Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

¹H NMR metabolomic signatures associated with oil-related bunch traits in within-family oil palm (Elaeis guineensis) progenies

Arfan Nazhri Simamora

arfan.nazhri@gmail.com

Indonesian Oil Palm Research Institute

Yurna Yenni

Indonesian Oil Palm Research Institute

Hernawan Yuli Rahmadi

Indonesian Oil Palm Research Institute

Research Article

Keywords: Elaeis guineensis, ¹H NMR metabolomics, Spear leaf, Metabolomics-assisted breeding, Progeny testing, Carbon–nitrogen metabolism, Oil yield, Candidate biomarkers

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11505630/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

^{1}H NMR metabolomic signatures associated with oil-related bunch traits in within-family oil palm (*Elaeis guineensis*) progenies

Arfan Nazhri Simamora[1,*], Yurna Yenni[1] and Hernawan Yuli Rahmadi[1]

[1]Plant Breeding Research Group, Indonesian Oil Palm Research Institute, Jl. Brigjend Katamso No. 51, Medan 20158, North Sumatera, Indonesia

*Corresponding author: arfan.nazhri@gmail.com

Abstract

Background Oil palm (*Elaeis guineensis* Jacq.) improvement relies on progeny testing, but bunch performance is expressed only after several years of field growth, and oil-related traits vary considerably even among genetically related progenies. Whether early-stage leaf metabolism is associated with this variation is unknown. We therefore compared the spear leaf metabolomes of top- and bottom-performing palms selected within the same biparental crosses.

Results Unopened spear leaves (leaf 0) were collected from 36 palms, one top and one bottom performer from each of 18 biparental crosses, and profiled by ^{1}H NMR. Top-class palms had higher mesocarp-to-fruit (+6.4%), oil-to-dry mesocarp (+4.4%) and oil-to-bunch (+22.4%) ratios than bottom-class palms (all $P < 0.001$). Of 123 metabolites identified, those with the highest variable importance in projection (VIP) scores, including guanidinoacetic acid, xylitol, methylguanidine, threonic acid and lysine, contributed to the class separation seen in partial least squares discriminant analysis. Several metabolites correlated with oil-related traits, notably L-arabitol (positively) and glutamine (negatively), but none remained significant after false discovery rate (FDR) correction. Six pathways were enriched (FDR < 0.01), led by alanine, aspartate and glutamate metabolism (impact 0.849) and including arginine biosynthesis, glyoxylate and dicarboxylate metabolism, and the TCA cycle. Five metabolite ratios discriminated the classes (AUC 0.790–0.814), the best being β-hydroxyisovaleric acid/methylguanidine (AUC = 0.814) and guanidinoacetic acid/pimelic acid (AUC = 0.806).

Conclusions Within biparental crosses, top- and bottom-performing palms differed in spear leaf profiles involving carbon-conserving and nitrogen-related (amino acid and guanidino compound) metabolism, a pattern consistent with coordinated carbon–nitrogen metabolism supporting oil accumulation. The candidate metabolite ratios are leads for non-destructive early selection, but the extreme-group design, the small sample size and the absence of FDR-significant individual correlations mean that they must be validated in independent, unselected populations and environments before use in breeding.

Keywords: Elaeis guineensis; ^{1}H NMR metabolomics; Spear leaf; Metabolomics-assisted breeding; Progeny testing; Carbon–nitrogen metabolism; Oil yield; Candidate biomarkers

Background

Oil palm (*Elaeis guineensis* Jacq.) is the world's most productive oil-bearing crop and supplies a major share of global demand for edible oils, cosmetics and industrial products [1]. As demand continues to rise, breeding programs are placing increasing emphasis on secondary traits such as abiotic and biotic stress tolerance and enhanced nutritional content [2].

Commercial oil palm breeding relies predominantly on reciprocal recurrent selection (RRS) [3]. In RRS, elite parental lines are identified through progeny testing, in which the performance of progenies from related crosses is used to estimate the breeding value of their parents [1, 4]. Because the performance of an individual palm is confounded by environment and age, comparing families that share ancestry allows additive and cross-specific genetic effects to be separated, so that the most promising parental populations can be selected and recombined to assemble the next cycle of planting material. In advanced breeding cycles, progenies are typically expected to show cumulative genetic gains over previous generations. However, field evaluation frequently reveals substantial phenotypic variation in yield-related traits among individual progenies, even within closely related full-sib families grown under similar environmental conditions [4–6]. Bunch performance is, moreover, expressed only after several years of field growth, which lengthens each selection cycle [1]. This suggests that considerable biological variation remains despite shared parentage and uniform field management. We hypothesized that this phenotypic divergence is not dictated solely by genetic variation but also reflects differences in functional metabolic regulation.

While genomics provides the biological blueprint of what a plant can do, the metabolome integrates the outcomes of transcriptional regulation, enzyme kinetics and environmental interactions [7, 8]. Because metabolomics captures these biochemical activities, it links genetic background to complex agronomic traits and offers a closer representation of the functional phenotype than genomic information alone [9–11]. In oil palm, metabolomic approaches have increasingly been used to study fruit development, abiotic stress responses and biomarker discovery, providing a functional framework for metabolomics-assisted breeding.

Total oil yield in oil palm results from structural sink capacity (mesocarp-to-fruit ratio, M/F) and biochemical efficiency (oil-to-dry mesocarp ratio, O/DM), which together determine the oil-to-bunch ratio (O/B). Metabolomics has captured the functional state underlying these

traits: sucrose and starch dynamics and organic acid remodeling have been proposed as early predictors of high oil yield and are linked to high O/DM [12–14], whereas early mesocarp development depends on active amino acid and organic acid metabolism [13, 15]. Maximizing lipid productivity further requires tight coordination between carbon and nitrogen pathways [16].

Previous metabolomic studies in oil palm have predominantly sampled developing fruit or mesocarp to study lipid biosynthesis [12, 14, 15, 17–19]. Such profiling gives valuable retrospective insight into fruit ripening but is of limited use for predictive breeding, because palms need several years to reach the reproductive phase. In addition, these studies typically compare highly divergent genetic backgrounds, such as tenera versus pisifera fruit types [20], which makes it difficult to separate metabolic drivers from genetic background effects.

This study addresses these gaps in two ways. First, we compared contrasting top and bottom performers selected within each of 18 biparental crosses, which reduces genetic and environmental background variation relative to comparisons among unrelated material. Second, instead of mature fruit, we profiled unopened spear leaves as an early-stage proxy. Based on the source–sink relationship, we hypothesized that the metabolic configuration of these actively developing source tissues reflects the capacity of the palm to allocate resources to the fruit sink. We therefore aimed to characterize the spear leaf metabolic differences associated with superior oil accumulation and to evaluate whether they could provide a non-destructive basis for metabolomics-assisted breeding in oil palm.

Results

Contrasting oil-related traits between performance classes

Top and bottom classes were defined within each cross from bunch performance data using PROMETHEE multi-criteria decision analysis (Table 1; see Methods). Linear mixed-effects models showed significant differences in bunch performance traits between the top and bottom classes (Table 2). The oil-related traits O/DM, M/F and O/B showed notable differences between the classes, with O/B showing the greatest difference (22.35%). The correlation heatmap showed that the oil-related traits, particularly M/F, O/DM and O/B, were significantly intercorrelated (Fig. 1).

