Development of Novel Coconut SSR Markers Derived From Genome-Wide Bioinformatics Prediction

Reina Esther S. Caro  
University of the Philippines Los Banos

Jesmar Cagayan  
University of the Philippines Los Banos

Roanne R. Gardoce  
University of the Philippines Los Banos

Anand Noel C. Manohar  
University of the Philippines Los Banos  
acmanohar@up.edu.ph

Alma O. Canama-Salinas  
University of the Philippines Los Banos

Ramon L. Rivera  
Philippine Coconut Authority

Darlon V. Lantican  
University of the Philippines Los Banos

Hayde F. Galvez  
University of the Philippines Los Banos

Consorcia E. Reaño  
University of the Philippines Los Banos

Research Article

Keywords: bioinformatics, CATD genome, coconut, marker-assisted breeding, SSR markers

DOI: https://doi.org/10.21203/rs.3.rs-723363/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

In the past, simple sequence repeat (SSR) marker development in coconut is achieved through microsatellite probing in bacterial artificial chromosome (BAC) clones or using previously developed SSR markers from closely related genomes. These coconut SSR markers are publicly available in published literatures and online databases; however, the number is quite limited. Here, we used a locally established, coconut genome-wide SSR prediction bioinformatics pipeline to generate a vast amount of coconut SSR markers. A total of 7,139 novel SSR markers were derived from the genome assembly of coconut ‘Catigan Green Dwarf’ (CATD). A subset of the markers, amounting to 131, were selected for synthesis based on motif filtering, contig distribution, product size exclusion, and success of in silico PCR in the CATD genome assembly. OligoAnalyzer-tool was also employed using the following desired parameters: %GC: 40–60%; minimum ΔG value for hairpin loop: -0.3 kcal/mol; minimum ΔG value for self-dimer: -0.9 kcal/mol; and minimum ΔG value for hetero-dimer: -0.9 kcal/mol. We have successfully synthesized, optimized, and amplified 131 novel SSR markers in coconut using ‘Catigan Green Dwarf’ (CATD), ‘Laguna Tall’ (LAGT), ‘West African Tall’ (WAT), and SYNVAR (LAGT x WAT) genotypes. Of the 131 SSR markers, 113 were polymorphic among the analyzed coconut genotypes. The development of novel SSR markers for coconut will serve as a valuable resource for mapping of quantitative trait loci (QTLs), assessment of genetic diversity and population structure, hybridity testing, and other marker-assisted plant breeding applications.

Declarations

Funding

The research leading to these results received funding from the Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development of the Department of Science and Technology (DOST-PCAARRD).

Conflicts of interests/competing interests

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Availability of data and material

Not applicable.
Introduction

Coconut (*Cocos nucifera* L.) is one of the most economically important crops in the Philippines. In 2017, the country produced 14.05 million metric tons of coconut and the value of production hit 120.3 million pesos (PSA 2018). The Philippines remained to be the top global supplier of coconut copra and desiccated coconut in both volume and total USD value as of 2010 (FAOSTAT 2013). Coconut oil, one of the many diversified products of coconut, ranked first among the top ten agricultural exports of the Philippines comprising 21.9 percent of the total agricultural exports in 2015 (PSA 2017).

Coconut is situated across the tropical and subtropical latitudes that are accessible to the equatorial Pacific Ocean current which possibly favored the evolution and dispersal of coconut. Coconut palms thrive well in humid coastal environments at about 18 degrees of latitude north or south of the equator where there is fertile soil, favorable temperature, and year-round rainfall (Foale 2003). Coconut belongs to the Indian center (II), and Indo-Malayan sub-center (II-A, where the Philippines belongs) in Vavilov's center of origin of cultivated plants (Vavilov 1926). It is generally classified into two types: tall and dwarf. The tall types are generally allogamous (heterozygous) or cross-pollinating, slow to mature; flower at 6-10 years with an economic life of 60-70 years. Dwarf types, on the other hand, are highly autogamous (homozygous), or mainly self-pollinating, early to flower at around 4-6 years after planting with a productive life of 30-40 years (Harries 1978; Meerow et al. 2012; Batugal et al. 2009).

Coconut is a diploid with 32 chromosomes (*2n* = 2*X* = 32). It belongs to the family Arecaceae (Palmaceae) in the subfamily Cocoideae and is the lone species of genus *Cocos* (Perera et al. 2007). The estimated genome size of coconut is approximately 2.6 Gbp comprising of 50-70% repetitive sequences (Alsaihati et al. 2014). Lantican et al. (2019) reported the estimated genome size of ‘CATD’ to be 2.14 Gbp. The abundance of repeat contents in the coconut genome becomes advantageous in the assessment and characterization of coconut varieties/populations using molecular marker techniques. The use of molecular tools offers a more accurate assessment than the conventional way of characterizing coconut which is through morphological and agronomical traits that are mostly influenced by many environmental factors (Perera et al. 2003).

