COCOS, 2010: ***19:*** *57-65*
Printed in Sri Lanka

RESEARCH ARTICLE

QTL Analysis in Coconut via Genome Mapping: Principles, Requirements and Prospects

S. A. C. N. Perera

Genetics & Plant Breeding Division, Coconut Research Institute, Lunuwila 61150, Sri Lanka
chandrikaperera2003@yahoo.com

ABSTRACT

Improvement of the coconut palm for desired traits is vital for producing high yielding and stress tolerant coconut cultivars. Many such economically important traits are governed by Quantitative Trait Loci (QTL). Coconut genome mapping is especially useful to overcome several difficulties related to coconut breeding and QTL analysis. At present, molecular markers have made it feasible to map QTL in plants. The paper describes the possibilities for mapping of the coconut genome to tag QTL using DNA markers. Genome mapping is the location of genes to positions along chromosomes. Genetic linkage between the genes located in the same chromosome is the basis for genome mapping. The distance between two genes is a function of the recombination frequency between the two genes measured in centi-Morgan (cM). Construction of an appropriate segregating population, selection of molecular marker systems and markers, use of appropriate mapping software are important considerations in coconut genome mapping. Genetic linkage mapping will result in a framework map and a QTL map for coconut based on genetic distances. Apart from genetic linkage maps there are some other types of genome maps possible for coconut, namely, association maps based on linkage disequilibrium, physical maps given by distances in base pairs and synteny maps based on comparative genomics. A fairly reliable genome map will mainly assist in marker assisted selection for important traits.

Key words: QTL analysis, genome mapping, coconut

INTRODUCTION

The genes governing most of the economically important traits of eukaryotes, such as yield, quality and stress tolerance in crop plants, milk yield, growth, fitness in farm animals and intelligence in humans are extremely complex. The phenotypes of such traits display continuous variation and are conditioned by allelic variation at several genetic loci, each with a relatively small effect compared with the effect of the environment. Such characters are referred to as quantitative traits and their inheritance is now known to be polygenic. The individual loci controlling a quantitative trait are referred to as Quantitative Trait Loci or QTL. The genetic dissection of a quantitative trait is termed QTL analysis, and is usually carried out with segregating populations.

Traditionally, QTL analysis of complex traits has been restricted to the statistical approaches of Biometry. However, these methods deal with the average effects of unresolved loci and provide little information about the number, location or relative effects of specific QTL (Kearsey and Pooni, 1996). The introduction of molecular markers in 1980's offered solutions to some of these problems marking a new era of scientific research with respect to QTL analysis. As a result, at present, molecular markers have made it feasible to map and characterize the polygenes underlying quantitative traits in natural populations.

Coconut is a major tropical plantation crop. Apart from being a major oil crop in the region, coconut has proved to be a multipurpose tree crop which plays an important economic role in many of the coconut growing countries. However, coconut cultivation suffers from low yields, drought and pest and diseases *etc.* resulting in low productivity. Consequently, breeding of the coconut palm for desirable traits is of utmost importance to blend desirable characters into cultivated material. As with many of the traits of other crop plants, breeding and improvement of coconut involves quantitative genetic analysis of the desired traits. However, the perennial nature, massive stature, high heterogeneity of cultivated material and the lack of an established vegetative propagation method make coconut one of the most difficult plants for QTL analysis and application of conventional breeding methods. In this context, coconut genome mapping using molecular marker technology can potentially play a key role in tagging important QTL in coconut. This paper discusses the basic principles, requirements and prospects for QTL mapping with relevance to coconut.

Gene mapping / QTL mapping

Gene mapping is the location of genes to positions on specific chromosomes (Kearsey and Pooni, 1996). Mapping and sequencing of chromosomal regions with genes affecting important characters has now become vital in breeding plants because it would help understand the gene function, gene regulation and their expression. More importantly, genome mapping will greatly accelerate the slower progress made with traditional breeding methods, in improvement of those complex characters such as grain or nut yield, yield quality and drought tolerance/resistance etc.

PRINCIPLES OF GENETIC LINKAGE MAPPING

Genetic linkage maps are entirely based on recombination frequencies. A genetic linkage map is an ordered set of loci with relative spacing determined from measured recombination frequencies. The phenomenon of genetic linkage between the genes located in the same chromosome provides the basis for the development of genetic linkage maps. Thus, genetic maps plot the estimated arrangement of genes/QTL along a chromosome with a distance related to the number of recombinations (crossovers) occurring along the specific chromosome. The distance between two genes on a genetic map is a function of the recombination frequency between the two genes measured in centi Morgan (cM), and not the actual physical distance.

