MJSAT

Malaysian Journal of Science and Advanced Technology

journal homepage: https://mjsat.com.my/

A Review on the Framework for Identifying High-yielding and Stress-tolerant Oil Palm with Molecular Markers

Fatini Nadhirah Mohd Nain[1], Najihah Ahmad Latif[1], Nurul Hashimah Ahamed Hassain Malim*[1], Rosni Abdullah[1], Muhamad Farid Abdul Rahim[2], Mohd Azinuddin Ahmad Mokhtar[2], and Nurul Syafika Mohamad Fauzi[2]

[1] *School of Computer Sciences, Universiti Sains Malaysia (USM), Gelugor 11800, Pulau Pinang, Malaysia.*

[2] *FGV Research and Development (R&D) Sdn Bhd, Unit Biak Baka Sawit, Pusat Penyelidikan Pertanian Tun Razak, Jengka 26400, Pahang, Malaysia.*

KEYWORDS

Oil palm
Oil palm breeding
Genetic markers
Molecular markers

ARTICLE HISTORY

Received 5 July 2023
Received in revised form 25 August 2023
Accepted 30 August 2023
Available online 1 September 2023

ABSTRACT

The various molecular markers employed in breeding programs have been identified in the last decade of the 20th century. Molecular marker development has significantly altered genetics. Molecular markers are not only utilised in connecting diagrams, but often involved during genetic differences in cultivars and germplasm evaluations in various plant breeding applications. Quantitative trait loci (QTL) and marker-assisted selection (MAS) are among the most interesting molecular markers utilisation. This review primarily aims to offer the knowledge required for the survival of plants developed via new molecular marker techniques and genetic modification implementation in oil palm breeding. Research in palm breeding is critical to gain insights into oil palm growth. This study could serve as a useful guide to understand oil palm development studies and improve existing prediction methods. Genetic markers crucial for oil palm development are also highlighted in the present study. Furthermore, a study gap was also identified.

1. INTRODUCTION

Malaysia and Indonesia are among the largest oil palm commodity countries worldwide [1]. Collectively, the nations supply 85% of the global palm oil [2]. In 2017, Indonesia contributed 12.3 million hectares of oil palm plantations, hence the primary global palm oil manufacturer and exporter [1]. Malaysia is the second-largest palm oil producer, considering 60% of its agricultural land was planted with 587 million ha of oil palm in 2020 [3]. Other nations, including Thailand, Colombia, Nigeria, and Ecuador, contribute 19 million acres, 0.36% of worldwide palm oil cultivation land [4].

Elaeis guineensis (*E. guineensis*), or African oil palm, and Elaeis oleifera (*E. oleifera*), also known as American oil palm, are two oil palm species. *E. guineensis* yields more oil as it produces a higher number of seeds. Although Tenera originated from Africa, it is primarily planted in Southeast Asia, particularly Indonesia and Malaysia, and contributes to over 50% of oil palm plantations globally.

Tenera, a Dura and Pisifera hybrid in Malaysia, is the most popular planted oil palm variety. Tenera produces a satisfactory quantity of palm and palm kernel oil. Ninety per cent of oil is employed traditionally, whereas the remaining 10% is in the oleochemical market. Consequently, palm oil with high production would satisfy the demand for vegetable oil due to the increasing world population.

The employment of molecular markers to identify DNA polymorphisms is a significant advancement in the genetic sciences field. The markers have been utilised to identify mutations and recombinations in a species or population. The technique provides a modern approach to creating new DNA profiling markings based on polymerase chain reaction (PCR).

*Corresponding author:
E-mail address: Nurul Hashimah Ahamed Hassain Malim <nurulhashimah@usm.my>
https://doi.org/10.56532/mjsat.v3i3.172

Random DNA-based markers were the first genomic methods developed and adopted for several plants. Molecular markers prove advantageous over traditional phenotype-based alternatives in all tissues independent of growth, diversion, proliferation, and cell defence. Therefore, reviewing previous studies on oil palm breeding methods and current challenges is integral to identifing potential approaches ad improving oil palm breeding via molecular markers. Furthermore, comprehension of oil palm breeding might also be applicable in other domains.

2. OIL PALM BACKGROUND

E. guineensis is a typical oil palm from the West African rainforest. The female and male *E. guineensis* oil palm trees are of one size and fluorescence in similar tree parts [5]. The oil palm species could grow up to over 30 m with an average life span between 25 to 30 years [6].

E. guineensis trees start to grow fruit bunches from three years old. Each tree contributes between eight and twelve fruit bunches per year [6]. *E. guineensis* Jacq is one of the most productive and abundant tropical oil plants globally and a major tropical oil harvest [7].

Oil palm is a monoecious crop following its ability to produce male and female flowers on a single plant [8]. The life spans of male and female flowers vary according to genotypes and environments, which could range from four to six months [9]. Inflorescence begins emerging as a palm tree matures at two to three years post-plantation. An oil palm could yield one to two fruit grapples (female phase) or male inflorescences monthly.

A comprehensive examination has revealed that every floral primordium comprises male and female reproductive structures [10], [11]. In the 14th month of growth, sexual distinction in oil palm trees occurs. The stigma in male flowers is suppressed, while the stamen in female flowers is underdeveloped during the sexual distinction period [12], [13]. The development of male and female organs leads to a hermaphrodite flora in young plants, however, during the floral cycle transition, the parts become normal [8].

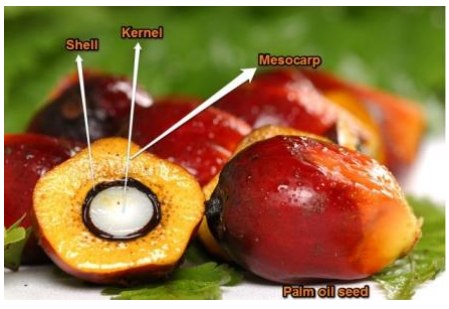

Fig. 1. The structure of an oil palm seed [14]

Thirty months post-plantation, oil palms produce fruits for 25 to 30 years, ensuring a steady oil supply. Each oil palm yields 1000–3000 fruitlets per compact bunch, weighing between 20 and 25 kg. Each fruitlet is elongated or almost spherical. The term "fresh fruit bunch" (FFB) describes a ripe bunch. The colour of the fruitlets turns orange red from dark purple, almost black when ripe (see Figure 1). The hard kernels (seed) in each fruitlet are enclosed in endocarpic shells, which are encircled within fleshy mesocarpic shells.

2.1 Oil palm types, fruits, and cross-sectional breeding genotypes

Oil palms are diploids ($2_n = 32$) with an approximately 1.8 billion base pairs (bp) genome [15]. Numerous oil palm agronomic features are polygenic [16]. Nevertheless, shell thickness is the most significant among the few monogenic parts of the plant. A locus with two alleles determines shell thickness; *Sh* and *sh* shell [17] (see Figure 2). Based on shell thickness, oil palm trees are classified into three fruit classes: Dura, Pisifera, and Tenera.

