



Genetic diversity and nonparametric statistics to identify possible ISSR marker association with fiber quality of pineapple

JULIANNA M. SILVA¹, PAULO R.L. LIMA², FERNANDA V.D. SOUZA³, CARLOS A.S. LEDO³, EVERTON H. SOUZA¹, KATIA N. PESTANA¹ and CLÁUDIA F. FERREIRA³

¹Universidade Federal do Recôncavo da Bahia, Centro de Ciências Agrárias, Ambientais e Biológicas, Campus Cruz das Almas, Rua Rui Barbosa, 710, Centro, 44380-000 Cruz das Almas, BA, Brazil

²Universidade Estadual de Feira de Santana, Departamento de Tecnologia, Avenida Transnordestina, s/n, Novo Horizonte, 44036-900 Feira de Santana, BA, Brazil

³Embrapa Mandioca e Fruticultura, s/n, Caixa Postal 007, Centro, 44380-000 Cruz das Almas, BA, Brazil

Manuscript received on July 25, 2018; accepted for publication on October 24, 2018

How to cite: SILVA JM, LIMA PRL, SOUZA FVD, LEDO CAS, SOUZA EH, PESTANA KN AND FERREIRA CF. Genetic diversity and nonparametric statistics to identify possible ISSR marker association with fiber quality of pineapple. *An Acad Bras Cienc* 91: e20180749. DOI 10.1590/0001-3765201920180749.

Abstract: Due to the increasing search for renewable resources, plant fibers have become an alternative when creating new products. Studies demonstrate the potential use of pineapple fibers in composites. The objective of this work was to evaluate the genetic diversity and verify any association between ISSR (Inter Simple Sequence Repeats) bands and quality of pineapple fibers for use in cements in the civil construction. The study analyzed the genetic variability of 11 pineapple genotypes, as well as the possible association of 131 bands from 16 ISSR markers with fiber quality characteristics. Eleven bands were selected based on their high correlations (0.64578* to 0.72457*) with three fiber quality variables. Of these, two bands were purified, sequenced, and blasted against sequences in GenBank at NCBI. These markers can be used in marker assisted selection to genetically improve the quality of pineapple fiber. Bands that returned no hits in the NCBI BLAST search can be deposited as new sequences in the GenBank. Therefore, the SCAR markers, once validated, can be useful in pineapple genetic breeding programs worldwide by using molecular marker assisted selection for fiber resistance, which could subsidize the development of more promising genotypes for industrial use and contribute to the sustainability of this new production sector.

Key words: *Ananas comosus* (L.) Merr., civil construction, genetic diversity, DNA markers, nonparametric statistics.

INTRODUCTION

Pineapples belong to the family Bromeliaceae, which comprises approximately 76 genera and 3.553 species (Butcher and Gouda 2018). It is one of the most consumed fruits worldwide, with global

production of 25.8 million tons and cultivated area of 1.04 million hectares in 2016 (Faostat 2018). In recent years, there has been a growing interest in use of materials from renewable and sustainable resources, including plant fibers instead of synthetic fibers to reinforce composites.

In civil construction, the use of plant fibers in cement matrices is notable as an excellent alternative because the fibers minimize the fragility

Correspondence to: Cláudia Fortes Ferreira
E-mail: claudia.ferreira@embrapa.br
ORCID: <https://orcid.org/0000-0001-5476-2236>

of the matrix (Aziz et al. 2005, Brandt 2008). Fibers from some taxa of *Ananas* are already in use, such as ‘Curauá’ (*Ananas comosus* var. *erectifolius*), but there is still a large number of pineapple cultivars which have never been evaluated, especially as a reinforcement in composite matrices.

Due to growing demand, ‘Curauá’ hybrids developed for ornamental purposes (Souza et al. 2014) have become potential candidates for fiber quality studies (Souza et al. 2017). These hybrids have shown promising results related to use in polymeric composites for automotive products (Sena Neto et al. 2013, 2015, 2017, Souza et al. 2017). However, little has been done in relation to using these fibers in cement matrices, which could be the basis of more sustainable civil construction products.

The Pineapple germplasm collection at Embrapa, located in Cruz das Almas, Bahia, with more than 700 conserved accessions of *Ananas* spp., is a potential source for these studies. However, phenotypic characterization for fiber quality is extremely laborious and expensive, making the characterization of a large number of accessions highly unfeasible.

Thus, given the difficulty to obtain the phenotypic characterization for fiber quality in pineapple, nonparametric tests becomes an interesting approach to try to identify associations between marker fragments and phenotypic variables to be used in marker assisted selection (MAS) in pineapple breeding programs aiming to develop pineapple with fiber quality for the purpose of uses in composites.

Molecular biology techniques can contribute to hasten genetic breeding programs by using molecular markers that detect genetic variability (or genetic polymorphisms) of DNA sequences (Arif et al. 2010). ISSR markers (Inter Simple Sequence Repeats) are based on PCR (polymerase chain reaction), do not require prior knowledge of DNA sequences of the target species, and produce

reproducible, highly polymorphic fragments at a low cost. These markers are broadly used in genetic diversity and variability studies (Gottardi et al. 2001, Barth et al. 2002, González et al. 2002).

Therefore, the objectives of the present study were to evaluate the genetic diversity between pineapple varieties already under cultivation and hybrids, and given the phenotypic characterization, try to identify possible associations between bands of ISSR markers and fiber quality variables for use in cement matrices using nonparametric statistics. Once validated, these markers can be used in MAS and hasten the pineapple genetic breeding program aiming fiber quality for use in the civil construction industry given that the phenotypic evaluation is extremely difficult, costly and time-consuming.

MATERIALS AND METHODS

GENETIC MATERIAL

This work evaluated pineapple varieties already under cultivation and hybrids with cultivation potential (close to being released as varieties), since one of the premises for using the fibers in cement matrices is production volume.

