SHORT COMMUNICATION

Journal of Oil Palm Research Vol. 29 (4) December 2017 p. 594 – 599
DOI: https://doi.org/10.21894/jopr.2017.0001

SHORT COMMUNICATION: TOWARDS DEVELOPMENT OF *Elaeis guineensis* CHROMOSOME-ARM SPECIFIC MARKERS AND THEIR UTILITY ACROSS THE *Elaeis* GENUS

NOORHARIZA MOHD ZAKI*; SINGH, R*; NORDIANA HANIM MOHD NOR*; MUHAMMAD AZWAN ZULKIFLI*; STEVEN W SMITH; TRUDE SCHWARZACHER‡; MADON, M* and HESLOP-HARRISON, J S‡**

ABSTRACT

The development of Elaeis guineensis *chromosome-arm specific markers is required for linking the genetic, sequence and chromosomal maps, valuable for comparative studies, hybridisation and breeding. Here, we aimed to develop strategies exploiting assembled short shotgun-sequence reads to identify markers for* in situ *hybridisation to identify arms and large-scale organisation of pseudo-chromosome 1. The conserved putative repetitive DNA sequence found through informatics analysis showed an extra intercalary band in one arm of pseudo-chromosome 1 in fluorescence the* in situ *hybridisation result. The results are a basis for establishing a North-South orientation of* E. guineensis *pseudo-chromosome. Furthermore, the ability of the newly developed markers for distinguishing both* Elaeis *species with* in situ *hybridisation showed their utility in identifying* Elaeis *hybrids.*

Keywords: *Elaeis guineensis*, chromosome markers, *in situ* hybridisation, repetitive DNA.

Date received: 3 May 2017; **Sent for revision:** 5 May 2017; **Received in final form:** 23 August 2017; **Accepted:** 7 September 2017.

INTRODUCTION

Chromosome identification in plants is important for the construction of cytogenetic maps. Moreover, the reliable chromosome identity can be used as reference for demarcating possible structural differences, both inter- and intra-species (Ren *et al*., 2012) and provide unique insight into genome organisation in the context of chromosomes (Heslop Harrison and Schwarzacher, 2011). For oil palm (*Elaeis guineensis*; 2n=32), with the exception of one chromosome pair, the sub-metacentric chromosomes show near-continuous variation in size, so they cannot be identified morphologically. Study and identification of oil palm chromosomes started as early as the 1940s, using a very limited range of techniques. Two approaches that have been applied by previous researchers are by measuring the chromosome length (Sato, 1949; Sharma and Sarkar, 1956; Madon *et al*., 1995) and by localising a few families of repetitive DNA to try to identify chromosomes in the *E. guineensis* karyotype (Castilho *et al*., 2000).

One strategy to construct a molecular cytogenetic map and determine identity of each chromosome is to localise directly the DNA sequences on metaphase or pachytene chromosomes by fluorescence *in situ* hybridisation (FISH) which gives reliable results regardless of the chromosome size. In the cereals, use of

* Malaysian Palm Oil Board, 6 Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia. E-mail: maria@mpob.gov.my

** Orion Genomics, 4041 Forest Park Avenue, Saint Louis, Missouri 63108, USA.

‡ Department of Genetics, University of Leicester, Leicester LE1 7RH, United Kingdom.

repetitive DNA sequences for *in situ* hybridisation has been valuable to identify each chromosome arm in breeding lines (Ali *et al*., 2016). For many species, including oil palm (Castilho *et al*., 2000), many repetitive DNA sequences have been isolated from genomic DNA using restriction-satellites, polymerase chain reaction (PCR) amplification, or selection from clone libraries. However, apart from the 45S which is located at the nucleolar organising region of the short chromosome and 5S rDNA hybridised on proximal of the longest chromosome (Castilho *et al*., 2000), *in situ* hybridisation with these repeats has not enabled chromosome arms to be identified uniquely in oil palm.

