Chloroplast Microsatellite Markers for Artocarpus (Moraceae) Developed from Transcriptome Sequences

Authors: Gardner, Elliot M., Laricchia, Kristen M., Murphy, Matthew, Ragone, Diane, Scheffler, Brian E., et al.

Source: Applications in Plant Sciences, 3(9)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500049

Applications in Plant Sciences 2015 3(9): 1500049

PRIMER NOTE

CHLOROPLAST MICROSATELLITE MARKERS FOR *ARTOCARPUS* (MORACEAE) DEVELOPED FROM TRANSCRIPTOME SEQUENCES[1]

ELLIOT M. GARDNER[2,3,7], KRISTEN M. LARICCHIA[2,3], MATTHEW MURPHY[4], DIANE RAGONE[5], BRIAN E. SCHEFFLER[6], SHERON SIMPSON[6], EVELYN W. WILLIAMS[2], AND NYREE J. C. ZEREGA[2,3]

[2]Department of Plant Science, Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, Illinois 60022 USA; [3]Plant Biology and Conservation, Northwestern University, 2205 Tech Drive, Hogan 2-144, Evanston, Illinois 60208 USA; [4]Department of Biology, Illinois College, 1101 West College Avenue, Jacksonville, Illinois 62650 USA; [5]Breadfruit Institute, National Tropical Botanical Garden, Kahaleo, Hawai'i 96741 USA; and [6]USDA, ARS, Genomics and Bioinformatics Research Unit, 141 Experiment Station Road, Stoneville, Mississippi 38776-0350 USA

- *Premise of the study:* Chloroplast microsatellite loci were characterized from transcriptomes of *Artocarpus altilis* (breadfruit) and *A. camansi* (breadnut). They were tested in *A. odoratissimus* (terap) and *A. altilis* and evaluated in silico for two congeners.
- *Methods and Results:* Fifteen simple sequence repeats (SSRs) were identified in chloroplast sequences from four *Artocarpus* transcriptome assemblies. The markers were evaluated using capillary electrophoresis in *A. odoratissimus* (105 accessions) and *A. altilis* (73). They were also evaluated in silico in *A. altilis* (10), *A. camansi* (6), and *A. altilis* × *A. mariannensis* (7) transcriptomes. All loci were polymorphic in at least one species, with all 15 polymorphic in *A. camansi*. Per species, average alleles per locus ranged between 2.2 and 2.5. Three loci had evidence of fragment-length homoplasy.
- *Conclusions:* These markers will complement existing nuclear markers by enabling confident identification of maternal and clone lines, which are often important in vegetatively propagated crops such as breadfruit.

Key words: *Artocarpus altilis*; *Artocarpus camansi*; *Artocarpus mariannensis*; *Artocarpus odoratissimus*; breadfruit; Moraceae.

Artocarpus J. R. Forst. & G. Forst. (Moraceae) contains approximately 70 species of monoecious trees with a center of diversity in Malesia (Zerega et al., 2010). Species include several underutilized crops that can improve food security (Jones et al., 2011). In addition to breadfruit (*A. altilis* (Parkinson) Fosberg) and jackfruit (*A. heterophyllus* Lam.), *Artocarpus* contains lesser-known crops like cempedak (*A. integer* (Thunb.) Merr.) and terap (*A. odoratissimus* Blanco), and more than a dozen other species with edible fruits whose potential remains largely unexplored (Zerega et al., 2010).

Nuclear microsatellites developed for *Artocarpus* (Witherup et al., 2013) have been used in characterizing genetic diversity of breadfruit germplasm (Zerega et al., 2015). We present primers for 15 chloroplast simple sequence repeat (SSR) loci from transcriptomes of *A. altilis* and *A. camansi* that will complement the nuclear markers in analyzing genetic diversity and population structure. Chloroplast SSRs are usually mononucleotide repeats, and as nonrecombinant, maternally inherited loci (Wheeler et al., 2014), they allow confident identification of maternal and clone lines—often important in vegetatively propagated crops such as breadfruit. These markers were developed from next-generation sequencing (NGS) transcriptome data. This approach enables rapid marker development directly from sequences in the target organisms. Primers were tested in *A. altilis* (diploid and triploid) and *A. odoratissimus*. We also constructed an in silico data set from additional transcriptomes of *A. altilis*, its wild progenitor (*A. camansi* Blanco), and *A. altilis* × *A. mariannensis* hybrids to test for fragment size homoplasy, a common problem with chloroplast SSRs that can overestimate relatedness by masking sequence variations that do not change allele sizes (Wheeler et al., 2014).