Eleven cultivars, hybrids and their background are detailed in Table I. The ‘Curauá’ variety was used as reference due to its excellent properties of fiber quality (Leão et al. 2009). The cultivated varieties and the hybrids were obtained from field crosses and mother plant nurseries at Embrapa, located in Cruz das Almas, Bahia.

DNA EXTRACTION

The DNA was extracted in the Advanced Biology Center (NBA, Núcleo de Biologia Avançada) at Embrapa using the CTAB (Cetyltrimethylammonium bromide) method proposed by Doyle and Doyle (1990).

Young leaves of the 11 pineapple genotypes were collected for DNA extraction. Briefly, approximately 300 mg of leaf tissue was macerated in liquid nitrogen

TABLE I
Genotypes evaluated for fiber quality for the use as reinforcement in cement composites.

Pineapple genotypes	Scientific name	Use
1. 'BRS Imperial'	<i>Ananas comosus</i> var. <i>comosus</i> (L.) Merrill	Cultivar/ Food
2. 'Pérola'	<i>A. comosus</i> var. <i>comosus</i>	Cultivar/ Food
3. 'Smooth Cayenne'	<i>A. comosus</i> var. <i>comosus</i>	Cultivar/ Food
4. 'BRS Vitória'	<i>A. comosus</i> var. <i>comosus</i>	Cultivar/ Food
5. 'BRS Ajubá'	<i>A. comosus</i> var. <i>comosus</i>	Cultivar/ Food
6. PE x SC-73	<i>A. comosus</i> var. <i>comosus</i>	New hybrid/ Food
7. 'BRS Anauê'	<i>Ananas comosus</i> (L.) Merr. var. <i>erectifolius</i> (L.B.Sm.) Coppens & F.Leal x <i>Ananas comosus</i> (L.) Merr. var. <i>bracteatus</i> (Lindl.) Coppens & F.Leal	Cultivar/ Ornamental
8. FIB-POT	<i>A. comosus</i> var. <i>erectifolius</i> x <i>A. comosus</i> var. <i>bracteatus</i>	New hybrid/ Fiber and Ornamental
9. 'BRS Boyrá'	<i>Ananas comosus</i> (L.) Merr. var. <i>microstachys</i> L.B.Sm. x <i>A. comosus</i> var. <i>erectifolius</i>	Cultivar/ Ornamental
10. FIB-GOR	<i>A. comosus</i> var. <i>erectifolius</i> x <i>A. comosus</i> var. <i>bracteatus</i>	New hybrid/ Fiber and Ornamental
11. 'Curauá'	<i>Ananas comosus</i> var. <i>erectifolius</i>	Cultivar/ Fiber and Ornamental

and transferred to 2.0 (mL) microtubes. Afterwards, an extraction buffer was added (1.7 [M] NaCl; 0.1 [M] Tris-HCl 1[M] pH 8.0; 20 [mM] EDTA 0.5; 2% [m/v] polyvinylpyrrolidone; 2.4% [m/v] CTAB; and 0.4% [v/v] β -mercaptoethanol), which was preheated to 65 °C, and the tubes homogenized for 5 min. Subsequently, the samples were incubated in a water bath at 65 °C for 45 min and then homogenized for 10 min.

Afterwards, chloroform:alcohol-isoamyl (24:1) was added to the samples, which were then centrifuged at 10.000 x g for 10 min (step conducted twice), followed by the addition of ice cold isopropyl alcohol to the supernatant. The material was placed in a freezer (- 20 °C) for 24 h and then centrifuged at 10.000 x g for 10 min. Following this, the DNA was washed with 70% ethanol and the pellet resuspended in TE (10 [mM] Tris-HC, pH 8.0, 1[mM] EDTA) plus ribonuclease (10 mg.mL⁻¹ RNase), placed in the

oven at 37 °C, and 3.0 (M) of sodium acetate was added. The material was then centrifuged for 20 s at 3.000 x g. Ice cold absolute ethanol was added to the supernatant and the samples centrifuged at 10.000 x g. The DNA as washed two more times with 70% ethanol and resuspended in nuclease free water. Finally, the extracted DNA was stored at - 20 °C. The quality and quantity of the DNA was evaluated with a comparative analysis of the samples in 1% agarose gel using markers of known molecular weight.

DNA AMPLIFICATION USING ISSR MARKERS

The DNA was amplified using 16 ISSR primers (Table II). The PCR reactions were completed to a final volume of 15 (μ L) with the following reagents: 1x buffer, 2.5 (mM) MgCl₂, 0.2 (mM) dNTPs, 0.2 (μ M) of each primer, 1 U of Taq Polymerase (LBM), and 25 (ng) of genomic DNA. The amplifications were carried out in the Applied Biosystems

TABLE II
Identification and sequences of the 16 ISSR primers used in the genetic diversity study of 11 pineapple genotypes.

Marker Name	Sequence (5'-3')	Annealing temperature (°C)	Number of polymorphic bands
ISSR 03 DiCA3'RG	CACACACACACACACARG	48	02
ISSR 07 DiCA5'CY	CYCACACACACACACACA	48	03
ISSR 24 DiGT5'A	AGTGTGTGTGTGTGTGT	48	17
ISSR 30 TriCAC3'YC	CACCACCACCACCACYC	48	07
ISSR 32 TriCAC5'CY	CYCACCACCACCACCAC	48	03
ISSR 34 TriCAG3'RC	CAGCAGCAGCAGCAGRC	48	04
ISSR 35 TriCAG3'YC	CAGCAGCAGCAGCAGYC	48	17
ISSR 36 TriCAG5'CR	CRCAGCAGCAGCAGCAG	48	05
ISSR 42 TriGTG5'CY	CYGTGGTGGTGGTGGTG	48	07
ISSR 53 TriATC3'RC	ATCATCATCATCATCRC	48	18
ISSR 72 TriTCC3'RC	TCCTCCTCCTCCTCCRC	48	09
ISSR 81 TriCTT3'RC	CTTCTTCTTCTTCTTRC	48	13
ISSR 84 TriCCT3'RC	CCTCCTCCTCCTCCTRC	48	17
ISSR 90 TriGAA3'RC	GAAGAAGAAGAAGAARC	48	02
ISSR 94 TriGTA3'RC	GTAGTAGTAGTAGTARC	48	04
ISSR 95 TriGTT3'RC	GTTGTTGTTGTTGTTTRC	48	03

