Molecular markers – New tools for an old science – The case of tea.

Francis N. Wachira

Tea Research Foundation of Kenya, P.O. Box 820, Kericho, Kenya.

Summary

Recent advances in molecular biology, principally in the development of the Polymerase Chain Reaction (PCR) for amplifying DNA have resulted in powerful techniques, which have variously been used to aid tea (*Camellia sinensis*) breeding in Kenya. This paper discusses progress made in the use of molecular markers to determine the level, structure and origin of genetic diversity among populations of tea and to develop marker assisted selection strategies in tea breeding programmes.

Key words: Molecular markers, germplasm, breeding, *Camellia sinensis*.

Introduction

Tea was first introduced into Kenya during the early 1900s. There is however, very scant historical data and information on the actual source of the germplasm from which pioneer plantations were developed though it is thought to have its origin in the North-East of India and possibly in Assam. Lack of such vital collection descriptors could mean that there is a likelihood that the germplasm was obtained from a restricted source and may therefore be of narrow genetic base which would necessitate its enrichment through introduction, breeding and conservation. The cultivation of clonal plants has become popular and such plants now accounts for 60% of all tea in Kenya. Like in several other countries, only a few clones and their progenies, which may represent a small fraction of the total available diversity are now grown. For example in Kenya, 14 clonal cultivars now account for 28% and 80% of all tea in the estate and small holder sectors, respectively. Since clones represent instantly fixed genotypes through vegetative propagation, over-reliance on a limited number has led to a reduction of the on-farm diversity. The old, yet diverse seedling teas in pioneer plantations are also gradually being uprooted and replaced with a few improved clones. A sustainable agricultural system, however, requires that components of biological diversity be used in a way and at a rate that will not lead to a long term decline of diversity, thus maintaining its potential to meet the needs and aspirations of present and future generations.

The threat posed by a diminishing genepool could be arrested by a deliberate conservation program. Despite the economic value of tea, there has been lack of a co-ordinated program for national and international collection, characterization and conservation of germplasm which may have led to lack of the effective utilization of tea genetic resources. Several tea producing countries including Kenya, have established *ex-situ* collections mostly made up of popular local and introduced germplasm which form active bases for clonal regeneration, multiplication and distribution [5,9,7]. The national repository centre at Kericho, Kenya has a collection of slightly over 200 conserved accessions [7]. This collection like others held elsewhere is not based on optimized collection strategies and genetic redundancy may be expected. A large genepool of tea is also held by farmers in pioneer plantations. Information on taxonomic characteristics, genetic diversity and biogeography of individual tea bushes and clones in these collections though vital in identifying sources of desirable genes, is scant and poorly documented. A large portion of the tea genetic resources however remains inaccessible to breeders in Kenya and those outside its center of distribution in the Yunnan and Guangxi provinces of China [1]. Accurate assessment of the level of genetic variation, characterization and evaluation of genotypes within the existing collection and plantations in Kenya is important in order to devise optimum management strategies for its efficient and sustainable utilization and conservation. This would also facilitate the establishment of a core collection that is representative of most of the available diversity. Among others, the effective conservation of tea germplasm will be dependent upon optimized strategies for identification.

Traditionally, plant genetic resources have been evaluated on the basis of morphological and agronomic traits, which do not necessarily reflect genetic diversity. The first step towards the effective management of such resources must therefore be the development and use of reliable and standardized genetic descriptors, which not only distinguish individuals and varieties but also reflect inherent variation and genetic relationships among collection holdings. A range of molecular techniques can now be used to detect variations at the DNA level. The array of techniques falls into three broad categories with respect

to strategy; a) Non-PCR (Polymerase Chain Reaction) based approaches, b) PCR arbitrary priming and, c) Targeted-PCR. Non-PCR based approaches are labour intensive and have low throughput and are therefore not widely used on tea. The assay procedures based on the PCR have been the choice markers for tea in most countries including Kenya. The following marker systems have been used in our studies; Randomly Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Inter Simple Sequence Repeat Polymorphism (ISSR), Sequence Tag Sites (STS), and Single Strand Conformation Polymorphism (SSCP) [4,6,7,8].

