Detection of Marker Trait Associations: A Step Towards Marker Assisted Selection in Coconut (*Cocos nucifera* L.)

S.A.C.N. Perera*, H.M.N.B. Herath, K.N.S. Perera and W.B.S. Fernando

Genetics and Plant Breeding Division
Coconut Research Institute of Sri Lanka, Lunuwila, 61150

** Corresponding author:*
chandrikaperera2003@yahoo.com

ABSTRACT

Coconut is the most important perennial palm in Sri Lanka. Genetic improvement of the coconut palm for desirable traits is important to increase and sustain the productivity amidst various environmental stresses. Marker Assisted Selection (MAS) is an appropriate option to overcome the challenges in conventional coconut breeding strategies. However, the use of MAS is restricted in tropical tree crops mainly due to the difficulty in developing segregating populations for QTL analysis. The aim of the present investigation was to study an available F_2 population of coconuts for phenotypic and genotypic segregation and to search for marker trait associations. The F_2 population consisting of 106 individuals were scored for 15 traits representing stem, leaf and nut morphologies and were genotyped at six SSR marker loci. Both phenotypic and genotypic data revealed the F_2 population to be sufficiently segregating for QTL analysis. Single marker analysis was performed in SAS to establish marker trait associations. The association of six QTL with SSR alleles were confirmed at $P<0.01$ level for stem girth at base, and leaf characters; petiole width, number of leaflets and length and width of leaflets. The information generated can be directly used in selecting parents for genetic improvement of coconut for desirable traits. This research marks a significant step as the first report on MAS in a perennial tree crop in Sri Lanka.

Keywords: *Coconut, Marker assisted selection, Phenotyping, Segregating population, SSR markers, Tree crop*

INTRODUCTION

Coconut is the most important perennial palm in Sri Lanka. It occupies the second largest extent of cultivated lands in Sri Lanka. The existence of the palm is more than justified by the numerous and diverse uses it offers, thousands of livelihoods the palm provides and its potential for poverty alleviation in rural communities of the tropics. However, the sustainable and profitable cultivation of this socially and economically important perennial plant is challenged with several factors including the global problem of changing climate, newly emerging biotic and abiotic stresses and the resulting low productivity.

Genetic improvement of the coconut palm for higher productivity and tolerance/resistance to environmental stresses has been identified as the most viable and environmentally sustainable option to overcome these problems. Yet, the improvement of traits by conventional breeding is challenging and time consuming due to its perennial nature, predominant cross-pollinating breeding behavior resulting in heterogenous populations, massive stature of the palm and the lack of a viable method for vegetative propagation.

Molecular biological tools have now been developed to the extent of providing practical options to complement the conventional plant breeding methods even in tropical tree crops such as coconut. Molecular markers such as RAPDs, AFLPs and SSRs have been used over the years to determine the genetic diversity of local as well as global coconut germplasm (Everard, 1996; Lebrun *et al*, 1998; Perera *et al*, 1999; Perera *et al*, 2000). Even a low density QTL map of coconut has been developed using AFLP and SSR markers (Lebrun *et al*, 2001). However, efforts towards marker assisted selection (MAS) in coconut still remains in its infancy.

A major problem in linkage mapping in coconut is the difficulty and the long duration needed for developing suitable segregating populations due to its long juvenile period (Perera, 2010). The lack of populations to tag QTL governing economically important traits has prevented the widespread use of MAS in coconut improvement. Therefore, the current investigation was undertaken with the objective of studying an available F_2 population of coconut for phenotypic and genotypic segregation and to search for marker trait associations for certain traits.

MATERIALS AND METHODS

One hundred and six individuals of an F_2 population comprising of several full sib and half-sib families of coconut were subjected for the assessment. This population had been derived from a cross between Sri Lanka Tall and Sri Lanka Green Dwarf parents. The palms were 32 years of age and planted in a Complete Randomized Design at Bandirippuwa Estate, Lunuwila (IL_{1a}) in Sri Lanka.

Phenotyping F_2 palms

Phenotypic data related to stem, leaf (measured in the most mature green leaf), fruit and nuts were scored in each individual F_2 palm (Table 1).