thermocycler (model Veriti® 96-Wells) using the following program: an initial denaturation step at 94 °C for 3 min, followed by 45 cycles of denaturation at 94 °C for 45 s for each one, 45 s for primer annealing at 48 °C, and an extension at 72 °C for 1 min. This was followed by a final extension at 72 °C for 7 min and 14 °C ∞.

The amplified products were separated using electrophoresis in 2.5% agarose gel, 90 V, in TBE 0.5X buffer, and stained with ethidium bromide (0.5 µg.mL⁻¹). The amplified fragments were viewed under UV light and photographed using a photodocumentation system. Fragment sizes were

determined based on a 1 kb molecular weight marker (Qiagen).

MOLECULAR DATA ANALYSIS

Fragments from the ISSR markers, for both the diversity and association studies, were evaluated for absence (0) and presence (1) of bands. The electrophoretic profile of primer ISSR-30 for the 11 pineapple genotypes evaluated is shown in Figure 1.

Based on the (0) and (1) data, a genetic dissimilarity matrix of the genotypes was generated, which was calculated with the Jaccard coefficient (complement of the Jaccard index = 1- c). Cluster

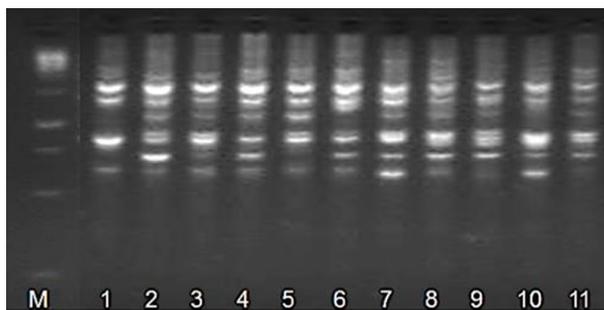


Figure 1 - Electrophoretic profile of primer ISSR-30 for the 11 pineapple genotypes in 2.5% agarose gel: 1- 'BRS Imperial'; 2- 'Pérola'; 3- 'Smooth Cayene'; 4- PE x SC-73; 5- 'BRS Vitória'; 6- 'BRS Ajubá'; 7- 'BRS Anauê'; 8- FIB-POT; 9- 'BRS Boyrá'; 10- FIB-GOR and 11- 'Curauá'. M = ladder 1 Kb (Qiagen™).

analysis was conducted using the UPGMA (Unweighted Pair-Group Method with Arithmetic Mean) method and the packages, *qlcmatrix*, *ade4* and *Nbclust*, of the R software (Development Core Team 2016) and the threshold defined according to the parameters described set by Mingoti (2005). Cluster validation was carried out by calculating the cophenetic correlation coefficient (CCC) (Sokal and Rolf 1962).

ANALYSIS OF PHENOTYPIC CHARACTERISTICS

Fiber quality data were obtained from previous tests (Silva 2016). The main variables measured are related to the mechanical properties of resistance to traction (RES), elasticity module (ELS), and water absorption (ABS), Crystallinity index (IC), fiber length (COM) and diameter of fiber (DIM).

As to the phenotypic characteristics, for all data regarding water absorption, the fibers were cut into 30 mm where approximately 1g of fiber weighed. These fibers were immersed in recipients with water and submitted to weighing and the first read was carried out after 3 h of immersion and at every 24 h during 5 consecutive days, according to Lopes et al. (2010). Four replicates were defined for each variety and the absorption calculated. Also the traction assay of fibers was carried out with 15 replicates for obtainment of RES and ELS values.

A combined analysis of the molecular (qualitative) and phenotypic (quantitative) data was also carried out to determine the genetic variability based on the Gower (1971) algorithm; groups were validated according to the CCC (Sokal and Rolf 1962). The qualitative data analyzed using the cluster, *Nbclust*, *ade4*, *agricolae* and *factoextra* packages in the R program (Development Core Team 2016) and the Gower algorithm analyzed using the *statmatch*, *cluster*, *ade4*, *Nbclust*, *clustOfvar* and *qlcmatrix* packages in the R software (Development Core Team 2016). The Principal Components Analysis (PCA) analysis using the phenotypic data was calculated using the *Factominer* package in the R software (Development Core Team 2016).

NONPARAMETRIC ANALYSIS

Since both genotypic and phenotypic data were available, the possibility of finding any correlation between a band and a phenotypic variable, although remote, was tested. The correlation analysis was carried out for 3 fiber quality variables (RES, ELS and ABS) and 131 bands from 16 ISSR molecular markers. Spearman's correlation and the nonparametric Kruskal-Wallis test were obtained by the statistical software SAS (SAS Institute 2010) using the following commands: *proc corr spearman* and *proc npar1way* and *anova*, respectively. The nonparametric test, was applied since it does not require a normal distribution of data and therefore, was most suitable for our data.