The oil palm genome paper (Singh *et al*., 2013) has re-assigned the 16 pairs of *E. guineensis* individual chromosomes into four groups; Group I: the largest chromosomes (hybridises to 5S rDNA at the proximal region); Group II: eight medium chromosomes; Group III: six small chromosomes and Group IV: small acrocentric chromosome (hybridise to 18S-25S rDNA). In the published 1.8 gigabase (Gb) oil palm sequence, Singh *et al*. (2013) reported approximately 57% of the genome was unassembled and is likely to include mainly repetitive DNA sequences. In this study, we report the first chromosome-arm specific marker for pseudo-chromosome 1 of oil palm *Elaeis guineensis* isolated from repetitive regions of *E. guineensis* which was found to be conserved in the genome. The utility of the developed marker showed a potential application for distinguishing both *Elaeis* species. The results are a starting point towards constructing a cytogenetic map. This also allows the association of the physical chromosomes of *E. guineensis* with the pseudo-chromosomes obtained through genome sequencing. Such a map is important for basic genomic research, comparative genomics, evolutionary studies, and further assists in understanding the inheritance of specific traits in oil palm.

MATERIALS AND METHOD

Plant Material

The oil palm (*E. guineensis*) materials used for developing the DNA probes were published by Singh *et al*. (2013) and are currently maintained at the MPOB Research Station, Kluang, Johor, Malaysia. Genomic DNA was extracted and purified from a spear leaf using the modified CTAB method as described by Doyle and Doyle (1990).

Universal Repetitive DNA Sequences

The 18S rDNA gene (1.7 kb) was amplified from *E. guineensis* genomic DNA using a primer pair published by Chang *et al*. (2010) (P1:5'-CGAACTGTGAAACTGCGAATGGC-3') and (P2:5'-TAGGAGCGACGGGCGGTGTG-3') and used as the probe to identify sites of the 45S rDNA. Clone pTa794, containing the 5S rRNA genes and intergenic spacer from wheat (Gerlach and Dyer, 1980), was used as the 5S rDNA probe. The synthetic oligomer $(CCCTAAA)_6$ was used to label the telomeres.

Identification of Repetitive DNA Sequences

A draft version of the P9 assembly [unpublished; an updated version of the assembly in Singh *et al*. (2013)] was analysed for repetitive elements with the REPET package (https://urgi.versailles.inra.fr/Projects/URGI-softwares/REPET) using the TEdenovo module which resulted in 0.54 GB of consensus repetitive sequence averaging 826 bp in length. These repetitive sequences were annotated using the TEannotate module, and were mapped back to the final P9 assembly.

Conversion of Repetitive Unique Sequence into FISH Probes

One primer pair flanking a putative repetitive sequence (Eg9CEN-length 354 bp) was designed; Eg9CEN-F: C C A T A T G G G T T G G T T G T C C and Eg9CEN-R: A C A G C G A C T C A T T C T T C T C C. The probe Eg9CEN was amplified from *E. guineensis* and *E. oleifera* genomic DNA by PCR using the following programme: 3 min 95°C, 30 cycles of (30 s 95°C, 30 s 58°C, 1 min 30 s 72°C), 1 min 72°C. PCR products were separated on a 1% agarose gel electrophoresis and isolated with the E.Z.N.A Gel Extraction Kit (Omega) as described by the manufacturer. Purified amplicon was stored at -20°C until use.

Probe Labelling

Synthetic oligomers and PCR products were labelled with digoxigenin-11-dUTP (Roche Diagnostic, Basel, Switzerland) or biotin-16-dUTP (Roche Diagnostic) using the BioPrime® Array CGH Labelling System (Invitrogen, California, USA) according to the manufacturer's instructions.