Table 2 Statistical comparison of bunch performance traits between top- and bottom-performing oil palm progenies

Trait	Bunch performance class		Δ (Top − Bottom)	% difference	*P*-value
	Top	Bottom			
F/B	65.687	64.546	1.141	1.768	0.024
M/F	**82.140**	**77.176**	**4.964**	**6.432**	**<0.001**
K/F	6.405	7.666	−1.261	−16.449	<0.001
K/B	4.204	4.959	−0.755	−15.225	<0.001
S/F	11.480	15.175	−3.695	−24.349	<0.001
O/DM	**75.834**	**72.608**	**3.226**	**4.443**	**<0.001**
O/B	**27.786**	**22.71**	**5.076**	**22.351**	**<0.001**

F/B, fruit-to-bunch ratio; M/F, mesocarp-to-fruit ratio; K/F, kernel-to-fruit ratio; K/B, kernel-to-bunch ratio; S/F, shell-to-fruit ratio; O/DM, oil-to-dry mesocarp ratio; O/B, oil-to-bunch ratio. Bold rows are the primary oil-determining traits, which were weighted three-fold in the PROMETHEE analysis. Mean value of bunch performance traits in percentage and based on model estimation. $P < 0.05$ was considered significant.

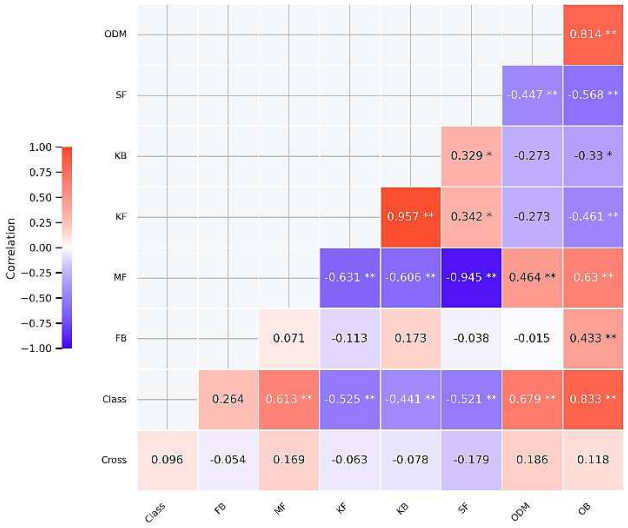
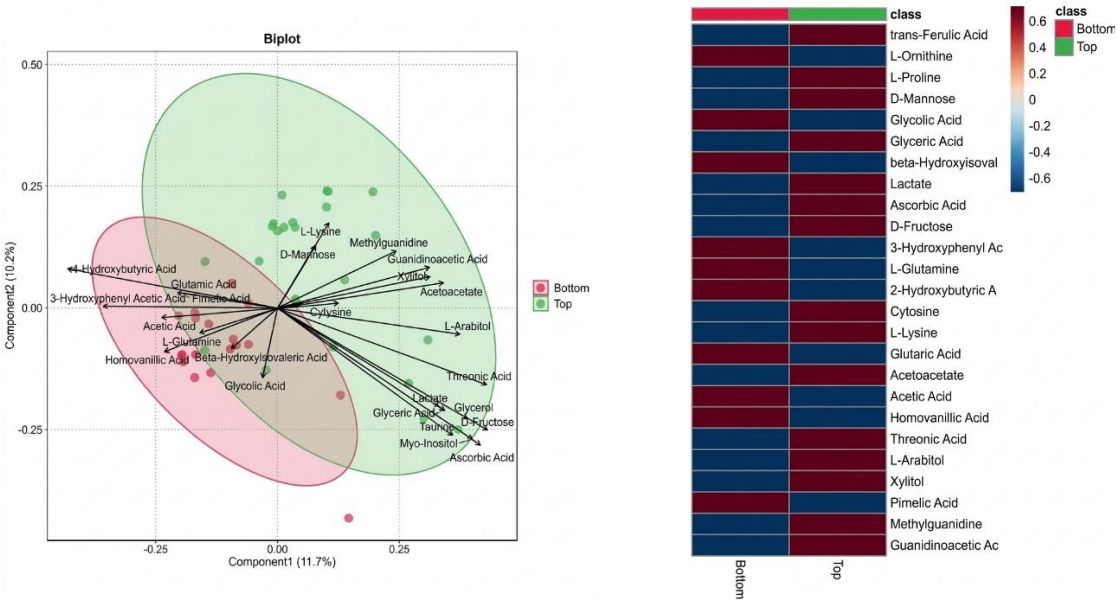


Fig. 1 Pearson correlation matrix of bunch performance traits across the evaluated oil palm progenies. The heatmap shows correlation coefficients (*r*) between traits; positive correlations are shown in red and negative correlations in blue. Asterisks denote significance (* $P < 0.05$; ** $P < 0.01$). Trait abbreviations are defined in Table 2.

Metabolomic profiling overview

A total of 123 metabolites were identified in the spear leaf samples. After filtering each class to include only metabolites with zero values in no more than 70% of the total samples, we identified 44 metabolites in the top-performing class and 32 in the bottom-performing class (Additional file 1). PLS-DA showed separation between the top and bottom classes, although with minor overlap (Fig. 2a).



(a) (b)

Fig. 2 Multivariate analysis of spear leaf metabolic profiles of top- and bottom-performing oil palm progenies. (a) PLS-DA score plot showing class separation. (b) Hierarchical clustering heatmap of differentially accumulated metabolites.

The heatmap (Fig. 2b) showed that the top-performing class was characterized by higher levels of simple sugars (D-fructose and D-mannose), sugar alcohols (xylitol and L-arabitol), amino acids (L-proline and L-lysine) and antioxidants such as ascorbic acid and threonic acid. Conversely, the bottom-performing class showed relatively higher levels of organic acids (glycolic, glutaric and acetic acid), homovanillic acid and the amino acid L-ornithine. The metabolites contributing most to class separation (VIP > 1) were guanidinoacetic acid, acetoacetate, xylitol, 2-hydroxybutyrate, 3-hydroxyphenylacetic acid, methylguanidine, arabitol, homovanillic acid, threonic acid and lysine (Fig. 3).