Molecular markers have established its importance as a modern breeding tool for crop improvement (Xu and Crouch 2008; Kesawat and Das 2009; Sindhumole and Ambili 2011). The use of molecular tools can significantly accelerate the overall duration of breeding programs for coconut improvement. One of the
extensively used markers in molecular breeding and genetic diversity analyses is the simple sequence repeats (SSR). SSRs are short tandem repeats that have repeating units of di-, tri-, tetra- and pentanucleotides (Powell et al. 1996). They are approximately 1-8 bp long, abundant, and well distributed throughout the genome on which repeat units can vary between genotypes/individuals which make it a very useful tool in fingerprinting, genotyping and genetic diversity analyses (Sharma et al. 2008).

In the past, SSR marker development in coconut was achieved through microsatellite probing in bacterial artificial chromosome (BAC) clones or using previously developed SSR markers from closely related genomes (Rivera et al. 1999; Perera et al. 2003). These coconut SSR markers are publicly available; however, the number and distribution across chromosomes is quite limited for quantitative trait loci (QTL) mapping and genetic diversity studies. Fortunately, with the current advancements in next-generation sequencing (NGS) technologies, it has now become possible to mine SSRs across the entire genome. By using genome-wide bioinformatics prediction, we can generate a vast amount of SSR markers efficiently.

This study aims to provide a valuable resource of SSR markers for potential use in marker assisted selection breeding for coconut.

**Materials And Methods**

**Plant materials and leaf collections**

Leaf samples of the coconut parental genotypes ‘Catigan Green Dwarf’ (CATD), ‘Laguna Tall’ (LAGT), ‘West African Tall’ (WAT) and a synthetic variety denoted as SYNVAR (LAGT × WAT) used in this study were obtained from the Philippine Coconut Authority–Zamboanga Research Center (PCA–ZRC) in San Ramon, Zamboanga City, Philippines. Coconut leaflets coming from the youngest frond or the “first leaf” and are free from any pest damage were carefully chosen as samples. Three (3) leaflets were gathered from each of the left and right portions of the midrib near the base of the frond. The samples were transported to the Genetics Laboratory at the Institute of Plant Breeding - University of the Philippines Los Baños (IPB-UPLB), Laguna, Philippines, for DNA extraction.

*Table 1.* Coconut genotypes used in the study for screening the SSR markers.
| Try Number | Coconut Cultivars          | Code | Palm Number | Origin     |
|------------|---------------------------|------|-------------|------------|
| 1          | Catigan Green Dwarf       | CATD | 1715        | Davao City |
| 2          | West African Tall         | WAT  | 0519        | Ivory Coast|
| 3          | West African Tall         | WAT  | 0610        | Ivory Coast|
| 4          | West African Tall         | WAT  | 0704        | Ivory Coast|
| 5          | West African Tall         | WAT  | 0720        | Ivory Coast|
| 6          | Laguna Tall               | LAGT | 0107        | Davao City |
| 7          | Laguna Tall               | LAGT | 0508        | Davao City |
| 8          | SYNVAR (LAGT x WAT)       | AN17 | 4017        | Zamboanga City |

**Genomic DNA extraction of Coconut Parental Genotypes**

*Genomic DNA extraction.* A total of eight (8) individuals/palms of the coconut genotypes (Table 1) were collected and extracted with genomic DNA following the procedure adapted from Doyle and Doyle (1990) with modifications. DNA quality and yield were determined by electrophoresis in 1% UltraPure™ agarose (Invitrogen Corp., Carlsbad, California, USA) in 1× Tris-borate EDTA (TBE) running buffer at 100 V for 40 min, 0.5 ug mL⁻¹ ethidium bromide staining, and UV illumination at 300 nm using the Enduro GDS Touch Imaging System (Labnet International, Inc, Edison, New Jersey, USA). DNA concentration was estimated by visual comparison of the intensity of fluorescence with known concentrations of lambda (λ) DNA molecular weight standards (Sigma-Aldrich Inc., St. Louis, Missouri, USA).

**Development of SSR markers using the genome assembly of coconut ‘Catigan Green Dwarf’ (CATD)**

A set of 7,139 novel SSR markers was previously automatically generated based on the SSR loci annotation of the genome assembly of coconut ‘Catigan Green Dwarf’ (CATD) using GMATA software package (Wang and Wang 2016; Lantican et al. 2019). Given the vast amount of the predicted SSR markers, selection criteria were employed to obtain high-quality markers for eventual use in coconut genotyping. Motif filtering, contig distribution, and product size exclusion were used to further filter the predicted markers by manual checking. Markers with AT/AT and TA/TA repeat motifs were excluded in the selection. *In silico* PCR in the ‘CATD’ genome assembly (Lantican et al. 2019) was then performed to ensure *in vitro* SSR amplification prior to synthesis (Rotmistrovsky et al. 2004). OligoAnalyzer-tool (Integrated DNA Technologies, Inc., Coralville, Iowa) was also employed using the following desired parameters: %GC: 40-60%; minimum ΔG value for hairpin loop: -0.3 kcal/mol; minimum ΔG value for self-dimer: -0.9 kcal/mol; and minimum ΔG value for hetero-dimer: -0.9 kcal/mol for further filtering of the SSR markers.
PCR analysis

