P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.24 No.2 April - June 2019
Page: [66-75]

Original research article

Intron Length Polymorphism in Candidate Genes for Secondary Growth and its Application in Diversity Assessment of Amazonian Accessions of *Hevea brasiliensis*

Gunlayarat Bhusudsawang and Kittipat Ukoskit*

Department of Biotechnology, Faculty of Science and Technology, Thammasat University, Pathum Thani 12120, Thailand.

Received 11 January 2019; Received in revised form 20 February 2019
Accepted 21 February 2019; Available online 23 May 2019

ABSTRACT

Gene-specific markers are important tools in genetic analysis, allowing direct estimation of functional diversity. The goal of this study was to develop intron length polymorphism (ILP) markers from candidate genes associated with the secondary growth related traits that underpin the economics of rubber tree cultivation. We performed a BLAST analysis of *Eucalyptus* spp. expressed sequence tag (EST) sequences for 23 candidate genes involved in secondary growth. From the total 142 shotgun contig hits, 531 introns were identified and 23 polymorphic ILP markers representing 23 candidate genes were selected for diversity assessment of 170 Amazonian accessions. The markers yielded 140 alleles with an average of 6.1 alleles per locus. Polymorphism information content (PIC) values ranged from 0.40 to 0.89, with an average of 0.64, indicating a high level of polymorphism in the markers. Analyses of population structure confirmed that the Amazonian accessions fell into two subpopulations grouped by catchment area, supporting previous reports. Among the 23 ILP markers analyzed, 11 revealed private alleles across Amazonian populations. The ILP markers developed in this study provide an immediate resource for the study of genetic diversity and establishment of marker-trait association for secondary growth traits in rubber trees.

Keywords: Intron length polymorphism; Secondary growth; Girth; Rubber tree; Diversity.

***Corresponding author:** ku@tu.ac.th **doi:** 10.14456/scitechasia.2019.15

1. Introduction

To improve productivity in rubber tree (*Hevea brasiliensis*) plantations, breeding programs attempt to exploit genetic variation within a population, to select for desired characteristics. In 1981, therefore, the International Rubber Research and Development Board (IRRDB) collected wild Amazonian rubber tree germplasm from several districts in the western states of Brazil [1]. This collection was initially planted in Malaysia and the Ivory Coast, then distributed more widely as the germplasm for rubber tree breeding across Southeast Asia [2]. To make full use of this germplasm in breeding programs, an understanding is needed of the genetic diversity among the acquisitions from different areas of the Amazon.

Secondary tree growth has been extensively studied, and several of the genes involved have been identified. Candidate genes involved in lignin biosynthesis and cell wall structure were reported for Loblolly Pine [3]. Transcriptome profiles of xylems and phloems isolated from the root-hypocotyl of Arabidosis demonstrated their role in the radial pattern observed during secondary growth. Many transcription factors are thought to play roles as regulators of xylem or phloem activity, cell differentiation, and cell wall development [4].

Introns are noncoding sequences that are widespread and abundant in eukaryotic genomes. Introns are more variable than exons, because selective pressure in intronic regions is much weaker than that in exonic regions [5]. The polymorphic length of introns has been investigated as a genetic marker, having the advantages of specificity and co-dominance, and offering reliable and convenient detection [6, 7]. Intron length polymorphism (ILP) is a robust marker derived from a large volume of genomic and transcriptome sequence data. By comparing EST sequences with genomic sequences, the intron position can be identified, allowing the development of ILP markers located in the flanking intron region. The primers are designed from the conserved nucleotides of the two exons flanking the intron region. The PCR products are then amplified across intron regions. The results reveal the polymorphism in intron size [6].

The objective of the current study was to develop gene- specific markers from candidate genes involved in secondary growth of the rubber tree. Twenty- three candidate genes were selected for their involvement in cellulose and lignin biosynthesis, cell expansion, and the expression of transcription factors regulating secondary cell wall development. To design ILP markers, the sequences of candidate genes were aligned with rubber tree genomic sequences. The selected polymorphic ILP markers were tested of their potential applications for evaluating the genetic diversity and population structure of the Amazonian accessions. This allows a better understanding to be gained of the germplasm resources available for use in rubber tree breeding programs.

