


RESEARCH

Open Access



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

Reina Esther S. Caro<sup>1,2</sup>, Jesmar Cagayan<sup>1,2</sup>, Roanne R. Gardoce<sup>1</sup>, Anand Noel C. Manohar<sup>1\*</sup> , Alma O. Canama-Salinas<sup>1</sup>, Ramon L. Rivera<sup>3</sup>, Darlon V. Lantican<sup>1</sup>, Hayde F. Galvez<sup>1,2</sup> and Consorcia E. Reaño<sup>4</sup>

## 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 bioinformatics 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  $\Delta G$  value for hairpin loop,  $-0.3$  kcal/mol; minimum  $\Delta G$  value for self-dimer,  $-0.9$  kcal/mol; and minimum  $\Delta G$  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  $\times$  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

\*Correspondence: acmanohar@up.edu.ph

<sup>1</sup> Institute of Plant Breeding (IPB), College of Agriculture and Food Science, University of the Philippines Los Baños, College, 4031 Los Baños, Laguna, Philippines

Full list of author information is available at the end of the article

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 ( $2n = 2 \times = 32$ ). 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<sup>-1</sup> ethidium bromide staining, and UV illumination at 300 nm using the Enduro GDS Touch

**Table 1** Coconut genotypes used in the study for screening the SSR markers

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 ( $\lambda$ ) DNA molecular weight standards (Sigma-Aldrich Inc., St. Louis, Missouri, USA).

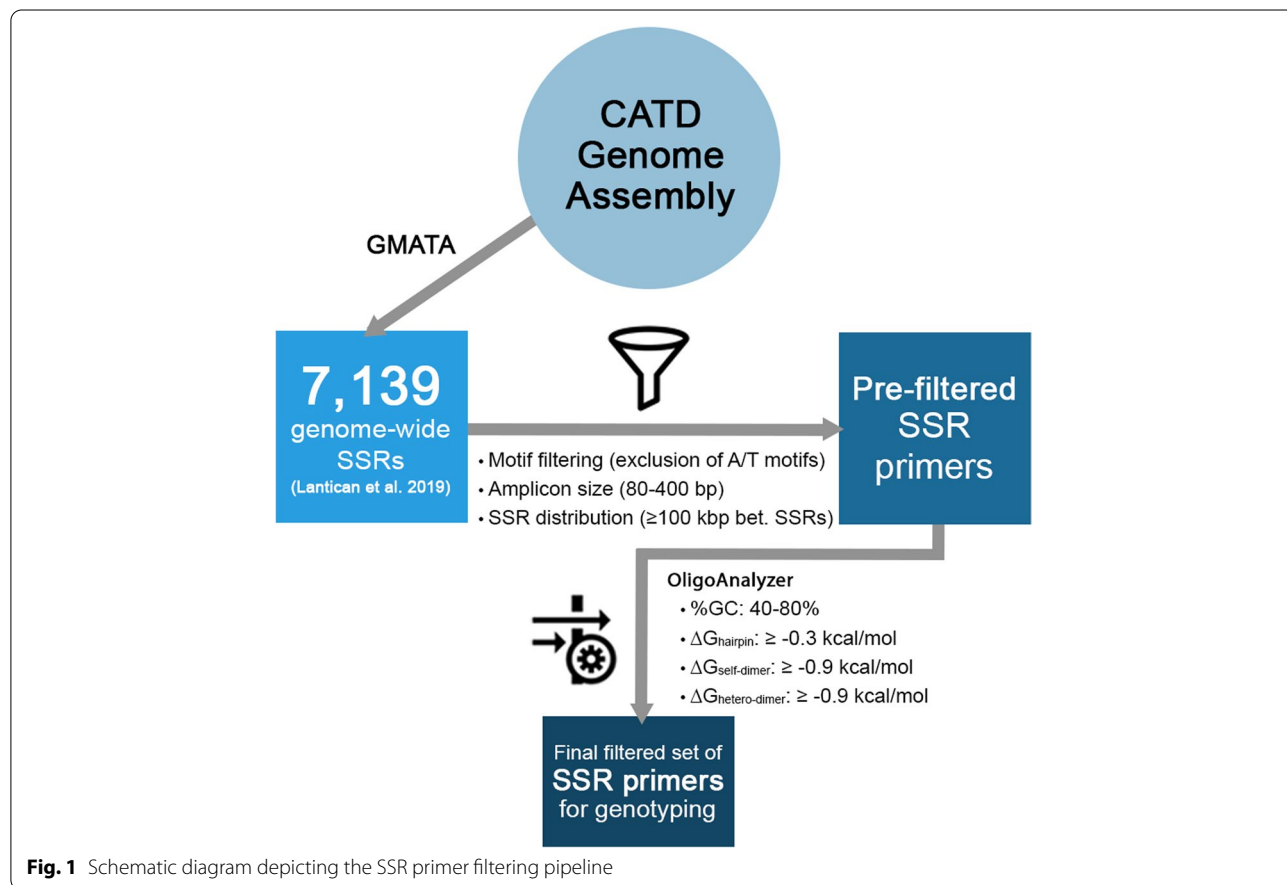
**Development of SSR markers using the genome assembly of coconut ‘Catigan Green Dwarf’ (CATD)**

