

In Situ Characterization of Fruit Components, Yield Parameters, and SSR-based Genetic Diversity of Select Tall Coconut Populations in Capiz, Philippines for Breeding and Conservation

Elaine P. Laurio

ep1aurio@up.edu.ph

University of the Philippines Los Baños

Andrea Jean A. Dolotina

University of the Philippines Los Baños

Salvacion J. Legaspi

Capiz State University

Speedy D. Crisostomo

University of the Philippines Los Baños

Mariecris Rizalyn R. Mendoza

University of the Philippines Los Baños

Anand Noel C. Manohar

University of the Philippines Los Baños

Mary Ann M. Abustan

University of the Philippines Los Baños

Research Article

Keywords: Cocos nucifera, tall coconut, genetic diversity, characterization, nut yield, fruit component analysis

Posted Date: July 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10382778/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Capiz is one of the major coconut-producing provinces in the Visayas, but horti-morphological and molecular characterization of *in situ* coconut populations in the locality remains limited. This study conducted an *in-situ* characterization of the fruit components and yield parameters as well as the genetic diversity and population structure of three tall coconut populations in Capiz using 16 SSR markers. Among these populations, CAP3 (1,543 g) has the highest total fruit weight, followed by CAP2 (1,419 g), while CAP1 (1,238 g) exhibited the lowest fruit weight. Furthermore, CAP3 showed relatively high values in terms of the number of leaves per palm, number of bunches, number of nuts per bunch, and nut yield per palm and per hectare. To complement these horti-morphological findings, SSR analysis revealed a total of 56 alleles, with an average of 3.5 alleles per locus. Observed heterozygosity ($H_o = 0.821$) exceeded expected heterozygosity ($H_e = 0.658$) and PIC values ranged from 0.520 to 0.669. Analysis of molecular variance (AMOVA) revealed that 93% of total genetic variation occurred within populations, while only 7% was distributed among populations. Genetic differentiation among populations was low to moderate but statistically significant ($\Phi_{ST} = 0.067$; $p < 0.001$), suggesting a high degree of genetic connectivity and gene flow, overriding the effects of geographic distance between and among sampling sites. Furthermore, hierarchical clustering and principal coordinates analysis (PCoA) showed extensive intermixing of individuals with no clear geographic structuring. Overall, the populations exhibited moderate to high genetic diversity, highlighting their value as genetic resources for future breeding, conservation, and replanting purposes.

INTRODUCTION

Coconut (*Cocos nucifera* L.) is a perennial crop from the family Arecaceae. It is a diploid organism with 16 pairs of chromosomes ($2n = 32$). Being monotypic, it is the sole species in the genus *Cocos* (Batugal et al. 2009). Because of this, products of hybrid breeding are generally via intra-specific crosses. These crosses are usually between the two types of coconut: the tall and dwarf type. Tall varieties (*typica*) can grow more than 30 m tall with a juvenile period of six to seven years and a life of productivity for 60 to 70 years, while dwarf varieties (*nana*) are shorter and grow up to 10 to 12 m height with a juvenile phase of two to three years and a productive life of 30 to 40 years (Neto et al. 2016; Bandupriya & Perera 2021).

In the Philippines, coconut remains as one of the most economically significant crops due to its wide range of products and consistent export demand. In fact, the Philippines is the second largest producer of coconut in the world. However, the coconut industry in the country continues to face long standing challenges. Many coconut farmers, typically small hold in nature, rely on aging trees that have been showing a decline in productivity over the past few years. This is further compounded by environmental factors such as typhoons, drought, and pest outbreaks which contribute to further reduction in yield and threaten farm stability (Philippine Coconut Authority 2023; Department of Agriculture 2023).

Still, throughout the years, coconut breeding has seen significant advancements, particularly with the integration of biotechnology. Techniques such as embryo culture, molecular marker-assisted selection, and whole genome sequencing have provided new opportunities to accelerate varietal improvement and conserve genetic resources (Arumugam & Hatta 2022; Deng et al. 2025). These tools have made it possible to study genetic variation more precisely and efficiently, especially in perennial crops like coconut where conventional breeding takes decades.

While several studies using molecular markers have been conducted in the Philippines (Uddin et al. 2002; Osio et al. 2019), most are concentrated in major research locations or coconut-producing provinces in Luzon and Mindanao. In contrast, much less is known about the genetic composition of coconut germplasm (*in situ* and/or *ex situ*) in the Visayas, particularly the western part, which consists of Aklan, Antique, Capiz, Guimaras, Iloilo, and Negros Occidental.

Addressing these persistent bottlenecks requires not only agronomic interventions but also a clearer understanding of the phenotypic and genetic diversity present in local tall populations. The tall coconut populations in Capiz remain largely uncharacterized molecularly. Characterizing these populations is crucial in creating baseline information to support breeding programs and conservation efforts.

Modern molecular markers, particularly DNA-based markers, have transformed genetic analysis in coconut. These are categorized into hybridization-based markers such as Restriction Fragment Length Polymorphisms (RFLP), and PCR-based markers including Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Inter-Simple Sequence Repeats (ISSR/ISTR), Single Nucleotide Polymorphisms (SNPs), and Simple Sequence Repeats (SSRs) (Duran et al. 2009; Lebrun et al. 2001; Rivera et al. 1999). Among these, SSR markers have gained prominence due to their codominant inheritance, abundance across the genome, high polymorphism, and reproducibility.

Simple Sequence Repeats (SSRs), also known as microsatellites, are short tandemly repeated DNA motifs typically consisting of 1–6 base pairs (Ellegren 2000). They are widely distributed across the genome and are found in both coding and non-coding regions, but most commonly in non-coding regions, such as introns and intergenic sequences. SSRs can be classified based on the length of their repeat units, commonly appearing as mono-, di-, tri-, or tetranucleotide repeats (Ellegren 2004). SSRs are highly reproducible, co-dominant and multi-allelic markers that are unaffected by environmental conditions and exhibit high levels of polymorphism (Marwal et al., 2014). Since their development, SSR markers have been employed extensively in coconut research, especially in genetic diversity studies for both tall and dwarf populations (Teulat et al. 2000; Rajesh et al. 2008; Ribeiro et al. 2010; Sudha et al. 2023).

In the Philippines, one of the earliest reports of SSR-marker development in coconut was by Rivera et al. (1999), which was able to isolate and characterize 41 polymorphic microsatellites using ‘Tagnanan Tall’ (TAGT) coconut germplasm. Subsequently, Perera et al. (1999) conducted a similar study in ‘Sri Lankan Tall’ cultivar, which generated 8 microsatellites and was successfully utilized to assess variation between and within populations of coconut. Most recently, Caro and colleagues (2022) were able to identify 131 novel coconut-specific SSR markers through an extensive search in the ‘Catigan Green Dwarf’ (CATD) genome assembly (GCA_006176705.1) using a bioinformatics pipeline.

The present study aimed to characterize the phenotypic variation in fruit components, yield-related traits, and SSR-based genetic diversity of *in situ* tall coconut populations the Philippines, especially in the Capiz region. By documenting the diversity present in these naturally occurring populations, this study provides baseline information to support coconut germplasm conservation, future genetic improvement efforts, and the identification of populations with potential breeding value.

MATERIALS AND METHODS

Study design and sampling of coconut populations

This study was conducted as an observational *in situ* characterization of naturally occurring tall coconut populations in Capiz, Philippines. The main objective was to characterize variations in fruit components, yield traits, and molecular diversity among existing farmer-managed populations. Since the study involved evaluation of naturally occurring populations, no experimental treatments were imposed, and no field plots were established. Consequently, conventional experimental field designs, such as randomized complete block or split-plot designs were not applicable.