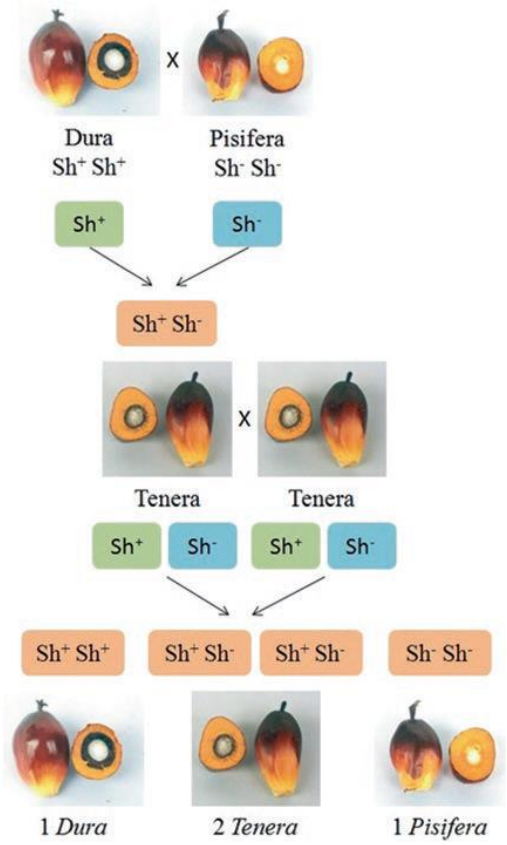

Fig. 2. Oil palm seed types and monogenic shell thickening gene heritage [18]

The homozygous genotype for the Dura trait (*ShSh*) is associated with fruits exhibiting a range of shell thickness, typically between 2 to 8 mm, and a diminished oil-bearing mesocarp. Conversely, Pisifera, which lacks a shell and has a mesocarp content of 95%, is as an optimal planting material. Nevertheless, Pisifera is typically characterised by female sterility due to the tendency of its pistillate inflorescence to undergo abortion during developmental stages. The characteristic renders the plant unsuitable for commercial employment, resulting in predominantly empty bunches production. Pisifera is also not the optimal female progenitor for breeding due to its propensity for fruit abortion.

Heterozygous Tenerad (*Shsh*) are obtained by breeding a Dura maternal parent and a Pisifera pollen donor. The fruits

from the resultant plant, Tenera, possess thinner shell, between 0.5 and 4 mm, and high mesocarp contents. Tenera fruits exhibit a notable characteristic, wherein approximately 30% of the shell is mesocarp, resulting in augmentation in oil production compared to Dura fruits [19], [20]. Consequently, Tenera is the most extensive oil palm variant cultivated commercially.

2.2 The oil palm breeding (OPB) genomic selection program

The two most common oil palm breeding (OPB) genomic selection programmes are the modified recurred selection (MRS), often applied in Far East programs held by Unilever, and the modified reciprocal selection (MRRS) led by the Center for International Cooperation (CIRAD), Indonesia [21]. In MRS, Duras (Ds) are chosen through a family and individual selection (FIS) process. The procedure involves collective and individual palm performance evaluations [17].

The MRS breeding plan primarily focuses on the general combining ability (GCA) effects with little emphasis on specific combining ability (SCA). The Pisiferas (Ps), which are female sterile, employed in the program are selected based on the performance of their Tenera (T) siblings in the T × T/P family. The chosen Ps are mated with designated sets of Ds to establish D × P progenies. When the average performance of the progenies is high, the parent (P) is deemed to possess a favourable GCA. Subsequently, the chosen Ds and Ps are employed in the commercial production of D × P seeds.

A benefit of MRS is the ability to conduct a higher number of recombinant crosses within a shorter timeframe. Consequently, less resources and effort are required compared to extensive progeny assessments. Nevertheless, a primary drawback is the absence of progeny testing for the D parents, raising concerns about the accuracy of the GCA effects observed in the D × D and T × T parent crosses. The hybrid performance of the D × P crosses might not completely reflect the GCA effects exhibited within the parental crosses [22]. Accordingly, several oil palm breeders have adopted the MRRS breeding system, in which two independent populations are maintained consisting of dura and pisifera genotypes.

Heterosis in hybrids (Tenera) is achieved by combining traits from different species. Offspring performance improvements could be traced back to the integration effects of desirable genes from the parents in inter-origin crosses. Introgression of selected materials from the germplasm developed by The Malaysian Palm Oil Board (MPOB) in existing materials is another means of increasing genetic diversity. Breeding programmes cross oil palms with a wide range of economically-desirable traits to diversify genetics. Nevertheless, the best materials require further evaluations through new combinations to maximise future commercial seed production.

Malaysian farmers profiting from the genetic potential of germplasms could ensure the long-term viability of the oil palm industry in the nation. The D × T is utilised in MRRS to detect all D and T parents during progeny evaluations. Optimal progeny assessments cross D and T parents. The resultant D and P palms are employed for commercial seed processing.

In MRRS, the parents and sibs are simultaneously raised and planted as progeny crosses for time efficiency. The approach involves GCA and SCA results. The only disadvantage is that the D x P crosses and sibs necessitate measurements in vast laboratory areas. For example, approximately 500 crosses and 180 self require plantations in over 600 ha and assessed over 15–25 years to yield 3 to 4 million commercial seeds within the top 15% of crosses.

3. MOLECULAR MARKERS

Historically, observational selection has been employed in traditional plant breeding to assess genetic diversity. Nonetheless, research in genetic diversity are currently conducted at the molecular level, focusing on DNA alterations and their impact on the phenotype in light of advancements in molecular biology.

After extracting DNA from plants, PCR or hybridisation techniques followed by agarose or acrylamide gel electrophoresis (AGE) and polyacrylamide gel electrophoresis (PAGE) are employed to detect sample alterations. The process allows the differentiation of various molecules based on size, chemical composition, and charges [23], [24].

Genetic markers enable the identification and monitoring of genetic variations within DNA specimens. Genetic markers are biological entities that could be recognised through allelic variations. The markers are experimental tools or indicators for tracing individuals, tissues, cells, nuclei, chromosomes, or genes. Consequently, genetic markers could facilitate heredity and variation studies.

Genetic polymorphism in classical genetics represents allele diversity, whereas, in contemporary genetics, it refers to the relative genetic loci variations of the genome. Some reports suggested that DNA markers could improve the effectiveness of traditional plant breeding programmes, thus improving global food production. Consequently, DNA markers are widely regarded as potentially beneficial crop-breeding tools.

A DNA marker is a small DNA sequence region that exhibits polymorphism and contains base deletion, insertion, and substitution between different individuals [25], [26]. Various DNA markers are frequently employed in plant breeding. Applying DNA markers in plant breeding has introduced a novel domain in agriculture named molecular breeding markers. The technique permits nucleotide sequences and polymorphism between various individual analyses [27]. Practical DNA markers are co-dominated, highly replicable, spread equally across the genome, and able to detect additional polymorphism [28]. The markers are utilised extensively due to their widespread occurrence and manifestation across different developmental phases in organisms.

According to Mendelian principles, genes are responsible for controlling inheritances. In early 20th century, scientists discovered that genes are arranged linearly on structures called chromosomes, which are defined cytogenetically. Studies demonstrated that gene inheritance could occur collectively, where they are linked together due to their proximity to the same chromosomes. Molecular DNA markers refer to specific genes in close proximity to each other within a defined interval. Co-dominance describes a genetic phenomenon, in which both alleles of a gene are expressed equally and distinctly, thus enabling differentiation between homozygous or heterozygous individuals at a particular locus.

Molecular markers represent discernible DNA sequences situated at precise positions within a genome. The sequences

correlate with transmitting a particular trait or a linked gene. Recently, several molecular markers have been planned for genetic association maps development and phylogenetic studies, including restriction fragment-length polymorphisms, single nucleotide polymorphisms, and simple sequence repeat in oil palm applications.

An advantage of molecular markers is their applicability in any region in the genome, including introns, exons, and regulation regions. Moreover, the DNA sequences do not possess pleiotropic or epistatic effects, are able to distinguish polymorphisms that do not result in phenotypic variations, and some are co-dominant. Genetic markers are also relatively resistant to environmental factors.

Numerous methodologies could be employed to evaluate multi- and single-locus markers. Multi-locus markers enable concurrent examinations of multiple genomic loci. The technique employs random chromosomal characteristic amplifications through oligonucleotide primers with arbitrary sequences. Multi-locus markers could be classified as dominant following their ability to indicate the presence or absence of a band at a particular locus. Nevertheless, the markers do not allow differentiation between heterozygote (a/-) and homozygote (a/a) conditions for the same allele. Conversely, single-locus markers utilise probes or primers specific to genomic loci. Consequently, the sequences enable chromosome traits with known sequence detection or amplification.

