

Descifrar la información del genoma de la palma de aceite



Rajinder Singh, PhD

Jefe de la Unidad de Genómica
del Malaysian Palm Oil Board
Malasia

rajinder@mpob.gov.my

Leslie Low Eng Ti
Maria Madon
Meilina Ong Abdullah
Leslie Ooi Cheng- Li
Chan Pek Lan
Ravigadev
Sambanthamurthi

Junta Malaya de Cultivadores de
Palma de Aceite, No. 6 Persiaran
Institusi, Bandar Baru Bangi,
43000 Kajang
Selangor, Malasia

Palabras CLAVE

Genoma, ADN, marcadores
moleculares, marcadores de
secuencias expresadas (EST),
mejoramiento genético

Genome, DNA, molecular markers,
Expressed Sequence Tags (EST),
breedin

Traducido por Fedepalma

Versión original en inglés
en el Centro de Información
de Fedepalma

Deciphering the Oil Palm Genome Information



Resumen

(Un modesto comienzo para usar la tecnología genómica en la palma de aceite se realizó con la secuenciación de clones de ADN complementario (ADNC) como marcadores de secuencias expresadas (EST, por su sigla en inglés). Con este método de descubrimiento rápido de genes, el MPOB generó más de 17.000 EST para la palma de aceite. Los EST demostraron ser valiosos como fuente de marcadores moleculares, con aplicación potencial para el mejoramiento genético de la palma de aceite y el cultivo de tejidos. En el esfuerzo para asignar funciones a los genes descubiertos en la palma de aceite, se inició también un programa de microarreglos de ADN. El EST y otras secuencias clonadas del gen de la palma de aceite se imprimieron en microarreglos o chips de ADN, y se estudiaron sus posibles funciones en diversos tejidos y fases de desarrollo, especialmente durante el cultivo de tejidos de la palma de aceite. Para sobrepasar a las deficiencias asociadas con los EST, la MPOB también inició un proyecto para generar la primera colección completa de secuencias de genes de un cultivo oleaginoso mediante la secuenciación de las regiones hipometiladas del genoma de la palma de aceite, empleando la técnica de filtración de metilación GeneThresher™. Más de 400.000 secuencias se generaron de nueve palmas individuales. También se minaron y se emplearon marcadores microsatelitales (repeticiones de secuencia simple, SSR, por su sigla en inglés) y polimorfismos de un solo nucleótido (SNP, por su sigla en inglés) para construir

mapas genéticos de alta densidad de la palma de aceite. A fin de obtener una cobertura más completa del genoma, el MPOB emprendió la secuenciación del genoma y del transcriptoma de la palma de aceite mediante el empleo de tecnologías de secuenciación de próxima generación (NGS, por su sigla en inglés). Asimismo se continúa activamente con la investigación citogenética molecular para obtener un panorama completo de la estructura y organización del genoma de la palma de aceite.

Abstract

A modest start to using genomics technology in oil palm was made with the sequencing of complementary DNA (cDNAs) clones as expressed sequence tags (ESTs). This method of rapid gene discovery led to more than 17.000 ESTs generated for oil palm by MPOB. The ESTs proved to be valuable as a source of molecular markers with potential application in oil palm breeding and tissue culture. In the effort to assign functions to the genes discovered in oil palm, a DNA microarray programme was also initiated. The ESTs and other cloned oil palm gene sequences were printed on DNA microarrays, or chips, and their possible functions in various tissues and developmental phases especially during oil palm tissue culture were studied.

To overcome the shortcomings associated with ESTs, MPOB also launched an effort to generate the first comprehensive collection of gene sequences of an oil crop by sequencing the hypomethylated regions of the oil palm genome using the GeneThresher™ methylation filtration technique. Over 400.000 sequences were generated from nine individual palms. Simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) markers were also mined and used to construct high density genetic maps for oil palm. In order to obtain a more comprehensive coverage of the genome, MPOB has embarked on sequencing of the oil palm genome and transcriptome using the next generation sequencing (NGS) technologies. Molecular cytogenetics research is also being actively pursued to obtain a complete picture of the oil palm genome structure and organization.