Preparation of Chromosome Spreads

Chromosome spreads preparation were made from root tips from both adults and seedlings of *E. guineensis* as well as adult *E. oleifera* by using technique modified from Madon *et al*. (1995) and Schwarzacher and Heslop-Harrison (2000). Root tips were pre-treated with 2 mM 8-hydroxyquinoline for 5-6 hr at 18°C and fixed in 3:1 ethanol: glacial acetic acid (v/v) and stored in 70% ethanol at 4°C until further use. The root tips were washed several times with citric acid-citrate buffer, pH 4.6 and digested at 37°C for up to 4 hr in enzyme

solutions containing 2%-4% (w/v) cellulase (Sigma C1184; final concentration 10-20 U ml^{-1}), 0.2% (w/v) 'Onozuka' RS cellulase (final concentration 10 U ml^{-1}), 3% (v/v) pectinase (Sigma P4716 from *Aspergillus niger*; solution in 40% glycerol, final concentration 15-20 U ml^{-1}) in citric acid-citrate buffer. Mitotic chromosomes were spread by squashing and heating onto pre-cleaned glass slide in a drop of 45% or 60% acetic acid under a cover slip, frozen before flicking off the cover slip, and left to air-dry before using for FISH.

Fluorescence *in situ* Hybridisation

In situ hybridisation was performed according to Schwarzacher and Heslop-Harrison (2000) with minor modifications. A total of 40 μl probe was applied per slide, containing 50% (v/v) formamide, 20% (w/v) dextran sulphate, 2X saline sodium citrate (SSC, 0.3 M NaCl, 0.03 M sodium citrate), 0.05 μg of salmon sperm DNA, 0.25% (w/v) sodium dodecyl sulphate, 0.25 mM ethylenediamine-tetraacetic acid and 25-100 ng probe. The probe mixture was denatured for 10 min at 80°C and immediately transferred to ice. Probe and chromosomal DNA were denatured together on a heated block (Thermo Fisher Scientific) at 72°C for 5 min under plastic cover slips, allowed to cool to hybridisation temperature of 37°C overnight. A series of post-hybridisation washes were carried out with 2X SSC and 0.1X SSC at 42°C. Hybridisation sites for biotin and digoxigenin-labeled probes were detected with 2 μg ml^{-1} streptavidin conjugated to Alexa 594 (Molecular Probes, Thermo Fisher Scientific) and 4 μg ml^{-1} anti-digoxigenin conjugated to fluorescein isothiocyanate (FITC, Roche Diagnostic) respectively. DAPI (4,6-diamidino-2-phenylindole) in CITIFLUOR AF1 (Chem Lab,) antifade solution was used to counterstain the chromosomes. At least two slides with 15 high quality metaphases were hybridised and analysed for each probe and species combination.

Image Acquisition

Slides were examined using a Nikon Eclipse 80i fluorescent microscope. Three Nikon filters were used for the observation, UV-2E/C (emission at 435-485) for DAPI, B-2E/C (emission at 515-555) for fluorescein and G-2E/C (emission at 590-650) for Alexa 594. Images were acquired with Nikon DS-Qi1 Digital camera and NIS elements AR, version 3.2 software. The individual channels were pseudo-coloured to visualise the sites of probe hybridisation. The images were processed using Adobe Photoshop CS5 software (Adobe System Inc., http://www.adobe.com) using cropping and functions that affect the whole image equally.

RESULTS AND DISCUSSION

A Chromosome-arm Specific Marker for *E. guineensis* Pseudo-chromosome 1

Analysis of the *E. guineensis* P9 sequence assembly identified a region (named Eg9CEN) of 354 bp with 1703 copies per genome. The amplification within the fragment by PCR showed a single band amplicon with expected size (c.330 bp) (*Figure 1*). *In situ* hybridisation with Eg9CEN showed a broad centromeric hybridisation signal observed on all 32 chromosomes, notably with a conserved and distinct intercalary site of hybridisation detected on the opposite arm to the 5S rDNA site (*Figure 2a*). Hybridisation with Eg9CEN and the telomere (CCCTAAA) (*Figure 2b*) also showed the intercalary site distal on the longest chromosome.