The primary categories for classifying basic marker techniques are non-PCR-based, hybridisation-based, and PCR-based approaches. On the other hand, three techniques are employed in molecular markers; PCR, hybridisation, and new generation-based method. Table 1 lists a comprehensive comparison of the markers frequently utilised in various contexts. This study reviewed literature that employed multiple techniques, including approaches derived from either nucleic acid restriction-hybridisation or PCR, or a combination of both.

Table 1. Key properties of the most widely UTILISED molecular markers in OPB Comparison [29], [30].

Characteristics	SNP*	SSR†	ISSR‡	RAPD§	AFLP**	CAPS††	RFLP‡‡	DArT§§
Co-dominant/ Dominant	Co-dominant	Co-dominant	Dominant	Dominant	Dominant	Co-dominant	Co-dominant	Dominant
Reproducibility	High	High	Medium–High	High	Intermediate	High	High	High
Degree of polymorphism	High	High	High	Very high	High	Low	Medium	High
Quality of DNA required	High	Low	Low	High	High	High	High	High
Quantity of DNA required	Low	Low	Low	Medium	Low	Low	High	Low
Level of genomic abundance	Very high	Medium	Medium	Very high	Very high	Low	High	Very high
Cost per assay	Variable	High	High	Less	High	Medium	High	Cheapest

3.1 PCR-based methods

PCR is an in vitro method that precisely replicates a selected DNA segment [31]. This section discusses seven types of PCR-based markers.

1) Single nucleotide polymorphism (SNP)

Single nucleotide polymorphism (SNP) is found in the genomic sequence of individuals within a population. The SNPs are the most recent generation of markers. In any organism, the most prevalent molecular markers might reveal hidden polymorphisms that other markers or techniques could not detect.

Although non-coding regions have higher SNP density than other parts, both are widely distributed throughout genomes. Theoretically, up to four alleles, each containing one of the four bases (A, T, C, or G), could be produced by SNP. Nonetheless, bi-allelic SNPs are more common in practice and frequently involve two pyrimidines, C/T, or two purines, A/G.

Although SNPs have a lower polymorphic information content (PIC) than multi-allelic microsatellites, the markers possess the potential to be the marker of choice. The limitation is more than offset by its number. Nugroho et al. [32] conducted a double digest restriction associated DNA (ddRAD) duplicate pairing on 236 palms from a series of genetic backgrounds to identify high-grade SNP, unweighted neighbour-joining (NJ) clustering, and allelic variant annotation. Resultantly, 8 459 674 SNP loci was generated from reads mapping when compared to the EG5 reference genome.The study also recorded over 133 048 mapped SNPs, with the highest density SNPs observed within the sixth pseudo-chromosome, documenting more than 531 SNPs per Mbp.

Genetic markers developed from selected SNPs could be further employed for oil palm breeding and marker-assisted selection programs and cultivar identification. Firthermore, high density map development is achieveable since SNPs are simple and cost effective. Studies on identifying oil palm

Note: * A single nucleotide polymorphism (SNP) refers to a genetic variation occurring at a specific base position within the DNA sequence.
† Simple sequence repeats (SSRs), commonly called microsatellites, are a prevalent form of repetitive DNA sequence found in most plant genomes.
‡ Inter simple sequence repeats (ISSR) refer to segments of DNA typically of 100 to 3000 bp.
§ The randomly amplified polymorphic DNA (RAPD) technique is a PCR-based method that employs arbitrary primers to selectively amplify DNA by binding to nonspecific sites on DNA molecules.
** The amplified fragment length polymorphism (AFLP) approach utilises PCR to selectively amplify a specific digested subset of DNA fragments.
†† Cleaved amplified polymorphic sequences (CAPS) refer to DNA fragments amplified through PCR.
‡‡ Restriction fragment length polymorphism (RFLP) identifies differences in DNA sequences at restriction enzyme recognition sites.
§§ The diversity arrays technology (DArT) is a genetic marker technique characterised by its high-throughput capabilities.

genomic sequences with SNP markers and their variants are summarised in Table 2.

Table 2. The Summary of Literature on Identifying oil palm genomic sequence with SNP markers

Author	Genomic DNA source	Method	Result
[33]	*E. guineensis* young leaves	SNP marker, SLAF tags, Bayesian clustering, and Linkage disequilibrium	1,261,501 SNPs and 17.81% genetic map
[34]	*E. guineensis* and *E. oleifera* young leaflet tissue.	SNP marker, Shapiro–Wilk and ANOVA assessments, QQ plot, Kinship matrix, principal component, analysis (PCA), and ion torrent personal genome machine (PGM)	167 gene candidates, 115 SNPs, and 12,200 SNPs
[35]	*E. guineensis* young leaves	SNP marker and unweighted pair group method with arithmetic mean (UPGMA)	768 SNPs
[36]	*E. oleifera* fresh leaves	SNP marker, principal coordinate analysis (PCoA), UPGMA, and Bayesian clustering	245 of 500 SNP markers mapped, $H_o^{***} = 0.190$ (0.001)

2) Simple sequence repeat (SSR)

The detection of simple sequence repeats (SSRs), alternatively microsatellites, short tandem repeats (STRs), or sequence-tagged microsatellite sites (STMS) rely on the PCR technique. The randomly occurring tandem repeats consist of short nucleotide motifs, typically ranging from 2 to 6 bp per nucleotide.

Mononucleotide (A), dinucleotide (GT), trinucleotide (ATT), tetranucleotide (ATCG), pentanucleotide (TAATC), and hexanucleotide (TGTGCA) are ubiquitously present in various plant species genomes [37]. Copy number variations of the repeats result in inter-individual variabilities and serve as a significant source of plant polymorphism. Primers specific to regions flanking microsatellite sequences are commonly designed for utilisation in PCR, as the sites tend to exhibit DNA sequence conservation.

Microsatellite loci are valuable and significant genetic markers as they possess high levels of allelic variation. The distinctive sequences flanking SSR motifs serve as templates for designing specific primers applicable for amplifying the SSR alleles through PCR. The SSR loci are amplified individually through PCR with oligonucleotide primer pairs specific to distinct DNA sequences surrounding the SSR sequence. The amplified products could be effectively separated via high-resolution electrophoresis systems, such as AGE and PAGE. The resulting bands are visually documented through fluorescent labelling or silver staining.

The SSR markers exhibit distinct characteristics, including high variability, reproducibility, co-dominance, and locus-specificity. Furthermore, the markers typically demonstrate random distribution throughout a genome. A notable benefit associated with SSR markers is their amenability to be analysed through PCR and subsequent detection through PAGE or AGE.

The SSR markers could be multiplexed, thus allowing simultaneous multiple loci analysis. The advantage enables high throughput genotyping, facilitating rapid and efficient genetic diversity assessments. Moreover, SSR markers are considerably suitable for automation, which further streamline the genotyping process.

Minimal DNA sample quantities are required for SSR assays, typically around 100 ng per individual. Furthermore, the markers necessitate relatively low initial expenses when employing manual assay techniques. Nevertheless, SSR markers require nucleotide information to facilitate primer design. Additionally, developing markers for SSR markers is labour-intensive with relatively high initial costs for implementing automated detections.

Since 1990, SSR markers have been widely employed in constructing genetic linkage maps, quantitative trait loci (QTL) mapping, marker-assisted selection, and plant germplasm analyses. Accordingly, numerous breeder-friendly SSR markers have been developed and are available for utilisation across various species.

Genetically diverse Tenera-hybrid progenies are crucial in oil palm breeding program development, particularly for high-yield Tenera growth. Consequently, this study assessed the genetic variations of 12 Tenera hybrid populations based on morphology and SSR markers [38]. Table 3 summarises reports identifying oil palm genomic sequences with SSR markers and their variants.