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

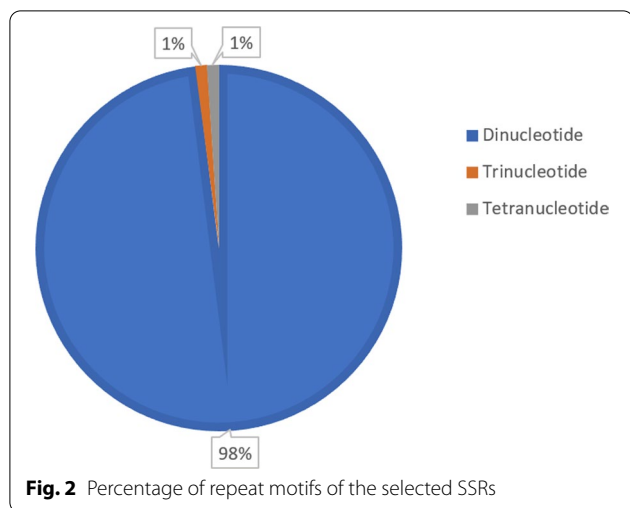
desired parameters: %GC, 40–60%; minimum  $\Delta G$  value for hairpin loop,  $-0.3$  kcal/mol; minimum  $\Delta G$  value for self-dimer,  $-0.9$  kcal/mol; and minimum  $\Delta G$  value for heterodimer,  $-0.9$  kcal/mol for further filtering of the SSRs (Fig. 1).

**PCR analysis**

PCR was carried out with 10  $\mu$ L reaction volume (15 ng genomic DNA, 1 $\times$  PCR buffer (10 mM Tris pH 9.1 at 20  $^{\circ}$ C, 50 mM KCl, 0.01% Triton<sup>TM</sup> X-100); Vivantis Technologies, Malaysia), 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs (Promega Corporation, Madison, Wisconsin, USA), 0.2  $\mu$ M 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  $^{\circ}$ C for 3 min, 30 cycles of denaturation (95  $^{\circ}$ C, 30 s), annealing (45–60  $^{\circ}$ C depending on the primer pair, 45 s), extension (72  $^{\circ}$ C, 1 min), and final extension at 72  $^{\circ}$ C for 5 min. Amplifications were carried out in the Applied Biosystems Veriti<sup>TM</sup> 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



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 ug mL<sup>-1</sup> 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,

