New Development and Validation of 50 SSR Markers in Breadfruit (Artocarpus altilis, Moraceae) by Next-Generation Sequencing

Authors: De Bellis, Fabien, Malapa, Roger, Kagy, Valérie, Lebegin, Stéphane, Billot, Claire, et al.

Source: Applications in Plant Sciences, 4(8)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1600021

Applications in Plant Sciences 2016 4(8): 1600021

PRIMER NOTE

NEW DEVELOPMENT AND VALIDATION OF 50 SSR MARKERS IN BREADFRUIT (*ARTOCARPUS ALTILIS*, MORACEAE) BY NEXT-GENERATION SEQUENCING[1]

FABIEN DE BELLIS[2,5], ROGER MALAPA[3], VALÉRIE KAGY[4], STÉPHANE LEBEGIN[4], CLAIRE BILLOT[2], AND JEAN-PIERRE LABOUISSE[2]

[2]Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UMR AGAP, F-34398 Montpellier, France; [3]Vanuatu Agricultural Research and Technical Centre (VARTC), P.O. Box 231, Santo, Vanuatu; and [4]Institut Agronomique néo-Calédonien (IAC), P.O. Box 73, 98890 Païta, New Caledonia

- *Premise of the study:* Using next-generation sequencing technology, new microsatellite loci were characterized in *Artocarpus altilis* (Moraceae) and two congeners to increase the number of available markers for genotyping breadfruit cultivars.
- *Methods and Results:* A total of 47,607 simple sequence repeat loci were obtained by sequencing a library of breadfruit genomic DNA with an Illumina MiSeq system. Among them, 50 single-locus markers were selected and assessed using 41 samples (39 *A. altilis*, one *A. camansi*, and one *A. heterophyllus*). All loci were polymorphic in *A. altilis*, 44 in *A. camansi*, and 21 in *A. heterophyllus*. The number of alleles per locus ranged from two to 19.
- *Conclusions:* The new markers will be useful for assessing the identity and genetic diversity of breadfruit cultivars on a small geographical scale, gaining a better understanding of farmer management practices, and will help to optimize breadfruit genebank management.

Key words: *Artocarpus altilis*; breadfruit; high-throughput sequencing; Illumina; microsatellites; Moraceae.

Breadfruit (*Artocarpus altilis* (Parkinson) Fosberg, Moraceae) is a multipurpose tree crop with a great potential for increasing food security, thanks to its nutritious and starchy fruit (Ragone, 1997). In the Pacific Islands, it is a traditional staple crop, typically grown in backyards and small holdings. Breadfruit's wild progenitor, *A. camansi* Blanco, is a native species of New Guinea where the first steps of breadfruit domestication occurred. Pacific seafarers migrated eastward carrying breadfruit in the form of seeds or cuttings (Kirch, 1997). Other events, such as accumulated somatic mutations and meiotic defects in diploid genomes of *A. altilis*, resulted in seedless triploid cultivars that predominate in eastern Polynesia (Zerega et al., 2004). Witherup et al. (2013) developed simple sequence repeat (SSR) loci from microsatellite-enriched libraries and validated 25 of them across a large number of *A. altilis* cultivars, wild congeners, and relatives. This traditional SSR isolation approach is a cost- and labor-intensive process that requires repeat enrichment, cloning, and Sanger sequencing. Next-generation sequencing (NGS) technologies allow a good coverage of large genomes, cost-effective identification, and rapid characterization of hundreds of SSRs in nonmodel organisms without previous genomic resources (Zalapa et al., 2012). Gardner et al. (2015) used this technology to develop 15 chloroplast SSRs from transcriptome data of *Artocarpus* spp. We report here on the development and validation of a new set of 50 nuclear SSR markers for breadfruit and related species using NGS Illumina technology.

[1]Manuscript received 23 February 2016; revision accepted 14 April 2016.

The authors thank H. Vignes for assistance with the construction of the library, X. Argout for assistance with bioinformatics analysis, C. Hamelin for data curation of TropGeneDB, X. Perrier for helpful comments on the manuscript, and P. Biggins for editorial input.

[5]Author for correspondence: fabien.de_bellis@cirad.fr

doi:10.3732/apps.1600021

METHODS AND RESULTS

Leaf fragments of 41 samples of *A. altilis* (33 diploids, six triploids), *A. camansi*, and *A. heterophyllus* Lam. originating from Vanuatu, New Caledonia, French Polynesia, Tonga, Samoa, and the Mariana Islands were collected from living trees conserved in field genebanks (Appendix 1) and stored in a drying agent (silica gel) at room temperature. DNA was extracted according to the mixed alkyl trimethylammonium bromide (MATAB) protocol described by Risterucci et al. (2000). Total genomic DNA isolated from *A. altilis* 'Novan' (sample VUT002; National Center for Biotechnology Information [NCBI] BioSample SAMN04508170) was used to generate the library with the Nextera DNA Library Preparation Kit (Illumina, San Diego, California, USA) according to the manufacturer's protocol. Paired-end sequencing was carried out at the Grand Plateau Technique Régional platform (Montpellier, France; http://www.gptr-lr-genotypage.com) on an Illumina MiSeq system using the MiSeq Reagent Kit version 3 (2 × 300 bp). The sequences were assembled using ABySS software (Simpson et al., 2009). SSRs were detected using MISA Perl script (Thiel et al., 2003) with search parameters set as follows: at least five repeats for dinucleotide motifs, four repeats for trinucleotide motifs, and three repeats for tetra-, penta-, and hexanucleotide motifs. Primers were designed with Primer3 software using standard settings (Rozen and Skaletsky, 1999). A total of 2,341,465 paired-end sequences were assembled into 1,281,784 contigs. Among them, 115,499 contigs exhibited at least one microsatellite locus and enabled us to define PCR primers on 46,504 contigs, totaling 47,607 SSR loci (Appendix S1). The cumulative length of these contigs was around 15.5 Mb, totaling approximately 6% of the sequence length generated in this study. Raw sequencing data were submitted to the NCBI Sequence Read Archive (accession SRP070931) under BioProject PRJNA312880.

