

STUDI ASOSIASI GENOM (GWAS) KOMPONEN TANDAN POPULASI INTERSPESIFIK *Elaeis oleifera* DAN *Elaeis guineensis*

GENOME-WIDE ASSOCIATION STUDY (GWAS) FOR BUNCH COMPONENTS OF THE INTERSPECIFIC POPULATION OF *Elaeis oleifera* AND *Elaeis guineensis*

Edy Suprianto, Heri Adriwan Siregar*, Sujadi, Hernawan Y. Rahmadi, Mohamad Arif, Retno Diah Setiowati, Nanang Supena, Rokhana Faizah, Sri Wening, Yurna Yenni, dan Abdul Razak Purba

Abstrak Program pemuliaan kelapa sawit spesies *Elaeis guineensis* maupun *backcross Elaeis oleifera* berjalan lambat karena kelapa sawit merupakan tanaman tahunan. Oleh karena itu, sangat diperlukan alternatif pendekatan yang mampu mempercepat program pemuliaan kelapa sawit. Pendekatan *genome-wide* SNP (*single nucleotide polymorphism*) kemudian digunakan untuk mempelajari asosiasi antara 18 fenotipe komponen tandan pada plasma nutfah kelapa sawit spesies *E. oleifera* origin Suriname dan Brazil Coari, beberapa hibrida interspesifik dan turunan elit *E. guineensis*. Metode *genotyping by sequencing* (GBS) menghasilkan total 459 juta atau sekitar 798 ribu bacaan per sampel dan 3.252 SNP memenuhi syarat untuk 456 genotipe. Dengan menggunakan berbagai model asosiasi, sebelas data fenotipe yang dinormalisasi berhasil menunjukkan asosiasi nyata dengan 29 SNP. Berdasarkan anotasi, 17 SNP diantaranya terkait gen dengan fungsi biologi tertentu. Tiga SNP ditemukan berada pada *exon* sebuah gen, yaitu SNP4416, SNP349 dan SNP3865, sedangkan 15 SNP lainnya berada pada daerah *intragenic* sebuah gen. Sebanyak 4 SNP diantaranya merupakan *common* SNP pada fenotipe C16:0 dan C18:1 serta C20:0 dan C20:1. Penelitian ini menunjukkan potensi SNP yang dapat digunakan sebagai pendekatan alternatif pemuliaan *backcross*

E. oleifera, walau diperlukan juga penelitian lanjutan dalam bentuk validasi.

Kata Kunci: Asosiasi, *Elaeis guineensis*, *Elaeis oleifera*, *Genotyping-by-sequencing*, Hibrida interspesifik

Abstract The oil palm breeding program for the species *Elaeis guineensis* and the *backcross Elaeis oleifera* is running slowly because oil palm is an annual plant. Therefore, it is necessary to have an alternative approach that can accelerate the oil palm breeding program. The SNP (*single nucleotide polymorphism*) *genome-wide* approach was then used to study the association between 18 phenotypes of bunch component in oil palm germplasm of *E. oleifera* from Suriname and Brazil Coari, some interspecific hybrids and some elite progeny of *E. guineensis*. The *genotyping by sequencing* (GBS) analysis produced a total of 459 million or approximately 798 thousand reads per sample and 3,252 SNPs were eligible for 456 genotypes. Using various association models, eleven normalized phenotypic data showed significant associations with 29 SNPs. Based on the annotations, 17 SNPs were related to genes with certain biological functions. Three SNPs were found to be at the exon of a gene, namely SNP4416, SNP349 and SNP3865, while the other 15 SNPs were at the intragenic to a gene. Four SNPs are common SNPs in phenotypes C16:0 and C18:1 as well as in C20:0 and C20:1. This research shows the potential of SNPs that can be used as an alternative approach to *E. oleifera* *backcross*

Penulis yang tidak disertai dengan catatan kaki instansi adalah peneliti pada Pusat Penelitian Kelapa Sawit

Heri Adriwan Siregar(✉)
Pusat Penelitian Kelapa Sawit
Jl. Brigjen Katamso No. 51 Medan 20158

Email: heri_adriwan@me.com



breeding, although further research is needed for validation purposes.

Keywords: Association, *Elaeis guineensis*, *Elaeis oleifera*, Genotyping-by-sequencing, Interspecific hybrid

PENDAHULUAN

Perjalanan panjang kelapa sawit spesies *Elaeis oleifera* atau dulu disebut juga *Elaeis melanococca* dan *Corozo oleifera* (Corley dan Tinker (2016) di Indonesia dimulai pada 1952-1953. Indonesia memperkenalkan kelapa sawit spesies tersebut dari Brazil Coari dan Suriname pada 1952-1953, serta Kolombia antara 1975-1979 (Lubis, 2008). Sejak saat itu lembaga penelitian yang saat ini disebut Pusat Penelitian Kelapa Sawit (PPKS) menjalankan program pemuliaan *backcross* hingga diperoleh populasi hibrida interspesifik, *pseudo-backcross* I, dan *pseudo-backcross* II. Hasil pengamatan terhadap populasi interspesifik secara umum tidak berbeda jauh dengan hasil yang diperoleh dari penelitian lainnya. Hal tersebut mencakup pertumbuhan meninggi lambat, arsitektur tanaman kompak, produktivitas lebih rendah dibanding *Elaeis guineensis*, peningkatan produktivitas pada populasi hibrida interspesifik dan *pseudo-backcross* dibanding *E. oleifera*, kandungan asam lemak tidak jenuh dan karoten tinggi, serta ketahanan terhadap penyakit tertentu (Lubis, 2008; Osorio-Guarin *et al.*, 2019; Montoya *et al.*, 2013; Mozzon *et al.*, 2013; Siregar *et al.*, 2018a; Siregar *et al.*, 2018b). Sementara itu, dari sisi fenologi ditemukan beberapa perbedaan dengan *E. guineensis* seperti pembentukan jumlah pelepah yang lebih rendah (Moreno dan Romero, 2015).

Kelapa sawit *E. oleifera* termasuk dalam kelompok tanaman tahunan, sehingga seperti yang terjadi pada spesies *E. guineensis*, program pemuliaan *backcross* *E. oleifera* juga berjalan lambat (Durand-Gasselin *et al.*, 2011). Oleh sebab itu, sangat diperlukan teknologi yang mampu mempercepat program pemuliaan kelapa sawit. Menjawab tantangan tersebut, PPKS telah berhasil melakukan *resequencing* terhadap sepuluh individu populasi *E. oleifera* dan dua *E. guineensis* menggunakan Illumina HiSeq 2500 System (www.illumina.com). Dua belas individu tersebut terdiri dari masing-masing satu individu spesies *E. oleifera* Brazil Coari dan Suriname, dua masing-masing hibrida

interspesifik Brazil dan Suriname, dua masing-masing *pseudo-backcross* 1 Brazil dan Suriname, serta dua individu *E. guineensis*. Hasil *resequencing* dengan jumlah bacaan yang sudah di-filter dan *trim* sebesar lebih kurang 2 milyar. Bacaan tersebut kemudian dipetakan terhadap genom acuan *E. guineensis* MPOB EG5 dan MPOB EO8 (Singh *et al.*, 2013) menggunakan Burrows-Wheeler Aligner (BWA) versi 0.5.9rc1 (Li dan Durbin, 2009). Nilai *coverage* yang diperoleh per individu berdasarkan ukuran genom *E. guineensis* sebesar 8,56-11,22X. Total SNP (*single nucleotide polymorphism*) yang diperoleh sebesar 34.410.224 dan 5.811.378 untuk MPOB EG5 dan MPOB EO8, secara berurutan. Semua SNP yang diperoleh berdasarkan MPOB EG5 telah dianotasi dan efeknya diprediksi menggunakan SnpEff versi 3.6c (Cingolani, *et al.*, 2012).

Beberapa riset dengan pendekatan genomik telah berhasil memperluas pemahaman dibidang tanaman seperti pada kelapa sawit (Singh *et al.*, 2013); mangga (Wang *et al.*, 2020); dan lotus (Li *et al.*, 2020). Hal ini disebabkan penurunan drastis biaya sequencing DNA melalui teknologi *next generation sequencing* (NGS) yang memungkinkan analisis dalam jumlah besar baik individu maupun loci. Aplikasi bioinformatika yang inovatif, pendekatan statistika dan berbagai model yang dapat mengekstrak informasi dari *big data* juga ikut berperan besar (Luikart *et al.*, 2018). *Genome-wide association studies* (GWAS) juga telah berhasil mengidentifikasi kontribusi loci terhadap variasi fenotipe diantara individu (Luikart *et al.*, 2018), dengan menggunakan model regresi seperti *linear mixed-effect* (Gibson, 2018). *Trait-associated loci* umumnya teridentifikasi di daerah yang menunjukkan perbedaan genetik yang nyata antar individu dan berbeda secara morfologi yang terdiri dari jumlah loci berkontribusi dan besaran kontribusinya (Korte dan Farlow, 2013). Pada perkembangan terkini dilaporkan bahwa GWAS pada kelapa sawit (Bai *et al.*, 2017; Teh *et al.*, 2016), padi (Kawahara *et al.*, 2013; Kim *et al.*, 2020), dan kedelai (Ravelombola *et al.*, 2019) telah berhasil mendeteksi asosiasi antara loci dengan fenotipe tertentu. Pada dasarnya, studi GWAS dapat dirancang untuk menilai penentu genetik untuk hampir semua sifat kualitatif atau kuantitatif (Scherer, 2017).