Table 3. The Summary of previous studies on Identifying oil palm genomic sequence with SSR markers

Author	Genomic DNA source	Method	Result
[39]	Dura (D) and Pisifera (P) nucleic acid leaf samples	SSR marker, Pedigree, and Hardy Weinberg equilibrium (HWE)	16 SSR markers, average polymorphic information content (PIC) = 0.605, and expected heterozygosity = 0.66
[40]	Tenera and Tenera, two Tenera and Pisifera crosses, and three Dura leaf	SSR marker, PCA, Neighbour-Joining clustering.	148 alleles, total PIC = 0.74, and total allele/locus = 8.38
[41]	Elaeidobius kamerunicus leaf	SSR marker and single-copy orthological versus benchmarking unit (BUSCO)	Total putative SSR = 281.668 and draft genome assembly = 59.9%

Note: *** Heterozygosity: Variants of a single gene (alleles).

3) Inter-simple sequence repeats (ISSR)

Inter-simple sequence repeat (ISSR) comprises DNA fragments consisting of approximately 100-3000 bp between neighbouring microsatellite regions. The markers, primers of a microsatellite core sequence with a few selective nucleotides, are boosted by PCR in neighbouring areas (16-18 bp). Accordingly, the number of primers that could be constructed with an anchor of a few bases for different combinations of di-, tri, tetra, and pentanucleotides is infinite. The resulting PCR also increases the sequence of two SSRs, hence providing a multi-locus marking system for fingerprinting, diversity analyses, and genomic mapping [42].

Not needing sequence data for the primary construction and low DNA amounts requirements are the primary advantages of ISSR. [43] conducted an *E. oleifera* × *E. guineensis* oil palm hybrid intersection with ISSR markers. The study employed eight ISSR markers in six of the 18 linkage groups involved, while 20 ISSRs consisting of seven separate small groups were utilised for the remaining samples. The report revealed that most of the ISSR markings were mapped at the distal ends of connecting units, which are areas inaccessible by other marking techniques.

4) Random amplification of polymorphic DNA (RAPD)

In 1991, Welsh and McClelland conducted a random, polymorphic diabetic DNA random amplification of polymorphic DNA (RAPD), a novel genetic PCR-based evaluation [44]. The RAPD is a PCR-oriented technology based on an arbitrary primer-enhancing target or random DNA segment. The technique employs a single primer of the arbitrary nucleotide sequence to identify DNA nuclear sequence polymorphisms.

During an RAPD reaction, the genomic DNA from two sites are utilised as supplementary strands for the DNA template via a primary annealing species. If the leading positions are in an enlarged range, thermocycle amplification forms a separate DNA product. On average, each primer amplifies several distinct loci in the genome, thus aiding effective nucleotide sequence polymorphism detection among individuals [45]. Nonetheless, DNA amplification that optimises and preserves coherent reaction words is critical following stochastic DNA amplifications with random sequence primers.

The RAPDs are PCR-amplified DNA fragments, which typically contain 10 bp primers. The oligonucleotides serve as frontal and reverse primers and commonly aid in improving genome sites with 1–10 fragments. The presence and absence of bands denote intensified products when viewed with ethidium bromide and ultraviolet light, typically within the 0.5–5 kb range [46].

The RAPDs are extremely rich in genomics and altered across the genome. The DNA sequences dominate markers, thus limiting their employment as mapping markers. Nevertheless, the disadvantage could be resolved by selecting appropriate connecting markers.

A study assessed the genetic variation of three *E.guineensis* Pisifera populations in Bangun Bandar, North Sumatra, Indonesia, with six RAPD markers [47]. The primary molecular diagnosis component involved in the report was the genetic material (germplasm) for improving palm oil seed quality and oil production.

Currently, Ganoderma resilient palm oil germplasms are available in Indonesia. Indonesian seedling manufacturers also offer commercially available germplasms, such as the Socfindo and PPKS MTG varieties. A molecular performance analysis of some RAPD-based palm oil MTG varieties involved 15 commercial variations of the oil palm [48]. The polymorphic PCR results revealed four RAPD markers.

5) Representational difference analysis (RDA)

Representational differential analysis (RDA) is a biological investigation tool employed in detecting gene differences between two genomic or cDNA samples. The technique involves PCR-amplified or variations analysis with subtractive DNA hybridisation genomes or cDNA sequences. The development of representational microarray analysis of oligonucleotides (ROMA), which utilises range technology, further enhanced the RDA technology. The ROMA approach is also applicable in identifying variations in DNA methylation.

The traditional subtractive hybridisation presented two main issues when analysing complex genomes. First, only the new adapter package will be connected to the test amplicons if the testing and driver amplicons have been installed following the removal of the adapters. During selective enhanced dual-stranded evaluations, PCR amplification is utilised. Excess DNA driver acts as competitive inhibitors to self-reanneal to the common sequences in the test drivers and the drivers. Consequently, only target sequences that could be selectively enriched at exponential rates during PCR would be present.

[49] reported successfully identifying a DNA fragment in the label region of the oil palm genomes that were highly susceptible to stress exerted on tissue culture and possessed high reorganisation and mutation levels. The authors proposed developing the DNA fragment into soma clonal detection packages. The study demonstrated the reason for the stability and efficiency of the approach in determining the distinctions between two samples. In another report, oil palm DNA fragments were utilised to assess the genetics of truncated leaf syndrome (TLS) ramets and healthy oil palms via genomic-representational difference analysis (G-RDA) [50]. The study developed oil palm clones from different variations with heterogenous primers package.

6) Amplified fragment length polymorphism (AFLP)

Numerous researchers have investigated the genetic variations in various plant species via amplified fragment length polymorphism (AFLP) molecular markers. Among molecular markers, AFLP produces notably more replicable marker bands. Therefore, the markers are essential in detecting in vitro genetic heterogeneity of plants [51], [52].

The primary characteristic of AFLP lies in its ability to reflect a genome. The technique is a highly versatile PCR-based method that could generate significant amounts of polymorphic loci [53]. Furthermore, AFLP connects RFLP to PCR, the break-in between DNA and PCR [54]. AFLP markers are also economical, and no prior sequence information is necessary. The AFLP process could be employed for high-quality and partially decayed DNA.

Palm oil variety assessments with AFLP markers demonstrated that the germplasm crosses of African and Deli palms differed with diversified crosses [55]. For instance, RFLP confirmed the genetic structures of geographically clustered accessions to *E. oleifera* in Brazil, French Guyana, Suriname, Peru, Northern Colombia, and Central America.

In an investigation, AFLP markers were employed to co-segregate the rough band of related genes in palm oil and classify a mixture of AFLP primers [56]. The study recorded that 22 of the 56 primers documented the standard rough bunch polymorphic markers. The results were also significantly similar to the retrotransposon copy Ty-1.

Karim and co-authors, constructed molecular markers via AFLP to cultivate prolific palm oil tissues [57]. The study obtained 44 polymorphic fragments with 33 non-prolific clones, one natural, and 10 prolific clone fragments. Further experiments involved 36 polymorphic fragments.

7) Cleaved amplified polymorphic sequence (CAPS)

Originally, cleaved polymorphic amplified sequence (CAPS) was known as PCR-RFLP, a hybrid [58], [59]. The procedure amplifies target DNA with PCR, followed by digestion via enzyme restriction [60], [61]. Accordingly, CAPS markers depend on PCR fragmentations caused by nucleotide polymorphism across samples and differences in restriction enzyme digestion pattern.

The benefits of CAPS markers include considerable ease and less time intensive to analyse alternate markers, such as south hybridisation, considering that PCR amplification is the basis for an interpretation of the polymorphic fragment-length restrictions. Secondly, CAPS expressed sequence tag (EST) primers are more valuable than their counterparts derived for comparative mapping studies from non-functional sequences, such as microsatellite markers. The CAPS markers are also inherited primarily co-dominantly. Moreover, CAPS markers could only be developed when mutations interrupt the enzyme recognition areas or establish limiting sites.

