Mining and validation of novel simple sequence repeat (SSR) markers derived from coconut (Cocos nucifera L.) genome assembly

Reina Esther S. Caro1,2, Jesmar Cagayan1,2, Roanne R. Gardoce1, Anand Noel C. Manohar1*†, Alma O. Canama-Salinas1, Ramon L. Rivera3, Darlon V. Lantican1, Hayde F. Galvez1,2 and Consorcia E. Reaño4

Abstract
Background: 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 SSRs 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 bio-informatics pipeline to generate a vast amount of coconut SSR markers.

Results: A total of 7139 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. The 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 heterodimer, −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 × WAT) genotypes. Of the 131 SSR markers, 113 were polymorphic among the analyzed coconut genotypes.

Conclusion: 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.

Keywords: Bioinformatics, Catigan green dwarf genome, Coconut (Cocos nucifera L.), Marker-assisted breeding, SSRs

Background
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 hits 120.3 million pesos [19]. 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 [4]. Coconut oil, one of the many diversified products of coconut, ranked first among the top ten agricultural exports of the Philippines comprising 21.9% of the total agricultural exports in 2015 [18].

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° of latitude north or south of the equator where there is fertile soil, favorable temperature, and year-round rainfall [5]. 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 [26]. 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 after planting, and 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 [2, 6, 12].

Coconut is a diploid with 32 chromosomes (2n = 2× = 32). It belongs to the family Arecaceae (Palmaceae) in the subfamily Cocoideae and is the lone species of genus Cocos [17]. The estimated genome size of coconut is approximately 2.6 Gbp comprising of 50–70% repetitive sequences. Lantican et al. [9] 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 [15].

Molecular markers have established its importance as a modern breeding tool for crop improvement [7, 24, 31]. 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 [20]. 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 [23].

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 [15, 21]. These coconut SSR markers are publicly available; however, the number and distribution across chromosomes are 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.

Methods

Plant materials and leaf collections

Leaf samples of the coconut parental genotypes ‘Catigan Green Dwarf’ (CATD), ‘Laguna Tall’ (LAGT), and ‘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 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.

Genomic DNA extraction of coconut parental genotypes

A total of eight individuals/palms of the coconut genotypes were collected (Table 1). Genomic DNA was extracted following the procedure adapted from Doyle and Doyle [3] 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 μg mL⁻¹ ethidium bromide staining, and UV illumination at 300 nm using the Enduro GDS Touch

| Entry 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 × WAT) | AN17 | 4017        | Zamboanga City |
Imaging System (Labnet International, Inc, Edison, New Jersey, USA). DNA concentration was estimated by visual comparison of gel fragments 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)

Previously, a set of 7139 novel SSRs was automatically generated based on the SSR loci annotation of the genome assembly of coconut ‘Catigan Green Dwarf’ (CATD) using GMATA software package [9, 27]. 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 [9] was then performed to ensure in vitro SSR amplification prior to synthesis [22]. 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 heterodimer, −0.9 kcal/mol for further filtering of the SSRs (Fig. 1).

PCR analysis

PCR was carried out with 10 μL 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 s), annealing (45–60 °C depending on the primer pair, 45 s), 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

![Fig. 1 Schematic diagram depicting the SSR primer filtering pipeline](image-url)
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**

A total of 131 SSR markers were synthesized, and 98% of these were comprised by dinucleotide repeats (or 2-mer), while the remaining 2% are tri- and tetranucleotide repeats comprising of 1% each, as shown in Fig. 2. 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. In addition, tri- and tetranucleotide repeats of AAG (1.0%) and ACAT (1.0%) were also observed.

All SSRs showed successful amplification in coconut genomic DNA. Of the 131 SSRs, 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 SSRs. Representative gels of polymorphic SSRs optimized among coconut genotypes are presented in Fig. 3 on which distinct and good amplification patterns were observed. The product size of these markers ranged from 130 to 690 bp. The summary of the characteristics of the selected SSRs are presented in Table 2 which includes the name of marker, annealing temperature, repeat motif, contig distribution, product size range, and number of alleles (Fig. 4).

**Discussion**

The work of Lantican et al. [9] was able to identify genome-wide SSRs based on de novo prediction of repeat loci across the CATD genome assembly. However, the predicted loci were not screened nor tested in actual wet lab conditions. Here, the SSR markers generated were subjected to various filtering parameters that are advantageous based on genome distribution, repeat motif, and ideal thermodynamic properties. 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 [9, 13].
Table 2  Characteristics of the selected coconut SSRs with name, primer sequence, annealing temperature, repeat motif, contig number, and expected allele

