

GENOMIC ANALYSIS IN THE STUDY OF DROUGHT TOLERANCE USING RECOMBINANT INBRED LINES AND RECURRENT SELECTION

¹Márcio Elias Ferreira; ¹Georgios Pappas Jr.; ²Alexandre A. Martins, ²Pedro Ítalo Tanno, ³Marco Pessoa-Filho, ¹Roberto C. Togawa; ¹Orzenil Silva Junior; ⁴Paulo H. N. Rangel

Introduction - Water use efficiency in agriculture is one of the major research challenges of this century. Crop yield losses caused by limitations on water use have been substantial in recent years (FAO, 2002). It is possible that these limitations will have even greater impacts as a consequence of global warming in a climate change scenario. It is estimated that 70% of the freshwater consumed by human populations are used for food production (WMO, 1997). This brings a perspective of a world where water will be increasingly scarce, and, therefore, more expensive, causing significant impact on crop production costs. Plant breeding for drought tolerance is one of the approaches that can be applied in order to improve water use efficiency in agriculture.

The genetic control of drought tolerance in cereals - such as rice - is quantitative, being determined by several genes with a strong interaction with the environment (Bennett, 2003). Because of their complexity, quantitative traits pose a great challenge for breeding programs. The number of genes involved in drought tolerance, in addition to genotype X environment interactions, makes the development of superior cultivars by breeding programs more difficult. In spite of the unquestionable advances in the development of improved crop varieties brought about by the use of conventional breeding strategies, the development of drought tolerant varieties by these methods has been recalcitrant and marginal (Tuberosa and Salvi, 2006). There are many factors contributing to this lack of efficiency, since there are several variables interfering on the water-plant system, from the capture of water by roots and their conduction to plant tissues, to its release in the atmosphere through transpiration. Throughout this journey, some environmental factors influence water use capacity by the plant: soil type, wind speed, air moisture, temperature, solar radiation, among others. In the plant itself, root system and stem architecture, cell and tissue resistance, transpiration, hormone regulation, photosynthesis, cell-wall composition and adjustment, plant developmental stage, among others, all affect water flow. Therefore, evaluating drought tolerance in plants is far from being a trivial task. Being such a complex trait to deal with, the definition of drought tolerance itself has had different interpretations by research groups. We use as a reference that drought tolerance is the capacity of a variety to present greater yields than other varieties of the same species under limiting water availability in the soil.

¹Embrapa Recursos Genéticos e Biotecnologia. Brasília, DF.

²Universidade de Brasília (UnB). Brasília, DF.

³Embrapa Cerrados. Planaltina, DF.

⁴Embrapa Arroz e Feijão. Santo Antônio de Goiás, GO. E-mail: phrangel@cnpaf.embrapa.br

Drought tolerance in rice – Rice (*Oryza sativa* L.) is a hydrophytic plant, adapted to water abundant environments. The Green Revolution promoted modifications in plant architecture in order to better adapt it to irrigated production systems, without drought stress. The definition of an ideotype for irrigated rice has culminated in a significant increase in yield, witnessed in several parts of the world in recent decades. This was a landmark in rice breeding. Most of the global rice production is done in banded fields or flood prone areas, using varieties highly responsive to crop inputs. It is important to notice, however, that in some countries, a significant share of rice production is carried in rainfed environments. Brazil is the largest rice producer among western countries, having a great proportion of its production receiving irrigation only by natural precipitation. Upland rice has a very important role in rice production in Brazil, representing about 60% of the cultivated area, but with less than 30% of total grain production (CONAB, 2010). Therefore, in a singular manner when compared to Asian countries, upland rice is of utmost importance in Brazil, where rice is one of the staple foods for almost 200 million people. Mean yield values for upland rice in Brazil are still low, around 2.2 ton/ha, which is lower than the average 6.8 ton/ha for irrigated rice (CONAB, 2010). Increasing yield for upland rice, with more than 1.6 million ha planted every year, is, therefore, fundamental and strategic for the country. Increasing water use efficiency in upland rice is part of the equation for making it more productive. Therefore, it is a fact that the Green Revolution occurred for irrigated rice, bringing innovation for plant ideotype and increasing crop yields. A “Green Revolution”, however, has not succeeded for upland rice yet. In order to contribute to the increase of rice yield in upland production systems and for the stability of production for this crop, a research project has been conceived which integrates conventional methods of plant breeding with new genomic tools for the understanding and investigation of drought tolerance control mechanisms in rice. The strategy being applied to this end is illustrated in Figure 1 and discussed below.