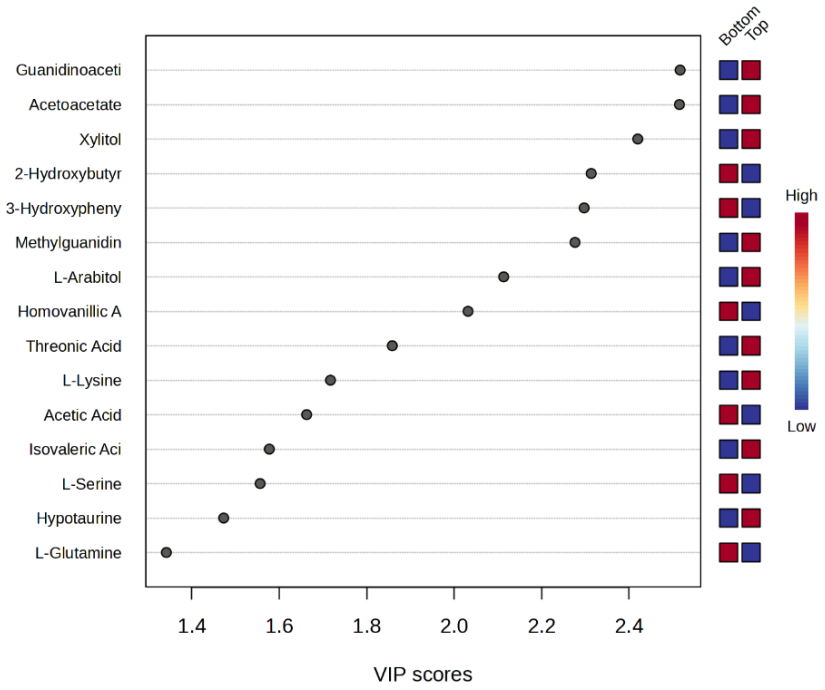


Fig. 3 Top-ranked discriminative metabolites contributing to class separation, based on variable importance in projection (VIP) scores (> 1.0).

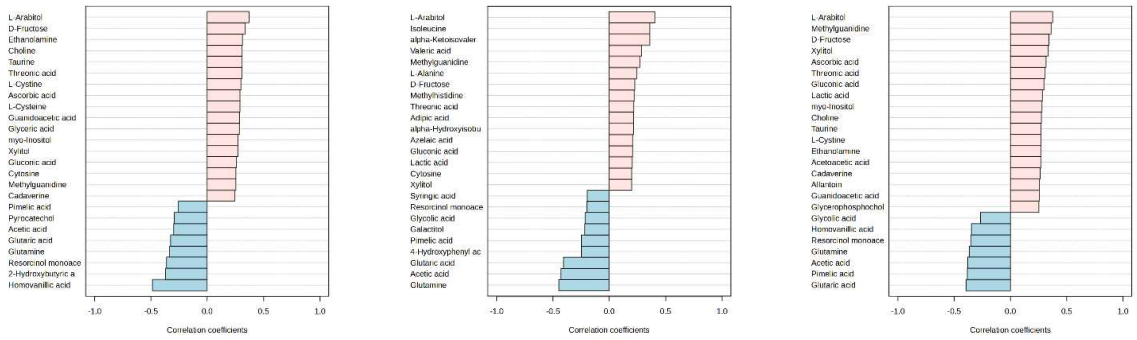
Metabolite–trait associations

Homovanillic acid showed the strongest negative correlation ($r = -0.488$, $P = 0.0014$), whereas L-arabitol ($r = 0.371$, $P = 0.018$) and D-fructose ($r = 0.338$, $P = 0.033$) were positively correlated with M/F. Glutamine was negatively correlated with M/F ($r = -0.333$, $P = 0.035$).

For O/DM, glutamine showed the strongest negative correlation ($r = -0.446$, $P = 0.0039$), followed by acetic acid ($r = -0.427$, $P = 0.0060$) and glutaric acid ($r = -0.406$, $P = 0.0093$). By contrast, L-arabitol was positively correlated with O/DM ($r = 0.404$, $P = 0.0097$), as were isoleucine ($r = 0.360$, $P = 0.0226$) and α-ketoisovaleric acid ($r = 0.359$, $P = 0.0227$). Metabolites associated with amino acid metabolism, including glutamine, isoleucine and α-ketoisovaleric acid, were among the strongest correlates of O/DM.

For O/B, glutaric acid ($r = -0.393$, $P = 0.0122$), pimelic acid ($r = -0.381$, $P = 0.0153$), acetic acid ($r = -0.379$, $P = 0.0157$) and glutamine ($r = -0.364$, $P = 0.0208$) were negatively correlated, whereas L-arabitol ($r = 0.375$, $P = 0.0172$), methylguanidine ($r = 0.361$, $P = 0.0220$), D-fructose ($r = 0.342$, $P = 0.0306$) and xylitol ($r = 0.336$, $P = 0.0338$) were positively correlated. Ascorbic acid was also positively correlated with O/B ($r = 0.317$, $P = 0.0465$), and threonic acid showed a similar trend that did not reach nominal significance ($r =$

0.309, $P = 0.0525$) (Fig. 4). None of these correlations remained significant after FDR adjustment (FDR > 0.05).



(a) (b) (c)

Fig. 4 Associations between metabolite abundances and oil-related bunch traits: (a) M/F, (b) O/DM and (c) O/B. Pearson correlations are unadjusted for multiple testing; none remained significant after FDR adjustment.

Across the three oil-related traits, several metabolites showed consistent patterns. L-arabitol was positively associated with M/F, O/DM and O/B, whereas glutamine was negatively associated with all three. D-fructose was positively associated with M/F and O/B, whereas isoleucine and α-ketoisovaleric acid, which are related to branched-chain amino acid metabolism, were associated specifically with O/DM (Fig. 4). These recurring patterns suggest coordinated changes in carbon and nitrogen metabolism accompanying variation in mesocarp development and oil accumulation among the progenies.

Metabolite set enrichment analysis (MSEA)

MSEA showed that both performance classes were significantly enriched in the primary metabolite classes carboxylic acids and derivatives, fatty acyls, hydroxy acids and derivatives, and keto acids and derivatives (FDR < 0.05; Table 3), with carboxylic acids and derivatives the most strongly enriched in both. Organooxygen compounds were enriched in the top-performing class only (nine hits, FDR = 2.20×10^{-6}) and not in the bottom-performing class (three hits, FDR = 0.989).

Table 3 Metabolite set enrichment analysis (MSEA) of spear leaf metabolic profiles in top- and bottom-performing oil palm progenies

Metabolite class	Matched metabolites (hits)		FDR	
	Top	Bottom	Top	Bottom
Carboxylic acids and derivatives	12	11	6.74×10^{-9}	2.77×10^{-9}
Fatty acyls	11	10	6.91×10^{-7}	3.26×10^{-7}
Hydroxy acids and derivatives	4	3	1.95×10^{-6}	1.32×10^{-4}
Organooxygen compounds	9	3	2.20×10^{-6}	0.989
Keto acids and derivatives	3	2	1.99×10^{-4}	0.017
Phenol esters	1	1	0.638	0.803
Dihydrofurans	1	0	1	–
Azoles	1	0	1	–
Organonitrogen compounds	1	0	1	–
Benzene and substituted derivatives	1	0	1	1
Phenols	0	1	–	1

FDR, false discovery rate. A dash indicates that no value was obtained.

Pathway enrichment analysis

Pathway enrichment analysis mapped the identified metabolites to six significantly enriched metabolic pathways (FDR < 0.01; Table 4). Alanine, aspartate and glutamate metabolism was the most significantly enriched (FDR = 8.78×10^{-7}) and had the highest pathway impact (0.849). Other amino acid-related pathways were also enriched: arginine biosynthesis (FDR = 1.92×10^{-5}; impact = 0.458), glycine, serine and threonine metabolism (FDR = 1.38×10^{-4}; impact = 0.324) and arginine and proline metabolism (FDR = 5.10×10^{-3}; impact = 0.507). Central carbon metabolism was represented by glyoxylate and dicarboxylate metabolism (FDR = 6.22×10^{-5}; impact = 0.282) and the citrate cycle (TCA cycle; FDR = 2.40×10^{-3}; impact = 0.291).