PCR was carried out with 10-uL reaction volume (15 ng genomic DNA, 1× PCR buffer (10 mM Tris pH 9.1 at 20 °C, 50 mM KCl, 0.01% Triton™ X-100; Vivantis Technologies, Malaysia), 1.5 mM MgCl₂, 0.2 mM dNTPs (Promega Corporation, Madison, Wisconsin, USA), 0.2 μM forward and reverse primer (Integrated DNA Technologies Pte. Ltd., Singapore) and Taq DNA polymerase (Vivantis Technologies, Malaysia). The temperature profile used is as follows: initial denaturation at 95 °C for 3 min, 30 cycles of denaturation (95 °C, 30 sec), annealing (45-60 °C depending on the primer pair, 45 sec), extension (72 °C, 1 min), and final extension at 72 °C for 5 min. Amplifications were carried out in the Applied Biosystems Veriti™ 96-well Thermal Cycler (Thermo Fisher Scientific, Madison, Wisconsin, USA). PCR products were resolved with electrophoresis using 8% non-denaturing polyacrylamide gel in 1× Tris-Borate EDTA buffer at 100 V for 60-75 min in the C.B.S. Scientific Triple Wide Mini-Vertical System™ (C.B.S. Scientific Company San Diego, California, USA), and visualized using 0.5 μg mL⁻¹ ethidium bromide staining and UV illumination using the Enduro GDS Touch Imaging System (Labnet International, Inc, Edison, New Jersey, USA). Gels were scored manually for the presence or absence of bands.

Results And Discussion

Local bioinformatics pipeline produces high-quality SSR markers

A vast amount of coconut SSR markers amounting to 7,139 was previously generated using a locally established bioinformatics pipeline (Lantican et al. 2019). A subset of these were pre-selected based on the characteristics of a good primer which includes sequence length (18-25 bp), GC content (40-60%), absence of runs of 4 or more of one base, absence of repetitive sequences, low dimerization capability and low hairpin loop formation (Zhao and Stodolsky 2004). Markers with AT/AT and TA/TA repeat motifs were excluded in the selection since these are the most common type of repeats found in the coconut/palm genome (Palliyarakkal et al. 2011; Xia et al. 2014; and Lantican et al. 2019) on which the high repeat content may hinder specificity of the markers and/or may result to non-specific amplification of products. Markers were also selected based on the distribution in the contig to cover the entire coconut genome.

In silico PCR in the CATD genome assembly was performed. This allows checking of contig specificity of the marker and ensures in vitro SSR amplification (Rotmistrovsky et al. 2004). Product size range of the markers was also limit to 80-400bp for easy visualization in gel and OligoAnalyzer-tool was used to check dimerization capability and formation of hairpin-loop of the primers to produce high-quality markers.

A total of 131 SSR markers were synthesized and 98% of these were comprised by dinucleotide repeats (or 2-mer) while the remaining 1 and 1% were composed of trinucleotide and tetranucleotide repeats, respectively, as shown in Figure 1. The predominance of dinucleotide repeats in coconut and other related species is supported by previous works of Rivera et al. (1999), Palliyarakkal et al. (2011), Xia et al. (2014), and Lantican et al. (2019). AG and GA motifs are the most abundant dinucleotide repeats found in the 131 SSR markers, with 29 and 18.3%, respectively. These are followed by CT (14.5%), TG (13.7%), TC
(11.5%), AC (7.6%), and GT (3.8%) repeats, while 1 and 1% of the markers contained a trinucleotide repeat of AAG and tetranucleotide repeat of ACAT, respectively. This result coincides with studies of Palliyarakkal et al. (2011) and Xia et al. (2014) on which AG/GA/TC/CT motifs were also the most common dinucleotide repeats found in coconut/palm genome.

The development of SSR markers using advanced bioinformatics tools in this study has become very efficient in generating high number of markers in coconut. The generated SSRs here are expected to contribute to the pool of available molecular markers (Lebrun et al. 1998; Perera et al. 1998; Xiao et al. 2013; Xia et al. 2014; Wu et al. 2019) for fingerprinting, genetic diversity analysis and QTL mapping as well as other relevant studies in coconut.

