Bioagro 31(1): 3-12. 2019

# ASSESSMENT OF MOLECULAR GENETIC DIVERSITY OF ECUADORIAN RICE CULTIVARS USING SIMPLE SEQUENCE REPEAT MARKERS

Iris Pérez-Almeida[1], Roberto Celi-Herán[2], Fernando Sánchez-Mora[3], Lenin Paz-Carrasco[4] and Belén Ramos-Viteri[4]

## ABSTRACT

Molecular markers are useful tools for evaluating genetic diversity and determining cultivar identity. Thirty simple-sequence-repeat (SSR) markers were selected in order to evaluate the genetic diversity within 76 cultivars of the Ecuadorian Rice Program for breeding. One-hundred-ninety-four alleles were detected in 22 SSR polymorphic markers, number of alleles per marker ranging from 2 to 24, with an average of 9 alleles per locus. The sizes of the alleles varied between 62 to 280 bp, with an average polymorphism information content value of 0.624, ranging from 0.202 (RM125) to 0.943 (RM413), indicating significant genetic diversity among and within the rice accessions. The average observed heterozygosity (Ho) was 0.086, while average expected genetic diversity (He) was 0.667. A set of eight of the polymorphic SSR markers produced seventeen unique alleles for the genotypes studied and could distinguish released cultivars from the rest of accessions. A dendrogram constructed using the unweighted pair-group method with arithmetic means (UPGMA) grouped the 76 rice materials in four well differentiated major clusters whereas the Structure program without any *a priori* information provided support for the existence of three genetically distinct clusters (K = 3).
**Additional key words:** Cluster analysis, molecular marker, *Oryza sativa*, rice breeding, variability

## RESUMEN

**Evaluación molecular de la diversidad genética de variedades de arroz de Ecuador utilizando marcadores de secuencia simple repetida**

Los marcadores moleculares son herramientas útiles para evaluar la diversidad genética y determinar la identidad del cultivar. Se seleccionaron 30 marcadores de secuencia simple repetida (SSR) para evaluar la diversidad genética en setenta y seis cultivares del Programa de Arroz Ecuatoriano con fines de mejoramiento. Se detectó un total de 194 alelos en 22 marcadores polimórficos SSR, el número de alelos por marcador osciló entre 2 y 24, con un promedio de 9 alelos por locus. Los tamaños de los alelos variaron entre 62 y 280 pb, con un valor de contenido de información polimórfica promedio de 0,624, variando entre 0,202 (RM125) y 0,943 (RM413), lo que indicó diversidad genética significativa entre y dentro de las accesiones de arroz. La heterocigosidad observada promedio (Ho) fue 0,086, mientras que la diversidad genética esperada promedio (He) fue 0,667. Un conjunto de ocho de los marcadores SSR polimórficos produjo diecisiete alelos únicos para los genotipos estudiados y pudo distinguir los cultivares liberados del resto de las accesiones. Un dendrograma construido utilizando el método de pares no ponderados con medios aritméticos (UPGMA) agrupó los 76 materiales de arroz en cuatro grupos principales bien diferenciados, mientras que el programa Structure sin información *a priori* proporcionó soporte para la existencia de tres grupos genéticamente distintos (K = 3).
**Palabras clave adicionales**: Análisis de conglomerados, marcador molecular; mejoramiento de arroz; *Oryza sativa*, variabilidad

## INTRODUCTION

Rice (*Oryza sativa* L.) is one of the staple foods in the world, directly feeding nearly half of the world's population (Travis et al., 2015). It is one of the most extensively cultivated cereal crops spreading across a wide range of geographical, ecological and climatic regions.

Received: July 3, 2018 Accepted: December 1, 2018
[1] Facultad de Ciencias Agrarias, Universidad de Guayaquil. Guayaquil EC090302 Ecuador. e-mail: iris.pereza@ug.edu.ec, ibperez1@gmail.com
[2] Instituto Nacional de Investigaciones Agropecuarias, Estación Experimental Litoral Sur, Rice Program, Guayaquil EC092406 Ecuador. e-mail:roberto.celi@iniap.gob.ec
[3] Facultad de Ingeniería Agronómica, Campo Experimental La Teodomira, Universidad Técnica de Manabí. Lodana, EC 13132, Ecuador
[4] INIAP - Instituto Nacional de Investigaciones Agropecuarias, Estación Experimental Litoral Sur. Guayaquil, EC092406 Ecuador. e-mail: lenin.paz@iniap.gob.ec; maria.ramos@iniap.gob.ec

It is key for food security in developing countries, such as Ecuador, where it is closely associated with traditional culture and customs. Rice is principally cultivated in the provinces of Guayas (67 %) and Los Ríos (29 %) (MAG, 2015). In 2016, approximately 366.194 ha were harvested with a total production of 1.53 million ton and an average yield level of 4.19 ton·ha$^{1}$ (FAOSTAT, 2018).

The rice breeding program in Ecuador, created in 1968 and led by the National Institute for Agricultural and Husbandry Research (INIAP), has developed several rice varieties between 1971 and 2017. The methodology has entirely been based on agro-morphological characterization through crossings, evaluation and selection of introduced materials.

The morphological characterization of the materials by itself is not very reliable, due to the fact that most of the traits of interest have low heritability and are genetically complex, thus requiring the use of more precise and modern techniques. Nowadays, there are more reliable molecular markers (nucleotide sequence located in a particular physical position in the genome) that are not affected by the environment, crop phenology, locality or agronomic practices. Their use can provide an extensive coverage of the rice genome, which has already been sequenced (Yu et al., 2002) and is available for research worldwide.