Table 4 Enriched metabolic pathways associated with bunch performance variation in the studied oil palm progenies

No.	Metabolic pathway	FDR	Pathway impact
1	Alanine, aspartate and glutamate metabolism	8.78×10^{-7}	0.849
2	Arginine biosynthesis	1.92×10^{-5}	0.458
3	Glyoxylate and dicarboxylate metabolism	6.22×10^{-5}	0.282
4	Glycine, serine and threonine metabolism	1.38×10^{-4}	0.324
5	Arginine and proline metabolism	5.10×10^{-3}	0.507
6	Citrate cycle (TCA cycle)	2.40×10^{-3}	0.291

FDR, false discovery rate.

Candidate metabolite-ratio biomarkers

ROC analysis was used to assess how well metabolite ratios distinguish the two performance classes. The top-ranked ratios had AUC values of 0.790–0.814 (Table 5; Fig. 5). β-Hydroxyisovaleric acid/methylguanidine had the highest AUC (0.814), followed by guanidinoacetic acid/pimelic acid (0.806). The remaining top-ranked ratios were acetic acid/guanidinoacetic acid, glycolic acid/threonic acid and glutaric acid/guanidinoacetic acid, with AUC values of 0.790 or higher (Table 5). Several of the metabolites making up these ratios were also highlighted by earlier analyses: guanidinoacetic acid and methylguanidine were among the VIP-ranked metabolites, and acetic, glutaric and pimelic acids were among the metabolite–trait correlates.

Table 5 Top-ranked metabolite-ratio candidate biomarkers for discriminating bunch performance classes in oil palm progenies

No.	Metabolite ratio (candidate biomarker)	AUC	*P*-value
1	β-Hydroxyisovaleric acid/methylguanidine	0.814	1.03×10^{-4}
2	Guanidinoacetic acid/pimelic acid	0.806	1.87×10^{-4}
3	Acetic acid/guanidinoacetic acid	0.793	3.68×10^{-4}
4	Glycolic acid/threonic acid	0.793	7.17×10^{-4}
5	Glutaric acid/guanidinoacetic acid	0.790	4.85×10^{-4}

AUC, area under the receiver operating characteristic curve.

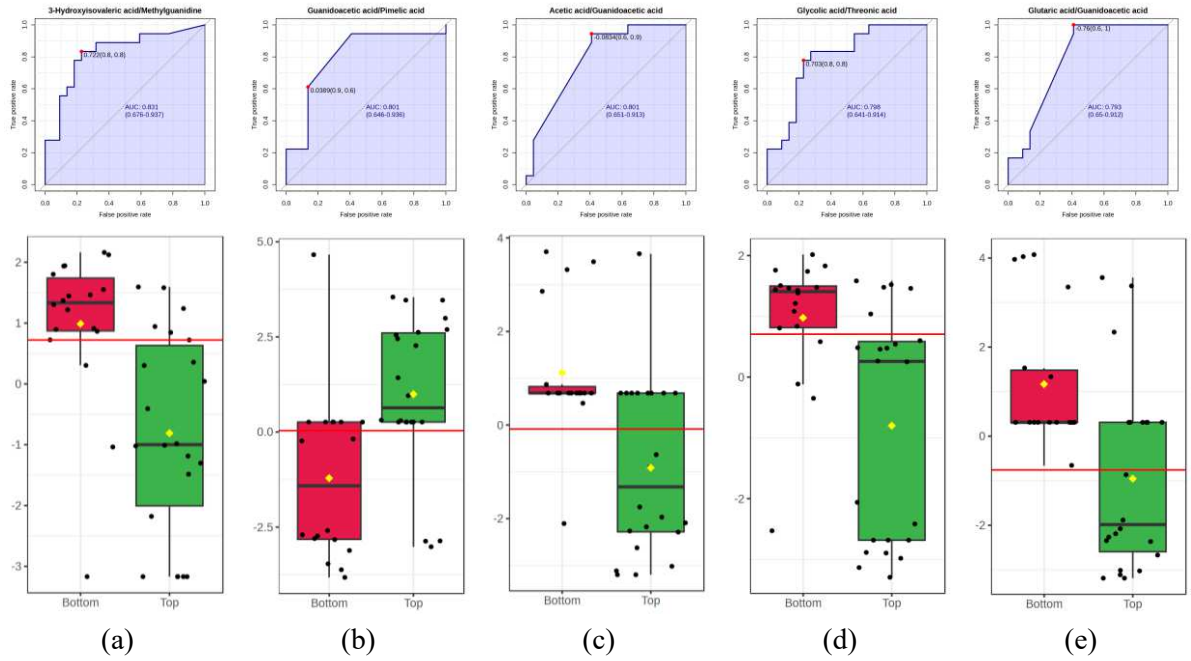


Fig. 5 Distribution and ROC performance of candidate metabolite-ratio biomarkers distinguishing top- and bottom-performing oil palm progenies: (a) β-hydroxyisovaleric acid/methylguanidine; (b) guanidinoacetic acid/pimelic acid; (c) acetic acid/guanidinoacetic acid; (d) glycolic acid/threonic acid; (e) glutaric acid/guanidinoacetic acid.

Discussion

This study compared the spear leaf metabolomes of top- and bottom-performing palms selected within 18 biparental crosses. The two classes differed in oil-related bunch traits (Table 2) and in spear leaf metabolite profiles involving carbon-conserving and nitrogen-related pathways (Tables 3 and 4), and several metabolite ratios discriminated between them (Table 5). We discuss these findings in terms of metabolic divergence among related progenies, carbon channeling, nitrogen buffering and carbon–nitrogen (C–N) coordination, and we consider the limitations of the study design.

Metabolic divergence among progenies from the same crosses

Although the top and bottom palms originated from the same biparental crosses and were grown in the same progeny trial, they differed markedly in mesocarp proportion (M/F) and oil-to-bunch ratio (O/B) (Table 2). Progeny of a biparental cross is not genetically identical, because alleles segregate within families, so part of this gap is expected to be genetic. However, the within-cross design reduces genetic and environmental background variation relative to comparisons among unrelated material, and the breadth of the metabolic differences observed suggests that regulatory and metabolic processes also differ between the classes. Genomic data establish the biological potential of a palm, but they do not capture the epigenetic, post-transcriptional or enzymatic flux differences that actively drive tissue development [21]. The metabolome is therefore a useful functional readout [22–24], showing how related genotypes can be translated into different agronomic outcomes.

This divergence is evident in the separation of the top and bottom classes in the PLS-DA and heatmap models (Fig. 2a, b). If yield variation were attributable to a single, isolated genetic change, only localized changes in a few downstream metabolites would be expected. Instead, VIP analysis (Fig. 3) identified a broad range of discriminatory metabolites spanning several biochemical classes, which suggests that contrasting bunch performance is associated with coordinated metabolic differences rather than with a single biochemical pathway [11, 25, 26].