METHODS AND RESULTS

Total RNA from two *A. altilis* accessions and one *A. camansi* accession (Appendix 1) was extracted using the QIAGEN RNeasy Universal Mini Kit (QIAGEN, Valencia, California, USA). Illumina TruSeq library preparation and sequencing in one lane of HiSeq 2000 (2 × 100, paired-end; Illumina, San Diego, California, USA) took place at Argonne National Laboratory (Lemont, Illinois, USA). Reads were de-multiplexed, quality-trimmed (>Q20 in a 5-bp window), and assembled using Trinity (Grabherr et al., 2011; Bolger et al., 2014). Chloroplast contigs were extracted using a BLAST search seeded with the *Morus indica* L. (Moraceae) chloroplast genome (GI: 89,574,460). Mono- and dinucleotide repeats were identified, aligned using BLAST, and screened for variability. Initially, primers for 16 chloroplast SSR loci were designed

[1]Manuscript received 28 April 2015; revision accepted 7 June 2015.

The authors thank the Sabah Forestry Department (J. Pereira), Sabah Agriculture Department (J. Linton, W. F. Au), and the Vanuatu Agricultural Research and Technical Center for coordinating fieldwork and providing access to living collections. This study was funded in part by the National Science Foundation (DEB REVSYS 0919119, DBI REU 1062675), the American Society of Plant Taxonomists, the Garden Club of America, Northwestern University, and USDA-ARS National Programs.

[7]Author for correspondence: egardner@u.northwestern.edu

doi:10.3732/apps.1500049

using Primer3 (Rozen and Skaletsky, 1999) (Table 1). Fifteen loci reliably amplified and were subjected to further testing.

To test for variability in *A. odoratissimus*, all loci were amplified in 105 accessions collected from four districts in Sabah, Malaysia (Appendix 2). PCR reactions were performed in two steps (Schuelke, 2000). For the first step, 10-μL reactions contained 5 μL of MyTaq Master Mix (Bioline USA, Taunton, Massachusetts, USA), 0.5 μL of 10 mg/mL bovine serum albumin (BSA), 0.25 μL of 10 μM forward primer with the M13 tail (5′-CAGGAAACAGCTATGAC-3′), 0.25 μL of 10 μM reverse primer, 3 μL of H_2O, and 1 μL of template DNA. PCR conditions for the first step were 94°C for 3 min; 13 cycles at 94°C for 30 s, 59.8°C for 30 s, and 72°C for 1 min; and 72°C for 10 min. The following were then immediately added: 2.5 μL MyTaq Master Mix, 0.25 μL of 10 mg/mL BSA, 0.125 μL of 2.5 μM $MgCl_2$, 0.25 μL of 10 μM labeled M13 primer (WellRED Dye D2, D3, or D4 [Beckman Coulter, Brea, California, USA]), and 1.875 μL of H_2O. PCR conditions for the second step were 94°C for 3 min; 27 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min; and 72°C for 10 min. Product was pooled as follows: 2 μL of D4-labeled product, 1 μL of D3, and 0.5 μL of D2. Pooled products were added to 30 μL of HiDi formamide (Azco Biotech, San Diego, California, USA) and 3.3 μL of 400-bp size standard (Beckman Coulter) and analyzed on a Beckman Coulter CEQ 8000 Genetic Analysis System. Alleles were scored using the CEQ 8000 software version 9.0 (Beckman Coulter).

To test for variability in *A. altilis*, all loci except AALTCP04, AALTCP07, AALTCP11, and AALTCP12 (which were less variable in transcriptomes) were amplified in 73 accessions of *A. altilis* from Vanuatu (Navarro et al., 2005, 2007), the Caribbean, and India (Appendix 1). Locus AALTCP14 followed the protocol described above. Other loci were amplified at the USDA in reduced PCR reaction volumes (step 1: 5 μL, step 2: 3 μL) without BSA using QIAGEN Multiplex PCR Master Mix (QIAGEN) and analyzed using ABI reagents on a 3730xl DNA Analyzer and GeneMapper 5 (Applied Biosystems, Foster City, California, USA).