The field characterization and sampling procedures performed in this study followed the Standard Coconut Research Techniques (StanTech) Manual (Santos et al. 1996), which provides standardized protocols for the *in situ* characterization of coconut genetic resources. Based on the StanTech recommendations, characterization was conducted on individual palms using random palm sampling to obtain representative samples of each population.

In coordination with the Philippine Coconut Authority (PCA) and Capiz State University (CapSU), field visits were conducted on three coconut plantations in Capiz. They were located in Brgy. Matangcong, Sigma (CAP1), Brgy. Agdahon, Cuartero (CAP2), and Brgy. Manoling, President Roxas (CAP3) (Table 1; Fig. 1), representing coconut plantations that had not yet been characterized by either PCA or CapSU. Within each plantation, 30 healthy, mature, and phenotypically normal palms were randomly selected, while diseased or atypical palms were excluded. Each sampled palm served as an observational unit for both the phenotypic characterization and molecular analyses.

Table 1
Summary of three tall coconut populations in Capiz, Philippines, included for horti-morphological and molecular analysis.

Population Code	N*	Age (years)	Variety	Location
CAP1	30	60	Tall	Brgy. Matangcong, Sigma
CAP2	30	15 and 50	Tall	Brgy. Agdahon, Cuartero
CAP3	30	15	Tall	Brgy. Manoling, President Roxas

*N = number of palms sampled per population

Characterization of yield parameters

Phenotypic characterization of fruit components and yield parameters of the coconut populations were collected, in accordance also with the StanTech Manual protocols by Santos et al. (1996), including fruit characters, number of leaves, number of bunches, number of tennis-ball-sized nuts, and the number of nuts per bunch. Table 2 lists these fruit components and yield parameters and explains how these data were collected.

Table 2
Yield parameters used to describe the coconut populations in Capiz, Philippines (Santos et al. 1996).

Yield Parameters	Method of data collection
Fruit weight (grams)	Two mature nuts were randomly collected from each sampled palm. Fruit weight was recorded using a calibrated scale. Each nut was manually dehusked and separated into its major components (whole fruit, nut, shell, and meat).
Nut weight (grams)	Obtained from the same sampled nuts used for fruit weight analysis. After dehusking, the whole nut was weighed using a calibrated digital weighing scale, and the mean value per palm was computed.
Shell weight (grams)	Shells from the sampled nuts were separated from the kernel and weighed individually using a calibrated digital weighing scale. Mean shell weight per palm was calculated.
Meat weight (grams)	Fresh kernel (meat) was separated from the shell and weighed using a calibrated digital weighing scale. The average kernel weight per palm was recorded.
Husk weight (grams)	Obtained by subtracting the nut weight from the fruit weight.
Water weight (grams)	Obtained by subtracting the sum of shell and meat weights from the nut weight.
Number of leaves	The total number of fully opened green leaves per palm was manually counted during a single field assessment period.
Number of bunches	All reproductive bunches per palm were counted, starting from the most mature bunch down to the youngest bunch bearing approximately tennis ball-sized nuts, during a single assessment period.
Number of tennis-sized ball nuts	All nuts approximately equivalent to a tennis ball size within each sampled palm were manually counted during the field assessment.
Number of nuts per bunch	The total number of nuts present in each bunch was manually counted during a single assessment period.
Annual nut yield (nuts/palm/year)	Annual nut yield was estimated by multiplying the total number of nuts counted across all bunches per palm by four to approximate four production cycles per year. Estimates were reported with an assumed $\pm 5\%$ error margin.
Nut yield per hectare (nuts/hectare/year)	Estimated by multiplying the annual nut yield per palm by 135 palms, representing the standard planting density under a 9 m triangular planting system per hectare. Estimates were reported with an assumed $\pm 5\%$ error margin.

DNA isolation and quantification of leaf samples

Genomic DNA from leaf samples was extracted using a modified cetyltrimethylammonium bromide (CTAB) extraction protocol by Doyle and Doyle (1990). Leaf samples were cut into small pieces and immersed in liquid nitrogen using a mortar and pestle to produce fine powder. Powdered tissue was transferred into 2 mL microcentrifuge tubes containing 700 μ L of freshly made CTAB extraction buffer (100 mM Tris-HCl, 1.4 M NaCl, 20 mM EDTA, 2% w/v CTAB, 2% w/v polyvinylpyrrolidone, 4 g/L sodium bisulfite, and sterile nanopure water). Samples were then incubated in a hot water bath at 65°C for 1 hour, with gentle inversion every 15 minutes to ensure complete lysis. Phase separation was carried out by adding one volume (700 μ L) of chloroform:isoamyl alcohol (24:1) to the samples inside the laminar flow hood, with mixing by gentle inversion for 3–5 minutes. Using a Beckman Coulter centrifuge, the samples were spun at 10,000 rpm for 5 minutes at room temperature. Four hundred (400) μ L of the upper aqueous phase was carefully transferred to a sterile 1.5 mL microcentrifuge tube. DNA was precipitated by adding 0.80 volume of cold isopropanol to the tubes containing the aqueous phase, with gentle inversions. Samples were then incubated at -20°C for less than 10 minutes. Afterwards, the tubes were gently inverted again until thawed and centrifuged at 10,000 rpm for 3 minutes to separate the DNA. Resulting DNA pellets were washed by adding 500 μ L of Wash 1 solution (76% v/v ethanol, 2.5 M sodium acetate, and distilled water) and incubated for 20 minutes. Once decanted, it was followed by the addition of Wash 2 (76% v/v ethanol, 1.0 M ammonium acetate, and distilled water) for 5 minutes. The DNA pellet was air-dried under a laminar flow hood and resuspended in 50–100 μ L of TE-RNase (0.80 mg/ml) solution. To check the

quality of the DNA, 2% UltraPure agarose gel was used. DNA concentration was quantified using a Multiskan SkyHigh Microplate Spectrophotometer. Each sample was adjusted to a final concentration of 30 ng/μL by diluting the required volume of DNA extract with nuclease-free water to a total volume of 100 μL.

Amplification of SSR Markers

Sixteen SSR markers were utilized for PCR analysis (Table 3). In a 96-well PCR plate, a 10-μL mixture per sample was used to perform amplification. This consists of 60 ng of DNA template, 1× PCR Buffer (Promega Corporation, Madison, Wisconsin, USA), 1.5 mM MgCl₂, 0.2 μM dNTPs, 0.2 μM oligonucleotide primers (forward and reverse), 0.03 U *Taq* polymerase (Vivantis Technologies, Malaysia), and sterile nanopure water as needed. PCR was carried out using the Bio-Rad T100™ 96-well Thermal Cycler. Amplification has an initial denaturation performed at 94°C for 3 mins, followed by 30 cycles of denaturation at 94°C for 30 secs, annealing at different temperatures and final extension at 72°C for 1 min. PCR products were resolved on 8% non-denaturing polyacrylamide gels prepared with 1× TBE buffer, 8% acrylamide/bis (19:1) solution, 0.1% w/v ammonium persulfate (APS), 0.1% v/v Tetramethylethylenediamine (TEMED), and sterile nanopure water. Two (2) μL of BlueJuice™ loading dye mixed with 2 μL of PCR product were loaded in the gels. Gels were stained using 0.3 μL GelRed® Nucleic Acid Stain. Electrophoresis was carried out using a C.B.S. Scientific Triple Wide Mini-Vertical System™ (C.B.S. Scientific Company San Diego, California, USA) for 50–60 mins at 100 volts. Gel visualization and documentation were conducted using the Enduro GDS Touch Imaging System (Labnet International, Inc., Edison, NJ, USA).