The observed metabolomic divergence suggests that top progenies may maintain a physiological state that is more favorable to oil accumulation, through more efficient metabolic flux, energy partitioning and stress responses, whereas bottom progenies may be subject to regulatory constraints that limit large-scale lipid accumulation [27, 28].

Carbon–nitrogen networks and mesocarp sink strength

PLS-DA and VIP analyses identified distinct metabolic signatures in the top and bottom progenies, and metabolite–trait correlations related these leaf-level differences to variation in the key oil-related traits (M/F, O/DM and O/B), although these correlations did not remain significant after FDR adjustment. This relationship is physiologically plausible in the context of source–sink dynamics: the unopened spear leaf, as a metabolically active source tissue, reflects the physiological and genetic status of the palm and may therefore influence resource allocation and the lipid biosynthetic capacity of the fruit sink.

Carbon channeling and carbon conservation

Although the central role of carbon metabolism in lipid biosynthesis is well established, the present data suggest that a key distinction between top and bottom progenies may be the efficiency of metabolic channeling rather than biosynthetic capacity alone. Given the shared parentage within each cross, the lower M/F and O/DM of bottom progenies may result from a regulatory bottleneck. Top progenies may possess greater sink strength in the developing mesocarp, which was not profiled here [29, 30], converting imported sucrose rapidly and continuously into lipid precursors [31]. This could prevent biochemical bottlenecks and sustain a strong concentration gradient that facilitates the movement of resources from source tissues into the fruit [32–34].

The enrichment of the TCA cycle and of glyoxylate and dicarboxylate metabolism (Table 4) is consistent with enhanced sink utilization. Enrichment of glyoxylate and dicarboxylate metabolism may reflect a strategy of carbon conservation: unlike the standard TCA cycle, in which carbon is lost as CO_2, the glyoxylate shunt and dicarboxylate pathways bypass decarboxylation steps and conserve carbon skeletons [35–39]. Top progenies might therefore reduce respiratory carbon loss and optimize the conversion of sugars into the acetyl-CoA pool required for *de novo* fatty acid synthesis.

In contrast, the lower O/DM and M/F of bottom progenies may reflect metabolic inefficiency. Without effective channeling mechanisms, imported carbon may encounter biochemical bottlenecks, and poor coordination of central carbon flux may impede sustained lipid accumulation in the mesocarp [27]. Carbon could be lost through increased respiratory activity [40, 41] or redirected into alternative structural sinks such as cell wall biosynthesis [42, 43], which is associated with reduced starch accumulation [44]. These findings suggest

that enhanced bunch productivity may depend not only on the up-regulation of lipid synthesis genes but also on a tightly regulated carbon economy that minimizes metabolic waste [45].

Nitrogen buffering and sink longevity

Although carbon forms the basis of oil formation, active lipid accumulation places a heavy demand on the enzymatic systems of the cell. Continuous fatty acid synthesis requires rapid turnover of metabolically expensive protein complexes such as acetyl-CoA carboxylase (ACCase) and fatty acid synthase (FAS) [46]. The duration and strength of oil deposition, known as sink longevity, are therefore closely controlled by the ability of the plant to maintain nitrogen homeostasis [47, 48]. The negative association of glutamine with oil-related traits, together with the enrichment of alanine, aspartate and glutamate metabolism and of glycine, serine and threonine metabolism (Table 4), suggests a flexible nitrogen network in high-performing palms [49–51]. Rather than consuming imported nitrogen, top progenies may convert and reassign amino acids to meet the demands of the enzymatic systems required for ongoing lipid synthesis.

Guanidino-related compounds, namely guanidinoacetic acid and methylguanidine, were repeatedly identified as candidate biomarkers. Their physiological roles are well described in animal energy metabolism, but their functions in plant storage tissues remain poorly characterized. Because they are derivatives of arginine, as shown mainly in animal studies [52–55], their accumulation may provide mechanistic information. It has been suggested that guanidino compounds may act as a concentrated, dynamic nitrogen store that temporarily holds excess nitrogen while keeping an accessible supply for rapid enzyme regeneration [55, 56]. Furthermore, a large influx of carbon into lipid biosynthesis can cause considerable cellular stress, with accumulation of reactive oxygen species (ROS) [57–61]. In several plant systems, amino acid metabolism, particularly of arginine and its downstream polyamine and guanidino derivatives, acts as a metabolic buffer that helps reduce abiotic stress [62–64]. The presence of these guanidino biomarkers in top progenies suggests that these palms may be better able to buffer both cellular stress and nitrogen limitation, although this remains to be tested.

Carbon–nitrogen coordination

Lipid synthesis serves not only as a mechanism for carbon storage but also as a significant physiological stress test that requires precise coordination between energy production and cellular maintenance [65, 66]. The carbon channeling and conservation described above

depend on complementary processes, and the nitrogen reservoirs and buffering systems described above may act as counterbalances to the substantial carbon demand. In integrated C–N networks, efficient carbon metabolism generates the ATP and reducing power required for nitrogen assimilation [65], while an adequate nitrogen supply maintains the enzymatic activity needed for continued carbon flow into oil synthesis.

The integration of the MSEA and pathway analyses is consistent with this coordination. Concurrent enrichment of the TCA cycle, glyoxylate and dicarboxylate metabolism and interconnected amino acid pathways (glutamate, aspartate and arginine) points to the nodes at which carbon and nitrogen metabolism communicate. In oil palm mesocarp, higher oil yield correlates with increased expression of genes involved in glycolysis, the TCA cycle and acetyl-CoA production, indicating that central carbon metabolism actively supports lipid synthesis rather than serving as a passive background process [45]. In addition, coexpression analysis links fatty acid synthesis with starch turnover and carbon recapture pathways, showing that oil accumulation is coordinated with carbon conservation or redirection mechanisms [67]. These pathways also generate the α-keto acid backbones, such as α-ketoglutarate and oxaloacetate, that are required for the assimilation of nitrogen into glutamine and aspartate, thereby linking carbon availability directly with nitrogen metabolism [68, 69]. This integration could enable top progenies to direct excess nitrogen efficiently into storage as a buffer, whereas bottom progenies might show decoupled C–N metabolism, leading to carbon inefficiency, nitrogen depletion and an early developmental bottleneck that impedes oil accumulation.

Taken together, these observations support a working model of coordinated C–N metabolism (Fig. 6) in which carbon channeling and nitrogen buffering operate as a linked system rather than as isolated pathways. Central carbon flux provides the acetyl-CoA precursors for lipid biosynthesis, whereas nitrogen-buffering networks sustain the enzymatic activity and structural integrity required for continuous mesocarp development. Sufficient carbon availability would support enzyme maintenance, and efficient nitrogen recycling would prevent metabolic bottlenecks, contributing to the high O/B observed in top progenies. This model is a hypothesis generated from correlative data and requires direct testing.