2. Materials and Methods

2.1 Plant materials

The Rubber Research Institute of Thailand introduced Amazonian rubber tree germplasm from collections in Malaysia and the Ivory Coast. The Amazonian collection used in the present study comprised 170 accessions from three Brazilian states [8]: 14 from Acre, 64 from Mato Grosso, 91 from Rondonia and one of unknown origin (Table 1). The trees were planted at the NongKhai Rubber Research Center (NKRR) of NongKhai Province, in the northeast region of Thailand. Genomic DNA was extracted from 500 mg of leaf material using the cetyltrimethylammonium bromide (CTAB) method of Gawel and Jarret [9].

Table 1. Amazonian accessions by geographical origin.

State	District	Code	No. of accessio
Acre (AC)	Assis-Brasil	AC/AB	1
	Feijo	AC/F	2
	Tarauaca	AC/T	5
	Sena	AC/S	6
Mato (MT)	Aracatuba	MT/A	1
	Cartriquacu	MT/C	19
	Itanba	MT/IT	38
	Vila Bela	MT/VB	6
Rondonia (RO)	Ariquemes	RO/A	20
	Calama	RO/C	4
	Costa arques	RO/CM	19
	Jaru	RO/J	30
	Ji-Parana	RO/JP	1
	Ouro Preto	RO/OP	1
	Pimenta	RO/PB	16
Other	Unknown	BRAZIL UN	1
Total			170

2.2 Development of ILP markers

Genes were selected using the available sequences in the public NCBI/gene database from four function classes: four from cellulose synthesis, eight from lignin synthesis, twelve associated with the transcription factors regulating secondary cell wall development, and three from cell expansion. A total of 36 *Eucalyptus* spp. ESTs from the 27 selected genes were aligned with whole-genome shotgun contigs of the rubber tree to predict the position and number of the introns. Introns of a manageable size (<800 bp) were extracted by aligning gene transcripts with their corresponding genome sequences using BLASTN program search of the whole-genome shotgun sequences in the NCBI database (WGS Project) of rubber tree cultivar RRIM 600 [10]. To avoid primers crossing the predicted exon/intron junctions, sequences of approximately 20 bp at the 5' and 3' ends of the predicted exon/intron junctions were excluded from the design of the intron-flanking primers. A pair of specific primers flanking each of the predicted intron positions was designed using the Primer3 program [11] with the default settings. Amplification by PCR was performed in a 20 µl reaction mixture containing 10 ng template DNA, 1×PCR buffer (20 mM Tris pH 9.0, 100 mM KCl, 3.0 mM MgCl2), 200 µM of each of the four dNTPs, 1.5 µM of each of the forward and reverse primers, and 0.5 U Taq DNA polymerase. The following PCR conditions were used: 94 ºC for 3 min, followed by 30 cycles of 94 ºC for 30 s, 55 ºC for 30 s, 72 ºC for 30 s, and 5 min at 72 ºC for the final extension. PCR products were separated on 6% denaturing polyacrylamide gels and silver-stained, following Benbouza, Jacquemin [12].

2.3 Data analysis

The genetic diversity parameters were determined from the genotypic profiles of the markers, using the PowerMarker 3.25 [13] and POPGENE 1.32 [14] software packages. The parameters calculated included the number of genotypes (Ng), number of alleles (Na), effective number of alleles (Ne), expected heterozygosity (He or gene diversity), observed heterozygosity (Ho); Shannon's information index (I; [15]), polymorphism information content (PIC; [16]), and inbreeding coefficient (f). Deviations from the Hardy-Weinberg equilibrium were assessed at each locus across all populations and their significance was evaluated using a chi-square test. A principal component analysis (PCA) was performed using NTSYS-pc v. 2.0 [17]. The population structure was assessed using model-based Bayesian analysis in the software STRUCTURE v. 2.3.4 [18]. The analysis was run ten times for each K value from K =1 to K = 10. The burn-in length was set to 100,000, and number of iterations to 100,000. The ad hoc statistics defined by [19] as Δ K. were used. Based on the posterior probability of membership (Q) of a given accession, the accession was classified as

admixture with a membership probability of $Q < 0.70$. Analysis of molecular variance (AMOVA) and calculation of Fst among groups were performed using ARLEQUIN 3.5 [20] with 100,000 permutations, for both the geographical location of the collection site and the cluster groups suggested by our Bayesian analysis. Gene flow across populations was estimated using an indirect method based on the number of migrants per generation (Nm) [21]. The number of different alleles per locus, number of effective alleles per locus, number of private alleles, observed heterozygosity, expected heterozygosity [22], and Shannon's Information Index were computed for each population using GenAlEx version 6.501 [23].