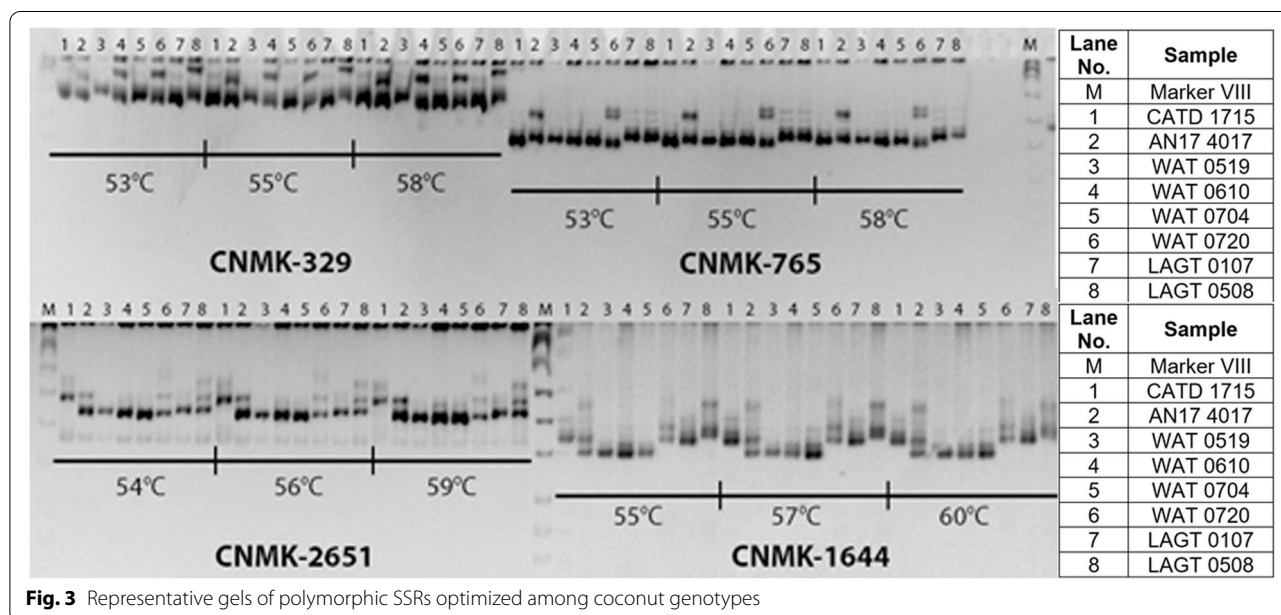


Fig. 3 Representative gels of polymorphic SSRs optimized among coconut genotypes

**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) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
1	CNMK7	F: GAGGGAGAGGCTGTGATG R: GCATGTTCAAGGCTTTGGTT	55.6	56	299	AG	0	242–260; 404–480	P	2
2	CNMK172	F: TTTCACATCTTGATCCACT R: TTGATGGCTGGTGAATGC	53.55	57	280	TG	1	280–320	P	3
3	CNMK329	F: AGTGGCTCAAGTGGTTCAG R: CACCATCTTGGCTTTCAAT	55.3	55	342	CT	1	320–360	P	2
4	CNMK653	F: AGCCAAATTCGATCAACTG R: ATACATGGCAATGGCACCTT	54.15	57	240	GA	4	130–170; 170–230	P	3
5	CNMK765	F: AATGTTATGGAAACCTTGC R: TGACAAGGTGGACACTTGA	55	53	211	GA	5	210–265	P	4
6	CNMK995	F: AAGACACGACACCCGATACC R: AAGTGCAACAGCAACTGG	56.55	54	196	AG	6	190–242	P	2
7	CNMK1095	F: CCTCATGGCTAACCTCAAG R: AGTCTGAACCGAGGTACGA	56.25	54	229	CT	7	230–250	P	4
8	CNMK3414	F: CCAGCTCACTGCCATACAGA R: AGCACCGGCTCTCATCTTC	56.85	50	289	AC	35	290–380	P	2
9	CNMK3683	F: ACCTCCAGTGAAGACATGC R: CCGTTACGTTATCCCTCCA	55.65	54	360	TG	40	320–370	P	2
10	CNMK4036	F: TGAGTTTCTGGACCAAAGC R: ATACGCTGCTAGGGTGGTGT	56.8	55	254	GA	45	290–320	P	3
11	CNMK4153	F: CATTGGATGTCATAGTGC R: AGGTATGCCCAAGGAAGGT	55.2	53	388	TC	46	380–400	P	2
12	CNMK4627	F: TCGGTGACC AACTACTCGAT R: GGAGCATGATGGAGAGAAA	55.25	51	237	AG	56	220–245	P	4
13	CNMK4772	F: TGCATCAAGACAGACCATCC R: TGCATGAATACACCCACATT	53.5	59	351	GT	59	380–450	P	3
14	CNMK4830	F: AAATGGCTTTGCATGCTCC R: GGTGTGCAAGCACCTCCAT	54.75	53	341	AG	60	320–400	P	2
15	CNMK4976	F: TGCCACCATATGAGCAGAGA R: CCCAGAGCTCCAAACTCAIC	55.65	58	199	AC	64	170–200; 215–250	P	6
16	CNMK5103	F: AAGGTGCCACTCTCCCAAG R: TTACCCGGCTTATCTCCT	56.15	54	307	CT	65	300–350	P	3
17	CNMK5211	F: AAGCTGACAAATGTGAAGAGC R: CAACAAGGCTTAGCACTGG	54.75	51	388	CT	68	190–210; 242–310	P	6
18	CNMK5746	F: CTGGCAATGACACAACAAG R: AACCGACACAGAGAGGATGG	55.45	50	152	CT	80	380–400	P	4
19	CNMK5910	F: CGATTGAAGCTGATGCTCTG R: TTGATGTGCGTGAATGGAAT	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) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
20	CNMK6063	F: CCACAAATGAAAGCAAGGA R: TGGACTAGGAGTGGGATTGG	54.7	53	394	AC	86	390–450	P	2
21	CNMK6206	F: TGCTAAAGGACCGAAGATG R: TCATGGAGAGGTGCATGTGT	55.4	53	321	GT	89	300–340	M	2
22	CNMK6376	F: CATCTCTTAAGCTGGCAACC R: TAGGATAGGTGAGCGCAACA	55.6	53	368	CT	92	380–450	P	3
23	CNMK6463	F: GATACAGGTGCCAGGTGAG R: TCCAGCACTCTGTGCGTTAT	57	57	304	GA	96	230–240; 300–320	P	4
24	CNMK6507	F: CCCAATACCTTTGGTTGCTC R: TGCATGTCCTGCATAAG	53.4	51	336	GA	98	150–160; 180–200	P	5
25	CNMK6571	F: TGAATGCCGTTGTTGTAGG R: GGATAAGGCTGCATGGCTAA	54.15	62	398	CT	99	400–450	M	1
26	CNMK6672	F: ATATAGGCAAGGCCAAGGT R: TGGTCATGGTGGAGTAGCAA	56.05	56	363	GA	102	330–370	P	4
27	CNMK7007	F: GGTGATTTGTCGTCGTT R: TGCAGAATGCTTGGCAATAG	54.3	57	381	TG	109	350–380	P	2