To test for variability in *A. camansi* and *A. altilis* × *A. mariannensis* hybrids and to explore the presence of homoplasy in these markers, loci were amplified in silico from the draft genome of *A. camansi*, the original four transcriptomes used for developing primers, and 18 additional transcriptome assemblies (Laricchia, 2014) (Appendix 1). Chloroplast contigs were extracted using the BLAST method described above, and amplification in silico took place following Bikandi et al. (2004). Some loci that failed to amplify because the region was split between two contigs or because a priming site was truncated were recovered using BLAST. Sequences were aligned using MUSCLE (Edgar, 2004), and a fragment-length data set was constructed. For both data sets, the number of alleles and a haplotype diversity index for each locus were calculated using GenAlEx (Table 2) (Peakall and Smouse, 2012).

Allele sizes were recovered from >60 individuals of *A. odoratissimus* for all loci but one (37 individuals for AALTCP05), and from >60 individuals of *A. altilis* for all 11 tested loci (Table 2). In silico capture recovered sequences and fragment sizes from most transcriptomes for all loci except AALTCP13, which tended to be absent from transcriptomes (Table 2, Appendix 3). All loci were polymorphic in the breadfruit complex (*A. altilis*, *A. camansi*, and *A. altilis* × *mariannensis* hybrids), with *A. camansi* showing the greatest unbiased haplotype diversity. Although the in silico sample size was small, this finding is consistent with a domestication bottleneck in *A. altilis* with respect to its wild progenitor, *A. camansi* (Zerega et al., 2005). The polymorphism in AALTCP04 in *A. camansi* was not in the repeat motif, but in a 22-bp indel. Six loci (AALTCP03, AALTCP05, AALTCP08, AALTCP10, AALTCP11, and AALTCP12) were monomorphic in *A. odoratissimus*. Average alleles per locus was 2.5 in *A. altilis*, 2.3 in hybrids and *A. odoratissimus*, and 2.2 in *A. camansi*. For comparison, average alleles per locus in the previously described nuclear markers using the same individuals as our in silico data set (with one parent-sibling substitution in *A. camansi*) were 2.1 in *A. camansi*, 5.0 in *A. altilis*, and 4.6 in hybrids (Zerega et al., 2015).

The in silico data revealed within-species homoplasy due to multiple SSRs in the same amplified fragment in loci AALTCP01, AALTCP09, and AALTCP10. All other loci showed no evidence of fragment-length homoplasy. We also identified single-nucleotide polymorphisms in flanking regions outside the target repeats in loci AALTCP01, AALTCP02, AALTCP07, AALTCP09,

Table 1. Chloroplast SSRs developed in this study, showing region, primers, motif, melting temperature, suggested pool and dye color for multiplexing, and GenBank accession number for sequences from *Artocarpus camansi* (NTBG 960,576.001).