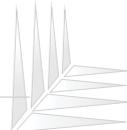


Introducción

La palma de aceite es un importante producto agrícola que contribuye de manera significativa a la producción de aceites y grasas en el ámbito mundial. La industria ha tenido un impacto importante en el desarrollo socioeconómico de los países productores en el Sudeste Asiático, África y Suramérica. El progreso en la investigación y desarrollo (I+D) ha sido fundamental para el éxito de la industria de la palma de aceite. Se han realizado mejoras, entre otras, en el mejoramiento genético, las prácticas agronómicas, los rendimientos, el control de plagas y enfermedades, el incremento de la productividad de la mano de obra, la eficiencia de las plantas de beneficio y el manejo general de plantaciones (Jalani, 1998).

Sin embargo, la I+D, especialmente en el mejoramiento genético para variedades nuevas y mejoradas, constituye un reto para este cultivo, que tiene un ciclo de mejoramiento largo (10-12 años) y requerimientos grandes de tierra para los ensayos en campo (Mayes *et al.*, 1997). Es por este motivo que la industria aún enfrenta muchos desafíos para avanzar, especialmente para reducir las brechas de rendimiento entre las parcelas experimentales, las plantaciones nacionales y las pequeñas explotaciones.

Los desarrollos en el ámbito de la biotecnología durante las tres últimas décadas han brindado oportunidades importantes para acelerar el mejoramiento de la palma de aceite. El advenimiento de las técnicas de ADN recombinante



hizo posible clonar fragmentos de ADN como sondas o marcadores, lo que proporcionó herramientas para aplicar en estudios genéticos y de mejoramiento. El posterior desarrollo de la reacción en cadena de la polimerasa (PCR, por su sigla en inglés) resultó en una mejora sistemática para la tecnología de marcadores moleculares, lo que llevó a una aplicación más amplia de los marcadores en la mayoría de los sistemas de cultivos. El advenimiento de la secuenciación del ADN a finales de la década de 1970, y su posterior automatización, sin duda ha tenido también una gran repercusión.

Con el rápido desarrollo de esta tecnología, los investigadores ya no están limitados a secuenciar solo el pequeño número de genes que directamente afectan su área de trabajo. Los métodos de secuenciación de alto rendimiento desarrollados, especialmente la secuenciación de nueva generación (NGS), ha posibilitado la secuenciación de genomas enteros a un costo razonable. Este trabajo dará una idea del esfuerzo de la Junta Malasia de Cultivadores de Palma de Aceite (MPOB) en el uso de herramientas de la biotecnología moderna para entender la composición genética de la palma de aceite, con miras a usar esta información para mejorar la eficiencia del mejoramiento genético y del cultivo de tejidos de palma de aceite.

Generación de etiquetas de secuencias expresadas (EST)

Las etiquetas de secuencias expresadas (EST), se generan mediante la secuenciación parcial de clones complementarios de ADN (ADNC) seleccionados de manera aleatoria. Adam *et al.* (1991) demostraron esto por primera vez y mostraron que de 150 a 400 bases eran suficientes para proporcionar una pareja con secuencias en las bases de datos públicas y proveer una identificación preliminar de los ADNC. La metodología es una forma económica de entender la expresión génica (Fields, 1994) y para el descubrimiento de nuevos genes (Verdun *et al.*, 1998). Antes de los proyectos de secuenciación de transcriptomas y genomas completos, la estrategia de las EST

fue eficaz en la identificación de nuevos genes, como se demuestra en *Arabidopsis* (Höfte *et al.*, 1993), arroz (Sasaki *et al.*, 1994) y soya (Shoemaker *et al.*, 2002).

A mediados de la década de 1990, el MPOB emprendió un programa para generar EST de palma de aceite como medio para identificar y marcar los genes (Cheah, 1999). En ese tiempo existía mínima información sobre los genes de palma de aceite en las bases de datos públicas, y el modesto inicio del MPOB fue ayudar a mejorar la base de conocimientos sobre la porción expresada del genoma de palma de aceite. Inicialmente, se secuenciaron cerca de 6.500 clones de ADNC a partir de diversos tejidos, tales como el mesocarpio, la almendra, la flor y la raíz (Singh *et al.*, 2001).

Posteriormente, en la medida en que la importancia de la palma de aceite clonal se hizo más evidente (Ong-Abdullah y Ooi, 2006 y 2007), el programa de EST T se concentró en entender la expresión génica durante el cultivo de tejidos. A este respecto, un total de 17.599 EST se generaron a partir de callos embriogénicos, no embriogénicos y tejidos embrioides (Low *et al.*, 2008). Se analizaron los genes para identificar cerca de 9.560 unigenes a los cuales se les asignaron funciones basadas en similitud y anotaciones de ontología génica. Más importante aún, mediante el análisis *Northern* digital, también se identificaron en el estudio anterior clases-grupos de genes diferencialmente expresados en callos embriogénicos y no embriogénicos.