TABLE 1. Characteristics of 50 genomic SSRs developed in *Artocarpus altilis*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size (bp)	T_a (°C)	GenBank accession no.
mAaCIR0016	F: TTGACCCCTAGATGACCC	$(AAAAG)_8$	119	53	KU129040
	R: AGCCTTGAGCCCATGA				
mAaCIR0019	F: TGACATTCCCGCAAAA	$(TTC)_{11}$	122	52	KU129037
	R: AAGTCTTCTGTTCCTACTGACAA				
mAaCIR0027	F: TGACTTCCAACCCAAAATC	$(TCT)_{13}$	134	53	KU129070
	R: GTGGACTTACGATGTGAGGA				
mAaCIR0033	F: CGGGTACAGGGTATTGGT	$(ATA)_{12}$	141	53	KU129028
	R: AGGAGAGCGTTTGAGGAA				
mAaCIR0034	F: AACAGCAATCACCTGAAAC	$(ATA)_{11}$	142	50	KU129032
	R: TTGTTCGTCTCTATGTTCGT				
mAaCIR0036	F: TTTATGGGAGTGTTTTAGTG	$(ATT)_{11}$	144	47	KU129023
	R: CTCTTATATTGCTTGCTCC				
mAaCIR0038	F: GGAAATTCTTCATCCTCCC	$(TTA)_{12}$	145	53	KU129029
	R: CAAGATTGGCTGTTTGGTT				
mAaCIR0047	F: TCCCATCATCATCACCTT	$(TAT)_{14}$	150	51	KU129045
	R: AGCAATGACCATGCAAA				
mAaCIR0048	F: CGAAATCGGAACAGAAAAC	$(AGA)_{11}$	151	53	KU129062
	R: GTCCTTGGCTACTATAATCCCT				
mAaCIR0049	F: TACATACAAGCCAACTTCCA	$(ATA)_{13}$	151	51	KU129035
	R: CCTTTGTGAGGAAGACCA				
mAaCIR0050	F: TTCCCTGCCTAGTTTTGTG	$(TTA)_{11}$	152	53	KU129052
	R: AATAAAGCGCGGACTTACA				
mAaCIR0053	F: GCAACACATTCATCAACA	$(TTA)_{13}$	153	48	KU129069
	R: GACTCACCAAGACTTTATTACC				
mAaCIR0075	F: CATTCTTGGGAAGAGTTGA	$(TAA)_{12}$	171	51	KU129072
	R: ATAGCGGTGAAAATGGAA				
mAaCIR0078	F: CTTCAACTATTACTACTGCTGCT	$(TAT)_{11}$	173	49	KU129025
	R: CTGTTCAGGTTGGTGCT				
mAaCIR0080	F: AACACGGCCTATTTTGGA	$(TTA)_{15}$	174	54	KU129067
	R: GGCCATACAGGTTACGACA				
mAaCIR0081	F: AATTGGCGGTATTCTATG	$(TAT)_{14}$	175	48	KU129058
	R: GGAGGCAGATAAATTAGAAA				
mAaCIR0089	F: CCTGAGTAGGACAAAGACTGAA	$(GAAAA)_8$	183	53	KU129041
	R: ATTGCGCTTTTCTTCCC				
mAaCIR0090	F: GGGTGTCCTCGCCTC	$(AAG)_{11}$	184	52	KU129059
	R: GGTGGATCATTCAGCAAA				
mAaCIR0097	F: TCTCCGGTAAGGAAGGG	$(TTA)_{11}$	191	53	KU129042
	R: CCGAAAGTTACCAACCAAG				
mAaCIR0098	F: GACTAGAATGAAGTTAGGTTTG	$(AAT)_{16}$	192	47	KU129061
	R: ATGCCTACCAAGGTTTT				
mAaCIR0099	F: CCTGTTACGTTTCCTCC	$(ATGT)_{13}$	192	48	KU129026
	R: ACAATTAGACCTCAATGGAT				
mAaCIR0104	F: AAAATTGTGTTCCAGCCA	$(TAA)_{15}$	196	52	KU129071
	R: CGTTTACAAAAGGGTAGGG				
mAaCIR0108	F: CAATATAGCAGGCACTAATTCA	$(AAT)_{11}$	199	51	KU129049
	R: TCTTCTTTCTCTCGTTCGTT				
mAaCIR0111	F: TGCAGGCATCACGAAAC	$(TTC)_{16}$	202	54	KU129065
	R: CTTCTGCATGAGCGGTG				
mAaCIR0113	F: CCAGTAAGCTCCTTTACTACCA	$(AAT)_{11}$	204	53	KU129050
	R: GCCAAGAGCCACGTAAA				
mAaCIR0115	F: ACAGCTTTGCACCGACAC	$(ATA)_{12}$	205	55	KU129030
	R: GCCCTCAACCACCCC				
mAaCIR0121	F: GTGAGAGAATTTGAGTGATGTG	$(ATA)_{15}$	212	52	KU129054
	R: GGAAATCCACTACCCACC				
mAaCIR0127	F: TGATTCTCTCTTTACAGGCAC	$(AAG)_{13}$	218	52	KU129047
	R: GCTCAGGTGCTTACTTGTTC				
mAaCIR0128	F: CAACCACTGATGGAGATAG	$(ATA)_{12}$	219	48	KU129046
	R: ACAACACCGTTTACTGAAG				
mAaCIR0129	F: TTGCGGGACAGTGATTT	$(ATT)_{15}$	221	52	KU129048
	R: GTACGGGTTTTGGGAGAG				
mAaCIR0130	F: ACACCTACCTCTTCGGG	$(TTC)_{12}$	223	50	KU129056
	R: AGGTCTAATCCCAACCCT				
mAaCIR0134	F: AGCTGCCAATGATCCC	$(TTA)_{11}$	228	52	KU129044
	R: ATGTGAAAAGGTTGGATTTG				
mAaCIR0141	F: TCAAGCCCCTCACTCAA	$(AAT)_{11}$	235	53	KU129057
	R: ATGGCATAGCACAACACAA				
mAaCIR0145	F: GAACAGTGGAGTGCTTGGT	$(TAT)_{14}$	239	53	KU129043
	R: GTGGATGTTTGCCATGAA				