Locus	Region	Primer sequences (5′–3′)	Repeat motif	T_a (°C)[a]	Pool/Dye	GenBank accession no.
AALTCP01	*ndhA*	F: TTGGGGCTTTACGTTGGTAG	$(T)_9$, $(C)_7(T)_7$	60.0	1/D4	KR185519
		R: CGTTTCTATTCTTCTTTCTCTTTCTG		58.5		
AALTCP02	*ndhA*	F: CAGAAAGAGAAAGAAGAATAGAAACG	$(A)_{10}$, $(TA)_7$	58.5	2/D4	KR185520
		R: AAACTCGTCCTTTTCACTTACGA		59.4		
AALTCP03	*petB*	F: ACCTCGTGGCCGGACTTT	$(T)_{14}$	63.0	4/D4	KR185521
		R: TCCTTGAGTAAGAACCGTTGG		59.2		
AALTCP04	*petB–petD*	F: TCACTTGGGGTAGGAACAATAG	$(TA)_6$, 22-bp indel	58.1	1/D3	KR185522
		R: TTCTGCATAGCCCAATCAAT		58.1		
AALTCP05	*psbE–petL*	F: TTCCAAGGATAGGGCTTTGT	$(A)_{11}$	58.7	4/D3	KR185523
		R: TTTTATTGTATCGCCGAATCC		59.0		
AALTCP06	*rpl16*	F: TGAATCATCCACCTTACCTTACA	$(T)_7A(T)_9$	58.5	1/D3	KR185524
		R: CATCGTTTCGCATTATCTGG		59.1		
AALTCP07	*rps8–rpl14*	F: TTTTTATTCATGTCAGCATTTCG	$(T)_{10}$	59.1	5/D4	KR185525
		R: AGGAAATTGTTTGTGTCTCACG		59.1		
AALTCP08	*rpl14–rp116*	F: TCAAATGGGTTTGAGGTTGA	$(A)_{11}$, $(T)_9$	59.0	3/D4	KR185526
		R: AGCGGTATCCAAAATGCCTA		59.6		
AALTCP09	*trnS–trnG*	F: TCCGACGCTTTAGTCCACTC	$(T)_{13}$, 5-bp indel	60.4	2/D3	KR185527
		R: GCCAAGCCGTGTAAAGAAAA		60.2		
AALTCP10	*trnS–trnG*	F: GGGCCTCTTTTGTTCTAACG	$(T)_9$, $(A)_9$	58.8	3/D2	KR185528
		R: TGTTCAAAAATTCATAGTTCTTTTGTT		58.7		
AALTCP11	*rps16*	F: GCCGTACGAGGAGAAAACTTC	$(G)_5A(G)_9$	60.3	4/D2	KR185529
		R: GCGCCCTTTTCAAGGAAATA		61.4		
AALTCP12	*rps16*	F: GCTCTTCGGAAAGTGGGTTT	$(AT)_6$	60.6	5/D4	KR185530
		R: TCATTCACACCTTAACGCTCT		57.5		
AALTCP13	*rps16*	F: GAAAGTGCTTTTGGCTCGAC	$(T)_{12}(G)_{10}$	60.0	3/D3	KR185531
		R: AGATTCTTGCCTCCGAAAAA		58.9		
AALTCP14	*trnT–trnE*	F: CGGATTTGAACCGATGACTT	$(TA)_5$	59.9	2/D2	KR185532
		R: TCGTTCCTGAGTGAACCACTA		58.3		
AALTCP15	*trnT–trnE*	F: TGGTTCACTCAGGAACGATAAA	$(A)_8$	59.6	1/D2	KR185533
		R: TGGATCTAGGTTGAATTGGTAGG		59.4		

Note: T_a = annealing temperature.

[a]All primers amplified with an annealing temperature of 59.8°C (step 1) and 55°C (step 2).

TABLE 2. Summary of allele size data for species in the breadfruit complex.

	A. odoratissimus				*A. camansi*[a]				*A. altilis*				*A. altilis* × *A. mariannensis*[a]			
Locus	*N*	*A*	ASR (bp)	*h*	*N*	*A*	ASR (bp)	*h*	*N*	*A*	ASR (bp)	*h*	*N*	*A*	ASR (bp)	*h*
AALTCP01	106	4	187–190	0.532	6	3	171–174	0.733	73	4	190–193	0.648	7	2	171–172	0.286
AALTCP02	105	2	186–187	0.074	6	2	175–178	0.600	73	2	193–194	0.027	7	3	174–176	0.667
AALTCP03	93	1	206	0.000	6	2	193–194	0.533	73	4	209–213	0.549	7	4	191–196	0.810
AALTCP04	98	3	229–233	0.099	6	2	210–232	0.600	10[a]	1	210	0.000	7	1	210	0.000
AALTCP05	61	1	247	0.000	6	3	231–233	0.600	66	2	248–249	0.142	7	2	231–232	0.286
AALTCP06	105	2	252–253	0.038	6	2	229–232	0.600	73	3	248–252	0.475	6	3	229–233	0.733
AALTCP07	104	2	204–208	0.019	6	2	183–184	0.533	11[a]	2	183–184	0.436	7	2	183–184	0.286
AALTCP08	105	1	232	0.000	6	2	216–218	0.600	73	2	233–234	0.465	7	3	214–217	0.667
AALTCP09	106	6	228–234	0.641	6	3	203–206	0.733	69	3	220–225	0.470	6	4	202–207	0.800
AALTCP10	89	1	278	0.000	5	2	280–281	0.600	71	3	296–299	0.481	7	2	278–281	0.286
AALTCP11	91	1	228	0.000	6	2	209–211	0.600	10[a]	1	211	0.000	5	1	211	0.000
AALTCP12	37	1	236	0.000	6	2	221–223	0.600	9[a]	1	221	0.000	7	2	221–223	0.476
AALTCP13	104	2	163–167	0.379	2	2	148–151	—	72	5	169–173	0.559	2	2	149–152	—
AALTCP14	103	3	215–219	0.246	4	2	198–200	0.500	65	2	218–220	0.031	5	1	198	0.000
AALTCP15	103	4	220–223	0.182	5	2	202–203	0.600	73	3	219–221	0.475	5	3	202–204	0.700