Genotyping the F_2 palms

Genomic DNA of the 106 F_2 individuals was extracted by the protocol explained by Doyle and Doyle (1990). Extracted DNA was amplified at six SSR marker loci developed by Perera *et al.* (1999) (Table 2) by the polymerase chain reaction (PCR). PCR was performed in a final volume of 10 µL which contained a mixture of 2 µL of (20 ng/µL) template DNA, 1 x *Taq* PCR green buffer containing 2 mM $MgCl_2$, 1.25 U of *Taq* DNA polymerase (Dream *Taq*- Fermentas), deoxynucleoside triphosphates (0.35 mM each; Geneshun Biotech) and 0.6 µM primer pair (1st BASE). PCR cycles consisted of 4 minutes initial denaturation at 94 °C followed by 35 cycles of 30 seconds at 94 °C, 30 seconds at annealing temperature (depending on primer) and 60 seconds at 72 °C, and final extension at 72 °C for 5 minutes. PCR products were subjected to 6% denaturing Polyacrylamide gel electrophoresis followed by Silver staining (Anolles and Petter 1994) to visualize amplified fragments of DNA.

Table 1. Phenotypic characters scored in every individual of the F_2 coconut palm population (all the length parameters were scored in cm except for height which was recorded in meters)

Stem parameters	Leaf parameters	Fruit and nut parameters
	Total number of leaves (NuoLvs)	Total number of bunches (NuoB)
Stem height (SH)	Rachis length (RL)	Average number of nuts per bunch (NPB)
Stem girth at 20cm above ground (G20)	Petiole length (PtL)	Average weight of fresh mature nuts (FNW)
Stem girth at 150 cm above ground (G150)	Petiole width (PtW)	Fruit polar circumference (PFC)
	Number of leaflets (NuoLlts)	Fruit equatorial circumference (EFC)
	Length of leaflets (LL)	
	Width of leaflets (LW)	

Table 2. Sequences of microsatellite primers used for genotyping individuals of the F_2 coconut population

Oligo Name		Sequence	Size (bp)	Annealing Temperature
CAC37	F	5'- CAAGACAAATCAGTTTCAAGTC -3'	161-183	52 °C
	R	5'- CTTTATTCCCTTGTCACTGG -3'		
CAC38	F	5'- ACCCTACTTCTAACTGTTCACTC -3'	160-176	56 °C
	R	5'- CAGCTTGATAAATATCATCCAT -3'		
CAC43	F	5'- TGCATGTTTCAGTACTTATTTTG -3'	193-201	55 °C
	R	5'- AAAAAGCCAGCACTACTCC -3'		
CAC55	F	5'- CTATGCATCAAAACAAGGAG -3'	194-204	55 °C
	R	5'- CAAATCTAGGTAGCTTGAGGT -3'		
CAC56	F	5'- ATTCTTTTGGCTTAAAACATG -3'	161-177	55 °C
	R	5'- TGATTTTACAGTTACAAGTTTGG -3'		
CAC72	F	5'- TCACATTATCAAATAAGTCTCACA -3'	90-108	54 °C
	R	5'- GCTCTCTTTCTCATGCACA -3'		

Analysis of data

Descriptive statistics were derived for the phenotypic data in MINITAB version 14. Summary statistics were derived for genotypic data and marker trait associations were determined by single marker analysis by classifying marker haplotypes and testing the null hypothesis of no association between each marker and trait by General Linear Models Procedure in SAS version 8.0

RESULTS AND DISCUSSION

Descriptive statistics derived for the morphological traits scored in the F_2 population are presented in Table 3.

Descriptive statistics; range, standard deviation and the variances indicated a wide spread from the means for all the traits scored indicating a higher within population variation of the measured quantitative parameters (Table 3). This variation among the individuals of the population is a pre-requisite for tagging QTL and thus the studied F_2 population was identified to have a satisfactory level of phenotypic segregation for tagging QTL.

Genotyping

Five out of the six SSR loci scored polymorphic alleles in the F_2 population while the SSR locus CAC72 was revealed to be heterozygous and monomorphic. Five polymorphic SSR loci scored a total of 19 alleles ranging from 3 to 5 alleles per marker. The observed polymorphisms and the genotypic information provided evidence for sufficient genotypic segregation of the F_2 population for QTL analysis.