28	CNMK7162	F: AGGCACACCTGAAAGGTACA R: GTTCCGAAAGGATTTGCACCTC	55.65	56	259	GT	113	250–320; 400–420	P	3
29	CNMK7449	F: TGCATCACACAGCTACACCA R: GTGCAGTTTGTGGCTTCA	56.3	54	246	TC	119	240–260	P	4
30	CNMK7553	F: TTGAGGTGTCCAACATGG R: GCTTGTAGGGCACGTTCAIT	54.5	50	267	AG	121	240–260	P	3
31	CNMK7710	F: TCAACTGCATCAGGCTTGG R: CTGACAGGTGGCACAGAGAA	56	54	301	GA	127	220–230	P	3
32	CNMK7859	F: ATGGTCTGGATGGCTTGC R: TCTCCGTTCACTCTGCAACA	55.65	54	355	AG	132	300–310	P	2
33	CNMK8015	F: GGCAGTTCCACTTACCCAAG R: TGCTTAAACAAGCGTTCGTG	54.9	55	391	TG	135	250–320; 400–480	P	2
34	CNMK8444	F: GGTCCGGATTACAGTTAGT R: AGCCAAAAGAACCCCTTGGAAAT	55.5	54	295	GT	152	200; 250–320	M	1
35	CNMK8741	F: ACCCGAGGTTTGAAGGAAC R: TTGGCAGCTTCACTTATTGC	54.35	57	248	TC	166	320	M	1
36	CNMK9091	F: GCTTGAATCCCTGGATACGA R: CTTGCTCTGTTCCCCATGAC	55.1	58	373	AG	185	320–380	P	2
37	CNMK9331	F: ATGCTTCGCTTGTTGTAGG R: AGTGAGGAATCCGATGCAAT	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) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
38	CNMK9514	F: TGAGATGAGATGGTGACA R: ATCAATGGAGGTCACAAGG	54.95	50	324	TG	200	150; 220–240	P	3
39	CNMK9655	F: TTGGCTAGTCTGCCATAGG R: CCAATCAACACCCACATTGA	54.55	50	381	AG	207	242–280; 370–400	P	2
40	CNMK9918	F: TTGGACTCCCAAGCACTAGG R: TTCCTTCCAAGCAGATGTCC	55.5	56	254	GT	224	230–242	P	2
41	CNMK10005	F: TGGTGCAGTCTTCTCAATCG R: CCATCTCTCCCTGATTCCA	54.35	54	251	CT	229	242–260; 320–350	P	2
42	CNMK10146	F: AATCGAATACGTGGCGAAC R: GCTTGTAGCAACTCCAAGTTC	54.9	58	208	GA	237	210–225; 240–250	P	3
43	CNMK10298	F: TCACTCCATACCCAAAGAA R: TTTAGTCCCAAGTGCCCATC	54.2	54	237	AG	245	230–245	P	2
44	CNMK10608	F: AGGAACCTATCGGTGTTGG R: GCATGATTTGTCATTGGAG	53.85	53	389	GA	265	350–380	P	2
45	CNMK10723	F: GTTTGCAGTGGAAAGTCGT R: AGCTTCTTGATGCCATAGCC	55.55	56	324	TC	274	280–350	P	3
46	CNMK10821	F: AATAGCCACGTTACCCITG R: TGCAGTGTGGAGACACCAT	56	59	326	AG	280	310–350	P	3
47	CNMK11095	F: GATCGGCAC TAGGGAACITG R: AATGCGAGGACAAC TGGAAC	55.4	55	361	AG	295	330–360	P	2
48	CNMK11349	F: ATGGCATTGAGGATGAGAA R: ACCGTTCTTTGGAAATGTG	52.55	62	280	AC	308	200; 300–350	P	3
49	CNMK16404	F: GGAATCTGAAGCAGGGACAG R: AGGGCATTGAAACACAGCAC	55.6	54	334	AG	1122	300–400	P	3
50	CNMK16553	F: GCCGAATGTCTTAATAGGTG R: AGGAATGCCATGTCAGGTTTC	53.45	48	339	GA	1168	200; 320–340	P	3
51	CNMK16634	F: CGAGCTTGAAGTCAGCTTTG R: GCAGCCTTACCTCTCAGTTC	55.95	54	357	AG	1203	400–500	P	2
52	CNMK17050	F: ATTGGCTGAGTGGAGGACAA R: ATGAGCAACCCATGTTGATG	54.85	52	397	CT	1370	400–470	P	2
53	CNMK17101	F: CCAGCCATGCTTACCAACTTA R: CGAGAACCACGTCATGAGA	54.9	50	299	AG	1401	250–290	P	3
54	CNMK17156	F: GTGTTCTGGCAATCATGCTC R: TGCTTGACATACGCACAAA	54.7	53	293	GA	1418	240–265	P	3
55	CNMK17229	F: CAAGCTGGAGAACACACAGG R: TGTATCAGCAACTGGAGA	55.75	56	400	AC	1462	280–320	P	2