Table 3
Information on the SSR markers used in this study.

Primer Name	Forward Sequence (5'-3')	Reverse Sequence	Repeat Motif	Annealing Temperature (°C)	References
CnCirE4	GCATGGTATTTCGGATTTG	ATGGTTCAGATTTGGACAGT	(CA) ₁₃ (TA) ₅	52	Lebrun 2004
CNZ33	GAGGTCAAAGTTATTTCCGAT	TTGCCCTATGACATTAAGA	(GT) ₁₁	53	Rivera et al. 1999; Lebrun et al. 2001
CnCirH4	TTAGATCTCCTCCCAAAG	ATCGAAAGAACAGTCACG	(CT) ₈ CGCT (CA) ₅ (CGCA) ₆	53	Lebrun 2002
mCnCIR187	CCAACATCTCTCGGT	AGTGAAAGCAGCATAAA	-	53	Lebrun 2004
CnCir167	GGTGGGTAAGTGAACATC	GTGATACAACGAACCCTC	(AG) ₁₈	57	Lebrun 2004
CNZ23	ATCAAAACATGACACCGTAAC	CTGATAGATGACAAGGTGTGG	(GA) ₁₈	55	Rivera et al. 1999; Lebrun et al. 2001
CN2A5	AAGGTGAAATCTATGAACACA	GGCAGTAACACATTACACATG	(CT) ₃₀	55	Rivera et al. 1999
CnCirD8	GCTCTTGATGTGGCTGCT	AGGCGTGTGAGATTGTGA	(AC) ₃₀	51	Lebrun 2004
CnCir46	GCTCGACTCAATAAACCT	CGAAGCAACCACACTT	-	53	Lebrun 2004
CnCir109	CCTACCACACCTTCCA	ATCATCTCAGTCCTTCTCA	(CTC) ₈ (CT) ₁₅	55	Lebrun 2004
CnCirE11	CAGAAACAGCCAATCAAGCAATA	ATCCATAATAGCCACTCAACAAAAA	(CA) ₁₀ (GA) ₂₃	51	Lebrun 2004
mEgCIR2433	CGAGGTGGTGATGCTTGTCTAC	CCGTGGGCCTTCTCTTTCTA	(GA) ₉	51	Billotte et al. 2005
CnCirE7	AACTACAACAAGAAACACG	AGAGAAGGGAAAAGGATA	(CA) ₈ (GA) ₁₉	53	Lebrun 2004
CnCir86	CCACTTGAGACTTGAAAC	ACTCACGCAAATATACTCA	(GA) ₁₀ (AG) ₁₄	55	Lebrun 2004
CnCirJ2	CCATTGTCATTGTTATTTTG	GTCACCATCTTCTCAGTTTC	(GT) ₉	53	Lebrun 2004
mCn16	AAGATTGAACATACTGCTGA	CTGAGATTCTGGCTG	-	53	Lebrun 2004
Note: Repeat motifs for some SSR loci were not available in the original publications or online databases.					

Data analysis

The horti-morphological data, including fruit component and yield parameters, were computed using descriptive statistics, including the mean and standard deviation. Differences among the three *in situ* populations were evaluated using the non-parametric Kruskal-Wallis test, followed by a pairwise comparisons among populations using Dunn's multiple comparisons test if significant differences were detected ($P < 0.05$). Statistical analyses were performed using GraphPad Prism version 10.6.1 (GraphPad Software, Boston, MA, USA).

Molecular data were analyzed using R (R Core Team; Posit Team 2025) for the estimation of genetic diversity parameters, including the average number of alleles per locus (N_a), effective number of alleles (N_e), expected heterozygosity (H_e), and observed heterozygosity (H_o). Polymorphic information content (PIC) was likewise calculated to evaluate marker informativeness. Genetic relationships were visualized through hierarchical clustering using the *hclust* package in R, based on Jaccard similarity coefficients and the unweighted pair group method with arithmetic mean (UPGMA). Principal coordinate analysis (PCoA) was also performed in R to further examine patterns of genetic variation. Analysis of Molecular Variance (AMOVA) was performed using GenAlEx 6.5 (Peakall & Smouse, 2005) to partition genetic variation within and among populations and to estimate fixation indices.

RESULTS

Horti-morphological analysis

Figure 2 shows the fruit component composition of the three tall populations, expressed as percentages of shell, meat, water, and husk weights relative to the overall fruit weight. Based on Fig. 2, the fruit weight (FW) of the three populations ranged from 1238 g to 1543 g. Among these populations, CAP3 has the highest fruit weight, followed by CAP2, while CAP1 exhibited the lowest fruit weight.

In CAP1, the husk weight represented the largest proportion of the fruit at 36.19%, followed by the meat weight at 33.84%. Shell weight accounted for 18.01%, while the water weight had the lowest weight at 11.95%. The CAP2 population showed a different composition, with meat weight comprising 31.78% of the total weight, husk weight 29.11%, water weight 23.04%, and shell weight 16.07%. Meanwhile, CAP3 exhibited a relatively balanced distribution of the husk and meat weight at 30.20% and 30.01%, respectively. Its water content accounted for 23.40% and shell weight for 16.40%.

Yield parameters of the three populations, including the number of leaves, number of tennis ball-sized nuts, number of bunches, number of nuts per bunch, nut yield per palm per year, and estimated nut yield per hectare per year, are presented in Fig. 3. Variations were observed among the populations across these yield-related traits. The number of leaves ranged from 27 to 29 leaves per palm, with CAP3 recording the highest value and CAP2 the lowest. The number of tennis ball-sized nuts varied from 5 to 10 nuts, with CAP2 exhibiting the highest number and CAP1 the lowest. However, no significant differences were observed among the three populations for this parameter.

For reproductive traits, CAP3 consistently outperformed the other populations. CAP3 produced the highest number of bunches (12), which was significantly greater than those of CAP1 (9) and CAP2 (8). Likewise, CAP3 recorded the highest number of nuts per bunch (6), significantly exceeding CAP1 (4) and CAP2 (3).

A similar trend was observed for yield performance. CAP3 produced the highest nut yield at 188 nuts/palm/year, whereas CAP1 and CAP2 produced 93 and 82 nuts/palm/year, respectively, both of which were significantly lower than CAP3. The estimated nut yield per hectare per year also differed significantly among populations. CAP3 attained the highest yield at 24,108 nuts/ha/year, while CAP1 and CAP2 produced 11,989 and 10,575 nuts/ha/year, respectively.

Genetic diversity analysis using SSR markers

Representative gel images of the amplified SSR fragments from the tall coconut populations are shown in Fig. 4. Across all loci, a total of 56 alleles were detected using 16 SSR markers. The number of alleles (N_a) ranged from 3 to 4, with an average of 3.5 alleles per locus (Table 4). The effective number of alleles per locus (N_e) ranged from 2.516 to 3.550 with a mean of 2.95. The expected heterozygosity (H_e) ranged from 0.602 to 0.718, with an overall mean of 0.658. On the other hand, the observed heterozygosity (H_o) ranged from 0.461 to 1.000, with a mean of 0.821. In most markers, H_o exceeded H_e , which indicates the presence of many heterozygotes. PIC values ranged from 0.520 to 0.669 with an average of 0.592, indicating that the SSR markers used were informative and capable of detecting high levels of polymorphism.