3. Results and Discussion

3.1 ILP marker development

Of total ESTs aligned to the reference sequences, 531 were identified. Of the total predicted introns, 405 had flanking exon sequences sufficiently long to allow the design of ILP primers and 237 predicted introns with sizes smaller than 800 bp were selected for primer synthesis and used in the evaluation of length polymorphism. Of the 237 ILP primer pairs examined, 230 pairs (97.05%) successfully amplified genomic DNA samples.

After screening for polymorphism using one accession from Acre, six from Rondonia, and three from Mato Grosso, a total of 23 polymorphic ILP markers representing 23 genes were selected for a population genetic analysis. The characteristics of the selected polymorphic 23 ILP markers are presented in Table 2. The 23 polymorphic ILP markers yielded 140 alleles, with an average of 6.09 alleles per locus. The number of alleles ranged from 3 to 13 and the number of genotypes from 4 to 55. PIC values were between 0.40 and 0.89, with an average of 0.64. One private allele was found in eight markers, and two in three markers, including Per173, NtLIM168, and MOR211. The observed heterozygosity ranged from 0.14 to 0.99 and the expected heterozygosity from 0.44 to 0.90, with average values of 0.55 and 0.68, respectively. This difference in average values suggested heterozygote deficiency. Departure from Hardy–Weinberg equilibrium was significant at $p < 0.001$ for all loci.

The goal of the present study was to develop ILP markers from candidate rubber tree genes involved in secondary growth. A total of 36 *Eucalyptus* spp. ESTs for the genes under study were used as queries, and the whole-genome shotgun sequences of the rubber tree were searched for prediction of intron regions. The number and density of introns varies significantly between organisms, due to differences in the rate of lineage-specific intron loss and/or gain [24]. Factors such as insertion of transposable elements [25] and the frequency and size of deletion events [26] may produce changes in intron size.

An ideal molecular marker should be reliable and polymorphic, quick and simple to apply [27], and inexpensive and functional. The 23 selected polymorphic ILP markers representing 23 genes were applied to a genetic analysis of the Amazonian population. The polymorphisms detected by these ILP markers ranged from informative to highly informative. The average PIC value generated by the ILP markers was 0.64 and the average number of alleles per locus 6.09, which is superior to previously reported ILP markers [28] and EST-SSR markers [29, 30], and comparable to the EST-SSR markers reported by Cubry et al. [31]. The range of polymorphism detected in these EST markers can be attributed in part to the level of diversity across the germplasm surveyed. PIC values greater than 0.5 have been suggested to indicate an informative marker, and loci with values greater than 0.7 as being appropriate for genetic mapping [32].

Table 2. Characteristics of 23 polymorphic ILP markers developed from candidate secondary growth genes.