Note: *A* = number of alleles; ASR = allele size range; *h* = unbiased haplotype diversity; *N* = number of individuals.
[a] In silico data, without the 17-bp M13 tail.

AALTCP12, and AALTCP14 (in *A. camansi* only for AALTCP02, AALTCP09, AALTCP12, and AALTCP14). These loci thus may provide additional resolution when a sequencing approach is used as opposed to a fragment-size approach.

CONCLUSIONS

These chloroplast SSR loci will be useful for rapid and low-cost genotyping in *Artocarpus* and possibly in other Moraceae species, given the level of conservation typical in chloroplast genomes. By enabling the isolation of maternal lineages, these markers can be applied to characterizing genetic diversity, tracing seed and vegetative dispersal history, and assessing relatedness of germplasm accessions. Even as NGS tools become more widespread, SSRs remain important, as they enable efficient genotyping with common laboratory equipment. This is particularly relevant for nonmodel, underutilized crops, which are often grown in less developed areas where only basic genotyping equipment is available.

LITERATURE CITED

BIKANDI, J., R. SAN MILLÁN, A. REMENTERIA, AND J. GARAIZAR. 2004. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics (Oxford, England)* 20: 798–799.

BOLGER, A. M., M. LOHSE, AND B. USADEL. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)* 30: 2114–2120.

EDGAR, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

GRABHERR, M. G., B. J. HAAS, M. YASSOUR, J. Z. LEVIN, D. A. THOMPSON, I. AMIT, X. ADICONIS, ET AL. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

JONES, A. M. P., D. RAGONE, N. G. TAVANA, D. W. BERNOTAS, AND S. J. MURCH. 2011. Beyond the bounty: Breadfruit (*Artocarpus altilis*) for food security and novel foods in the 21st century. *Ethnobotany Research and Applications* 9: 129–149.

LARICCHIA, K. 2014. Transcriptome analysis of domesticated breadfruit and its wild relatives. Master's thesis, Northwestern University, Evanston, Illinois, USA.

NAVARRO, M., J.-P. LABOUISSE, S. MALRES, D. RAGONE, AND O. ROUPSARD. 2005. Vanuatu Breadfruit Project. Report prepared for Vanuatu Agricultural Research and Technical Center, Wellington, New Zealand; Secretariat of the Pacific Community, Noumea, New Caledonia; and Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Paris, France.

NAVARRO, M., S. MALRES, J.-P. LABOUISSE, AND O. ROUPSARD. 2007. Vanuatu Breadfruit Project: Survey on botanical diversity and traditional uses of *Artocarpus altilis*. *Acta Horticulturae* 757: 81–88.

PEAKALL, R., AND P. E. SMOUSE. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—An update. *Bioinformatics* 28: 2537–2539.

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.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

WHEELER, G. L., H. E. DORMAN, A. BUCHANAN, L. CHALLAGUNDLA, AND L. E. WALLACE. 2014. A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. *Applications in Plant Sciences* 2: 1400059.