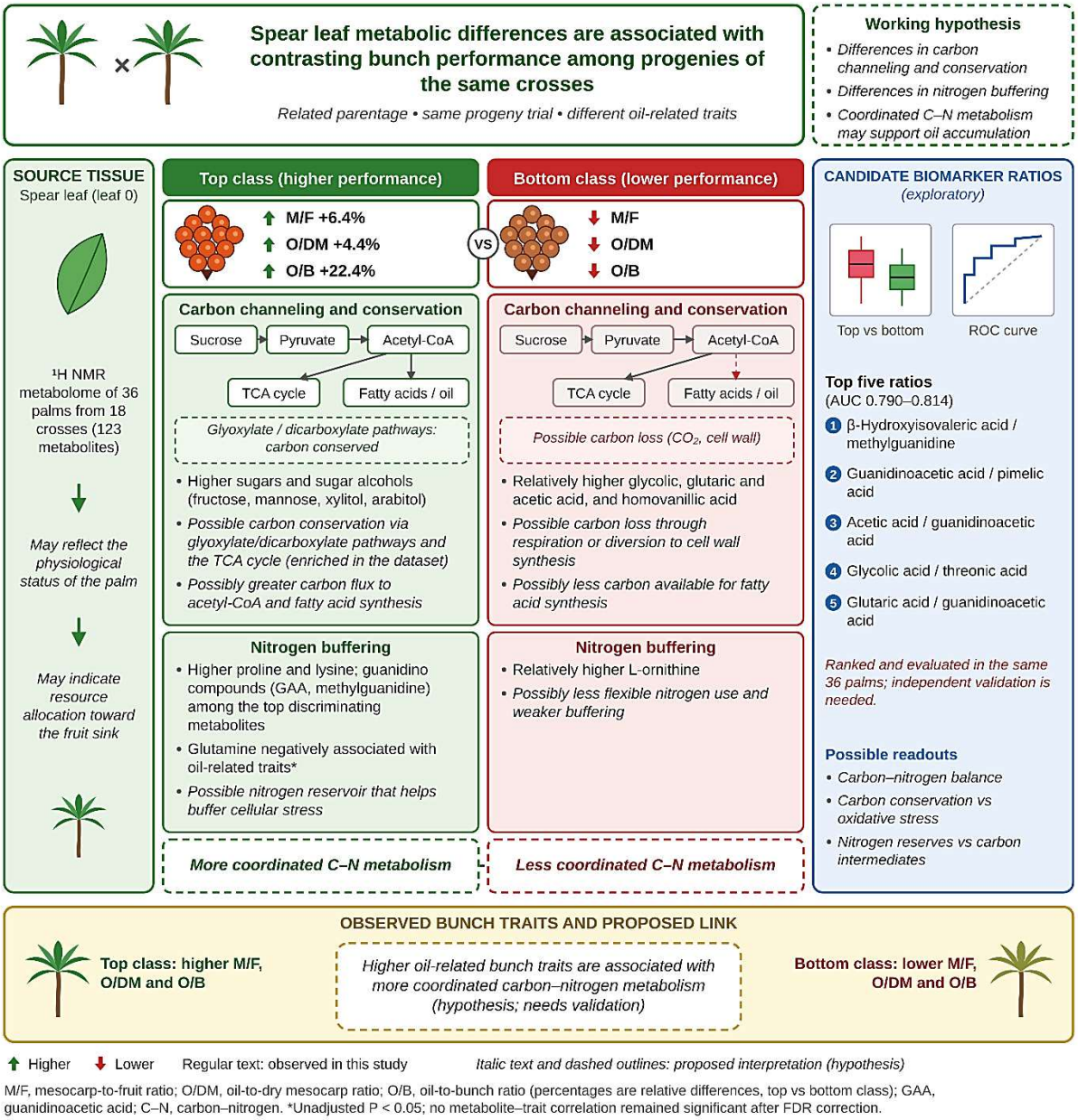


Fig. 6 Conceptual framework for metabolic divergence and C–N coordination associated with bunch performance in oil palm progenies from shared parentage.

Metabolite-ratio biomarker candidates

Plant metabolomics is increasingly used to identify biomarkers that link metabolic states to crop traits, stress responses, cultivar identity and bioactivity [70–72]. Biomarker pairs selected according to their metabolic pathways can reveal the balance, interactions and shifts between connected pathways, while also reducing sample-to-sample analytical variability [11, 73–75]. In this study, ROC analysis identified five candidate metabolite ratios with AUC values of 0.790–0.814 ($P < 0.001$), which is consistent with a link between coordinated carbon–nitrogen metabolism and oil palm productivity.

The composition of the top-ranked pairs fits the proposed model. Acetic, glutaric and pimelic acids are paired with guanidinoacetic acid, an arginine-derived guanidino compound implicated in nitrogen recycling. The highest-ranking pair, β-hydroxyisovaleric acid/methylguanidine (AUC = 0.814), links an intermediate of leucine catabolism, which feeds into acetyl-CoA and the TCA cycle, with a nitrogen compound derived from arginine. The prominence of such cross-pathway pairs suggests that balanced carbon partitioning and nitrogen utilization, rather than lipid production alone, may be important for bunch performance, with lower-performing palms showing a metabolic mismatch.

The glycolic acid/threonic acid ratio points to a possible protective dimension of the proposed mechanism. Glycolic acid is an intermediate of the glyoxylate and photorespiratory pathways, which are relevant to carbon conservation. In contrast, threonic acid is a recognized breakdown product of ascorbic acid [76–78] and may indicate oxidative stress and antioxidant depletion under conditions of high metabolic demand [78]. The discriminatory value of this ratio suggests that the capacity to balance carbon recovery with protection against oxidative damage may be important for mesocarp longevity and sink capacity.

These metabolite ratios may therefore provide insight into cellular efficiency as well as serving as statistical indicators. Monitoring the balance between carbon utilization and nitrogen protection could be a useful strategy for identifying high-performing progenies, provided that the ratios are validated as discussed below.

Limitations and outlook

Several limitations should be considered. First, palms were selected as the top and bottom performers within each cross, an extreme-group design that increases the contrast between the classes and can inflate discrimination metrics such as AUC relative to what would be expected in unselected populations. Second, the study comprised 36 palms from a single progeny trial sampled at a single time point, and the metabolite–trait correlations did not remain significant after FDR adjustment. The metabolite ratios were ranked and evaluated in the same dataset, and no independent validation set was used. Third, only spear leaves were profiled, so statements about mesocarp sink strength are inferences. The study therefore identifies associations rather than demonstrating prediction or causation. Future work should test the candidate ratios in independent, unselected progenies across environments and generations and combine leaf profiling with mesocarp measurements.

Conclusions

Within biparental crosses, top- and bottom-performing oil palm progenies differed in spear leaf metabolite profiles as well as in oil-related bunch traits. The differences involved carbon-conserving pathways (glyoxylate and dicarboxylate metabolism, TCA cycle) and nitrogen-related metabolism (alanine, aspartate and glutamate metabolism; arginine-derived guanidino compounds), which is consistent with a model in which coordinated carbon–nitrogen metabolism supports oil accumulation.

Metabolite ratios such as β-hydroxyisovaleric acid/methylguanidine and guanidinoacetic acid/pimelic acid discriminated the two classes (AUC 0.790–0.814) and are candidate markers for non-destructive early selection using spear leaves rather than fruit. Because of the extreme-group design, the small sample and the absence of FDR-significant individual correlations, these markers need validation in independent, unselected populations, across environments and across generations before they can be applied in breeding programs.