Statistical methods for QTL mapping

Several statistical methods are available for data analysis for detecting marker-trait associations. The basic criterion for selecting a method is the number of markers that can be used to perform a test of significance. Statistical methods, Single Point Regression, (Weller, 1986), Interval Mapping (IM) (Lander and Botstein, 1989) Composite Interval Mapping (CIM) (Jansen, 1993) and Multiple Interval Mapping (MIM) (Kao et al. 1999) have been widely used in genome mapping in plants.

Mapping software

Development of genetic linkage maps essentially requires the deployment computer software developed for genome mapping. The most widely used genetic mapping software is Mapmaker (Lander et al. 1987). JoinMap (Stam, 1993) is a computer package used to construct an integrated genetic map from different sets of mapping populations where population sizes can be limiting. It searches for an arrangement of like markers between previously established genetic maps and suggests a linear arrangement of all the markers dependent on recombination values. Currently, there are more QTL mapping software made available for QTL mapping. Some packages such as QTL Café and QTL cartographer are available free of charge on the internet.

QTL ANALYSIS IN COCONUT BY GENETIC LINKAGE MAPPING

Availability of a genome map for coconut will pave way for more accurate insight into the coconut genome with respect to the economically important QTL. As with any other crop, construction of a linkage map in coconut involves several major steps. Quite a few of these tasks are highly challenging in a crop such as coconut due to the unavailability of clearly characterized mapping populations and also the limited availability of informative molecular markers. Three genetic linkage maps have so far been produced in the world for coconut (Rhode et al, 1999; Herran et al, 2000, Lebrun et al 2001). However, these maps have been developed with mapping populations consisting of a few number of individuals and hence are needed to be further improved for obtaining reliable information with respect to the reliable location of QTL.

CONSTRUCTING A MAPPING POPULATION FOR COCONUT

Selection of a mapping population is one of the most critical decisions in constructing a linkage map with DNA markers. A mapping population should comprise of the parents and their segregating progeny. Segregating populations such as F2s and back crosses or advanced segregating populations such as Recombinant Inbred Lines (RILs) or Doubled Haploid Lines (DHLs) are the commonly used mapping populations at present for linkage mapping in self pollinated crops. The choice of a mapping population depends upon the biology of the organism and the power of the different methods under the heritability value of the trait of interest. However, the development of the above mentioned typical mapping populations for coconut is hindered by the predominantly cross pollinating nature of coconut.

Genetic linkage mapping in outbred species such as coconut, thus poses additional challenges in QTL mapping. As a result, in QTL mapping ventures in outbred species, (e.g. most of the trees, humans etc.) the information is restricted to existing pedigree populations. The problems involved in obtaining information from pedigree populations are the substantially smaller family sizes compared with the commonly used segregating populations and the direct unavailability of information about the linkage phase of genes at the marker loci and QTL. To solve these problems, respectively a large number of families and complicated statistical tools are needed for modelling the inheritance of genes within a multiple generation pedigree (Mackay, 2001). In considering the breeding behaviour of coconut, it has been established that tall coconuts are cross pollinating while dwarf coconuts are self breeding. Therefore, a segregating population in coconut can be obtained by crossing self pollinating and thus homozygous dwarf coconuts with cross pollinating and thus heterozygous tall coconuts. However, one critical consideration in such an occasion is the selection of tall parents possessing sufficient levels of heterozygosity resulting in a considerable level of polymorphism in the subsequent progeny.

The major constraint in producing a fairly large mapping population in coconut is the limited number of seeds produced from a particular mother palm and also the low rate of success in artificial pollination. As a result, obtaining a sufficient number of progeny for mapping will take a long time which is not desirable because there will be a considerable age gap among the individuals of the resulting mapping population. Combining several separate half sib families to produce a single mapping population could be taken as an alternative for increasing progeny size within a shorter time span. In view of this understanding, it is appropriate to choose several dwarf palms from a population as female parents and a single highly heterozygous male tall coconut as pollen donor to construct the mapping population.