Table 4
Summary of diversity parameters using 16 SSR markers across the three tall coconut populations in Capiz, Philippines.

Locus	Na	Ne	He	Ho	PIC
CnCirE4	3	2.621	0.618	0.954	0.540
CNZ33	3	2.996	0.666	0.977	0.592
CnCirH4	4	3.493	0.714	0.978	0.661
mCnCIR187	4	3.490	0.713	0.800	0.662
CnCir167	3	2.525	0.604	0.818	0.520
CNZ23	4	3.550	0.718	0.558	0.669
CN2A5	3	2.960	0.662	0.512	0.588
CnCirD8	4	2.982	0.665	1.000	0.603
CnCir46	4	3.173	0.685	0.774	0.626
CnCir109	3	2.732	0.634	0.930	0.562
CnCirE11	4	2.870	0.652	0.767	0.594
mEgCIR2433	3	2.516	0.602	0.461	0.535
CnCirE7	4	2.985	0.665	0.988	0.599
CnCir86	3	2.854	0.650	0.727	0.576
CnCirJ2	4	2.797	0.642	0.989	0.581
mCn16	3	2.743	0.635	0.909	0.564
Mean	3.50	2.955	0.658	0.821	0.592

AMOVA based on binary distance matrix on GenAlEx 6.5 indicated that most of the genetic variation occurred within populations with 93% whereas only 7% was attributed to variation among populations (Table 5; Fig. 5). Φ_{ST} , an analogue of F_{ST} derived from AMOVA, ranges from 0 to 1, with lower values indicating little genetic differentiation and extensive gene flow, whereas higher values indicate strong genetic divergence due to limited allelic exchange. The obtained Φ_{ST} value of 0.067 ($p < 0.001$) indicates low to moderate but statistically significant genetic differentiation among the three tall coconut populations.

Table 5
Summary of AMOVA table showcasing percent variation among and within populations.

Source of Variation	df	SS	MS	Estimated Variance	% Variation	Φ_{ST}	p-value
Among Populations	2	77.611	36.806	0.837	7%	0.067	0.001
Within populations	87	1016.733	11.687	11.687	93%		
Total	89	1090.344			100%		

Hierarchical clustering was performed using *hclust* package in R based on Jaccard distance. The palms were grouped into three genetic clusters (Fig. 6), but these clusters did not correspond to their geographic origin. Cluster 1 contained the majority of individuals from all three populations, while cluster 2 consisted of a few palms from CAP1 and CAP3, and Cluster 3 consisted of two palms from CAP2. This result indicates a lack of clear population-level structuring and suggests high genetic similarity among populations.

PCoA based on Jaccard genetic distances was performed to visualize patterns of genetic variation and relatedness among the 90 tall individuals analyzed. The first two principal coordinates (PCoA1 and PCoA2) explained 9.88% and 9.0% of the total genetic variation, respectively, accounting for a combined 18.88% of the total variation (Fig. 7). This low percentage indicates that the majority of genetic variation is dispersed across the unplotted dimensions in the graph. This result further suggests that the genetic divergence between and among the analyzed populations is minimal.

DISCUSSION

The present study evaluated the yield performance and genetic diversity of the three *in situ* tall coconut populations in Capiz to better understand their productivity, genetic structure, and potential value for conservation and breeding. The results revealed substantial variation in fruit and yield characteristics among populations, alongside moderate to high levels of genetic diversity and low genetic differentiation. These findings provide insights into the factors influencing coconut production and highlight the importance of maintaining genetically diverse populations for future coconut improvement and conservation programs.

The observed variation in fruit component composition and yield parameters among the three tall populations indicates differences in fruit development and productivity. Among the populations, CAP3 consistently exhibited superior performance, recording the highest fruit weight, greater numbers of bunches and nuts per bunch, and the highest nut yield per palm and per hectare. Its balanced distribution of fruit components also suggests favorable fruit development and overall fruit quality. In contrast, CAP1 had the highest husk proportion and lowest water content, indicating greater biomass allocation to protective tissues. This trait is particularly important in the fiber industry, specifically for coco coir and coco peat. Likewise, CAP1 also exhibited the highest shell proportion, making it an ideal parent for shell-based by-products such as coconut charcoal briquettes and activated charcoal. Meat weight remained relatively consistent across populations, while shell weight showed minimal variation, suggesting that these traits are relatively stable across the sampled populations. Meat weight and thickness are usually the most economically valuable component in terms of coconut production. Hence, CAP3, which exhibits both high fruit weight and meat content represents a promising population for the identification of elite palms for copra and oil yield improvement.

Although CAP2 produced the highest number of tennis ball-sized nuts, this advantage did not translate into the highest actual yield, possibly due to unsuccessful fruit development, premature nut fall, or fruit abortion (Malarvannan 2012). Meanwhile, the slightly higher number of leaves observed in CAP3 may have contributed to greater photosynthetic capacity, thereby supporting enhanced fruit development and higher nut production. This observation was supported by the study of Natarajan et al (2010) in which the nut yield has a significant positive correlation with the number and length of leaves. In addition, higher leaf count also highly correlates directly to leaf emission rate which translates to higher number of reproductive bunches, since every leaf axil has the potential to generate a spadix. In terms of plant architecture, it is also observed that fronds provide structural support to developing nuts as they gain weight on the bunch. These results suggest that both reproductive and vegetative traits influence yield performance and may serve as useful indicators for selecting superior palms.

Using 16 SSR markers across three tall populations, a total of 56 alleles with an average of 3.5 alleles per locus were detected. The expected heterozygosity ranged from 0.602 to 0.718 (mean $H_e = 0.658$), while observed heterozygosity was consistently higher ($H_o = 0.821$), demonstrating moderate to high genetic diversity expressed at the individual level. Other studies also reported higher heterozygosity in tall coconut populations ranging from 0.35–0.59 (Rajesh et al. 2008; Sudha et al. 2023). This result can be attributed to the predominantly outcrossing mating system of tall coconuts, which is associated with their protandrous flowering habit (Foale 2020). PIC values which ranged from 0.520 to 0.669 (mean = 0.592) indicate that the markers were moderately to highly informative for detecting genetic variation among the tall populations.

AMOVA revealed that 93% of total genetic variation occurred within populations, whereas only 7% was attributed to variation among or between populations. Consistent with this pattern, the fixation index showed low to moderate but statistically significant genetic differentiation ($\Phi_{ST} = 0.067$; $p < 0.001$), reflecting a high degree of genetic connectivity among the three populations. Conventional guidelines (Hartl & Clark 1997 as cited by Sudha et al. 2023) interpret F_{ST} values < 0.05 as little genetic differentiation, 0.05–0.15 as moderate differentiation, 0.15–0.25 as great differentiation, and > 0.25 as very great differentiation. The Φ_{ST} value of 0.067 ($p < 0.001$), which is an analogue of F_{ST} , indicates little to moderate genetic differentiation among the three LAGT populations.

The predominance of within-population variation is consistent with the reproductive biology of tall coconuts and the extensive movement of planting materials among farming communities. In a similar study, Namia et al. (2008), using inverse sequence-tagged repeat analysis, studied indigenous tall coconut palms across eight locations in the Southern Tagalog (Batangas, Quezon, Laguna). They found an overall F_{ST} value of 0.24 among populations. The majority of genetic variation they obtained (97.05%) was found within populations, and some populations have negligible differences from each other due to the movement of seednuts between localities. Likewise, anecdotal evidence from farmers in Capiz reveals that seedlings are commonly derived from the progeny of older, more established palms and subsequently replanted across different sites. Such agricultural practices may have facilitated the ongoing gene flow among the populations despite the perceived physical distance among the sampling sites.

