

## **Happy mapping as an alternative to overcome the problems in coconut genome mapping**

C.K. Bandaranayake<sup>1</sup>

### **Abstract**

An excellent way of producing a reliable mapping population for quantitative trait loci analysis and marker assisted selection was considered. A physical mapping method known as 'Happy Mapping' was discussed to make a framework map as an alternative to overcome the problems associated with meiotic mapping.

**Key words:** Coconut, framework map, genome mapping, happy mapping, meiotic mapping, mapping population, polymorphic marker and segregation.

**Abbreviations:** Amplified Fragment Length Polymorphism (AFLP), Log of Odds (LOD), Marker Assisted Selection (MAS), Polymerase Chain Reaction (PCR), Quantitative Trait Loci (QTL) and Simple Sequence Repeats (SSR).

---

<sup>1</sup> Genetics and Plant Breeding Division, Coconut Research Institute, Lunuwila, Sri Lanka  
Email: [champakumarib@hotmail.com](mailto:champakumarib@hotmail.com)

## Introduction

Development of a genome map is very advantageous in a crop like coconut because of its particular limitations such as, a long juvenile period, large crop size, and inherent heterozygosity, out breeding nature and poor vegetative propagation ability. Therefore, a high-density linkage map and thereby incorporating MAS are imperative to increase the efficiency of coconut breeding. The first genome map for coconut has been developed for EATxLAGT (East African Tall x Laguna Tall) F<sub>1</sub> population (Rohde *et al.*, 1999). This work was extended with MYDxLAGT (Malayan Yellow Dwarf x Laguna Tall) and six QTL for early germination were detected (Herran *et al.*, 2000). This is the first opportunity for MAS because germination is genetically correlated to early flowering and then to yield. Further, QTL for other traits such as, leaf production, girth and height were identified for the same mapping population (Ritter *et al.*, 2000). In addition to that, another coconut genome map has been produced using half-sib families of CRDxRIT (Cameroon Red Dwarf x Rennell Island Tall) cross and some QTL were detected for several yield characters such as number of bunches and number of nuts (Lebrun *et al.*, 2001).

However, most of the above mapping populations were small in size such as, 52 F<sub>1</sub>s in MYDxLAGT (Herran *et al.*, 2000) and 67 half-sibs in CRDxRIT (Lebrun *et al.*, 2001). In addition to that, the low polymorphism was observed in those studies using MYDxLAGT (Herran *et al.*, 2000) and CRDxRIT (Lebrun *et al.*, 2001). Herran *et al.* (2000) has discussed that this low polymorphism in parents is much lower than that observed in many other crops. According to Lebrun *et al.* (2001), only a small number of Amplified Fragment Length Polymorphism (AFLP) bands were polymorphic in parents and those did not segregate in the progeny. Small population sizes initiate errors while molecular marker mapping, since the resolution of a map and the ability to determine marker order is largely dependent on population size (Young, 1994) and also may cause problems

whilst QTL mapping such as, the effects of QTL detected may be overestimated and the QTL of little effect may not be detected (Beavis, 1994).

Hence, it is obvious that there are two main difficulties, such as, the unavailability of an appropriate mapping population and a limited number of polymorphic markers, associated with coconut linkage mapping (Bandaranayake, 2002). The second difficulty, *i.e.* the low number of polymorphic markers, entirely depend on the mapping population. If we could choose genetically diverse parents to produce a particular mapping population, many loci could be made polymorphic. On top of that, the first thing to do is to produce a mapping population from genetically distinct parents. Sri Lankan tall coconuts and Sri Lankan dwarf coconuts come under two main genetically distinct groups of coconuts known as *Typica* and *Nana*. Therefore, Sri Lankan Dwarf and Sri Lankan Tall coconuts are one of the best selections as parents to produce a mapping population.