Las etiquetas de secuencia expresada (EST) como fuente de marcadores moleculares

Además de proporcionar un mejor entendimiento de los genes involucrados en los diferentes tejidos y etapas de desarrollo de la palma de aceite, las EST también resultaron ser una valiosa fuente de marcadores moleculares. Los clones de ADNC fueron una excelente fuente de marcadores de longitud de fragmentos de restricción (RFLP, por su sigla en inglés), tanto para la huella genética como para la cartografía genética.



Figura 1. Evaluación de la fidelidad de un cruce autofecundado controlado. Las palmas que tienen bandas (indicadas por ←) que no se encuentran en la palma parental (P) probablemente son ilegítimas debido a la contaminación por polen.

En experimentos de la huella genética, las sondas de ADNC utilizadas como marcadores pudieron identificar fácilmente las palmas “ilegítimas” en un cruce controlado (Figura 1). Esta es una sencilla pero importante aplicación de los marcadores en el mejoramiento de palma de aceite para evaluar la fidelidad de los cruces controlados sembrados en el campo. Las sondas ADNC-RFLP resultaron especialmente útiles para limpiar la población para el mapeo genético al identificar las palmas ilegítimas (Singh *et al.*, 2008a). Más importante aún, los clones de ADNC mostraron suficiente polimorfismo para ayudar a construir un mapa genético para una palma Ténera autofecundada, lo que proporcionó el primer mapa publicado de palma de aceite con marcadores específicos por genes (Singh *et al.*, 2008a).

La emergente popularidad de los ensayos basados en PCR impulsó al MPOB para analizar detalladamente la colección de EST en busca de marcadores microsatélites (SSR). Con base en el análisis de cerca de 9.560 unigenes, Low *et al.*, (2008) identificaron 648 microsatélites

no redundantes. Entre los motivos de repeticiones de dinucleótidos identificados, las repeticiones AG/CT fueron las más comunes en el conjunto de datos, mientras que AAG/CTT fue la repetición de trinucleótidos de palma de aceite que ocurrió con más frecuencia. Singh *et al.*, (2008b) y Ting *et al.*, (2010) demostraron la eficacia de los marcadores EST-SSR para evaluar la diversidad genética del germoplasma de palma de aceite y también mostraron su transferibilidad entre las especies y taxones de palma.

Es interesante notar que los marcadores EST-SSR también pueden servir como una herramienta de control de calidad en el proceso del cultivo de tejidos de palma de aceite. Algunas veces son inevitables los errores humanos que hacen que los cultivos se mezclen, dado el gran número de cultivos que se manejan a diario en el laboratorio. Estos errores, sin embargo, deben rectificarse y los cebadores SSR empleados pudieron distinguir entre los *ramets* y los “solitarios (*rogues*)” cuando sus perfiles SSR se compararon con el *ortet* (Figura 2).

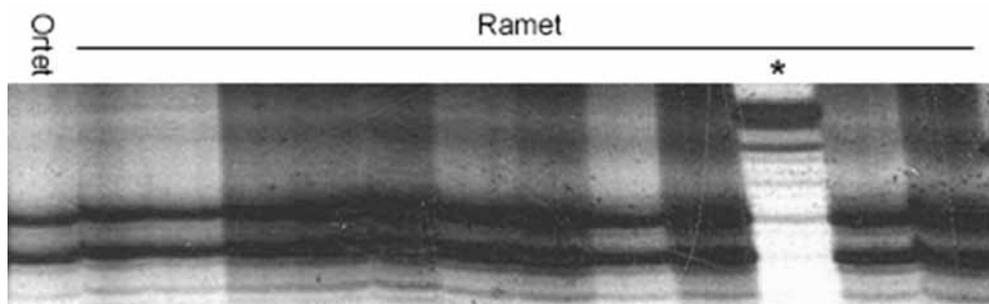


Figura 2. Detección de mezcla de cultivos. El símbolo * denota una muestra que no se originó del *ortet*.