Methods

Plant material and sampling

Plant material was collected from an advanced-cycle oil palm progeny trial at the RB03S Experimental Station, Rambutan Estate, PT Perkebunan Nusantara III (PTPN III), North Sumatra, Indonesia. To minimize developmental bias, unexpanded spear leaves (leaf 0) were sampled from 36 palms age 12 year after planting, representing 18 crosses (Table 1). In each cross, the palms with the highest (top class) and lowest (bottom class) bunch performance were selected based on historical progeny evaluation data. Immediately after collection, the spear leaves were quenched in liquid nitrogen and stored at −80 °C until analysis.

Table 1 Oil palm progenies selected for metabolic profiling from the RB03S progeny trial

Cross	Top-performer (field ID)	Bottom-performer (field ID)
Cross 01	30-21	44-16
Cross 02	11-8	32-12
Cross 03	33-23	49-1
Cross 04	30-9	39-16
Cross 05	10-9	37-23
Cross 06	53-16	40-22
Cross 07	26-4	26-2
Cross 08	38-18	43-10
Cross 09	18-4	50-14
Cross 10	13-1	49-6
Cross 11	28-7	50-22
Cross 12	31-18	31-20
Cross 13	21-24	47-20
Cross 14	8-12	9-12
Cross 15	40-25	15-20
Cross 16	27-21	34-14
Cross 17	13-28	63-12
Cross 18	60-6	3-4

Performance classes were determined from the bunch performance traits F/B, M/F, K/F, K/B, S/F, O/DM and O/B using PROMETHEE multi-criteria decision analysis. In total, 36 palms (18 per class) were sampled.

Metabolite extraction and ^{1}H NMR profiling

Metabolites were extracted using a modified biphasic solvent system of methanol:chloroform:water (2:0.8:1, v/v/v) [79]. For each sample, frozen tissue was ground to a fine powder in liquid nitrogen, and the extraction solvent was added at 10 times the sample mass. The mixture was macerated overnight on an orbital shaker at 150 rpm and then filtered. The filtrate was concentrated using a rotary vacuum evaporator to obtain the crude extract.

For NMR analysis, 50 mg of crude extract was dissolved in 1 mL of deuterated chloroform ($CDCl_3$; Merck-Sigma, USA) containing 0.03% (v/v) tetramethylsilane (TMS) as an internal standard. The mixture was vortexed for 1 min, sonicated for 5 min and centrifuged at 5,000 rpm for 5 min, and the supernatant was transferred into a standard NMR tube. Proton (^{1}H) NMR spectra were acquired on a JEOL JNM-ECZR spectrometer (Tokyo, Japan) operating at a proton frequency of 500 MHz. Acquisition parameters included 128 scans, with water suppression applied to the residual $CDCl_3$ signal at 7.24 ppm.

NMR data processing and metabolite identification

Raw ^{1}H NMR spectra were processed using MestReNova software (version 14.3, Mestrelab Research), including automatic Fourier transformation, phase correction and baseline adjustment [80, 81]. Chemical shifts were calibrated against the TMS signal at 0.0 ppm. Metabolites were identified and quantified from the deconvoluted spectra using the ASICS (Automatic Statistical Identification in Complex Spectra) package in R [82], which performs automated spectral deconvolution based on reference libraries.

Phenotypic evaluation and classification

Individual phenotypic rankings were established from bunch analysis data collected during the progeny trials. The evaluated traits were fruit-to-bunch (F/B), mesocarp-to-fruit (M/F), kernel-to-fruit (K/F), kernel-to-bunch (K/B), shell-to-fruit (S/F), oil-to-dry mesocarp (O/DM) and oil-to-bunch (O/B) ratios. To stratify each palm into the top or bottom performance class within its cross, multi-criteria decision analysis (MCDA) was applied using the PROMETHEE method [83]. The primary oil-determining traits (M/F, O/DM and O/B) were assigned a weight three times greater than the other traits.

Differences in bunch performance traits between the top and bottom classes were assessed with linear mixed-effects models fitted by restricted maximum likelihood (REML). Performance class (top vs. bottom) was included as a fixed effect and cross as a random effect to account for family-level variance. Fixed effects were tested using the Kenward–Roger approximation for degrees of freedom. A correlation heatmap was constructed to visualize trait-to-trait associations. Phenotypic tests were considered significant at $P < 0.05$.

Multivariate metabolomics and biomarker analysis

The quantified metabolite data matrix was sum-normalized and Pareto-scaled before statistical modeling in MetaboAnalyst 6.0 [73]. PLS-DA was used to classify samples and visualize the global metabolic divergence between top and bottom progenies. VIP scores and hierarchically clustered heatmaps were generated to identify the metabolites that most strongly discriminated the groups.

To link metabolomic profiles with phenotypic outcomes, metabolite–trait associations were evaluated using Pearson's correlation analysis (two-tailed, $P < 0.05$). MSEA and pathway analysis were performed to identify biochemical networks associated with bunch performance, using the Kyoto Encyclopedia of Genes and Genomes (KEGG) reference database. Significant pathways were selected based on an impact score > 0.1 and FDR < 0.05.

Finally, ROC curve analysis, together with VIP scores, was used to evaluate the robustness of candidate biomarkers for high oil yield traits.

Abbreviations

ACCase, acetyl-CoA carboxylase
ASICS, Automatic Statistical Identification in Complex Spectra
AUC, area under the curve
C–N, carbon–nitrogen
FAS, fatty acid synthase
FDR, false discovery rate
F/B, fruit-to-bunch ratio
K/B, kernel-to-bunch ratio
K/F, kernel-to-fruit ratio
KEGG, Kyoto Encyclopedia of Genes and Genomes
M/F, mesocarp-to-fruit ratio
MCDA, multi-criteria decision analysis
MSEA, metabolite set enrichment analysis
NMR, nuclear magnetic resonance
O/B, oil-to-bunch ratio
O/DM, oil-to-dry mesocarp ratio
PLS-DA, partial least squares discriminant analysis
PROMETHEE, Preference Ranking Organization Method for Enrichment Evaluation
REML, restricted maximum likelihood
ROC, receiver operating characteristic
ROS, reactive oxygen species
RRS, reciprocal recurrent selection
S/F, shell-to-fruit ratio
TCA, tricarboxylic acid
TMS, tetramethylsilane
VIP, variable importance in projection

Supplementary information

Additional file 1. Supplementary Table 1_Metabolite profile of progeny RB03S_Complete.xlxs

Acknowledgements

The author gratefully acknowledges PT. Perkebunan Nusantara III, Kebun Rambutan for providing field trail in progeny testing programs.

Authors' contributions

ANS performed data curation, formal analysis, methodology, investigation, visualization, writing original draft and provide resources. YRN contributed to conceptualization, methodology, validation, writing review and also as project administrator. HYR contributed to methodology, writing review and editing, funding acquisition and supervision. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by Indonesian Oil Palm Research Institute as part of regular internal research program in Plant Breeding Research Group.