The major constraint, in producing a fairly large mapping population to overcome the first difficulty, *i.e.* the unavailability of an appropriate 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, and as a result an insufficient family size. Therefore, the next limitation is the population size that has been tried to solve by combining several separate families using both simulated data (Bandaranayake, 2003) and real data. It has been found that the simulated data of a combined mapping population came from several separate families could map the coconut genome without any difficulty (Bandaranayake, 2003). However, this has been found to be impossible with real data due to errors in simultaneous handling of different types of pollen and several palms (Bandaranayake, 2002). There is an alternative to overcome these errors using a single pollen type from a specified parent. Dwarf coconuts are highly homozygous and therefore it may be possible to produce identical genotypes by selfing over several generations. These identical

genotypes can then be simultaneously crossed with pollen from a specific tall parent to generate large quantities of seeds in a short period of time. This was the system used to construct a linkage map of RIT coconuts by Lebrun *et al.* (2001). Therefore, it is possible to obtain a reasonable size progeny having a little difference in age among individuals so that it would not complicate field evaluations for QTL analysis.

Considering the above points, the main problem is the production of a proper segregating population because of the long generation interval, low success rate in artificial pollination and rather unmanageable physical structure of coconut palms. Therefore it will take a long time to produce a segregating population and also there is a possibility to have some human or technical errors along the long process (Bandaranayake, 2002). The other problem is the difficulty of obtaining a reasonable number of polymorphic markers to produce the map. It was very difficult to cover whole genome with a considerable number of SSR markers due to the large size of genome in coconut. In order to collect more markers, AFLP analysis has been tried in addition to those of SSR markers (Bandaranayake, 2002). But it was a really challenging task with very few polymorphic markers with each AFLP primer combination. Therefore, there would need to be a massive number of primer pairs to achieve a sufficient number of polymorphic markers. Hence, it is clear that almost all difficulties in coconut meiotic mapping can be overcome with the use of ‘happy mapping’ (Table 1).

### Principle behind the process “Happy Mapping”

The term ‘happy mapping’ has come from its theoretical approach that includes haploid cells and the polymerize chain reaction (PCR). This approach is analogous to classical linkage mapping where meiosis both breaks DNA by crossing-over and segregates it into aliquots containing haploid amounts of DNA (gametes). These processes are replaced by *in vitro* analogues by means of ‘happy mapping’, breaking DNA physically and then diluting it into

aliquots that contain ~1 haploid equivalent. Generally here, it is possible to introduce many more breaks into the chromosome than meiosis and achieve far higher resolution than genetic maps. Consequently ‘happy mapping’ is an *in vitro* linkage technique based on screening approximately haploid amounts of DNA by the PCR.

**Table 1. Meiotic mapping versus happy mapping**

Meiotic mapping	Happy mapping
Two copies of a genome must be distinguished from one another by DNA polymorphisms.	A single copy of the genome of interest is enough for analysis.
A segregating population is needed.	A single plant of interest is sufficient.
Only polymorphic markers can be used to genotype.	Both polymorphic and non-polymorphic DNA markers can be used.
A limited number of DNA markers can be used.	A wider spectrum of DNA markers can be used.
Some markers are not informative in some progenies.	All markers are informative in every aliquot.
Resolution power of mapping is less.	High resolution of mapping can be achieved.
Takes time to produce mapping populations and accuracy depends on several different factors.	Fast and accurate.
Suffer from the inaccuracies in meiotic recombination due to recombination hotspots.	Not subject to distortions of meiotic recombination.

The ‘happy mapping’ strategy includes several steps (Figure 1).

1. DNA is extracted from a particular organism.
2. Intact genomic DNA is cut by irradiation to get a pool of random segments (This step is analogous to meiosis and crossing-over in classical linkage mapping).
3. The random fragments are size selected by Pulsed Field Gel Electrophoresis (PFGE) to

eliminate the smallest and largest fragments that contain no mapping information.

4. Fragments are divided into aliquots (typically 100-200) containing ~1 haploid genome equivalent, so each marker is present in only a subset of the aliquots (This step is analogous to segregation in classical mapping).
5. This mapping panel (segregating progeny in classical mapping) is pre-amplified a hundred-fold by primer extension pre-amplification.
6. Sub-fractions of the pre-amplified panel are screened for markers specific for that individual using nested PCR followed by gel analysis to determine which markers are present in each aliquot.
7. A table is constructed in which LOD scores reflect the tendency of pairs of markers to occur together in the aliquots. A high LOD score reflect high co-segregation frequency of markers.
8. A map is constructed using co-segregation frequencies of marker pairs.