**Table 2** (continued)

No.	Marker ID	Primer sequence	Calculated $T_m$ (°C)	Optimum $T_a$ (°C) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
56	CNMK17487	F: TGCCAATGTTAAGGATGC R: TCCATACGAAGCAACTGTG	54.25	52	288	CT	1550	310–350	P	3
57	CNMK17639	F: AAATCTGGGTGGCTCTCTCA R: GCCAGCAGCTATGGAAGAAG	55.85	54	303	TC	1621	280–340	P	4
58	CNMK17725	F: GCTGAGTGTCTGTGTCCA R: CCTGAACCAAGGGATGAGAA	55.1	55	234	AG	1655	170–190; 220–250	P	6
59	CNMK17797	F: GGTGCCTTAGTGCCTTCTCTT R: AATCCGTTGGACGTTATTC	55.25	53	370	GA	1688	315–350	M	2
60	CNMK17875	F: GGCCTGGGTTCAAATCTT R: TTGAAGGCACCTAAGGCACCT	55.1	50	397	AG	1750	300–400	P	2
61	CNMK18331	F: TGCAAAGTAAGGACCCCTGGA R: CTGATTCATGGTTGGCTCCT	55.45	55	368	AG	1987	320–500	M	1
62	CNMK18501	F: CTGGACGAAACAATGGTCTCT R: CTCACAGAGGTATCCATCCA	55.25	55	380	TC	2084	350–400	P	2
63	CNMK18573	F: TGTGTTCACTCGGTCTGTTA R: AGGCCCTTCTCGATCACTAGA	55.45	57	383	TG	2130	400–440	P	2
64	CNMK18799	F: GCCTCGATTATTCACCTGGA R: GTCAGGAGGCAGTGGAGAAG	56.25	56	255	AG	2306	240–320	P	2
65	CNMK18903	F: CAGGACTCGGAGATAGCTG R: TTGGCTGCTAATGCTGCAC	56.2	56	396	AG	2370	350–400	P	2
66	CNMK18972	F: GGTGGTTGGCATCCATGTT R: TACGTGGGACACCAAGAGT	56.6	59	358	TC	2409	230–260; 330–400	P	5
67	CNMK19118	F: TACCCATCCCAAAATGGAC R: TCAGGGTGGCATGATGAATA	54.1	52	393	CT	2535	230–320	P	4
68	CNMK19193	F: ATGTTGTGGGACGATGAGT R: ACCTCGCATGATGAAACTG	55.75	54	398	TC	2585	360–400; 690	P	3
69	CNMK19386	F: AAGGGTTTGAATGGTGGTG R: CCTAACCCAGGCAAAGGACAA	55.25	52	354	AG	2748	300–330; 400–420	P	3
70	CNMK19611	F: ATCCATCCAATGCTATCAGG R: GACCGCATTAGCTCTGGTACT	54.2	50	364	AC	2872	300–340	P	3
71	CNMK19799	F: CGTCTGGGATAGCCTTCAGT R: CCAAGCAACGGAGAACTTG	55.2	55	337	TC	2988	320–410	P	3
72	CNMK20018	F: TGACAAGTTTCAGGGCATCA R: TGCAGATCTGCCAAAGCTA	54.5	53	362	AG	3251	300–330	P	3
73	CNMK20227	F: GCAGCACACTCATGAAAAT R: TTAGTGGCAGAGAGGTTGG	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) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
74	CNMK20739	F: ACACGATTGATGAGGA R: GCCATGAGCCACATCTATT	54.6	58	270	AG	4154	250–330	P	3
75	CNMK21015	F: CATGCCATTGTCAATCCA R: TGCAGAGGAGTCCAAGTATG	54.05	57	333	CT	4606	130–160; 310–500	P	2
76	CNMK21174	F: CATGACTGACCGCTACTCC R: TCCTAATCCTCCATGTGCTG	55.65	62	392	AG	4907	380–400	P	4
77	CNMK21493	F: AGGCATAGTCTCTCGGCTTG R: CAAAGTGAATTGCTCGTGAA	55	55	315	TC	5710	220–230; 320–400	P	2
78	CNMK318	F: GGCAACCTTCTAAATGACC R: CTTTGTCCAGCCGTACCTGT	55.6	60	358	TG	1	310–330	P	2
79	CNMK425	F: GGATGTAGGTGGGCTCTTG R: GCCACTAGAGGTCATTGGA	55.9	59	217	TG	2	130–160; 180–200	P	3
80	CNMK808	F: CCATGCCACACCTTACTCCT R: AACAAACGCCACCTACTCTG	56.1	55	225	GA	5	280–320; 400–430	P	3
81	CNMK3765	F: GAGAGAGTGGTCGGCTTCAC R: ATTCGGATGTCGATTTGGA	54.7	55	372	TC	40	290–380	P	3
82	CNMK4127	F: TCAAGGCATCAATACCCAAG R: GAAGTCCAAGCAACCCAGCTC	54.65	58	274	TG	46	160; 350–390	M	1
83	CNMK5054	F: CTTTCATGTTTCATGGTGCAG R: ATGGATATGAACAGCGAGCA	54.1	57	391	TG	64	400–440	P	2
84	CNMK5329	F: GCTGGTCGGAAATGCTAGAC R: CATGCAAAAGCCTCACTCAAG	55.05	53	320	ACAT	70	300–320; 360–410	P	2
85	CNMK5632	F: TAGCCCTTTCAGGACCTCT R: GAGGAAGTCATCCGAAAGCAG	56.4	56	180	TG	78	180–200	P	5
86	CNMK6746	F: AAGCACTTCTTACGCCAGA R: GCCTTGGTGGTGAAGATTGT	56.05	59	382	TG	103	380–400	P	2
87	CNMK6908	F: AGATTGCCGGAGTTGATTG R: CCTATTCGGTCGCAATTGAT	53.05	58	344	CT	107	200–300	P	3
88	CNMK7627	F: TTAGTCTGGAATCGTTAGAGG R: CATGGTGGCAACTGTGATA	54.45	54	356	TG	124	240–350	M	1
89	CNMK7985	F: GAAATGAGACCGCCATTGAT R: CGGACCGTTAGACAGATTGC	54.4	57	323	AC	135	320–340	P	2
90	CNMK8083	F: GGGGTTTCGGTAGCATCTC R: CTCCAGCACAGATGGAACA	55.15	58	247	GA	137	200–250	P	3
91	CNMK8371	F: AAGGACTTGTGATGCCTTGG R: GTCACCATAGCCGACAACT	56.2	54	310	AG	148	310–400	P	3