Molecular markers are powerful tools for the detection of genetic variation, the interpretation of genetic relationships within and between species, and a rational management of plant genetic resources.

Several types of molecular marker methods have been developed to study genetic diversity. SSRs are technically efficient PCR-based markers, cost-effective and are the most commonly used due to their simplicity, reproducibility, high polymorphism, multi-allelic nature, codominant inheritance, abundance and extensive genome coverage (Temnykh et al., 2002; Garris et al., 2005; Orjuela et al., 2010). The SSR markers have been widely used for various purposes such as diversity analysis, genome mapping and varietal identification.

Assessing genetic diversity is an essential component in the development of improved rice cultivars for resistance to diseases and pests, and tolerance to abiotic stresses (Wei et al., 2009). It is important when selecting appropriate donors but also for protecting unique rice varieties (Parikh et al., 2012). In this regard, the genetic and molecular diversity and relationships among Ecuadorian rice cultivars have not been characterized but germplasm analyses and information about levels and patterns of genetic diversity can provide invaluable information in a successful breeding program. Knowledge regarding the extent of genetic variation and genetic relationships between genotypes are vital for designing effective breeding and conservation strategies (Roy et al., 2015). Scant information is available on the genetic diversity of traditional rice varieties kept by farmers in Ecuador.

Genetic analysis of rice landraces collected from a region also helps in understanding the complex interaction between rice diversity and human cultivation practices and culture, as the cultivar structure is shaped by the interplay between adaptation to the local environment and artificial selection imposed by the farmers (Roy et al. 2015). They represent the main sources with which to undertake genetic improvement for the tolerance of biotic and abiotic stresses (Xu et al., 2016). Moreover, identifying the genetic diversity of local varieties compared with improved or introduced ones would assist the breeding of elite varieties for use in sustainable agriculture (Xu et al., 2016).

The genetic diversity assessment is a key component in germplasm characterization, utilization, and conservation. Rice breeding programs aim to obtain genotypes with special traits by introducing improved varieties of narrow genetic base, which in the long run may result in genetic vulnerability caused by a genetic uniformity (Day et al., 2005; Keneni et al., 2012). The continuous utilization of local germplasm in breeding programs leads to saturation of desirable alleles in cultivating varieties and loss of diversity, resulting in low variability and a narrow genetic base, which has been reported in rice cultivars for Latin American (Cuevas et al., 1992), or individual countries such as Brazil (Guimaraes et al., 1996; Rangel et al., 1996), Cuba (Fuentes et al., 1999), Chile (Aguirre et al., 2005), and Venezuela (Ghneim et al., 2008).

In the present study, a set of 30 microsatellite markers distributed on 12 different chromosomes of the rice genome were used for DNA profiling of 76 rice genotypes, including landrace varieties, commercial varieties and nursery breeding lines, with the objectives of characterizing their genetic

diversity and assessing the genetic relationships among them in order to modeling the breeding strategies.

## MATERIALS AND METHODS

**Plant materials.** The study was conducted at the Biotechnology and Phytopathology Departments of the National Institute for Agricultural and Husbandry Research (INIAP) South Coast Experimental Station, Guayas, Ecuador. Seventy-six materials (Table 1) were selected, that is, 13 publicly released rice varieties, 35 nursery observation lines and 24 landrace varieties. Their breeding programs and releasing institutes are listed in Table 1. Four international rice varieties, 'Fanny', 'Azucena', 'Fedearroz 50' and 'IR-64', were used as controls. All the accessions seeds were provided by INIAP's Rice Program.

**Table 1**. Rice germplasm studied in this research

| Group | Type | Code | No. |
|---|---|---|---|
| 1 | INIAP varieties | INIAP-10[1], INIAP-11[1], INIAP-12[1], INIAP-14[2], INIAP-15[3], INIAP-16[3], INIAP-17[3], INIAP-18[3], INIAP-19[3] (GO39590), INIAP-FL01[4], INIAP415[1], INIAP-FL0202[4] (Arenillas)*, INIAP-FL1480[4] (Cristalino)* | 13 |
| 2 | Nursery observation lines | GO00226[4], GO00559[4], GO00904[1], GO00918[1], GO00919[1], GO00920[1], GO00921[1], GO02576[4], GO02745[4], GO03169[4], GO03232[4], GO03526[4], GO03495[4], GO03530[4], GO03532[4], GO03568[4], GO03574[4], GO01415[4], GO39085[4], GO39783[4], GO39789[4], GO01384[4], GO00201[4], GO00623[4], GO01346[4], GO01377[4], GO01604[4], GO01671[4], GO01684[4], GO01827[4], GO01866[4], GO39599[4], GO03600[4], GO03494[4], GO03497[4] | 35 |
| 3 | Landrace varieties | Oriente-2, Oriente-3, Pana , Pancho Vera, Papayo, Paraguillo, Malcriado, Donato, Fama, F-Canuto, Gallinazo, Pico Negro, Piedad, Rabo de Yegua, Sacaclavos, Tinajones, Brasilero, Cafuringa 1, Canilla, Canuto, Cenit, Chato, Chileno, Chato Rayado | 24 |
| 4 | International varieties | Fanny[1], Azucena[5], Fedearroz 50[6], IR-64[2] | 4 |

[1] International Center for Tropical Agriculture (CIAT); [2] International Rice Research Institute (IRRI); [3] National Institute for Agricultural Research (INIAP); [4] Latin American Fund for Irrigated Rice (FLAR); [5]Bolivia; [6] Fedearroz, Colombia; *Commercial name

**DNA extraction and PCR reaction.** Ten seeds of each accession were grown for eight days in aseptic conditions using 240 mL clear glass plant tissue culture vessels with a bed of sterile water soaked tissue paper as support. DNA was extracted from 10-d old fresh rice leaves according to the method of Ferreira and Grattapaglia 1998. DNA was quantified using a Quantus Fluorometer (Promega) and diluted to 10 $ng \cdot uL^{-1}$ working concentration.

