Plant Cell Biotechnology and Molecular Biology 17(7&8):267-278; 2016 ISSN: 0972-2025

GENETIC DIVERSITY, ANCESTRY RELATIONSHIPS AND CONSENSUS AMONG PHENOTYPE AND GENOTYPE IN BANANA (*Musa acuminata*) CLONES FROM FORMOSA (ARGENTINA) FARMERS

J. L. ERMINI*, G. TENAGLIA AND G. R. PRATTA

Consejo Nacional de Investigaciones Científicas y Técnica (CONICET), Cátedra de Genética, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario (UNR), CC14 S2125ZAA Zavalla, Provincia de Santa Fe, Argentina [JLE, GRP]

Instituto de Investigación y Desarrollo Tecnológico para la Pequeña Agricultura Familiar Región NEA (IPAF), Ruta Prov. Nº 8, Paraje Isla Puén S/N (3611), Laguna Naineck, Provincia de Formosa, Argentina [GT]

[*For Correspondence: E-mail: joseluis.ermini@unr.edu.ar]

ABSTRACT

Though banana is a tropical crop, the Argentinean Northeastern Province of Formosa has extensively promoted its production. Genetic variation was assessed by the Amplified Fragment Length Polymorphism (AFLP) in a set of allotriploid and autotriploid clones from Formosa farms. Clones were also evaluated for perimeter and height of the pseudo-stem, number of leaves before flowering, number of flower clusters and length of fruits. Four international autotriploid varieties were used as samples. An AMOVA test widely discriminated among allotriploid clones, autotriploid clones and check varieties. The ancestry relationships between the autotriploids genotypes verified that clones were derived from the check varieties. Principal Coordinates Analysis was applied to assess the genetic diversity, demonstrating that 45% of the total molecular variation was explained by the three first principal coordinates. The first two principal components explained 77% of total phenotypic variability according to a Principal Components Analysis with the mean phenotypic values of clones and check varieties. Procrustes Analysis verified a high consensus (71.3%) among phenotypic and genotypic characterizations, suggesting that putative associations could be found among both sets of data. The 79.6% of the total variation was explained by the two principal components in this Procrustes Analysis.

Keywords: Allotriploid; autotriploid; amplified fragment length polymorphism (AFLP); molecular characterization; multivariate analysis; procrustes analysis.

INTRODUCTION

Banana (*Musa* genus) is one of the most popular fruits all over the world. It is the first fruit crop in quantity and the second, after the citric family, in economic value (Galan Sauco, 1992). In Argentina, banana production is an important activity in Formosa province. As a polyploid species, it presents a high level of sterility, being mainly multiplied by asexual propagation. Though it is a typical tropical crop, Formosa province – located in the Argentinean subtropical Northeastern– has promoted banana production since the ’60, which hence acquired a great local economic impact. During adaption of banana to this marginal region mainly by natural selection for abiotic and biotic factors and also by artificial selection intuitively practiced by farmers for fruit quality and other agronomical desired traits in their productive environment, a wide biodiversity whose molecular basis is completely unknown should have been originated in Formosa. This genetic variation largely generated by spontaneous mutation is essential for present and future banana breeding, biotechnology and industry demands.

Many authors (Al-Saady, Al-Lawati, Al-Subhi, & Khan, 2010; Wang, Chiang, Roux, Hao, & Ge, 2007; Wong, y otros, 2001) have previously reported the diversity and relationships among

Musa genus by different molecular markers. AFLP analysis is an interesting approach due to the generation of a high number of selectively amplified DNA bands that cover a large proportion of the genome under study (Noyer , Causse, Tomekpe, Bouet, & Baurens, 2005). The AFLP technique has advantages over other molecular based techniques for DNA fingerprinting including Restriction Fragment Length Polymorphism (RFLP) and Random Amplified Polymorphic DNA (RAPD). For example, AFLP technique can be used for DNA samples of any origin or complexity, and small sequence variations can be detected using only small quantities of genomic DNA (0.05–0.5 µg). The capacity to generate many polymorphic bands in one lane is its major advantage because a greater number of *loci* may be simultaneously revealed by AFLP than by other currently available PCR-based techniques. Therefore polymorphism detected per reaction is much higher. AFLP is superior in terms of the number of sequences amplified per reaction and their reproducibility. The markers produced are reliable and reproducible within and between laboratories, and are relatively easy and inexpensive to generate (Blears, De Grandis, Lee, & Trevors, 1998).

The general objective of this work was to use AFLP technique to assess the genetic diversity in 52 banana clones collected in Formosa from different farmers' fields, which was the first reported molecular characterization of this crop in Argentina. The specific aims were to characterize banana clones at two levels of genetic diversity (molecular and phenotypic), to estimate the ancestry relationships of autotriploids genotypes according to the molecular polymorphism detected by AFLP characterization, and to detect putative associations among molecular and phenotypic variations in this unique germplasm of banana adapted to subtropical environmental conditions.

MATERIALS AND METHODS

Plant Material and Data Generation

Banana clones were grown in the northeastern of Argentina, a productive banana zone located at 24º59'S, 58º51'W in Formosa province. Some of the clones were autotriploid (AAA genome) belonging to the Cavendish subgroup (Simmonds & Weatherup, 1990). The autotriploid genotypes included 4 common international varieties of banana (Williams, Jaffa, Grand Naine, and Gal) used as experimental checks. Other clones were allotriploid (AAB genome). List of genotypes assayed is provided in the table (Online Resource 1). Phenotypic traits (perimeter and height of the pseudo-stem, number of leaves before flowering, number of hands -clusters of bananas by inflorescence-, length of finger -individual fruit-, and bunch weight) were evaluated at the farmers' fields.