**Table 2** (continued)

No.	Marker ID	Primer sequence	Calculated $T_m$ (°C)	Optimum $T_a$ (°C) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
92	CNMK8904	F: GTTTGCCCGTACTTCTCAGC R: TGGCAGCATCCTCTTTAC	55.85	56	364	GA	178	320–350	P	2
93	CNMK9440	F: TGGGACCTGTCTGCATATT R: TATCGGCACATTCGATTTCA	53.9	54	333	GA	199	300–340	P	2
94	CNMK9988	F: CTCGCAAAATGCAATATAGGC R: CGCAAATTCGGTTGATCTTA	51.45	49	383	GA	228	230–320	P	4
95	CNMK10263	F: TTCAGGACAAATGGAAGTGTGG R: AGAATGCCCAAGTCAAGCAG	54.5	55	313	CT	244	290–320; 400–500	P	5
96	CNMK10632	F: TTCACGTTGCCAAATGACAC R: CAGGCATGCCTCAAGATG	54.25	52	376	AC	268	150	M	1
97	CNMK10681	F: CGACCTCCATACATGGCTCT R: TAACTGGCTTTGGTTGGTC	55.85	57	337	TG	269	320–400	P	3
98	CNMK11807	F: AGTGAAGATCTGCCCGAGAA R: TTTGAACCTCACGCTTGTTC	55	55	305	CT	346	242–330	P	3
99	CNMK12241	F: AGTGCTAGCCAGACCCATGT R: CTCCTCAAAGTACTGTCAAT	56.9	55	315	AG	385	240–330	P	4
100	CNMK12746	F: TAGATCGAGGCATGCGAGTA R: AGAATGGATATTGGCCCTCA	54.3	54	383	TG	441	380–450	P	2
101	CNMK13043	F: CCTTCAGGGTTAAGTGCAGA R: TCAAACCTGGCTGATCCTTCA	55.25	55	383	AG	472	320–340	P	3
102	CNMK13232	F: GGAAGTCTCAGTCGTGCTC R: GGAATCAAGAAAGGCATTGG	54.7	58	165	TC	495	170–230	P	3
103	CNMK13632	F: CGAGGTCTCAGCGAATTAG R: GCTGGACCTTTGGTGAAG	57.2	57	379	TG	561	300–330	P	3
104	CNMK13852	F: TTGGAAGAAATGGCAGTGGT R: GGATATGGATGGATGGATGG	53.55	52	184	AG	595	390–450	M	2
105	CNMK13946	F: GGCAGAGGTAGTGAACGAG R: CCTGACGGAGGACTGTTGAT	56.9	57	363	TC	605	330–400	P	3
106	CNMK14272	F: AAGGGTGCATGATGGTTAGG R: CAAACATCTCCGTGTG	54.55	53	371	AG	643	320–410	P	2
107	CNMK14692	F: GGAGGCTACCAGCCATAGTG R: CACAGTCTCTCGCATGAGA	56.85	53	370	GA	700	330–380	P	2
108	CNMK15137	F: TTGGTCCGATGATGCTTTC R: CTGAGCACCCCTGGTAGGT	56.1	56	337	TC	796	330	M	1
109	CNMK15508	F: ATAGCTGGTGGAGTCGGCAAG R: GGCCTACTGATGGACTGGA	56.75	57	374	AG	868	200–220; 320–400	P	3

