful to the Rockefeller Foundation, USA, for the financial assistance, in the form of a Project, FS 031 (2002-03) to Shailaja Hittalmani.

References

1. **Paterson A. H.** 1995. Molecular dissection of quantitative traits: Progress and prospects. Genome Research, **5**: 321-333.
2. **Sax K.** 1923. Association of size differences with seed-coat pattern and pigmentation in Phaseolus vulgaris. Genetics, **8**: 552-560.
3. http://www.gramene.org
4. http://www.library.nuigalway.ie/help/jargon/
5. http://www.bambooweb.com
6. http://www.cdc.gov/tobacco/evaluation_manual/ glossary.html