QTL analysis and marker trait associations

QTL analysis was carried out by single marker analysis method by classifying the marker genotypes and testing the null hypothesis of no marker trait association by general linear models procedure. Single marker analysis calculates whether phenotype values differ among genotypes for a given molecular marker and significant differences in the F statistics suggest that the marker genotype and phenotype are associated. In this analysis, the presence of a QTL was considered to be indicated at $P<0.05$ while the presence of a QTL was confirmed at $P<0.01$ (Table 4).

The effects of the presence of each allele contributing to the detected QTL were calculated, using the overall mean value of the population and the means of the two samples possessing and lacking the particular allele having QTL effect (Table 4). Although this method does not provide the exact QTL effect quantitatively, it provides sufficient evidence to determine whether the QTL effect is positive or negative on the trait.

In total 18 QTL having either negative or positive effects on the phenotype were detected for 12 traits studied in the present study. More than one QTL was detected for traits LW, NuoLlts, G20 and PtW recording the presence of 3, 2, 2 and 3 QTL respectively for each trait.

Out of them QTL having positive effects on the phenotype were confirmed for traits G20 at marker allele CAC37E, PtW and LW at SSR allele CAC55C, NuoLlts at CAC56B while QTL having negative effect on the phenotype were confirmed for traits LL at allele CAC56C and PtW at SSR allele CAC56D. Histograms developed for the population data of the traits with confirmed QTL are presented in Figure 1.

Table 3. Descriptive statistics of the phenotypic data scored in individual palms of the F_2 coconut population

Trait variable	Range	Mean (µ)	Standard deviation (α)	Variance (α^2)	Standard error of the mean
SH	2.1-15.45	8.47	2.544	6.47	0.251
G20	59-163	104.12	25.17	633.53	2.52
G150	55.4-99	76.04	8.447	71.35	0.841
NuoLvs	12-41	28.14	5.954	35.45	0.587
RL	76-845	433	87.1	7586.41	10.4
PtL	70.5-145	109.9	16.86	284.23	2.01
PtW	4.6-8.7	6.56	0.845	0.71	0.101
NuoLlts	88-138	112.79	8.1	65.61	0.97
LL	75.3-139.3	110.12	14.73	216.97	2.22
LW	4.1-6.4	5.10	0.5591	0.31	0.084
NuoB	6-17	11.70	2.89	8.35	0.285
NPB	0-13.5	4.7	2.639	6.96	0.26
FNW	538-3370	1676.5	613.7	376627.7	64
PFC	44.6-69.8	60.19	5.279	27.87	0.544
EFC	35.4-62.7	51.91	5.977	35.72	0.617

A population derived from a cross between diverse parents is required for QTL analysis for obtaining maximum segregation for important traits to facilitate the tagging of a higher number of QTL. Sri Lanka Tall and Sri Lanka Green Dwarf are the parents used in developing the studied F_2 population. These two coconut varieties represent South Asian, African and South East Asian origins (Gunn *et al.* 2011) of coconut providing maximum phenotypic and genotypic segregation for QTL analysis resulting in the revelation of a high number of QTL in the present study.