Some researchers produced an alternate marker, derivative CAPS (dCAPS), to overcome CAPS-related issues by creating discontinuity in the PCR mains. Subsequently, the dCAPS are utilised to produce a polymorphism dependent on the goal mutation. [62] utilised CAPS on an oil palm fruit detected two alleles in Dura, three in the Tenera genotype, and one allele in the Pisifera genotype. The resultant CAPS markers were validated with 300 genotypes, 25 Pisifera, 80 Pisifera, and 60 Tenera/Tenera/cross progeny lines.

3.2 Hybridisation-based method

Hybridisation-based markers named as such due to the polymorphism discovery procedure that involved hybridising the DNA markers in the samples assessed. Currently, restriction fragment length polymorphism (RFLP) and diversity array technology (DArT)are the hybridisation-based approaches applied in oil palm. The techniques are discussed in the following sections.

1) Restriction fragment length polymorphism (RFLP)

Restriction fragment length polymorphism (RFLP) markers are co-dominant, highly polymorphic, and reproducible. Consequently, the markers are transferable between labs. Nevertheless, the approach requires vast high-quality DNA amounts, relies on sample growth, and time, work, and cost efficient [63], [64]. The PCR-based RFLP method was implemented to increase RFLP detection and enable fast and high-performance analyses [65]. Nonetheless, known sequences are necessary to produce each marker.

The RFLPs were the early molecular markers for the first DNA generation-based genetics. The method also contributed to genetically-related marking identification. In its early conceptions, RFLPs were employed in plant genetic mappings. For instance, Singh et al. (2007) proposed a map containing 60 cDNA RFLP markers, which was applied during an oil palm genetic map linkage construction.

Zulkifli et al. [67], proposed utilising RFLP markers in determining oil palm hybrid interspecific crosses. The study identifies a CB75A marker for oil palm breeding in diverse autonomous populations. The results documented 20 out of 60 backcross palmitic acid (C16:0) samples in the form 2 groups of genotypes (AA and AB). Nevertheless, the large quantities of fatty acid in the specimens renders identification with RFLP challenging and time consuming.

2) Diversity array technology (DArT)

Diversity array technology (DArT) is a microarray-based technique that simultaneous typing of several hundred genomic polymorphism loci. The effectiveness of DArT markers has been demonstrated in various plants, genetic diversity and population composition studies, connection mappings, and connecting maps.

In a study, Gan et al. [68] proposed generating two closely-related populations of auto pollinated oil palm, AA0768 and AA0769, with 48 and 58 progeny with the first DArTseQ platform. The results revealed that 82% of the DArTseQ sequence marker tags recognised one site, focusing on gene-rich regions.

The DArT platform does not rely on sequence information availability compared to other SNP genotype platforms. Furthermore, DArT was re-developed for next-generation sequencing (NGS) advent by integrating it with NGS complexity reduction [69]. The resultant technique is commonly known as genotyping-by-sequencing (GBS). Currently, the improved approach covers over 130 different organisms, including 284 oil palm genotypes derived from Malaysian breeding programs.

3.3 New generation method

1) Next-generation sequencing (NGS)

The next-generation sequencing (NGS) strategy possesses the capability to parallel millions of analogous sequence readings in one run, offering an advanced sequence technology. The availability of the approach has swiftly transformed the life sciences landscape. The three major available systems on the market include Roche/454 FLX, Illumina/Solexa Genome Analyzer, and the SOLiDTM System for Applied Biosystems. Recently, two new systems were launched [70].

Array-based NGS provides much greater parallelism than traditional capillary sequences. The technology allows the simultaneous processing of hundreds of millions of readings instead of 96 at a time. Consequently, NGS substantially decrease sequencing costs by several orders per base.

Among the primary constraints of the new technology are shorter reading lengths of between 35 and 500 bp, the raw accuracy of base calls from new platforms are at least 10-fold less accurate than those produced by Sanger sequencers depending on the platform, and low precision of base calls from new platforms [71].

2) EST

The EST is obtained from cDNA clones. Commonly, the information generated is stored in databases. Expressed DNA sequence data disposal has allowed the creation of markers physically related to genome coding regions. Accordingly, the sequences are applicable for designing polymorphic markers. The resultant markers are part of the cDNA/EST sequences, thus, a part of the genome [72]. Subsequently, the sequences are developed into genetic markers by identifying polymorphic sequences affecting plant phenotypes in the genes and validating their associations and trait fluctuations.

According to recent analyses, EST markers are higher than random DNA markers utilised in marker-assisted selections. The raw sequence information also aids during EST-SSR or EST-SNP sequence occurrence screenings. The step also allows the genomic transcript regions to be designed.

3.4 Applications of marker technologies in oil palm breeding

Molecular markers are applied in various genomic analyses, including gene location and improving plant varieties by QTL, linkage mapping, genomic wide association (GWAS), genotyping by sequencing (GBS), and marker-assisted selection (MAS). Genomic analysis provides a wide range of information and numerous databases to maintain and popularise the data. Table 4 lists marker application trends in oil palm breeding between 2016 and 2021.

Table 4. The Trend of Molecular Marker Applications in Oil Palm Breeding (2016–2021)

Author	GWAS	GBS	MAS	QTL
[73]		/		
[74]	/			
[75]	/			/
[76]	/			
[77]				/
[78]		/		
[79]		/		

1) QTL and linkage mapping

Developing linkage maps for different types of crops is one of the most common utilisations of DNA markers in agricultural research. The relative positions of chromosomal molecular markers are critical to enable the utilisation of genetic information given by the markers. Chromosomal regions containing single gene traits dominated by a single gene are identified via QTL mapping, while linkage maps aid in detecting chromosomal sites with single gene traits.

Linkage mapping demonstrates the path involving two parent chromosomes. Similarly, genetic correlation maps reveal the relative lengths of symptoms and signals along the roads between markers along chromosomes. Primarily, the application of connection maps necessitates recognising chromosome areas with QTLs and genes correlated with the desired characteristics, referred to as QTL maps. Alternatively, QTL mapping is referred to as creating connection maps and analysing QTL to classify genomic regions involving desirable features.

2) Genomic-wide association (GWAS)

Genomic-wide association (GWAS) studies of different organisms are possible due recent advances in sequence and genotyping. The technique is a potent tool mainly employed to investigate the biology and characteristics of natural variants. Consequently, several companies have commercially established GWAS platforms.

Commonly, inbred lines are utilised for GWAS. Several phenotyping times are also performed post-genotyping the lines. A large population scale is required for QTL detections to produce a high resolution. The number of SNPs produced by GWAS is increasing due to technological advancements. Consequently, GWAS permits higher resolution and capacities to investigate small-scale haplotype blocks that are connected together. Furthermore, the approach is considerably effective and inexpensive. Almost every economically significant cultivation, such as maise, sorghum, millet, rice, and especially oil palm conducts GWAS.

3) Genotyping by sequencing (GBS)

One of the most effective plant breeding methods in genome applications is GBS. The method is employed during GWAS research, GS, gd studies, genetic relation analysis, and non-model plant marker discoveries. The GBS is also an excellent forum that enables a wide range of assessment, from a single gene to a diverse whole genome.

The GBS is becoming an outstanding method for various applications and research on plant reproduction and genetics in different food and industry crops due to its simplicity and low cost. The technique is a valid instrument for performing genomic studies. For instance, gdGBS is modern and exclusive to other Illumina-based GBS protocols. The approach provides more SNP genotype data with fewer observations than the Roche 454 platform. The main features of GD-GBS include the utilisation of Illumina multiplexing indexes for barcoding and specific bioinformatics, diploid population genotyping, and two restrictive enzymes to minimise genomic complexity.

4) Marker-assisted selection (MAS)

In marker-assisted selection (MAS), the genotype of a marker is employed to obtain a phenotypical collection. The molecular breeding technique aids in preventing issues related to traditional plant breeding. Moreover, the selection level has completely improved. Accordingly, plant breeding companies primarily employ MAS to identify supreme or recessive alleles in a generation and identify the most desirable samples in the segregating progeny.