**Table 2** (continued)

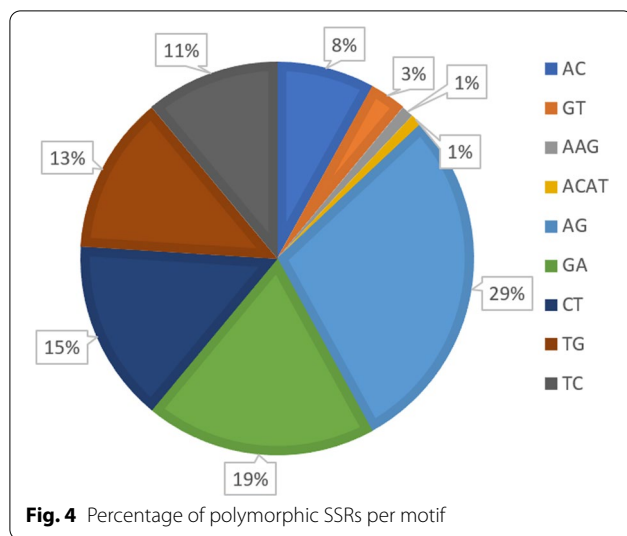
No.	Marker ID	Primer sequence	Calculated $T_m$ (°C)	Optimum $T_a$ (°C) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
110	CNMK15694	F: AAGACTGTTGCCCTGGAAGA R: TCGATGATGCAGAGATCAGG	55.15	55	353	AG	908	320–390	P	3
111	CNMK15970	F: CGTGTGGTATTGTTGCAT R: ATTGCGGGTAAGGAGAAGT	55.15	53	294	TG	994	220–280	P	3
112	CNMK5852	F: ACCCACTAGCACTTGCACCT R: CCTGAGTCAACAAGCCATC	57.4	60	310	AG	82	310–400	P	3
113	CNMK17532	F: AAGTTGGCTCACCATCAC R: GATGGGATACATCCAATGC	54.4	55	388	AG	1570	350–500	P	2
114	CNMK17684	F: TAGCCGTCGGATATTC AAGG R: TGCATTCTAAGGGAATGGATG	53.3	53	176	CT	1636	170	P	2
115	CNMK18364	F: TCCCAATGGCAGTCCTAGTC R: GAAAACCCATCCTTGTGGAGA	55.45	58	347	GA	2021	260–320; 470	P	4
116	CNMK19333	F: ACCTGCCTATTATGCACAA R: GAGTGACGCCAAGACAATCTCC	55.15	53	304	AC	2726	280–420	P	4
117	CNMK19929	F: AACTGAGCAGATGGGCTTTG R: AGCCTCTGTGACGAACGAAT	56	56	321	AG	3153	260–330	P	3
118	CNMK20075	F: GTTGTGCCTCCAATGTTCTC R: GATGGCCTGAATCCTGTAA	54.95	55	280	GA	3333	300–320	P	4
119	CNMK1433	F: GGTGATTGACTCTGGCACT R: TTCACCTCTGGATCTTGGTG	55.85	50	183	GA	10	150–170	M	1
120	CNMK1524	F: GCTCCTGGTACAGGCACATA R: ACATCGGGATGGTTCAAT	55.3	55	266	GA	11	230–250	M	1
121	CNMK1644	F: TAGGACGTTTACC CGCAGGAG R: CTGTAGGGTAGGGTGCATCG	56.8	57	205	TC	12	190–210; 270	P	4
122	CNMK1809	F: CGGGACTTGGGAGTCACTA R: TGGCACTTCGCTGTGTAGC	56.6	57	386	AG	14	350–500	M	1
123	CNMK2363	F: CAAGACACAGCTTCGAGATCC R: GATTCCTCCGCTATGACCT	55.85	56	400	AG	22	350–380	M	1
124	CNMK2470	F: GCAATCGGCCCAAGAACTAC R: CCCAACCTTCCACCAATATG	54.6	55	369	CT	23	250–300	M	1
125	CNMK2651	F: CCTCCCTTACCCTTGCATAA R: CTGCAGTCTCACCGTAIGT	56.05	54	306	GA	25	280–320	P	3
126	CNMK2960	F: GAGGAGTGGAGCGGATTGGA R: CGATCCAAGATTGGTACTGGA	55.05	55	300	CT	30	260–300	P	3
127	CNMK3730	F: CGATTGAAGCCAGTCTCTC R: AACGACATCTTCAACGCAA	54.95	53	296	AAG	40	250–280	P	3