The Eg9CEN sequence, derived from the bioinformatics analysis, showed 95% homology to five of the repetitive DNA sequences isolated by screening clones for repetitive sequences from genomic DNA, pEgKB23, pEgKB15, pEgKB1, pEgKB14, and pEgKB19 families. In the consensus region, there were no indels between any of the five sequences, and no unique single nucleotide

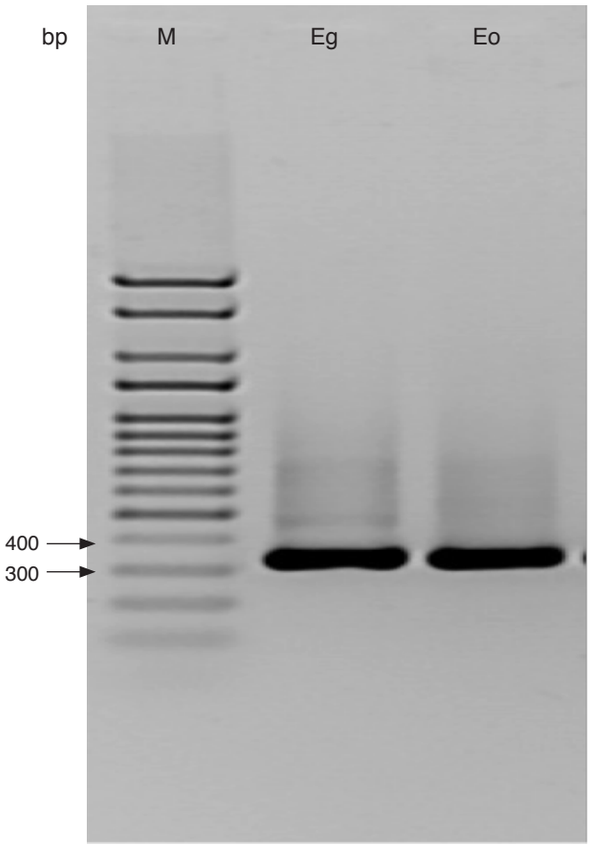

Figure 1. Polymerase chain reaction (PCR) amplification of Eg9CEN from genomic DNA of E. guineensis *(Eg) and* E. oleifera *(Eo). A similar band (c. 330 bp) is amplified by the primer pair from both species.*