4. OIL PALM MOLECULAR MARKERS TOOLS AND DATABASE SYSTEMS

Data mining essentially includes raw data entry databases. Nonetheless, efficient database development is a notable challenge due to continuous database content updates

requirements. Databases offer simple, user-friendly frameworks. Furthermore, software readily accessible to third parties are potential assets as programmers could modify and integrate them. The techniques are also significantly more relevant in whole genome sequence programs. Novel technologies and systems are required for the ever-expanding data extraction needs to deal with the complete genome data and the complexities of research.

Several databases have been developed to manage and share oil palm sequencing data to allow exploring, retrieving, and analysing large expressed sequence tags (ESTs) and whole genome sequence databases available (see Table 5). For example, the Genomsawit website offers sequences from the Malaysian Oil Palm Genome Programme (MyOPGP) downloads. Genome sequence, gene prediction, and transcriptome data, GeneThresher sequences (genomic sequences derived from the hypomethylated or gene-rich regions of the oil palm genome), SNP variants, sequences, and mapping locations (linkage groups) are the oil palm genome data accessible in MyOPGP. The NCBI has also received the oil palm genome sequences available in MyOPGP under the BioProject accession numbers PRJNA217845 for *E. guineensis* and PRJNA217846 for *E. oleifera*.

Elaeis guineensis oil palm (PalmXplore) is a public-domain archive listing predicted genes. The PalmXplore database contains information on genes related to significant traits, including the genes involved in fatty acid biosynthesis and disease resistance and annotations derived from external databases, such as Pfam, Gene Ontology, and Kyoto Encyclopaedia of Genes and Genomes.

Several oil palm databases offering genome sequences, proteins, genes, gene expression omnibus (GEO), pathways, substances, and compounds information are accessible via NCBI. A website allowing SSR identifications based on repeat motif and repeat types and primer details is pSatdb, an oil palm microsatellite database.

Various bioinformatics databases have examined the genes involved in vitamin E production in oil palms. The protein sequences deduced were aligned and characterised with SDSC Biology Workbench tools. Amino acid composition and isoelectric point were determined with ExPASy Proteomics tools, while the ProtParam tool served to examine the physical and chemical properties of the amino acid sequences. SignalP and ChloroP were employed to predict the targeted locations of the signal peptides of the deduced proteins. Furthermore, the Simple Modular Architectural Research Tool (SMART) database was utilised in identifying the protein domain, while the subcellular localisation prediction was conducted with WoLF PSORT. The Phylogeny.fr web services offer phylogenetic and molecular evolutionary analyses.

Table 5. The Oil palm genome databases created in Malaysia

Database	Description
Oil Palm SSR Resource Interface (OPSRI)	Provide exploring SSR specialised web-based tools [80].
Genomsawit	Produce an oil palm genome draught sequence [81].
PalmXplore	A public domain predictable oil palm (*E. guineensis*) archive [82].
FGV Integrated Breeding System (FIBS)	Stores FGV OPB study data collection and knowledge [83].

5. CONCLUSION

The OPB has played a vital role in developing crops. Numerous DNA data are obtained daily, thus enabling genes and genomes to be cross-referenced with map tools and sequence. Based on this literature review, molecular markers are the most advantageous diagnosis method in establishing genetic or DNA sequences.

Traditional markers, including biological, cytological, morphological, and phytochemical have been employed to identify new oil palm breeds. Nevertheless, the accuracy of the findings from the approaches is unknown. The invention of molecular markers has rendered identification methods straightforward. Overall, each molecular marker allows various detection methods and is employed for particular purposes.

AUTHOR CONTRIBUTIONS

The conceptualisation of the present study is by Fatini Nadhirah Mohd Nain, Najihah Ahmad Latif, and Nurul Hashimah Ahamed Hassain Malim. Fatini Nadhirah Mohd Nain performed original draft preparation and visualisation. Fatini Nadhirah Mohd Nain, Najihah Ahmad Latif, Nurul Hashimah Ahamed Hassain Malim, Rosni Abdullah, Muhamad Farid Abdul Rahim, Mohd Azinuddin Ahmad Mokhtar, and Nurul Syafika Mohamad Fauzi wrote, reviewed, and edited this article. Nurul Hashimah Ahamed Hassain Malim, and Rosni Abdullah conducted supervision, project administration, and funding acquisition. The authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGEMENTS

The authors are grateful for the support provided by Universiti Sains Malaysia and collaborator FGV Research and Development (R&D) Sdn Bhd. This study was funded via the Long-term Research Grant (LRGS) under the Malaysian Research University Network (MRUN)—203.PKOMP. 6777002.

REFERENCES

[1] H. Herdiansyah, H. A. Negoro, N. Rusdayanti, and S. Shara, “Palm oil plantation and cultivation: Prosperity and productivity of smallholders,” Open Agric., vol. 5, no. 1, pp. 617–630, 2020.

[2] Sime Darby Plantation, “Raising the Bar on Sustainable Growth,” Annu. Rep. 2019, 2019.

[3] I. D. Division and N. Sembilan, “Oil palm planted area,” vol. 2012, no. December, pp. 2–3, 2012.

[4] A. Kushairi et al., “Oil palm economic performance in Malaysia and r&d progress in 2017,” J. Oil Palm Res., vol. 30, no. 2. pp. 163–195, 2018.

[5] F. N. Maluin, M. Z. Hussein, and A. S. Idris, “An overview of the oil palm industry: Challenges and some emerging opportunities for nanotechnology development,” Agron., vol. 10, no. 3, 2020.

[6] E. Barcelos et al., “Oil palm natural diversity and the potential for yield improvement,” Front. Plant Sci., vol. 6, no. MAR, pp. 1–16, March 2015.

[7] D. J. Murphy, "The future of oil palm as a major global crop: Opportunities and challenges," J. Oil Palm Res., vol. 26, no. MAR, pp. 1–24, 2014.

[8] S. T. Gan, "The Development and Application of Molecular Markers for Linkage Mapping and Quantitative Trait Loci Analysis of Important Agronomic Traits in Oil Palm (Elaeis Guineensis Jacq.)," no. August, p. 482, 2014.

[9] E. Meijaard and D. Sheil, "The Moral Minefield of Ethical Oil Palm and Sustainable Development," Front. For. Glob. Chang., vol. 2, 2019.

[10] C. Darmawan, N. M. A. Wiendi, C. Utomo, and T. Liwang, "Electroporation-mediated genetic transformation of oil Palm (Elaeis guineensis)," Biodiversitas, vol. 21, no. 8, pp. 3720–3726, 2020.

[11] T. S. Wandita and S. Ashari, "Genetic Distance Between 47 Accessions Of Palm Oil's Germplasm (Elaeis guineensis Jacq.) Introduction From Cameroon Based On Morphological Character," PLANTROPICA J. Agric. Sci., vol. 4, no. 1, pp. 86–93, 2019.

[12] A. Rival, "Breeding the oil palm (Elaeis guineensis Jacq.) for climate change," OCL - Oilseeds fats, Crop. Lipids, vol. 24, no. 1, 2017.

[13] J. A. Teixeira da Silva and F. Engelmann, "Cryopreservation of oil palm (Elaeis guineensis Jacq.)," Cryobiology, vol. 77. pp. 82–88, 2017.

[14] info malaysia123, "Plam oil seed and Fresh fruit bunches - New Malaxi," 2014. [Online]. Available: https://www.newmalaxi.com/info/plam-oil-seed-and-fresh-fruit-bunches/. [Accessed: 11-Apr-2021].

[15] M. D. Bennett and I. J. Leitch, "Nuclear DNA amounts in angiosperms: Targets, trends and tomorrow," Ann. Bot., vol. 107, no. 3, pp. 467–590, 2011.

[16] E. T. L. Low et al., "The oil palm genome revolution," J. Oil Palm Res., vol. 29, no. 4, pp. 456–468, 2017.