Microarreglos de ADN

La disponibilidad de un gran número de secuencias EST generadas de diversos tejidos de palma de aceite brindaron la oportunidad de establecer un programa de microarreglos de ADN en el año 2000 (Ooi *et al.*, 2001). La tecnología de arreglos de ADN fue desarrollada para explorar la función putativa de las EST disponibles mediante un análisis de expresión global. El arreglo de ADN desarrollado en el MPOB tuvo 6.714 fragmentos de genes y 363 controles impresos en láminas de vidrio (Singh *et al.*, 2011). Como la mayoría de los fragmentos de genes fueron identificados a partir de las diferentes etapas de desarrollo del cultivo de tejidos, el arreglo resultó útil para el descubrimiento de genes potencialmente ligados a la amenidad (*amenity*) del cultivo de tejidos y la anomalía floral (Chan *et al.*, 2005; Low *et al.*, 2006; Sambanthamurthi *et al.*, 2009).

Los estudios de expresión mediante este arreglo han mejorado nuestro entendimiento de los eventos moleculares asociados con la calogénesis y embriogénesis del cultivo de tejidos de palma de aceite. Se identificaron para investigaciones posteriores un número de genes que exhiben perfiles de expresión interesantes durante las primeras etapas de desarrollo de los explantes (experimento de curso temporal), y también en las etapas posteriores de desarrollo del cultivo de tejidos, tales como callo embriogénico (CE), callo no embriogénico (CNE), poliembrioides (EMB) y brotes.

En el experimento de curso temporal, la perfilación de la expresión en las primeras etapas de los explantes de hojas del cultivo de tejidos en el Día 0, Día 7, Día 14 y Día 21 resultó en la identificación del 7 al 10% de los genes diferencialmente expresados en cada punto de tiempo (Low *et al.*, 2006). La expresión de genes, tales como la proteína de choque térmico, una proteína relacionada con la patogénesis, la MAP quinasa 5 y una proteína de floración temprana, aumentó después de siete días en el medio de inducción de callos (Sambanthamurthi *et al.*, 2009).

El estudio de expresión se extendió a las etapas posteriores de desarrollo del cultivo de

tejidos. Un grupo de genes se preseleccionó a partir del análisis de agrupamiento jerárquico de los datos del microarreglo. Estos genes podrían usarse para diferenciar los CNE de los otros materiales del cultivo de tejidos (Low *et al.*, 2006). Se están realizando esfuerzos para validar el patrón de expresión de estos genes mediante la PCR cuantitativa en tiempo real.

Secuenciación del espacio génico de palma de aceite

A pesar de su popularidad y rápido éxito, un exhaustivo análisis de los datos secuenciados reveló que la secuenciación de EST es de valor limitado después de que las primeras miles de secuencias se han obtenido, y solo puede alcanzar una cobertura parcial (menos del 50% de todos los genes) incluso cuando se generan millones de secuencias de EST. Además, las secuencias reguladoras que controlan la actividad de los genes no se muestrean en absoluto por este procedimiento.

Así pues, el MPOB estableció que, además del esfuerzo de las EST, era necesario acelerar la secuenciación de los genes de palma de aceite a través de un método más rápido y eficaz. En este sentido, el MPOB empleó una tecnología de secuenciación conocida como *GeneThresher*[™], que permite la selección preferencial de esas regiones del genoma que codifican los genes y su regiones reguladoras (Rabinowicz *et al.*, 1999). Esta técnica elimina las secuencias repetitivas, mientras el espacio génico hipometilado mucho más pequeño es secuenciado, lo que lleva a un descubrimiento génico completo a un costo muy reducido.

El proyecto generó alrededor de 400.000 secuencias de fragmentos genómicos a partir de genotecas filtradas *GeneThresher*[®] o de genomas completos (no filtradas). Las genotecas no filtradas se crearon para determinar la eficacia de las genotecas filtradas *GeneThresher*[®]. Luego se analizaron detalladamente el ensamble y las secuencias individuales en busca de microsatélites (SSR) y polimorfismos de un solo nucleótido (SNP). A la fecha, se han identificado cerca de 86.000 secuencias que contienen SSR y 29.913 SNP de buena calidad.

Los marcadores SNP y SSR se están explotando como marcadores codominantes para su asignación en los mapas genéticos existentes. Los mapas de alta densidad han permitido la identificación de marcadores ligados a importantes caracteres como el color del fruto. Las secuencias también ofrecen una gran oportunidad para que el MPOB realice investigaciones en biología *in silico* para identificar a los promotores de ADNC parciales, ADNC de longitud completa y genes de interés (Ooi, Feshah y Ong-Abdullah, 2010).