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

ZEREGA, N., T. WIESNER-HANKS, D. RAGONE, B. IRISH, B. SCHEFFLER, S. SIMPSON, AND F. ZEE. 2015. Diversity in the breadfruit complex (*Artocarpus*, Moraceae): Genetic characterization of critical germplasm. *Tree Genetics & Genomes* 11: 4.

ZEREGA, N. J. C., M. N. NUR SUPARDI, AND T. J. MOTLEY. 2010. Phylogeny and recircumscription of Artocarpeae (Moraceae) with a focus on *Artocarpus*. *Systematic Botany* 35: 766–782.

ZEREGA, N. J. C., D. RAGONE, T. MOTLEY, AND W. ZOMLEFER. 2005. Systematics and species limits of breadfruit (*Artocarpus*, Moraceae). *Systematic Botany* 30: 603–615.

APPENDIX 1. Accession and locality information for *Artocarpus altilis*, *A. camansi*, and *A. altilis* × *A. mariannensis*. Individuals labeled "NTBG" are part of a living germplasm collection at the National Tropical Botanical Garden's Breadfruit Institute (Kalaheo, Hawai'i, USA). Germplasm source localities appear in parentheses. Individuals labeled "VUT" were collected as part of the Vanuatu Breadfruit Project; detailed accession information appears in Navarro et al. (2005), and additional information about 36 accessions comprising a living collection at the Vanuatu Agricultural Research and Technical Center appears in Navarro et al. (2007). Individuals labeled "CHIC" refer to vouchers deposited at the Chicago Botanic Garden Nancy Poole Rich Herbarium (CHIC). Asterisks denote the individuals used for the initial marker development. FSM = Federated States of Micronesia.

Artocarpus altilis *NTBG 030042.001 (Society Islands), *NTBG 040063.001 (Samoa), NTBG 900261.001 (Fiji), NTBG 790487.001 (Society Islands), NTBG 890479.002 (Pohnpei, FSM), NTBG 890167.002 (Pohnpei, FSM), NTBG 880690.001 (Tonga), NTBG 790485.001 (Society Islands), NTBG 900265.001 (Fiji), NTBG 890455.001 (Samoa), VUT001, VUT002, VUT003, VUT004, VUT005, VUT006, VUT007, VUT008, VUT009, VUT010, VUT011, VUT012, VUT013, VUT014, VUT015, VUT016, VUT017, VUT018, VUT019, VUT020, VUT021, VUT022, VUT023, VUT024, VUT025, VUT026, VUT027, VUT028, VUT029, VUT030, VUT031, VUT032, VUT033, VUT034, VUT035, VUT036, VUT037, VUT038, VUT039, VUT040, VUT041, VUT042, VUT043, VUT044, VUT045, VUT046, VUT047, VUT048, VUT049, VUT050, VUT051, VUT052, VUT053, VUT054, VUT055, VUT056, VUT057, VUT058, VUT059, VUT060, VUT061, VUT062, VUT063, VUT064, VUT065, VUT066, VUT067, VUT068, VUT069, *N. Zerega 955* (India, photo voucher at CHIC), *N. Zerega 958* (India, photo voucher at CHIC), *N. Zerega 959* (India, photo voucher at CHIC), *N. Zerega 960* (Caribbean, CHIC), *N. Zerega 961* (Caribbean, CHIC), *N. Zerega 962* (Caribbean, CHIC).

Artocarpus camansi *EG 140 (CHIC), seed offspring of NTBG 000501.005 (Papua New Guinea), NTBG 910280.001 (Pohnpei, FSM), NTBG 000389.001 (Papua New Guinea), NTBG 980212.001 (Palau), NTBG 770444.001 (Tahiti), NTBG 960576.001 (Honduras).

Artocarpus altilis* × *A. mariannensis NTBG 890174.001 (Tokelau), NTBG 890173.002 (Tokelau), NTBG 890184.001 (Yap, FSM), NTBG 790490.001 (Society Islands), NTBG 890183.001 (Palau), NTBG 910269.001 (Chuuk, FSM), NTBG 910265.001 (Society Islands).

Appendix 2. Voucher and locality information for *Artocarpus odoratissimus* collected in Sabah, Malaysia. At least one voucher was made per site, with the exception of two sites in Sandakan District for which only photographic vouchers were taken. All voucher specimens were deposited at the Chicago Botanic Garden Nancy Poole Rich Herbarium (CHIC).