[17] A. C. Soh, S. Mayes, and J. Roberts, "Plant genetics," in Oil Palm Breeding: Genetics and Genomics, 1st ed., 2017, pp. 57–83.

[18] F. A. Malike et al., "Oil palm (Elaeis spp.) breeding in Malaysia," in Advances in Plant Breeding Strategies: Industrial and Food Crops, vol. 6, Cham: Springer International Publishing, 2019, pp. 489–535.

[19] R. H. V. Corley and C. H. Lee, "The physiological basis for genetic improvement of oil palm in Malaysia," Euphytica, vol. 60, no. 3, pp. 179–184, 1992.

[20] S. S. Masura et al., "Post-genomic technologies for the advancement of oil palm research," J. Oil Palm Res., vol. 29, no. 4, pp. 469–486, 2017.

[21] A. C. Soh, G. Wong, T. Y. Hor, C. C. Tan, and P. S. Chew, "Oil Palm Genetic Improvement," in Plant Breeding Reviews, 2010, pp. 165–219.

[22] Q. Bin Kwong et al., "Evaluation of methods and marker Systems in Genomic Selection of oil palm (Elaeis guineensis Jacq.)," BMC Genet., vol. 18, no. 1, 2017.

[23] M. Kordrostami and M. Rahimi, "Molecular markers in plants: Concepts and applications," Genet. Third Millenn., vol. 13, no. 2, pp. 4024–4031, 2015.

[24] R. Bernardo, "Molecular markers and selection for complex traits in plants: Learning from the last 20 years," Crop Sci., vol. 48, no. 5, pp. 1649–1664, Sep. 2008.

[25] A. W. George and C. Cavanagh, "Genome-wide association mapping in plants," Theor. Appl. Genet., vol. 128, no. 6, pp. 1163–1174, 2015.

[26] A. Grover and P. C. Sharma, "Development and use of molecular markers: Past and present," Crit. Rev. Biotechnol., vol. 36, no. 2. pp. 290–302, 2016.

[27] A. H. Paterson, S. D. Tanksley, and M. E. Sorrells, "DNA Markers in Plant Improvement," Adv. Agron., vol. 46, no. C, pp. 39–90, 1991.

[28] M. E. D. Michael W. Bruford, Neil Davies and D. P. F. and M. Walters, "The GEO Handbook on Biodiversity Observation Networks," in The GEO Handbook on Biodiversity Observation Networks, M. W. • R. J. Scholes, Ed. Springer Nature, 2017, pp. 107–128.

[29] M. Kumar et al., "Advances in biochemical and molecular marker techniques and their applications in genetic studies of orchid: A review," Int. J. Chem. Stud., vol. 6, no. 6, pp. 806–822, 2018.

[30] M. A. Nadeem et al., "DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing," Biotechnol. Biotechnol. Equip., vol. 32, no. 2. pp. 261–285, 2018.

[31] K. B. Mullis and F. A. Faloona, "Specific Synthesis of DNA in Vitro via a Polymerase-Catalyzed Chain Reaction," Meth. Enzymol., vol. 155, no. C, pp. 335–350, 1987.

[32] Y. A. Nugroho et al., "Genome-wide SNP-discovery and analysis of genetic diversity in oil palm using double digest restriction site associated DNA sequencing," in IOP Conference Series: Earth and Environmental Science, 2019, vol. 293, no. 1.

[33] W. Xia et al., "Development of high-density snp markers and their application in evaluating genetic diversity and population structure in elaeis guineensis," Front. Plant Sci., vol. 10, 2019.

[34] M. Astorkia et al., "Association mapping between candidate gene SNP and production and oil quality traits in interspecific oil palm hybrids," Plants, vol. 8, no. 10, 2019.

[35] L. C. L. Ooi et al., "SureSawitTM TRUE-TO-TYPE - A high throughput universal single nucleotide polymorphism panel for DNA fingerprinting, purity testing and origin verification in oil palm," J. Oil Palm Res., vol. 31, no. 4, pp. 561–571, 2019.

[36] V. M. Pereira et al., "American oil palm from Brazil: Genetic diversity, population structure, and core collection," Crop Sci., vol. 60, no. 6, pp. 3212–3227, 2020.

[37] K. Shrivastava, R. Sinha, M. Singh Thakur, and C. Rebeka Sinha, "Molecular DNA markers in conservation of wildlife: A brief review," J. Entomol. Zool. Stud., vol. 6, no. 3, pp. 1519–1522, 2018.

[38] T. Khomphet, T. Eksomtramage, and S. Duangpan, "Genetic variation of improved oil palm tenera hybrid populations using morphological and SSR markers," Songklanakarin J. Sci. Technol., vol. 40, no. 6, 2018.

[39] L. F. Budiman, A. Apriyanto, A. Pancoro, and Sudarsono, "Illegitimacy testing of Elaeis guineensis population based on simple sequence repeat markers," Agrivita, vol. 41, no. 3, pp. 504–512, 2019.

[40] L. F. Budiman, A. Apriyanto, A. D. I. Pancoro, and S. Sudarsono, "Genetic diversity analysis of Tenera × Tenera and Tenera × pisifera crosses and D self of oil palm (Elaeis guineensis) parental populations originating from Cameroon," Biodiversitas, vol. 20, no. 4, pp. 937–949, 2019.

[41] A. Apriyanto and V. B. Tambunan, "Draft genome sequence, annotation, and SSR mining data of Elaeidobius kamerunicus Faust., an essential oil palm pollinating weevil," Data Br., vol. 34, 2021.

[42] I. D. Godwin, E. A. B. Aitken, and L. W. Smith, "Application of inter simple sequence repeat (ISSR) markers to plant genetics," Electrophoresis, 1997, vol. 18, no. 9, pp. 1524–1528.

[43] H. M. N. Nordiana, T. Ngoot-Chin, R. Singh, M. M. Clyde, and M. Madon, "Evaluation of inter-simple sequence repeat (ISSR) markers for genetic mapping of an oil palm interspecific hybrid mapping population," J. Oil Palm Res., vol. 26, no. 3, pp. 214–225, 2014.

[44] S. MARAKLİ, "A Brief Review of Molecular Markers to Analyse Medically Important Plants," Int. J. Life Sci. Biotechnol., vol. 1, no. 1, pp. 29–36, 2018.

[45] Y. Chang, E. U. Oh, M. S. Lee, H. B. Kim, D. G. Moon, and K. J. Song, "Construction of a genetic linkage map based on RAPD, AFLP, and SSR markers for tea plant (Camellia sinensis)," Euphytica, vol. 213, no. 8, 2017.

[46] A. Younis, F. Ramzan, Y. Ramzan, F. Zulfiqar, M. Ahsan, and K. B. Lim, "Molecular markers improve abiotic stress tolerance in crops: A review," Plants, vol. 9, no. 10. pp. 1–16, 2020.

[47] M. Basyuni et al., "RAPD markers on genetic diversity in three populations of pisifera type of oil palm (elaeis guineensis)," in IOP Conference Series: Earth and Environmental Science, 2018, vol. 130, no. 1.

[48] L. A. P. Putri et al., "Molecular performance of commercial MTG variety oil palm based on RAPD markers," in IOP Conference Series: Earth and Environmental Science, 2018, vol. 122, no. 1.

[49] N. Sarpan, S. E. Ooi, M. Ong-Abdullah, C. L. Ho, C. F. Chin, and P. Namasivayam, "Representational difference analysis (RDA) for the identification of DNA marker's associated with tissue culture amenity in oil palm," J. Oil Palm Res., vol. 25, no. DEC, pp. 305–313, 2013.

[50] K. Silvarajoo, P. Namasivayam, C. L. Ho, and S. S. R. S. Alwee, "Genomic differences detected in the oil palm truncated leaf syndrome (TLS) ramets (elaeis guineensis Jacq.) using the representational difference analysis (RDA) approach," J. Oil Palm Res., vol. 30, no. 2, pp. 219–235, 2018.