Marker utility in the determination of diversity of tea germplasm resources.

In a study using 48 tea cultivars from different tea growing regions of the world, the AFLP assay was revealed to detect a higher level of polymorphism as measured by the expected heterozygosity compared to the RAPD assay. It also had a larger marker index (Table 1). Among other considerations, the choice of a marker system is dependent upon its information content and the extent to which it can detect multiple polymorphism. Based on these two parameters, the AFLP assay could be considered more robust than the RAPD assay.

Table 1. The average expected heterozygosity for polymorphic markers, $Hav_{(p)}$ of the fraction of polymorphic markers (β), multiplex ratio (n), the effective multiplex ratio (E), and the marker index (M.I) for RAPD and AFLP (non radio - labelled markers) calculated on the basis of experimental data from 48 tea cultivars collected from different regions of the world.

Marker	No. of loci	$Hav_{(p)}$	β	N	E	MI
RAPD	176	0.311 (0.012)	8.23*	1+	8.23	2.56
AFLP	90	0.361 (0.014)	0.89	51	45.39	16.39
Combined (RAPD+AFLP)	266	0.317 (0.009)	-	-	-	-

*Standard error of Hav is shown in parenthesis; *Mean number of polymorphic bands per primer on the 48 representative tea cultivars; +Primers used per gel lane.*

Studies to determine the structure of genetic diversity in tea, revealed that most of the variation resided within populations (Table 2). Significant genetic variation ($P<0.001$) among countries was also revealed which is a reflection of disjunction due to geographical isolation. This is an indication that after the initial introductions from the centers of distribution into the respective growing countries, there has been very minimal exchange of tea germplasm resources and therefore populations from the different countries have evolved independently as islands of genetic diversity thus being highly differentiated. Tea growing countries therefore need to exchange germplasm more.

Table 2. Nested Analysis of Molecular Variance for 48 individuals (n) of tea from 12 populations (p) derived from 8 countries/groups (g). Degrees of freedom (df), sum of squares (SS) and the significance of the variance components are shown and the partitioning of the total diversity into population components.

Source of variation	df	SS	variance component	p-value	% of total
Among countries.	7 (g-1)	464.21	6.722	<0.001	23.5
Among populations Within countries.	4 (p-g)	95.25	1.271		4.4
Among individuals Within populations.	36 (n-p)	742.88	20.64	<0.001	72.1

Analysis of molecular diversity using RAPD and ISSR markers revealed that the 14 widely grown clonal cultivars accessed only about 41% of the total available diversity in pioneer tea plantations in Kenya (data not presented). As is evident in the genetic diversity tree presented (Figure 1) these clones have a high genome homology an indication that selection has been skewed and that common gene alleles have been selected for in the said clones. The existing genepool in pioneer tea plantations in Kenya has therefore not been efficiently used and thus the need to broaden the genetic base of improved commercial cultivars by selecting divergent clones. Likewise, several tea producing countries may be accessing only a small proportion of the total available diversity within their borders. It is prudent that the existing germplasm resource be used efficiently before elaborate germplasm introduction programs are considered. Figure 1 also reveals the presence of some interesting germplasm from some pioneer plantations in Western Kenya which merit conservation (most distant on the tree). Several such plantations are under the threat of being uprooted and replaced with improved clonal cultivars and thus the apparent need for their conservation.