In view of that, the entire idea behind 'happy mapping' technique is that the closely linked markers are more likely to be found in the same aliquot than unlinked markers. The number of times that any two markers co-segregate is used to construct the map because closely linked markers will co-segregate from most of the cells while unlinked markers will do so rarely. Though happy mapping is also a form of segregation analysis, it is an entirely *in vitro* process. It has been used for mapping human genome (Dear and Cook, 1989, 1993; Water *et al.* 1983; Dear *et al.* 1998; Biunno *et al.* 2000) and non-human genomes (Piper *et al.* 1998; Williams *et al.* 2000).

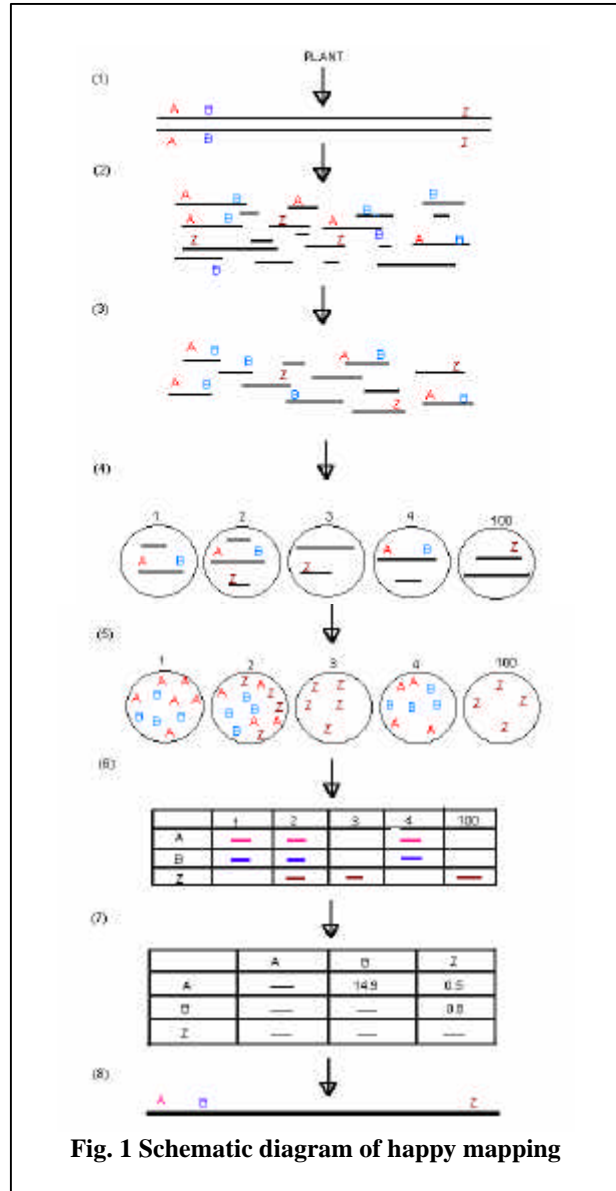


Fig. 1 Schematic diagram of happy mapping

In view of that, the entire idea behind 'happy mapping' technique is that the closely linked markers are more likely to be found in

### Justification in favour of coconut

It is very easy to circumvent the first problem in coconut, *i.e.* a proper segregating population, using 'happy mapping' because it doesn't need a segregating population and almost certainly any one single plant is sufficient. The 'happy mapping' panel (analogous to segregating population) is prepared as aliquots using DNA fragments obtained from that single plant. In this case, there is no need to bother about the long generation time, low success rate in artificial pollination and human or technical errors due to the hard process of producing multi-generations. Therefore, it is practicable to obtain a precise 'happy' panel quickly compared to genetic mapping especially in coconut. The second problem, polymorphic markers, is also easily overcome in 'happy mapping' because it can use both polymorphic and monomorphic markers in the same way due to scoring of presence or absence of particular bands. The only thing is to find a sufficient number of DNA markers specific for coconut. Therefore, it is feasible to acquire many markers to construct the coconut map because each marker is equally informative in every aliquot unlike genetic mapping.

Accordingly, the material need for happy mapping are a single coconut plant and an adequate number of DNA markers specific for coconut. Finding a coconut palm is not more than nothing and also it is not difficult to get a substantial number of DNA markers considering all molecular marker systems currently available. The foremost advantage here is the ability to use both polymorphic and monomorphic markers rather than ignoring monomorphic markers in most of other genotyping processes.