| No. | Marker ID  | Primer sequence                  | Calculated $T_m$ (°C) | Optimum $T_a$ (°C)$^a$ | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|------------|----------------------------------|-----------------------|------------------------|---------------------------|-------|--------|--------------------------|---------------|--------------------------|
| 1   | CNMK7      | F: GAGGGAGAGGCTGTGATGR: GCATGTCAAGGGCTTGTT | 55.6                  | 56                     | 299                       | AG    | 0      | 242–260; 404–480 P       | 2             | 2                        |
| 2   | CNMK172    | F: TTCCACCTTACATCCACCT  R: TTATGGGTCGTAATGCG | 53.55                 | 57                     | 280                       | TG    | 1      | 280–320 P                 | 3             |                          |
| 3   | CNMK329    | F: AGTGCTTCAAGTTGGTGAGTG  R: CACCACTCCTTGCTTCTAAAT | 55.3                  | 55                     | 342                       | CT    | 1      | 320–360 P                 | 2             |                          |
| 4   | CNMK653    | F: AGCCAAATTTGGCATCAACTG  R: ATACATGGCAATGGGACCTT | 54.15                 | 57                     | 240                       | GA    | 4      | 130–170; 170–230 P       | 3             |                          |
| 5   | CNMK765    | F: AATGTTCACCAAGGGCTTTG  R: TGCAAGATTGGGACACTTTG | 55                    | 53                     | 211                       | GA    | 5      | 210–265 P                 | 4             |                          |
| 6   | CNMK995    | F: AAGACGAGAGACACCCGATAR  R: AAGTGCAACACAGCAACCTTGG | 56.55                 | 54                     | 196                       | AG    | 6      | 190–242 P                 | 2             |                          |
| 7   | CNMK1095   | F: CTCATTGGGCTTAACCTCAAG  R: AGTCTTGAAACGGGATACCA | 56.25                 | 54                     | 229                       | CT    | 7      | 230–250 P                 | 4             |                          |
| 8   | CNMK3414   | F: CCACGTCATGCGCATAAGA  R: AGCAACAGGCTCCTCATCTTC | 56.85                 | 50                     | 289                       | AC    | 35     | 290–380 P                 | 2             |                          |
| 9   | CNMK3683   | F: ACCTCCACAGTGGAAGCATGC  R: CCGTATTGCTATCCCTCCA | 55.65                 | 54                     | 360                       | TG    | 40     | 320–370 P                 | 2             |                          |
| 10  | CNMK4036   | F: TGAGTTTCCGGAGCAGAAAAG  R: ATACGCTTGAGGTGGTTG | 56.8                  | 55                     | 254                       | GA    | 45     | 290–320 P                 | 3             |                          |
| 11  | CNMK4153   | F: CATTCTGTGATGCGGATATG  R: AGGTATGGCAAGGAAAGGT | 55.2                  | 53                     | 388                       | TC    | 46     | 380–400 P                 | 2             |                          |
| 12  | CNMK4627   | F: TGCGTGAGCACAATCTCTGAT  R: GAGACGTAGTGGGAGAGAAA | 55.25                 | 51                     | 237                       | AG    | 56     | 220–245 P                 | 4             |                          |
| 13  | CNMK4772   | F: TGAGGTTACAGCATTCCTCC  R: TGAGTGATATACCCGCTTCC | 53.5                  | 59                     | 351                       | GT    | 59     | 380–450 P                 | 3             |                          |
| 14  | CNMK4830   | F: AAAAGGGTTGCGTATGCTCTCC  R: GGTGTTGACAGCCCTCTCC | 54.75                 | 53                     | 341                       | AG    | 60     | 320–400 P                 | 2             |                          |
| 15  | CNMK4976   | F: TGCCACCATATGAGCACAGA  R: CCCAGAGGTCCAAACTACAT | 55.65                 | 58                     | 199                       | AC    | 64     | 170–200; 215–250 P       | 6             |                          |
| 16  | CNMK5103   | F: AACGTTGCTCCTCTCGATAAG  R: TTACCCCGCTCTATCTTCT | 56.15                 | 54                     | 307                       | CT    | 65     | 300–350 P                 | 3             |                          |
| 17  | CNMK5211   | F: AAGCTGCAAAATGGGAGAGG  R: CAACAACTGTCGACTGG | 54.75                 | 51                     | 388                       | CT    | 68     | 190–210; 242–310 P       | 6             |                          |
| 18  | CNMK5746   | F: CTTGCGCAATGACACAAACAAG  R: AACCAGACAGAGGAGATGG | 55.45                 | 50                     | 152                       | CT    | 80     | 380–400 P                 | 4             |                          |
| 19  | CNMK5910   | F: GATTGGAGGCTGATGCTCTG  R: TTATGGGTCGGTAATGGAAT | 53.25                 | 53                     | 338                       | AG    | 82     | 330 M                      | 1             |                          |
Table 2 (continued)