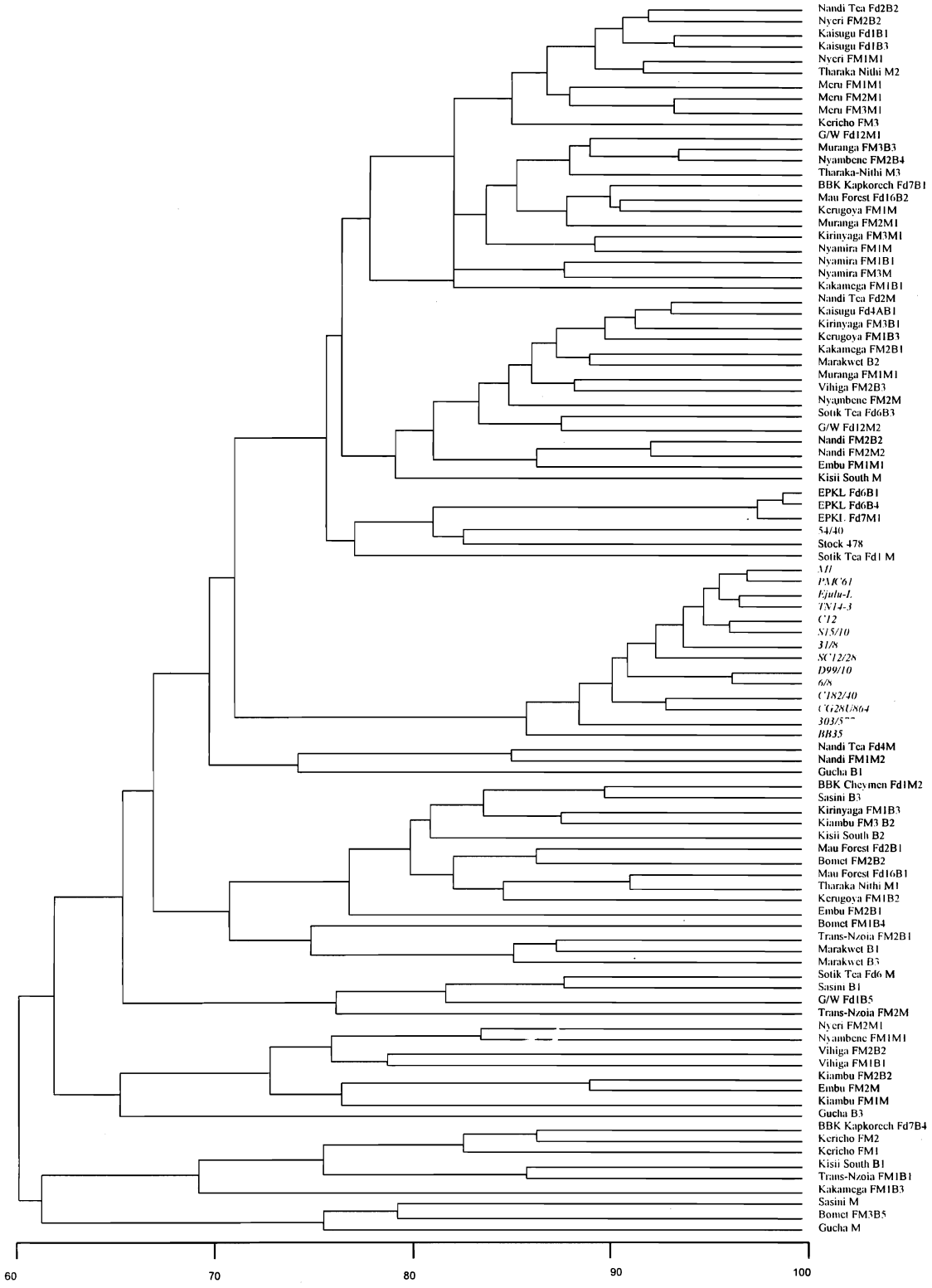

Fig 1: Dendrogram of genetic similarity among individual bushes representing 14 popular commercial clonal cultivars (in italics) and from pioneer tea plantations in 26 tea producing districts of Kenya. The dendrogram was constructed according to the UPGMA method using the Nei and Li similarity coefficient.

Tea breeding strategies.

Genetic linkage maps provide a great potential for increasing the speed and precision of tea improvement programs. Because of its long juvenile period, predominantly allogamous nature and high heterozygosity, conventional breeding requires large investments in time and land for progeny trials. A linkage map could provide a vital tool for relating genotype to phenotype, and thus facilitate the identification and selection of recombinant individuals with desirable attributes. For a long time, however, mapping of plant genomes was restricted to linkage groups that caused discrete and visible changes in morphological characters which unfortunately are very rare. Molecular markers now provide a means of constructing saturated maps.