Thematic Core Collection for Drought Tolerance - Initially, an assessment of rice landraces deposited in the Rice Germplasm Bank of Embrapa was performed. Among the 3,200 accessions of rice collected by Embrapa starting in the 1970's all around the Brazilian territory, a list of approximately 700 accessions of rice landraces showing a *tropical japonica* genetic background was delimited. These landraces are adapted to different environmental conditions in the country, almost always cropped for subsistence by small farmers. This germplasm stands as a true genetic treasure, given its diversity for root system and stem architecture, plant height, type and color of grains, plant cycle, etc. The accessions were genotyped with panels of microsatellite markers distributed throughout the rice genome (Pessoa-Filho et al., 2007). With the information gathered by this genotyping process, a thematic core collection for drought tolerance was obtained (Pessoa-Filho, 2010). It is composed of approximately 90 *tropical japonica* landraces. This core collection presents high levels of diversity for drought tolerance, conserving about 84% of the gene diversity of the source collection in a sample with ~ 18% of its original size.

Phenotyping of the thematic core collection for the definition of an ideotype

– Due to its high level of genetic diversity and variability in drought stress responses, the thematic core collection for drought tolerance is being submitted to evaluations of several traits, including: morphological descriptors for the species (IRRI/IPGRI), leaf anatomy by scanning microscopy, canopy temperature, date of flowering, plant height, weight of 100 grains, number of tillers, number of panicles, yield, root length, root diameter, root volume, among others. Special attention is being given to root development, under the rationale that drought tolerant plants explore water more efficiently from deeper soil layers. Phenotyping experiments are being performed in field conditions and in the greenhouse, with evaluation of plots in treatments with and without drought stress. The main goal is to detect morphological, structural, physiological and genetic variations associated with drought tolerance. The collected phenotypical data are being compiled in order to assist on the definition of a new ideotype for upland rice, focusing on a plant with an architecture that promotes greater yields in this production system. Data indicate, for instance, a significant negative genetic correlation between the drought susceptibility index (DSI) (Fischer and Maurer, 1978) and plant height, as well as a significant negative correlation between DSI and root length (20-40 cm and 40-60cm layers). This suggests that the definition of a possible upland rice ideotype should take into consideration greater plant heights than observed in semi-dwarf varieties and longer root lengths than current upland varieties. Apparently, the irrigated rice ideotype from the Green Revolution was directly transferred to current upland rice varieties when, realistically, a new upland rice ideotype still claims to be built.

Development of segregating populations for the identification of genomic regions associated with drought tolerance

– Four Brazilian landraces, belonging to the *tropical japonica* group (Chorinho, Puteca, Catetão and Ligeiro) were selected in the thematic core collection in order to develop populations for genetic studies concerning drought tolerance. Three other *tropical japonica* varieties well-known for showing drought tolerance (Azucena, Moroberekan and IAC 165) and a variety with desirable traits for grain quality (Primavera), have also been selected as parental varieties. Some of these varieties are sources of genes for blast resistance, such as Azucena, Moroberekan and Chorinho. Tolerance to abiotic (e.g. drought tolerance) and biotic stresses (e.g. blast resistance) is a limiting factor for the development of upland rice varieties with higher yields in Brazil. These new populations segregate not only for drought tolerance, but also for blast resistance.

The crosses were designed in order to develop three types of populations:(1) 8-Way RIL population - a population of ~ 1200 recombinant inbred lines derived from the cross between eight *tropical japonica* rice varieties (Chorinho, Puteca, Primavera, IAC 165, Catetão, Azucena, Moroberekan, and Ligeiro), presenting variations in mechanisms for tolerance or susceptibility to drought and for blast resistance. Dozens of individuals of each variety were genotyped with a panel of microsatellite markers distributed throughout the rice genome. One individual representing the standard fingerprinting of each variety was selected in order

to obtain paired crosses (Chorinho x Puteca; Primavera x IAC 165; Catetão x Azucena; Moroberekan x Ligeiro). This population is being used in genetic and genomic analyses, since it is expected to be composed of highly homozygous recombinant inbred lines with reduced extensions of linkage disequilibrium. From generation G4 on, up till at least generation F12, the population is being conducted by SSD. When finished, this population will be adequate for genetic association studies for the discovery and selection of drought tolerance and blast resistance genes in rice. It will also be useful for similar studies regarding other economically important traits segregating in the population.