| No. | Marker ID | Primer sequence | Calculated $T_m$ (°C) | Optimum $T_a$ (°C)* | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-----------|-----------------|-----------------------|---------------------|--------------------------|-------|--------|------------------------|---------------|-------------------------|
| 20  | CNMK6063  | F: CCACAACATGAAAGGAGGAR TGGACATGGAGTGGATTTGG | 54.7               | 53                  | 394                      | AC    | 86     | 390–450               | P             | 2                      |
| 21  | CNMK6206  | F: TGGCTAAAGGAGCCAGAGATG R: TCATGGAGAGCTGATGTTG | 55.4               | 53                  | 321                      | GT    | 89     | 300–340               | M             | 2                      |
| 22  | CNMK6376  | F: CATCATCTAAGCTGGCAACC R: TGGATAGGAGCCAGCAACA | 55.6               | 53                  | 368                      | CT    | 92     | 380–450               | P             | 3                      |
| 23  | CNMK6463  | F: GATACAGGTGGCAGGTGAAG R: TCCAGCCTCTGTGCGTTAT | 57                  | 57                  | 304                      | GA    | 96     | 230–240; 300–320     | P             | 4                      |
| 24  | CNMK6507  | F: CCAATACCTTTGGTGAGTC R: TGGATCTCCGTGCGATAAG | 53.4               | 51                  | 336                      | CT    | 99     | 150–160; 180–200     | P             | 5                      |
| 25  | CNMK6571  | F: TGAAGCTGATGTCGGCTACG R: GGAATAGTCTCAGGCATAAA | 54.15              | 62                  | 398                      | CT    | 99     | 400–450               | M             | 1                      |
| 26  | CNMK6672  | F: ATATAGGCAAGGGCCAGAAGT R: TGGCTAGTGGAGTGAATGCAA | 56.05              | 56                  | 363                      | GA    | 102    | 330–370               | P             | 4                      |
| 27  | CNMK7007  | F: GTGATTTTCGCTCTGCTTT T R: TGGAGAAAGCTTGGCAAATAAG | 54.3               | 57                  | 381                      | TG    | 109    | 350–380               | P             | 2                      |
| 28  | CNMK7162  | F: AGCCACAACCTGGAAAGGATCA R: GTGGAGAAAGATTGTACCTC | 55.65              | 56                  | 259                      | GT    | 113    | 250–320; 400–420     | P             | 3                      |
| 29  | CNMK7449  | F: TGCATACAAAGCTAACAAC R: GTGAGTTGTGGCGCTCTCTTTCA | 56.3               | 54                  | 246                      | TC    | 119    | 240–260               | P             | 4                      |
| 30  | CNMK7553  | F: TTAGGATGCTGCAACATGG R: GCTTGGTAGGGCCAGCTCATT | 54.5               | 50                  | 267                      | AG    | 121    | 240–260               | P             | 3                      |
| 31  | CNMK7710  | F: TCAACTGCGATGCTGTTGG R: CTGACGAGGGCGACAGAGAA | 56                  | 54                  | 301                      | AG    | 127    | 220–230               | P             | 3                      |
| 32  | CNMK7859  | F: ATGGCGCTGGATATGGCTGTC R: TCTCAGTCTCAGTGAAACA | 55.65              | 54                  | 355                      | AG    | 132    | 300–310               | P             | 2                      |
| 33  | CNMK8015  | F:GGCAGTTCTCAGTGTCCTAC R: TGCTTTAAACAAAGGCTTTGTG | 54.9               | 55                  | 391                      | TG    | 135    | 250–320; 400–480     | P             | 2                      |
| 34  | CNMK8444  | F: GGTCGGGAGATTGAGTCTTGA R: AGCAGAGAGGAGAGTGGAT | 55.5               | 54                  | 295                      | GT    | 152    | 200–250; 320        | M             | 1                      |
| 35  | CNMK8741  | F: ACCCGAGTGGTTGGGAAGAAC R: TGGCAAGCATCCTTAGTGG | 54.35              | 57                  | 248                      | TC    | 166    | 320                   | M             | 1                      |
| 36  | CNMK9091  | F: GCCTGTACCATCAGGATCTGC R: GCTGCTGTCCTCGATCAGAC | 55.1               | 58                  | 373                      | AG    | 185    | 320–380               | P             | 2                      |
| 37  | CNMK9331  | F: AGTGTTCCTGTTGGTGTAGG R: AGTGGAGAATCGGAATGCAA | 54.8               | 45                  | 374                      | GA    | 196    | 350–380               | P             | 2                      |
Table 2 (continued)