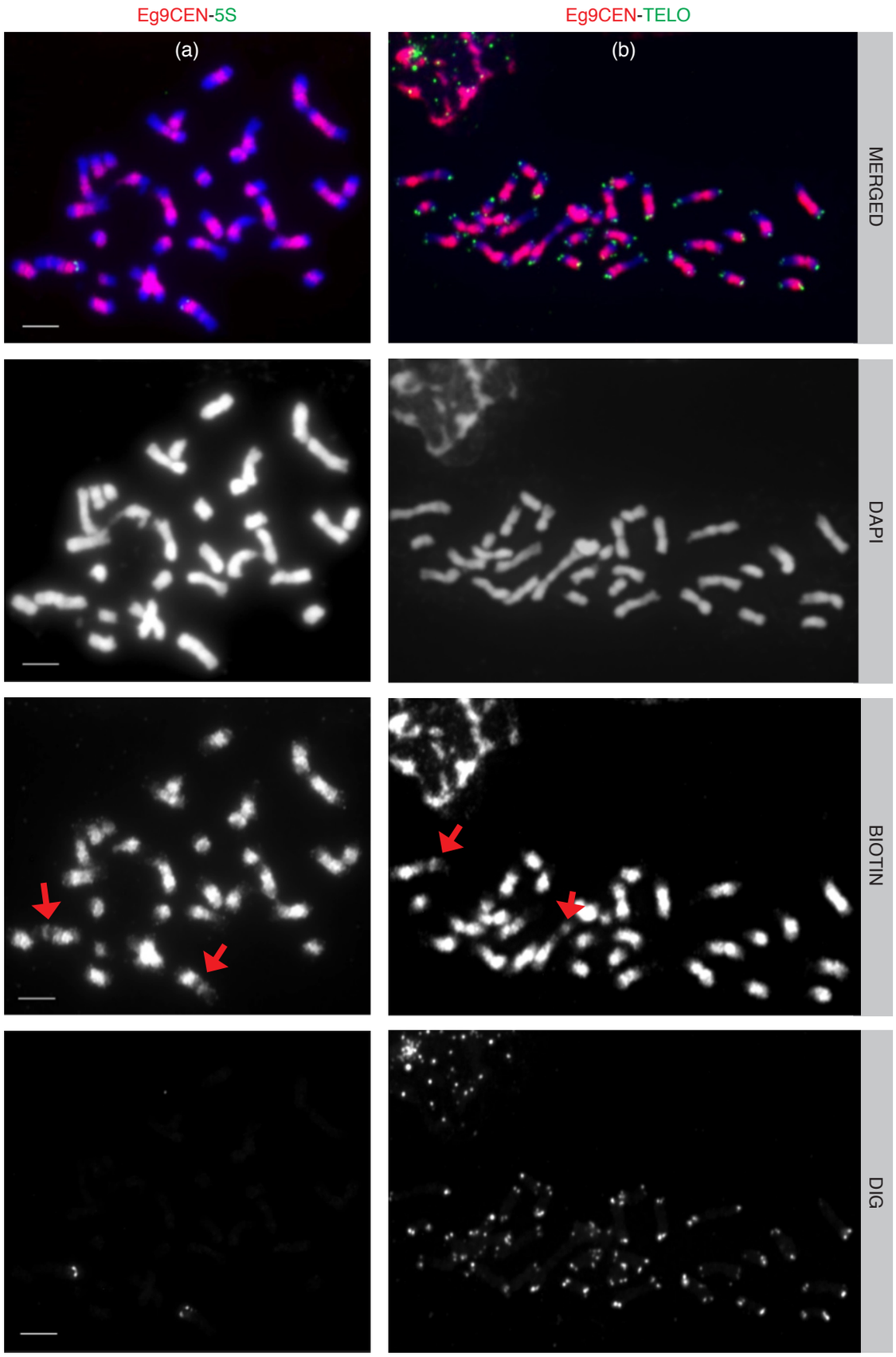

Figure 2. In situ *hybridisation on metaphase chromosomes of* Elaeis guineensis *(2n=32). Eg9CEN (red signal) probe is hybridised with the 5S rDNA (green signal) probe (a) and a synthetic DNA telomere (green signal) probe (b). DAPI counterstaining shows the 32 metaphase chromosomes. Red arrow shows the intercalary site of Eg9CEN.*

polymorphisms in Eg9CEN, although pEgKB23 was some 130 bp longer.

Both the 5S rDNA and the Eg9CEN intercalary site enabled identification of pseudo-chromosome 1, and its North and South or long and short arm orientation. Thus, the repetitive probe identification strategy from the whole genome assembly helps to validate the current linkage map and assembly.

Comparative Analysis to Differentiate Two *Elaeis* Species

PCR amplification of the Eg9CEN unique repetitive probe produced a similar single band, length (c. 330 bp) in both *E. guineensis* and *E. oleifera* (*Figure 1*). However, the localisation of Eg9CEN by *in situ* hybridisation on the *E. oleifera* physical chromosome along with the synthetic telomere, 18S-25S rDNA and 5S rDNA showed a contrasting distribution between the two *Elaeis* species (*Figures 3* and *4*). Both rDNA hybridisation sites on *E. oleifera* were similar as reported for *E. guineensis* (Castilho *et al*., 2000), however, it was noticeable that 5S hybridisation signal on *E. oleifera* chromosome is much stronger compared to *E. guineensis*. As for Eg9CEN, the probe that showed a unique hybridisation pattern on pseudo-chromosome 1 of *E. guineensis*, exhibited a dispersed distribution over the whole of the 16 *E. oleifera* chromosomes. In 1999, Madon *et al*. successfully discriminated the two sets of chromosomes in interspecific hybrid palms via genomic *in situ* hybridisation (GISH) approach, where a total genomic DNA was used as a probe. The repetitive DNA sequence of Eg9CEN established in this study is most likely a component of the genomic DNA probe that allows *Elaeis* discrimination. Moreover, the defined nature of the Eg9CEN probe, showing substantial differences in the hybridisation between the species, makes it a valuable species-specific marker for identifying parental or ancestral chromosomes in hybrids as well as looking for introgression, or identifying any inter-specific translocation events in oil palm hybrid derivatives.