TABLE 1. Continued.

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size (bp)	T_a (°C)	GenBank accession no.
mAaCIR0146	F: CTTGCACCATCGTCATTT R: GTTAATTGGAAGTTGTGTCTCC	$(ATT)_{13}$	239	52	KU129066
mAaCIR0147	F: TGGCAAGAAAAGCCAAA R: GTTTCAACCACAGTCCAAAA	$(TCT)_{11}$	241	53	KU129024
mAaCIR0149	F: CGACCGGGACCTAACA R: ACCTGGTGTCATTATCGTTTT	$(ATG)_{11}$	243	53	KU129038
mAaCIR0152	F: TAGTTCTGGTATGGCATTTT R: AAAGGCACAGGGACTAAC	$(TTA)_{16}$	246	49	KU129051
mAaCIR0154	F: TCGAGGCCCTTGTTG R: GGAAATTCACCTTTCCTTG	$(AAT)_{11}$	250	52	KU129064
mAaCIR0164	F: GTTATCCGGCACCACC R: GAGTTAGGCAAAAGGGACA	$(TAA)_{15}$	265	52	KU129031
mAaCIR0167	F: AAGTGTTTGACAATGTGGTTAG R: ATGGAGCCTTGCTTTTG	$(TTA)_{14}$	267	51	KU129055
mAaCIR0169	F: GAAGCTATTTTCAAGGGA R: ATGTAAGGAAGTGTAGCAAA	$(TAT)_{14}$	270	47	KU129063
mAaCIR0172	F: GCTGTGAGAATGGTGTGG R: TCCGTCTTCGTACTGGTG	$(TTA)_{11}$	276	52	KU129033
mAaCIR0178	F: CCAGACCCAATCAACCA R: CAAGGACTCGCCCAAA	$(AAT)_{11}$	283	53	KU129034
mAaCIR0179	F: GTAGCACATGGCCCTACTC R: ATATACCCGTTGATGCCC	$(ATA)_{13}$	284	53	KU129053
mAaCIR0192	F: TGGGCTATTAAATTGGACTTTGG R: GCATCATGTTTGATTGCAGTTT	$(ATT)_{12}$	298	57	KU129068
mAaCIR0193	F: ACAAACCAACTCCGCCT R: GCCAGGGACGCATTT	$(ATT)_{11}$	298	53	KU129060
mAaCIR0195	F: AAAAGACCAGCCAAATCC R: TTGCTTTTCACGCTCTTC	$(AAT)_{11}$	313	52	KU129039
mAaCIR0204	F: TTTAGGGTCCGTTGAAGA R: GAAGTCTTGTTATTTGTGGAAG	$(TAT)_{11}$	330	50	KU129027
mAaCIR0205	F: TTAATAGGGCTTCTTCCCTT R: CACTGTGTTGATTGATCCC	$(AAG)_{11}$	337	52	KU129036

Note: T_a = annealing temperature.

[a]Additional information can be found in TropGeneDB, a multitropical crop information system, hosted by the SouthGreen bioinformatics platform (http://tropgenedb.cirad.fr/).

As a first step, 96 loci were selected according to the following criteria for motif type, repeat length, and amplicon size. We firstly excluded dinucleotide motifs, because these are prone to enzyme slippage during amplification, which may make allele designation difficult (Guichoux et al., 2011). Only perfect motifs were selected, as they are more likely to follow the stepwise mutation model. We selected loci with lengths of 11 to 16 repeats, as recommended by van Asch et al. (2010). Lastly, we selected loci with amplicon sizes ranging from 100 to 400 bp to facilitate the construction of multiplex sets.