Walau belakangan ini permintaan sangat berfluktuatif yang disebabkan berbagai faktor,

namun penggunaan minyak sawit terus meningkat setiap tahun (Reuters, 2020; Nurkhoiry, 2017). Diversifikasi penggunaan minyak sawit menjadi berbagai bahan baku industri makanan, kesehatan, kecantikan bahkan energi membuat industri kelapa sawit tetap menjanjikan sekaligus merupakan tantangan baru bagi industri kelapa sawit itu sendiri untuk mendesain dan menyediakan minyak sawit yang sesuai peruntukannya. Dengan demikian, menemukan marka DNA yang bertanggung jawab terhadap karakter-karakter agronomis tertentu menjadi relatif penting untuk mempercepat program pemuliaan kelapa sawit yang dapat mendukung berbagai industri tersebut.

BAHAN DAN METODE

Material genetik

Sebanyak 500 tanaman dari berbagai populasi interspesifik *E. oleifera* dan *E. guineensis* koleksi PPKS telah dianalisis. Seluruh tanaman ditanam di Kabupaten Simalungun Sumatera Utara dan berada pada 14 nomor percobaan berbeda dengan berbagai umur. Sampel terdiri dari empat pohon *E. oleifera* origin Brazil Coari, tiga pohon *E. oleifera* origin Suriname, 18 pohon *E. guineensis*, 339 pohon hibrida interspesifik, dan 136 pohon *pseudo-backcross* 1. Projeni hibrida interspesifik diperoleh dari persilangan antara 19 *E. oleifera* termasuk origin Brazil Coari dan Suriname dengan 30 individu *E. guineensis*. Projeni *pseudo-backcross* 1 diperoleh dari persilangan antara 14 individu hibrida interspesifik dengan 15 individu *E. guineensis* yang berbeda saat menghasilkan projeni hibrida interspesifik, sehingga disebut *pseudo-backcross* (Montoya *et al.*, 2013). Sementara itu, populasi *E. guineensis* diperoleh dari persilangan antara tiga individu dura Deli dan lima tenera Afrika.

Pengamatan fenotipe

Tanaman diamati antara dua sampai empat ulangan yang bergantung dari ketersediaan tandan dapat dipanen selama lebih kurang 24 bulan. Fenotipe yang diamati sebanyak 18 variabel terdiri dari komponen tandan meliputi persentase buah per tandan, daging per buah, minyak per mesokarp, minyak per tandan, inti per buah, rendemen, dan berat buah, serta kualitas minyak yang meliputi

kandungan karoten, persentase asam lemak laurat, miristat, palmitat, palmitoleat, stearat, oleat, linoleat, linolenat, arakidat, dan gadoleat. Analisis komponen tandan dilakukan berdasarkan metode pengamatan standar di PPKS dengan mengadopsi metode yang disampaikan Corley dan Tinker (2016). Komposisi asam lemak dianalisis menggunakan metode GC (gas chromatography) yang mengacu pada AOCS Official Method Ce 1b-89. Kandungan karoten dianalisis dengan acuan MPOB Test Methode p.2.6. part 2: 2004, PORIM Test Methode 1995 dan Metode Kurva baku.

Penetapan SNP

Data SNP diperoleh dari hasil analisis *resequencing* dua belas sampel yaitu, masing-masing satu dari *E. oleifera* origin Brazil dan Suriname, masing-masing dua dari F1 interspesifik dan *pseudo-backcross* 1 Brazil dan Suriname, serta masing-masing satu dari *E. guineensis* tipe tenera dan dura. Analisis *resequencing* menggunakan teknologi Illumina HiSeq2500 yang dipetakan terhadap genom acuan *E. guineensis* MPOB EG5 dan MPOB menggunakan Burrows-Wheeler Aligner (BWA) versi 0.5.9rc1. Sebanyak lima ribu SNP dari puluhan juta yang dihasilkan kemudian ditetapkan berdasarkan *impact* yang diprediksi menggunakan aplikasi SnpEff versi 3.6c antara lain *high*, *moderate*, *low*, *modifier upstream*, dan *intron modifier*.

Genotyping by sequencing

Platform targeted genotyping by sequencing-SeqSnp yang dikembangkan oleh *Biosearch Technologies Genomic Analysis* by LGC digunakan karena efektifitas dan efisiensinya (Hickey *et al.*, 2017). Proses genotyping dilakukan di laboratorium LGC yang secara garis besar kegiatan terdiri dari pengumpulan daun sampel dengan kit yang tersedia, ekstraksi DNA, desain *probe*, *libraries preparation and sequencing*, *sequence QC*, *sequence data analysed*, *genotyping data*, hingga diperoleh data *FASTQ sequence reads*.

Proses dan analisis data sequencing

Hasil bacaan Illumina di-*demultiplexing* menggunakan aplikasi *bcl2fastq* 2.17.1.14, kemudian



sisa *adapter*, bacaan terdapat N, dan bacaan dengan panjang bp <65 dibuang. Untuk memastikan rerata minimal nilai Phred diatas 30, dibuang juga 10 basa pada daerah 3'-end. Selanjutnya, hasil bacaan dipetakan ke genom acuan (Singh, *et al.*, 2013) menggunakan aplikasi Bowtie2 (Langmead, *et al.*, 2009). *Variant discovery* dan *genotyping* sampel menggunakan Freebayes v1.0.2-16 (Garrison E dan Marth G., 2012) dengan parameter ploidy= 2 dan minimal bacaan= 8. Data SNP di-*filter* menggunakan aplikasi Plink Galaxy versi 1.0 (Purcell, *et al.*, 2007) untuk membuang *missing data* dan memastikan polimorfisme hanya SNP dengan parameter minimum MAF= 0.05 dan *maximum MAF*= 0.5 serta *missing data proportion per SNP*= 0.05 dan *missing data proportion per Ind*= 0.3.

Principal component analysis

Struktur populasi dianalisis dengan aplikasi TASSEL versi 5.2.60 (Bradbury *et al.*, 2007) dan divisualisasi dengan aplikasi CurlyWhirly versi 1.19.09.04 (<https://ics.hutton.ac.uk/curlywhirly/>) (Milne *et al.*, 2014) menggunakan data SNP yang telah diperiksa dengan aplikasi plink Galaxy versi 1.0 (Purcell, *et al.*, 2007).

Analisis data fenotipe dan asosiasi

Data fenotipe diuji normalitas, ditransformasi dan dianalisis signifikansinya menggunakan aplikasi R versi 4.0.0, kemudian aplikasi GAPIT (*Genome Association and Prediction Integrated Tool*) (Lipka, *et al.*, 2012) digunakan untuk mempelajari asosiasi antar SNP dengan fenotipe. Selain model *general linear model* (GLM) digunakan juga *mixed linear model* (MLM), *Compressed MLM* (CMLM), *Enriched CMLM* (ECMLM), *Settlement of MLM Under Progressively Exclusive Relationship* (SUPER), *multiple loci mixed model* (MLMM), *Fixed and random model Circulating Probability Unification* (FarmCPU), dan *Bayesian-information and Linkage disequilibrium Iteratively Nested Keyway* (BLINK) sebagai pembanding dan menghindari pengaruh struktur populasi. Kalkulasi dengan QQ plots dan *false discovery rate* (FDR) juga disertakan untuk mempelajari dan mengkoreksi hasil (*false positive*) yang diperoleh dari model analisis GWAS yang digunakan. Kemudian, seluruh SNP yang berasosiasi nyata dengan fenotipe tertentu dicari posisi dan anotasi fungsi biologi

putative-nya dan *database genome kelapa sawit* (*E. guineensis*) pada situs <http://www.ncbi.nlm.nih.gov/> dan www.arabidopsis.org.

HASIL DAN PEMBAHASAN

Analisis data fenotipe

Hasil uji normalitas data fenotipe berdasarkan Skewness and kurtosis, Shapiro-Wilk, dan Shapiro-Francia menunjukkan bahwa semua variabel tidak terdistribusi normal, kecuali variabel C18:2. Selanjutnya, hasil transformasi data fenotipe yang tidak terdistribusi normal dapat dibagi atas dua grup, yaitu: 1. distribusi data tidak normal, namun dapat ditransformasi dengan salah satu metode *cubic*, *square*, *identity*, *square root*, *log*, *1/(square root)*, *inverse*, *1/square*, dan atau *1/cubic*; serta 2. distribusi data tidak normal dan tidak atau belum bisa dinormalisasi. Semua variabel yang bisa dinormalisasi dilanjutkan dengan analisis asosiasi, sedangkan variabel-variabel yang tidak atau belum bisa dinormalisasi tidak dilanjutkan ke analisis asosiasi.