Marker	Gene	Ng	Na	Pa	Ne	He	Ho	PIC	f*	Forward primer	Reward primer	Tm (°C)	Expected Size (bp)
CesA016	Cellulose synthase2	12	6	1	3.01	0.67	0.37	0.61	0.45	taggcttgataggaggttca	agaatgatcgatatgccaac	55	292
CesA026_1	Cellulose synthase2	11	5	0	2.73	0.63	0.31	0.6	0.51	ttcatcatgccagaggtaat	gcaatgaagaaaagacttgc	55	317
COBL4057	COBRA-like protein 4	21	7	0	4.7	0.79	0.28	0.76	0.65	aatgggatgtaatgtcttgg	ggtacatttggaagttgctc	55	390
CaS106	Callose synthase 10	36	10	1	5.59	0.82	0.72	0.8	0.12	gcaacttcagttgtgttcag	gacggactacaaagcctcta	54.7	272
4CL069	4-coumarate:CoA ligase	15	6	1	4.43	0.77	0.68	0.74	0.13	ctagcacatcctgacatatcc	cctttggaatggattctatg	54.8	385
CAD072	cinnamyl alcohol dehydrogenase	11	5	1	3.07	0.67	0.67	0.62	0.01	acatcccctctggtattttt	agttgcaacccatgtaaatc	55.1	291
COMT084	caffeic O-methylransferase1	18	9	1	3.51	0.72	0.52	0.68	0.27	gccattttcatgaaggtaag	attttaagcagtgctcatcg	55.7	167
CCoAMT090_2	caffeoyl-CoA 3-O-methyltransferase	10	6	1	2.7	0.63	0.59	0.58	0.06	agaagcatgctcaagaattg	gaattgacagcaaatcatcc	55.7	196
CCR110	cinnamoyl-CoA reductase	9	5	0	2.81	0.64	0.69	0.59	-0.07	gtgatggagttttccacact	agctgctatgatcacgtttt	55.2	534
C4H113	cinnamate 4-hydroxylase 1	15	5	0	3.24	0.69	0.38	0.65	0.45	gttgtgcttgaaagtttcct	aagatctgcaaggaagtgaa	55.1	283
PAL116	phenylalanine ammonia-lyase	6	4	0	2.28	0.56	0.14	0.52	0.76	ctcttgttgcagagtgtgac	cagagggagctcattaggta	54.6	844
Per173	peroxidase 2	40	12	2	7.38	0.86	0.46	0.85	0.46	agtcttttcccctatgaagg	catctttgctcagattgcat	56.4	191
LIM118	proteins Lin11, Isl-1 & Mec-3 transcription factor	10	4	0	3.14	0.68	0.5	0.62	0.27	gagttggatatggtgatgct	aggcaccagagaaaaatgt	55	266
MYB126_1	MYB transcription factor 1	7	4	0	2.74	0.63	0.8	0.57	-0.26	tgttcttcctggtaatctcc	tcattatcaaactgcacagc	54.8	174
HDzip134_2	Homeobox-leucine zipper protein	7	4	0	3.09	0.55	0.86	0.61	0.24	tcctccaaggaagaagagat	ggtgaggacactctttgaac	55.5	165
NAC149	NAC domain-containing protein 100-like	4	4	1	2.41	0.59	0.99	0.5	-0.69	ctcctaaaggagagaaaacca	tcttgtggaaaaccctacac	55.4	349
APL155_2	myb family transcription factor APL	6	3	0	2.19	0.54	0.41	0.48	0.24	ttacccttaccacctcaaa	gttgctttccaagtctgaat	54.8	219
KNAT161	homeobox protein knotted-1-like	10	4	0	2.98	0.66	0.41	0.59	0.38	caagtttggatggttcagat	tcctctctgatgtcagcaat	55.8	299
NtLIM168	Eglim1 gene for transcription factor lim1	7	6	2	1.8	0.44	0.28	0.4	0.36	gatcaaatcttcaagcgaac	acggggtcaagtgaatataa	55	222
MOR211	protein MOR1	55	13	2	9.67	0.9	0.72	0.89	0.2	cctttctttctgaccacttg	actcctttggagaaatctgg	55.9	197
PIN222	auxin efflux carrier component 1-like	11	5	0	4.12	0.76	0.59	0.72	0.22	cctgccattatagcaaaatc	gcaaagccatgaacagac	55	440
COB185	COBRA-like protein (COBL4)	26	9	1	4.71	0.79	0.39	0.76	0.51	gcatatgacccattagatcc	ggtacatttggaagttgctc	54.8	430
KORR194	korrigan	4	4	0	2.34	0.57	0.96	0.48	-0.67	catactgaagtggctgaggt	acaaccagaccagcataatc	55.3	209
Mean		15.3	6.09	0.61	0.68	0.68	0.55	0.64					

*Departure from Hardy–Weinberg equilibrium was significant at $p < 0.001$ for all loci. Ng = number of genotypes; Na = number of alleles; Pa = Number of private alleles; Ne = Effective number of alleles; He = Expected heterozygosity; Ho = Observed heterozygosity; PIC = Polymorphism information content; f = inbreeding coefficient

In the current study, 20 of the ILP markers showed PIC values greater than 0.5, and seven greater than 0.7, suggesting high levels of polymorphism. These ILP markers for genes associated with secondary growth in the rubber tree provide an immediate resource for the study of genetic diversity, the establishment of marker-trait association, and the identification of genes that regulate secondary growth traits. This will help accelerate marker-assisted selection in rubber tree breeding programs.