Hacia la secuenciación de genomas completos

La introducción de la tecnología de secuenciación de nueva generación (NGS) y con ella una reducción sustancial en los costos de secuenciación, han permitido secuenciar genomas completos. El MPOB y sus colaboradores (Orion Genomics, Universidad de Washington en San Luis, Mogene (todas en Estados Unidos) y Macrogen (Corea del Sur) emprendieron un programa para secuenciar tres genomas de palma de aceite mediante la tecnología NGS.

La iniciativa del MPOB añade valor a los esfuerzos anunciados por algunas empresas privadas. Además de los tres genomas, el MPOB y sus colaboradores también están secuenciando los transcriptomas de más de treinta tejidos de palma de aceite.

No hay duda de que el MPOB tendrá en su poder una colección completa de información del genoma de palma de aceite que se explorará para identificar la ubicación de los genes que controlan caracteres agronómicos importantes como el color del fruto. La explotación de los datos del genoma completo constituirá la base para el desarrollo de variedades nuevas y mejoradas de palma de aceite.

Citogenética molecular

La investigación en el área de la citogenética molecular se inició en el MPOB para facilitar el programa de mejoramiento de híbridos. La técnica de hibridación *in situ* genómica (GISH, por su sigla en inglés) se desarrolló para distinguir

los juegos de cromosomas tanto de *E. oleifera* (o) como de *E. guineensis* (g) en los híbridos OxG (Madon, Clyde y Cheah, 1999). En estos híbridos, igual número de cromosomas de *E. oleifera* y *E. guineensis* (dieciséis cada uno) se heredan. Las progenies de retrocruzamiento, en cambio, pueden tener un genoma con todos los 32 cromosomas originalmente de *E. guineensis* ($2n=32$) o un número variable de cromosomas de *E. oleifera* para un máximo de dieciséis.

La variabilidad en el número de cromosomas parentales se evalúa mediante la técnica GISH, seguido de la selección de palmas adecuadas para entrecruzamiento posterior. Esto podría ayudar a reducir el número de ciclos de entrecruzamiento necesarios para introducir los caracteres deseables de *E.oleifera* a *E. guineensis*. Un número de progenies de retrocruzamiento que se han identificado como de alto rendimiento, se están tamizando para seleccionar palmas adecuadas para posterior entrecruzamiento con el parental *E. guineensis*.

El conocimiento sobre el contenido de ADN es también importante en el ámbito de la investigación genómica. Mediante citometría de flujo y fronda -1, Madon *et al.* (2008) establecieron en $3,86 \pm 0,26\text{pg}$ el contenido promedio de ADN 2C de *E. guineensis*. Aunque se reportó que el ADN de una *E. oleifera* procedente de Surinam era cerca de la mitad del de *E. guineensis*, esto probablemente se debió a dificultades técnicas encontradas durante el procesamiento de las muestras de *E. oleifera* de Surinam. La evidencia reciente de los esfuerzos de secuenciación apuntan al hecho de que el tamaño del genoma de *E. oleifera* no es muy diferente del de *E. guineensis*.

Conclusión

Hasta hace poco, los científicos agrícolas, incluidos los mejoradores de palma de aceite y los agrónomos, han utilizado principalmente el fitomejoramiento híbrido convencional para mejorar continuamente el rendimiento. Aunque los resultados han sido sobresalientes, se están exigiendo beneficios adicionales en



la productividad agrícola a un ritmo cada vez más rápido, debido a los continuos cambios en las prácticas agrícolas y en las preferencias de los consumidores.

Las herramientas genómicas modernas permiten alcanzar amplios conocimientos sobre los caracteres biológicos de una especie que puede explotarse para mejorar la productividad y obtener mejor calidad de los productos. El desarrollo de marcadores moleculares, los cuales no son específicos por tejidos ni por fases, servirá de base para el desarrollo de métodos objetivos de selección e introgresión asistida por marcadores. La disponibilidad de sondas para caracteres de interés en palma de aceite permitirá a los mejoradores seleccionar en la etapa de vivero, y así reducir los cos-

tos y el periodo de tiempo de los programas de mejoramiento.