Molecular characterization was made in the *In vitro* Culture and Molecular Biology Lab of the Cátedra de Genética, Facultad de Ciencias Agrarias (UNR), located at Zavalla, Santa Fe Province (33º01'S, 60º88'W), Argentina. Genomic DNA from 60 mg of young leaves was extracted by commercial kit (Wizard ® Genomic DNA Purification Kit, Promega). Leaves were collected at farmers' field and conserved in –80 °C and DNA was extracted in duplicate to increase reliability of the experiments. The extracted DNA was visualized on 1% agarose gel to verify the quality and quantity in comparison to λ phage. Protocol reported by Blears et al. (1998) was followed with minor modifications. The genomic DNA (30 µL) was digested with 6 units of EcoRI and 6 units of MseI in a final volume of 40 µL incubated at 37◦C for 2 h. Ligation of the digested fragments to the specific universal EcoRI and MseI adapters (EcoRI adapter, final concentration 0.25 µM and MseI adapter, final concentration 0.25 µM) was achieved by incubation at 37°C for 2 h with 3.6 units of T4 ligase in a final volume of 10 µL. The preamplification samples were prepared using 1 µL of the digested-ligated DNA plus 75 ng of each primer+1 (EcoRI+1: 5'-GACTGCGTACCAATTCA-3' and MseI+1: 5'-GATGAGTCCTGAGTAAC-3'), 0.2 mM dNTPs and 1 unit of Taq polymerase in a final volume of 25 µL. The PCR conditions were 30 cycles of 30 s at 94°C, 1 min at 56°C and 1 min at 72°C each one. Once preamplified, the solution was 1:10 diluted in sterile water. The selective amplification samples were prepared using 3 µL of this dilution and the same mix used for preamplification, though six primer+3 combinations reported by Ermini, Tenaglia, & Pratta (2013) were used in this step (Table 1).

Table 1. Combinations of EcoRI and MseI primer +3 selected by Ermini, Tenaglia, & Pratta (2013) and used for selective amplification in this experiment

Combination code	EcoRI primer sequence	Primer code	MseI primer sequence	Primer code
B	5'-GACTGCGTACCAATTCACT-3'	E38	5'-GATGAGTCCTGAGTAACAG-3'	M32
H	5'-GACTGCGTACCAATTCAGA-3'	E39	5'-GATGAGTCCTGAGTAACAG-3'	M32
J	5'-GACTGCGTACCAATTCAGA-3'	E39	5'-GATGAGTCCTGAGTAACTA-3'	M34
Q	5'-GACTGCGTACCAATTCAGC-3'	E40	5'-GATGAGTCCTGAGTAACTG-3'	M35
R	5'-GACTGCGTACCAATTCAGC-3'	E40	5'-GATGAGTCCTGAGTAACGT-3'	M36
3	5'-GACTGCGTACCAATTCATG-3'	E45	5'-GATGAGTCCTGAGTAACTG-3'	M35

Also, a touchdown PCR was programmed, starting with a cycle of 30 s at 94°C, 30 s at 65°C and 1 min at 72°C. During the next 10 cycles, the annealing temperature was reduced to 1°C per cycle, until reaching 56°C. Then, 23 cycles of 30 s at 94°C, 1 min at 56°C and 1 min at 72°C each were repeated. The amplified fragments were boiled for 5 min and loaded into a 6% poly-acrylamide denaturing gel. The molecular weight marker CincuentaMarker (®Biodynamics SRL) was loaded in each gel, electrophoretic separation was done for 2.5 hs at 1900 V. Then, the gel was stained with the silver nitrate technique.

Data Analysis

For molecular characterization, the AFLP profiles were first compared among genotypes for each primer combination. Total number of amplified fragments, number of polymorphic fragments and polymorphism percentage were calculated, the polymorphism being expressed as the presence or absence of a given fragment. Polymorphic fragments present in a given genotype were assigned 1 and those absent were assigned 0. Genetic diversity was estimated by an AMOVA test with allotriploid clones, autotriploid clones and check varieties as source of variation, and the Jaccard's distance (Jaccard, 1908) calculated from the binary matrix of 1 / 0 as dependent variables. This binary matrix was also used to perform an ancestry relationships analysis among autotriploid genotypes (clones and check varieties), and Dice coefficient (Dice & Goldberg, 1975) was chosen to find associations among the whole set of allotriploid and autotriploid bananas by a Principal Coordinates Analysis.

Phenotypic data were compared by a Principal Component Analysis in order to detect the traits greatly contributing to total variation. Finally, consensus among genotype (molecular variation) and phenotype (agronomic quantitative traits) was measured by a Procrustes Analysis (Gower, 1975), which summarizes and gets the better adjustment for the information provided by both Principal Coordinates and Principal Components Analyses. The Procrustes Analysis were also generated by molecular and phenotypic data. The mantel test was made to estimate linear correlation between the both set of array. All statistical analyses were made with Info-gen Software® (Balzarini & Di Rienzo, 2012).