(2) RIL (recombinant inbred lines) populations, derived from bi-parental crosses (Chorinho x Puteca; Primavera x IAC 165; Catetão x Azucena; Moroberekan x Ligeiro). The four RIL populations, each having ~200 recombinant lines, are being used for QTL mapping studies for the identification of genomic regions associated with the control of drought tolerance and blast resistance.

(3) Recurrent Selection Population, derived from the 8-way cross between varieties mentioned above, with selection cycles derived from generation G4. This population is segregating for drought tolerance, grain quality, yield, plant type, blast resistance, among other economically important traits. Drought tolerance is the main focus of the project, but this population allows the integration between programs of selection for drought tolerance and blast resistance. Because it is a recurrent selection population, it will be adequate for the use of genome wide selection principles (Meuwissen et al., 2001). In this case, the thematic core collection for drought tolerance and the bi-parental RIL populations, after thorough phenotyping for drought tolerance traits and genotyping with molecular markers (SSR's and SNP's) are being used as "training populations" in order to estimate the value of each molecular marker on the composition of the drought tolerant phenotype. These data will be used in the selection of superior plants in each selection cycle for recombination and obtainment of the next cycle progeny.

Rice genome re-sequencing and SNP detection – Genome wide selection requires discovering and genotyping of a large number of molecular markers scattered throughout the genome. In order to cope with this challenge, a second generation DNA sequencing technology (Illumina) has been used to re-sequence the genome of the eight parental varieties used to obtain the segregating populations (bi-parental RIL, 8-way RIL and recurrent selection populations). This approach allows for a broad search for SNP polymorphism throughout the rice genome. It also makes it possible to characterize all sources of genetic variation at the nucleotide level between the parental varieties. Genome re-sequencing has been accomplished with an Illumina GAIIx sequencer, based on DNA from the same plant used to obtain the segregating populations. A total of 66 million short reads (76 base long) have been obtained for the variety Chorinho, covering ~ 5 Gigabases of the variety genome (~ 13.5X average genome coverage). The following values have been observed for the other parental varieties: Puteca (80,443,550

reads; average coverage = 16.4X); Primavera (60,197,846 reads; average coverage = 12.3X); IAC 165 (50,470,079 reads; average coverage = 10.3X); Catetão (62,978,225 reads; average coverage = 12.9X); Azucena (57,482,501 reads; average coverage = 13.4X); Moroberekan (52,052,047 reads; average coverage = 10.6X); Ligeiro (52,804,710 reads; average coverage = 10.8X).

The analysis aiming at SNP discovery and marker genotype selection indicated that the number of SNP´s identified between each parental genome and the reference genome (Nipponbare) is high, varying from 504,337 to 584,996 SNP´s in the pairwise comparisons evaluated. The strategy used also included the comparison between the parental varieties (Chorinho x Puteca; Primavera x IAC 165; Catetão x Azucena; Moroberekan x Ligeiro) (Figure 2). For the recurrent selection population and the 8-way RIL population, SNP´s have been selected based on confirmed polymorphism in at least 2 out of 8 varieties. Several filters have been tested to identify and select the SNP´s employed in the genotyping experiments.

Besides genome re-sequencing of the eight varieties to promote SNP discovery and selection, SNP mining was also pursued through comparison of rice ESTs (Expressed Sequence Tags) deposited on international databases. By doing so, 158 SNPs have been extracted after screening 1,251,304 EST sequences of the rice genome deposited on the dbEST database (<http://www.ncbi.nlm.nih.gov/dbEST/>). Also, 1,797 SNPs have been selected after screening 100.1 Mb deposited at the OryzaSNP database (<http://www.oryzasnp.org/>). This last database has been recently developed by a consortium lead by IRRRI (International Rice Research Institute) focused on genome re-sequencing of 20 rice cultivars (McNally et al., 2009).

Structural variants with 4 bp or more have been detected in large amounts on each pairwise comparison with the reference genome, varying from 51.332 to 80.092 indels per comparison. New methods of detection and validation of indels have been tested. Indel based genetic mapping and the study of the role of structural variants on plant phenotype diversity is still an underexplored area, potentially promising for the understanding of the genetic diversity of the species, with eventual application in selection for traits of economic importance.

