International Training-Workshop (2007)
The Conservation and Utilization of Tropical/Subtropical Plant Genetic Resources
131-134

Molecular Diversity and Evolution of Aromatic Rice in Thailand

By

Apichart Vanavichit[1]

Introduction

The rice world market ranks aromatic rice at the top. For the 4-5 million Tons of aromatic rice worth 2-2.5 billion US Dollars, Thai Jasmine rice shares almost half-half with Basmati rice from India and Pakistan with small players like USA, Vietnam, and recently Cambodia. The lucrative market will become more competitive in the near future from new players like USA, Australia and Vietnam that has launched new plant type that combines high grain yield with aromatic quality.

Biodiversity of Aromatic Rice

Existing aromatic rice cultivars belongs to three groups according to isozyme analysis (Khush et al, 2000). Group I includes Thai Hom Mali and some aromatic rice varieties from China, Vietnam and Cambodia. Group V includes aromatic rice varieties from indian subcontinent including Basmati. The center of origin of the Group V aromatic rice is the foothill of Himalayas in India where large variation of aromatic rice exists. From the foothill, aromatic rice spreads northwestward to as far as Myanmar (Khush et al., 2000). No member of Group V was found in the Southeast and East Asia. The last is Group VI including aromatic rice from Indonesia, the Philippines and China. Azucena and Milagrosa are two well known aromatic rice varieties from the Philippines. Aromatic rice found in Italy, France and others are believed to be introduced from Asia when European convoys reached South and Southeast Asia several hundred years ago.

In Thailand, more than 600 aromatic landraces were distributed in four ecotypes, the deep-water, lowland rainfed, irrigated and upland ecotypes. In lowland rainfed rice ecotype, most aromatic rice varieties found here is tall and photoperiod-sensitive in contrast to the

[1] Associate Professor, Rice Gene Discovery and Rice Science Center, Kasetsart University, Kamphang Saen Thailand; e-mail: vanavichit@gmail.com

ones found in the upland. The infamous Jasmine rice was first reported in Chacherng Sao province during the seed purification program for elite landrace. In order to determine the population structure of these invaluable germplasm, different molecular markers were used.

Molecular Marker Systems

Phenotypic markers are not an ideal for phylogenetic studies due to the limited abundance, low degree of polymorphism, influenced by environments (GxE), and low throughput. Desirable molecular markers must abide with these properties:

- Highly abundant and evenly distributed throughout the genome,
- Highly polymorphism or polymorphic information content (PIC),
- High multiplexing ratio,
- Co-dominant,
- Neutral.

In addition, researchers must consider the development cost of the molecular markers, robustness and high reproducibility, and laboratory transfer guaranteed. A large number of molecular markers is now available, each has different advantages and drawbacks. Among these, five molecular markers commonly used are RFLP (Restriction Fragment Polymorphism), RAPD (Random Amplified Polymorphic DNA), SSR (Simple Sequence Repeat) and SSCP (Single Strand Conformational Polymorphism).

Molecular Markers for Grain Aroma

Genetic markers for grain aroma were developed from molecular mapping. The locus namely *fgr* was first discovered by molecular tagging of gene controlling aroma (Ahn *et al*., 1992). Near-isogenic lines were developed from introgression a candidate segment from Della, the aromatic donor. The molecular mapping located genes for grain aroma at a likelihood of 4.5 cM near RG28, a RFLP marker located on chromosome 8. In Thai Hom Mali Rice, Khoa Dawk Mali, using bulk segregant analysis (BSU), a PCR fragment, Jasmine 500, a PCR fragment generated from Random Amplified Polymorphic DNA (RAPD) in an F2 population developed from Thai Hom Mali Rice and CT9993 (Tragoonrung *et al*., 1995) was linked to grain aroma. Because of the qualitative scoring used in those studies, they were not possible to determine quantitative trait loci for the intensity of grain aroma. The quantitative trait loci (QTL) analysis was firstly reported for grain aroma in a doubled haploid population (DH) developed from Azucena, the upland fragrance rice from the Philippines and IR64

(Lorieux *et al*., 1996). In this study, sensory test and the accumulation of 2-AP were perfectly mapped on the same chromosomal region flanked by RG28 and RG1.