Patterns of population structure further support this interpretation. Hierarchical clustering based on Jaccard distance grouped the 90 palms into three genetic clusters; however, these clusters did not correspond to geographic origin but instead showed intermixing among individuals from different populations. Cluster 1 which consisted of the majority of the palms from all three sampling sites (Brgy. Matangcong, Sigma; Brgy. Agdahon, Cuartero; Brgy. Manoling, President Roxas) provides insights into the shared genetic background of tall populations in Capiz.

PCoA further reinforced these findings by showing substantial overlap among individuals from different populations and no distinct population separation along the first two principal coordinates. The absence of a strong population structure indicates that geographic distance alone has not been sufficient to restrict gene flow among the sampled populations. Consequently, the combined evidence from AMOVA, F_{ST} estimates, clustering analysis, and PCoA suggests that the three tall populations function as a largely interconnected genetic system.

Although this study provides the first integrated phenotypic and SSR-based characterization of *in situ* tall coconut populations in Capiz, the phenotypic characterization presented was based on a single assessment of naturally growing palms under existing farmer-managed conditions. Consequently, the observed variation in fruit components and yield-related traits may reflect the combined influence of genetic differences and environmental factors, including site-specific management practices and growing conditions. Nevertheless, the SSR-based molecular analyses complement the phenotypic observations by providing independent evidence of the underlying genetic diversity and population structure of the sampled populations. Future studies incorporating repeated assessments across years, common-garden evaluations, or multi-location trials would help distinguish genetic from environmental sources of phenotypic variation and provide a more robust assessment of the stability and genetic basis of the observed quantitative traits.

Despite these limitations, the findings of this study have important implications for coconut conservation and genetic improvement. The high levels of genetic diversity and heterozygosity observed within populations indicate that these tall populations represent valuable genetic resources that should be conserved both *in situ* and *ex situ*. Because most genetic variation is contained within populations rather than between them, conservation strategies should prioritize the maintenance of a large number of genetically diverse individuals within each population to minimize genetic erosion and preserve adaptive potential. The observed genetic connectivity among populations also suggests that farmer-mediated movement of planting materials has contributed to maintaining diversity, although continued monitoring is needed to prevent the loss of unique genotypes through genetic homogenization.

Furthermore, from a breeding perspective, the superior phenotypic performance of CAP3 under the conditions of this study suggests that it may represent a useful source population selecting elite palms and developing improved planting materials. At the same time, the moderate genetic differentiation detected among populations suggests that crossing individuals from different populations may still provide opportunities to capture useful genetic variation and enhance productivity. Integrating phenotypic performance data with molecular diversity information can therefore support more informed selection decisions and help establish breeding populations that combine high yield with broad genetic diversity. Such an approach would contribute to the development of resilient coconut varieties capable of sustaining production under changing environmental and management conditions.

CONCLUSION

The study analyzed the horti-morphological and molecular characteristics of *in situ* tall populations in Brgy. Matangcong, Sigma (CAP1), Brgy. Agdahon, Cuartero (CAP2), and Brgy. Manoling, President Roxas (CAP3) in Capiz. Variation in fruit components and yield parameters was observed among the three populations, with CAP3 recording relatively high values in terms of fruit weight, number of bunches, number of nuts per bunch, and nut yield per palm and per hectare. Molecular findings showed a moderate to high genetic diversity within populations and exhibited minimal population differentiation among them. Based on clustering and PCoA, high gene flow is likely occurring across the populations. This pattern is consistent with the open-pollinated and predominantly outcrossing nature of tall coconuts and is further reinforced by long-standing farmer-mediated seed exchange in the area. The findings from this study may help breeders or local farmers in Capiz to understand the diversity of tall coconuts in their locality, which may be used as a baseline for future breeding, conservation, and replanting purposes.

Declarations

Statement of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contribution

E.P.L. and A.J.A.D.: Methodology, Formal Analysis, Visualization, Writing – Original Draft. E.P.L.: Supervision. S.J.L.: Resources, Methodology, Supervision, Project Coordination. S.D.C. and M.R.R.M.: Conceptualization, Methodology, Formal Analysis, Supervision, Writing – Review & Editing. A.N.C.M.: Conceptualization, Supervision, Writing – Review & Editing. M.A.M.A.: Conceptualization, Methodology, Supervision, Project Administration, Funding Acquisition, Writing – Review & Editing. All authors have read and approved the final version of the manuscript.

Acknowledgement

The authors gratefully acknowledge the support of the Coconut Hybridization Program (CHP) of the Department of Science and Technology–Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (DOST-PCAARRD) for funding this research under the project titled “Agro-morphological and Genetic Characterization of In Situ Philippine Coconut Populations Towards Hybrid Breeding for Improved Yield and Special Traits.” A.J.A.D. likewise acknowledges the thesis funding support provided by the DOST Science Education Institute (SEI) through her scholarship grant. The authors also extend their sincere appreciation to the Philippine Coconut Authority and the Capiz State University for assisting and providing the coconut germplasm resources used in this study, and to the Genetics Laboratory of the Institute of Plant Breeding, University of the Philippines Los Baños (UPLB), for access to laboratory facilities and equipment.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Arumugam T, Hatta MA. Improving coconut using modern breeding technologies: Challenges and opportunities. *Plants* 2022 11(24):3414. <https://doi.org/10.3390/plants11243414>.
2. Bandupriya HD, Perera SA (2021) Coconut Genomics. In *Oil Crop Genomics* (pp. 161–188). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-70420-9_9.
3. Batugal P, Benigno D, Oliver J (2005) Coconut Genetic Resources International Plant Genetic Resources Institute-Regional Office for Asia, the Pacific and Oceania (IPGRI-APO). Serdang, Bioversity International.
4. Batugal P, Bourdeix R, Baudouin L (2009) Coconut breeding. In *Breeding plantation tree crops: Tropical species* (pp. 327–375). New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-71201-7_10.
5. Billotte N, Marseillac N, Risterucci AM, Adon B, Brottier P, Baurens FC, Singh R, Herran A, Asmady H, Billot C, Amblard P (2005) Microsatellite-based high density linkage map in oil palm (*Elaeis guineensis* Jacq.). *Theoretical and Applied Genetics* 110(4):754–65. <https://doi.org/10.1007/s00122-004-1901-8>.
6. Caro RE, Cagayan J, Gardoce RR, Manohar AN, Canama-Salinas AO, Rivera RL, Lantican DV, Galvez HF, Reaño CE (2022) Mining and validation of novel simple sequence repeat (SSR) markers derived from coconut (*Cocos nucifera* L.) genome assembly. *Journal of Genetic Engineering and Biotechnology* 20(1):71. <https://doi.org/10.1186/s43141-022-00354-z>.
7. Department of Agriculture (2023) DA wants to boost coconut industry productivity to benefit millions of farmers - Official Portal of the Department of Agriculture. Accessed from <https://context.ph/2023/11/27/da-wants-to-boost-coconut-industry-productivity-to-benefit-millions-of-farmers>.
8. Deng K, Yang S, Sisunandar S, Tran BM, Kottekkate M, Shaftang N, Mu Z (2025) Biotechnologies for Promoting Germplasm Resource Utilization and Preservation of the Coconut and Important Palms. *Horticulturae* 2025 11(12):1461. <https://doi.org/10.3390/horticulturae11121461>.
9. Duran C, Appleby N, Edwards D, Batley J (2009) Molecular genetic markers: discovery, applications, data storage and visualisation. *Current bioinformatics* 4(1):16–27. <https://doi.org/10.2174/157489309787158198>.