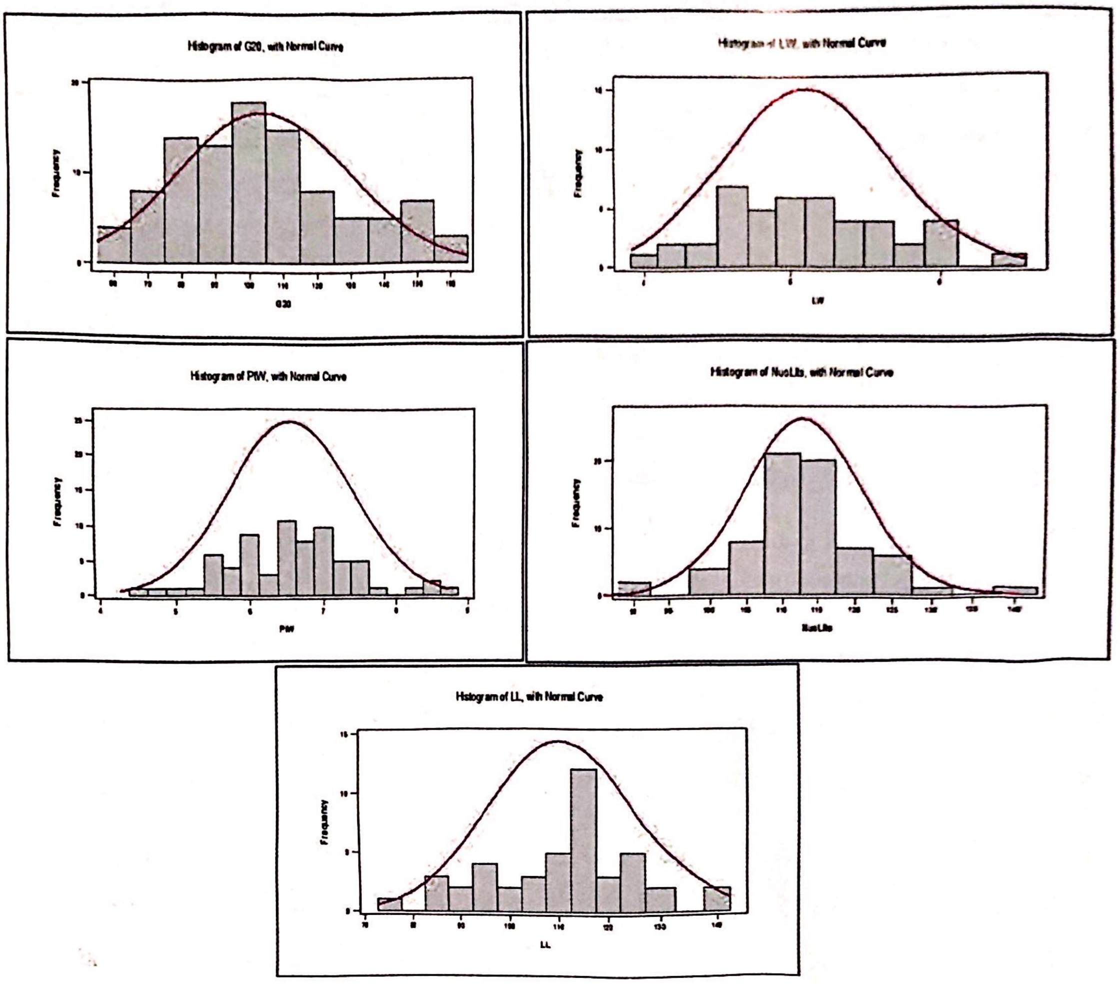

Figure 1. Histograms overlapped with normal curves for traits recording confirmed QTL in the segregating coconut palm population

In a previous study carried out with the coconut variety Rennel Island Tall, three QTL were detected for each trait number of bunches and number of nuts (Lebrun *et al.,* 2011). In addition, Shalini *et al* (2007) reported SSR and RAPD marker association for resistance of several coconut varieties in India for *Aceria* mite infestation. Compared to these studies a higher number of QTL was detected in the current study due to both phenotypic and genotypic segregation in the F2 population used in the current study.

Table 4. Significant QTL associations detected by single marker analysis (QTL confirmed based on P<0.01 are in bold letters) in the F_2 coconut palm population