Thirty SSR markers covering all the 12 chromosomes were selected for the panel of genetic diversity analysis based on the published rice SSR Universal Core Map (Orjuela et al., 2010). The loci, chromosome position and repeat motifs for these markers are presented in Table 2. SSR primers were then obtained from Integrated DNA Technologies (USA). Primer pairs were hydrated and diluted to 10 mM working concentration. Melting temperatures were estimated according to manufacturer's data and optimal temperatures for hybridization during PCR were established using a temperature gradient ramp.

PCR amplification was carried out in 7.5 µL reaction mixtures, each containing 10 ng of template DNA, 0.13 mM of each primer, 2.5 mM of each dNTPs, 1X PCR buffer (100 mM Tris-HCl, pH 8.8, 500 mM KCl, 1 % Triton X-100, 1 % BSA), 2 mM $MgCl_2$ and 0.1 U of Taq DNA polymerase. A Techne thermal cycler was used along with the following PCR profile: an initial denaturation step of 3 min at 94 ºC, followed by 30 cycles of 30 s at 94 ºC, 30 s at 48-65 ºC (according to the primer melting temperature), 1 min at 72 ºC, and a final extension at 72 ºC for 5 min.

**Analysis of polymorphisms.** The PCR products were separated by agarose-based electrophoresis on a 2.5 % MetaPhor Agarose (Lonza) gel in 0.5X TBE buffer at 105 V for 2.3 h, stained with Diamond Nucleic Acid Dye (Promega) for

20 min and visualized under UV trans-illumination using an Enduro GDS gel documentation system (Labnet). The size of the amplification products was estimated by comparisons with a 20-bp molecular ruler from BioRad. Each gel was photographed and digital images of the gels were analyzed using Image J analysis software (National Institutes of Health) to determine the molecular weight of the amplicons.

**Table 2**. Summary statistics of the 22 simple sequence repeat (SSR) markers used in this study

| Locus* | Chr | Repeat motif | Range (bp) | n | AN | AE | RA | Ho | He | PIC |
|---|---|---|---|---|---|---|---|---|---|---|
| RM125 | 7 | (GCT)8 | 105-137 | 76 | 2 | 1.296 | 0 | 0 | 0.230 | 0.202 |
| RM536 | 11 | (AG)16 | 235-250 | 76 | 5 | 1.315 | 3 | 0 | 0.241 | 0.230 |
| RM316 | 9 | (GT)8(TG)9-(TTTG)4(TG)4 | 178-198 | 76 | 2 | 1.650 | 0 | 0.171 | 0.397 | 0.316 |
| RM455 | 7 | (TTCT)5 | 123-131 | 74 | 2 | 1.513 | 0 | 0 | 0.341 | 0.320 |
| RM495 | 1 | (CTG)7 | 142-156 | 76 | 2 | 1.999 | 0 | 0.079 | 0.503 | 0.375 |
| RM284 | 8 | (GA)8 | 125-148 | 76 | 5 | 1.719 | 1 | 0 | 0.421 | 0.396 |
| RM133 | 6 | (CT)8 | 158-161 | 68 | 2 | 1.882 | 0 | 0.044 | 0.472 | 0.484 |
| RM152 | 8 | (GGC)10 | 130-154 | 76 | 4 | 2.531 | 1 | 1 | 0.609 | 0.530 |
| RM215 | 9 | (CT)16 | 150-156 | 50 | 4 | 2.018 | 2 | 0.08 | 0.509 | 0.601 |
| RM514 | 3 | (AC)12 | 226-256 | 76 | 6 | 2.937 | 3 | 0.039 | 0.664 | 0.605 |
| RM237 | 1 | (CT)18 | 089-099 | 65 | 4 | 2.931 | 1 | 0.015 | 0.664 | 0.681 |
| RM283 | 1 | (GA)18 | 150-166 | 76 | 10 | 4.034 | 6 | 0 | 0.757 | 0.715 |
| RM447 | 8 | (CTT)8 | 100-116 | 76 | 6 | 4.132 | 2 | 0 | 0.763 | 0.718 |
| RM161 | 5 | (AG)20 | 165-190 | 76 | 11 | 4.457 | 7 | 0 | 0.781 | 0.745 |
| RM408 | 8 | (CT)13 | 112-130 | 76 | 12 | 4.310 | 6 | 0 | 0.773 | 0.747 |
| RMOSR13 | 3 | (GA)n | 092-128 | 76 | 13 | 5.096 | 8 | 0.079 | 0.809 | 0.787 |
| RM44 | 8 | (GA)16 | 068-105 | 76 | 12 | 6.670 | 6 | 0 | 0.856 | 0.834 |
| RM162 | 6 | (AC)20 | 151-280 | 76 | 17 | 7.149 | 11 | 0.026 | 0.866 | 0.847 |
| RM154 | 2 | (GA)21 | 128-206 | 76 | 20 | 7.784 | 15 | 0.224 | 0.877 | 0.861 |
| RM277 | 12 | (GA)11 | 114-129 | 76 | 14 | 10.540 | 3 | 0 | 0.911 | 0.897 |
| RM271 | 10 | (GA)15 | 087-156 | 76 | 17 | 10.898 | 10 | 0.039 | 0.914 | 0.901 |
| RM413 | 5 | (AG)11 | 062-121 | 76 | 24 | 18.337 | 15 | 0.132 | 0.952 | 0.943 |
| Total | | | | | 194 | 105.198 | 100 | 1,928 | 14,31 | 13,735 |
| Mean | | | | | 9 | 4.782 | 4.545 | 0.086 | 0.667 | 0.624 |