The 96 primer pairs were then tested for amplification with a subset made up of four samples (two *A. altilis*, one *A. camansi*, and one *A. heterophyllus*). Only 15 failed to amplify. The remaining 81 primer pairs were classified according to their polymorphism and the overall quality of the profile. Among them, we chose to select only 50 polymorphic single-locus markers with no ambiguity in allele size determination (Table 1). These 50 SSRs were assessed using the 41 samples listed in Appendix 1. For comparison, we genotyped the same samples with 18 SSRs developed by Witherup et al. (2013). PCR reactions were performed in a solution A (25-μL total volume) containing 2.5 μL of PCR buffer (10 mM Tris-HCl, 50 mM KCl, 2 mM $MgCl_2$, 0.001% glycerol), 2.5 μL of dNTP (Jena Bioscience GmbH, Jena, Germany), 0.25 μL of $MgCl_2$, 0.2 μL of 10 μM forward primer with an M13 tail at the 5′-end (5′-CACGACGTTGTAAAACGAC-3′), 0.25 μL of 10 μM reverse primer, 0.25 μL of fluorescently labeled M13-tail (6-FAM, NED, VIC, or PET [Applied Biosystems, Foster City, California, USA]), 0.1 units of *Taq* DNA polymerase (Sigma-Aldrich, St. Louis, Missouri, USA), 5 μL of template DNA (5 ng/μL), and 14 μL of H_2O. The PCR conditions were as follows: an initial denaturation at 94°C for 5 min; 30 cycles at 94°C for 45 s, 55°C for 45 s, and 72°C for 1 min; and a final extension at 72°C for 10 min. PCR products were pooled in a solution B containing: 2 μL of 6-FAM, 2 μL of VIC, 2.5 μL of NED, and 3.5 μL of PET. From this solution B, a volume of 4 μL was taken and added to 10 μL of Hi-Di formamide and 0.12 μL of GeneScan 600 LIZ Size Standard and analyzed on an ABI 3500xL Genetic Analyzer (Life Technologies, Carlsbad, California, USA). Alleles were scored using GeneMapper version 4.1 software (Applied Biosystems). Basic statistics were computed using PowerMarker software (Liu and Muse, 2005).

Of the 50 loci assessed, all amplified and were polymorphic in *A. altilis*, 44 in *A. camansi*, and 21 in *A. heterophyllus*. The number of alleles per locus ranged from two (mAaCIR0167) to 19 (mAaCIR0121), with an average of seven alleles per locus (Table 2). When genotyping the samples with 18 of the SSRs developed by Witherup et al. (2013), we obtained similar results, but with a smaller number of alleles, ranging from one (MAA3) to 10 (MAA156) with an average of six alleles per locus (Appendix 2).

The Hardy–Weinberg equilibrium (HWE) test was only performed on diploids from Vanuatu and revealed that 20 of the new SSRs exhibited significant deviation from HWE (Table 2). This is not surprising as we did not sample populations but cultivated varieties, most of them clonally propagated and maintained in the form of a few trees planted in backyards or gardens. In the triploids, we calculated the percentage of heterozygous individuals and gave the number of individuals harboring one, two, or three alleles for each microsatellite locus. For 60% of the microsatellite loci, we observed unambiguous genotypes (i.e., with three alleles), ranging from one individual (mAaCIR0178) to five individuals (mAaCIR0080). Fifty percent of the loci were highly informative with a polymorphism information content value (PIC; Botstein et al., 1980), calculated on diploid data, greater than 0.7; only seven had a PIC less than 0.5, with a minimum value of 0.29 for mAaCIR0078. Although less informative, this latter category of loci may have characteristics, such as private alleles, useful for detecting admixture between species.

CONCLUSIONS

These 50 new nuclear SSR loci will be useful for assessing the identity and genetic diversity of breadfruit cultivars on a

TABLE 2. Genetic results of the 50 newly developed SSR markers of *Artocarpus altilis* and congeners.