| No. | Marker ID | Primer sequence | Calculated $T_m$ (°C) | Optimum $T_a$ (°C) | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-----------|-----------------|-----------------------|-------------------|--------------------------|-------|--------|------------------------|---------------|------------------------|
| 38  | CNMK9514  | F: TGAGATGAGATGGGTGGACAGTAGG<br>R: ATGAATGGGAGGTGACAGTAGG | 54.95 | 50 | 324 | TG | 200 | 150, 220–240 | P | 3 |
| 39  | CNMK9655  | F: TTGGACTGATCCGCAACTAGGG<br>R: CCAATCCAACCCAGGACCTAGG | 54.55 | 50 | 381 | AG | 207 | 242–280, 370–400 | P | 2 |
| 40  | CNMK9918  | F: TTGGACTCCACCCAGGACTAGG<br>R: TCCTTCCAGGACAGGTGCC | 55.5 | 56 | 254 | GT | 224 | 230–242 | P | 2 |
| 41  | CNMK10005 | F: TGCTGACGTCATCCATCTG<br>R: CACATCTCCCTCGATTCA | 54.35 | 54 | 251 | CT | 229 | 242–260, 320–350 | P | 2 |
| 42  | CNMK10146 | F: AATCGAATACGTGGGCAAC<br>R: GCTTGTAGCACAACACCTG | 54.9 | 58 | 208 | GA | 237 | 210–225, 240–250 | P | 3 |
| 43  | CNMK10298 | F: TCACACATCATCCATACCAAGAA<br>R: TTATGCACAGTGCCAAT | 54.2 | 54 | 237 | AG | 245 | 230–245 | P | 2 |
| 44  | CNMK10608 | F: AGGAATCTACGGTCTGTGTTGG<br>R: GCTGATATTGTCAGTTGGAG | 53.85 | 53 | 389 | GA | 265 | 350–380 | P | 2 |
| 45  | CNMK10723 | F: TTGTTCAGGTGGAAGTCTG<br>R: AGCTTTCTGTGCAACTAGCC | 55.55 | 56 | 324 | TC | 274 | 280–350 | P | 3 |
| 46  | CNMK10821 | F: AATACGGACCTGTTACTCTTG<br>R: TGCAATGTGGGAAGACACAT | 56 | 59 | 326 | AG | 280 | 310–350 | P | 3 |
| 47  | CNMK11095 | F: GATGCGCAGTGGGAAC<br>R: AATGCCGAGCAACACTGGAAC | 55.4 | 55 | 361 | GC | 295 | 330–360 | P | 2 |
| 48  | CNMK11349 | F: ATGGCAATTGGAGGATGGA<br>R: ACCGTTCTTGGGAATAGTG | 52.55 | 62 | 280 | AC | 308 | 200, 300–350 | P | 3 |
| 49  | CNMK16404 | F: GGAAATCTGAAAGCAAGCAG<br>R: AGGCATGTTAGAAGAAGACAC | 55.6 | 54 | 334 | AG | 1122 | 300–400 | P | 3 |
| 50  | CNMK16553 | F: GCGGAAATGTGCTTAAGGATTTG<br>R: AGGAATGTGAGTAGGTCT | 53.45 | 48 | 339 | GA | 1168 | 200, 320–340 | P | 3 |
| 51  | CNMK16634 | F: CGACCTGAAGTCAGCTTGG<br>R: GCACCGTATTACCCTTACGTC | 53.95 | 54 | 357 | GC | 1203 | 400–500 | P | 2 |
| 52  | CNMK17050 | F: ATTGGCAGTGAGGGGCAAC<br>R: ATGAGCAACCCATGTGTGAT | 54.85 | 52 | 397 | CT | 1370 | 400–470 | P | 2 |
| 53  | CNMK17101 | F: CAGGAACTGCTTACCAAACT<br>R: CGAGAAACCGCTGAATGAGA | 54.9 | 50 | 299 | AG | 1401 | 250–290 | P | 3 |
| 54  | CNMK17156 | F: GTGTCTGCGAAATCTGCT<br>R: TGCTTGAATACGCGGACACCAAA | 54.7 | 53 | 293 | GA | 1418 | 240–265 | P | 3 |
| 55  | CNMK17229 | F: CAACAGTGGAGCAACACAGG<br>R: TGTCATCGGCAACTGAG | 55.75 | 56 | 400 | AC | 1462 | 280–320 | P | 2 |
| No. | Marker ID | Primer sequence | Calculated T<sub>m</sub> (°C) | Optimum T<sub>a</sub> (°C)* | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-----------|-----------------|-----------------------------|-----------------------------|--------------------------|-------|--------|----------------------------|---------------|------------------------|
| 56  | CNMK17487 | F: TGCCA ATG TTA AGG GAT GC  
R: TCC ATG CAA AGG CAC ACT GTG | 54.25 | 52 | 288 | CT | 1550 | 310–350 | P | 3 |
| 57  | CNMK17639 | F: AAATG TGG TGG GTC CTC TCA  
R: GCC AGG GAG CACT GAG AAG | 55.85 | 54 | 303 | TC | 1621 | 280–340 | P | 4 |
| 58  | CNMK17725 | F: GCT GAT GTC GTT GTT TC A  
R: CCT GAC ACG AGG GAT GAG AA | 55.1 | 55 | 234 | AG | 1655 | 170–190; 220–250 | P | 6 |
| 59  | CNMK17797 | F: GTG GCC TCT AGT GCC CTC TT  
R: ATC CGT TGC GAG CT TAT C | 55.25 | 53 | 370 | GA | 1688 | 315–350 | M | 2 |
| 60  | CNMK17875 | F: GGC TGG GTT CAA AT TCT TT  
R: TGA AGG GCC ACC TAA GGC ACT | 55.1 | 50 | 397 | AG | 1750 | 300–400 | P | 2 |
| 61  | CNMK18331 | F: TGA AAG TGA GAC CTT GGA  
R: TGT AAT CAG TCG GTT CAG A | 55.45 | 55 | 368 | AG | 1987 | 320–500 | M | 1 |
| 62  | CNMK18501 | F: CTG GAC GAA AAG TGT GTC T  
R: CTC CAC GGG TAT TGA CTA CCA | 55.25 | 55 | 380 | TC | 2084 | 350–400 | P | 2 |
| 63  | CNMK18573 | F: TGT GTT CGC AGT TCG TG CTA  
R: AGG CTC TTG TCG TC A ACT G | 55.45 | 57 | 383 | TG | 2130 | 400–440 | P | 2 |
| 64  | CNMK18799 | F: GCC TTG CAT ATT CC TAC CGA  
R: GTG AGG AGC AGT GGA AAG | 56.25 | 56 | 255 | AG | 2306 | 240–320 | P | 2 |
| 65  | CNMK18903 | F: CGG ACT CCG GGA GAT AGC TG  
R: TGG C TGC TAA GT TGC CAC C | 56.2 | 56 | 396 | AG | 2370 | 350–400 | P | 2 |
| 66  | CNMK18972 | F: GGT GTT TGCC AT CTT CTA TT  
R: TAC GGT TGG GAC CAA AGG AGT | 56.6 | 59 | 358 | TC | 2409 | 230–260; 330–400 | P | 5 |
| 67  | CNMK19118 | F: ACC C AT CCA AAT GGG AC  
R: TCAG GGT GGC CAT GTA AATA | 54.1 | 52 | 393 | CT | 2535 | 230–320 | P | 4 |
| 68  | CNMK19193 | F: ATG TGT TGG GGG ACG ATG AGT  
R: ACC TCG CAT GAT GGA AACT G | 55.75 | 54 | 398 | AG | 2585 | 360–400; 690 | P | 3 |
| 69  | CNMK19386 | F: AAG GGT GTT GAG TGT GGG TTG  
R: CTA AAC AGG GCA AAG GCA A | 55.25 | 52 | 354 | AG | 2748 | 300–330; 400–420 | P | 3 |
| 70  | CNMK19611 | F: ATC C AT CCA AT C AT CTC GAG  
R: GAC CGC TAA GCT CTT GAC AC | 54.2 | 50 | 364 | AC | 2872 | 300–340 | P | 3 |
| 71  | CNMK19799 | F: CTT GTC TGG GAT TGC TCA GT  
R: CCA AGG CAC CGG GAA GAC TTT | 55.2 | 55 | 337 | TC | 2988 | 320–410 | P | 3 |
| 72  | CNMK20018 | F: TGA CAA GGT TCT TGA CCA CCA  
R: TGC GTAT GTC TGG CAA ACG TGA | 54.5 | 53 | 362 | AG | 3251 | 300–330 | P | 3 |
| 73  | CNMK20227 | F: GCA GAC AACT CAT GCA AAT  
R: TTA TTG TGG GAG GAG ATT | 55.55 | 54 | 278 | TG | 3455 | 320–400 | M | 1 |
Table 2 (continued)