*Taken from the panel of 50 standard SSR Universal Core Map (Orjuela et al., 2010). Chr, Rice chromosome; n, number of loci; AN, Number of allele per locus; AE, effective allele number; RA, rare allele number; Ho, Observed heterozygosity; He, Gene diversity or expected heterozygosity; PIC, Polymorphism information content

**Data analysis.** The polymorphism of each amplified SSR locus was considered as a unit character and was scored qualitatively in the form of a binary data matrix presented in discrete variables such as 1 and 0 for presence and absence, respectively. Another matrix was prepared using the molecular size of the observed alleles for the genetic analysis considering only the more informative primers for the diversity analysis.

Basic statistics such as major allele frequency, total number of alleles, number of rare alleles, accession-specific alleles, observed and expected heterozygosity and Polymorphic Information Content (PIC) were calculated using GenAlEx 6.41 (Peakall and Smouse, 2006) and the Cervus 3.0 sofware (Kalinowski et al., 2007) to measure diversity for each SSR marker. The analysis of molecular variance (AMOVA) was also performed.

Basing it on the genetic distance matrix between pairs of accessions (using the Rst parameter by GenAlEx 6.41), an unweighted pair group method of arithmetic averages (UPGMA) (Sneath and Sokal, 1973) cluster analysis was used to assess the pattern of diversity among the 76 rice genotypes and a dendrogram was constructed and visualized with the R software.

The model-based program Structure v. 2.3.4 (Pritchard et al., 2000) was used to infer the INIAP's materials structure and estimate

probable accessions groups (K). At least six runs of Structure were conducted by setting the number of groups from 1 to 5. Each run was implemented with a burn-in period of 500,000 steps followed by 150,000 Monte Carlo Markov Chain replicates. Ideal K value was calculated using the online software Structure Harvester v 6.92 (Earl and von Holdt, 2012) and then plotted to find the plateau of the Delta K values.

## RESULTS AND DISCUSSION

**Levels of polymorphisms.** Only 22 (Table 2) of the 30 microsatellites, used to assess the genetic diversity among 76 genetic materials from INIAP Rice Program, showed distinct bands while the rest (RM118, RM124, RM338, RM431, RM433, RM452, RM484, RM507) were monomorphic for these genotypes.

SSR loci diversity data are summarized in Table 2. A total of 194 alleles was detected, and the number of alleles generated per marker ranged from 2 (RM125, RM133, RM316, RM455 and RM495) to 24 (RM413), with an average of 9 alleles per locus. The sizes of the alleles varied between 62 and 280 bp.

The average PIC value was 0.624, varying from 0.202 (RM125) to 0.943 (RM413). Out of 22 primers used for characterization of INIAP´s rice materials, six SSR primers (RM44, RM154, RM162, RM271 and RM413) exhibited the maximum PIC (Table 2). The primers RM125 and RM 536 were found to be the least informative with PIC values equaling 0.202 and 0.23, respectively. In general, the more informative loci were those registering allele numbers above 10, while one of the markers with 6 alleles showed a PIC of 0.718 (RM447) similar to RM283 with 10 alleles and a PIC value of 0.715. PIC is the measure of allelic diversity at a given locus. A hundred rare alleles (RA) (51.5 %), defined as those alleles with a frequency less than 5 %, were identified at 17 of the 22 loci, with an average of 4.5 rare alleles per locus. Rice markers that showed high PIC values, allele number (AN) and rare alleles (RA) were capable of detecting polymorphism and could be used in further studies of rice genetic diversity. Markers RM413 and RM154 showed the highest PIC value, AN and RA per locus. PIC values and rare alleles provide useful information regarding genetic diversity analysis of genotypes. It is important to include rare alleles in a breeding program in order to maximize variation (Yanchuk, 2001).

Average observed heterozygosity was Ho=0.086, with a range of 0 to 1, showing high genetic diversity; one marker (RM152) out of the 22 used exhibited 100 % heterozygous loci (Ho=1), and ten markers (45.5 %) showed homozygous loci. Average expected genetic diversity was He=0.667, while RM125 registered the lowest expected heterozygosity (He=0.230). We found that the allelic richness per locus varied widely and PIC, observed heterozygosity and expected heterozygosity values were relatively high, indicating a high level of genetic diversity

**Genetic distance-based clustering analysis.** A dendrogram constructed using the unweighted pair-group method with arithmetic means (UPGMA) grouped the 76 rice materials in four well differentiated major clusters (Figure 1), in which the oldest INIAP varieties (launched on 1972 to 1990s) comprised a group of eight varieties (INIAP-10 to INIAP-11), showing that they share in their genetic base some ancestors, with agronomical characteristics of high tilling, precocity, and yielding potential. The branch divides in three well defined groups: INIAP-10 remained separated; this material presents dark green foliage, compact plant structure, long panicles with a large grain number and excellent milling quality, while in the other branch INIAP-11, INIAP-14 and INIAP-15 share a node; these varieties show very similar agronomical traits such as precocity, high tilling, good yield, as well as INIAP-16, INIAP-17 and INIAP-18. INIAP-17 and INIAP-18 are siblings, with similar plant type and good grain quality.