3.2 Genetic structure of the Amazonian population

Two methods were applied: model-based Bayesian analysis and PCA. The model-based Bayesian analysis, combined with computation of the Evanno ΔK statistics (Fig 1), suggested two clusters (K=2) with admixed accessions. Cluster 1 comprised 13 accessions from Acre, one from Mato Grosso, and 63 from Rondonia. Cluster 2 comprised 50 accessions from Mato Grosso and two from Rondonia. Forty accessions had a membership probability (Q value) below 0.7, and were defined as admixtures. Of these, one originated from Acre, 13 from Mato Grosso, and 26 from Rondonia (Table 3). The accession with unknown origin (BRAZIL_UN) was placed into Cluster 2. The results were in agreement with those from the PCA (Fig. 2), in which the first two principal components identified two clusters and admixed accessions from Rondonia.

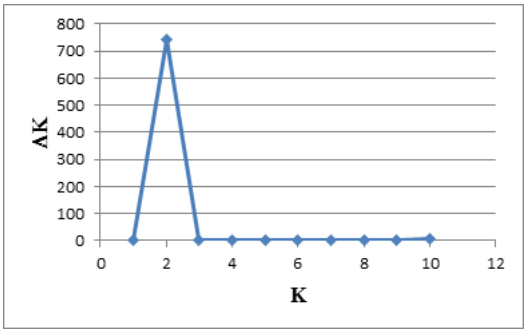

Fig. 1. The genetic structure of 170 Amazonian accessions inferred through Bayesian analysis of the most probable number of clusters (K), estimated using ad-hoc ΔK statistics.

Table 3. Division into two clusters (K = 2) from the model-based Bayesian analysis.

Cluster (C)	Number	C1	C 2	Admixture[1]
Acre	14	13	0	1
Mato Grosso	64	1	50	13
Rondonia	91	63	2	26
BRAZIL_UN	1	0	1	0
Total	170	77	53	40

[1]Accessions with a membership probability (Q value) below 0.7 were classified as Admixture