**SSR markers exhibit polymorphism in test coconut varieties**

All SSR markers showed successful amplification in coconut genomic DNA. Of the 131 SSR markers, 113 (86%) were polymorphic among the test coconut varieties while the remaining 18 (14%) were monomorphic. An average of 2.70 alleles per locus was observed across test varieties, implying a high degree of polymorphism of the selected SSR markers. The results obtained here are consistent with previous studies on which high levels of polymorphism are likely attributed to phenotypic variation and differences in the breeding behaviors of the dwarf and tall varieties which are said to be generally autogamous (self-pollinating) and allogamous (cross-pollinating), respectively (Perera et al. 1999; Rivera et al. 1999; Teulat et al. 2000). Representative gels of polymorphic SSR markers optimized among coconut genotypes are presented in Figure 2 on which distinct and good amplification patterns were observed. Also, majority of the polymorphic markers have AG repeats (29%) and GA repeat motif (19%) as shown in Figure 3. The product size of these markers ranged from 130 to 690 bp. The summary of the characteristics of the selected SSR markers are presented in Table 2 which includes the name of marker, annealing temperature, repeat motif, contig distribution, product size range and number of alleles.

**Table 2.** Characteristics of the selected coconut SSR markers with name, annealing temperature, repeat motif, contig number, product size range, and number of alleles.
| Ker ID | Tm (°C) | Optimized Ta (°C) | Predicted Length (bp) | Motif | Contig | Observed Size range (bp) | Polymorphisms | Number of alleles per locus |
|--------|---------|------------------|-----------------------|-------|--------|--------------------------|--------------|--------------------------|
| 7      | 55.6    | 56               | 299                   | AG    | 0      | 242-260; 404-480          | P            | 2                        |
| 172    | 53.55   | 57               | 280                   | TG    | 1      | 280-320                   | P            | 3                        |
| 329    | 55.3    | 55               | 342                   | CT    | 1      | 320-360                   | P            | 2                        |
| 653    | 54.15   | 57               | 240                   | GA    | 4      | 130-170; 170-230          | P            | 3                        |
| 765    | 55      | 53               | 211                   | GA    | 5      | 210-265                   | P            | 4                        |
| 995    | 56.55   | 54               | 196                   | AG    | 6      | 190-242                   | P            | 2                        |
| 1095   | 56.25   | 54               | 229                   | CT    | 7      | 230-250                   | P            | 4                        |
| 3414   | 56.85   | 50               | 289                   | AC    | 35     | 290-380                   | P            | 2                        |
| 3683   | 55.65   | 54               | 360                   | TG    | 40     | 320-370                   | P            | 2                        |
| 4036   | 56.8    | 55               | 254                   | GA    | 45     | 290-320                   | P            | 3                        |
| 4153   | 55.2    | 53               | 388                   | TC    | 46     | 380-400                   | P            | 2                        |
| 4627   | 55.25   | 51               | 237                   | AG    | 56     | 220-245                   | P            | 4                        |
| 4772   | 53.5    | 59               | 351                   | GT    | 59     | 380-450                   | P            | 3                        |
| 4830   | 54.75   | 53               | 341                   | AG    | 60     | 320-400                   | P            | 2                        |
| 4976   | 55.65   | 58               | 199                   | AC    | 64     | 170-200; 215-250          | P            | 6                        |
| 5103   | 56.15   | 54               | 307                   | CT    | 65     | 300-350                   | P            | 3                        |
| 5211   | 54.75   | 51               | 388                   | CT    | 68     | 190-210; 242-310          | P            | 6                        |