Positional Cloning of Grain Aroma

High resolution mapping was developed from isogenic lines differing only for grain aroma. Map-based cloning came close to only few candidate genes. Two PCR-based markers were developed flanking the genomic segment responsible for grain aroma. Several SNPs were discovered by additional comparative sequencing between a pair of ISLs. Three open reading frames (ORF) were identified by annotation and homology search. Subsequent segregation for grain aroma from a single heterozygous F11 plants were completely associated with an 8 bp deletion in the ISLs and Thai Hom Mali Rice. This specific indel located about 729 bp from the initiation codon and within an aldehyde dehydrogenase domain of the ORF. We named this ORF Os2AP. The Os2AP consists of 15 exons. Ten more SNPs including a mutation at a 5'splice site were identified in the exon 2 but no association was found with grain aroma. This 8 bp indel was used to develop PCR-based markers, designated as 'Aromarker', for efficient marker-assisted selection in rice. Transformation of Nipponbare, a non-aromatic japonica rice, to the aromatic pair clearly demonstrated that the suppression of the Os2AP gene bring out aromatic compound.

Phylogenetic Analysis of Aromatic Rice

We compared phylogenetic analysis based on AFLP, SCAR and micro-satellite (SSR) of the aromatic rice germplasm. AFLP was used to classified aromatic rice germplasm. Selective primers were screened for polymorphic information content. Two selective primer combinations selected were MS + AAC/ER + CAA and MS + CAA/ER + CAG. The genomic DNAs were restriced with EcoR1 and Tru91 and ligated with AFLP adapters. Further selective amplification was accomplished with the second primers. For SCAR, Jasmine 500 was developed from specific amplicon that consistently aligned with aromatic rice in a mapping population. SSRs as the only co-dominant markers were screened for high PIC values Cluster analysis using Jaccard's coefficient for similarity index generated phylogenetic tree for aromatic rice germplasm. The phylogenic tree (Dendrogram) was calculated using unweighted pair-group method arithmetic analysis (UPGMA).

The results from these three type II molecular markers were that aromatic rice varieties did not form a single cluster as expected. AFLP was considered the best molecular marker because it can cluster rice into several sub-clusters. In particular, APLP can differentiate

Jasmine rice into a separate cluster. It can be concluded that aromatic germplasm was highly diversified.

The Origin of the Aromatic Gene

How many aromatic genes are found in the rice genome? This is quite an interesting questions when breed rice for aroma quality. We extensively survey all aromatic rice germplasm representing different ecosystems in Thailand including Basmati and Acuzena from India and the Philippines, respectively; using the *Aromarker* we developed here. All aromatic rices from all three isozyme groups shared the same 8 bp deletion in the *Os2AP*. It is almost conclusive to state that all aromatic genes in cultivate rices are the same. Searching for the ancestor is on the way. One possibility of the ancestor is wild rice. By comparing the limited accessions of *Oryza glaberrima* and *Oryza officinalis*, no aromatic allele of *Os2AP* was detected. This means that it is possible that *O. longistaminata* and *O. breviligulate,* the ancestor of the cultivated rice in Africa *O. glaberrima*, was not the ancestor of the aromatic gene. It is likely that the common ancestor aromatic gene could be found in *O.nivara* and *O.rufipogon*. We surveyed about 120 accessions for *O.rufipogon* and *O.nivara* using the *Aromarker* and with surprise; we found many of them carried 8 bp deletion of the *Aromarker*. All wild rices carrying the aromatic allele of *Os2AP* produced aromatic grains. We also found that many of them are heterozygous. It is now conclusive to state that the common ancestor of aromatic gene can be found in *O.rufipogon,* the wild perenial ancestor of *O.nivara*. In order to answer if it is one or many populations of wild relatives, genomic sequence of this gene must be compared.

Combining the AFLP results with the functional marker of aromatic gene, it is possible that the 8 bp deletion was naturally mutated in wild progenitors such as *Oryza nivara* and *O. rufipogon*. As the genomic sequence of most aromatic wild species and landraces are completed, we can be able to determine the origin of the jasmine rice.