Because 'happy mapping' is an entirely *in vitro* process, any gross distortion of the map due to biological activity of centromeres, telomeres or other sequences or to the effect of chromosome structure could not be expected.

Therefore this may be of specific benefit in coconut genome mapping having little information on chromosome structure on locations of centromeres and telomeres. Whereas most genetic mapping methods are limited in the resolution that they can achieve, the reverse is true for happy mapping. Resolution down to a few kilo bases is possible by using small DNA fragments (Dear and Cook, 1993; Walter *et al.* 1993), but it is only to a few mega bases in genetic mapping (Cox *et al.* 1990). Therefore, 'happy mapping' can easily be used as an alternative for meiotic mapping to produce a framework map.

### Acknowledgement

I am kindly indebted to Professor. M. J. Kearsey, School of Biosciences, University of Birmingham, UK for directing me to find a substitute to overcome the problems associated with coconut genome mapping.

### References

- Bandaranayake, C.K. 2002. Genome mapping of *Arabidopsis thaliana* and *Cocos nucifera*. PhD thesis. The University of Birmingham, UK, 227pp.
- Bandaranayake, C.K. 2003. Investigation of the feasibility of constructing a map for coconut with several F2 families using computer-simulated data. *Cord* **19**(1): 59-68.
- Beavis, W.D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. In: *Proceedings of the 49<sup>th</sup> Annual Corn and Sorghum Industry Research Conference*, pp. 250-266. Chicago, American Seed Trade Association, Washington D.C.
- Biunno, I., Bernard, L., Dear, P., Cattaneo, M., Volorio, S., Zannini, L., Bankier, A. and Zollo, M. 2000. SEL1L, the human homologue of *C. elegans* sel-1: refined physical mapping, gene structure and

- identification of polymorphic markers. *Human Gen.* **106**: 227-235.
- Cox, D.R., Burmeister, M., Price, E.R., Kim, S. and Myers, R.M. 1990. Radiation hybrid mapping: A somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science* **250**: 245-250.
- Dear, P.H. and Cook, P.R. 1989. Happy mapping: a proposal for linkage mapping the human genome. *Nucleic Acids Res.* **17**: 6795-6897.
- Dear, P.H. and Cook, P.R. 1993. Happy mapping: linkage mapping using a physical analogue of meiosis. *Nucleic Acids Res.* **21**: 13-20.
- Dear, P.H., Bankier, A.T. and Piper, M.B. 1998. A high-resolution metric HAPPY map of human chromosome 14. *Genomics* **48**: 232-241.
- 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. App. Gen.* **101**: 292-300.
- 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.
- Piper, M.B., Bankier, A.T. and Dear, P.H. 1998. A HAPPY map of *Cryptosporidium parvum*. *Genome Res.* **8**: 1299-1307.
- Ritter, E., Rodriguez, M.J.B., Herran, A., Estioko, L., Becker, D. and Rohde, W. 2000. Analysis of quantitative trait loci (QTL) based on linkage maps in coconut (*Cocos nucifera* L.). pp. 42-48. In: *Plant Genetic Engineering towards the Third Millennium*, A. Arencibia (Ed.). Elsevier Science B. V. Amsterdam, The Netherlands.
- Rohde, W., Becker, D., Kullaya, A., Rodriguez, M.J.B., Herran, A. and Ritter, E. 1999. Analysis of coconut germplasm biodiversity by DNA marker technologies and construction of a first genetic linkage map, pp. 99-120. In: *Current Advances in Coconut Biotechnology*, C. Oropeza, J.L. Verdeil, G.R. Ashburner, R. Cardega and J.M. Santamaria (Eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Walter, G., Tomlinson, I.M., Cook, G.P., Rabbits, T.H. and Dear, P.H. 1993. HAPPY mapping of a YAC reveals alternative haplotypes in the human immunoglobulin V<sub>H</sub> locus. *Nucleic Acids Res.* **21**: 4524-4529.
- Williams, J.G. and Firtel, R.A. 2000. HAPPY days for the *Dictyostelium* genome project. *Genome Res.* **10**: 1658-1659.
- Young, N.D. 1994. Constructing a plant genetic linkage map with DNA markers, pp. 39-57. In: *DNA Based Markers in Plants*, R.L. Phillips and I.K. Vasil (Eds.). Kluwer Academic Publishers, The Netherlands.