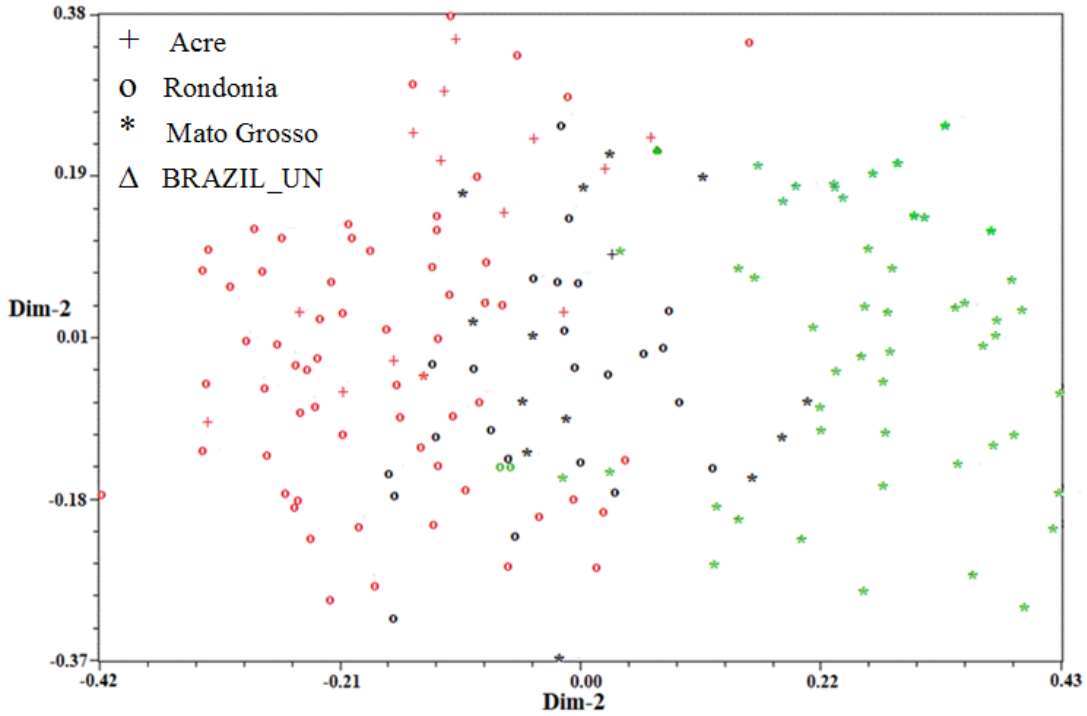

Fig. 2. Principal component analysis (PCA) of 170 Amazonian accessions based on 23 ILP markers. Accessions shown in red were assigned to Cluster 1, those in green to Cluster2, and those in black as admixtures, corresponding to Bayesian analysis based on K = 2.

As a further test of their potential applications, the 23 polymorphic ILP markers were used to evaluate the genetic diversity and population structure of the 170 Amazonian accessions. The population was divided into two clusters. Analyses of pairwise genetic distance suggested that the Mato Grosso population was genetically more distant than the Acre and Rondonia populations. This can be partly explained by the hydrographical network of the Amazonian region [33-36]. The relatively low Fst value suggested little differentiation between populations. This might be attributed to gene flows, reflected in the high rate of migration. The results reported in previous studies suggest a history of gene flows between populations, possibly due to the outcrossing mode of reproduction of these species and the dispersal of seeds through the network of rivers in the Amazon basin [8, 35-36]. Our results confirmed the conclusions of these previous studies that Amazonian accessions are grouped by catchment area. The substantial admixture observed in the Rondonia population was attributed to its borders with the Mato Grosso and Acre regions. Genetic differentiation occurs, but is attenuated by the highly connected nature of the Amazon basin, and the associated gene flow

3.3 Diversity and differentiation among population

Diversity statistics across all 170 accessions and within each of the three populations based on geographical origins (Table 4) were computed. The number of different alleles per ILP locus ranged from 4.30 (Acre) to 5.87 (Rondonia) with an average of 5.19. The observed heterozygosity (Ho) ranged from

Table 4. Genetic diversity within each population assessed using 23 ILP markers.

Population	N	Na	He	Ho	I	F	Pa
Acre	14	4.30	0.60	0.53	1.10	0.12	0
Rondonia	91	5.87	0.68	0.58	1.35	0.14	11
Mato Grosso	65	5.39	0.64	0.52	1.24	0.19	2
Overall population	170	5.19	0.64	0.54	1.23	0.15	13

N = Sample Size; Na = Mean number of alleles per marker; He = Expected heterozygosity;
Ho = Observed heterozygosity; I = Shannon's Information index;
F = Fixation index; Pa = Number of private alleles

0.52 (Mato Grosso) to 0.58 (Rondonia), with a mean of 0.54. The genetic diversity (He) was higher than Ho in all cases, with values ranging from 0.60 (Acre) to 0.68 (Rondonia), and a mean of 0.64. Wright's fixation index (F) values were positive, with a mean of 0.15 across the three populations. No private allele was found in the Acre accessions, whereas those from Rondonia had 11 private alleles and those from Mato Grosso had two. For Per173 and MOR211, one private allele was identified in each of the Mato Grosso and Rondonia populations. NtLIM168 had two private alleles in the Rondonia accessions. CesA016, CaS106, 4CL069, CAD072, COMT084, CCoAMT090_2, NAC149, and COB185 had one private allele in the Rondonia accessions. These private alleles were always found at low frequencies (< 4%) The genetic differentiation among populations (F_{ST} = 0.062 at P< 0.001) was significant as indicated by the permutation test. The pairwise genetic distances between geographical regions were 0.12 for Mato Grosso–Acre, 0.28 for Mato Grosso–Rondonia, and 0.14 for Acre–Rondonia. These results supported the population structure that grouped Acre and Rondonia in the same cluster.