	Diploid *A. altilis* ($n = 33$)																				
	Vanuatu ($n = 27$)						New Caledonia ($n = 6$)					Triploid *A. altilis* ($n = 6$)					*A. camansi* ($n = 1$)		*A. heterophyllus* ($n = 1$)		
Locus	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	HWE[b]	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	Obs	*A*	Allele size range (bp)[a]	% Het[c]	IndDis[d]	*A*	Allele size range (bp)[a]	*A*	Allele size range (bp)[a]	PIC
mAaCIR0016	25	6	93–122	0.258	0.646	*	5	3	92–122	0.600	0.549	6	3	97–122	50.0	0/3/3	1	104	—	—	0.66
mAaCIR0019	27	3	108–120	0.545	0.478	ns	6	1	117	0.000	0.000	6	3	102–120	50.0	3/0/3	1	108	1	104	0.44
mAaCIR0027	27	9	122–165	0.879	0.754	*	6	3	134–162	0.833	0.586	6	8	119–165	100.0	3/3/0	1	122	—	—	0.75
mAaCIR0033	27	6	128–149	0.545	0.708	*	6	4	137–146	1.000	0.612	6	3	134–146	66.7	0/4/2	1	134	2	123–135	0.73
mAaCIR0034	24	7	125–152	0.500	0.692	*	5	3	125–139	0.600	0.549	4	5	131–164	100.0	3/1/0	1	125	—	—	0.70
mAaCIR0036	26	8	132–156	0.781	0.766	*	5	4	137–150	0.400	0.419	6	4	122–156	83.3	0/5/1	1	144	—	—	0.77
mAaCIR0038	26	8	131–162	0.844	0.779	ns	6	3	131–153	0.833	0.540	6	4	134–156	50.0	0/3/3	—	—	—	—	0.75
mAaCIR0047	26	6	132–151	0.625	0.591	ns	5	2	132–147	0.200	0.162	6	6	135–163	100.0	3/3/0	1	135	—	—	0.53
mAaCIR0048	27	3	137–145	0.485	0.376	ns	6	1	137	0.000	0.000	6	2	137–142	33.3	0/2/4	1	136	2	115–124	0.36
mAaCIR0049	25	10	140–171	0.484	0.797	*	5	4	143–159	0.600	0.644	6	4	140–162	50.0	2/1/3	1	179	1	122	0.83
mAaCIR0050	22	6	147–166	0.857	0.685	ns	4	3	154–166	0.750	0.529	5	4	147–166	80.0	2/2/1	1	154	1	171	0.67
mAaCIR0053	27	15	141–191	0.909	0.805	ns	5	4	148–185	1.000	0.653	6	4	151–179	100.0	0/6/0	1	148	1	179	0.83
mAaCIR0075	25	7	162–192	0.516	0.584	ns	4	2	162–168	0.750	0.439	6	3	168–192	100.0	0/6/0	1	162	1	158	0.62
mAaCIR0078	25	3	158–170	0.355	0.302	ns	5	1	158	0.000	0.000	6	2	158–170	33.3	0/2/4	1	158	2	147–166	0.29
mAaCIR0080	25	15	146–199	0.806	0.835	*	5	5	146–171	0.800	0.680	6	7	149–196	83.3	5/0/1	1	211	—	—	0.85
mAaCIR0081	13	5	157–179	0.889	0.700	*	—	—	—	—	—	5	4	157–173	100.0	2/3/0	1	175	1	149	0.73
mAaCIR0089	13	2	170–180	0.053	0.470	*	5	2	170–180	0.000	0.256	6	2	170–180	50.0	0/3/3	1	165	1	170	0.41
mAaCIR0090	27	3	174–180	0.545	0.578	ns	6	3	174–180	0.833	0.586	6	3	171–177	100.0	0/6/0	1	177	1	166	0.57
mAaCIR0097	26	2	188–197	0.531	0.435	ns	5	1	188	0.000	0.000	6	4	188–197	83.3	3/2/1	1	188	—	—	0.32
mAaCIR0098	9	6	174–194	0.615	0.763	ns	2	2	180–189	0.000	0.250	4	4	162–192	100.0	2/2/0	1	162	—	—	0.80
mAaCIR0099	25	9	171–207	0.903	0.836	*	6	3	171–195	0.667	0.457	6	4	171–195	100.0	3/3/0	1	183	—	—	0.83
mAaCIR0104	26	11	177–216	0.813	0.715	ns	5	4	177–207	0.400	0.419	5	5	182–195	100.0	0/5/0	1	173	—	—	0.71
mAaCIR0108	22	9	188–228	0.821	0.814	*	5	4	198–213	1.000	0.622	5	6	192–216	100.0	2/3/0	1	188	1	176	0.82
mAaCIR0111	26	10	169–206	0.781	0.752	ns	5	3	184–206	0.400	0.308	6	6	175–197	100.0	3/3/0	1	200	2	175–177	0.75
mAaCIR0113	8	4	201–210	0.462	0.462	ns	2	2	195–204	0.500	0.281	3	3	195–207	66.7	0/2/1	1	195	—	—	0.50
mAaCIR0115	27	6	180–205	0.818	0.757	ns	6	3	180–201	0.167	0.417	6	3	180–214	83.3	0/5/1	1	201	2	180–201	0.72
mAaCIR0121	23	19	198–252	0.759	0.800	ns	5	7	210–243	1.000	0.747	5	8	231–252	100.0	4/1/0	—	—	—	—	0.84
mAaCIR0127	27	8	203–243	0.727	0.590	ns	6	3	208–226	1.000	0.601	6	3	208–226	66.7	3/1/2	1	206	1	200	0.60
mAaCIR0128	25	9	197–245	0.806	0.805	ns	5	4	209–232	0.400	0.419	6	6	197–227	83.3	2/3/1	1	200	1	218	0.83
mAaCIR0129	22	7	195–223	0.643	0.788	*	5	3	204–232	1.000	0.591	4	6	195–210	75.0	3/0/1	1	210	—	—	0.78
mAaCIR0130	27	4	219–228	0.697	0.705	ns	6	2	219–228	0.000	0.231	6	3	219–228	83.3	0/5/1	1	222	—	—	0.64
mAaCIR0134	24	9	215–249	0.733	0.786	ns	5	4	215–246	0.800	0.648	5	4	215–252	80.0	2/2/1	1	215	1	210	0.81
mAaCIR0141	25	5	228–243	0.806	0.643	ns	5	2	240–243	1.000	0.500	6	4	228–243	100.0	0/6/0	1	213	1	252	0.65
mAaCIR0145	26	5	215–245	0.469	0.623	*	5	3	215–236	1.000	0.561	6	3	215–242	50.0	0/3/3	1	224	—	—	0.63
mAaCIR0146	25	5	232–245	0.871	0.733	*	5	3	236–242	1.000	0.591	6	3	232–245	33.3	2/0/4	1	236	1	214	0.75
mAaCIR0147	26	4	229–242	0.844	0.680	*	6	3	236–242	1.000	0.590	6	4	229–242	100.0	2/4/0	1	229	—	—	0.66
mAaCIR0149	26	5	227–239	0.250	0.479	*	6	2	230–236	0.333	0.257	6	2	227–236	16.7	0/1/5	1	239	—	—	0.44
mAaCIR0152	26	8	222–247	0.719	0.721	*	5	3	228–234	0.800	0.538	6	5	216–238	100.0	2/4/0	2	259–262	—	—	0.70
mAaCIR0154	23	6	238–258	0.931	0.742	ns	5	5	238–261	1.000	0.669	5	6	231–258	100.0	2/3/0	1	243	1	225	0.75
mAaCIR0164	6	5	252–270	0.727	0.704	ns	3	4	258–270	0.333	0.522	5	5	243–270	60.0	2/1/2	1	243	—	—	0.76
mAaCIR0167	1	2	247–265	0.000	0.281	ns	1	1	247	0.000	0.000	1	1	247	0.0	0/0/1	—	—	—	—	0.36
mAaCIR0169	22	5	249–273	0.893	0.690	ns	5	4	249–270	1.000	0.622	6	4	249–273	66.7	2/2/2	—	—	—	—	0.65
mAaCIR0172	23	6	261–279	0.379	0.543	ns	5	3	264–276	0.800	0.507	6	3	264–276	100.0	2/4/0	1	277	—	—	0.60
mAaCIR0178	11	6	271–292	0.750	0.755	ns	4	2	277–280	0.750	0.439	4	5	271–289	100.0	1/3/0	—	—	—	—	0.72
mAaCIR0179	25	7	268–285	0.774	0.774	ns	5	3	268–288	0.400	0.403	5	5	268–282	100.0	4/1/0	1	274	—	—	0.77
mAaCIR0192	25	7	285–303	0.613	0.690	ns	5	3	285–297	0.800	0.585	6	4	282–294	83.3	0/5/1	1	285	—	—	0.67
mAaCIR0193	22	9	294–322	0.964	0.775	*	5	2	300–306	0.200	0.162	5	6	291–312	100.0	4/1/0	1	283	—	—	0.76
mAaCIR0195	20	10	293–317	0.538	0.806	*	5	3	293–309	0.000	0.448	5	4	290–305	60.0	2/1/2	—	—	—	—	0.79