The next important point is the size of the mapping population from which the ultimate resolution of the map would be determined. However, the size of the mapping population will again depend on the choice of the mapping population. Yet, a mapping population with less than 50 individuals would not be sufficient for constructing a reliable map. It is difficult to separate two QTL located on the same chromosome with the generally used mapping techniques in a mapping population without an adequate number of progeny. For a reasonable map resolution it is necessary to score more recombinations which can only be done by increasing the population size and genotyping with four to five well-spaced markers along a chromosome, rather than increasing the number of markers along the chromosome (Darvasi et al. 1993; Kearsey and Pooni, 1996). Simulation studies have shown that the optimum population size (developed using the method described above) for coconut is between 200 to 400 individuals (Bandaranayake, 2006) assuming 100% heterozygosity at marker loci.

Selection of Molecular Marker Techniques and Genotyping the mapping population

Genotyping in genome mapping is the process that is used to distinguish individuals of the progeny who are homozygous for female parent (AA), homozygous for male parent (BB) and heterozygous (AB). Easily recognizable genetic markers are needed to identify and locate different genes along the chromosomes. Such markers have to exist in two or more allelic forms in order to detect polymorphism between them and to be used in mapping (Kearsey and Pooni, 1996). In the early days, all the attempts at developing genetic linkage maps were based on morphological markers. Although the information on these conventional maps is important to know the location of genes corresponding to phenotypical traits, their usefulness is limited by the low number of morphological markers, available. In addition, the expression of morphological markers is affected by environmental conditions. The advent of biochemical markers, followed by the molecular or DNA markers allowed more advanced techniques for producing genome maps densely covered with markers. Selection of molecular marker techniques for genome mapping depends on the breeding habit & genome size of the organism, already known information on genome organization of the organism and funds and facilities available.

MOLECULARMARKERSFORCOCONUT GENOME MAPPING

Coconut genome is considerably large containing a haploid number of 16 chromosomes. Consequently, a fairly large number of polymorphic markers are needed in genotyping the mapping population to achieve a dense marker coverage for efficient use of the genetic map. Molecular marker systems RAPDs, Microsatellites and AFLPs have so far been used in coconut. Use of RAPDs will not be practical for genotyping coconut mapping population due to very low levels of polymorphism. Approximately, about 300 microsatellite markers (Perera et al, 1999b; Rivera et al. 1999) have so far been developed for coconut and also there is the opportunity for the use of AFLP markers (Perera et al. 1999a) for genotyping of the mapping population. Microarray based Diversity Array Technology (DArT) marker system has been validated for coconut and about 350 coconut specific DArT markers have been identified (Perera and Kilian, 2008). Yet, some of these molecular markers may become uninformative to capture the segregating loci in an F1 mapping population. Accordingly, it will be advantageous to employ all the available molecular marker systems, especially SSRs, AFLPs and DArTs in combination to ensure sufficiently dense marker coverage for the coconut genome.

Framework map for coconut

Genotypic data of the mapping population are fed into mapping software to generate the framework map denoting the order of molecular markers in chromosomes which are called linkage groups at this stage. A mapping software such as JoinMap is highly suitable to correctly depict the structure of the segregating population in coconut.

QTL ANALYSIS IN COCONUT VIA QTL MAPPING

When establishing a QTL map, it should have two essential steps namely, the mapping of the molecular markers (framework map) and establishing the association of the traits (QTL) with markers (QTL map). Accurate marker data (genotypes) and trait scoring data (phenotypes) from a segregating population are vital for QTL analyses and it is possible to map only those QTL which are segregating in the mapping population. This segregating population should be raised using a proper experimental design and score the phenotypes of the traits of interest. Both phenotypic data and genotypic data should be analysed in combination by an appropriate statistical method to locate QTL linked with molecular markers on the framework map.

Several factors should be specially considered in field planting a mapping population of coconut for scoring quantitative morphological data. As described above a sufficiently large mapping population of coconut will consist of a minimum of 200 individuals which will occupy a large area in the field. Extreme attention is needed in the field to ensure uniform environmental conditions to each coconut palm in order to minimize the experimental error variation for unbiased scoring of the quantitative data.

OTHER MAPS FOR COCONUT

Maps covering genomes are of several types. They vary depending on factors such as the unit of measurement, the method of development and or the uses offered by the map. The above described genetic linkage maps resulting in framework and QTL maps are widely developed as part of QTL analysis with molecular markers. There are however, certain other types of genome maps which are applicable in coconut also.