RESULTS

Estimation of Genetic Diversity

Selective amplifications of banana clones with the six primer combinations revealed a total of 85 reliable fragments, 50 of them were discrepant among clones which resulted in a 72% of total polymorphism among this allotriploid and autotriploid set of genotypes. Online resource 2 shows the AFLP profiles generated by Q combination of primer+3 in a sample of representative clones. AMOVA test with Jaccard's distances calculated with polymorphic AFLP as dependent variables detected significant within (33%) and between (67%) variance when comparing allotriploid clones, autotriploid clones and check varieties. Total variance had a value of 7.55, the within component being 5.05 and the between component, 2.50. The Phi coefficient of AMOVA was 0.33.

Ancestry Relationships by Molecular Polymorphism for Autotriploid Genotypes

The ancestry relationships among autotriploid genotypes (clones and check varieties) were estimated considered the variety Williams as the

common ancestor, as it was proposed by Novak et al. (1990) Fig. 1 shows that the four check varieties were close in the basis of the diagram, confirming their narrow co-ancestry. Just one of the clones collected at farmers' field was near to the check varieties. The other clones were located far away, indicating that they related but genetically diverse. The allotriploid clones were not included in this analysis because they have a different genetic origin than the autotriploid ones.

Molecular and Phenotypic Characterization and Consensus among Both Sets of Data

The Principal Coordinates biplot obtained using Dice´s distances is shown in Fig. 2. The 45% of the total molecular variation was explained by the three first principal coordinates (25.1% by the first, 11.5% by the second and 8% by the third, as shown in the Table (Online Resource 3). Associations among genotypes according to molecular variation showed a clear differentiation between autotriploid and allotriploid clones, with the four check varieties being also distinguished in a third group upper located in Fig. 2.

Fig. 3 shows the biplot from the Principal Components Analysis after the mean values of phenotypic traits. The biplot had a cophenetic correlation of 0.96, indicating a good adjustment of results to original data. A clear discrimination among the group of check varieties with some autotriploid clones and the other clones (including allotriploid and autotriploid) was achieved. This multivariate analysis showed that the first two principal components (PC) explained 77% of total phenotypic variability, PC1 and PC2 accounting for 43% and 34%, respectively as shown in the Table (Online Resource 4). The traits having major relative contributions to PC1 were perimeter of the pseudo-stem, number of hands, and bunch weight. All traits had positive contribution to PC1 (Table 2). For PC2, contributions were either positive or negative. Traits mostly correlated to PC2 were height of pseudo-stem, number of leaves before flowering and length of fingers, whose coefficient was negative (Table 2).

Table 2. Composition of the first (CP1) and the second (PC2) principal components. Composition of the first (CP1) and the second (PC2) principal components according to each trait contribution (TC) and correlation coefficient (CC) among each trait and the corresponding principal component

Trait	PC1		PC2	
	TC	CC	TC	CC
Perimeter of the pseudo-stem	0.49	0.78	0.21	0.31
Height of the pseudo-stem	0.29	0.46	0.56	0.83
Number of leaves before flowering	0.13	0.21	0.59	0.87
Number of hands	0.52	0.83	-0.09	-0.13
Length of fingers	0.32	0.51	-0.46	-0.77
Bunch weight	0.53	0.85	-0.28	-0.42

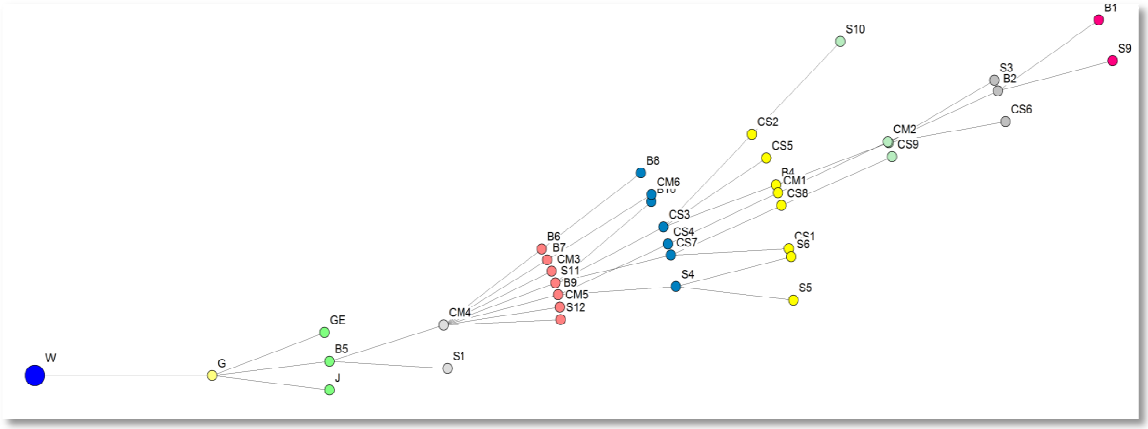


Fig. 1. Ancestry relationships among autotriploid genotypes showed by analysis of minimum spanning tree (prim algorithm)