More recently released varieties (1990 - present) and eligible breeding lines were placed in the next cluster (18 genotypes: INIAP-FL1480 to GO00918), these materials were developed with a new plant type, with less tilling, good panicle size and weight and large grain number, good milling and cooking qualities. A large cluster containing landrace varieties and several nursery observation lines was also obtained; those lines share phytosanitary tolerance to *Pyricularia* and *Burkholderia*, grain quality (low percentage of white belly in the grain), and high yielding potential including promissory selection lines.

The dendrogram (Figure 1) also shows that landrace varieties allocated in a large group at the opposite end of currently breeding lines (nursery lines). Only GO03497, with characteristics of long grains, and presumable tolerance to bacterial damage (unpublished

data), groups in this cluster. In general, the landraces are tall, susceptible to flattening and shelling, with fast germination (low dormancy) and high adaptability to diverse stress conditions. The control lines used in this study have been at least two decades in the field. They have very different physical appearance among them. Azucena belonged to the biofortification program for iron and zinc; Fedearroz 50 was a great variety used as a progenitor in several Latin America breeding programs; IR64 was an emblematic variety obtained at IRRI, and Fanny is a control line for *Pyricularia* resistance studies.

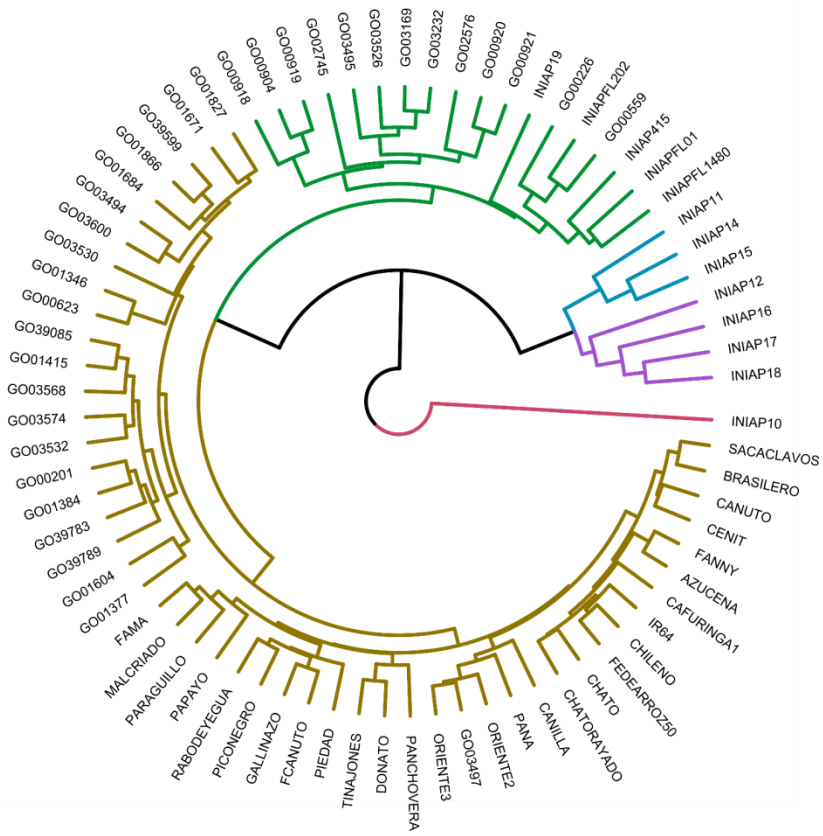


**Figure 1**. UPGMA dendrogram based on polymorphisms of 22 SSR loci showing the genetic among the 76 rice accessions

The analysis of molecular variance (AMOVA) among populations indicated that 15 % of the variation was due to differences among them, with the remaining 73 % due to difference among individuals (data not shown). Around 12 % of the total genetic variance was explained by differences at the individual level with permutation value of 999 for all the studied germplasm.

Bayesian analysis of INIAP Rice Program materials structure using the model-based approach in Structure program without any *a priori* information provided support for the existence of three genetically distinct clusters (K=3) (Figure 2a; 2b). Genotypes were classified into three groups (blue, green and red color bars in Figure 2) which include 26, 21 and 29 rice materials (Figure 2). The first group (blue color) comprises commercial varieties released by INIAP (13 genotypes) and about 37 % of nursery lines (13 genotypes). A second group (green color bars) consists of 49 % of nursery lines (17 genotypes) and the international varieties. In the third group (red color) landrace varieties were included with 14 % of the nursery lines, that is, five genotypes: GO01866, GO39599, GO03494, GO03600 and GO03497. GO03600 is usually employed in crossings because of its tolerance to fungal diseases. The nursery lines GO03494 and GO03497 showed remarkable genetic similarity with the landraces. It may be important to highlight that this latter line is presumable tolerant to biotic stress due to bacteria.

This research for the first time assesses the diversity among the genetic materials managed by the Ecuadorian Rice Program. The results showed that the group of rice commercial

varieties, nursery lines and landrace varieties studied constitute a diverse genetic pool, whose variability is yet be exploited. This knowledge is a key component for germplasm characterization, breeding and conservation planning.

