DNA fingerprint based on nuclear and chloroplast genome, combine analysis on Sulawesi cacao (*Theobroma cacao* L.)

I N Suwastika¹*, Umrah¹, Muslimin², Y Ishizaki³, A F Cruz³, Z Basri⁴ and T Shiina³

¹ Biology Department, Faculty of Sciences, Tadulako University, Palu 94118, Indonesia
² Faculty of Forestry, Tadulako University, Palu 94118, Indonesia.
³ Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto 606-8522, Japan
⁴ Faculty of Agricultural, Tadulako University, Palu 94118, Indonesia

*E-mail: isuwastika@yahoo.com.au

**Abstract.** High genetic diversity and high similarity on phenotypic performance of local Sulawesi cacao (*Theobroma cacao* L.) was challenge in elucidating of identification including developing of DNA fingerprint approaches. Identification of nuclear genome polymorphism, based on RAPD and SSR marker are well established. Recently, genetic variation on chloroplast genome was introduced as alternative in this polymorphism identification works. Here we analysis the possibility of application of SSR and chloroplast markers independently, also in combination of both approaches. The combination analysis showed the accuracy in analyzing sub species level on local Sulawesi cacao.

**1. Introduction**

Genetic diversity of cacao (*Theobroma cacao* L.) was commonly assessed based on approaches of RAPD (Random Amplified Polymorphic DNA)[1][2] or minisatellite sequences (SSR, simple sequences repeat)[3][4][5]. Morphological characteristics cannot be used easily, it due to the high rate of evolution of this species. Genetic diversity was also reported that it close related to geography in where the plant cultivated/distributed [6][3]. However, it importance to keep update the genetic approaches in resolving of this genetic diversity. Recently, DNA fingerprint based on diversity of chloroplast genome was also developed in detection of variety for several species [7][8]. Several reports suggesting that, chloroplast genome have enough unique regions, as a tool in genetic diversity assay at various taxonomic levels.

Chloroplast is well known as organelle in plant cell which responsible in converting solar energy into carbohydrate through photosynthesis process. It is commonly accepted that the organelle are derived from green photosynthetic cyanobacteria, through a single endosymbiotic event which engulfed in more than one billion years ago. In current period, chloroplast genomes of land plants have only 50 protein-coding genes, in addition to 30 tRNA genes and full sets of rRNA genes, which involved in photosynthesis, gene expression, lipid metabolism and other physiology and development processes. It predicted that plant has about 0.15Mbp of chloroplast genome, and it is quite small versus 3 Mbp in cyanobacteria [9]. Chloroplast genome has less diversity, it was derivate through endosimbiotic and maternally inherited, hence their genome can only transferred through
females line. In fact, their polymorphism was shown that it reasonable enough as DNA markers, including in between sub species level of cacao [10].

Recently studies on Local Sulawesi cacao, reported that there are more than 20 clones can be detected in their genetic variation [11][12]. In local farm around Sigi, Poso, and Parigi districts, cacao seedling were done used parental seeds from natural breeding, and it showed high variation of cross breeding in various different clones. Therefore its importance to keep updating the DNA fingerprint approaches and need to be standardized, reliable method to genetically identify different clones/varieties and to estimate their overall genetic diversity. Here we introduce new method in this issue, by combining the genetic assay based on genetic marker of nuclear and chloroplast genomes. The result showed that this assay can be used in practice genetic diversity assay.

2. Materials and Method

Sample of 16 cacao were collected based on our previous works, it from local farm in districts of Sigi, Parigi and Poso of Central Sulawesi[5]. Total genome of each sample was purified from 100 mg leaf sample after grinded in liquid nitrogen and further purified by using Qiagen- DNeasy® Plant Mini Kit. PCR amplification was done using KOD tag polymerase enzyme (Toyobo) and set of primers: TCss4/mTcCIR 24 (tttggggtgatttcttctga- tctgtctgcttcttttggtga) and TCss5/mTcCIR 26 (gcattcatcaatcatc- gcactcaaagtt cat act ac) for Nuclear genome SSR markers, and rpl32 (gcgtattcgtaaaaatatttggaa) – trnL (ttctaaagcagctgtctac) for chloroplast genome marker. PCR product were purified and subjected to DNA sequencer.

Bioinformatic analysis was performed on CLC sequence Veizer ([https://www.qiagenbioinformatics.com/products/clc-sequence-viewer/](https://www.qiagenbioinformatics.com/products/clc-sequence-viewer/)).

3. Result and Discussion

Local cacaos in Sulawesi have very high genetic diversity. It due to natural breeding in the field, and it became root stock on the commercial practices. Recently up grade superior cultivars were introduced from high quality selected clones as entries in grafting. However, various clones and cultivars grow in farming area. High similarity of phenotypic performance caused the in-accurate identification of each clone/cultivar. Previously we proposed the application of SSR marker in DNA barcoding among these local cacao[5]. Based on the outcome data from previous observation, here we perform analysis on numerous clones, which collected from the same field, and further assayed genetically, either using same SSR primers or based on chloroplast marker.

As shown in Figure 1, this phylogenetic trees shows the high similarity among clones and several of each can be distinguished clearly, for instance, ICCRI (product of professional breeding) and local lines. Since this analysis based on nuclear chromosome, it mean the product of unification of half chromosome from male and female parental of each. Consequently, high genetic variations in outcome of those naturally cross breeding.