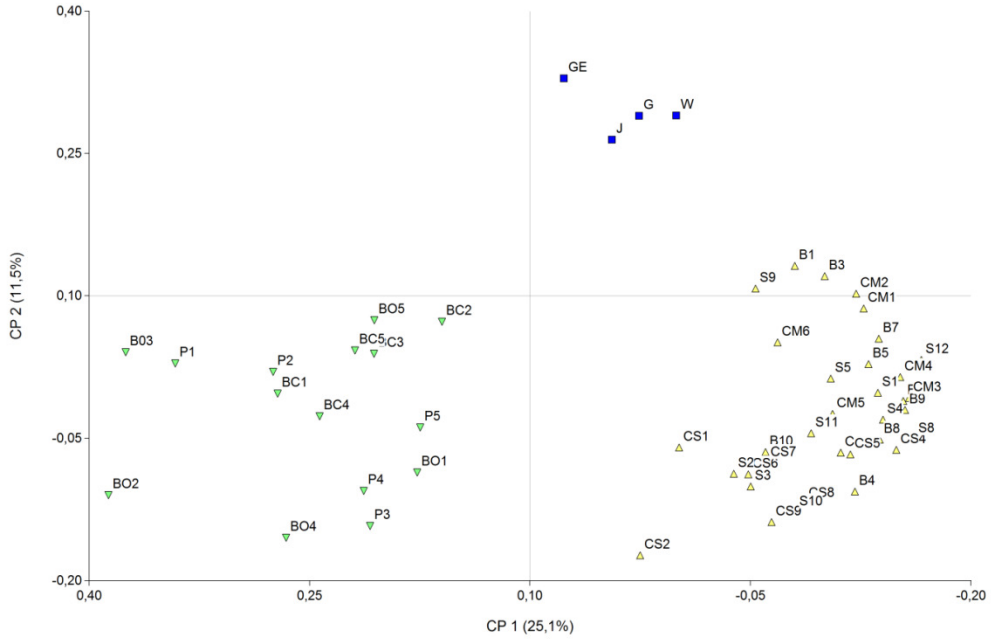


Fig. 2. Position of banana allotriploid and autotriploid clones and check genotypes in the biplot generated by the first (PCo1) and the second (PCo2) Principal Coordinates calculated from AFLP profiles with Dice coefficient as distance metrics

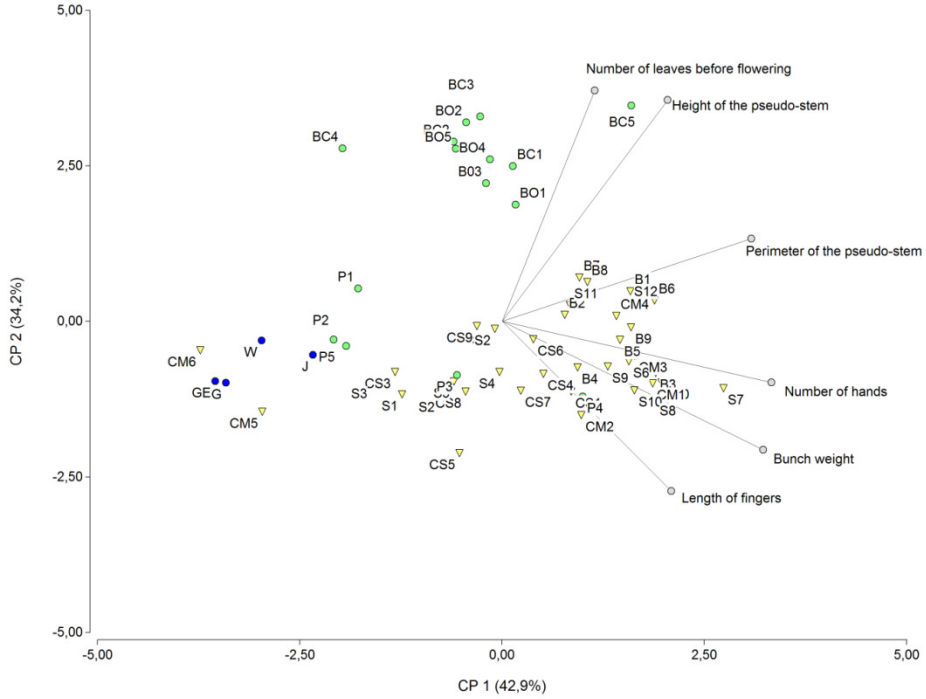


Fig. 3. Position of banana allotriploid and autotriploid clones and check genotypes in the biplot generated by the first (PC1) and the second (PC2) Principal Components calculated from mean values of phenotypic traits (also located in the biplot)