Knowledge regarding the magnitude of the genetic base of Ecuadorian irrigated rice cultivars may help in the selection of divergent parents in breeding programs, contributing to the broadening of the genetic base of cultivars and cultivated species. Rice breeders have the responsibility to exploit this potential in significant ways. For genetic breeding purposes, a narrow genetic base results in low variability, which may limit the possibilities for recombination and genetic segregation, and hinder the achievement of gains from selection.

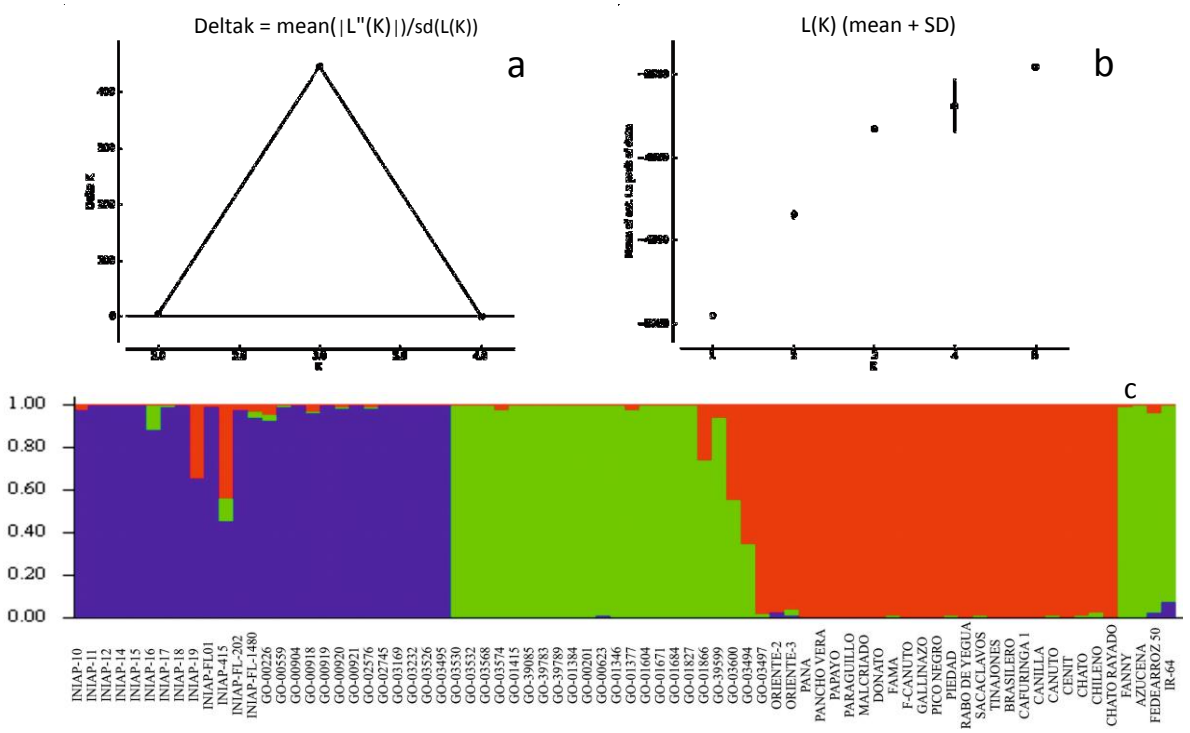


**Figure 2**. a) Variation of values L(K) and b) Delta K derived ΔK for K from 1 to 5, for 76 rice materials, with bayesian calculations without *a priori* information on the origin of genotypes based on SSR markers. c) Results of grouping for K = 3, obtained without *a priori* information from Bayesian analysis. Each individual is represented by a single vertical line broken into K (K=3) colored segments, with lengths proportional to each of the K inferred clusters. Lateral numbers indicate genetic proportion of each grouping. Genotypes are differentially colored according to the analysis results

Landraces constitute a vast genetic resource for breeding purposes. Farmers in Ecuador have planted them for decades, although their origins are unknown. Unfortunately, regardless their adaptation, productivity and quality, until now their use in the Ecuadorian breeding program have been limited. Their plant type differs to that of the green revolution (short stems, erected leaves) generated during the sixties by the International Rice Research Institute (IRRI). It is important to characterize these traditional varieties to avoid genetic erosion produced through introduction and planting of foreigner rice material. It would be useful to employ them because of their productivity, grain quality, and great adaptation, among others; planting them directly in specific recommended areas or breeding with other genetic materials to obtain enough variability for selecting superior genotypes. In this way the breeding program could deliver new cultivars to the Ecuadorian rice sector, with less supplies requirements and adapted to the national production systems, basically avoiding the loss of existing landraces by using this germplasm in future research to generate technology that provides solution to problems of low productivity, intensive use of supplies, low grain quality, low rentability.

Several studies of genetic diversity in Latin America have estimated a narrow genetic base among rice materials. Cuevas et al. (1992) and Montalvan et al. (1998) presented results on irrigated and upland rice varieties in Latin America, respectively, indicating that commercial varieties released for both systems had a narrow genetic base. A low coancestry coefficient among the Argentinian materials compared to cultivars from other Latin American countries was confirmed more recently by Giarrocco et al. (2007). Brazilian cultivars also presented a narrow genetic base, according to Rangel et al. (1996) and Guimaraes et al. (1996). Venezuelan varieties presented an average coancestry coefficient of 0.16 (Cuevas et al., 1992) and were similar to Colombian germplasm. Low genetic diversity found among the Venezuelan accessions evidenced the narrow genetic bases used in breeding programs in the region, later confirmed by Ghneim et al. (2008) and Pérez et al. (2011).