| No. | Marker ID   | Primer sequence                      | Calculated $T_m$ (°C) | Optimum $T_a$ (°C) | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-------------|--------------------------------------|-----------------------|-------------------|---------------------------|-------|--------|--------------------------|---------------|-------------------------|
| 74  | CNMK20739   | F: ACA CGATTTGATGCTATGAGGAAGG     | 54.6                  | 58                | 270                       | AG    | 4154   | 250–330                  | P             | 3                       |
| 75  | CNMK21015   | F: CATGCGATTTGCTATCAGAAGG     | 54.05                 | 57                | 333                       | CT    | 4606   | 130–160, 310–500         | P             | 2                       |
| 76  | CNMK21174   | F: CATGACGCTCCGCTTCATCCGTTG     | 55.65                 | 62                | 392                       | AG    | 4907   | 380–400                  | P             | 4                       |
| 77  | CNMK21493   | F: AGCCATATCTCTCCGCTTTG         | 55                     | 55                | 315                       | TC    | 5710   | 220–230, 320–400         | P             | 2                       |
| 78  | CNMK318     | F: GCCAAAACCTTCTAACATGACG     | 55.6                  | 60                | 358                       | TG    | 1      | 310–330                  | P             | 2                       |
| 79  | CNMK425     | F: GGATGATGTTGGGCTCTTGGG     | 55.9                  | 59                | 217                       | TG    | 2      | 130–160, 180–200         | P             | 3                       |
| 80  | CNMK808     | F: CCTCCTGCAACCATCTCCTG         | 56.1                  | 55                | 225                       | GA    | 5      | 280–320, 400–430         | P             | 3                       |
| 81  | CNMK3765    | F: GAGAGAGCTGCTGGCTCTCAC     | 54.7                  | 55                | 372                       | TC    | 40     | 290–380                  | P             | 3                       |
| 82  | CNMK4127    | F: TAACGCCGATATCACACTCAG       | 54.65                 | 58                | 274                       | TG    | 46     | 160, 350–390            | M             | 1                       |
| 83  | CNMK5054    | F: CTTAGTGTTGATGGTGGCAAG      | 54.1                  | 57                | 391                       | TG    | 64     | 400–440                  | P             | 2                       |
| 84  | CNMK5329    | F: GGCTGGTCGAAATGGCTGAGAC     | 55.05                 | 53                | 320                       | ACAT  | 70     | 300–320, 360–410         | P             | 2                       |
| 85  | CNMK5632    | F: TAGGCTTTAGGCACCTTCTCCTTGCAG | 56.4                  | 56                | 180                       | TG    | 78     | 180–200                  | P             | 5                       |
| 86  | CNMK6746    | F: AGAGCTCTctgTGCGGAAAGAC     | 56.05                 | 59                | 382                       | TG    | 103    | 380–400                  | P             | 2                       |
| 87  | CNMK6908    | F: GATGCGCGAGTGATCTGAGAC     | 53.05                 | 58                | 344                       | CT    | 107    | 200–300                  | P             | 3                       |
| 88  | CNMK7627    | F: GTCGGCTGCTGATCTGAGAC     | 54.45                 | 54                | 356                       | TG    | 124    | 240–350                  | M             | 1                       |
| 89  | CNMK7985    | F: GAAATGGACAACGCGATTTGAT       | 54.4                  | 57                | 323                       | AC    | 135    | 320–340                  | P             | 2                       |
| 90  | CNMK8083    | F: CGGGTATCCCTGCTGATCTC       | 55.15                 | 58                | 247                       | GA    | 137    | 200–250                  | P             | 3                       |
| 91  | CNMK8371    | F: AGGACCTTTGGATGCTGCTG     | 56.2                  | 54                | 310                       | AG    | 148    | 310–400                  | P             | 3                       |
| No. | Marker ID   | Primer sequence                  | Calculated T$_m$ (°C) | Optimum T$_s$ (°C)a | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-------------|----------------------------------|------------------------|---------------------|--------------------------|-------|--------|--------------------------|---------------|--------------------------|
| 92  | CNMK8904    | F: GTTTGCCCCTACGTTCAGC R: TGGGAGCACTACGCCTTTTTCAC | 55.85                  | 56                  | 364                      | GA    | 178    | 320–350                  | P             | 2                        |
| 93  | CNMK9440    | F: TGCGGACGCTGCTTGATGCATT R: TATGCGGGGCAATGTTGGAATCA  | 53.9                   | 54                  | 333                      | GA    | 199    | 300–340                  | P             | 2                        |
| 94  | CNMK9988    | F: CTCGCAATGCAAMATAGGCR R: CGAAATCTGCGTTGGATCTCA | 51.45                  | 49                  | 383                      | GA    | 228    | 230–320                  | P             | 4                        |
| 95  | CNMK10263   | F: TCTGGAACAAATTGGGAATGTTG R: AGATGCTGCCCAAGGTCAAGGAG | 54.5                   | 55                  | 313                      | CT    | 244    | 290–320; 400–500         | P             | 5                        |
| 96  | CNMK10632   | F: TCTGAGGTGGAAGATACGAC R: CAGGCATGACCTCAAAAGTTCG | 54.25                  | 52                  | 376                      | AC    | 268    | 150                      | M             | 1                        |
| 97  | CNMK10681   | F: CGACCTCAATCTGCTCTCTR: TAATCTGCTTTTGCTTGTC | 55.85                  | 57                  | 337                      | TG    | 269    | 320–400                  | P             | 3                        |
| 98  | CNMK11807   | F: AGTGAAGATCTGGCCGAGAA R: TTTGAATCGCCGCTTGTCG | 55.5                   | 55                  | 305                      | CT    | 346    | 242–330                  | P             | 3                        |
| 99  | CNMK12241   | F: AGTGCTCGCAGGGCCAGATGT R: CTCTGGCAGCTGCCTTCAGT | 56.9                   | 55                  | 315                      | AG    | 385    | 240–330                  | P             | 4                        |
| 100 | CNMK12746   | F: TAGATCGGAGGACATGGGAGTA R: AGAATGGGATGTGCGCCCTCA | 54.3                   | 54                  | 383                      | TG    | 441    | 380–450                  | P             | 2                        |
| 101 | CNMK13043   | F: CTCGAGGTGCTGGACAGTAC R: TCAATGCTGGACACCTTCTCA | 55.25                  | 55                  | 383                      | AG    | 472    | 320–340                  | P             | 3                        |
| 102 | CNMK13232   | F: GGAGATCTCAGATGCTGC TGTTCR: GAATATCAGGGATGGGATG | 54.7                   | 58                  | 165                      | TC    | 495    | 170–230                  | P             | 3                        |
| 103 | CNMK13632   | F: CAGGGAGTCAGCAGAGATTAG R: GCTTGACCTTGTTGGTTGAAG | 57.2                   | 57                  | 379                      | GC    | 561    | 300–330                  | P             | 3                        |
| 104 | CNMK13852   | F: TGGAGAGAATGGCGAGATG R: GGGATGGAATGGCGAGATG | 53.55                  | 52                  | 184                      | AG    | 595    | 390–450                  | M             | 2                        |
| 105 | CNMK13946   | F: GCCGACAGAGTAGGCAAGCAG R: CCTGACGAGACCTGTGTATG | 56.9                   | 57                  | 363                      | TC    | 605    | 330–400                  | P             | 3                        |
| 106 | CNMK14272   | F: AAGGGTGTAGTGGATTGAC R: CAAAATCTACCGGTTTGTTGTG | 54.55                  | 53                  | 371                      | AG    | 643    | 320–410                  | P             | 2                        |
| 107 | CNMK14692   | F: GGAGGCTACAGCAGATAGT G R: CACAGATCTGCCTGGGATG | 56.85                  | 53                  | 370                      | GA    | 700    | 330–380                  | P             | 2                        |
| 108 | CNMK15137   | F: TGGTGCGATGATGCTGCT R: CTTGAGCAGGCTTGTCGTTG | 56.1                   | 56                  | 337                      | TC    | 796    | 330                      | M             | 1                        |
| 109 | CNMK15508   | F: ATAGGTGCGAGTGCGCAG R: GGCTACTCATGGCGATG | 56.75                  | 57                  | 374                      | AG    | 868    | 200–220; 320–400         | P             | 3                        |
Table 2 (continued)