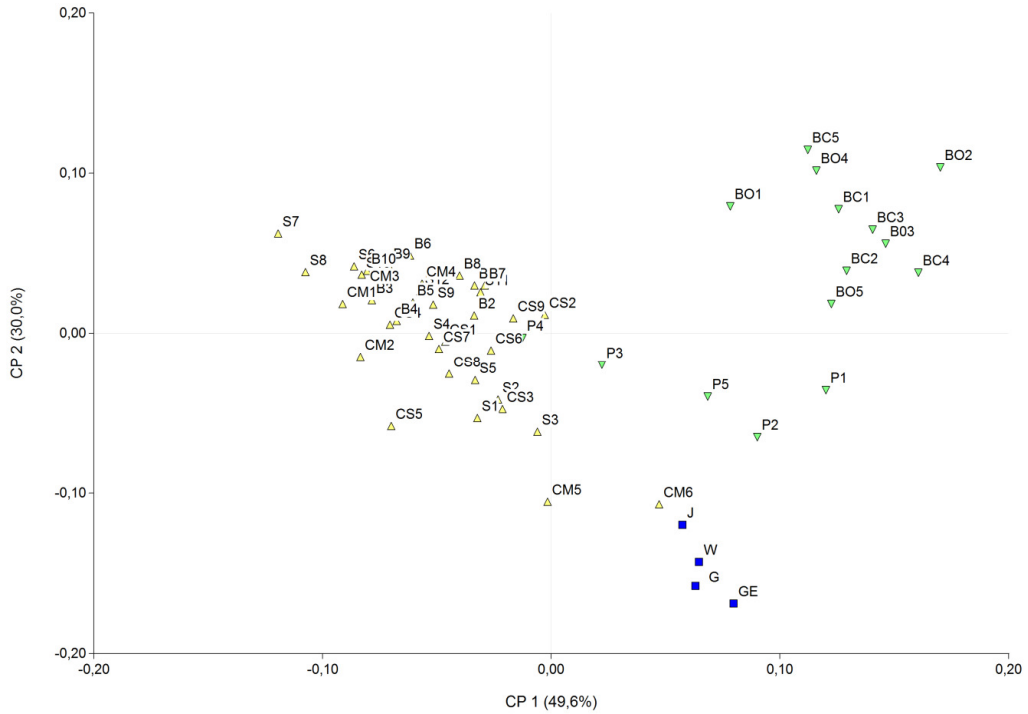


Fig. 4. Consensus position among phenotypic and genotypic data of banana allotriploid and autotriploid clones and check genotypes in the biplot generated by the first (PC1) and the second (PC2) principal components calculated by Procrustes analysis

Location of banana allotriploid and autotriploid clones and check varieties in the biplot according to the Procrustes Analysis are presented in Fig. 4. The first two principal components obtained by this multivariate analysis jointly explained the 79.6% (49.6% for CP1 and 30.0% for CP2) of the total variation; additional data are given in the Table (Online Resource 5). The total proportion of consensus was 71.3%, ranging from 38.9% to 92.8%. This high proportion of consensus indicated significant associations among molecular and phenotypic diversity. The Procrustes Analysis by both set of data (molecular and phenotypic) shows a comparable grouping that Fig. 4 (Online Resource 6). Mantel test is also show a significant linear correlation of 0.23 between the phenotypic and molecular arrays (Online Resource 7).

DISCUSSION

Assessing genetic diversity of *Musa* spp. provides scientific basis for the identification and evaluation of genetic variability. The percentage of polymorphism detected in this work was similar to previous report of other authors (Noyer , Causse, Tomekpe, Bouet, & Baurens, 2005; Al-Saady, Al-Lawati , Al-Subhi , & Khan, 2010) but larger to that obtained when characterizing only the 4 check varieties (16% of polymorphism) with the aim of validating the AFLP protocol in local conditions and selecting the more appropriate primer+3 combinations (Ermini, Tenaglia, & Pratta, 2013). In this work, selected primer combinations were applied to genotypes collected at farmer fields, allowing us to explain the greatest polymorphism by the addition of biodiversity (autotriploid and allotriploid clones) and by eventual spontaneous mutation followed by both natural and artificial selection at farmers field for adaptation to Formosa subtropical conditions and for productive traits, respectively. Results from AMOVA test were consistent with previous reports (Wang, Chiang, Roux, Hao, & Ge, 2007) and clearly verified the effect of broadening genetic variability by including the collected

clones, since between molecular variance is larger than within variance.

The ancestry relationships among autotriploid genotypes indicated that all clones were derived from at least two check varieties (Williams and Gal varieties) and agreed to AMOVA results because they also suggest that accumulation of different gene variations during cultivation in framers' field was followed by natural and artificial selection that resulted in a great genetic homogeneity among autotriploid clones, which are noticeably different from the autotriploid check varieties from which they were derived.

Differentiation among allotriploid clones, autotriploid clones and check varieties was also shown by Principal Coordinates Analysis given that position of these plant germplasm in Fig. 2 reflected the molecular variation detected by AFLP bands (Al-Saady, Al-Lawati , Al-Subhi , & Khan, 2010). Instead, Principal Component Analysis (Fig. 3) performed with phenotypic data indicated different associations among this germplasm, which could be due to the intuitive artificial selection by farmers previously mentioned that favored the expression of desired agronomic traits. Contributions of each trait to the first and the second Principal Components revealed that important productive attributes (bunch weight, number of hand, length of finger) were involved in determining phenotypic variability. These traits would have certainly been under artificial selection at farmers' field to increase crop yield. The relatively high proportion of consensus (0.713) of the Procrustes Analyses indicates that, though positions of banana genotypes according to molecular and phenotypic variability appears to be different in the respective biplots, there were some noticeable associations among both sets of data. The Procrustes made from de data and the Mantel test (Sokal, 1979) support the result from the Procrustes made from synthetic variables. In fact, positions of genotypes in the biplot generated by Procrustes Analysis (Fig. 4) are the results of simultaneously considering molecular and phenotypic diversity, and suggest that in this unique banana germplasm, a high proportion of molecular variation revealed by AFLP is involved in the genetic determination of phenotypic traits favored by both natural and artificial selection in farmer's fields (Mahuad, Pratta, Rodriguez, Zorzoli, & Picardi, 2013).