SEQUENCING BANDS TO DESIGN SCAR MARKERS

After the correlation analysis of the data using nonparametric tests, some relatively high correlations were found between bands and the phenotypic characteristics of interest for fiber quality for use in cement composites/civil construction industry. However, although all 11 bands were excised from the agarose gel, only 2 presented PCR with a single band. The

single bands were purified using the BIONEER Invitrogen Accuprep gel purification kit (<http://us.bioneer.com/Protocol/AccuPrep%20Gel%20Purification%20Kit.pdf>) and sequenced by capillary electrophoresis - ABI3730 sequencer, with the POP7 polymer and BigDye v3.1, by the Myleus Biotechnology company (www.myleus.com).

After obtaining the FASTA files with the contigs from the sequencing, the SCAR primers were designed using the PRIMER3PLUS software (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>).

RESULTS AND DISCUSSION

Studies using plant fibers to reinforce composites have been conducted using different plants, such as sisal, jute, coconut, soybeans, and bananas (Mohanty et al. 2000, Mishra et al. 2004, Liu et al. 2005, Cao et al. 2006, Romanzini et al. 2013, Zhu et al. 2013).

GENETIC DIVERSITY BASED ON THE MOLECULAR DATA

The 16 ISSR primers tested on the 11 genotypes generated 131 polymorphic bands. The number of bands per primer varied from 2 (ISSR 03) to 18 (ISSR 53), with an average of 8 polymorphic bands per primer.

The dendrogram, generated by the Jaccard index using the genetic dissimilarity matrix of the genotypes, contained 2 groups (Fig. 2).

G1 comprises of 6 genotypes ('BRS Imperial', 'Pérola', 'Smooth Cayenne', PE x SC-73, 'BRS Vitória' and 'BRS Ajubá') and G2, 5 genotypes ('BRS Anauê', FIB-POT, FIB-GOR, 'BRS Boyrá' and 'Curauá'). In G1, all the varieties are edible, and G2 contains the ornamental 'Curauá' hybrids and 'Curauá' itself. FIB-POT demonstrated excellent resistance when used in the cement composite, corroborating studies by Sena Neto et al. (2015)

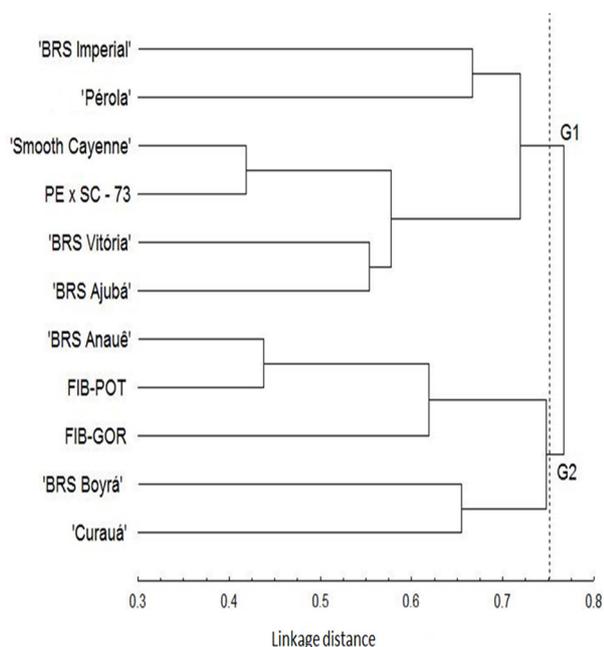


Figure 2 - Dendrogram of the 11 pineapple genotypes based on the Jaccard index and the 131 bands from 16 ISSR markers. The UPGMA method was used to define the groups and cutoff point was defined according to the parameters indicated by Mingoti (2005).

that identified this genotype as extremely good for reinforcing polymer matrices. This characteristic might have been inherited from 'Curauá', which produces fibers that are already being used in the automotive industry to reinforce polymer composites.

Based on the dissimilarity matrix of the molecular data, the genetic distance values varied from 0.42 to 0.86. Dissimilarity was the highest between genotypes 'BRS Imperial' and 'BRS Anauê' and the lowest between 'Smooth Cayenne' (SC) and PE x SC-73, which is probably because the latter is a hybrid of 'Smooth Cayenne.'

Souza et al. (2017) studied and characterized the diversity of 13 pineapple genotypes, in relation to using their fibers as mechanical reinforcement in polymer matrices, and obtained 217 bands from 17 ISSR markers, with a genetic distance ranging from 0.09 to 0.64.

Vanijajiva (2012) evaluated the diversity of 15 pineapple accessions in Thailand, using 4 ISSR markers, and obtained 27 polymorphic bands and a genetic distance ranging from 0.32 to 0.97. These studies reveal the existence of genetic variability to be explored among pineapple varieties using ISSR markers.

GENETIC DIVERSITY BASED ON PHENOTYPIC CHARACTERISTICS

Measuring phenotypic characteristics for fiber quality in pineapple is one of the most excruciating and time consuming efforts. Generally it is unfeasible to evaluate more than 6 characteristics in 10 genotypes in one year. The data used to evaluate fiber quality are listed in Table III. As to the quantitative variables: RES, ELS, and ABS, IC, COM and DIM, two groups were formed based on the mean Euclidean distance (Figure 3).

Group 1 comprised 10 genotypes: ‘BRS Imperial’, ‘BRS Vitória’, ‘BRS Ajubá’, ‘BRS Anauê’, FIB-POT, ‘Pérola’, ‘Smooth Cayenne’, ‘BRS Boyrá’, PE x SC-73 and FIB-GOR. G2 comprised only of ‘Curauá,’ which is probably due to the higher RES (480 MPa) and ELS (18 GPa) values of its fibers in comparison to the other genotypes.

The genetic distance values ranged from 0.23, between ‘BRS Vitória’ and ‘BRS Ajubá’ to 3.06, between FIB-GOR and ‘Curauá’. In principle, the genotypes, which have ‘Curauá’ as one of their parents, should be considered good candidates for having interesting fiber characteristics to be exploited by the cement composite civil construction industry.