A recent study among varieties released by the Latin American Fund for Irrigated Rice (FLAR) in thirteen Latin American countries and carried out between 2003-2014 (Berrio et al., 2016), revealed that average coancestry values equalled 0.19. Country values varied reporting a minimum for Ecuador ($r_{xy}$ = 0.13) and maximum for Venezuelan varieties ($r_{xy}$ = 0.31). Coancestry values ($r_{xy}$) measure genetic diversity, where $r_{xy}$ = 0 is the maximum diversity. Our data generated a large cluster containing landrace varieties and several nursery observation lines showing that the Ecuadorian rice varieties possesses variability that can yet be used by the breeder.

We were able to generate a fingerprinting for identifying the INIAP´s varieties based upon eight microsatellites (RM154, RM413, RM162, RM447, RM408, RM283, RM237 and RM277) with unique alleles in these materials (Table 3). These markers produced 17 genotype-specific alleles that distinguished the thirteen Ecuadorian rice cultivars (INIAP-10, INIAP-11, INIAP-12, INIAP-14, INIAP-15, INIAP-16, INIAP-17, INIAP-18, INIAP-19, INIAP-FL01, INIAP-415, INIAP-FL202 and INIAP-FL1480).

**Table 3**. Genotype-specific alleles identified for 13 rice cultivars released by the INIAP Program (1968 -2017) that were analyzed in this study

| Locus | Alelle | Size (bp) | Variety |
|---|---|---|---|
| RM154 | 1 | 187 | INIAP-10, INIAP-11, INIAP-15 |
| RM162 | 1 | 280 | INIAP-10, INIAP-11 |
| | 2 | 151 | INIAP-10 |
| RM237 | 1 | 95 | INIAP-FL0202 |
| RM283 | 1 | 161 | INIAP-15 |
| | 2 | 159 | INIAP-19 |
| | 3 | 157 | INIAP-FL0202 |
| RM277 | 1 | 129 | INIAP415 |
| | 2 | 125 | INIAP-FL1480 |
| RM408 | 1 | 130 | INIAP-14 |
| | 2 | 127 | INIAP-10, INIAP-11, INIAP-12, INIAP-15 |
| | 3 | 125 | INIAP-FL01 |
| | 4 | 124 | INIAP-16 |
| RM413 | 1 | 121 | INIAP-11 |
| | 2 | 95 | INIAP-10, INIAP-12, INIAP-18 |
| | 3 | 80 | INIAP-15, INIAP-14 |
| RM447 | 1 | 102 | INIAP-11, INIAP-14, INIAP-15 |

The set of microsatellite markers used in the present study provided a positive assessment of the ability of SSR markers to produce unique DNA profiles and establish discrete identities of rice genotypes, which otherwise, would have not been possible using morphological traits. We detected genotype-specific markers that can be used for the molecular identification/characterization of the Ecuadorian germplasm (Table 3). For example, SSR 154 allele 1 distinguishes INIAP-10, INIAP-11 and INIAP-15 varieties of the rest. INIAP-11 presents a unique band at 121 bp with marker RM413. We would like to underline that two recently released varieties each presented a unique allele, that is to say INIAPFL1480 amplifies an 80 bp band with RM413 and a 125 bp with RM277, whereas INIAPFL0202 showed

a fragment of 157 bp with marker RM283. It would be interesting to develop a protocol reaction to amplify simultaneously different microsatellite markers standardizing melting temperatures and expected base pairs in the same PCR mix, to discriminate in few steps these genetic lines.

The results from this study would be useful when planning breeding strategies to reduce the genetic vulnerability of crops and for conservation genetics. This will also help breeding programs to plan crosses to incorporate this variability into the genetic background of elite rice germplasm, which in turn will generate new rice cultivars with more efficient traits. Having genetic diversity available allows breeders the possibility of developing useful products that would have an impact at farmers' field level.

## CONCLUSION

Useful genetic parameters were determined showing variability in the set of varieties currently under breeding and establishing that there is diversity among the rice materials in the main groups that can still be used by the breeder. Our results provide preliminary insight into the degree of diversity among the rice germplasm managed by the Ecuadorian Rice Program.