| 5746   | 55.45   | 50               | 152                   | CT    | 80     | 380-400                   | P            | 4                        |
| 5910   | 53.25   | 53               | 338                   | AG    | 82     | 330                       | M            | 1                        |
| 6063   | 54.7    | 53               | 394                   | AC    | 86     | 390-450                   | P            | 2                        |
| 6206   | 55.4    | 53               | 321                   | GT    | 89     | 300-340                   | M            | 2                        |
| 6376   | 55.6    | 53               | 368                   | CT    | 92     | 380-450                   | P            | 3                        |
| 6463   | 57      | 57               | 304                   | GA    | 96     | 230-240; 300-320          | P            | 4                        |
| 6507   | 53.4    | 51               | 336                   | GA    | 98     | 150-160; 180-200          | P            | 5                        |
| 6571   | 54.15   | 62               | 398                   | CT    | 99     | 400-450                   | M            | 1                        |
| 6672   | 56.05   | 56               | 363                   | GA    | 102    | 330-370                   | P            | 4                        |
| 7007   | 54.3    | 57               | 381                   | TG    | 109    | 350-380                   | P            | 2                        |
| 7162   | 55.65   | 56               | 259                   | GT    | 113    | 250-320; 400-420          | P            | 3                        |
| 7449   | 56.3    | 54               | 246                   | TC    | 119    | 240-260                   | P            | 4                        |
| 7553   | 54.5    | 50               | 267                   | AG    | 121    | 240-260                   | P            | 3                        |
| 7710   | 56      | 54               | 301                   | GA    | 127    | 220-230                   | P            | 3                        |
| 7859   | 55.65   | 54               | 355                   | AG    | 132    | 300-310                   | P            | 2                        |
| 8015   | 54.9    | 55               | 391                   | TG    | 135    | 250-320; 400-480          | P            | 2                        |
| 8444   | 55.5    | 54               | 295                   | GT    | 152    | 200; 250-320              | M            | 1                        |
| 8741   | 54.35   | 57               | 248                   | TC    | 166    | 320                       | M            | 1                        |
| 9091   | 55.1    | 58               | 373                   | AG    | 185    | 320-380                   | P            | 2                        |
| K9331 | 54.8  | 45   | 374  | GA   | 196  | 350-380 | P  | 2   |
| K9514 | 54.95 | 50   | 324  | TG   | 200  | 150-220; 240 | P  | 3   |
| K9655 | 54.55 | 50   | 381  | AG   | 207  | 242-280; 370-400 | P  | 2   |
| K9918 | 55.5  | 56   | 254  | GT   | 224  | 230-242 | P  | 2   |
| 10005 | 54.35 | 54   | 251  | CT   | 229  | 242-260; 320-350 | P  | 2   |
| 10146 | 54.9  | 58   | 208  | GA   | 237  | 210-225; 240-250 | P  | 3   |
| 10298 | 54.2  | 54   | 237  | AG   | 245  | 230-245 | P  | 2   |
| 10608 | 53.85 | 53   | 389  | GA   | 265  | 350-380 | P  | 2   |
| 10723 | 55.55 | 56   | 324  | TC   | 274  | 280-350 | P  | 3   |
| 10821 | 56    | 59   | 326  | AG   | 280  | 310-350 | P  | 3   |
| 11095 | 55.4  | 55   | 361  | AG   | 295  | 330-360 | P  | 2   |
| 11349 | 52.55 | 62   | 280  | AC   | 308  | 200-300-350 | P  | 3   |
| 16404 | 55.6  | 54   | 334  | AG   | 1122 | 300-400 | P  | 3   |
| 16553 | 53.45 | 48   | 339  | GA   | 1168 | 200-320-340 | P  | 3   |
| 16634 | 55.95 | 54   | 357  | AG   | 1203 | 400-500 | P  | 2   |
| 17050 | 54.85 | 52   | 397  | CT   | 1370 | 400-470 | P  | 2   |
| 17101 | 54.9  | 50   | 299  | AG   | 1401 | 250-290 | P  | 3   |
| 17156 | 54.7  | 53   | 293  | GA   | 1418 | 240-265 | P  | 3   |
| 17229 | 55.75 | 56   | 400  | AC   | 1462 | 280-320 | P  | 2   |
| 17487 | 54.25 | 52   | 288  | CT   | 1550 | 310-350 | P  | 3   |
| 17639 | 55.85 | 54   | 303  | TC   | 1621 | 280-340 | P  | 4   |
| 17725 | 55.1  | 55   | 234  | AG   | 1655 | 170-190; 220-250 | P  | 6   |