CONCLUSION

In concluding, a high level of phenotypic and molecular polymorphism was detected among autotriploid and allotriploid clones. The origin of the autotriploid clones from the check varieties was verified, and the importance of mutation and selection in a subtropical region as Formosa Province for generating genetic diversity was evidenced. Associations among agronomic data and molecular polymorphisms could be used as a first approach to identify genome regions involved in the expression of superior phenotypes in this unique genetic banana material.

ACKNOWLEDGEMENT

This work was supported by the Centro de Validación de Tecnologías Agropecuarias (CEDEVA) de Misión Tacaaglé, Formosa., and Fundación Ciencias Agrarias (NUR, PICT 2014).

AUTHORS' CONTRIBUTIONS

This work was carried out in collaboration between all authors. Author JLE designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Authors GT and GRP managed the analyses of the study. Author GRP managed the literature searches and also designed the study. All authors read and approved the final manuscript.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

Al-Saady, N.A., Al-Lawati, A.H., Al-Subhi, A.M., Khan, A.J. 2010. Evaluation of genetic diversity in Omani banana cultivars (Musa cvs.) using AFLP markers. Journal of Plant Sciences. **5**(4): 404-413.

Balzarini, M.G., Di Rienzo, J.A. 2012. InfoGen Software. Cordoba, Cordoba, Argentina. Obtenido de http://www.info-gen.com.ar/: http://www.info-gen.com.ar/

Blears, M.J., De Grandis, S.A., Lee, H., Trevors, J.T. 1998. Amplified fragment length polymorphism (AFLP): A review of the procedure and its applications. Journal of Industrial Microbiology and Biotechnology. **21**(3):99-114.

Dice, J.F., Goldberg, A.L. 1975. Relationship between *in vivo* degradative rates and isoelectric points of proteins. Proceedings of the National Academy of Sciences. **72**(10):3893-3897.

Ermini, J.L., Tenaglia, G., Pratta, G.R. 2013. Marcadores de AFLP en el cultivo de banana: selección de combinación de cebadores y caracterización de la biodiversidad. Revista de la Facultad de Ciencias Exactas, Químicas y Naturales-Universidad de Morón. **11:**93-109.

Galan Sauco, V. 1992. Los frutales tropicales en los subtrópicos. II Plátano (Banano). Madrid, España: Mundi prensa.

Gower, J.C. 1975. Generalized procrustes analysis. Psychometrika. **40**(1):33-51.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. 44.

Mahuad, S.L., Pratta, G.R., Rodriguez, G.R., Zorzoli, R., Picardi, L.A. 2013. Preservation of Solanum pimpinellifolium genomic fragments in recombinant genotypes improved the fruit quality of tomato. Journal of Genetics. **92**(2): 195-203.

Novak, F.J., Afza, R., Duren, M.V., Omar, M.S. 1990. Mutation induction by gamma irradiation of in vitro cultured shoot-tips of banana and plantain (Musa cvs). Tropical Agriculture. **67**(1):21-28.

Noyer, J.L., Causse, S., Tomekpe, K., Bouet, A., Baurens, F.C. 2005. A new image of plantain diversity assessed by SSR, AFLP and MSAP markers. Genetica. **124**(1):61-69.

Simmonds, N.W., Weatherup, S.T. 1990. Numerical taxonomy of the wild bananas (Musa). New Phytologist. **115**(3):567-571.

Sokal, R.R. 1979. Testing statistical significance of geographic variation patterns. Systematic Zoology. **28**(2):227-232.

Wang, X.L., Chiang, T.Y., Roux, N., Hao, G., Ge, X.J. 2007. Genetic diversity of wild banana (*Musa balbisiana* Colla) in China as revealed by AFLP markers. Genetic Resources and Crop Evolution. **54**(5):1125-1132.

Wong, C., Kiew, R., Loh, J.P., Gan, L.H., Set, O., Lee, S., Gan, Y.Y. 2001. Genetic diversity of the wild banana Musa acuminata Colla in Malaysia as evidenced by AFLP. Annals of Botany. **88**(6):1017-1025.