## LITERATURE CITED

1. Aguirre, C., R. Alvarado and P. Hinrichsen. 2005. Identificación de cultivares y líneas de mejoramiento de arroz de Chile mediante amplificación de fragmentos polimórficos (AFLP). Agricultura Técnica 65(4): 356-369.
2. Berrio, L.E, E.A Torres, J. Barona and J.B. Cuasquer. 2016. Diversidad genética de las variedades de arroz FLAR liberadas entre 2003-2014. Agrononomia Mesoamericana 27(2): 217-231.
3. Cuevas-Pérez, F., E.P. Guimaraes, L.E. Berrio and D. Gonzales. 1992. Genetic base of the irrigated rice in Latin America and the Caribbean. Crop Science 32(4): 1054-1059.
4. Day-Rubenstein, K.A., P. Heisey, R.A. Shoemaker, J. Sullivan and G. Frisvold. 2005. Crop genetic resources: An economic appraisal. Washington: USDA Economic Information Bulletin http://ageconsearch.umn.edu/record/59388/files/eib2.pdf (retrieved on Jul 30, 2018).
5. Earl D.A. and B.M. von Holdt. 2012. Structure Harvester: a website and program for visualizing Structure output and implementing the Evanno method. Conservation Genetics Resources 4(2): 359-361.
6. FAOSTAT (Statistics Division Food and Agriculture Organization of the United Nations). 2018. http://www.fao.org/faostat/en/#data (retrieved on May 5, 2018).
7. Ferreira M.E and D. Grattapaglia. 1998. Introduction to the use of molecular markers in the genetic analysis. EMBRAPA, Brasil, 220 p.
8. Fuentes, J., F. Escobar, A. Alvarez, G. Gallego, M.C. Duque, M. Ferrer, J. Deus and J. Tohme. 1999. Analyses of genetic diversity in Cuban rice varieties using isozyme, RAPD and AFLP markers. Euphytica 109(2): 107-115.
9. Garris, A.J., T.H. Tai, J. Coburn, S. Kresovich and S.R. McCouch. 2005. Genetic structure and diversity in *Oryza sativa* L. Genetics 169(3): 1631-1638.
10. Ghneim, T., D. Posso, I. Pérez-Almeida, G. Torrealba, A. Pieters, C. Martinez and J. Tohme. 2008. Assessment of genetic diversity in Venezuelan rice cultivars using simple sequence repeats markers. Electronic Journal of Biotechnology 11(5): 3-4.
11. Giarrocco L.E., M.A. Marassi and G.L. Salerno. 2007. Assessment of genetic diversity in Argentine rice cultivars with SSR markers. Crop Science 47(2): 853-860.
12. Guimaraes E.P., J. Borrero and Y. Ospina-Rey. 1996. Genetic diversity of upland rice germplasm distributed in Latin America. Pesquisa Agropecuaria Brasileira 31(3): 187-194.
13. Keneni, G., E. Bekele, M. Imtiaz and K. Dagne. 2012. Genetic vulnerability of modern crop cultivars: causes, mechanism and remedies. International Journal of Plant Research 2(3): 69-79.
14. MAG (Ministerio de Agricultura y Ganadería). 2015. Boletín situacional de arroz. Coordinación General del Sistema de Información Nacional. Quito. 6 p.
15. Montalvan, R., D. Destro, E.F. Silva and J.C. Montano. 1998. Genetic base of Brazilian

upland rice cultivars. Journal of Genetics & Breeding 52(1): 203-209.

16. Orjuela, J., A. Garavito, M. Bonioul, J. Arbelaez, L. Moreno, J. Kimball, G. Wilson, J. Rami, J. Tohme, S. McCouch and M. Lorieux. 2010. A universal core genetic map for rice. Theoretical and Applied Genetics 120(3): 563-572.
17. Parikh, M., K. Motiramani, N. Rastogi and B. Sharma. 2012. Agro-morphological characterization and assessment of variability in aromatic rice germplasm. Bangladesh Journal of Agricultural Research 37(1): 1-8.
18. Peakall, R. and P.E. Smouse. 2006. Genalex 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6(1): 288-295.
19. Pérez-Almeida, I., E. Torres, L. Angulo and M. Acevedo. 2011. Genetic diversity among Venezuelan rice cultivars based on parentage coefficient estimation and analysis using microsatellite molecular markers (SSR). Interciencia 36(7): 545-551.
20. Pritchard, J.K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2): 945-959.
21. Rangel, P.H., E.P. Guimaraes and P.C. Neves. 1996. The genetic base of Brazilian irrigated rice (*Oryza sativa* L.) cultivars. Pesquisa Agropecuaria Brasileira 31(2): 349-357.
22. Roy, S., A. Banerjee, B. Mawkhlieng, A.K. Misra, A. Pattanayak, G.D. Harish et al. 2015. Genetic diversity and population structure in aromatic and quality rice (*Oryza sativa* L.) landraces from North Eastern India. PLoS ONE 10(6): e0129607.
23. Sneath, P.H. and R.R. Sokal. 1973. Numerical taxonomy: The principles and practice of numerical classification. W.H. Freeman. San Francisco, USA. 573 p.
24. Temnykh, S., W. Park, N. Ayres, S. Cartinhour, N. Hauck, L. Lipovich and S.R. McCouch. 2002. Mapping and genome organization of microsatellites sequences in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 100(5): 697-712.
25. Travis, A., G. Norton, S. Datta, R. Sarma, T. Dasgupta, F. Savio, M. Macaulay, P. Hedley, K. McNally, M. Sumon, M. Islam and A. Price. 2015. Assessing the genetic diversity of rice originating from Bangladesh, Assam and West Bengal. Rice 8(1): 35.
26. Wei, X., X. Yuan, H. Yu, Y. Wang, Q. Xu and S. Tang. 2009. Temporal changes in SSR allelic diversity of major rice cultivars in China. Journal of Genetics & Genomics 36(6): 363-370.
27. Xu, Q., X. Yuan, Li, X., Wang, S., Feng, Y., Yu, H., Wang, Y., Y. Yang and X. Wei. 2016. The genetic diversity and structure of *Indica* rice in China as detected by single nucleotide polymorphism analysis. BMC Genetics 17: 53.
28. Yanchuk, A.D. 2001. A quantitative framework for breeding and conservation of forest tree genetic resources in British Columbia. Canadian Journal of Forest Research 31(4): 566-576.
29. Yu, J., S. Hu, J. Wang, G. Wong, S. Li, B. Liu, et al. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296 (5565): 79-92.