| 17797 | 55.25 | 53   | 370  | GA   | 1688 | 315-350 | M  | 2   |
| 17875 | 55.1  | 50   | 397  | AG   | 1750 | 300-400 | P  | 2   |
| 18331 | 55.45 | 55   | 368  | AG   | 1987 | 320-500 | M  | 1   |
| 18501 | 55.25 | 55   | 380  | TC   | 2084 | 350-400 | P  | 2   |
| 18573 | 55.45 | 57   | 383  | TG   | 2130 | 400-440 | P  | 2   |
| 18799 | 56.25 | 56   | 255  | AG   | 2306 | 240-320 | P  | 2   |
| 18903 | 56.2  | 56   | 396  | AG   | 2370 | 350-400 | P  | 2   |
| 18972 | 56.6  | 59   | 358  | TC   | 2409 | 230-260; 330-400 | P  | 5   |
| 19118 | 54.1  | 52   | 393  | CT   | 2535 | 230-320 | P  | 4   |
| 19193 | 55.75 | 54   | 398  | TC   | 2585 | 360-400; 690 | P  | 3   |
| 19386 | 55.25 | 52   | 354  | AG   | 2748 | 300-330; 400-420 | P  | 3   |
| 19611 | 54.2  | 50   | 364  | AC   | 2872 | 300-340 | P  | 3   |
| 19799 | 55.2  | 55   | 337  | TC   | 2988 | 320-410 | P  | 3   |
| 20018 | 54.5  | 53   | 362  | AG   | 3251 | 300-330 | P  | 3   |
| 20227 | 55.55 | 54   | 278  | TG   | 3455 | 320-400 | M  | 1   |
| 20739 | 54.6  | 58   | 270  | AG   | 4154 | 250-330 | P  | 3   |
| 21015 | 54.05 | 57   | 333  | CT   | 4606 | 130-160; 310-500 | P  | 2   |
| K21174 | 55.65 | 62 | 392 | AG | 4907 | 380-400 | P | 4 |
|--------|-------|----|-----|----|------|--------|---|----|
| K21493 | 55    | 55 | 315 | TC | 5710 | 220-230; 320-400 | P | 2 |
| 318    | 55.6  | 60 | 358 | TG | 1    | 310-330 | P | 2 |
| 425    | 55.9  | 59 | 217 | TG | 2    | 130-160; 180-200 | P | 3 |
| 808    | 56.1  | 55 | 225 | GA | 5    | 280-320; 400-430 | P | 3 |
| 3765   | 54.7  | 55 | 372 | TC | 40   | 290-380 | P | 3 |
| 4127   | 54.65 | 58 | 274 | TG | 46   | 160; 350-390 | M | 1 |
| 5054   | 54.1  | 57 | 391 | TG | 64   | 400-440 | P | 2 |
| 5329   | 55.05 | 53 | 320 | ACAT | 70 | 300-320; 360-410 | P | 2 |
| 5632   | 56.4  | 56 | 180 | TG | 78   | 180-200 | P | 5 |
| 6746   | 56.05 | 59 | 382 | TG | 103  | 380-400 | P | 2 |
| 6908   | 53.05 | 58 | 344 | CT | 107  | 200-300 | P | 3 |
| 7627   | 54.45 | 54 | 356 | TG | 124  | 240-350 | M | 1 |
| 7985   | 54.4  | 57 | 323 | AC | 135  | 320-340 | P | 2 |
| 8083   | 55.15 | 58 | 247 | GA | 137  | 200-250 | P | 3 |
| 8371   | 56.2  | 54 | 310 | AG | 148  | 310-400 | P | 3 |
| 8904   | 55.85 | 56 | 364 | GA | 178  | 320-350 | P | 2 |
| 9440   | 53.9  | 54 | 333 | GA | 199  | 300-340 | P | 2 |
| 9988   | 51.45 | 49 | 383 | GA | 228  | 230-320 | P | 4 |
| 10263  | 54.5  | 55 | 313 | CT | 244  | 290-320; 400-500 | P | 5 |
| 10632  | 54.25 | 52 | 376 | AC | 268  | 150 | M | 1 |
| 10681  | 55.85 | 57 | 337 | TG | 269  | 320-400 | P | 3 |
| 11807  | 55    | 55 | 305 | CT | 346  | 242-330 | P | 3 |
| 12241  | 56.9  | 55 | 315 | AG | 385  | 240-330 | P | 4 |
| 12746  | 54.3  | 54 | 383 | TG | 441  | 380-450 | P | 2 |
| 13043  | 55.25 | 55 | 383 | AG | 472  | 320-340 | P | 3 |
| 13232  | 54.7  | 58 | 165 | TC | 495  | 170-230 | P | 3 |
| 13632  | 57.2  | 57 | 379 | TG | 561  | 300-330 | P | 3 |
| 13852  | 53.55 | 52 | 184 | AG | 595  | 390-450 | M | 2 |
| 13946  | 56.9  | 57 | 363 | TC | 605  | 330-400 | P | 3 |
| 14272  | 54.55 | 53 | 371 | AG | 643  | 320-410 | P | 2 |
| 14692  | 56.85 | 53 | 370 | GA | 700  | 330-380 | P | 2 |
| 15137  | 56.1  | 56 | 337 | TC | 796  | 330 | M | 1 |
| 15508  | 56.75 | 57 | 374 | AG | 868  | 200-220; 320-400 | P | 3 |
| 15694  | 55.15 | 55 | 353 | AG | 908  | 320-390 | P | 3 |
| 15970  | 55.15 | 53 | 294 | TG | 994  | 220-280 | P | 3 |
| 5852   | 57.4  | 60 | 310 | AG | 82   | 310-400 | P | 3 |
| 17532  | 54.4  | 55 | 388 | AG | 1570 | 350-500 | P | 2 |
| 17684  | 53.3  | 53 | 176 | CT | 1636 | 170 | P | 2 |
| 18364  | 55.45 | 58 | 347 | GA | 2021 | 260-320; 470 | P | 4 |
| 19333  | 55.15 | 53 | 304 | AC | 2726 | 280-420 | P | 4 |
Microsatellites or SSR markers are a very useful molecular tool for studying genetic diversity and genotyping of coconut (Lebrun et al. 1998; Perera et al. 1998; Perera et al. 2003; Konan et al. 2007; Xiao et al. 2013). It has been extensively used in these analyses since SSR markers are abundant and well distributed throughout the genome, multi-allelic, co-dominant, highly polymorphic, and highly reproducible (Powell 1996; Mason 2015). Previous studies like Rivera et al. (1999), Perera et al. (2003), Xiao et al. (2013), and Wu et al. (2019) have already developed SSR markers in coconut for genetic diversity studies and these markers showed high levels of polymorphism as well.