Las herramientas citogenéticas como la GISH ayudarán al programa de mejoramiento de híbridos, mediante la eliminación de segmentos cromosómicos no deseados del parental dominante o la selección de más regiones cromosómicas de los parentales recurrentes. La exitosa secuenciación del genoma de palma de aceite tiene un significado adicional, ya que proporciona una base sólida para el mejoramiento guiado por la genómica de la palma de aceite.

Agradecimientos

Los autores agradecen al director general del MPOB por su permiso para presentar este trabajo.



Bibliografía

- Adam, M. D.; Kelley, J. M.; Gocayne, J. D.; Dubnick, M.; Polymeropoulos, M. H.; Xiao, H.; Merril, C. R.; Wu, A.; Olde, B.; Moreno, R. F.; Kertavague, A. R.; McCombie, W. R.; Venter, J. C. 1991. Complementary DNA sequencing: expressed sequenced tags and human genome project. *Science*, 252: 1651-1656.
- Cheah, S. C. 1999. Biotech in the 21st century: what's in it for the oil palm? *Palm Oil Developments*, 31: 20-25.
- Chan, P. L.; Low, E. T. L.; Boon, S. H.; Rozana, R.; Ooi, L. C-L.; Rajinder, S.; Cheah, S. C. 2005. Monitoring gene expression changes in different stages of oil palm tissue culture via cDNA microarray. Proc. of the 2005 Conference on Biotechnology of Plantation Commodities. Miec, Mines Resort City, Selangor, 9-11 agosto 2005: 445-453.
- Fields, C. 1994. Analysis of gene expression by tissue and developmental stage. *Current Opinion in Biotechnology*, 5: 595-598.
- Höfte, H.; Desprez, T.; Amselem, J.; Chiapello, H.; Caboche, M.; Moisan, A.; Jourjon, M-F.; Charpentreau, J-L.; Berthomieu, P.; Guerrier, D.; Giraudat, J.; Quigley, F.; Thomas, F.; Yu, D-Y., Mache, R., Raynal, M., Cooke, R., Grellet, F., Delseny, M., Parmentier, Y., De Marcillac, G., Gigot, C., Fleck, J., Philipps, G., Axelos, M., Bardet, C., Tremousaygue, D.; Lescure, B. 1993. An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. *The Plant Journal*, 4(6): 1051-1061.
- Jalani, B. S. 1998. Research and Development of Oil Palm towards the next millennium. En: International Oil Palm Conference, Indonesian Oil Palm Research Institute (Iopri) and Indonesian Palm Oil Producers Association (Gapki), Bali (Indonesia), 23-25 septiembre 1998: 21 p.
- Low, E. T. L.; Tan, J. S.; Chan, P. L.; Boon, S. H.; Wong, Y. L.; Rozana, R.; Ooi, L. C-L.; Ma, L. S.; Ong-Abdullah, M.; Cheah, S. C.; Rajinder, S. 2006. Developments toward the application of DNA chip technology in oil palm tissue culture. *J. Oil Palm Research* (Special Issue): 87-98.
- Low, E. T. L.; Alias, H.; Boon, S-H.; Shariff, E. M.; Tan, C-Y. A.; Ooi, L. C-L.; Cheah, S. C.; Raha, A-R.; Wan, K-L; Singh, R. 2008. Oil palm (*Elaeis guineensis* Jacq.) tissue culture ESTs: Identifying genes associated with callogenesis and embryogenesis. *BMC Plant Biology*, 8: 62.
- Madon, M.; Clyde, M. M.; Cheah, S. C. 1999. Application of genomic in situ hybridization (GISH) on *Elaeis* hybrids. *Journal of Oil Palm Research*, Special Issue (octubre, 1999): 74-80.
- Madon, M.; Phoon, L. Q.; Clyde, M. M.; Mohd Din, A. 2008. Application of flow cytometry for estimation of nuclear DNA content in *Elaeis*. *J. Oil Palm Research*, 20: 447-452.
- Mayes, S.; Jack, P. L.; Marshall, D. F.; Corley, R. H. V. 1997. Construction of a RFLP genetic linkage map for oil palm (*Elaeis guineensis* Jacq.). *Genome*, 40: 116-122.
- Ong-Abdullah, M.; Ooi S. E. 2006. Biomarkers: Finding a Niche in Oil Palm Tissue Culture. Part 1-Laying the Foundation. *Oil Palm Bulletin*, 53: 36-48