CONCLUSION

Various studies have identified tandemly repeated and dispersed transposon-related repetitive elements from both genomic DNA and informatic studies (Biscotti *et al*., 2015). This study shows that informatic analysis allows sequence-independent identification of abundant repetitive DNA sequences whole genome assemblies. We were able to identify an abundant sequence, with no known homology to known repetitive element classes, that acted as a chromosome-arm specific marker for *Elaeis guineensis* pseudo-chromosome 1. Notably, the Eg9CEN sequence, a consensus extracted from the bioinformatics analysis rather than being a single cloned sequences, showed a more uniform hybridisation pattern than the 95% similar pEgKB23 clone published by Castilho *et al*. (2000). Previously, the intercalary site on pseudo-chromosome 1 was not distinguished when the clones were used in the *in situ* hybridisation, despite the similarity.

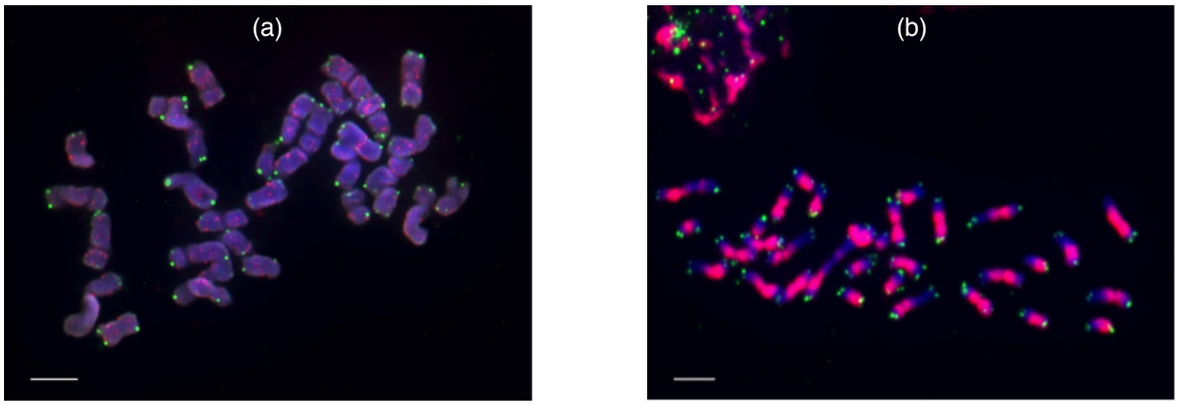

Figure 3. In situ *hybridisation of Eg9CEN (red signal) on* E. oleifera *(2n=32) (a) and* E. guineensis *(b) chromosomes. Different hybridisation patterns of Eg9CEN are observed in these two* Elaeis *species.*

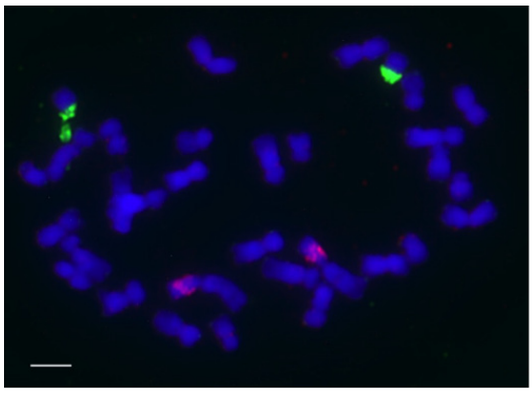

Figure 4. The 5S (red) and 45S rDNA (green) signals observed at a single pair of chromosomes on different E. oleifera *metaphase chromosomes.*

Several authors have discussed the occurrence of either chromosome-specific or region-specific variants of repetitive DNA sequences (Kuhn *et al*., 2007), and we have considered the value of making a 'consensus' sequences which may never exist in a genome. Given the similarity (95%) between the sequences, even with highly stringent hybridisation and wash conditions, *in situ* (or indeed Southern) hybridisation is unlikely to distinguish robustly between the probes, so it is interesting that the cloned probes gave a less specific signal.

