



Rice drought tolerance QTLs: meta-analysis improves resolution to a few candidate gene level

Brigitte Courtois¹, Nourollah Ahmadi¹, Farkhanda Khawaja², Adam Price², Jean-François Rami¹, Julien Frouin¹, Chantal Hamelin¹, Manuel Ruiz¹
¹Cirad Montpellier, France; ²University Aberdeen, UK

Objective

During the last 15 years, a large number of QTLs contributing to drought tolerance in rice have been detected in a broad range of mapping populations. Meta-analyses of QTLs could be extremely helpful in determining the likely number of true QTLs (or meta-QTLs) underlying those QTLs observed in a regions, and narrowing down the meta-QTL confidence intervals and, therefore, the number of underlying candidate genes. This approach has already been successfully used on root traits (Courtois et al. 2009) and is here extended to leaf parameters

Methods

We constructed a web accessible database of all rice drought QTLs (root and leaf traits) published between 1995 and 2007 (<http://tropgeneDB.cirad.fr>) by browsing the literature. The database includes notably information on the experimental conditions of QTL detection, and the physical position of their flanking markers on the pseudo-chromosomes. The database can be queried on these parameters (Figure 1). The QTL list can be exported as an Excel file.

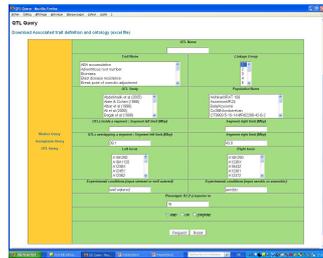


Figure 1: Query form to retrieve QTLs from TropGeneDB

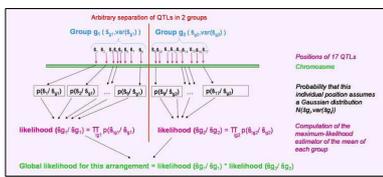
We extracted QTLs for leaf rolling, leaf drying, relative water content, osmotic adjustment and biomass from the database and subjected them to a meta-analysis (Table 1). We compared the QTL and meta-QTL positions and confidence intervals with those of QTLs and meta-QTLs for root parameters (Courtois et al. 2009).

Table 1: List of the studies and QTL included in the meta-analysis

Trait name	No pop.	No QTLs	No studies	References
Leaf rolling	4	85	6	1, 2, 5, 14, 19, 20
Leaf drying	3	48	6	2, 5, 6, 14, 19, 20
Relative water content	3	25	3	2, 11, 14
Osmotic adjustment	3	25	4	11, 13, 16, 21
Biomass	7	87	9	3, 5, 7, 8, 9, 10, 12, 15, 19

Principle of the meta-analysis:

The meta-QTL analysis was run using the software package MetaQTL (Veyrieras et al. 2007). The method relies on a clustering algorithm based on a Gaussian mixture model and enables determination of the likely number of clusters considered as the "true" QTLs underlying the QTLs observed in a given region.



The main parameters needed are the QTL position and confidence interval. To homogenise the data set, we computed the IC in cM of all QTLs using the formula $530/Nr^2$ for BC, F₂ and DH populations (Darvasi and Sollers, 1997), and $163/Nr^2$ for recombinant inbred line populations (Guo et al. 2006) and with N, the population size and r², the percentage of phenotypic variance explained by the QTL. The cM were then converted to Mb using a ratio of 1 cM per 244 kb. The optimal number of clusters is chosen by means of an information based criterion, which should be minimized; the position and confidence interval of the meta-QTLs can, then, be estimated. Among the 5 possible criteria, we choose the Akaike Information Content criterion for model choice since it performed best in simulations.

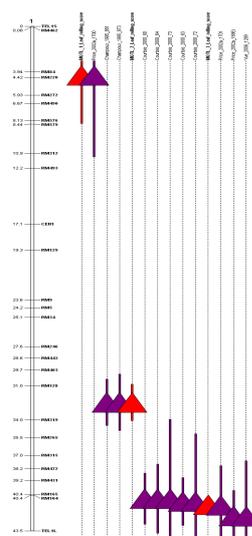


Figure 2: Example of leaf rolling QTLs (in violet) and meta-QTLs (in red) on chromosome 1

Results and discussion

The QTLs were organized in "hot spots", with many QTLs for several target traits localized on the same chromosome segment (Table 2). The concentration of QTLs for leaf parameters often corresponded to concentration of QTLs for biomass and root parameters (e.g. segment 35-40 Mb on chr1 where *sd1*, the gene controlling semi-dwarfism, is located; or segment 25-35 Mb on chr2). In a few cases, they were located on segments independent from biomass that carried root QTLs (e.g. segment 0-5 Mb on chr3; or segment 20-30 Mb on chr7; or segment 15-20 Mb on chr9). These co-localizations may reflect the effect of transpiration, which increases with biomass, and is permitted by a good water uptake favored by deep and thick roots when there is water in the deeper layers.