Berdasarkan hasil pengukuran fenotipe kualitas dan kuantitas komponen tandan kedua populasi, *E. oleifera* origin Brazil dan Suriname tampak berkontribusi positif terhadap kualitas minyak pada populasi projeni hibrida interspesifik dan *pseudo-backcross*, sedangkan populasi *E. guineensis* berkontribusi positif terhadap kuantitas produksi minyaknya. Kualitas minyak pada populasi hibrida interspesifik dan *pseudo-backcross* lebih baik dibanding populasi *E. guineensis* seperti persentase asam oleat yang lebih tinggi dan asam palmitat yang lebih rendah, namun tidak lebih baik dibanding kedua populasi *E. oleifera* yang dianalisis pada penelitian ini (Tabel 1 dan Gambar 1). Selain total asam lemak tak jenuh yang lebih tinggi di banding *E. guineensis*, kandungan total karoten juga cukup tinggi pada populasi hibrida interspesifik dengan rerata >1000 ppm (Gambar 3). Komposisi asam lemak pada hibrida interspesifik pada penelitian ini masih pada kisaran yang sama dengan laporan Mozzon, Foligni dan Mannozi (2020); Bayona-Rodriguez dan Romero (2019). Sementara itu, Yaakub *et al.* (2020) melaporkan bahwa populasi *pseudo-backcross* 2 bahkan masih menunjukkan kualitas minyak yang baik khususnya pada fenotipe karoten, asam palmitat, stearat, oleat dan linoleat.

Tabel 1. Deskripsi statistik berbagai fenotipe pada berbagai populasi
 Table 1. Statistical description of various phenotypes in various populations

No.	Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Populasi F1 Brazil						Populasi F1 Suriname					
1	Buah per tandan (%)	636	56.29	10.64	9.90	86.60	56	56.58	11.45	29.00	75.20
2	Mesokarp per buah (%)	636	65.24	11.46	34.40	97.70	56	64.49	6.10	50.74	82.30
3	Minyak per mosakarp (%)	636	41.96	9.73	4.03	69.00	56	36.45	6.95	18.34	52.09
4	Minyak per tandan (%)	636	15.60	5.84	0.99	37.51	56	13.20	3.67	5.89	25.31
5	Inti per buah (%)	636	8.07	3.02	0.50	19.30	56	14.00	3.53	3.89	22.19
6	Rendemen (%)	636	13.34	4.99	0.85	32.07	56	11.29	3.14	5.04	21.64
7	Berat buah (%)	636	11.65	3.49	5.10	31.70	56	9.79	4.03	6.26	35.50
8	Karoten (ppm)	624	906	404.16	51.69	2,756	56	1,163	541.63	127.27	2,644
9	C12:0 (%)	603	0.08	0.17	0.01	1.40	56	0.10	0.37	0.01	2.70
10	C14:0 (%)	621	0.43	0.23	0.02	1.80	56	0.55	0.52	0.19	3.80
11	C16:0 (%)	623	34.64	4.86	21.62	48.30	56	37.23	3.32	28.10	45.71
12	C16:1 (%)	620	0.20	0.15	0.03	1.70	56	0.22	0.13	0.05	0.80
13	C18:0 (%)	622	2.97	1.13	0.40	14.89	56	3.39	1.66	1.00	12.28
14	C18:1 (%)	623	49.24	5.80	10.20	64.43	56	47.46	3.37	38.40	57.49
15	C18:2 (%)	622	11.71	1.80	0.20	17.98	56	10.50	2.10	7.15	15.10
16	C18:3 (%)	616	0.28	0.23	0.10	4.30	56	0.26	0.12	0.10	0.90
17	C20:0 (%)	617	0.25	0.10	0.09	1.10	56	0.24	0.08	0.10	0.40
18	C20:1 (%)	623	0.08	0.09	0.00	2.06	56	0.06	0.03	0.00	0.11

Keterangan: F1: hibrida interspesifik

No.	Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
Populasi BC1 Brazil						Populasi BC1 Suriname					
1	Buah per tandan (%)	195	62.29	12.49	15.80	89.30	69	54.50	14.34	16.40	85.00
2	Mesokarp per buah (%)	195	61.05	13.64	28.10	100.00	69	71.49	13.08	43.85	100.00
3	Minyak per mosakarp (%)	194	46.63	11.14	12.23	64.47	69	47.42	6.32	28.29	62.25
4	Minyak per tandan (%)	195	17.47	6.66	3.49	39.90	69	18.46	6.82	5.49	40.86
5	Inti per buah (%)	192	8.58	3.18	1.01	17.88	66	8.87	4.11	0.85	19.76
6	Rendemen (%)	195	14.94	5.69	2.98	34.11	69	15.78	5.83	4.69	34.94
7	Berat buah (%)	195	10.33	4.07	3.86	35.11	69	9.16	2.70	3.62	17.13
8	Karoten (ppm)	188	863.65	444.64	141.00	2,868	69	648.41	313.05	123	1,305
9	C12:0 (%)	183	0.10	0.21	0.01	1.24	69	0.07	0.14	0.01	1.04
10	C14:0 (%)	189	0.73	0.39	0.03	3.00	69	0.57	0.26	0.02	1.24

(continued)

No.	Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
		Populasi BC1 Brazil					Populasi BC1 Suriname				
11	C16:0 (%)	190	39.98	4.15	30.00	51.34	69	37.58	5.07	27.21	46.95
12	C16:1 (%)	189	0.16	0.13	0.02	1.00	69	0.12	0.06	0.03	0.27
13	C18:0 (%)	190	3.84	1.31	1.40	8.88	69	4.68	1.64	2.19	9.79
14	C18:1 (%)	190	43.83	4.20	33.42	55.00	69	44.33	5.85	31.81	60.30
15	C18:2 (%)	189	10.70	1.72	6.50	18.28	69	11.99	2.27	6.80	16.95
16	C18:3 (%)	190	0.26	0.12	0.10	1.10	69	0.25	0.06	0.14	0.44
17	C20:0 (%)	189	0.30	0.12	0.10	1.20	69	0.31	0.08	0.19	0.56
18	C20:1 (%)	165	0.10	0.03	0.06	0.22	69	0.10	0.03	0.05	0.29

Keterangan: BC: *pseudo-backcross*

No.	Variable	N	Mean	Std. Dev.	Min	Max	N	Mean	Std. Dev.	Min	Max
		Populasi <i>E. oleifera</i> Brazil					Populasi <i>E. oleifera</i> Suriname				
1	Buah per tandan (%)	9	34.72	22.46	3.90	62.82	7	62.87	8.66	54.55	80.00
2	Mesokarp per buah (%)	9	55.93	9.43	45.80	71.04	7	41.08	4.04	36.60	47.50
3	Minyak per mesokarp (%)	9	32.71	11.07	11.19	50.42	7	16.51	2.14	14.10	19.41
4	Minyak per tandan (%)	9	6.77	6.65	0.76	20.87	7	4.30	1.07	3.18	5.89
5	Inti per buah (%)	9	6.64	3.86	2.20	13.93	7	10.27	3.13	5.20	14.40
6	Rendemen (%)	9	5.79	5.68	0.65	17.84	7	3.67	0.91	2.72	5.03
7	Berat buah (%)	9	12.00	3.73	5.10	17.85	7	4.06	0.60	3.45	5.06
8	Karoten (ppm)	8	683.29	585.12	31.22	1,799	7	1,450	867.84	624	2,583
9	C12:0 (%)	7	0.22	0.25	0.01	0.60	7	0.11	0.11	0.01	0.30
10	C14:0 (%)	8	0.64	0.53	0.18	1.80	7	0.15	0.07	0.09	0.29
11	C16:0 (%)	8	28.66	6.69	22.05	42.30	7	24.23	2.36	21.02	27.00
12	C16:1 (%)	8	0.42	0.37	0.19	1.30	7	0.47	0.13	0.25	0.61
13	C18:0 (%)	8	1.87	0.69	1.20	3.20	7	2.10	1.28	0.50	4.07
14	C18:1 (%)	8	54.24	5.12	43.60	60.26	7	68.28	2.67	65.34	72.20
15	C18:2 (%)	8	13.08	2.84	8.80	16.36	7	3.98	0.72	2.96	5.14
16	C18:3 (%)	8	0.49	0.35	0.20	1.30	7	0.42	0.23	0.10	0.63
17	C20:0 (%)	8	0.35	0.55	0.07	1.70	7	0.17	0.04	0.10	0.22
18	C20:1 (%)	8	0.05	0.05	0.00	0.11	7	0.08	0.05	0.00	0.12

No.	Variable	N	Mean	Std. Dev.	Min	Max
Populasi <i>E. guineensis</i>						
1	Buah per tandan (%)	33	66.37	6.02	41.47	75.00
2	Mesokarp per buah (%)	33	78.54	14.90	45.44	94.77

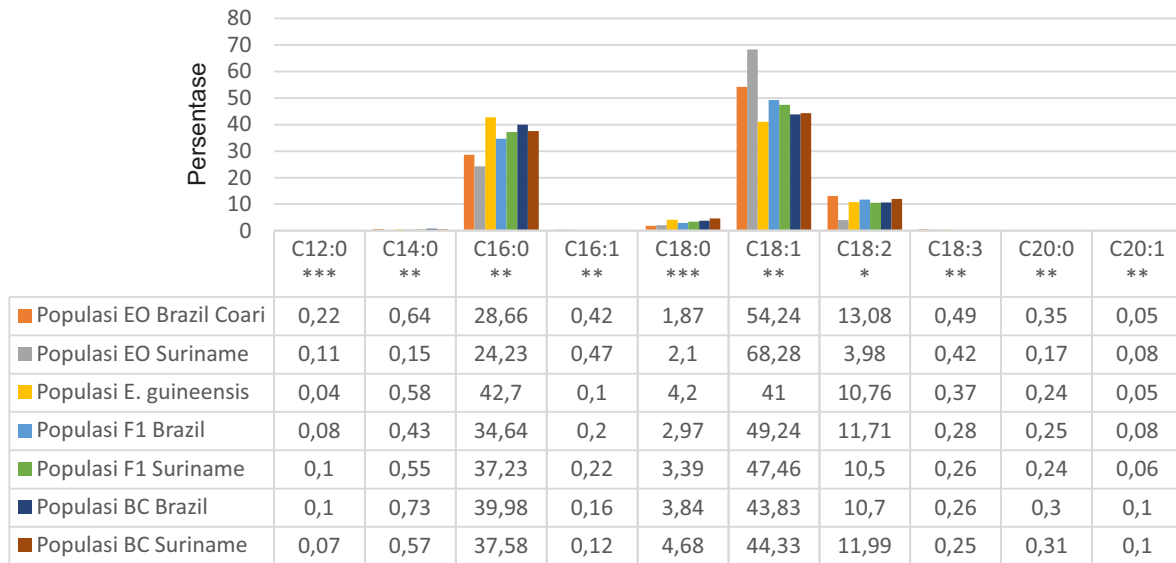
(continued)