Trait	SSR locus	Allele size (bp)	P value	Population mean	Mean of the sample having QTL	Mean of the sample lacking QTL	QTL effect
LW	CAC37	A- 178	0.047	5.10	4.9676	5.2966	Negative
NuoLlts	CAC37	E- 204	0.049	112.79	115.65	111.70	Positive
G20	CAC37	E-204	0.007	104.12	108.59	101.22	Positive
NuoB	CAC38	C-164	0.037	11.70	11.767	11.804	Negative
PFC	CAC43	B-197	0.022	60.19	58.350	60.554	Negative
EFC	CAC43	B-197	0.043	51.91	49.750	52.428	Negative
G20	CAC55	A-182	0.014	104.12	97.11	114.44	Negative
NuoLvs	CAC55	B-194	0.026	28.14	30.833	26.922	Positive
PtW	CAC55	C-196	0.008	6.56	6.8667	6.2867	Positive
LW	CAC55	C-196	0.010	5.10	5.3167	4.9542	Positive
PtW	CAC56	A-161	0.043	6.56	7.1800	6.3471	Positive
PtW	CAC56	B-173	0.019	6.56	6.7828	6.3432	Positive
NuoLlts	CAC56	B-173	0.010	112.79	113.86	111.46	Positive
NPB	CAC56	C-175	0.023	4.7	3.994	4.885	Negative
PtL	CAC56	C-175	0.035	109.9	105.94	110.53	Negative
LL	CAC56	C-175	0.002	110.12	98.20	113.63	Negative
LW	CAC56	C-175	0.043	5.10	4.8000	5.2320	Negative
PtW	CAC56	D-183	0.010	6.56	6.2590	7.0280	Negative

CONCLUSIONS

In the current study, we were able to confirm the presence of six QTL in the segregating F_2 population of coconut. The presence of 12 other QTL were indicated including the QTL for yield parameters, number of bunches and fruit size also. One of the confirmed traits is girth of stem at the base of the palm while all the rest were leaf characters contributing to the photosynthetic capacity of the plant. The information generated can be directly used in selecting parents for genetic improvement of coconut for desirable traits.

Search for QTL in perennial tree crops is a challenging task due to constraints mentioned above and these factors are common for coconut also. However, the application of molecular biological tools such as MAS is extremely important for expediting and enhancing the conventional coconut breeding strategies. In this context the research described in this paper marks a significant step as the first report in moving towards marker assisted selection and breeding in perennial tree crops in Sri Lanka.

ACKNOWLEDGEMENTS

The authors acknowledge Dr. Lalith Perera, Head of the Genetics and Plant Breeding Division of CRISL for facilitating this research.

REFERENCES

Anolles, G.C., and Petter, M.G. (1994). Staining nucleic acids with Silver: An alternative to radio isotopic and fluorescent labelling. Promega Notes Magazine. 45: 13-14.

Doyle, J.J., and Doyle, J.L. (1990). Isolation of plant DNA from fresh tissues. Focus 12:13-15.

Everard, J.M.D.T. (1996). Use of molecular markers for breeding of the coconut palm (*Cocos nucifera* L.). M.Sc Thesis. University of New England, Armidale. Australia.

Gunn, B.F., Baudouin L., and Olsen, K.M. (2011). Independent origins of cultivated coconut (*Cocos nucifera* L.) in the old world tropics. PLoS One 6(6):e21143.

Lebrun, P., N'Cho, N.P., Seguin, M., Grivet, L., and Baudouin, L. (1998). Genetic diversity in coconut (*Cocos nucifera* L.) revealed by restriction fragment length polymorphism (RFLP) markers. Euphytica 101:103-108.

Lebrun, P., Baudouin, L., Bourdeix, R., Louis Konan, J., Barker, J.H.A., Aldam, C., Herran, A., and Ritter, E. (2001). Construction of a linkage map of the Rennell Island Tall coconut type (*Cocos nucifera* L.) and QTL analysis for yield characters. Genome 44:962-970.

Perera, L., Russell, J.R., Provan, J., and Powell, W. (1999). Identification and characterization of microsatellites in coconut (*Cocos nucifera* L.) and the analysis of coconut populations in Sri Lanka. Molecular Ecology 8:344-346.

Perera, L., Russell, J.R., Provan, J., and Powell, W. (2000). Use of Microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L.). Genome 43:15-21.

Perera. S.A.C.N. (2010). QTL analysis in coconut via Genome mapping; Principles, requirements and prospects. Cocos. 19:57-65.

Shalini, K. V., Manjunatha, S., Lebrun, P., Berger, A., Baudouin, L., Pirany, N., Ranganath, R.M., and Prasad, D.T. (2007). Identification of molecular markers associated with mite resistance in coconut (*Cocos nucifera* L.) Genome 50: 35-42.