APPENDIX 1

Legends to figures

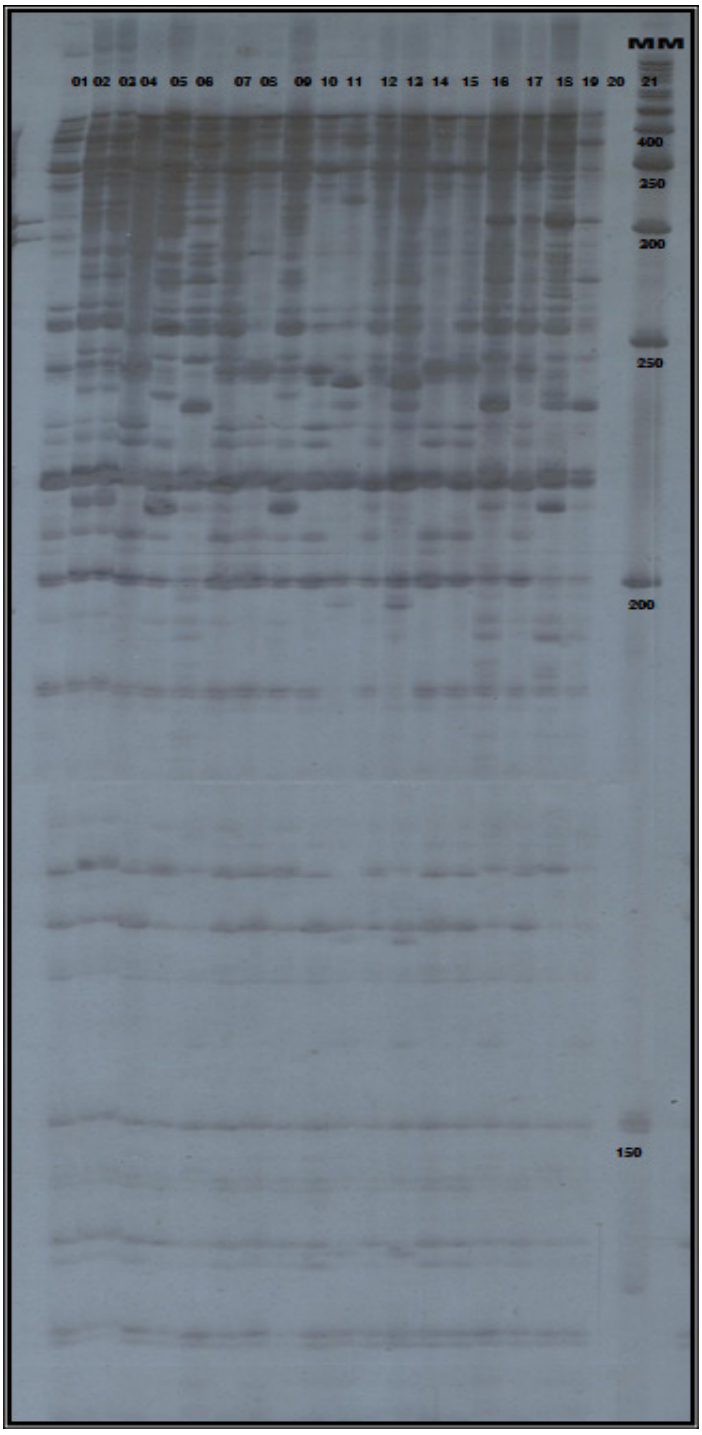
Abbreviations of each genotype in the biplots are shown below. Check varieties: Williams (W), Jaffa (J), Grand Naine (GE) and Gal (G). Farmers´ Clones: Sanchez (S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12); Bondaruck (B1, B2, B3, B4, B5, B6, B7, B8, B9, B10); CEDEVA MISIONES (CM, CM1, CM2, CM3, CM4, CM5); CEDEVA SALTA (CS1, CS3, CS4, CS5, CS6, CS7, CS8, CS9, CS10); Allotriploids Clones: BANANA DE ORO (BO1, BO2, BO3, BO4, BO5); PRATA (P1, P3, P4, P5); BANANA DE COCCIÓN (BC1, BC2, BC3, BC4, BC5). Some clones are hidden due to overlapped positions in the biplot.

Online resources

Online resource 1: List of banana genotypes analyzed in this paper, according to its type of polyploidy and experimental status

Polyploidy	
Autotriploid (AAA) genotypes	**Allotriploid (AAB) genotypes**
Check genotypes	BANANA DE ORO 1
WILLIAMS	BANANA DE ORO 2
JAFFA	BANANA DE ORO 3
GRAN ENANO	BANANA DE ORO 4
GAL	BANANA DE ORO 5
Farmers' clones (AAA)	PRATA 1
SANCHEZ 1	PRATA 2
SANCHEZ 2	PRATA 3
SANCHEZ 3	PRATA 4
SANCHEZ 4	PRATA 5
SANCHEZ 5	BANANA DE COCCIÓN 1
SANCHEZ 6	BANANA DE COCCIÓN 2
SANCHEZ 7	BANANA DE COCCIÓN 3
SANCHEZ 8	BANANA DE COCCIÓN 4
SANCHEZ 9	BANANA DE COCCIÓN 5
SANCHEZ 10	
SANCHEZ 11	
SANCHEZ 12	
CEDEVA SALTA 1	
CEDEVA SALTA 3	
CEDEVA SALTA 4	
CEDEVA SALTA 5	
CEDEVA SALTA 6	
CEDEVA SALTA 7	
CEDEVA SALTA 8	
CEDEVA SALTA 9	
CEDEVA SALTA 10	
BONDARUCK 1	
BONDARUCK 2	
BONDARUCK 3	
BONDARUCK 4	
BONDARUCK 5	
BONDARUCK 6	
BONDARUCK 7	
BONDARUCK 8	
BONDARUCK 9	
BONDARUCK 10	
CEDEVA MISIONES	
CEDEVA MISIONES 1	
CEDEVA MISIONES 2	
CEDEVA MISIONES 3	
CEDEVA MISIONES 4	
CEDEVA MISIONES 5	

Online resource 2: AFLP profiles of some banana clones obtained by selective amplification with primer+3 combination Q.