No.	Variable	N	Mean	Std. Dev.	Min	Max
Populasi <i>E. guineensis</i>						
3	Minyak per mosakarp (%)	33	51.02	12.84	15.53	66.33
4	Minyak per tandan (%)	33	27.68	9.78	7.58	39.80
5	Inti per buah (%)	33	5.78	3.52	1.38	14.33
6	Rendemen (%)	33	23.67	8.36	6.48	34.03
7	Berat buah (%)	33	8.03	1.51	5.72	11.39
8	Karoten (ppm)	29	499.50	187.03	181	861.06
9	C12:0 (%)	19	0.04	0.05	0.01	0.20
10	C14:0 (%)	30	0.58	0.16	0.30	0.94
11	C16:0 (%)	30	42.70	4.40	32.40	50.80
12	C16:1 (%)	27	0.10	0.05	0.03	0.20
13	C18:0 (%)	30	4.20	1.38	1.50	7.67
14	C18:1 (%)	30	41.00	6.77	18.80	54.30
15	C18:2 (%)	30	10.76	3.88	7.80	24.10
16	C18:3 (%)	30	0.37	0.77	0.10	3.20
17	C20:0 (%)	29	0.24	0.09	0.10	0.44
18	C20:1 (%)	30	0.05	0.05	0.00	0.14

Rendemen atau kuantitas minyak pada populasi *E. guineensis* memang menunjukkan rerata paling tinggi sebesar 23.67%, sebaliknya populasi *E. oleifera* memiliki rendemen paling rendah sebesar 3.67% dan 5.79% pada *origin* Suriname dan Brazil Coari, berturut-turut. Sementara itu, populasi interspesifik menunjukkan fenotipe rendemen minyak berada di antara kedua populasi tetuanya, baik pada populasi hibrida interspesifik maupun pada populasi *pseudo-backcross* (Gambar 2). Dibandingkan penelitian ini, Arias *et al.* (2015) melaporkan rentang fenotipe mesokarp per buah yang lebih luas dan menarik pada populasi *E. oleifera* origin Peru, Brazil dan Ekuador, yaitu 42.51-63.71%. Sedangkan, fenotipe minyak per tandan pada populasi Brazil Coari pada penelitian ini masih sedikit lebih baik khususnya pada buah normal, namun tidak lebih baik pada populasi hibrida interspesifik bila dibandingkan dengan penelitian Astorkia *et al.* (2019).

Kualitas tandan *E. oleifera* relatif kurang baik dibanding *E. guineensis* seperti pada fenotipe

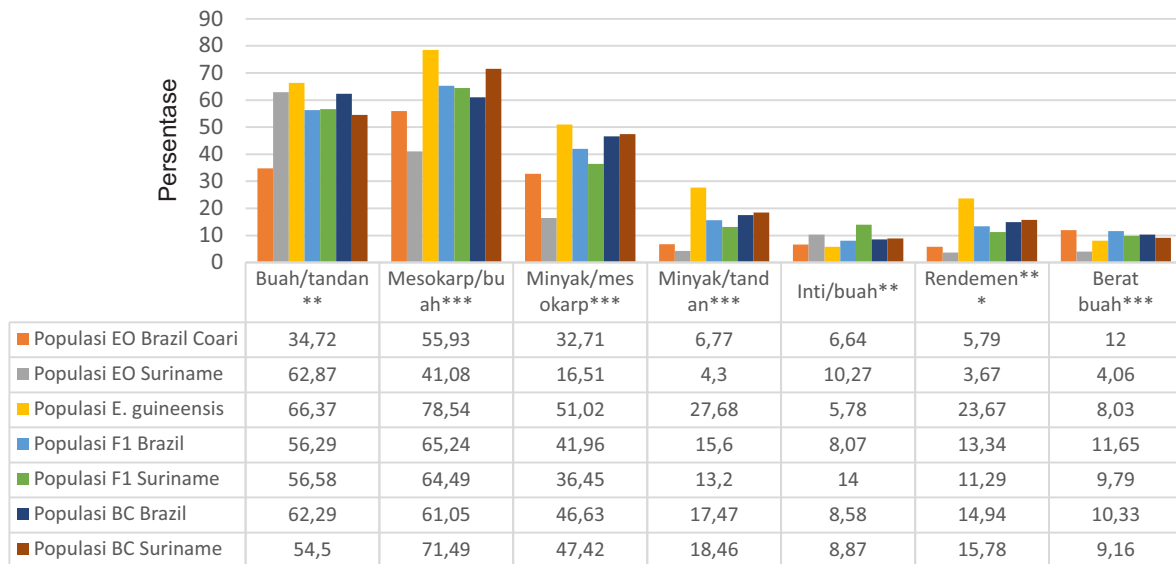
persentase buah jadi yang rendah, buah partenokarpi yang tinggi, berat tandan yang rendah, serta tandan buah aborsi yang tinggi. Fenotipe-fenotipe tersebut tampaknya ikut diwariskan pada sebagian projeni hibrida interspesifik, namun sedikit membaik pada populasi *pseudo-backcross* yang lebih mirip dengan fenotipe tandan *E. guineensis*. Fenotipe tersebut juga ditemukan oleh Arias *et al.* (2015) yang melaporkan bahwa sebanyak 5.28 – 11.88% buah tidak jadi (*parthenocarpy fruits*) pada populasi *E. oleifera* origin Peru, Brazil dan Ekuador. Socha *et al.* (2019) melaporkan bahwa nilai *fruit set* yang rendah pada *E. oleifera* tampaknya dipengaruhi oleh jumlah tepung sari yang digunakan pada percobaan penyerbukan buatan yang mereka lakukan. Sementara itu González *et al.* (2013) berhasil menganalisis buah partenokarpi merah pada populasi hibrida interspesifik termasuk buah bagian luar dan dalam yang mencapai persentase minyak per mesokarp kering maksimal pada waktu yang sama.



Keterangan: F1: hibrida interspesifik; BC: *pseudo-backcross*; EO: *E. oleifera* *data terdistribusi normal; ** data dapat dinormalisasi; *** data belum dapat dinormalisasi

Description: F1: interspecific hybrid; BC: *pseudo-backcross*; EO: *E. oleifera*; * normally distributed data; ** data can be normalized; *** currently data can not be normalized

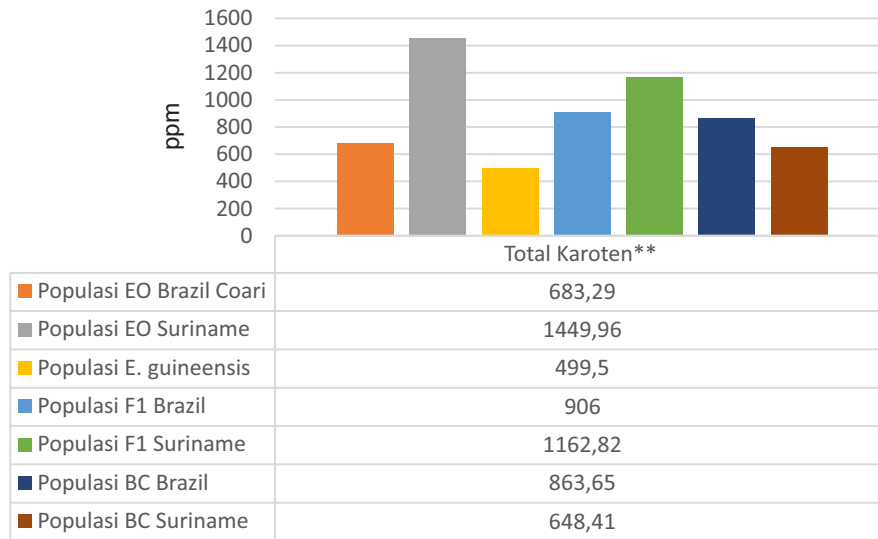
Gambar 1. Rerata fenotipe asam lemak
Figure 1. Fatty acids composition of palm oil



Keterangan: F1: hibrida interspesifik; BC: *pseudo-backcross*; EO: *E. oleifera*; *data terdistribusi normal; ** data dapat dinormalisasi; *** data belum dapat dinormalisasi

Description: F1: interspecific hybrid; BC: *pseudo-backcross*; EO: *E. oleifera*; * normally distributed data; ** data can be normalized; *** currently data can not be normalized