The average estimated gene flow among populations (Nm) was 3.52. A hierarchical analysis of variance revealed statistically significant values (P<0.001). Between-cluster variation accounted for 6.04% of the total genetic variance, whereas within-cluster variation accounted for 93.96%.

An assessment using the ILP markers revealed high levels of genetic diversity between the Amazonian accessions, which is consistent with previous reports [34, 35]. An average Ho of 0.54 was detected across all accessions. The estimated genetic variability (heterozygosity) was lower than that reported in previous studies using SSR markers [34, 35], a result that could be anticipated from the specific molecular nature of the two marker types. The number of different alleles per ILP locus identified across the three populations was proportional to the sample size, and some of the differences seen here may be attributed to sampling differences. Of the 23 ILP markers analyzed, 11 revealed private alleles. The presence of unique alleles is a measure of the genetic individuality of plant populations, which reflects genetic adaptation to ecological conditions. The ILP markers displaying private alleles therefore constitute valuable markers for genetic diversity and DNA fingerprinting, and can be used to guide selection of parents in a breeding program, as discussed by Chen et al.[37].

4. Conclusion

Polymorphic ILP markers were developed from candidate genes involved in secondary growth of rubber trees. These were then used to investigate the genetic diversity and population structure of the Amazonian population. The results suggested that the intronic regions were sufficiently varied to be used as marker resources for genetic analyses of the rubber tree. These markers will support the

construction of genetic linkage maps for identification of QTL and candidate gene association mapping. The ability to identify markers associated with secondary growth related traits will allow selection for such traits at the seedling stage, accelerating the progress of rubber tree breeding programs.

Acknowledgements

This work was supported by the Faculty of Science and Technology, Thammasat University and the Rubber Research Institute of Thailand, Rubber Authority of Thailand.

References

[1] Onokpise OU. Natural rubber, *Hevea brasiliensis* (Willd. ex a. Juss) Mull Arg, Germplasm collection in the Amazon basin, Brazil: A retrospective. Econ Bot 2004;58:544-55.

[2] Sedgley M, Attanayake DPSTG. The breeding system of rubber (*Hevea brasiliensis*): an evaluation of controlled hand pollination methods. Euphytica 1988;39:83-91.

[3] Brown GR, Bassoni DL, Gill GP, Fontana JR, Wheeler NC, Megraw RA. Identification of Quantitative Trai Loci Influencing Wood Property Traits in Loblolly Pine (*Pinus taeda* L.). III. QTL Verification and Candidate Gene Mapping. Genetics 2003;164:1537-46.

[4] Zhao C, Craig JC, Petzold HE, Dickerman AW, Beers EP. The Xylem and Phloem Transcriptiome from Secondary Tissues of the Arabiodopsis Root-Hypocotyl. Plant Physiol 2005;138:803-18.

[5] Choi H-K, Kim D, Uhm T, Limpens E, Lim H, Mun JH, et al. A sequencebased genetic map of medicago truncatula and comparison of marker colinearity with *M. sativa*. Genetics 2004;166:1463-502.

[6] Yang L, Jin G, Zhao X, Zheng Y, Xu Z, Wu W. PIP: a database of potential intron polymorphism markers. Bioinformatics 2007;23(16):2174-7.

[7] Saisug W, Ukoskit K. Comparative analysis of EST- derived markers for allelic variation in *Jatropha curcas* L. and cross transferability among economically important species of Euphorbiaceae. Genes Genom 2013;35:1-12.

[8] Chanroj V, Rattanawong R, Phumichai T, Tangphasornruang S, Ukoskit K. Genome- wide association mapping of latex yield and girth in Amazonian accessions of *Hevea brasiliensis* grown in a suboptimal climate zone. Genomics 2017;109:475-84.

[9] Gawel NJ, Jarret RL. A modified CTAB DNA extraction procedure for *Musa* and *Ipomoea*. Plant Mol Biol Rep 1991;9(3):262-6.