- Ong-Abdullah, M.; Ooi, S. E. 2007. Biomarkers: Finding a Niche in Oil Palm Tissue Culture. Part 2 - Targeting the Transcriptome. *Oil Palm Bulletin*, 54:68- 74.
- Ooi, L.C. L.; Cheah, S. C.; Rahimah, A. R.; Bresnahan, P. 2001. Preliminary microarray gene profiling and differential gene expression studies in oil palm. Poster presented at *Microarray Users Group (MUG) Meeting*, 12-14 agosto 2001, Victoria, B.C. (Canadá).
- Ooi, S. E.; Feshah, I.; Ong-Abdullah M. 2010. Cloning and *in silico* analysis of promoters of highly expressed genes in oil palm embryogenic cultures. *Journal of Oil Palm Research*, 22: 856-868.
- Rabinowicz, P. D.; Schutz, K.; Dedhia, N.; Jordan, C.; Parnell, L. D.; Stein, L.; Mccombie, W. R.; Martienssen, R. A. 1999. Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. *Nature Genetics*, 23 (3): 305-308.
- Sambanthamurthi, R.; Rajinder, S.; Parveez, G. K. A.; Ong-Abdullah, M.; Kushairi, A. 2009. Opportunities for the oil palm via breeding and biotechnology. *Breeding Plantation Tree Crops* (Mohan Jain, S. y Priyadarshan, P. M. eds.). Springer Science & Business Media: 377-421).
- Sasaki, T.; Song, J.; Koga-Ban, Y.; Matsui, E.; Fang, E.; Higo, F.; Nagasaki, H.; Hori, M.; Miya, M.; Murayamakayano, E. 1994. Toward cataloging all rice genes: large-scale sequencing of randomly chosen cDNA from a callus cDNA library. *Plant Journal*, 6(4): 615-624.
- Shoemaker, R.; Keim, P.; Vodkin, L.; Retzel, E.; Clifton, S. W.; Waterson, R.; Smoller, D.; Coryell, V.; Khanna, A.; Erpelding, J.; Gai, X.; Brendel, V.; Raph-Schmidt, C.; Shoop, E. G.; Vielweber, C. J.; Schmatz, M.; Pape, D.; Bowers, Y.; Theising, B.; Martin, J.; Dante, M.; Wylie, T.; Granger, C. 2002. A compilation of soybean ESTs: generation and analysis. *Genome*, 45(2): 329-338.
- Singh, R.; Cheah, S. C.; Madon, M.; Ooi, L. C-L.; Rahimah, A. R. 2001. Genomic Strategies for Enhancing the value of oil palm. En: *Proceedings MPOB International Palm Conference*, Malaysian Palm Oil Board, 20-27 agosto 2001, Kuala Lumpur (Malasia): 3-7.
- Singh, R.; Madon, M.; Low, E. T. L.; Ooi, L. C- L.; Chan, P. L.; Rosli, R.; Ting, N. C.; Ithnin, M. 2011. Oil palm genomics: a foundation for improved agricultural productivity. *Further Advances*. En: *Oil Palm Research (2000-2010)* (Mohd Basri, W; Choo, Y. M. y Chan, K. W. (eds.)). 1. MPOB, Bangi: 202-251.
- Singh, R.; Tan S. G.; Panandam, J.; Rahimah, A. R.; Cheah, S. C. 2008a. Identification of cDNA-RFLP markers and their use for molecular mapping in oil palm (*Elaeis guineensis*). *Asia Pacific Journal of Molecular Biology & Biotechnology*, 16(3): 53-63.
- Singh, R.; Zaki, N. M.; Ting, N-C.; Rosli, R.; Tan, S-G.; Low, E-T. L.; Ithnin, M.; Cheah, S-C. 2008b. Exploiting an oil palm EST database for the development of gene-derived SSR markers and their exploitation for assessment of genetic diversity. *Biologia*, 63(2): 227-235.
- Ting, N. C.; Noorhariza, M. Z.; Rozana, R.; Low, E. T. L.; Maizura, I.; Cheah, S. C.; Tan, S. G.; Singh, R. 2010. SSR mining in oil palm EST database: application in oil palm germplasm diversity studies. *Journal of Genetics*, 89: 135-145.
- Verdun, R. E.; Di Paolo, N.; Urmenyi, T. P.; Rondinelli, E.; Frasch, A. C.; Sanchez, D. O. 1998. Gene discovery through expressed sequence tag sequencing in *Trypanosoma cruzi*. *Infection and Immunity*, 66: 5393-5398.