Here, we demonstrated that a locally established bioinformatics pipeline can mine SSR markers from NGS data with actual utility in terms of amplification and distinguishing power across several varieties of coconut. The advantage of using a genome-wide bioinformatics prediction approach in marker development is its relatively fast and cost-effective way of generating vast amounts of markers (Anderson and Lubberstedt 2003; Gupta et al. 2013). SSRs and SNPs can be easily generated automatically in the genome sequences with the use of these programs or pipelines (Ching et al. 2002; Lantican et al. 2019).

Polymorphic markers in this study will be further used to genotype the coconut mapping population generated from a three-way cross of 'Pacific' LAGT and CATD, and 'Indo-Atlantic' WAT coconut for QTL mapping analysis. The development of novel SSR markers for coconut will serve as a valuable resource for mapping QTLs, assessment of genetic diversity and population structure, hybridity testing, and other marker-assisted plant breeding applications.

Acknowledgements
We express our gratitude to the Department of Science and Technology - Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development (DOST-PCAARRD) for funding the project “QTL mapping in coconut for high yield outstanding quality of copra oil and other coconut major by-products” under the program “Improvement of Coconut Varieties through Genomics, Genetics, and Breeding for a Competitive and Sustainable Philippine Coconut Industry (Genomics-Assisted Molecular Breeding)”. This research also has been made possible by the commitment and support of the Philippine Genome Center, and the Philippine Coconut Authority-Zamboanga Research Center (PCA-ZRC), which we also thank for providing the plant materials. We likewise acknowledge the valuable technical services rendered by Ms. Desiree Diaz for the completion of this work.

**References Cited**

ANDERSON JR, LUBBERSTEDT T. 2003. Functional markers in plants. *Trend Plant Science*. 8:554-560.

ALSAIHATI B, LIU W, LIN Q, AL-MSALLEM IS. 2014. Coconut genome *de novo* sequencing. In Proceedings of the Plant and Animal Genome XXII Conference, San Diego, CA, USA, 873 January 2014.

BATUGAL P, BOURDEIX R, BAUDOUIN L. 2009. Coconut Breeding. In: Jain SM, Priyadarshan PM (eds.) Breeding Plantation Tree Crops: Tropical Species, Springer Science Business Media, LLC, pp 327-373.

CHING A, CALDWEll KS, JUNG M, DOLAN M, SMITH OS, TINGEY S, MORGANTE M, RAFALSKI AJ. 2002. SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, 3:19-24.

DOYLE JJ, DOYLE JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.

FAOSTAT Database. 2013. http://faostat.fao.org/

FOALE M. 2003. The coconut odyssey: the bounteous possibilities of the tree of life. ACIAR Monograph No. 101, 132p.

GUPTA S, BHARALEE R, DAS R, THAKUR D. 2013. Bioinformatics tools for development of fast and cost effective simple sequence repeat (SSR), and single nucleotide polymorphisms (SNP) markers from expressed sequence tags (ESTs). *African journal of biotechnology*. 12. 4713-4721. 10.5897/AJB12.1410.

HARRIES HC. 1978. Evolution, dissemination and classification of *Cocos nucifera* L. *Botanical Review* 44(3): 265-320.

KESAWAT MS, KUMAR BD. 2009. Molecular markers: It's application in crop improvement. *Journal of Crop Science and Biotechnology* 12(4): pp169-181.

KONAN KJN, KOFFI KKE, KONAN JL, LEBRUN P, DERY SK, SANGARE A. 2007. Microsatellite gene diversity in coconut (*Cocos nucifera* L.) accessions resistant to lethal yellowing disease. *African Journal of Biotechnology* 6(4): 341-347.
LANTICAN D, STRICKLER S, CANAMA A, GARDOCE R, MUELLER L, GALVEZ H. 2019. De Novo Genome Sequence Assembly of Dwarf Coconut (Cocos nucifera L. ‘Catigan Green Dwarf’) Provides Insights into Genomic Variation Between Coconut Types and Related Palm Species. G3 (Bethesda). 2019 Aug 8;9(8):2377-2393. doi: 10.1534/g3.119.400215. PMID: 31167834; PMCID: PMC6686914.

LEBRUN P, N’CHO YP, SEGUIN M, GRIVET L, BAUDOUIN L. 1998. Genetic diversity in coconut (Cocos nucifera L.) revealed by restriction fragment length polymorphism (RFLP) markers. Euphytica 101(1): 103–108. https://doi.org/10.1023/a:1018323721803

MASON AS. 2015. SSR Genotyping. In: Batley J, editor. Plant Genotyping. Springer; New York, NY: 2015. pp. 77–89.

MEEROW AW, KRUEGER RR, SINGH R, LOW ETL, MAIZURAITHNIN M, OOI LCL. 2012. Coconut, Date, and Oil Palm Genomics. In: Schnell RJ and Priyadarshan PM (eds.), Genomics of Tree Crops. DOI 10.1007/978-1-4614-0920-5_10, © Springer Science Business Media, LLC, pp. 299-351.