Construction of such maps nevertheless, requires that suitable segregating genetic stocks of tea be first developed. The availability of multigeneration pedigrees of tea is however limited. Available pedigrees involve only two parents and their full-sib progeny, or maternal half-sib families. There has therefore been need to explore alternative approaches to mapping in tea. One applicable approach involves the use of already existing and commonly generated pedigrees in breeding programs. This design termed "pseudo- testcross", involves crossing two non-inbred parents, one of which is known to segregate for the trait of interest [3]. We have used such a pedigree to study the segregation of RAPD and AFLP markers and to construct a partial genetic linkage map of tea [2]. The location of putative Quantitative Trait Loci (QTLs) controlling some traits was determined by regressing the marker locus data against the trait values. Marker locus prediction of trait expression was then determined from the R^2. Significant ($P<0.05$) associations were obtained in several pairs of combinations of quantitative trait and marker locus indicating potential for identification of markers useful in early indirect selection for elite plants (Table 3). Several such segregating marker loci can provide an effective means of identifying genes and/or genomic regions influencing a wide array of traits in tea. However, since most individual markers accounted for only a small proportion of the total phenotypic variance in trait expression, there is need to continue identifying many more segregating markers and to test the populations over a long time in different environments (genotype x environment interactions).

Table 3. Numbers of markers with significant effect ($P<0.05$) on four quantitative traits in a mapping population of tea and the magnitudes of their effects (note: population was tested in one environment).

Trait	No. of markers[a]	Locus R squared[b] Min.	Max.
Caffeine content	18	3.4	10.5
Shoot density	42	3.3	14.0
Shoot dry weight	58	3.2	14.2
Yield	61	3.3	24.1

[a]- including markers not placed on the map., [b]- minimum and maximum percent of the phenotypic variation explained by genotypic classes at marker loci with significant association with trait expression.

Conclusion

Breeding programs in Kenya and elsewhere are adopting emerging technologies to enhance their effectiveness. Molecular characterization of germplasm will establish the level of useful variation available to tea breeders. Such variation may also be crucial to identification of specific heterotic groups. With the help of molecular markers integrated with phenotypic assessment, core collections representative of as much of the total available diversity in tea will be established thereby improving on the conservation and management of *Camellia* genetic resources. The potential use of molecular markers for indirect selection of elite plants will continue to be evaluated in relation to the costs involved.

Acknowledgements

Part of this work was funded by the Agricultural Research Fund (contract No. ARF/CSWP/3008001/1)

References

[1] Chang, H.T. & Bartholomew, B. (1984). Camellias, B.T. Batsford Ltd., London. U.K.

[2] Hackett, C. A.; Wachira, F.N.; Paul, S.; Powell, W. & Waugh, R. (2000). *Hered.* **85**:346-355.

[3] Hemmat, M. & Weeden N.F. (1994). *J. Hered.* **85**:4-11.

[4] Paul, S.; Wachira, F.N.; Powell, W. & Waugh, R. (1997). *Theor. Appl. Genet.* **94**:255-263.

[5] Singh, I.D. (1999). In: Global Advances in Tea Science. (N.K. Jain (Ed)) pp 427-448. Aravali Books International, New Dehli, India.

[6] Wachira, F.N.; Powell, W. & Waugh, R. (1997). *Hered.* **78**:603-611.

[7] Wachira, F.N.; Tanaka, J. & Takeda, Y. (2001). *J. Hort. Sci. Biotec.* **76 (5)**: in press.

[8] Wachira, F.N.; Waugh, R.; Hackett, C.A. & Powell, W. (1995). *Genome* **38**:201-210.

[9] Yu, F. & Xu, N.(1999). In: Global Advances in Tea Science. (N.K. Jain (Ed)) pp 393-412, Aravali Books International, New Dehli, India.