| No. | Marker ID | Primer sequence | Calculated $T_m$ (°C) | Optimum $T_a$ (°C)* | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-----------|----------------|-----------------------|---------------------|--------------------------|-------|--------|--------------------------|---------------|-------------------------|
| 110 | CNMK15694 | F: AAG ACT GTT GCC CCT GGA AGA R: TCG ATG TG CAG AGA GCC | 55.15 | 55 | 353 | AG | 908 | 320–390 | P | 3 |
| 111 | CNMK15970 | F: CGT GTT GTT GTT GGT GAT | 55.15 | 53 | 294 | TG | 994 | 220–280 | P | 3 |
| 112 | CNMK5852 | F: ACC CAC ACT GTC GAC CT R: CTT GAG GCA AAG GGC CT | 57.4 | 60 | 310 | AG | 82 | 310–400 | P | 3 |
| 113 | CNMK7532 | F: AGT GGT GCC TCC ACA ATG GAC R: GAT GGG GAT ACA TCA GTC | 54.4 | 55 | 388 | AG | 1570 | 350–500 | P | 2 |
| 114 | CNMK17684 | F: TAG CC GTG CGA ATC TCA AGG R: TCC ATT TCA AGG GAT GAG GAT | 53.3 | 53 | 176 | CT | 1636 | 170 | P | 2 |
| 115 | CNMK18364 | F: TCC AAA GGT GCA TAC GCT R: GAA ACC CAT CTC TGT GGA GAA | 55.45 | 58 | 347 | GA | 2021 | 260–320, 470 | P | 4 |
| 116 | CNMK19333 | F: AC TGG CTT ATT CAT GC ACA AA R: GAT GAC GCA AGA CAA ATA CTC C | 55.15 | 53 | 304 | AC | 2726 | 280–420 | P | 4 |
| 117 | CNMK19929 | F: AAC TGA CGA AGT GGG CTT TG R: AGC CTC CTC TTC TGT TAC | 56 | 56 | 321 | AG | 3153 | 260–330 | P | 3 |
| 118 | CNMK20075 | F: GTT GTG CTC CTA GAT TGC TC | 54.95 | 55 | 280 | GA | 3333 | 300–320 | P | 4 |
| 119 | CNMK1433 | F: GTG TAT TGT ACT CTC GTG GC ACT R: TTC ACC TCT GTG AAT TGC | 55.85 | 50 | 183 | GA | 10 | 150–170 | M | 1 |
| 120 | CNMK1524 | F: GTT CTC TGT TAC AGG GC ACC AATA R: CAT GTG GCA AAG AGG GAT | 55.3 | 55 | 266 | GA | 11 | 230–250 | M | 1 |
| 121 | CNMK1644 | F: TAG GAC GTT TAC CCG AGA AG GAG R: CTG AGG GTT GCTG CAT GGC | 56.8 | 57 | 205 | TC | 12 | 190–210, 270 | P | 4 |
| 122 | CNMK1809 | F: GG GAC CTT GGG GAT CTC TA R: TGC ACT TCG CTA GTT GAC | 56.6 | 57 | 386 | AG | 14 | 350–500 | M | 1 |
| 123 | CNMK2363 | F: CA AGA AC ACG CT CAG GAT GAC R: GAT TCC TCG GT CTA GAC T | 55.85 | 56 | 400 | AG | 22 | 350–380 | M | 1 |
| 124 | CNMK2470 | F: GA AAT CGG CCA CAA GAT TAC R: CCA AAT TCC CAA AAT ATG | 54.6 | 55 | 369 | CT | 23 | 250–300 | M | 1 |
| 125 | CNMK2651 | F: CT TCC TCC CT TCC GCA AA R: CTC ACT CTC CAA CAG GT A | 56.05 | 54 | 306 | GA | 25 | 280–320 | P | 3 |
| 126 | CNMK2960 | F: GAG GAG GAC GAC GAG GAG GAG R: GAT CCA AAG ATG TCT GAC TCA TGA | 55.05 | 55 | 300 | CT | 30 | 260–300 | P | 3 |
| 127 | CNMK3730 | F: GAG TAA GGG CCA CAG CTC TCT R: AAC GAC AT CTC ACC AAG | 54.95 | 53 | 296 | AAG | 40 | 250–280 | P | 3 |
| No. | Marker ID | Primer sequence | Calculated Tm (°C) | Optimum Ta (°C)* | Expected allele size (bp) | Motif | Contig | Observed size range (bp) | Polymorphisms | No. of alleles per locus |
|-----|-----------|-----------------|-------------------|-----------------|--------------------------|-------|--------|--------------------------|---------------|------------------------|
| 128 | CNMK3865  | F: AGGAGTAGCTCCGGCCCCTCTA  
R: CCC TCAGATTAAAGCAGAGAAG | 57.05 | 57 | 323 | AG | 43 | 300–320 | P | 2 |
| 129 | CNMK4080  | F: AGTTTCGTACGGGCGTGTGCTG  
R: AGACACTACAGCAAGCAAGG | 57.9 | 53 | 327 | AG | 45 | 300–320 | P | 4 |
| 130 | CNMK4336  | F: CGGACGTGTACACGGCTCATA  
R: AAACCTTTCGGAGAACATCAC | 55.35 | 55 | 229 | AC | 49 | 230–300 | P | 4 |
| 131 | CNMK5287  | F: CCAACACAGACCAACTCAAAT  
R: TGTGGAAGATGTGGATGGA | 55.4 | 53 | 205 | CT | 69 | 230–340 | P | 4 |