10. Ellegren H (2000) Microsatellite mutations in the germline:: implications for evolutionary inference. *Trends in genetics* 16(12): 551–558. [https://doi.org/10.1016/s0168-9525\(00\)02139-9](https://doi.org/10.1016/s0168-9525(00)02139-9).
11. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nature reviews genetics* 5(6):435–45. <https://doi.org/10.1038/nrg1348>.
12. Foale M, Biddle J, Bazrafshan A, Adkins S (2020) Biology, ecology, and evolution of coconut. In *Coconut Biotechnology: Towards the Sustainability of the 'Tree of Life'* (pp. 17–27). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-44988-9_2.
13. Hartl DL & Clark AG (1997) *Principles of population genetics*. Sunderland, MA: Sinauer associates.
14. Lebrun P, Baudouin L, Bourdeix R, Konan JL, Barker JH, Aldam C, Herran A, Ritter E (2001) Construction of a linkage map of the Rennell Island Tall coconut type (*Cocos nucifera* L.) and QTL analysis for yield characters. *Genome* 44(6):962–70. <https://doi.org/10.1139/g01-085>.
15. Malarvannan S (2012) Button shedding in coconut. CABI PlantwisePlus Knowledge Bank. <https://doi.org/10.1079/pwkb.20157800004>.
16. Marwal A, Gaur RK (2020) Molecular markers: tool for genetic analysis. In *Animal biotechnology 2020 Jan 1* (pp. 353–372). Academic Press. <https://doi.org/10.1016/B978-0-12-811710-1.00016-1>.
17. Namia MT, Sangalang JB, Altoveros NC, Rodriguez MJ, Fajardo NN (2008) Genetic structure and diversity of coconut (*Cocos nucifera* Linn.) in Southern Tagalog, Philippines as revealed by inverse sequence-tagged repeat (ISTR) analysis. *Philippine agricultural scientist* 90(3): 196–205. <https://www.cabidigitallibrary.org/doi/full/10.5555/20083155618>.
18. Natarajan C, Ganesamurthy K, Kavitha M (2010) Genetic variability in coconut (*Cocos nucifera*). *Electronic Journal of Plant Breeding* 1(5): 1367–1370.
19. Neto MF, Pereira TN, Geronimo IG, Azevedo AO, Ramos SR, Pereira MG (2016) Coconut genome size determined by flow cytometry: Tall versus Dwarf types. *Genet Mol Res.* 15(1):1–9. <https://doi.org/10.4238/gmr.15017470>.
20. Osio CA, Aquino GM, Aala WF, Lado JP, Tendero BJ, Cueto CA, Diaz MG, Laude RP (2019) Transcriptome analysis of 'Philippine Lono Tall' coconut (*Cocos nucifera* L.) endosperm reveals differential expression of genes involved in oil biosynthesis. *Philippine agricultural scientist* 102.
21. Perera L, Russell JR, Provan J, Powell W (1999) Identification and characterization of microsatellite loci in coconut (*Cocos nucifera* L.) and the analysis of coconut populations in Sri Lanka. *Molecular Ecology* 8(2).
22. Philippine Coconut Authority (2023) Coconut Statistics. <http://pca.gov.ph>.
23. Posit Team (2025) RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. <http://www.posit.co>.
24. Rajesh MK, Arunachalam V, Nagarajan P, Lebrun P, Samsudeen K, Thamban C (2008) Genetic survey of 10 Indian coconut landraces by simple sequence repeats (SSRs). *Scientia Horticulturae* 118(4):282–7. <https://doi.org/10.1016/j.scienta.2008.06.017>.
25. Ribeiro FE, Baudouin L, Lebrun P, Chaves LJ, Brondani C, Zucchi MI, Vencovsky R (2010) Population structures of Brazilian tall coconut (*Cocos nucifera* L.) by microsatellite markers. *Genetics and Molecular Biology* 33(4):696–702. <https://doi.org/10.1590/S1415-47572010005000077>.
26. Rivera R, Edwards KJ, Barker JH, Arnold GM, Ayad G, Hodgkin T, Karp A (1999) Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L. *Genome* 42(4):668–75. <https://doi.org/10.1139/g98-170>.
27. Santos GA, Batugal PA, Othman A, Baudouin L, Labouisse JP (1996) Manual on standardized research techniques in coconut breeding.
28. Sudha R, Rajesh MK, Niral V, Samsudeen K, Shil S, Ramchander S, Diwakar Y (2023) Analysis of genetic diversity and population structure in worldwide coconut germplasm (*Cocos nucifera* L.) using microsatellite markers. *Scientia Horticulturae* 309:111681. <https://doi.org/10.1016/j.scienta.2022.111681>.
29. Teulat B, Aldam C, Trehin R, Lebrun P, Barker JH, Arnold GM, Karp A, Baudouin L, Rognon F (2000) An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. *Theoretical and applied genetics* 100(5):764–71. <https://doi.org/10.1007/s001220051350>.
30. Uddin MS, Sangalang JB, Borromeo TH (2005) Study on diversity of selected coconut populations in Northern Luzon, the Philippines. *Bangladesh Journal of Agricultural Research* 31(1):139–52.