[10] Rahman AYA, Usharraj AO, Misra BB, Thottathil GP, Jayasekaran K, Feng Y, et al. Draft genome sequence of the rubber tree *Hevea brasiliensis*. BMC Genomics 2013;14:75.

[11] Rozen S, Skaletsky HJ. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

[12] Benbouza H, Jacquemin J-M, Baudoin J-P, Mergeai G. Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels. Biotechnol Agron Soc Environ 2006;10:77-81.

[13] Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 2005;21(9):2128-9.

[14] Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX, Yang R, et al. PopGene: the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta Canada, 1997.

[15] Lewontin R. The Apportionment of Human Diversity. Evol Biol 1972;6:381-98.

[16] Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 1980;32(3):314-31.

[17] Rohlf FJ. NTSYSpc: Numerical Taxonomy and Multivariate Analysis System Version 2.0. N Y1998.

[18] Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in

structured population. Am J Hum Genet 2000;67(1):170-81.

[19] Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 2005;14(8):2611-20.

[20] Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 2010;10(3):564-7.

[21] Slatkin M, Barton NH. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 1989;43(7):1349-68.

[22] Nei M, Li W-H. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A 1979;76(10):5269-73.

[23] Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 2012;28(19):2537-9.

[24] Wang H, Devos KM, Bennetzen JL. Recurrent loss of specific introns during angiosperm evolution. PLos Genetics 2014;10(12):e1004843.

[25] Bartolomé C, Maside X, Charlesworth B. On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. Mol Biol Evol 2002;19(6):926-37.

[26] Petrov DA, Sangster TA, Johnston JS, Hartl DL, Shaw KL. Evidence for DNA loss as a determinant of genome size. Science 2000;287(5455):1060-2.

[27] Agarwal M, Shrivastava N, Padh H. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 2008;27:617-31.

[28] Li D, Xia Z, Deng Z, Liu X, Dong J, Feng F. Development and characterization of intron-flanking EST-PCR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). Mol Biotechnol 2012;51(2):148-59.

[29] Feng SP, Li WG, Huang HS, Wang JY, Wu YT. Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). Mol Breed 2009;23(1):85-97.

[30] Perseguini JMKC, Romão LRdC, Briñez B, Junior EJS, Gonçalves PdS, Benchimol LL. Genetic diversity of cultivated accessions and wild species of rubber tree using EST-SSR markers. Pesqui Agropecu Bras 2012;47(8).

[31] Cubry P, Pujade-Renaud V, Garcia D, Espeout S, Le Guen V, Granet F, et al. Development and characterization of a new set of 164 polymorphic EST-SSR markers for diversity and breeding studies in rubber tree (*Hevea brasiliensis*Müll. Arg.). Plant Breeding 2014;133(3):419-26.

[32] Liu W, Jia X, Liu Z, Zhang Z, Wang Y, Liu Z, et al. Development and Characterization of Transcription Factor Gene-Derived Microsatellite (TFGM) Markers in *Medicago truncatula* and Their Transferability in Leguminous and Non-Leguminous Species. Molecules 2015;20(5):8759-71.

[33] Chevallier MH. Genetic variability of *Hevea brasiliensis* germplasm using isozyme markers. Journal of natural rubber research 1988;3:42-53.

[34] deSouza LM, Guen VL, Cerqueira-Silva CBM, Silva CC, Mantello CC, Conson ARO, et al. Genetic diversity strategy for the management and use of rubber genetic resources: more than 1,000 wild and cultivated accessions in a 100-genotype core collection. PLoS One 2015;10:e0134607.

[35] LeGuen V, Doare F, Weber C, Seguin M. Genetic structure of Amazonian populations of *Hevea brasiliensis* is shaped by hydrographical network and isolation by distance. Tree Genet Genomes 2009;5(4):673-83.

[36] Lam LV, Thanh T, Chi VT, Tuy LM. Genetic diversity of Hevea IRRDB'81 collection assessed by RAPD markers. Mol Biot 2009;42(3):292-8.

[37] Chen S, Nelson MN, Ghamkhar K, Fu T, Cowling WA. Divergent patterns of allelic diversity from similar origins: the case of oilseed rape (*Brassica napus* L.) in China and Australia. Genome 2008;51(1):1-10.