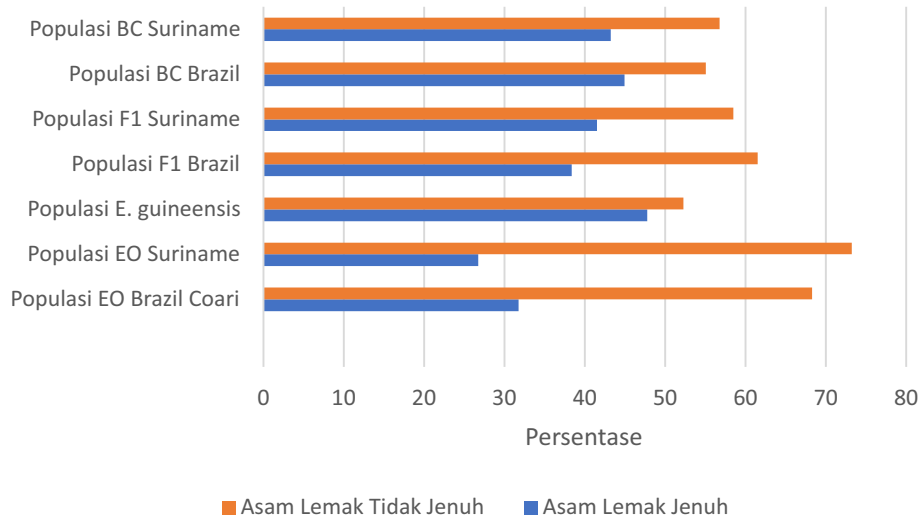
Gambar 2. Rerata fenotipe komponen tandan dan kuantitas minyak
Figure 2. Average of bunch components and oil extraction ratio



Keterangan: F1: hibrida interspesifik; BC: *pseudo-backcross*; EO: *E. oleifera*; *data terdistribusi normal; ** data dapat dinormalisasi; *** data belum dapat dinormalisasi

Description: F1: *interspecific hybrid*; BC: *pseudo-backcross*; EO: *E. oleifera*; * *normally distributed data*; ** *data can be normalized*; *** *currently data can not be normalized*

Gambar 3. Rerata fenotipe total karoten
Figure 3. Mean of total carotene



Gambar 4. Perbandingan persentase asam lemak jenuh dan asam lemak tidak jenuh. BC: *pseudo-backcross*; F1: hibrida interspesifik; dan EO: *E. oleifera*.

Figure 4. Comparison of the percentage of saturated fatty acids and unsaturated fatty acids. BC: *pseudo-backcross*; F1: *interspecific hybrid*; and EO: *E. oleifera*

Genotyping by sequencing dan analisis SNP

Jumlah bacaan yang dihasilkan dari analisis *genotyping* lebih dari 459 juta atau sekitar 798 ribu

bacaan per sampel dengan persentase bacaan pada target sebesar 98%. Rerata liputan (*coverage*) per sampel dan target sebesar 249x dengan panjang 75 bp *single read*. Analisis *genotyping by sequencing*

menghasilkan 4,969 SNP dari 5,000 SNP yang dianalisis terkait spesifitas probe SNP yang rendah, kemudian hanya 3,252 atau 65.44% SNP yang lolos pemeriksaan *in silico* dan QC. Frekuensi SNP per kromosom mulai dari 3.29% pada kromosom 16 hingga 9.75% pada kromosom 2 dengan rerata 1 SNP per setiap 220,142 pasang basa terhadap

genome (Tabel 2). Sementara itu, individu yang lolos pemeriksaan sebanyak 456 dari 500 genotipe, sedangkan fenotipe yang berdistribusi normal setelah ditransformasi sebanyak 11 variabel dari 18 variabel yang diukur. Tujuh fenotipe yang tidak atau belum bisa dinormalisasi tidak dilanjutkan untuk analisis asosiasi.

Tabel 2. Distribusi SNP yang digunakan berdasarkan kromosom
Table 2. SNP distribution based on chromosomes

Kromosom	Jumlah SNP	Persentase SNP	Rasio terhadap kromosom
1	236	7.26	289,970
2	317	9.75	206,802
3	242	7.44	248,174
4	149	4.58	384,215
5	133	4.09	390,630
6	300	9.23	147,849
7	210	6.46	206,920
8	207	6.37	194,168
9	127	3.91	299,644
10	313	9.62	101,884
11	310	9.53	89,619
12	116	3.57	248,270
13	131	4.02	212,337
14	230	7.07	105,994
15	124	3.81	196,077
16	107	3.29	199,725
Jumlah	3,252		
Rerata			220,142

Struktur populasi

Hasil analisis menunjukkan bahwa populasi *E. oleifera* Suriname dan Brazil Coari benar-benar terpisah dengan populasi *E. guineensis* dan masing-masing populasi *E. oleifera* mengelompok pada lokasi yang berbeda. Sementara itu, populasi hibrida interspesifik baik Brazil maupun Suriname berada di antara populasi *E. oleifera* dan *E. guineensis*. Seperti dugaan awal, populasi *pseudo-backcross* tampak lebih dekat ke kelompok *E. guineensis* dibanding

populasi hibrida interspesifik, karena populasi ini merupakan persilangan antara hibrida interspesifik dengan *E. guineensis*.

Analisis asosiasi

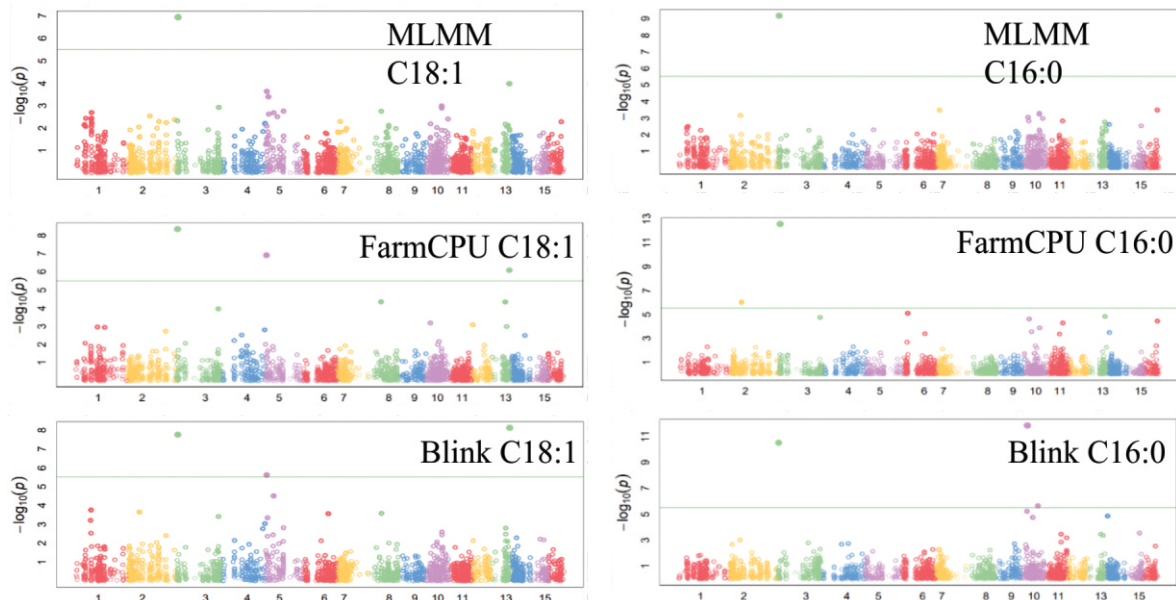
Analisis asosiasi menggunakan model MLM, FarmCPU (Liu *et al.*, 2016) dan Blink (Huang *et al.*, 2019) serta didukung analisis Q-Q plot menunjukkan bahwa 48 SNP berasosiasi nyata dengan fenotipe

tertentu, namun 19 SNP diantaranya merupakan *common* SNP atau SNP yang sama berasosiasi dengan fenotipe tertentu. Dengan demikian, analisis asosiasi hanya menemukan 29 SNP berasosiasi nyata dengan fenotipe tertentu dan telah berdasarkan koreksi dari analisis FDR *Adjusted p-values* (Tabel 3). Analisis model MLMM menghasilkan lebih sedikit SNP berasosiasi nyata dibanding model FarmCPU dan Blink, namun konsistensi hasil asosiasi tetap terjaga (Gambar 5 dan 6). SNP berasosiasi nyata yang dihasilkan dari model MLMM juga dapat ditemukan baik dari model FarmCPU maupun Blink. Hal ini menunjukkan bahwa model FarmCPU dan Blink lebih baik dalam hal menghasilkan jumlah SNP berasosiasi nyata. Sementara itu, analisis asosiasi menggunakan model lainnya seperti GLM, MLM, CMLM, ECMLM, dan SUPER memberikan hasil asosiasi yang tidak nyata antara SNP terhadap semua fenotipe.

Penelitian menggunakan 3,252 SNP dengan berbagai model asosiasi telah menunjukkan asosiasi nyata antara SNP dengan fenotipe tertentu (Tabel 3 dan 4). Selain pada daerah exon, lokasi SNP tersebut juga berada pada daerah *intragenic* dan *intergenic* (Tabel 3). Pencarian posisi basa pada *database Elaeis* sp. di situs <http://www.ncbi.nlm.nih.gov/> terhadap 29 SNP menemukan bahwa 17 SNP

berada pada gen-gen yang memiliki fungsi biologi tertentu (Tabel 4). Walau demikian, fungsi beberapa gen *putative* tersebut masih belum terdeskripsi dengan baik pada *database Arabidopsis* (www.arabidopsis.org), yang juga merupakan sumber acuan anotasi.