TABLE 2. Continued.

	Diploid *A. altilis* ($n = 33$)											Triploid *A. altilis* ($n = 6$)				*A. camansi* ($n = 1$)		*A. heterophyllus* ($n = 1$)		
	Vanuatu ($n = 27$)					New Caledonia ($n = 6$)														
Locus	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	HWE[b]	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	Obs	*A*	Allele size range (bp)[a]	% Het[c]	IndDis[d]	*A*	Allele size range (bp)[a]	*A*	Allele size range (bp)[a]	PIC
mAaCIR0204	25	5	320–332	0.645	0.571	ns	5	2	320–329	0.200	0.353	6	3	323–329	66.7	0/4/2	1	323	1	301	0.58
mAaCIR0205	26	6	332–352	0.781	0.720	*	5	2	329–332	1.000	0.500	6	5	329–341	50.0	3/0/3	1	328	—	—	0.73

Note: *A* = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; HWE = Hardy–Weinberg equilibrium; Obs = number of observations; PIC = polymorphism information content value, calculated on diploid data, as described by Botstein et al. (1980).

[a] Including the 19-bp M13-tail.

[b] To test HWE, a χ^2 goodness-of-fit method was used; * = significant at $P < 0.05$, ns = no significant deviation from HWE.

[c] Due to allele ambiguity in the triploids, the $H_e : H_o$ ratio is replaced by the percentage of heterozygous genotypes.

[d] Number of triploid individuals with three, two, or one allele(s).

small geographical scale and for gaining a better understanding of farmer management practices (seed or vegetative propagation methods, exchanges, and dispersal). They will help to optimize the management of national genebanks by identifying duplicates and guiding future collecting activities. Of the 47,607 SSR loci identified, a very large number of additional markers could be further developed to address future research needs (genetic mapping, QTL, and association studies).

LITERATURE CITED

BOTSTEIN, D., R. L. WHITE, M. SKOLNICK, AND R. W. DAVIS. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314–320.

GARDNER, E. M., K. M. LARICCHIA, M. MURPHY, D. RAGONE, B. E. SCHEFFLER, S. SIMPSON, E. W. WILLIAMS, AND N. J. C. ZEREGA. 2015. Chloroplast microsatellite markers for *Artocarpus* (Moraceae) developed from transcriptome sequences. *Applications in Plant Sciences* 3: 1500049.

GUICHOUX, E., L. LAGACHE, S. WAGNER, P. CHAUMEIL, P. LÉGER, O. LEPAIS, C. LEPOITTEVIN, ET AL. 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11: 591–611.

KIRCH, P. V. 1997. The Lapita peoples: Ancestors of the Oceanic world, 203–212. Wiley-Blackwell, Cambridge, Massachussets, USA.

LIU, K., AND S. V. MUSE. 2005. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics (Oxford, England)* 21: 2128–2129.

NAVARRO, M., S. MALRES, J. P. LABOUISSE, AND O. ROUPSARD. 2007. Vanuatu breadfruit project: Survey on botanical diversity and traditional uses of *Artocarpus altilis. In* M. B. Taylor, J. Woodend, and D. Ragone [eds.], Proceedings of the first international symposium on breadfruit research and development, Nadi, Fiji, April 16–19, 2007, 81–87. Acta Horticulturae, 757. IHSH, Louvain, Belgium.

RAGONE, D. 1997. Breadfruit. *Artocarpus altilis* (Parkinson) Fosberg. IPGRI, Rome, Italy.

RISTERUCCI, A. M., L. GRIVET, J. A. K. N'GORAN, I. PIERETTI, M. H. FLAMENT, AND C. LANAUD. 2000. A high-density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* 101: 948–955.

ROZEN, S., AND H. J. SKALETSKY. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

SIMPSON, J. T., K. WONG, S. D. JACKMAN, J. E. SCHEIN, S. J. M. JONES, AND I. BIROL. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research* 19: 1117–1123.