PCA AND GOWER ALGORITHM

Since we had the phenotypic data for 6 characteristics for fiber quality, PCA analysis was carried out to identify the most important variables; those which were contributing more to the phenotypic variation among the genotypes (Fig. 4).

TABLE III
Values for resistance to traction, modulus of elasticity, and absorption of water indices for the leaf fibers of the pineapples studied.

Pineapple genotypes	Resistance to traction (MPa)	Elasticity module (MPa)	Absorption (g)
‘BRS Imperial’	321.47	12358.67	2.27
‘Pérola’	257.48	8915.12	2.30
‘Smooth Cayenne’	396.12	18464.48	2.10
PE x SC-73	451.04	16706.74	2.21
‘BRS Vitória’	376.91	11952.57	2.25
‘BRS Ajubá’	370.13	13982.35	2.26
‘BRS Anauê’	273.01	14619.78	2.27
FIB-POT	336.70	18170.52	2.23
‘BRS Boyrá’	272.95	16046.04	2.05
FIB-GOR	224.15	6963.66	2.14
‘Curauá’	480.23	27426.84	1.99

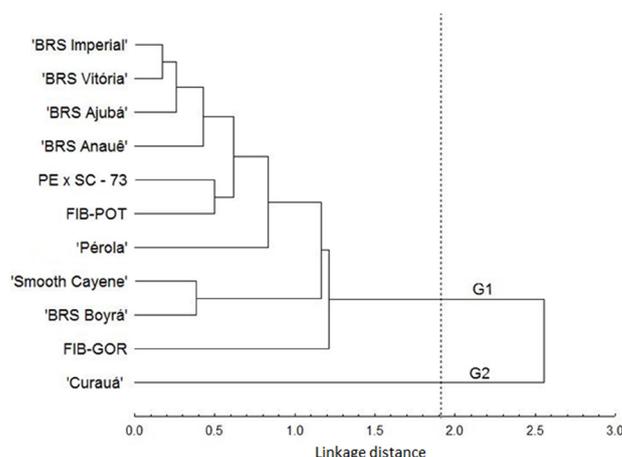


Figure 3 - Dendrogram of the 11 pineapple genotypes based on the mean Euclidean distance and the fiber quality variables. The UPGMA method was used to define the groups and cutoff point was defined according to the parameters indicated by Mingoti (2005).

In our work, the PCA showed that PC1 and PC2, in total, correspond to approximately 70% of the phenotypic variation with contributions of 49.14 and 19.29%, respectively (Fig. 3). The variables which most contributed were ELS, IC, RES and DIM, contributing 29.11, 20.14, 19.69 and 19.51%, respectively. Variables RES, IC and ELS, showed high and positive correlations, and are directly related with fiber quality, all contributing to desired mechanical properties. Regarding ABS, DIM and COM, these contributed with negative correlations, confirming that when the fiber absorbs more water, it suffers dimensional variations, which is a characteristic which hinders the contact of the fiber with the matrix and for a fiber to be considered of quality, it must have low ABS (Fig. 3). The COM variable contributed least to the phenotypic variation between the genotypes: 4.3%, and therefore, was excluded from the Gower algorithm analysis.

The UPGMA grouping method was based on the Gower algorithm (Gower 1971) (Fig. 5). The number of groups was calculated by the pseudo t^2 index, proposed by Duda and Hart (1973) in the “NbClust” package (Charrad et al. 2013) of the R software program (R Development Core Team 2016). The genetic distance varied from 0.16 to 0.44. Dissimilarity was lowest between ‘BRS Anauê’ and FIB-POT, probably because they originated from the same cross. Genetic dissimilarity was highest between ‘BRS Imperial’ and ‘BRS Boyrá’, as well as ‘BRS Imperial’ and FIB-GOR.

G1 comprised of 6 genotypes: ‘BRS Imperial’, ‘Pérola’, ‘Smooth Cayenne’, PE x SC-73, ‘BRS Vitória’ and ‘BRS Ajubá’ and G2 of 5 genotypes: ‘BRS Anauê’, FIB-POT, FIB-GOR, ‘BRS Boyrá’ and ‘Curauá’. Considering the botanical varieties, all the edible pineapples of *Ananas comosus* var. *comosus* were allocated in G1.

Souza et al. (2017) conducted a combined group analysis using the UPGMA method and the Gower algorithm and thermal and mechanical data for

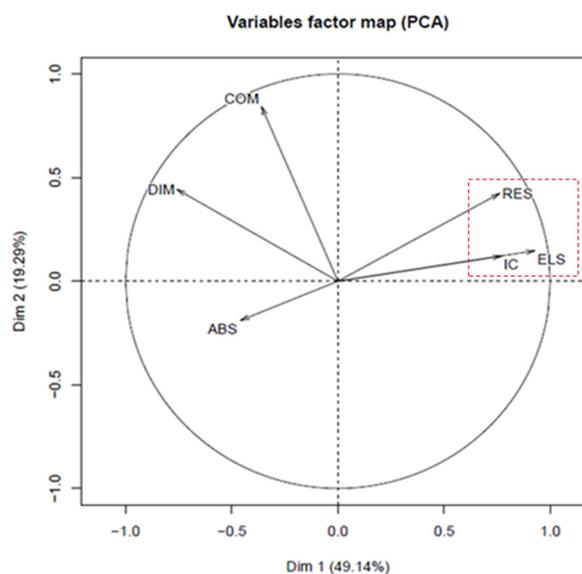


Figure 4 - Principal components analysis using 6 phenotypic variables for pineapple fiber quality: Resistance to traction (RES), elasticity module (ELS), and absorption (ABS), Crystallinity index (IC), fiber length (COM) and diameter of fiber (DIM). *Factominer* package in the R software (Development Core Team 2016).

fiber quality of pineapple for use in the automobile industry, which resulted in 4 groups with genetic distance values that varied from 0.14 to 0.50.