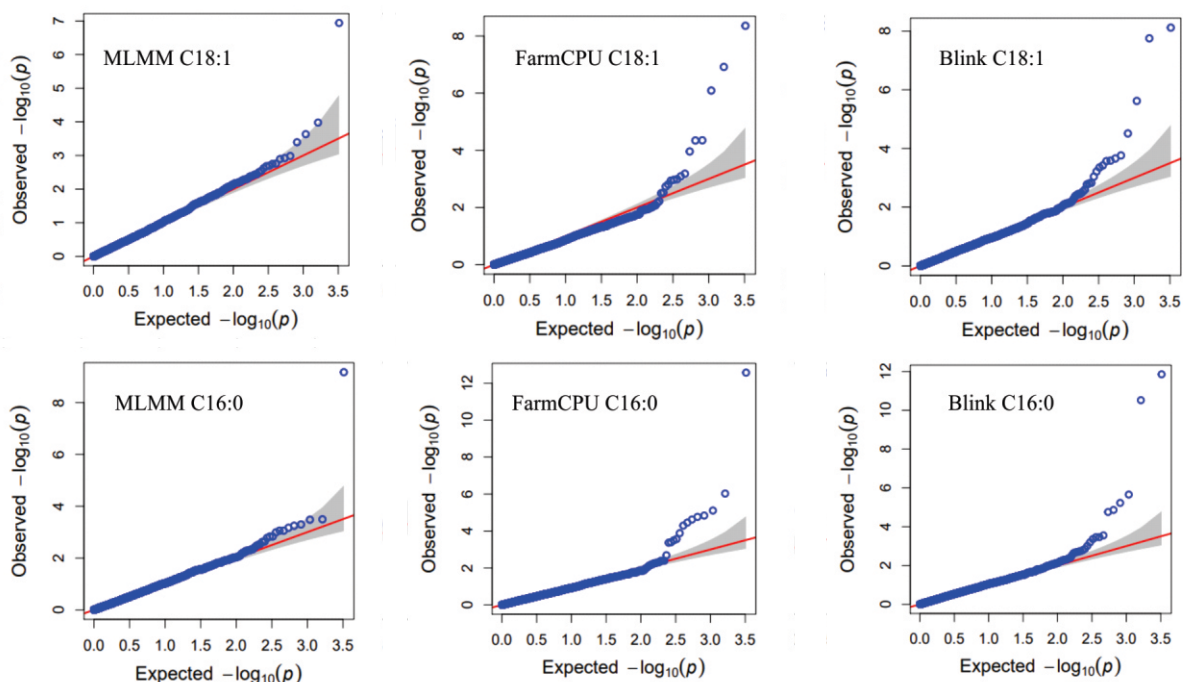
Tiga SNP ditemukan berada pada exon gen tertentu, yaitu SNP4416, SNP349 dan SNP3865, sedangkan 15 SNP lainnya berada pada *intragenic* gen tertentu. Ithnin *et al.* (2021) juga melaporkan bahwa sebagian besar SNP yang berasosiasi nyata dengan fenotipe tertentu berada pada daerah *intergenic*, sedangkan hanya ada 8 SNP yang berada pada daerah exon. Sementara itu, sebanyak 4 SNP diantaranya merupakan *common* SNP pada fenotipe C16:0 dan C18:1 serta C20:0 dan C20:1. *Common* SNP biasanya ditemukan pada fenotipe yang berkorelasi relatif tinggi, baik berkorelasi negatif maupun positif seperti fenotipe C16:0 dengan C18:1 (SNP3052) dan fenotipe C20:0 dengan C20:1 (SNP371, SNP2396, dan SNP3052). Ithnin *et al.* (2021) melaporkan bahwa sebanyak 8 *common* SNP dari *single locus* berasosiasi dengan fenotipe C18:1 dan *iodine value* (IV), serta satu SNP dapat berasosiasi dengan lima fenotipe yang saling berkorelasi baik positif maupun negatif.



Gambar 5. Manhattan plot fenotipe C18:1 dan C16:0 dengan analisis model MLMM, FarmCPU dan Blink.
 Figure 5. Manhattan plots of the C18:1 and C16:0 using the MLMM, FarmCPU and Blink model analysis.

Beberapa fenotipe pada kelapa sawit telah dilaporkan berasosiasi nyata dengan pendekatan GWAS, bahkan beberapa SNP telah menjalani uji validasi. Fenotipe tersebut antara lain minyak per mesokarp segar dan minyak per mesokarp kering (Bai *et al.*, 2017; Teh *et al.*, 2016), tinggi batang dan berat

tandan (Pootakham *et al.*, 2015), berat kering daun, panjang pelepah dan jumlah anak daun (Osorio-Guarín *et al.*, 2019), pertumbuhan meninggi (Babu *et al.*, 2019), minyak per tandan, produktivitas CPO, dan panjang pelepah (Ithnin *et al.*, 2017), serta kualitas minyak (Astorkia *et al.*, 2019).



Gambar 6. QQ plot fenotipe C18:1 dan C16:0 dengan analisis model MLMM, FarmCPU dan Blink.

Figure 6. QQ plots of the C18:1 and C16:0 using the MLMM, FarmCPU and Blink model analysis.

Tabel 3. Marka SNP yang berasosiasi nyata berdasarkan model MLMM, FarmCPU, dan Blink

Table 3. Significantly associated SNP markers based on the MLMM, FarmCPU, and Blink models analysis

No	Feno	SNP	Krm	Posisi	O	P values	MAF	FDR		Efek	Model asosiasi	Lokasi pada gen
								Adjusted	P values			
1	FB	SNP4774	8	32,184,986	+	1.88E-10	0.29	6.11.E-07	NA	Blink	NA	
		SNP4987	9	33,371,105	+	1.81E-07	0.4	5.89.E-04	452.3	FarmCPU	intragenic	
		SNP4774	8	32,184,986	+	1.06E-06	0.29	1.73.E-03	-362.2	FarmCPU	NA	
2	KF	SNP4416	7	11,061,375	+	3.14E-07	0.45	1.02.E-03	0.13	FarmCPU	exon	
		SNP4907	9	22,646,424	+	8.27E-07	0.49	1.34.E-03	0.23	FarmCPU	NA	
		SNP4594	8	11,732,185	+	1.95E-06	0.44	2.11.E-03	-0.27	FarmCPU	intragenic	

(continued)

No	Feno	SNP	Krm	Posisi	O	P values	MAF	FDR		Efek	Model	Lokasi pada gen
								Adjusted	P values			
3	Karoten	SNP349	10	204,255	+	2.16E-06	0.43	6.78.E-03	NA	Blink	exon	
		SNP5138	3	31,369,218	+	4.17E-06	0.03	6.78.E-03	NA	Blink	NA	
		SNP266	1	36,038,888	+	5.71E-07	0.48	1.53.E-03	2.45	FarmCPU	intragenic	
		SNP3626	6	2,719,936	+	9.41E-07	0.26	1.53.E-03	-2.42	FarmCPU	intragenic	
4	C14:0	SNP557	10	13,888,305	+	1.11E-07	0.48	3.61.E-04	NA	Blink	NA	
		SNP2356	15	23,722,659	+	2.83E-07	0.31	4.60.E-04	NA	Blink	intragenic	
		SNP1299	12	1,953,788	+	1.69E-06	0.39	1.83.E-03	NA	Blink	intragenic	
		SNP3054	3	2,747,142	+	1.74E-08	0.48	5.65.E-05	-0.21	FarmCPU	intragenic	
		SNP4620	8	21,334,892	+	2.02E-06	0.43	2.01.E-03	-0.14	FarmCPU	NA	
		SNP1299	12	1,953,788	+	2.20E-06	0.39	2.01.E-03	-0.1	FarmCPU	intragenic	
		SNP1734	14	301,423	-	2.83E-06	0.35	2.01.E-03	0.15	FarmCPU	NA	
		SNP2356	15	23,722,659	+	3.08E-06	0.31	2.01.E-03	0.12	FarmCPU	intragenic	
5	C16:0	SNP412	10	6,517,273	+	1.41E-12	0.49	4.58.E-09	NA	Blink	intragenic	
		SNP3052	3	1,730,558	+	3.04E-11	0.45	4.94.E-08	NA	Blink	NA	
		SNP753	10	20,529,991	+	2.22E-06	0.45	2.40.E-03	NA	Blink	intragenic	
		SNP3052	3	1,730,558	+	2.68E-13	0.45	8.71.E-10	-160.37	FarmCPU	NA	
		SNP2789	2	13,998,128	+	9.38E-07	0.44	1.53.E-03	122.98	FarmCPU	NA	
		SNP3052	3	1,730,558	+	6.74E-10	0.45	2.19.E-06	NA	MLMM	NA	
6	C18:1	SNP1716	13	22,965,019	-	7.67E-09	0.42	2.49.E-05	NA	Blink	NA	
		SNP3052	3	1,730,558	+	1.77E-08	0.45	2.88.E-05	NA	Blink	NA	
		SNP3478	5	2,333,663	+	2.40E-06	0.23	2.61.E-03	NA	Blink	NA	
		SNP3052	3	1,730,558	+	4.40E-09	0.45	1.43.E-05	0	FarmCPU	NA	
		SNP3478	5	2,333,663	+	1.21E-07	0.23	1.97.E-04	0	FarmCPU	NA	
		SNP1716	13	22,965,019	-	8.12E-07	0.42	8.80.E-04	0	FarmCPU	NA	
		SNP3052	3	1,730,558	+	1.15E-07	0.45	3.74.E-04	NA	MLMM	NA	
7	C20:0	SNP3865	6	29,329,963	+	2.01E-12	0.4	6.55.E-09	NA	Blink	exon	
		SNP371	10	4,239,504	+	8.16E-12	0.4	1.33.E-08	NA	Blink	NA	
		SNP2396	16	7,712,541	+	5.58E-09	0.5	6.05.E-06	NA	Blink	intragenic	
		SNP4651	8	25,214,899	+	2.87E-07	0.46	2.33.E-04	NA	Blink	intragenic	
		SNP4654	8	25,300,364	+	1.17E-06	0.44	7.60.E-04	NA	Blink	intragenic	
		SNP4670	8	26,050,353	+	1.96E-06	0.48	9.49.E-04	NA	Blink	intragenic	
		SNP4683	8	26,062,104	+	2.04E-06	0.46	9.49.E-04	NA	Blink	intragenic	