M monomorphic, P polymorphic, N/A not applicable

* Based on gradient PCR optimization
become very efficient in generating high number of markers using advanced bioinformatics tools in this study has 10, 15, 16, 30]. It has been extensively used in these analyses since SSR markers are abundant and well distributed for studying genetic diversity and genotyping of coconut [8, QTL mapping, and other relevant studies in coconut. The generated SSRs here are expected to contribute to the pool of available molecular markers [10, 16, 28–30] for fingerprinting, genetic diversity analysis and QTL mapping, and other marker-assisted plant breeding applications.

The predominance of dinucleotide repeats in coconut and other related species is supported by previous works of Rivera et al. [21], Palliyarakkal et al. [13], Xia et al. [29], and Lantican et al. [9]. This result coincides with studies of Palliyarakkal et al. [13] and Xia et al. [29] on which AG/GA/TC/CT motifs were also the most common dinucleotide repeats found in coconut/palm genome. 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 [14, 21, 25]. The development of SSRs 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 [10, 16, 28–30] for fingerprinting, genetic diversity analysis and QTL mapping, and other relevant studies in coconut.

Microsatellites or SSRs are a very useful molecular tool for studying genetic diversity and genotyping of coconut [8, 10, 15, 16, 30]. 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 [11, 20]. Previous studies like Rivera et al. [21], Perera et al. [15], Xiao et al. [30], and Wu et al. [28] have already developed SSRs in coconut for genetic diversity studies, and these markers showed high levels of polymorphism as well.