Online resource 3: Results for principal coordinates analysis after Dice's distances estimated from polymorphic AFLP: value and proportion of explained and cumulative variance by each eigenvalue

Eigenvalues	Value	Proportion of variance	
		Explained	Cumulative
1	1.49	0.25	0.25
2	0.68	0.11	0.37
3	0.5	0.08	0.45
4	0.45	0.08	0.53
5	0.35	0.06	0.59
6	0.29	0.05	0.63
7	0.24	0.04	0.67
8	0.22	0.04	0.71
9	0.2	0.03	0.75
10	0.16	0.03	0.77
11	0.15	0.03	0.8
12	0.14	0.02	0.82
13	0.12	0.02	0.84
14	0.11	0.02	0.86
15	0.1	0.02	0.88
16	0.09	0.01	0.89
17	0.08	0.01	0.9
18	0.07	0.01	0.92
19	0.07	0.01	0.93
20	0.06	0.01	0.94
21	0.05	0.01	0.95
22	0.05	0.01	0.96
23	0.04	0.01	0.96
24	0.04	0.01	0.97
25	0.03	0.01	0.97
26	0.03	0	0.98
27	0.02	0	0.98

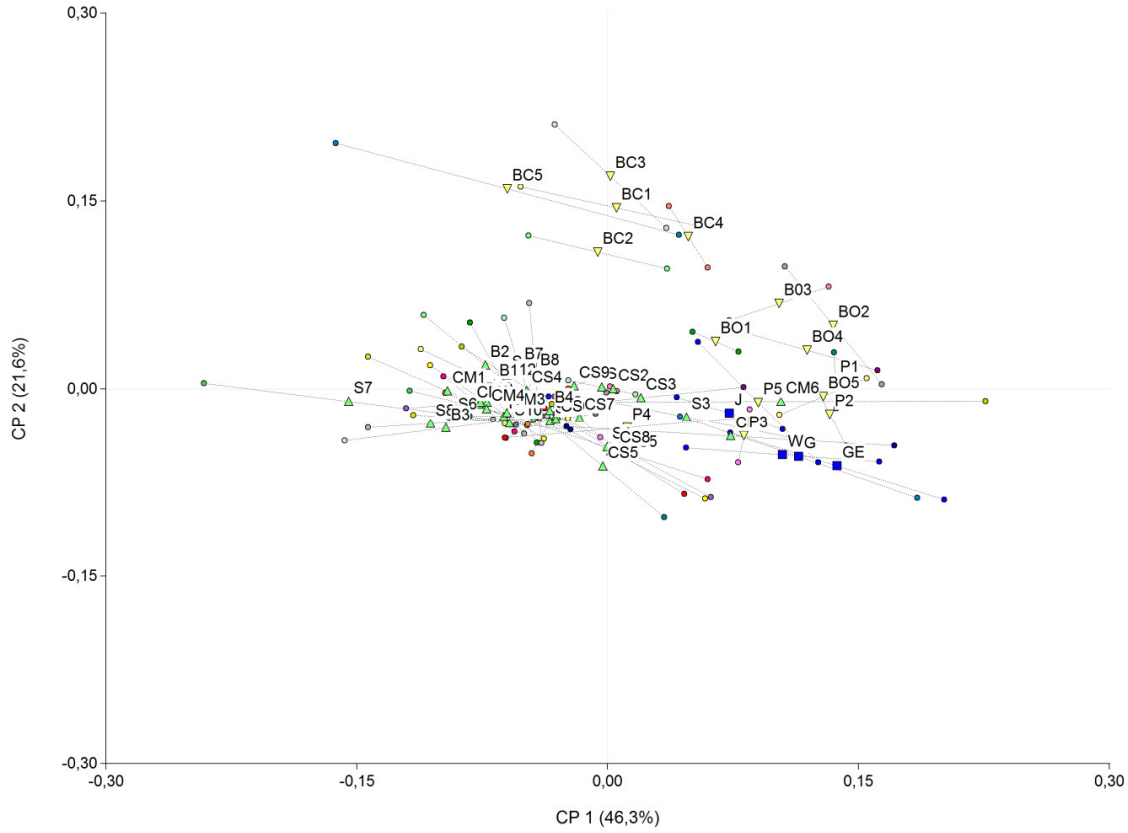
Online resource 4: Results for principal components analysis after mean of phenotypic traits evaluated at farmers's field: value and proportion of explained and cumulative variance by each eigenvalue

Eigenvalues	Value	Proportion of variance	
		Explained	Cumulative
1	2.58	0.43	0.43
2	2.05	0.34	0.77
3	0.70	0.12	0.89
4	0.36	0.06	0.95
5	0.18	0.03	0.98
6	0.13	0.02	1.00

Online resource 5: Results for Procrustes analysis made from synthetic variables. 10 eigenvalues – molecular data - and phenotypic data - the first and the second eigenvalue -. The value and proportion of explained and cumulative variance by each eigenvalue

Eigenvalues	Value	Proportion of variance	
		Explained	Cumulative
1	0.353	0.496	0.496
2	0.214	0.300	0.796
3	0.032	0.044	0.840
4	0.027	0.037	0.877
5	0.021	0.030	0.907
6	0.019	0.026	0.933
7	0.014	0.020	0.954
8	0.012	0.017	0.971
9	0.012	0.017	0.988
10	0.009	0.012	1.000

Online resource 6: Procrustes analysis made from molecular and phenotypic data show the same group pattern that Procrustes Analysis from synthetic variables. The first two principal components obtained by this multivariate analysis jointly explained the 67.9% (46.3% for CP1 and 21.6% for CP2)



Online resource 7: Mantel test result showed a linear correlation within the phenotypic array and the molecular array. (P=.0001)

	1	2
1	1.00	
2	0.23	1.00