(continued)

No	Feno	SNP	Krm	Posisi	O	P values	FDR		Efek	Model	Lokasi pada gen
							MAF	Adjusted P values			
		SNP3865	6	29,329,963	+	1.17E-08	0.4	3.82.E-05	-0.18	FarmCPU	exon
		SNP2396	16	7,712,541	+	1.54E-07	0.5	2.50.E-04	0.09	FarmCPU	intragenic
8	C20:1	SNP3865	6	29,329,963	+	6.21E-13	0.4	2.02.E-09	NA	Blink	exon
		SNP371	10	4,239,504	+	1.63E-10	0.4	2.65.E-07	NA	Blink	NA
		SNP2396	16	7,712,541	+	6.51E-10	0.5	7.06.E-07	NA	Blink	intragenic
		SNP4655	8	25,436,057	+	6.72E-07	0.45	5.47.E-04	NA	Blink	intragenic
		SNP2396	16	7,712,541	+	1.29E-10	0.5	4.19.E-07	0.03	FarmCPU	intragenic
		SNP3865	6	29,329,963	+	9.02E-10	0.4	1.47.E-06	-0.05	FarmCPU	exon
		SNP371	10	4,239,504	+	3.31E-07	0.4	3.59.E-04	-0.02	FarmCPU	NA
		SNP3865	6	29,329,963	+	3.42E-06	0.4	1.11.E-02	NA	MLMM	exon

Keterangan= FB: persentase buah per tandan; KF: persentase inti per buah; C14:0: miristat; C16:0 palmitat; C18:1: oleat; C20:0: arakidat; dan C20:1: gadoleat; Krm: kromosom; O: orientasi

Description: FB: percentage of fruit per bunch; KF: percentage of kernel to fruit; C14:0: myristate; C16:0 palmitate; C18:1: oleic; C20:0: arachidate; and C20:1: gadoleic; Krm: chromosome; O: orientation

Tabel 4. Asosiasi 17 SNP dengan gen terkait dan anotasi pada database www.arabidopsis.org dan www.ncbi.nlm.nih.gov

Table 4. Association of 17 SNPs related to genes and annotations on the www.arabidopsis.org and www.ncbi.nlm.nih.gov databases

No	Fenotipe	SNP	Simbol Gen	Gen Terkait
1	FB	SNP4987	LOC105051957	CMP-sialic acid transporter 5; termasuk keluarga protein membran mengangkut gula nukleotida dari sitoplasma ke vesikula golgi
2	KF	SNP4416	LOC105048536	Coiled-coil domain-containing protein 130
		SNP4594	LOC105049739	Ubiquitin thioesterase otubain-like
3	Karoten	SNP349	LOC105052147	Laccase-7; Degradasi lignin dan detoksifikasi produk turunan lignin.
		SNP266	LOC105054133	Tetratricopeptide repeat protein 38
		SNP3626	LOC105046526	Calcium-dependent protein kinase 20
4	C14:0	SNP2356	LOC105058794	Alpha-glucan water dikinase, chloroplastic; dibutuhkan untuk degradasi pati
		SNP1299	LOC105054590	Ribosomal RNA small subunit methyltransferase, chloroplastic, RNA binding site

(continued)

No	Fenotipe	SNP	Simbol Gen	Gen Terkait
		SNP3054	LOC105040289	<i>V-type proton ATPase subunit F</i> ; Subunit kompleks V1 perifer dari vacuolar ATPase penting untuk perakitan atau fungsi katalitik. V-ATPase bertanggung jawab untuk mengasamkan berbagai kompartemen intraseluler dalam sel eukariotik
5	C16:0	SNP412	LOC105052282	<i>Probable serine/threonine-protein kinase WNK5</i> ; ATP binding dan aktifitas protein kinase
		SNP753	LOC105052908	<i>Chaperone protein dnaJ 10</i>
6	C20:0	SNP3865	LOC105047059	<i>Eukaryotic translation initiation factor 5-like</i>
		SNP2396	LOC105059273	<i>Growth-regulating factor</i>
		SNP4651	LOC105050107	Protein TPR3; regulasi transkripsi, DNA-template
		SNP4670	LOC105050151	<i>Kinesin-like protein KIN-4A</i> ; Nucleotide binding; ATP binding
		SNP4654	LOC105050114	<i>Pyrophosphate-energized vacuolar membrane proton pump</i>
7	C20:1	SNP3865	LOC105047059	<i>Eukaryotic translation initiation factor 5-like</i>
		SNP2396	LOC105059273	<i>Growth-regulating factor 5</i> ; termasuk keluarga gen GRF, yang mengandung domain penargetan nuklir. Mutan menghasilkan daun yang lebih kecil yang menunjukkan peran gen dalam perkembangan daun. Terekspresi pada akar, pucuk dan bunga
		SNP4655	LOC105050121	Voltage-gated <i>chloride channel-like protein CLC-g</i>

KESIMPULAN

Hasil analisis menemukan 29 SNP berasosiasi nyata dengan fenotipe tertentu, namun hanya 17 SNP yang berada pada daerah gen-gen yang memiliki fungsi biologi tertentu. Tiga SNP diantaranya ditemukan berada pada exon gen tertentu, sedangkan 15 SNP lainnya berada pada *intra-genic* gen tertentu. Empat SNP diantaranya merupakan *common* SNP pada fenotipe C16:0 dan C18:1 serta C20:0 dan C20:1.

Penelitian ini menggambarkan potensi penggunaan GWAS sebagai pendekatan alternatif dalam program pemuliaan *backcross* *E. oleifera*. Penelitian lanjutan termasuk validasi marka sangat layak dilakukan sebelum penerapannya dalam seleksi berbantuan marka.

UCAPAN TERIMA KASIH

Penulis mengucapkan penghargaan atas kerja keras dari para Pemulia kelapa sawit terdahulu di PPKS yang telah menanam banyak nomor percobaan ini. Penghargaan juga kami sampaikan kepada Badan Pengelola Dana Perkebunan Kelapa Sawit yang telah mendukung penelitian ini sehingga dapat berjalan dengan baik. Tidak lupa juga ucapan penghargaan kepada Manajemen PPKS yang telah mendukung tiada henti atas keberlanjutan penelitian ini.

DAFTAR PUSTAKA

Arias, D., González, M., Prada, F., Ayala-Diaz, I., Montoya, C., Daza, E. and Romero, H.M., 2015. Genetic and phenotypic diversity of natural American oil palm (*Elaeis oleifera* (HBK)