**Table 2** (continued)

No.	Marker ID	Primer sequence	Calculated $T_m$ (°C)	Optimum $T_a$ (°C) <sup>a</sup>	Expected allele size (bp)	Motif	Contig	Observed size range (bp)	Polymorphisms	No. of alleles per locus
128	CNMK3865	F: AGGAGTAGCTCCGCCCTCTA R: CCTCGAATGACCCAGAGAAG	57.05	57	323	AG	43	300–320	P	2
129	CNMK4080	F: AGTTTCGTAGCGGCTGATGT R: AGAGCACTCAGCAAGCAGGT	57.9	53	327	AG	45	300–320	P	4
130	CNMK4336	F: CCGACGTGTGACAGCTCTA R: AAAACCTTTCGCACGAATCAC	55.35	55	229	AC	49	230–300	P	4
131	CNMK5287	F: CCCAACAGACCCAACTCAAT R: TGTGAAGATGTGGAGTGGA	55.4	53	205	CT	69	230–340	P	4
			N/A	N/A	N/A	N/A	N/A	N/A	N/A	2.7

M monomorphic, P polymorphic, N/A not applicable

<sup>a</sup> Based on gradient PCR optimization



29] on which the high repeat content may hinder specificity of the markers and/or may result to nonspecific 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 [22]. Allele size range of the markers was also limited to 80–400 bp 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.

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**

<sup>1</sup>Institute of Plant Breeding (IPB), College of Agriculture and Food Science, University of the Philippines Los Baños, College, 4031 Los Baños, Laguna, Philippines. <sup>2</sup>Philippine Genome Center — Program For Agriculture, Livestock, Fisheries and Forestry, University of the Philippines Los Baños, College, Los Baños 4031, Laguna, Philippines. <sup>3</sup>Philippine Coconut Authority — Zamboanga Research Center, San Ramon, 7000 Zamboanga City, Philippines. <sup>4</sup>Institute 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**

- Anderson JR, Lubberstedt T (2003) Functional markers in plants. *Trend Plant Sci* 8:554–560
- 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
- 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* 101:132
- Harries HC (1978) Evolution, dissemination and classification of *Cocos nucifera* L. *Bot Rev* 44(3):265–320
- Kesawat MS, Kumar BD (2009) Molecular markers: it's application in crop improvement. *J Crop Sci Biotechnol* 12(4):169–181
- 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. *Afr J Biotechnol* 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)* 9(8):2377–2393. <https://doi.org/10.1534/g3.119.400215> PMID: 31167834; PMCID: PMC6686914
- Lebrun P, N'cho Y, 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 (ed) *Plant genotyping*. Springer, New York, 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, 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](https://doi.org/10.1007/978-1-4614-0920-5_10)
- Palliyarakkal MK, Ramaswamy M, Vadivel A (2011) Microsatellites in palm (Arecaceae) sequences. *Bioinformatics*. 7(7):347–351
- Perera L, Russell JR, Provan J, Powell W (1999) 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
- Perera L, Russell JR, Provan J, Powell W (2003) Studying genetic relationships among coconut varieties/populations using microsatellite markers. *Euphytica* 132:121–128
- 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 PIP, Hoher V, Verdeil JL, Yakandawala DMD, Weerakoon LK (2007) Recent advances in another 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. 2017. Agricultural foreign trade statistics of the Philippines: 2015. <https://psa.gov.ph>
- Philippine Statistics Authority. 2018. Selected Statistics on Agriculture 2013–2017. <https://psa.gov.ph>
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 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. <https://doi.org/10.1093/nar/gkh450>
- Sharma A, Namdeo AG, Mahadik KR (2008) Molecular markers: new prospects in plant genome analysis. *Pharmacogn Rev* 2(3):23–34
- Sindhumole P, Ambili SN (2011) Marker assisted breeding in coconut (*Cocos nucifera* L.). *Gregor Mendel Foundation Proceedings* 2011:30–32
- Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, Arnold GM, Karp A, Baudouin L, Rognon 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>
- 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. *Front Plant Sci* 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. *Trop Plant Biol* 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*. *Mol Breed* 34(2):1–11
- 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 Omics J* 6(3):193–200
- Xu Y, Crouch JH (2008) Marker-assisted selection in plant breeding: from publications to practice. *Crop Sci* 48:391–407. <https://doi.org/10.2135/cropsci2007.04.0191>

**Publisher's Note**

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

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)