PALLIYARAKKAL M.K., RAMASWAMY M., AND VADIVEL A. 2011. Microsatellites in palm (Arecaceae) sequences. Bioinformation. 7(7): 347–351.

PERERA L, RUSSELL RJ, PROVAN J, MCNICOL WJ, POWELL W. 1998. Evaluating genetic relationships between indigenous coconut (Cocos nucifera L.) accessions from Sri Lanka by means of AFLP profiling. Theor Appl Genet 96(3):545–550. https://doi.org/10.1007/s001220050772

PERERA L, RUSSELL JR, PROVAN J, POWELL W. 1999. Identification and characterization of microsatellites in coconut (Cocos nucifera L.) and the analysis of coconut population in Sri Lanka. Mol Ecol. 1999;8:344–346.

PERERA L, RUSSELL JR, PROVAN J, POWELL W. 2003. Studying genetic relationships among coconut varieties/populations using microsatellite markers. International Journal of Plant Breeding, Euphytica pp. 1-8.

PERERA PIP, HOCHER V., VERDEIL J.L., YAKANDAWALA DMD, WEERAKOON LK. 2007. Recent Advances in Anther Culture of Coconut (Cocos nucifera L.). In: Xu Z, Li J, Xue Y, Yang W. (eds) Biotechnology and Sustainable Agriculture 2006 and Beyond. Springer, Dordrecht, p 451

Philippine Statistics Authority. 2018. Selected Statistics on Agriculture 2013-2017. https://psa.gov.ph

Philippine Statistics Authority. 2017. Agricultural Foreign Trade Statistics of the Philippines: 2015. https://psa.gov.ph

POWELL W, MACHRAY GC, PROVAN J. 1996. Polymorphism revealed by simple sequence repeats. Trends in plant science 1(7): 215-222.
RIVERA R, EDWARDDS KJ, BARKER JHA, ARNOLD GM, AYAD G, HODGKIN T, KARP AA. 1999. Isolation and characterization of polymorphic microsatellites in Cocos nucifera L. Genome 42: 668–675.

ROTMISTROVSKY K, JANG W, SCHULER GD. 2004. A web server for performing electronic PCR. Nucleic acids research, 32 (Web Server issue), W108–W112. doi:10.1093/nar/gkh450

SHARMA A, NAMDEO AG, MAHADIK KR. 2008. Molecular markers: New prospects in plant genome analysis. Pharmacognosy Reviews 2(3): 23-34.

Sindhumole. P and Ambili SN. 2011. Marker assisted breeding in Coconut (Cocos nucifera L.). Gregor Mendel Foundation Proceedings, 30-32.

TEULAT B, ALDAM C, TREHIN R, LEBRUN P, BARKER JHA, ARNOLD GM, KARP A, BAUDOUIN L, ROGNON F. 2000. An analysis of genetic diversity in coconut (Cocos nucifera L.) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. Theor Appl Genet. 2000; 100:764–771. doi: 10.1007/s001220051350.

VAVILOV NI. 1926. Centres of Origin of Cultivated Plants. Bull. Appl. Bot. Genet. Plant Breed. 16: 1-248.

WANG X, WANG L. 2016. GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing. Frontiers in Plant Science 7, 1350.

WU Y, YAODONG Y, QADRI R, IQBAL A, LI J, FAN H, WU Y. 2019. Development of SSR Markers for Coconut (Cocos nucifera L.) by Selectively Amplified Microsatellite (SAM) and Its Applications. Tropical Plant Biology 12(1): 32–43.

XIA W, XIAO Y, LIU Z, LUO Y, MASON A, HAIKUO F, YANG Y, ZHAO S, PENG M. 2014. Development of gene-based simple sequence repeat markers for association analysis in Cocos nucifera. Molecular Breeding 34(2):1-11.

XIAO Y, LUO Y, YANG Y, FAN H, XIA W, MASONs AS, ZHAO S, SAGER R, QIAO F. 2013. Development of microsatellite markers in Cocos nucifera and their application in evaluating the level of genetic diversity of Cocos nucifera. Plant Omics Journal 6(3), pp 193-200.

XU Y, CROUCH JH. 2008 Marker-assisted selection in plant breeding: from publications to practice. Crop Science 48:391–407. doi: 10.2135/cropsci2007.04.0191

ZHAO S, STODOLSKY M. 2004. Bacterial Artificial Chromosomes: Library Construction, Physical Mapping (Volume 1). Human Press Inc., New Jersey, p 265.

Figures
Figure 1

Percentage of repeat motifs of the selected SSR markers.

Figure 2

Representative gels of polymorphic SSR markers optimized among coconut genotypes.
Figure 3

Percentage of polymorphic SSR markers per motif.