The strategy to generate repetitive probes with characteristic chromosomal hybridisation patterns provides a basis for establishing the North-South orientation of the *E. guineensis* pseudo-chromosomes. It is likely that extensions of the bioinformatic analysis based on high-coverage sequence assemblies can be extended to many if not all other chromosome arms. Furthermore, the utility of the developed marker showed a potential application for distinguishing chromosome in the interspecific hybrid *E. oleifera* x *E. guineensis*. This finding is important for basic genomic research comparative genomics, evolutionary studies and assists in understanding the inheritance of specific traits in oil palm.

ACKNOWLEDGEMENT

The authors would like to thank the Director-General of MPOB for permission to publish this article and research assistants of MPOB involved in this work, particularly Zainab Anip and Mohd Razali Mohd Nor. The first author would like to thank MPOB for Ph.D fellowship.

REFERENCES

ALI, N; HESLOP-HARRISON, J S; AHMAD, H; GRAYBOSCH, R A; HEIN, G L and SCHWARZACHER, T (2016). Introgression of chromosome segments from multiple alien species in wheat breeding lines with wheat streak mosaic virus resistance. *Heredity, 117(2)*: 114-123.

BISCOTTI, M A; OLMO, E and HESLOP-HARRISON, J S (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Research, 23(3)*: 415-420.

CASTILHO, A; VERSHININ, A V and HESLOP-HARRISON, J S (2000). Repetitive DNA and the chromosomes in the genome of oil palm (*Elaeis guineensis*). *Annals of Botany, 85(6)*: 837-844.

CHANG, K D; FANG, S A; CHANG, F C and CHUNG, M C (2010). Chromosomal conservation and sequence diversity of ribosomal RNA genes of two distant *Oryza* species. *Genomics, 96(3)*: 181-90.

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

GERLACH, W L and DYER, T A (1980). Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. *Nucleic Acids Res., 8*: 4851-4865.

HESLOP-HARRISON, J S and SCHWARZACHER, T (2011). Organisation of the plant genome in chromosomes. *Plant J., 66(1)*: 18-33.

KUHN, G C S; FRANCO, F F; MANFRIN, M H; MOREIRA-FILHO, O and SENE, F M (2007). Low rates of homogenization of the DBC-150 satellite DNA family restricted to a single pair of micro -chromosomes in species from the *Drosophila buzzatii* cluster. *Chromosome Research, 15*: 457-46.

MADON, M; CLYDE, M M and CHEAH, S C (1999). Application of genomic *in situ* hybridization (GISH) on *Elaeis* hybrids. *J. Oil Palm Res*. (*Special Issue*): 74-80.

MADON, M; CLYDE, M M and CHEAH, S C (1995). Cytological analysis of *Elaeis guineensis* (*tenera*) chromosomes. *Elaeis, 7(2)*: 122-134.

REN, Y; ZHAO, H; KOU, Q; JIANG, J; GUO, S; ZHANG, H; HOU, W; ZOU, X; SUN, H; GONG, G; LEVI, A and XU, Y (2012). A high resolution genetic map anchoring scaffolds of the sequenced watermelon genome. *PLoS One, 7(1)*: e29453.

SATO, D (1949). Karyotype alteration and phylogeny VI; karyotype analysis in Palmae. *Cytologia, 12*: 174-186.

SCHWARZACHER, T and HESLOP-HARRISON, J S (2000). *Practical* in situ *Hybridization*. BIOS Scientific Publishers Ltd, Abindgdon, UK.

SHARMA, A K and SARKAR, S K (1956). Cytology of different species of palms and its bearings on the solutions of the problems of phylogeny and speciation. *Genetica, 28*: 361-488.

SINGH, R; ONG-ABDULLAH, M; LOW, E T L; MANAF, M A A; ROSLI, R; NOOKIAH, R; OOI, C L; OOI, S E; CHAN, K L; HALIM, M A; AZIZI, N; NAGAPPAN, J; BACHER B; LAKEY, N; SMITH, S W; HE, D; HOGAN, M; BUDIMAN, M A; LEE, E K; SALLE, R D; KUDRNA, D; GOICOECHEA, J L; WING, R A; WILSON, R L; FULTON, R S; ORDWAY, J M; MARTIENSSEN, R A and SAMBANTHAMURTHI, R (2013). Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature. 500 (7462)*: 335-339.