In our study, the combined group analysis (Fig. 5) better separated the genotypes in comparison to the individual analyses for each type of variable: molecular (Fig. 1) and phenotypic (Fig. 2). In previous studies, cluster analyses have helped identify promising genotypes for use in genetic improvement programs by providing information about the potential of existing variability (Souza et al. 2012, 2017).

The cophenetic correlation coefficient values were 0.86, 0.84 and 0.85 for the data evaluated using the Jaccard index, mean Euclidean distance, and Gower distance (Gower 1971), respectively. According to Sokal and Rohlf (1962), values above 0.80 indicate a good fit between the distance matrices and derivatives of the graphic distances, giving credibility to the groups formed using the UPGMA method.

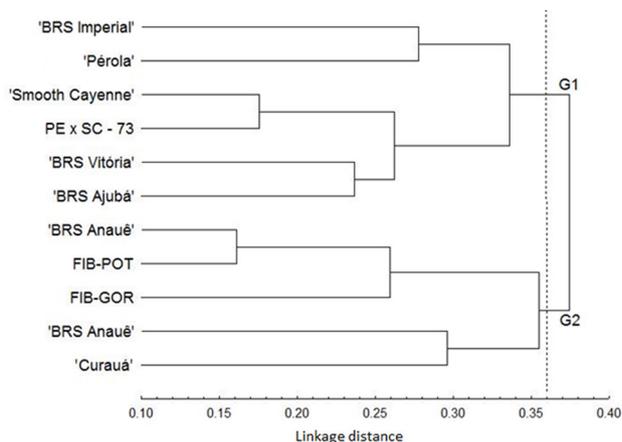


Figure 5 - Dendrogram of the 11 pineapple genotypes based on the Gower algorithm (Gower 1971): combined molecular and phenotypic data for fiber quality. The UPGMA method was used to define the groups and cutoff points were defined according to the pseudo t^2 parameters (Duda and Hart 1973).

The results of our work are in agreement with those of Souza et al. (2017) and Sena Neto et al. (2013, 2015), who found significant values for thermal and mechanical properties of fibers from different pineapple genotypes, which demonstrates the possibility of using these fibers to reinforce composites.

ANALYSIS VIA NONPARAMETRIC TESTS

Nonparametric tests, although limited, are used when there is no normal distribution of data and therefore, were chosen to increase information using molecular and phenotypic data already obtained, especially due to the difficulty in obtaining phenotypic data for fiber quality for a large number of pineapple genotypes.

In association studies, it is important to evaluate many individuals of a population of crosses and to carefully measure phenotypic characteristics of interest. For pineapples, however, measuring phenotypic characteristics of fiber quality is unfeasible due to the difficulty and high cost of the tests that make it impossible to evaluate numerous individuals at one single time. Therefore, this possible association was analyzed using nonparametric methods, Spearman's correlation

and the Kruskal-Wallis test, as successfully done by Souza et al. (2017).

In our study, we did not expect to obtain any correlations between markers and phenotypic characteristics of interest, mainly due to the quantitative nature of the variables for fiber quality. However, interesting results were found that can be explored to improve the culture, with the goal of identifying promising materials that can be used in the composite in civil construction industry.

For the evaluation of the association between the 131 polymorphic bands and 3 phenotypic variables, 11 high (+ and -) and significant correlations, which varied from -0.72457^* to 0.74903^{**} for fiber quality (Table IV), were encountered.

One plant fiber characteristic is the capacity to absorb high amounts of water, which results in high dimensional variation. This is one of the main phenomena that can influence the behavior of fibers when they are used in cement composites, because the natural variation in humidity can reduce the diameter of the fibers causing them to lose contact with the matrix to which they are adhered. For this variable, there was a correlation with 4 bands for different ISSR markers that varied from -0.64578^* to 0.72457^* .

Resistance to traction is the maximum tension of the fiber and is a determinant variable to measure high performance as a mechanical reinforcement. In relation to this variable, bands M10 of the ISSR 24 marker, M8 of the ISSR 53 marker, M7 of the ISSR 84 marker, and M3 ISSR 95 marker had correlation values from -0.63960^* to 0.74903^{**} .

The elasticity module presented 3 bands, M4 of the ISSR 81 marker, M5 of the ISSR 53 marker and M13 of the ISSR 84 marker, which had correlation values of 0.72175^{**} , -0.6931^* and -0.72457^* , respectively. The highest correlation found between a band and characteristic of interest for fiber quality was band M3 of marker ISSR 95 and resistance to traction, with a correlation value of 0.74903^{**} . Two

TABLE IV
Spearman and Kruskal-Wallis correlation used to analyze the association between molecular markers and fiber quality variables of 11 pineapple genotypes.

Marker	Variable	Spearman	χ^2	pr > χ^2
M4 – ISSR 24	Water absorption (g)	-0.69631*	4.36	0.0253
M10 – ISSR 24	Resistance to traction (MPa)	-0.63960*	3.68	0.0464
M6 – ISSR 72	Water absorption (g)	-0.64578*	3.75	0.0437
M4 – ISSR 81	Elasticity module (MPa)	-0.72175**	4.68	0.018
M4 – ISSR 53	Water absorption (g)	0.69631*	4.36	0.0253
M5 – ISSR 53	Elasticity module (MPa)	-0.69631*	4.36	0.0253
M8 – ISSR 53	Resistance to traction (MPa)	-0.64578*	3.75	0.0437
M4 – ISSR 84	Water absorption (g)	0.72457*	4.20	0.0272
M7 – ISSR 84	Resistance to traction (MPa)	-0.72457*	4.20	0.0272
M13 – ISSR 84	Elasticity module (MPa)	-0.72457*	4.20	0.0272
M3 – ISSR 95	Resistance to traction (MPa)	0.74903**	5.12	0.0127

*Significant at 5% probability; ** Significant at 1% probability, χ^2 (chi square), pr > χ^2 .

other high correlations were found for the ISSR 84 marker, band M7 and resistance to traction and band M23 and elasticity module, which both had correlation values of -0.72457*.