Figures

Map of Capiz, Philippines

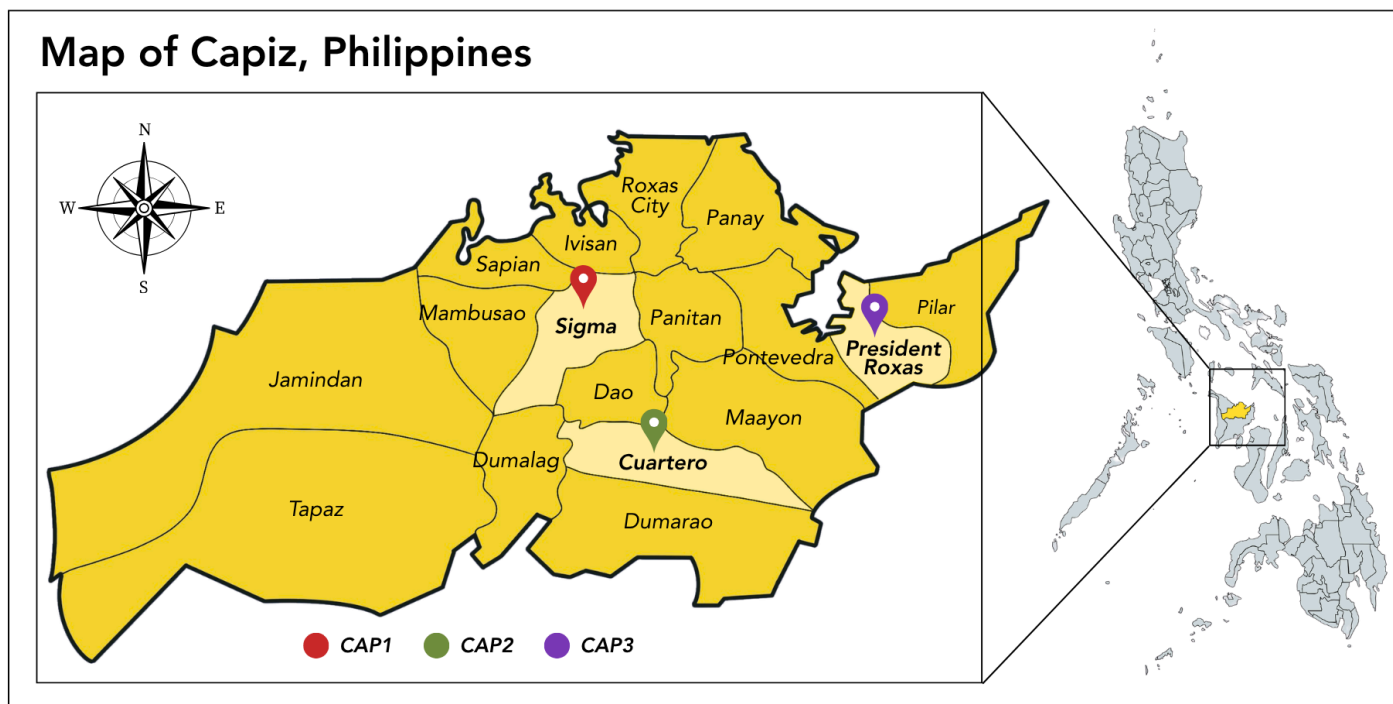


Figure 1

Geographic distribution of three *in situ* tall coconut populations (CAP1-CAP3) in Capiz, Philippines.

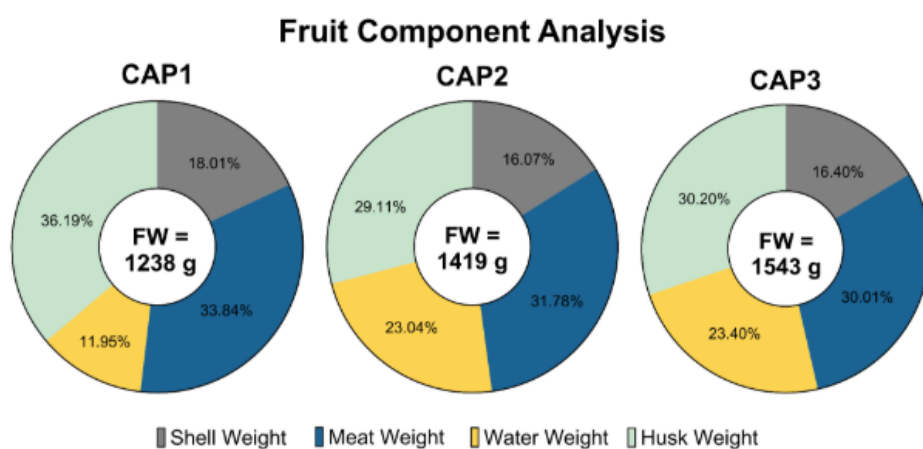


Figure 2

Fruit component analysis (FCA) of the three coconut populations in Capiz, Philippines. FW means fruit weight.

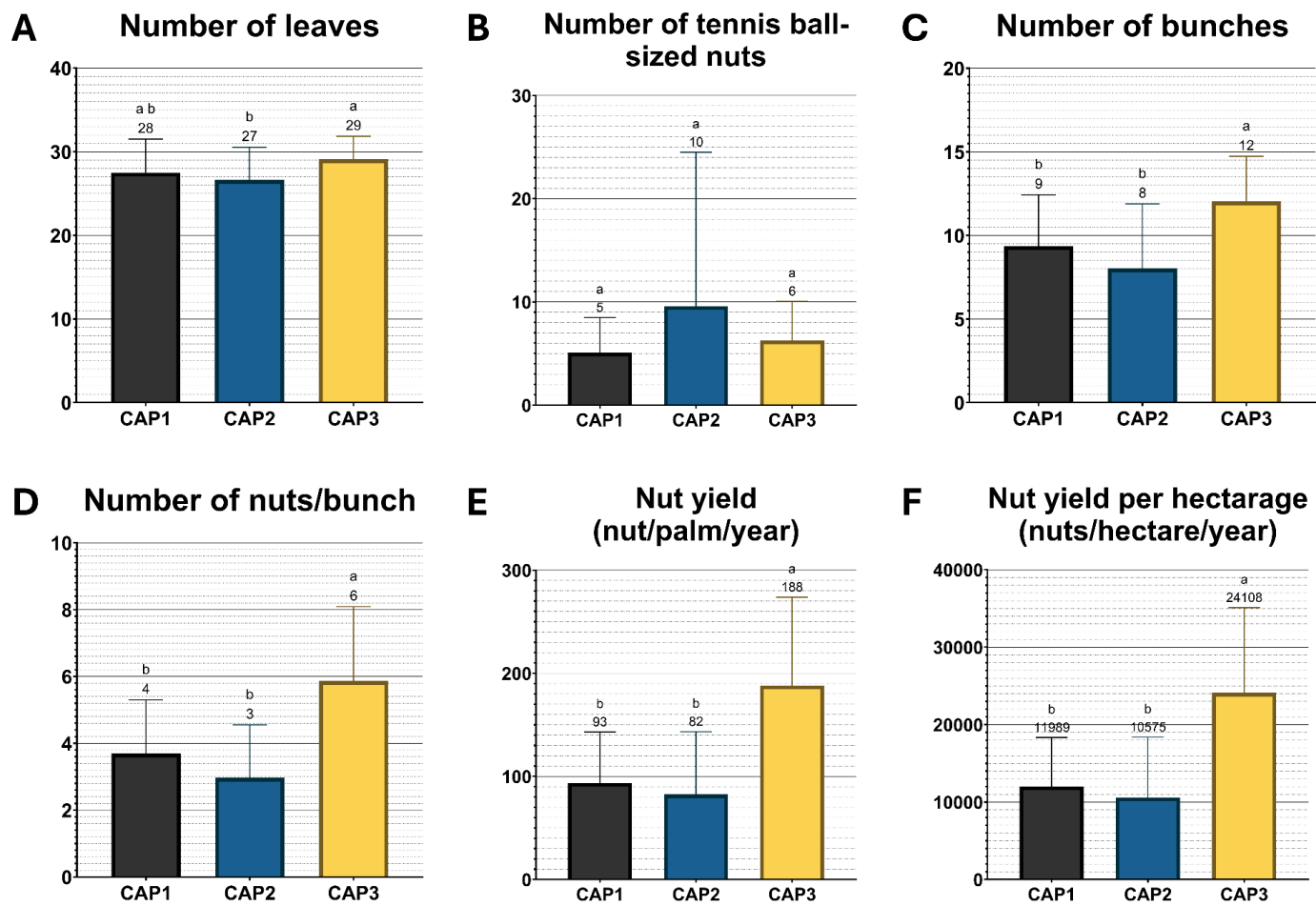


Figure 3

Yield parameters of the three coconut populations in Capiz, Philippines. A) Number of leaves; B) Number of tennis ball-sized nuts; C) Number of bunches; D) Number of nuts/bunch; E) Nut yield (nut/palm/year); F) Nut yield per hectare (nuts/hectare/year).