[51] R. Aversano, F. Di Dato, A. Di Matteo, L. Frusciante, and D. Carputo, "AFLP analysis to assess genomic stability in Solanum regenerants derived from wild and cultivated species," Plant Biotechnol. Rep., vol. 5, no. 3, pp. 265–271, 2011.

[52] R. Mehta, V. Sharma, A. Sood, M. Sharma, and R. K. Sharma, "Induction of somatic embryogenesis and analysis of genetic fidelity of in vitro-derived plantlets of Bambusa nutans Wall., using AFLP markers," Eur. J. For. Res., vol. 130, no. 5, pp. 729–736, 2011.

[53] P. Vos et al., "AFLP: A new technique for DNA fingerprinting," Nucleic Acids Res., vol. 23, no. 21, pp. 4407–4414, 1995.

[54] M. Lynch and B. Walsh, "Genetics and analysis of quantitative traits. Sinauer Assocs," Inc., Sunderland, MA, vol. 980, 1998.

[55] A. R. Purba, J. L. Noyer, L. Baudouin, X. Perrier, S. Hamon, and P. J. L. Lagoda, "A new aspect of genetic diversity of Indonesian oil palm (Elaeis guineensis Jacq.) revealed by isoenzyme and AFLP markers and its consequences for breeding," Theor. Appl. Genet., vol. 101, no. 5–6, pp. 956–961, 2000.

[56] , R., , S., S. Yahya, N. Toruan-Mathius, and T. Liwang, "Identification of Gene Related to Hard Bunch Phenotype in Oil Palm (Elaeis guineensis Jacq.)," J. Agron. Indones. (Indonesian J. Agron., vol. 43, no. 2, p. 147, 2015.

[57] S. K. A. Karim and N. M. Sidik, "Construction of molecule markers for prolific oil palm tissue culture," Sains Malays., vol. 47, no. 8, pp. 1701–1708, August 2018.

[58] F. Khan, "Molecular Markers: An Excellent Tool for Genetic Analysis," J. Mol. Biomark. Diagn., vol. 06, no. 03, 2015.

[59] M. Maeda et al., "A simple and rapid method for HLA-DP genotyping by digestion of PCR-amplified DNA with allele-specific restriction endonucleases," Hum. Immunol., vol. 27, no. 2, pp. 111–121, 1990.

[60] C. Palomino, M. D. Fernández-Romero, J. Rubio, A. Torres, M. T. Moreno, and T. Millán, "Integration of new CAPS and dCAPS-RGA markers into a composite chickpea genetic map and their association with disease resistance," Theor. Appl. Genet., vol. 118, no. 4, pp. 671–682, 2009.

[61] X. Yu et al., "Integration of est-Caps markers into genetic maps of eucalyptus urophylla and e. tereticornis and their alignment with e. grandis genome sequence," Silvae Genet., vol. 61, no. 6, pp. 247–255, 2012.

[62] B. Kalyana Babu et al., "Development, identification & validation of CAPS marker for SHELL trait which governs dura, pisifera & tenera fruit forms in oil palm (Elaeis guineensis Jacq.)," PLoS One, vol. 12, no. 2, 2017.

[63] B. Mittal, P. Chaturvedi, and S. Tulsyan, "Restriction Fragment Length Polymorphism," Brenner's Encycl. Genet. Second Ed., pp. 190–193, 2013.

[64] R. Ahmad, M. A. Anjum, and R. M. Balal, "From markers to genome based breeding in horticultural crops: An overview," Phyton, vol. 89, no. 2. pp. 183–204, 2020.

[65] H. Hussain and M. Nisar, "Assessment of plant genetic variations using molecular markers: A review," J. Appl. Biol. Biotechnol., vol. 8, no. 5. pp. 99–109, 2020.

[66] R. Singh, T. S. Guan, J. Panandam, M. Sharma, and C. S. Choo, "Preliminary analysis of quantitative trait loci associated with oil quality in an interspecific cross of oil palm," Pertanika J. Trop. Agric. Sci., vol. 30, no. 1, pp. 31–44, 2007.

[67] Y. Zulkifli, I. Maizura, S. Rajinder, M. Marhalil, A. Mohd Din, and M. Ong-Abdullah, "An RFLP marker, CB75A, predicts for palmitic acid content across selected oil palm germplasm and interspecific-hybrids," J. Oil Palm Res., vol. 31, no. 2, pp. 315–320, 2019.

[68] S. T. Gan et al., "High density SNP and DArT-based genetic linkage maps of two closely related oil palm populations," J. Appl. Genet., vol. 59, no. 1, pp. 23–34, February 2018.

[69] N. T. P. Phung et al., "Genome-wide association mapping for root traits in a panel of rice accessions from Vietnam," BMC Plant Biol., vol. 16, no. 1, 2016.

[70] D. Vlk and J. Řepková, "Application of next-generation sequencing in plant breeding," Czech J. Genet. Plant Breed., vol. 53, no. 3. pp. 89–96, 2017.

[71] A. Roychoudhury, "Next Generation Sequencing: Prospects in Plant Breeding and Crop Improvement," 2020.

[72] J. Parkinson and M. Blaxter, "Expressed sequence tags: An overview," Methods Mol. Biol., vol. 533, pp. 1–12, 2009.

[73] W. Pootakham, C. Sonthirod, C. Naktang, N. Jomchai, D. Sangsrakru, and S. Tangphatsornruang, "Effects of methylation-sensitive enzymes on the enrichment of genic SNPs and the degree of genome complexity reduction in a two-enzyme genotyping-by-sequencing (GBS) approach: a case study in oil palm (Elaeis guineensis)," Mol. Breed., vol. 36, no. 11, 2016.

[74] Q. Bin Kwong et al., "Development and Validation of a High-Density SNP Genotyping Array for African Oil Palm," Mol. Plant, vol. 9, no. 8, pp. 1132–1141, 2016.

[75] B. K. Babu, R. K. Mathur, G. Ravichandran, and M. V. B. Venu, "Genome-wide association study (GWAS) for stem height increment in oil palm (Elaeis guineensis) germplasm using SNP markers," Tree Genet. Genomes, vol. 15, no. 3, 2019.

[76] W. Xia et al., "Identification and Validation of Candidate Genes Involved in Fatty Acid Content in Oil Palm by Genome-Wide Association Analysis," Front. Plant Sci., vol. 10, no. October, pp. 1–10, 2019.

[77] P. W. Ong et al., "Association of SNP markers with height increment in mpob-angolan natural oil palm populations," J. Oil Palm Res., vol. 30, no. 1. pp. 61–70, 2018.

[78] B. G. Díaz et al., "Genome-wide SNP analysis to assess the genetic population structure and diversity of Acrocomia species," PLoS One, vol. 16, no. 7, July 2021.

[79] A. Nyouma et al., "Genomic predictions improve clonal selection in oil palm (Elaeis guineensis Jacq.) hybrids," Plant Sci., vol. 299, 2020.

[80] N. C. Ting et al., "SSR mining in oil palm EST database: Application in oil palm germplasm diversity studies," J. Genet., vol. 89, no. 2, pp. 135–145, 2010.

[81] K. . Rosli, R., Ab Halim, M.-A., Chan, "Welcome to Malaysian Oil Palm Genome Programme Website," MPOB Information Series, 2014. [Online]. Available: http://genomsawit.mpob.gov.my/index.php?track=1. [Accessed: 31-Mar-2021].

[82] N. S. N. M. Sanusi et al., "PalmXplore: Oil palm gene database," Database, vol. 2018, no. 2018, 2018.

[83] N. S. Mohamad Fauzi, M. F. Abd Rahim, and M. N. Hj Mohamad, "Implementation of information system in oil palm breeding research: FGV's experiences," Int. J. Eng. Trends Technol., no. 1, pp. 104–108, 2020.