Association mapping (Linkage Disequilibrium mapping) in coconut

Association mapping is a method for identifying loci in the inheritance of complex traits. In its simplest form, this method involves identification of markers with significant allele frequency differences with the phenotype of interest and a set of unrelated control individuals. A statistical association between genotypes at a marker locus and the phenotype is usually considered to be evidence of close physical linkage between a marker and a trait locus (Jonathan et al. 2000). Association mapping studies are based on existing populations/germplasm collections, which is a major advantage for species such as coconut where experimental populations are difficult to access and/or to construct. Association mapping studies are in progress for coconut at French Agricultural Research Centre for International Development (CIRAD) with a Vanuatu tall population comprising of several generations.

Physical maps for coconut

Unlike genetic maps, physical maps are based directly on the measurements of DNA structure in terms of the number of base pairs (in kilo bases). Yeast Artificial Chromosomes (YAC), Bacterial Artificial Chromosomes (BAC) and cosmids are used as vectors for cloning in order to produce physical maps of different organisms (Terryn et al 1999). Physical maps focusing on the region of interest are a requirement to characterize and isolate genes of desirable characters once they are mapped on a linkage map. Research on physical mapping of coconut using cosmid clones is underway (Snaidy et al. 2003) with the objective of providing an efficient prerequisite tool to enable marker assisted selection for important traits.

Synteny maps / Comparative genomics

Synteny maps are co-linear maps of distantly related species for certain portions in the genomes. Synteny maps are possible because of the existence of highly conserved portions of the genomes between related species in many crops. Rice and maize (Schmidt, 2000) Arabidopsis and Brassica species (Bohuon et al. 1998), Medicago and soybean (Li et al. 2008) are examples for such related species in which synteny studies have been reported via comparative genomics. Oil palm is a good example for a distantly related species of coconut and comparative genomics between oil palm and coconut is thus possible using genome maps to reveal important QTL.

Prospects for QTL analysis of coconut by genome mapping

A population developed by crossing Sri Lanka Red Dwarf (SLRD) and Sri Lanka Tall (SLT) respectively as maternal and paternal parents have been produced and field planted in Sri Lanka (Perera, 2007). This population proved not to be segregating adequately for producing a framework map although the previous coconut map have been produced using F1 segregating populations. As mentioned previously alleles segregating in tall and dwarf parents should be determined prior to producing an F1 segregating population for genome mapping in coconut. However, a segregating population such as a backcross or an F2 between tall and dwarf coconut varieties would be segregating for many different loci and thus would facilitate mapping of such loci. The major practical problem in developing conventional segregating populations in coconut is the long time duration taken for the construction of them. An F2 population planted at the Bandirippuwa Estate of the Coconut Research Institute has been shown to be segregating at marker loci in addition to display high levels of morphological polymorphism (Perera et al, 2009). Therefore, there may be a possibility of using this population for marker assisted selection in coconut and QTL mapping.

Once a relatively high density coconut genome map is produced it is expected that information from the previously constructed maps also can be incorporated by integrating the maps from different populations. The reliability of the integrated map depends on the availability of a subset of markers, which are common to all individual maps. It will also be possible to incorporate information generated from physical mapping, synteny studies with oil palm and association mapping in QTL analysis and identification of marker loci for important traits in coconut.

The main use of coconut genome map will be marker assisted selection for economically important traits in coconut breeding. Other advanced applications of genome maps such as gene cloning and genetic modifications will have to be delayed for coconut until advanced biotechnological techniques such as transformation becomes a practical possibility. However, the continuation of studies already started on genome mapping in coconut will ensure the availability of a large volume of important genetic information with respect to quantitative traits governed by QTL. QTL analysis via a coconut genome map will thus greatly assist in the genetic improvement of the coconut palm for incorporating desirable traits into cultivated material.

REFERENCES

Bohuon, E.J., Ramsay, L.D., Craft, J.A., Arthur, A.E., Marshall, D.F., Lydiate, D. J. and Kearsey M. J. (1998). The association of flowering time quantitative trait loci with duplicated regions and candidate loci in Brassica oleracea, Genetics 150: 393-401.

Bandaranayake, C.K. (2006). An effective population size for reliable map resolution of coconut (Cocos nucifera L.), CORD 22 (2): 41-47.

Darvasi, A., Weinreb, A., Minke, V., Weller, J.I. and Soller. M. (1993). Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map, Genetics 134: 943-951.