District	Locality	*N*	Geographic coordinates	Collection no.	Collection date	Voucher no.
Beaufort	Beaufort Hill	3	5°20′48″N, 115°44′59.82″E	NZ 839, 841	June 23, 2013	NZ 839
				SAN 156751	May 20, 2014	SAN 155751
Beaufort	Ganui Forest Reserve	7	4°59′42.96″N, 115°41′19.86″E	NZ 879, 884–886, 888, 892, 893	June 25, 2013	NZ 884–886
Beaufort	Near Binsuluk Forest Reserve	7	5°29′36″N, 115°38′21″E (estimated)	NZ 895–901	June 26, 2013	NZ 895
Beaufort	Sianggau Forest Reserve	11	5°10′44.4″N, 115°36′26.46″E	NZ 855–857, 862, 866, 867, 870–873, 876	June 24, 2013	NZ 855, 866, 867
Beluran	Along Sungai Telupid	4	5°37′14.58″N, 117°6′12.42″E	NZ 735, 741, 742, 744	June 18, 2013	NZ 735
Papar	Kampung Kopozon	10	5°42′30″N, 116°00′59.94″E	NZ 789–791, 797, 802, 805–809	June 21, 2013	NZ 789
Ranau	Kinabalu Park, Poring Springs	14	6°2′42.48″N, 116°42′10.86″E	NZ 749–752, 755, 760, 764, 765, 768–770, 772–774	June 19, 2013	NZ 755, 769
Sandakan	Kampung Sungai Batang	1	5°56′7.9″N, 118°0′41.5″E	NZ 706	June 17, 2013	Photo only
Sandakan	Kinabatangan	1	5°30′13.2″N, 118°13′9.24″E	NZ 951	June 29, 2013	Photo only
Sandakan	Sepilok	4	117°56′27.7″N, 117°56′27.7″E	NZ 614, 704, 706, 714, 720	June 13 & 17, 2013	NZ 614, 714
Sandakan	Ulu Dusun ARS	24	5°47′25.96″N, 117°46′31.56″E	NZ 618–631, 678–685 EG 94, 131	June 14, 2013 May 15 & 29, 2014	NZ 618, EG 94
Tambunan	Kipundu Butterfly Park	8	5°52′16.2″N, 116°15′1.44″E	NZ 810, 811, 816, 817, 819–822	June 21, 2013	NZ 810
Tenom	Sabah Agriculture Park and ARS Tenom	16	5°11′11.4″N, 116°00′1.6″E	NZ 912, NZ 935–937 EG 60, 62, 102–111	June 27, 2013 May 6 & 19, 2014	NZ 912, EG 102, EG 106

Note: ARS = Agriculture Research Station; *N* = number of individuals.

APPENDIX 3. GenBank accession numbers for sequences from the in silico data set.