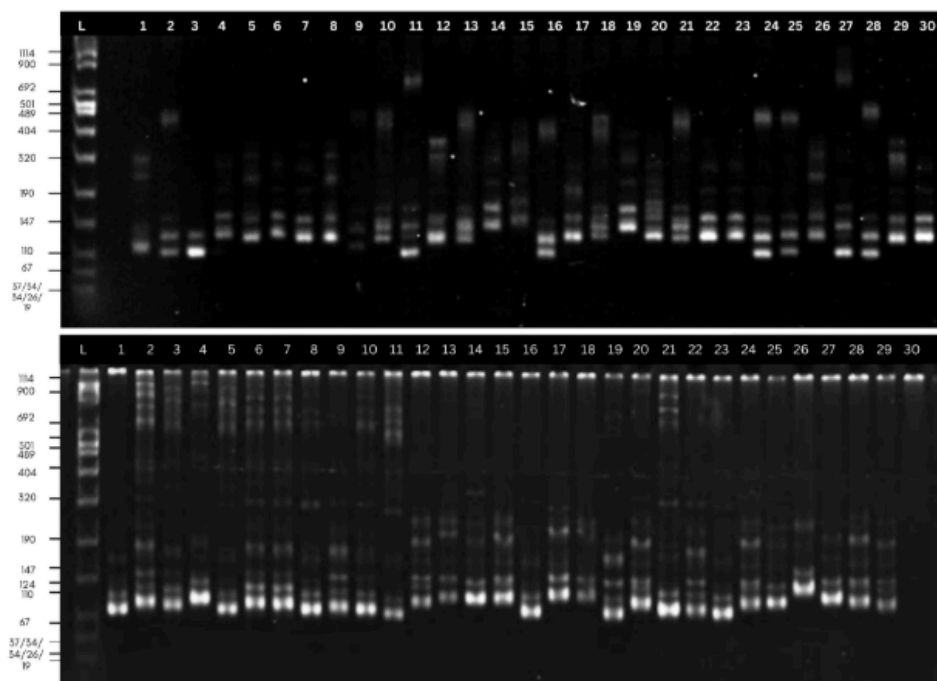


Figure 4

Representative gel profiles showing PCR amplification of simple sequence repeat (SSR) loci using primers CnCirE11 and CnCir109. L indicates the 1 kb Plus DNA Molecular Weight Marker VIII ladder, while lanes 1–30 represent amplified SSR fragments from individual tall coconut populations.

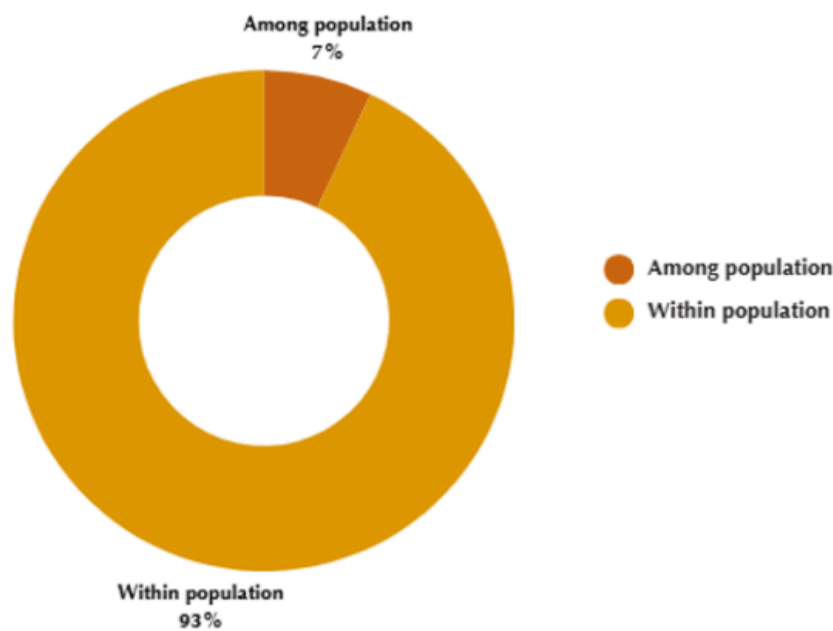


Figure 5

Analysis of Molecular Variance (AMOVA) illustrating the distribution of genetic variation among and within tall populations in Capiz.

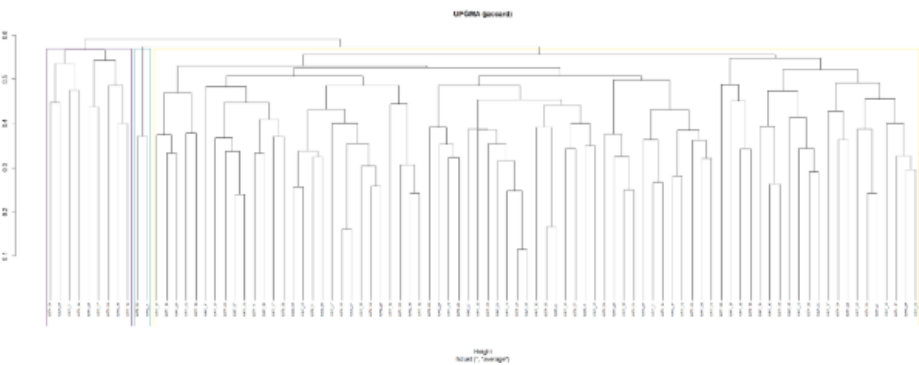


Figure 6

Hierarchical clustering of 90 samples from 3 barangays in Capiz, revealing 3 genetic clusters. Cluster 1 predominantly consists of palms from CAP2 (Brgy. Agdahon, Cuartero) and CAP1 (Brgy. Matangcong, Sigma) followed by palms from CAP3 (Brgy. Manoling, President Roxas).

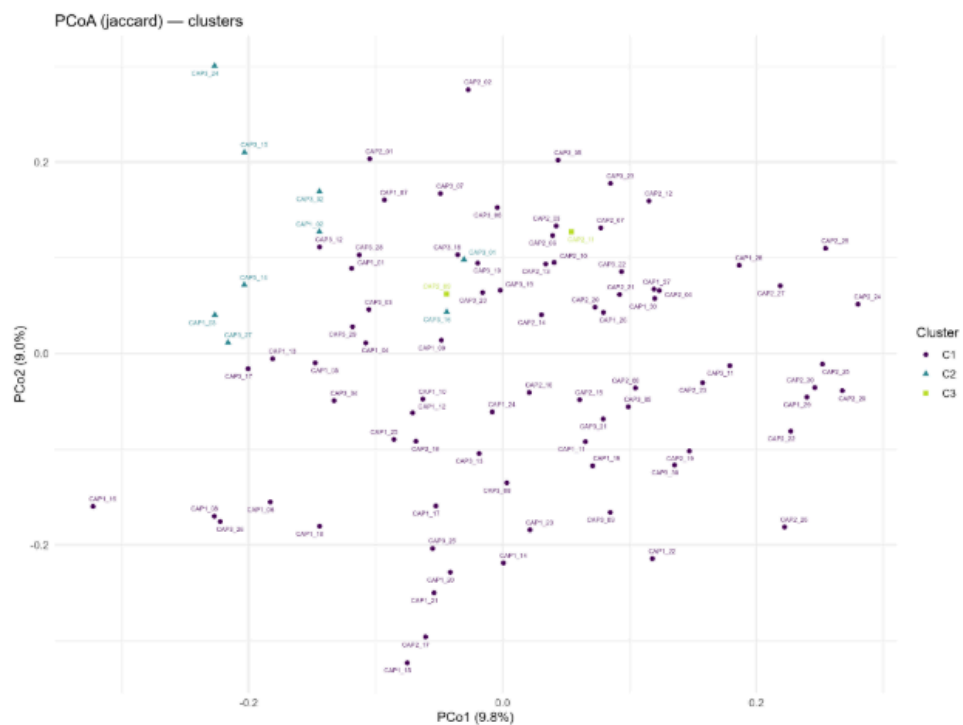


Figure 7

Principal coordinate analysis of tall samples in Capiz based on SSR data.