In next step, we also perform genetic variation chloroplast genome of the above clones. In this work we use set primer which detect the polymorphism in the region in between of rpl32 region as forward and rpL region as reverse primer. As predicted, there was high in conservation among the genome samples. This locus was suggested as marker for species identification[7], but fortunately, here we shows there was at least three groups can be made (Figure 2). Even though based on only one region of genome, it indicates this locus promising for identification in sub species level.
Figure 1. Phylogenetic relationship between different clones of local cacao. Unrooted trees were constructed by CLC sequence viewer based on DNA sequences data of PCR amplified fragment using two SSR marker primers of mTcCIR26/ klon COL_219 and mTcCIR24/ klon COL_510 locus.

Figure 2. Phylogenetic relationship between different clones of local cacao. Unrooted trees were constructed by CLC sequence viewer based on DNA sequences data of PCR amplified fragment using two chloropas marker primers of rpl32-trnL region. There are three groups (left, center and right) can be formed among tested clones.

Finally, we try to combine the sequence data of two SSR motives (mTcCIR26/ klon COL_219 and mTcCIR24/ klon COL_510), and sequence data of of rpl32-trnL regions. By applying alignment and contructing the phylogenetic analysis, here we can show more broad clades and groups among the samples (Figure 3.). This combination analysis showed more diversity on set of genome which derivated maternaly-non recombinant (chloroplast genome), and it derivated through male-female (uniparentally) inherited proceses. In addition, chloroplast genome identification have more benefit in less of cost, due to the high conserved regions of primer annaling site. This mean that set of primer can
be use universally to many different species and taxa, instead of SSR marker primer, which unique according to species.

Figure 3. Phylogenetic relationship between different clones of local cacao. Unrooted trees were constructed CLC sequence viewer based on DNA sequences data of PCR amplified fragment using both set of nuclear-SSR marker and chloroplas marker.

4. Conclusion

Based on this analysis, here we show the possibility of genetic analysis in combination between nuclear genome marker and chloroplast genome marker. This method is promising as a tool in determination work on genetic diversity, even in sub species level on broad range of taxa. For high degree of similarity/conservation, more specific chloroplast primers can be considered as additional marker.

References

[1] Russell J R, Hosein F, Johnson E, Waugh R and Powell W 1984 Genetic differentiation of cocoa (Theobroma cacao L) populations revealed by RAPD analysis no. 7 993, pp. 89–97
[2] Goran J A K N, Laurent V, Risterucci A M and Lanaud C 1994 Theobroma cacao L. using RFLP and RAPD markers vol. 73, no. February, pp. 589–597
[3] Motamayor J C et al. 2008 Geographic and genetic population differentiation of the Amazonian chocolate tree (Theobroma cacao L) PLoS One, vol. 3, no. 10, p. e3311, Jan
[4] Saunders J, Mischke S, Leamy E and Hemeida A 2004 Selection of international molecular standards for DNA fingerprinting of Theobroma cacao Theor. Appl. Genet., vol. 110, no. 1, pp. 41–7, Dec
[5] Suwastika I N, Aisyah N, Ishizaki Y, Basri Z and Shiina T 2015 Genotyping based on SSR marker on local cacao (Theobroma Cacao L.) from Central Sulawesi Procedia Environ. Sci., vol. 28, no. May 2016, pp. 88–91
[6] Efombang M I B, Sounigo O, Nyassé S and Cilas C 2006 Genetic diversity in cocoa germplasm of southern Cameroon revealed by simple sequences repeat (SSRS) markers vol. 5, no. August, pp. 1441–1449
[7] Dong W, Liu J, Yu J, Wang L and Zhou S 2012 Highly variable chloroplast markers for evaluating plant phyleny at low taxonomic levels and for DNA barcoding PLoS One, vol. 7, no. 4, p. e35071, Jan
[8] Chang C C et al. 2006 The chloroplast genome of Phalaenopsis aphrodite (Orchidaceae): Comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications Mol. Biol. Evol., vol. 23, no. 2, pp. 279–291
[9] Yagi Y and Shiina T 2014 Recent advances in the study of chloroplast gene expression and its evolution Front. Plant Sci., vol. 5, no. February, p. 61, Jan

[10] Yang J Y, Motilal L A, Dempewolf H, Maharaj K and Cronk Q C B 2011 Chloroplast microsatellite primers for cacao (Theobroma Cacao ) and other Malvaceae,” Am. J. Bot., pp. 372–374

[11] Rifka, Aisyah N, Muslimin and Suwastika I N 2014 Polimorfisme Klon Kakao (Theobroma cacao L.) Dari Perkebunan Rakyat Kabupaten Sigi vol. 3, no. December, pp. 269–277

[12] Rahmansyah I N S, Mutmainah and Muslimin 2014 Variasi Genetik Klon Kakao (Theobroma cacao L.) Di Desa Sausu Peore Kab. Parigi Moutong J. Nat. Sci., vol. 3, no. December, pp. 239–246

Acknowledgments
This work was supported by KEMENRISTEKDIKTI of Indonesian Government Research Grant of PKLN DIKTI to INS, MUS and ZB, and PUSN DIKTI to UM and INS.