THIEL, T., W. MICHALEK, R. VARSHNEY, AND A. GRANER. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

VAN ASCH, B., R. PINHEIRO, R. PEREIRA, C. ALVES, V. PEREIRA, F. PEREIRA, L. GUSMÃO, AND A. AMORIM. 2010. A framework for the development of STR genotyping in domestic animal species: Characterization and population study of 12 canine X-chromosome loci. *Electrophoresis* 31: 303–308.

WITHERUP, C., D. RAGONE, T. WIESNER-HANKS, B. IRISH, B. SCHEFFLER, S. SIMPSON, F. ZEE, ET AL. 2013. Development of microsatellite loci in *Artocarpus altilis* (Moraceae) and cross-amplification in congeneric species. *Applications in Plant Sciences* 1: 1200423.

ZALAPA, J. E., H. CUEVAS, H. ZHU, S. STEFFAN, D. SENALIK, E. ZELDIN, B. MCCOWN, ET AL. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99: 193–208.

ZEREGA, N. J. C., D. RAGONE, AND T. J. MOTLEY. 2004. Complex origins of breadfruit (*Artocarpus altilis*, Moraceae): Implications for human migrations in Oceania. *American Journal of Botany* 91: 760–766.

Appendix 1. Accession information for this study.[a]

Genebank location	Species	Original label	Cultivar name	Country of collection	Island	Village (County town)
VARTC/Santo	*A. altilis*	VUT001	Tiomal	Vanuatu	Malekula	Rano
VARTC/Santo	*A. altilis*	VUT002*	Novan	Vanuatu	Malekula	Rano
VARTC/Santo	*A. altilis*	VUT003	Baewok	Vanuatu	Malekula	Rano
VARTC/Santo	*A. altilis*	VUT013	Wawahisao	Vanuatu	Malo	Avunatari
VARTC/Santo	*A. altilis*	VUT029	Namnerlap	Vanuatu	Mota Lava	Gnerenigmen
VARTC/Santo	*A. altilis*	VUT061	Gortsaro	Vanuatu	Santo	Port Olry
VARTC/Santo	*A. altilis*	VUT102	Brobwe	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT103	Endoum	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT106	Koveuteap	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT107	Koveutniewe	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT109	Limbwedeng	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT118	Shienbase	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT127	Temelopsa	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT130	Teupanmei	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT131	Tewakon	Vanuatu	Ambrym	Port Vato
VARTC/Santo	*A. altilis*	VUT140	Matualelei	Vanuatu	Efate	Epule
VARTC/Santo	*A. altilis*	VUT141	Nako	Vanuatu	Efate	Tokararik
VARTC/Santo	*A. altilis*	VUT143	Nambatav-ani-Franck	Vanuatu	Efate	Epule
VARTC/Santo	*A. altilis*	VUT148	Lof	Vanuatu	Efate	Erakor
VARTC/Santo	*A. altilis*	VUT149	Naliu	Vanuatu	Efate	Erakor
VARTC/Santo	*A. altilis*	VUT150	Nasul	Vanuatu	Efate	Erakor
VARTC/Santo	*A. altilis*	VUT153	Pakala	Vanuatu	Efate	Mele
VARTC/Santo	*A. altilis*	VUT161	Aveloloa	Vanuatu	Efate	Pohi
VARTC/Santo	*A. altilis*	VUT167	Betima	Vanuatu	Malekula	Brenwei
VARTC/Santo	*A. altilis*	VUT200	Nakut-ulcecerea	Vanuatu	Nguna	Rewoka
VARTC/Santo	*A. altilis*	VUT205	Narongrong	Vanuatu	Nguna	Rewoka
VARTC/Santo	*A. altilis*	VUT212	Napeere	Vanuatu	Nguna	Urapua
IAC/Saint-Louis	*A. altilis*	Kea	Kea	Tonga	Tongatapu	—
IAC/Saint-Louis	*A. altilis*	Hamoa	Hamoa	French Polynesia	Tahaa, Society Islands	—
IAC/Saint-Louis	*A. altilis*	Ma'afala B3	Ma'afala	Samoa	—	—
IAC/Saint-Louis	*A. altilis*	Lemae B5	Lemae	Mariana Islands	—	—
IAC/Saint-Louis	*A. altilis*	Rotuma 15C2	Rotuma	French Polynesia	Tahaa, Society Islands	—
IAC/Saint-Louis	*A. altilis*	Puo	Puo	Tonga	Tongatapu	—
IAC/Saint-Louis	*A. altilis*	A1	—	New Caledonia	Grande Terre	Tiéti (Poindimié)
IAC/Saint-Louis	*A. altilis*	A2	—	New Caledonia	Grande Terre	Parawié (Houailou)
IAC/Saint-Louis	*A. altilis*	A3	—	New Caledonia	Grande Terre	Parawié (Houailou)
IAC/Saint-Louis	*A. altilis*	A7	—	New Caledonia	Grande Terre	Tyé (Touho)
IAC/Saint-Louis	*A. altilis*	A9	—	New Caledonia	Grande Terre	Sainte-Marie (Pouébo)
IAC/Pocquereux	*A. altilis*	A11	—	New Caledonia	Maré	Medhu
IAC/Pocquereux	*A. camansi*	A13-camansi	—	New Caledonia	Grande Terre	Vallée des Colons (Nouméa)
IAC/Pocquereux	*A. heterophyllus*	heterophyllus	—	New Caledonia	Grande Terre	Appala (Koumac)

Note: — = not available.