Souza et al. (2017) conducted the first work related to the use of nonparametric methods to evaluate fiber quality of pineapple and obtained results similar to the present study. The study selected 11 bands based on the correlation of 17 ISSR molecular markers with four fiber quality variables and obtained high correlations (0.63434* to 0.76169**). According to the same author, nonparametric tests are a viable alternative for this type of evaluation, since pineapple fiber has superior mechanical properties that should be explored as a reinforcement of polymeric (Sena Neto et al. 2015) and cement composites.

In our work, the correlation between bands of DNA markers and phenotypic characteristic was evaluated based on three fiber quality variables (RES, ELS, and ABS) and ISSR markers. The three variables evaluated are of fundamental importance to fiber quality studies and the evaluation of fibers for use in cement composites. Fibers for this purpose need to have low ABS and high RES and ELS. Two highly correlated bands were sequenced

(M6 ISSR 72 and M13 ISSR 84; 4 – PE x SC-73; 10- ‘FIB-GOR’) and, from this, 5 SCAR (Sequence Characterized Amplified Regions) primers were designed for each ISSR marker, which will be validated in accessions of the Pineapple collection at Embrapa (Table V).

The contigs were blasted against the NCBI website (BLAST:< https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch>) but they were classified as no hits (i.e., there were no similar sequences in this database). However, these sequences could be exploited in the future by validation of the primers. What is interesting about this study is the combination of markers that can be used giving more credibility when choosing plants that provide interesting materials for use as cement composites. Furthermore, the more bands that are sequenced and validated, the higher the credibility of the MAS.

The fact that the sequenced bands had an output of “no hits” in the BLAST (Basic Local Alignment Search Tool) search, opens new opportunities in regard to validation, because they can be deposited as new sequences in the NCBI GenBank if they are associated with fiber quality for the purpose proposed in this work. Therefore, this work will

TABLE V

Design of new primers/ SCAR molecular markers to be tested and validated for association with fiber quality in pineapples.

Primer name	Sequence (5'-3')
ISSR 72 TriTCC3'RC	F' TTGCCACAAAGTAACCACCA
	R' GAGCTAATGCACGATGACGA
	F' AAGCGAGATACGAAGGACGA
	R' ACGTGACTCGCCTCAGAAAT
	F' TCGGACCTAACCCCTTCCTTT
	R' GCGAGAGACCTGTTTGATCC
	F' TTCGGACCTAACCCCTTCCTT
	R' GCGAGAGACCTGTTTGATCC
	F' AGTTAGCGGAAGGGAGGAAG
	R' TGTAAGCTCGCCAGTGCATC
ISSR 84 TriCCT3'RC	F' TTGCCACAAAGTAACCACCA
	R' GAGCTAATGCACGATGACGA
	F' AAGCGAGATACGAAGGACGA
	R' ACGTGACTCGCCTCAGAAAT
	F' TCGGACCTAACCCCTTCCTTT
	R' GCGAGAGACCTGTTTGATCC
	F' TTCGGACCTAACCCCTTCCTT
	R' GCGAGAGACCTGTTTGATCC
	F' AGTTAGCGGAAGGGAGGAAG
	R' TGTAAGCTCGCCAGTGCATC

significantly contribute to provide markers that can be used in MAS for cement composites in pineapples for use in the civil construction industry.

Therefore, the SCAR markers, once validated, can be useful in pineapple genetic breeding programs worldwide by using MAS selection for fiber quality, which could subsidize future improvement in the development for more promising genotypes for industrial use and contribute to the sustainability of this new production sector.

ACKNOWLEDGMENTS

The authors would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/ Embrapa/ 2014 and CAPES/ Programa Nacional de Pós-Doutorado) for the scholarship and Fundação de Amparo à Pesquisa do Estado da

Bahia (FAPESB) – TSC 0013/2014) for the funding of the project.

AUTHOR CONTRIBUTIONS

JMS carried out the research as part of her master's degree and contributed to the data analysis, prepared the tables and figures and interpreted the results. PRL, EHS, KNP contributed to molecular analysis, performed data analysis and interpreted results in the discussion; FVDS contributed to discussion of the results and commented on the manuscript. CASL carried out the statistical analysis and data interpretation; CFF was the supervisor of the project, and contributed to the experimental design, data analysis, writing and revision of the manuscript, as well as the correspondence with the editors and reviewers.