The SNP´s selected in the three databases (genome re-sequencing of eight parental varieties, Oryza dbEST and OryzaSNP) are in the process of experimental validation with samples of the different segregating populations and accessions of the Rice Germplasm Bank. The objective is to obtain a minimal set of hundreds of SNP markers that will be used in large scale population genotyping and in genome wide selection of superior individuals in the recurrent selection population. The integration of classic breeding methods with new genomic tools opens the perspective of improving the efficiency of selecting superior genotypes based on DNA polymorphism analysis, aiming at the development of plants that are more tolerant to drought stress.

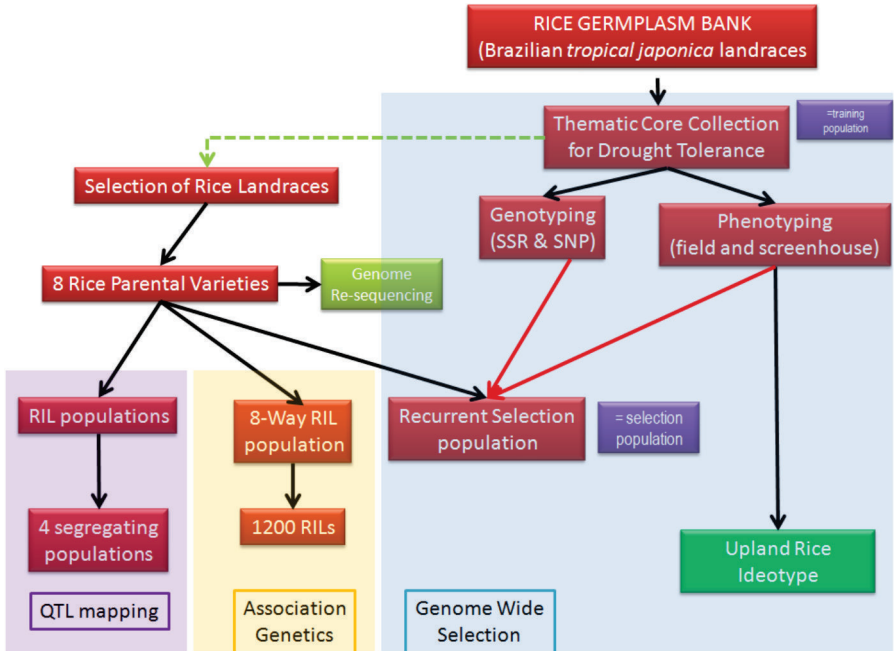


Figure 1. Strategy applied for the detection, isolation and investigation of regions of the rice genome associated with drought tolerance for the development of drought tolerant varieties.



Figure 2. Detection of a SNP (A/T) locus mapped on rice chromosome 1, based on genomic re-sequencing data of varieties Primavera and IAC 165.

References

Bennett, J. 2003 Opportunities for increasing water productivity of CGIAR crops through plant breeding and molecular biology. In. Improving water productivity in agriculture. Comprehensive Assessment of Water Management in Agriculture Series. Jacob W. Kijne, Randolph Barker and David Molden (eds). Chapter 7:103-126

CONAB-Companhia Nacional de Abastecimento 2010 Acompanhamento da Safra Brasileira: Grãos – 4º Levantamento, janeiro 2010, CONAB, 39 pp

FAO. 2002. Crops and Drops: Making the Best Use of Land and Water. FAO, Rome.

Fischer, R. A.; Maurer, R. 1978 Drought resistance in spring wheat cultivars. I grain yield responses, Australian Journal of Agricultural Research. v. 29, p. 897-912.

McNally et al., 2009 Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc Natl Acad Sci U S A. 2009 July 28; 106(30): 12273–12278.

Meuwissen, T.H.E.; Goddard, M.E. and Hayes, B.J. 2001 Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829

Pessoa-Filho, M. A. C. ; Belo, A. ; Alcochete, A. A. N. ; Rangel, P. H. N.; Ferreira, M. E. 2007 A set of multiplex panels of microsatellite markers for rapid molecular characterization of rice accessions. BMC Plant Biology, 7:23

Pessoa-Filho, M.A.C.P.; Rangel, P.H.N. and Ferreira, M.E. 2010 Extracting samples of high diversity from thematic collections of large gene banks using a genetic-distance based approach. BMC Plant Biology 10:127

Tuberosa, R. and Salvi. S. 2006 Genomics-based approaches to improve drought tolerance of crops. Trends in Plant Science 11(8) 405-412

WMO 1997 Comprehensive assessment of the freshwater resources of the world. World Meteorological Organization, Geneva, Switzerland, 52 pp.