Accession no.	Species	AALTCP01	AALTCP02	AALTCP03	AALTCP04	AALTCP05	AALTCP06	AALTCP07	AALTCP08	AALTCP09	AALTCP10	AALTCP11	AALTCP12	AALTCP13	AALTCP14	AALTCP15
NTBG 030,042.001	*altilis*	KR185384	KR185385	KR185386	KR185387	KR185388	KR185389	KR185390	KR185391	KR185392	KR185393	KR185394	—	—	KR185395	KR185396
NTBG 040,063.001	*altilis*	KR185397	KR185398	KR185399	KR185400	KR185401	KR185402	KR185403	KR185404	KR185405	KR185406	KR185407	KR185408	—	KR185409	KR185410
NTBG 790,485.001	*altilis*	KR185411	KR185412	KR185413	KR185414	KR185415	—	KR185416	KR185417	KR185418	KR185419	KR185420	—	—	KR185421	KR185422
NTBG 790,487.001	*altilis*	KR185423	KR185424	KR185425	KR185426	KR185427	KR185428	KR185429	KR185430	KR185431	KR185432	KR185433	—	—	KR185434	KR185435
NTBG 880,690.001	*altilis*	KR185436	KR185437	KR185438	KR185439	KR185440	KR185441	KR185442	KR185443	KR185444	KR185445	KR185446	KR185447	KR185448	KR185449	KR185450
NTBG 890,167.002	*altilis*	KR185451	KR185452	KR185453	KR185454	KR185455	KR185456	KR185457	KR185458	KR185459	KR185460	KR185461	—	—	KR185462	KR185463
NTBG 890,455.001	*altilis*	KR185464	KR185465	KR185466	KR185467	KR185468	KR185469	KR185470	KR185471	KR185472	KR185473	KR185474	KR185475	—	KR185476	KR185477
NTBG 890,479.002	*altilis*	KR185478	KR185479	KR185480	KR185481	KR185482	KR185483	KR185484	KR185485	KR185486	KR185487	KR185488	—	—	KR185489	KR185490
NTBG 900,261.001	*altilis*	KR185491	KR185492	KR185493	KR185494	KR185495	KR185496	KR185497	KR185498	KR185499	KR185500	KR185501	KR185502	—	KR185503	KR185504
NTBG 900,265.001	*altilis*	KR185505	KR185506	KR185507	KR185508	KR185509	KR185510	KR185511	KR185512	KR185513	KR185514	KR185515	KR185516	—	KR185517	KR185518
NTBG 000,389.001	*camansi*	KR185534	KR185535	KR185536	KR185537	KR185538	KR185539	KR185540	KR185541	KR185542	KR185543	KR185544	KR185545	—	—	KR185546
NTBG 770,444.001	*camansi*	KR185547	KR185548	KR185549	KR185550	KR185551	KR185552	KR185553	KR185554	KR185555	—	KR185556	KR185557	—	KR185558	KR185559
NTBG 910,280.001	*camansi*	KR185560	KR185561	KR185562	KR185563	KR185564	KR185565	KR185566	KR185567	KR185568	KR185569	KR185570	KR185571	KR185572	—	—
NTBG 960,576.001	*camansi*	KR185519	KR185520	KR185521	KR185522	KR185523	KR185524	KR185525	KR185526	KR185527	KR185528	KR185529	KR185530	KR185531	KR185532	KR185533
NTBG 980,212.001	*camansi*	KR185573	KR185574	KR185575	KR185576	KR185577	KR185578	KR185579	KR185580	KR185581	KR185582	KR185583	KR185584	—	KR185585	KR185586
EG 140	*camansi*	KR185587	KR185588	KR185589	KR185590	KR185591	KR185592	KR185593	KR185594	KR185595	KR185596	KR185597	KR185598	—	KR185599	KR185600
NTBG 790,490.001	*altilis* × *mariannensis*	KR185601	KR185602	KR185603	KR185604	KR185605	KR185606	KR185607	KR185608	KR185609	KR185610	KR185611	KR185612	KR185613	KR185614	KR185615
NTBG 890,173.002	*altilis* × *mariannensis*	KR185616	KR185617	KR185618	KR185619	KR185620	KR185621	KR185622	KR185623	—	KR185624	—	KR185625	—	—	—
NTBG 890,174.001	*altilis* × *mariannensis*	KR185626	KR185627	KR185628	KR185629	KR185630	KR185631	KR185632	KR185633	KR185634	KR185635	—	KR185636	KR185637	—	—
NTBG 890,183.001	*altilis* × *mariannensis*	KR185638	KR185639	KR185640	KR185641	KR185642	KR185643	KR185644	KR185645	KR185646	KR185647	KR185648	KR185649	—	KR185650	KR185651
NTBG 890,184.001	*altilis* × *mariannensis*	KR185652	KR185653	KR185654	KR185655	KR185656	KR185657	KR185658	KR185659	KR185660	KR185661	KR185662	KR185663	—	KR185664	KR185665
NTBG 910,265.001	*altilis* × *mariannensis*	KR185666	KR185667	KR185668	KR185669	KR185670	—	KR185671	KR185672	KR185673	KR185674	KR185675	KR185676	—	KR185677	KR185678
NTBG 910,269.001	*altilis* × *mariannensis*	KR185679	KR185680	KR185681	KR185682	KR185683	KR185684	KR185685	KR185686	KR185687	KR185688	KR185689	KR185690	—	KR185691	KR185692