The 375 QTLs were resolved into 90 metaQTLs (Figure 1). The average confidence interval decreased in parallel from 9.58 Mb for QTLs to 4.25 Mb for meta QTLs. Using the possibility offered by OrygenesDB (<http://orygenesDB.cirad.fr>) to retrieve all the genes between two points, it was possible to list the putative candidate genes below the meta-QTL confidence intervals. In one of the cases (a biomass meta-QTL on chr2 with a confidence interval below 60 kb), the number went down to 8 genes. Because they involved more studies and more QTLs, the results obtained with root QTLs were more precise with a resolution of 3 genes in the best case.

Table 2: Physical position of the QTLs and metaQTLs for leaf parameters compared to those for biomass and root parameters

Chr (Mb)	LEAF PARAMETERS								BIOMASS			ROOTS			
	LR	LD	RWC	OA	LR	LD	RWC	OA	OTLs	MOTLs	OTLs	MOTLs	OTLs	MOTLs	
1	0-10	1	1	1	1	1	1	1	2	1	1	1	5	5	
1	10-15	1	1	1	1	1	1	1	2	1	1	1	4	3	
1	15-20	1	1	1	1	1	1	1	2	1	1	1	4	3	
1	20-25	2	2	1	1	1	1	1	4	2	1	1	7	6	
1	25-30	1	1	1	1	1	1	1	2	1	1	1	4	3	
1	30-35	2	2	1	1	1	1	1	4	2	1	1	7	6	
1	35-40	9	1	2	1	1	1	1	12	4	1	1	15	14	
1	40-44	1	1	1	1	1	1	1	2	1	1	1	4	3	
2	0-5	1	1	1	1	1	1	1	2	1	1	1	4	3	
2	5-10	1	1	1	1	1	1	1	2	1	1	1	4	3	
2	10-15	1	1	1	1	1	1	1	2	1	1	1	4	3	
2	15-20	2	1	1	1	1	1	1	4	2	1	1	7	6	
2	20-25	2	1	1	1	1	1	1	4	2	1	1	7	6	
2	25-30	2	1	1	1	1	1	1	4	2	1	1	7	6	
2	30-35	2	1	1	1	1	1	1	4	2	1	1	7	6	
2	35-40	2	1	1	1	1	1	1	4	2	1	1	7	6	
3	0-5	2	3	1	1	1	1	1	2	3	1	1	3	3	
3	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
3	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
3	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
3	20-25	2	1	1	1	1	1	1	4	2	1	1	5	4	
3	25-30	2	1	1	1	1	1	1	4	2	1	1	5	4	
3	30-35	2	1	1	1	1	1	1	4	2	1	1	5	4	
3	35-40	2	1	1	1	1	1	1	4	2	1	1	5	4	
4	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
4	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
4	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
4	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
4	20-25	2	2	1	1	1	1	1	4	2	1	1	5	4	
4	25-30	2	2	1	1	1	1	1	4	2	1	1	5	4	
4	30-35	2	2	1	1	1	1	1	4	2	1	1	5	4	
5	0-5	3	2	1	1	1	1	1	5	3	1	1	8	7	
5	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
5	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
6	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
7	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
8	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
9	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	0-5	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	5-10	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	10-15	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	15-20	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	20-25	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	25-30	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	30-35	1	1	1	1	1	1	1	2	1	1	1	3	2	
10	35-40	1	1	1	1	1	1	1	2	1	1	1	3	2	
Total		85	48	25	25	12	22	17	12	270	63	87	27	637	477

Conclusion

QTL meta-analysis is an efficient method to narrow down the number of candidate genes underlying a set of QTLs. The quality of the analysis, however, depends on the number and quality of the QTLs included in the analysis. It is therefore useful, when possible, to increase the number of studies included but care as to be taken to make a screening of the QTL set based on map quality, QTL confidence interval and LOD score before undertaking the analysis.

References

- Champoux et al (1995) Theor Appl Genet 90:969-981
- Courtois et al (2000) Molecular Breeding 6:55-66
- Courtois et al (2003) Euphytica 134: 335-345
- Courtois et al (2009) Rice 2(2):115-128
- Gomez et al (2006) Am J Bioch & Biotech 2(4):161-169
- Hemamalini et al (2000) Euphytica 112:69-78
- Hittalmani et al (2003) Theor Appl Genet 107:679-690
- Kamoshita et al (2002) Theor Appl Genet 104:880-893
- Lanceras et al (2004) Plant physiology 109:1237-1246
- Lafitte & Courtois (1999) In: Molecular approaches for the genetic improvement of cereals for stable production in water-limited environment, CIMMYT, p97-102
- Lilley et al (1996) J Exp Bot 47:1427-1436
- McMillan et al (2006) Theor Appl Genet 113:935-964
- Nguyen et al (2004) Mol Gen Genomics 272:35-46
- Price et al (2002) Plant Mol Biol 48:683-695
- Price et al (2002) Field Crop Research 76:25-43
- Robin et al (2003) Theor Appl Genet 107:1288-1296
- Veyrieras et al (2007) BMC Bioinformatics 8:49
- Xu et al (2004) Theor Appl Genet 109: 640-647
- Yue et al (2005) Theor Appl Genet 111: 1127-1136
- Yue et al (2006) Genetics 154:885-891
- Zhang et al (2001) Theor Appl Genet 103:19-29