Jansen, R.C. (1993). Interval mapping of multiple Quantitative Trait Loci. Genetics 135: 205-211.

Jonathan, K, Pritchard, M. S., Rosenberg, N.A. and Donelly, P. (2000). Association mapping in structured populations, Amer. J. human Genetics 67: 170-181

Herran A., Estioko L., Becker D., Rodriguez, M.J.B., Rohde W. and Ritter E. (2000). Linkage mapping and QTL analysis in coconut (Cocos nucifera L.). Theor. and Appl. Genetics 101: 292-300

Kao, C.H., Zeng, Z.B. and Teasdale, R.D. (1999). Multiple interval mapping for Quantitative Trait Loci, Genetics 152 1203-1216.

Kearsey, M.J. and Pooni H.S. (1996). The Genetical Analysis of Quantitative Traits. Chapman & Hall.

Kebede, H., Subudhi, P.K., Rosenow, D.T. and Nguyen, H.T. (2001). Quantitative Trait Loci influencing drought tolerance in grain sorghum (Sorghum bicolor L. Moench), Theor. and Appl. Genetics 103: 266-276.

Lander, E., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S., Newburg, L. and Green, P. (1987). Mapmaker a computer package for constructing genetic-linkage maps, Cytogenetics and Cell Genetics 46: 642-652.

Lander, E. S. and Botstein D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps, Genetics 121: 185-199.

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 Rennel Island Tall coconut type (Cocos nucifera L.) and QTL analysis for yield characters, Genome 44: 962-970.

Li, L., He, H., Zhang, J., Wang, X., Bai, S., Stolc V., Tongprasit, W., Young, N.D., Yu, O. and Deng, X. (2008). Transcriptional analysis of highly syntenic regions between Medicago trunculata and Glycine max using tiling microarrays, Genome Biol. 9: 57

Mackay, T.F.C. (2001). The genetic architecture of quantitative traits, Ann. Rev. Genetics 35: 303-339.

Perera, K.N.S., Perera, S.A.C.N, and Vivehananthan, K. (2009). Detection of Microsatellite marker segregation in an F2 population of coconut (Cocos nucifera L.) Proc. 9th Symp. Agric. Res. Wayamba University of Sri Lanka 249-253

Perera, L., Russel, J.R., Provan, J., McNicol J.W. and Powell, W. (1999a) Evaluating genetic relationships between indigenous coconut (Cocos nucifera L.) accessions from Sri Lanka by means of AFLP profiling, Theor. and Appl. Genetics 96: 545-550

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

Perera, L. (2007). Coconut genome mapping, Report of the Genetics and Plant Breeding Division, Ann. Rept. Coconut Res. Inst. of Sri Lanka. 59

Perera, S. A. C. N. and Kilian, A. (2008). Diversity Arrays Technology; A high throughput molecular marker system for coconut, Pragna xix(1) 60-64

Rivera, R., Edwards, K.J., Barker, J.H.A., Arnold, G.M., Ayad, G., Hodgkin, T. and Karp, A. (1999). Isolation and characterization of polymorphic microsatellites in Cocos nucifera L. Genome 42: 668-675.

Rohde, W., Becker, D., Kullaya, A., Rodriguez, M.J.B., Herran, A., Ritter, E. (1999). Analysis of coconut germplasm biodiversity by DNA marker technologies and construction of a first genetic linkage map. In: Current advances in coconut biotechnology. Eds. C. Oropeza, J.L. Verdeil, G.R. Ashburner, R. Cardega and J.M. Santamaria. Kluwer Academic Publishers, The Netherlands. 99-120

Schmidt, R. (2000) Synteny; Recent advances and future prospects, Current Opinion in Molecular Biol. 3: 97-102

Snaidy, V., Rodhe, W., Ritter, E., Becker, D. and Herran A. (2003). A rapid way pf physical mapping in coconut and oil palm, Proc. Deutcher Tropentag "Technological and industrial innovations for sustainable rural development" Gottingen.

Stam, P. (1993). Construction of integrated genetic-linkage maps by means of a new computer package - Joinmap, Plant J. : 3 739-744.

Terryn, N., Rouze, P. and Van Montagu, M. (1999). Plant genomics. Febs Letters 452: 3-6.

Weller, J.I. (1986). Maximum-likelihood techniques for the mapping and analysis of quantitative trait loci with the aid of genetic-markers, Biometrics 42: 627-640.