- Cortés) accessions. *Tree Genetics & Genomes*, 11(6), p.122.
- Astorkia, M., Hernandez, M., Bocs, S., Lopez de Armentia, E., Herran, A., Ponce, K., León, O., Morales, S., Quezada, N., Orellana, F. and Wendra, F., 2019. Association mapping between candidate gene SNP and production and oil quality traits in interspecific oil palm hybrids. *Plants*, 8(10), p.377.
- Babu, B.K., Mathur, R.K., Ravichandran, G. and Venu, M.V.B., 2019. Genome-wide association study (GWAS) for stem height increment in oil palm (*Elaeis guineensis*) germplasm using SNP markers. *Tree Genetics & Genomes*, 15(3), p.40.
- Bai, B., Wang, L., Lee, M., Zhang, Y., Alfiko, Y., Ye, B.Q., Wan, Z.Y., Lim, C.H., Suwanto, A., Chua, N.H. and Yue, G.H., 2017. Genome-wide identification of markers for selecting higher oil content in oil palm. *BMC plant biology*, 17(1), p.93.
- Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., Buckler, E.S. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- Bayona-Rodriguez, C.J. and Romero, H.M., 2019. Physiological and agronomic behavior of commercial cultivars of oil palm ('*Elaeis guineensis*') and OxG hybrids ('*Elaeis oleifera*'x'*Elaeis guineensis*') at rainy and dry seasons. *Australian Journal of Crop Science*, 13(3), p.424.
- Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*, 6(2), pp.80-92.
- Corley, R.H.V. and Tinker P.B. 2016. *The Oil Palm*, Fifth Edition. R.H. Published by John Wiley & Sons, Ltd.
- Durand-Gasselín, T., De Franqueville, H., Amblard, P., Breton, F., Jacquemard, J.C., Syaputra, I., Cochard, B., Louise, C. and Nouy, B., 2011. Breeding for sustainable palm oil.
- Garrison E dan Marth G., 2012. Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907 [q-bio.GN]*.
- Gibson, G., 2018. Population genetics and GWAS: a primer. *PLoS biology*, 16(3), p.e2005485.
- González, D.A., Cayón, G., López, J.E. and Alarcón, W.H., 2013. Development and maturation of fruits of two *Indupalma* OxG hybrids (*Elaeis oleifera* x *Elaeis guineensis*). *Agronomía Colombiana*, 31(3), pp.343-351.
- Hickey, J.M., Chiurugwi, T., Mackay, I., Powell, W., Eggen, A., Kilian, A., Jones, C., Canales, C., Grattapaglia, D., Bassi, F. and Atlin, G., 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature genetics*, 49(9), p.1297.
- HiSeq 2500 System (illumina.com). Diakses pada 24 Maret 2021.
- Huang, M., Liu, X., Zhou, Y., Summers, R. M. & Zhang, Z., 2019. BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *G i g a s c i e n c e g i y 1 5 4*. doi:10.1093/gigascience/giy154
- Ithnin, M., Vu, WT., Shin, MG., Suryawanshi, V., Sherbina, K., Zolkafli, SH., Serdari, NM., Amiruddin, MD., Abdullah, N., Mustaffa, S., Marjuni, M., Nookiah, R., Kushairi, A., Marjoram, P., Nuzhdin, SV., Chang, PL., Singh, R., 2021. Genomic diversity and genome-wide association analysis related to yield and fatty acid composition of wild American oil palm. *Plant Sci*, 304:110731. doi: 10.1016/j.plantsci.2020.110731. Epub 2020 Oct 24. PMID: 33568284.
- Ithnin, M., Xu, Y., Marjuni, M., Serdari, N.M., Amiruddin, M.D., Low, E.T.L., Tan, Y.C., Yap, S.J., Ooi, L.C.L., Nookiah, R. and Singh, R., 2017. Multiple locus genome-wide association studies for important economic traits of oil palm. *Tree Genetics & Genomes*, 13(5), p.103.
- Kawahara, Y., de la Bastide, M., Hamilton J. P., Kanamori, H., McCombie, W. R., Ouyang, S.,

- Schwartz, D. C., Tanaka, T., Wu, J., Zhou, S., Childs, K. L., Davidson, R. M., Lin, H., Quesada-Ocampo, L., Vaillancourt, B., Sakai, H., Lee, S. S., Kim, J., Numa, H., Itoh, T., Buell, C. R., Matsumoto, T. 2013. Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4
- Kim, E.J., Park, S.W., Hong, W.J., Silva, J., Liang, W., Zhang, D., Jung, K.H., and Kim, Y.J. 2020. Genome-wide analysis of *RopGEF* gene family to identify genes contributing to pollen tube growth in rice (*Oryza sativa*). *BMC Plant Biol* 20, 95. <https://doi.org/10.1186/s12870-020-2298-5>
- Korte, A. and Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, 9(1), p.29.
- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>.
- Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, 25(14), pp.1754-1760.
- Li, Y., Zhu, F. L., Zheng, X. W., Hu, M. L., Dong, C., Diao, Y., Wang, Y.W., and Hu, Z. L., 2020. Comparative population genomics reveals genetic divergence and selection in lotus, *Nelumbo nucifera*. *BMC genomics*, 21(1), 1-13. <https://doi.org/10.1186/s12864-019-6376-8>
- Liu, X., Huang, M., Fan, B., Buckler, E. S. & Zhang, Z., 2016. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLoS Genet*. 12, e1005767 (2016).
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Xie, K.Q ... & Zhang, Z., 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics*, 28(18), 2397-2399.
- Lubis, A. U., 2008. Kelapa sawit (*Elaeis guineensis* Jacq.) di Indonesia Edisi 2. Pusat Penelitian Kelapa Sawit, Medan Indonesia.
- Luikart, G., Kardos, M., Hand, B.K., Rajora, O.P., Aitken, S.N. and Hohenlohe, P.A., 2018. Population genomics: advancing understanding of nature. In *Population Genomics* (pp. 3-79). Springer, Cham.
- Milne, I., Stephen, G., Bayer, M., Shaw, P.D., Raubach, S., Hearne, S., Singh, S., Wenzl, P., Marshall, D. 2014. Graphical applications for visualization and analysis of genotype data sets. *Plant and Animal Genome XXII*, 2014. San Diego.
- Montoya, C., Lopes, R., Flori, A., Cros, D., Cuellar, T., Summo, M., Espeout, S., Rivallan, R., Risterucci, A.M., Bittencourt, D. and Zambrano, J.R., 2013. Quantitative trait loci (QTLs) analysis of palm oil fatty acid composition in an interspecific pseudo-backcross from *Elaeis oleifera* (HBK) Cortés and oil palm (*Elaeis guineensis* Jacq.). *Tree genetics & genomes*, 9(5), pp.1207-1225.
- Moreno, L.P. and Romero, H.M., 2015. Phenology of the reproductive development of *Elaeis oleifera* (Kunth) Cortes. *Agronomía Colombiana*, 33(1), pp.29-35.
- Mozzon, M., Pacetti, D., Lucci, P., Balzano, M. and Frega, N.G., 2013. Crude palm oil from interspecific hybrid *Elaeis oleifera* × *Elaeis guineensis*: Fatty acid regiodistribution and molecular species of glycerides. *Food chemistry*, 141(1), pp.245-252.
- Mozzon, M., Foligni, R. and Mannozi, C., 2020. Current Knowledge on Interspecific Hybrid Palm Oils as Food and Food Ingredient. *Foods*, 9(5), p.631.
- Nurkhoiry, R., 2017. Daya saing ekspor minyak kelapa sawit Indonesia. *J. Pen. Kelapa Sawit*, 25(2): 105 – 116.
- Osorio-Guarín, J.A., Garzón-Martínez, G.A., Delgadillo-Duran, P., Bastidas, S., Moreno, L.P., Enciso-Rodríguez, F.E., Cornejo, O.E. and Barrero, L.S., 2019. Genome-wide association study (GWAS) for morphological and yield-related traits in an oil palm hybrid (*Elaeis oleifera* × *Elaeis guineensis*) population. *BMC Plant Biology*, 19(1), p.533.
- Pootakham, W., Jomchai, N., Ruang-areerate, P., Shearman, J.R., Sonthirod, C., Sangsrakru, D.,



- Tragoonrung, S. and Tangphatsornruang, S., 2015. Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics*, 105(5-6), pp.288-295.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC, 2007. American Journal of Human Genetics.
- Ravelombola, W.S., Qin, J., Shi, A. et al., 2019. Genome-wide association study and genomic selection for soybean chlorophyll content associated with soybean cyst nematode tolerance. *BMC Genomics* 20, 904. <https://doi.org/10.1186/s12864-019-6275-z>.
- <https://de.reuters.com/article/indonesia-palm-oil-mistry-idINKBN28D24L>. Diakses pada 23 Desember 2020.
- Scherer, A., 2017. GWAS, 3rd Edition. Golden Helix, Inc.
- Singh R, Ong-Abdullah M, Low E-TL, Manaf MAA, Rosli R, Nookiah R, et al., 2013. Oil palm genome sequence reveals divergence of interfertile species in old and new worlds. *Nature*. 500 : 335 . <https://doi.org/10.1038/nature12309>.
- Siregar, Heri., Rahmadi, H., Setiowati, R., Suprianto, E., 2018. Vegetative and bunch components of first pseudobackcross of *Elaeis oleifera* Brazil and Suriname origin in North Sumatra, Indonesia. *Jurnal Penelitian Kelapa Sawit*. 26(1). 23-28.
- Siregar, Heri., Rahmadi, H., Wening, S., Suprianto, E., 2018. Fatty acid and carotene composition in *elaeis oleifera*, interspecific hybrid, and the first pseudo-backcross in North Sumatra, Indonesia. *Jurnal Penelitian Kelapa Sawit*. 26(2). 91-101.
- Socha, J., Cayón, D., Ligarreto, G., & Chaves, G., 2019. Effect of pollen doses on fruit formation and oil production in two hybrid palm genotypes (*Elaeis oleifera* HBK Cortes x *Elaeis guineensis* Jacq.). *Agronomía Colombiana*, 37(1), 12-17.
- Teh, C.K., Ong, A.L., Kwong, Q.B., Apparow, S., Chew, F.T., Mayes, S., Mohamed, M., Appleton, D. and Kulaveerasingam, H., 2016. Genome-wide association study identifies three key loci for high mesocarp oil content in perennial crop oil palm. *Scientific reports*, 6(1), pp.1-7.
- Wang, P., Luo, Y., Huang, J., Gao, S., Zhu, G., Dang, Z., ... & Chen, Y., 2020. The genome evolution and domestication of tropical fruit mango. *Genome biology*, 21(1), 1-17. <https://doi.org/10.1186/s13059-020-01959-8>
- Yaakub, Z., Kamaruddin, K., Singh, R., Mustafa, S., Marjuni, M., Ting, N.C., Amiruddin, M.D., Leslie, L.E.T., Cheng-Li, O.L., Sritharan, K. and Nookiah, R., 2020. An Integrated Linkage Map of Interspecific Backcross 2 (BC2) Populations Reveals QTLs Associated with Fatty Acid Composition and Vegetative Parameters Influencing Compactness in Oil Palm.