REFERENCES

- ARIF IA, BAKIR MA, KHAN HA, FARHAN AHA, HOMAIDAN AAA, BAHKALI AH, SADOON MA AND SHOBRAK M. 2010. A brief review of molecular techniques to assess plant diversity. *Int J Mol Sci* 11: 2079-2096.
- AZIZ SH, ANSELL MP, CLARKE SJ AND PANTENY SR. 2005. Modified polyester resins for natural fibre composites. *Compos Sci Technol* 65: 525-535.
- BARTH S, MELCHINGER AE AND LÜBBERSTEDT T. 2002. Genetic diversity in *Arabidopsis thaliana* L. Heynh. Investigated by cleaved amplified polymorphic sequence (CAPS) and inter-simple sequence repeat (ISSR) markers. *Mol Ecol* 11: 495-505.
- BRANDT AM. 2008. Fibre reinforced cement-based (FRC) composites after over 40 years of development in building and civil engineering. *Compos Struct* 86: 3-9.
- BUTCHER D AND GOUDA EJ. 2018. The New Bromeliad Taxon List. University Botanic Gardens, Utrecht (cont. updated) (8 February 2018). <https://botu07.bio.uu.nl/bcg/taxonList.php?>
- CAO Y, SHIBATA S AND FUKUMOTO I. 2006. Mechanical properties of biodegradable composites reinforced with bagasse fibre before and after alkali treatments. *Composites Part A* 37: 423-429.
- CHARRAD M, GHAZZALI N, BOITEAU V AND NIKNAFS A. 2013. NbClust: An examination of indices for determining the number of clusters. R package version 1.4. <http://cran.rproject.org/web/packages/NbClust/index.html>.
- DEVELOPMENT CORE TEAM. 2016. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- DOYLE JJ AND DOYLE JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.
- DUDA RO AND HART PE. 1973. Pattern classification and scene analysis. J Wiley & Sons, Hoboken.
- FAOSTAT - FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. 2018. Agricultural Production. (9 February 2018). <http://faostat.fao.org/site/339/default.aspx>.
- GONZÁLEZ A, COULSON M AND BRETTELL R. 2002. Development of DNA markers (ISSRs) in mango. *Acta Hort* 575: 139-143.
- GOTTARDI MVC, LEMOS EGM AND RUGGIERO C. 2001. Avaliação de plantas matrizes de abacaxizeiro cultivar smooth cayenne utilizando marcadores RAPD e padrões isoenzimáticos. *Rev Bras Frutic* 23: 463-467.
- GOWER JC. 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-874.
- LEÃO AL, MACHADO IS, SOUZA SF AND SORIANO L. 2009. Production of 'curauá' (*Ananas erectifolius* L.B. Smith) fibers for industrial applications: characterization and micropropagation. *Acta Hort* 822: 227-238.
- LIU W, MISRA M, ASKELAND P, DRZAL LT AND MOHANTY AK. 2005. Green composites from soy based plastic and pineapple leaf fiber: fabrication and properties evaluation. *Polym* 46: 2710-2721.
- LOPES FFM, ARAÚJO GT, NASCIMENTO JWB, GADELHA TS AND SILVA VR. 2010. Estudo dos efeitos da acetilação em fibras de sisal. *Rev Bras Eng Agríc Ambient* 14(7): 783-788.
- MINGOTI AS. 2005. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: UFMG, 297 p.
- MISHRA S, MOHANTY AK, DRZAL LT, MISRA M AND HINRICHSEN G. 2004. A Review on Pineapple Leaf Fibers, Sisal Fibers and Their Biocomposites. *Macromol Mater Eng* 289: 955-974.
- MOHANTY AK, KHAN MA AND HINRICHSEN G. 2000. Influence of chemical surface modification on the properties of biodegradable jute fabrics-polyester amide composite. *Composites Part A* 31: 143-150.
- ROMANZINI D, LAVORATTIA, ORNAGHI JR HL, AMICO SC AND ZATTERA AJ. 2013. Influence of fiber content on the mechanical and dynamic mechanical properties of glass/ramie polymer composites. *Mater Des* 4: 9-15.
- SAS INSTITUTE INC. 2010. SAS/STAT: user's guide, v.9.2. Cary, NC.
- SENA NETO AR, ARAUJO MAM, BARBOZA RMP, FONSECA AS, TANOLI G, SOUZA FVD, MALTOSO LHC AND MARCONCINI JM. 2015. Comparative study of 12 pineapple leaf fiber varieties for use as mechanical reinforcement in polymer composites. *Ind Crops Prod* 64: 68-78.
- SENA NETO AR, ARAUJO MAM, SOUZA FV, MATTOSO LH AND MARCONCINI JM. 2013. Characterization and comparative evaluation of thermal, structural, chemical, mechanical and morphological properties of six pineapple leaf fiber varieties for use in composites. *Ind Crops Prod* 43: 529-537.
- SENA NETO AR, CLARO PIC, SOUZA FVD, MATTOSO LHC AND MARCONCINI JM. 2017. Poly(lactic acid) composites reinforced with leaf fibers from ornamental variety of hybrid pineapple (Potyra). *Polym Compos* 38: 1228-1235.
- SILVA JM. 2016. Caracterização de fibras em germoplasma de abacaxi para a formulação de compósitos cimentícios. Dissertation (Mestrado em Recursos Genéticos Vegetais). Universidade Federal do Recôncavo da Bahia, Cruz das Almas, Bahia, Brazil. (Unpublished).
- SOKAL RR AND ROHLF FJ. 1962. The comparison of dendrograms by objective methods. *Taxon* 11: 33-40.
- SOUZA CPF, FERREIRA CF, SOUZA EH, NETO ARS, MARCONCINI JM, LEDO CAS AND SOUZA FVD. 2017. Genetic diversity and ISSR marker association with

- the quality of pineapple fiber for use in industry. *Ind Crops Prod* 104: 263-268.
- SOUZA EH, COSTA MAPC, SANTOS-SEREJO JA AND SOUZA FDV. 2014. Selection and use recommendation in hybrids of ornamental pineapple. *Revista Cienc Agrar* 45: 409-416.
- SOUZA EH, SOUZA FVD, COSTA MAP, COSTA JUNIOR DS, SANTOS-SEREJO JA, AMORIM EP AND LEDO CAS. 2012. Genetic variation of the *Ananas* genus with ornamental potential. *Genet Resour Crop Evol* 59: 1357-1476.
- VANIJAJIVA O. 2012. Assessment of genetic diversity and relationships in pineapple cultivars from Thailand using ISSR marker. *J Agric Technol* 8: 1829-1838.
- ZHU J, ZHU H, NJUGUNA J AND ABHYANKAR H. 2013. Recent Development of Flax Fibres and Their Reinforced Composites Based on Different Polymeric Matrices. *Mater Res* 6: 5171-5198.