**Conclusion**

Here, we demonstrated that a locally established bioinformatics pipeline can mine SSRs 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. SSRs and SNPs can be easily generated automatically in the genome sequences with the use of these programs or pipelines.

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.

**Authors’ contributions**

RESC and JC conducted the wet lab experiments and drafted original manuscript. RESC and DVL conducted bioinformatics analyses. RRG and ANCM supervised wet lab experiments and confirmed validation procedures. RLR provided the leaf samples for analysis. AOCS, HFG, and CER secured funding for the project. DVL, ANCM, and RRG conceptualized the hypothesis and methodology of the study. All authors have read and approved the final manuscript for publication.

**Funding**

The Department of Science and Technology — Philippine Council for Agriculture, Aquatic and Natural Resources Research and Development of the Department of Science and Technology (DOST-PCAARRD)

**Availability of data and materials**

The dataset(s) supporting the conclusions of this article is (are) included within the article (and its additional file(s)).

**Declarations**

**Ethics approval and consent to participate**

Not applicable.
Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Institute of Crop Science, College of Agriculture and Food Science, University of the Philippines Los Baños, College, 4031 Los Baños, Laguna, Philippines. 2Philippine Genome Center — Program For Agriculture, Livestock, Fisheries and Forestry, University of the Philippines Los Baños, College, Los Baños 4031, Laguna, Philippines. 3Philippine Coconut Authority — Zamboanga Research Center, San Ramon, 7000 Zamboanga City, Philippines. 4Institute of Crop Science, College of Agriculture and Food Science, University of the Philippines Los Baños, College, 4031 Los Baños, Laguna, Philippines.

Received: 26 January 2022  Accepted: 2 May 2022
Published online: 16 May 2022

References
1. Anderson JR, Lubberstedt T (2003) Functional markers in plants. Trend Plant Sci 8:554–560
2. 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
3. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15
4. FAO STAT Database. 2013. http://faostat.fao.org/
5. Foale M (2003) The coconut odyssey: the bounteous possibilities of the Cocos nucifera L. Bot Rev 43(3):265–320
6. Harries HC (1978) Evolution, dissemination and classification of Cocos nucifera L. Euphytica 129:169–181
7. Konan KJ, 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. Afr J Biotechnol 6(4):341–347
8. Lantican D, Stricker 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) 9(8):2377–2393. https://doi.org/10.1534/g3.119.400215 PMID: 31167834; PMCID: PMC6686914
9. Lebrun P, Nicho Y, Seguin M, Gritet 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
10. Mason AS (2015) SSR genotyping. In: Batley J (ed) Plant genotyping. Springer, New York, pp 77–89
11. Meerow AW, Krueger RR, Singh R, Low ETL, Maizuraithnin M, Ooi LCL (2012) Coconut, date, and oil palm genomics. In: Schnell RJ, Priyadarshan PM (eds) Genomics of tree crops. © Springer Science Business Media, LLC, pp 299–351. https://doi.org/10.1007/978-1-4614-0920-5_10
12. Paliyarakkal MK, Ramaswamy M, Vadivel A (2011) Microsatellites in palm (Areaceae) sequences. Bioinformation. 7(7):347–351
13. 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 8:344–346
14. Perera L, Russell JR, Provan J, Russell W (2003) Studying genetic relationships among coconut varieties/populations using microsatellite markers. Euphytica 132:121–128
15. Perera L, Russell JR, 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
16. Perera LP, Hoehn Y, Verdel J, 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
17. Perera PIP, Hocher V, Verdeil JL, Yakandawala DMD, Weerakoon LK (2007) Molecular markers: it’s application in crop improvement. J Crop Sci Biotechnol 12(4):169–181
18. Philippine Statistics Authority. 2017. Agricultural foreign trade statistics of the Philippines. 2015. https://psa.gov.ph
19. Philippine Statistics Authority. 2018. Selected Statistics on Agriculture 2013-2017. https://psa.gov.ph
20. Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1(7):215–222
21. Rivera R, Edwards 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
22. Rotmistrovsky K, Jang W, Schuler GD (2004) A web server for performing electronic PCR. Nucleic acdes research, 32 (web server issue), W108–W112. https://doi.org/10.1093/nar/gkh450
23. Sharma A, Namdeo AG, Mahadik KR (2008) Molecular markers: new prospects in plant genome analysis. Pharmacogn Rev 2(3):23–34
24. Siddhurumale P, Ambili SN (2011) Marker assisted breeding in coconut (Cocos nucifera L.). Gregor Mendel Foundation Proceedings 2011:30–32
25. Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, Arnold GM, Karp A, Baudouin L, Rogonin F (2000) 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 100:764–771. https://doi.org/10.1007/s001220051350
26. Vavilov NI (1926) Centres of origin of cultivated plants. Bull Appl Bot Genet Plant Breed 16:1–248
27. Wang X, Wang L (2016) GAMA: an integrated software package for genome-scale SSR mining, marker development and viewing. Front Plant Sci 7:1350
28. Wu Y, Yaoqiong Y, Qadri A, 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. Trop Plant Biol 12(1):32–43
29. 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. Mol Breeding 34(2):1–11
30. Xiao Y, Luo Y, Yang Y, Fan H, Xia W, As M, 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 Ottomics J 6(3):193–200
31. Xu Y, Cloudh JH (2008) Marker-assisted selection in plant breeding: from publications to practice. Crop Sci 48:391–407. https://doi.org/10.2135/cropsci200704.0191

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.