[a]The study was conducted with 41 living accessions forming two subsets of the collections of the Vanuatu Agricultural Research and Technical Centre (VARTC) and the Institut Agronomique néo-Calédonien (IAC). The VARTC genebank is located on the island of Santo, Vanuatu (15.453°S, 167.184°E). The 27 VARTC accessions (VUT label) were collected during two surveys conducted across the Vanuatu archipelago in 2004–2005 by Navarro et al. (2007) and in 2009 by Mies (no published data) with the support of the Pacific Plant Genetic Resources Network under the auspices of the Pacific Community (SPC). The 14 IAC accessions are conserved at the research stations of Saint-Louis (22.232°S, 166.538°E) and Pocquereux (21.731°S, 165.886°E) in New Caledonia. They comprise six accessions of *A. altilis*, one *A. heterophyllus*, and one *A. camansi*, all of them collected in New Caledonia, and six seedless cultivars sent to IAC by the National Tropical Botanical Garden of Hawaii (NTBG) from the Kahanu Garden, Maui Island in 1999. Leaf fragments were collected from living trees conserved in the field genebanks and stored in a drying agent (silica gel) at room temperature before performing DNA extraction.

*Sample used to generate the genomic library (NCBI BioSample SAMN04508170).

Appendix 2. Genetic properties of 18 SSR markers developed by Witherup et al. (2013) tested on *Artocarpus altilis* and congeners.

	Diploid *A. altilis* ($n = 33$)											Triploid *A. altilis* ($n = 6$)					*A. camansi* ($n = 1$)		*A. heterophyllus* ($n = 1$)		
	Vanuatu ($n = 27$)						New Caledonia ($n = 6$)														
Locus	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	HWE[b]	Obs	*A*	Allele size range (bp)[a]	H_o	H_e	Obs	*A*	Allele size range (bp)[a]	% Het[c]	IndDis[d]	*A*	Allele size range (bp)[a]	*A*	Allele size range (bp)[a]	PIC
MAA3	26	1	234	0.000	0.000	—	5	1	234	0.000	0.000	6	1	234	0.0	0/0/6	1	234	1	236	0.05
MAA40	26	6	195–210	0.875	0.810	*	5	2	208–210	0.800	0.461	6	3	201–210	100.0	0/6/0	1	199	1	189	0.81
MAA54a	26	8	191–212	0.813	0.819	ns	5	5	191–212	1.000	0.700	4	5	191–212	75.0	0/3/1	1	196	1	181	0.83
MAA54b	25	3	222–232	0.613	0.613	ns	5	2	222–224	0.800	0.461	4	3	220–232	75.0	1/2/1	1	224	1	217	0.57
MAA71	26	9	172–195	0.875	0.779	ns	5	5	180–198	1.000	0.669	5	6	170–195	100.0	2/3/0	1	174	2	142–183	0.79
MAA85	22	5	174–187	0.679	0.707	ns	5	3	178–187	0.200	0.385	5	3	174–179	40.0	0/2/3	—	—	—	—	0.69
MAA96	23	7	220–230	0.310	0.727	*	5	4	220–228	0.800	0.585	6	5	222–230	83.3	0/5/1	1	226	—	—	0.76
MAA122	25	7	297–311	0.774	0.717	ns	5	3	305–311	1.000	0.607	5	4	297–309	100.0	0/5/0	1	299	1	272	0.73
MAA135	24	9	287–312	0.900	0.806	ns	5	4	289–308	1.000	0.622	4	4	287-308	100.0	3/1/0	1	318	1	370	0.81
MAA140	27	8	150–179	0.545	0.774	*	6	4	150–169	0.833	0.586	6	3	165–179	83.3	2/3/1	1	159	1	151	0.78
MAA145	23	6	294–323	0.679	0.690	*	5	4	294–308	0.800	0.632	5	5	305–323	100.0	2/3/0	1	286	1	294	0.71
MAA156	27	10	290–319	0.697	0.699	ns	6	4	290–311	1.000	0.680	6	4	295–324	83.3	5/0/1	1	292	1	292	0.72
MAA182	27	8	197–230	0.303	0.566	*	6	1	224	0.000	0.000	6	3	221–227	50.0	0/3/3	1	226	1	191	0.53
MAA201	26	4	281–300	0.406	0.573	ns	6	2	281–282	0.000	0.231	6	7	281–311	83.3	2/3/1	1	294	2	258–282	0.53
MAA219	24	4	279–291	0.733	0.607	ns	5	3	279–300	1.000	0.561	6	3	279–291	50.0	0/3/3	1	294	1	279	0.57
MAA251	26	3	190–208	0.625	0.536	ns	6	2	190–208	0.000	0.231	5	5	190–217	100.0	2/3/0	1	214	1	211	0.49
MAA287	27	7	204–241	0.242	0.683	*	6	1	212	0.000	0.000	6	5	200–228	83.3	3/2/1	1	196	2	196–228	0.63
MAA293	27	4	176–188	0.727	0.616	ns	6	3	176–184	1.000	0.590	6	3	178–184	83.3	0/5/1	1	180	1	184	0.56

Note: *A* = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; HWE = Hardy–Weinberg equilibrium; Obs = number of observations; PIC = polymorphism information content value, calculated on diploid data, as described by Botstein et al. (1980).

[a] Including the 19-bp M13-tail.

[b] To test HWE, a χ^2 goodness-of-fit method was used; * = significant at $P < 0.05$, ns = no significant deviation from HWE.

[c] Due to allele ambiguity in the triploids, the H_e:H_o ratio is replaced by the percentage of heterozygous genotypes.

[d] Number of triploid individuals with three, two, or one allele(s).