Availability of data and materials

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Corley RHV, Tinker PB. The Oil Palm: Fifth Edition. Chichester, UK: John Wiley & Sons, Ltd; 2015. https://doi.org/10.1002/9781118953297.
2. Murphy DJ, Goggin K, Paterson RRM. Oil palm in the 2020s and beyond: challenges and solutions. CABI Agric Biosci. 2021;2:39. https://doi.org/10.1186/s43170-021-00058-3.

3. Barcelos E, Rios SDA, Cunha RNV, Lopes R, Motoike SY, Babiychuk E, et al. Oil palm natural diversity and the potential for yield improvement. Front Plant Sci. 2015;6. https://doi.org/10.3389/fpls.2015.00190.
4. Tupaz-Vera A, Ayala-Diaz I, Barrera CF, Romero HM. Genetic gains for obtaining improved progenies of oil palm in Colombia. Euphytica. 2023;219:38. https://doi.org/10.1007/s10681-022-03137-2.
5. Abu Bakar D, Tahir NI, Ramli U, Ong-Abdullah M, Hashim AT. Metabolome characterization of three oil palm tissue culture- derived morphotypes. Kuala Lumpur: MPOB-ISBAB; 2019.
6. Marhalil M, Rafii MY, Afizi MMA, Arolu IW, Noh A, Din AM, et al. Genetic variability in yield and vegetative traits in elite germplasm of MPOB-Nigerian dura × AVROS pisifera progenies. J Food Agric Environ. 2013;11:515–9.
7. Fernie AR, Schauer N. Metabolomics-assisted breeding: a viable option for crop improvement? Trends Genet. 2009;25:39–48. https://doi.org/10.1016/j.tig.2008.10.010.
8. Katam R, Lin C, Grant K, Katam CS, Chen S. Advances in Plant Metabolomics and Its Applications in Stress and Single-Cell Biology. Int J Mol Sci. 2022;23:6985. https://doi.org/10.3390/ijms23136985.
9. Sharma K, Sarma S, Bohra A, Mitra A, Sharma NK, Kumar A. Plant Metabolomics: An Emerging Technology for Crop Improvement. In: Çelik Ö, editor. New Visions in Plant Science. InTech; 2018. https://doi.org/10.5772/intechopen.76759.
10. Sharma V, Gupta P, Priscilla K, SharanKumar S, Hangargi B, Veershetty A, et al. Metabolomics Intervention Towards Better Understanding of Plant Traits. Cells. 2021;10:346. https://doi.org/10.3390/cells10020346.
11. Shen S, Zhan C, Yang C, Fernie AR, Luo J. Metabolomics-centered mining of plant metabolic diversity and function: Past decade and future perspectives. Mol Plant. 2023;16:43–63. https://doi.org/10.1016/j.molp.2022.09.007.
12. Bourgis F, Kilaru A, Cao X, Ngando-Ebongue G-F, Drira N, Ohlrogge JB, et al. Comparative transcriptome and metabolite analysis of oil palm and date palm mesocarp that differ dramatically in carbon partitioning. Proc Natl Acad Sci. 2011;108:12527–32. https://doi.org/10.1073/pnas.1106502108.
13. Neoh BK, Teh HF, Ng TLM, Tiong SH, Thang YM, Ersad MohdA, et al. Profiling of Metabolites in Oil Palm Mesocarp at Different Stages of Oil Biosynthesis. J Agric Food Chem. 2013;61:1920–7. https://doi.org/10.1021/jf304561f.
14. Apriyanto A, Compart J, Zimmermann V, Alseekh S, Fernie AR, Fettke J. Indication that starch and sucrose are biomarkers for oil yield in oil palm (Elaeis guineensis Jacq.). Food Chem. 2022;393:133361. https://doi.org/10.1016/j.foodchem.2022.133361.
15. Tranbarger TJ, Dussert S, Joët T, Argout X, Summo M, Champion A, et al. Regulatory Mechanisms Underlying Oil Palm Fruit Mesocarp Maturation, Ripening, and Functional Specialization in Lipid and Carotenoid Metabolism. Plant Physiol. 2011;156:564–84. https://doi.org/10.1104/pp.111.175141.
16. Chen H, Zheng Y, Zhan J, He C, Wang Q. Comparative metabolic profiling of the lipid-producing green microalga Chlorella reveals that nitrogen and carbon metabolic pathways contribute to lipid metabolism. Biotechnol Biofuels. 2017;10:153. https://doi.org/10.1186/s13068-017-0839-4.

17. Teh HF, Neoh BK, Hong MPL, Low JYS, Ng TLM, Ithnin N, et al. Differential Metabolite Profiles during Fruit Development in High-Yielding Oil Palm Mesocarp. PLoS ONE. 2013;8:e61344. https://doi.org/10.1371/journal.pone.0061344.
18. John Martin JJ, Xu W, Li X, Liu X, Zhou L, Li R, et al. Deciphering the molecular mechanisms of oil palm lipid metabolism through combined metabolomics and transcriptomics. Food Chem. 2025;482:144154. https://doi.org/10.1016/j.foodchem.2025.144154.
19. Hassan H, Tahir NI, Rozali NL, Lau BYC, Othman A, Weckwerth W, et al. Integrative tissue-resolved proteomics and metabolomics analysis of oil palm (Elaeis guineensis Jacq.) fruit provides insights into stilbenoid biosynthesis at the interface of primary and secondary metabolism. Biocatal Agric Biotechnol. 2024;60:103308. https://doi.org/10.1016/j.bcab.2024.103308.
20. Wei L, Yang C, John Martin JJ, Li R, Zhou L, Cheng S, et al. Metabonomics and Transcriptomic Analysis of Free Fatty Acid Synthesis in Seedless and Tenera Oil Palm. Int J Mol Sci. 2024;25:1686. https://doi.org/10.3390/ijms25031686.
21. Arbona V, Manzi M, Ollas C, Gómez-Cadenas A. Metabolomics as a Tool to Investigate Abiotic Stress Tolerance in Plants. Int J Mol Sci. 2013;14:4885–911. https://doi.org/10.3390/ijms14034885.
22. Fiehn O. Metabolomics – the link between genotypes and phenotypes. Plant Mol Biol. 2002;48:155–71.
23. Jansson JK, Baker ES. A multi-omic future for microbiome studies. Nat Microbiol. 2016;1:16049. https://doi.org/10.1038/nmicrobiol.2016.49.
24. Sakurai N. Recent applications of metabolomics in plant breeding. Breed Sci. 2022;72:56–65. https://doi.org/10.1270/jsbbs.21065.
25. Hall RD. Plant metabolomics: from holistic hope, to hype, to hot topic. New Phytol. 2006;169:453–68. https://doi.org/10.1111/j.1469-8137.2005.01632.x.
26. Manickam S, Rajagopalan VR, Kambale R, Rajasekaran R, Kanagarajan S, Muthurajan R. Plant Metabolomics: Current Initiatives and Future Prospects. Curr Issues Mol Biol. 2023;45:8894–906. https://doi.org/10.3390/cimb45110558.
27. Zeng W, Xie B, Chen Y, Chen J, Li P, Jiang L, et al. Regulatory mechanism of carbohydrate metabolism pathways on oil biosynthesis of oil plant Symplocos paniculata. Front Plant Sci. 2025;16:1452533. https://doi.org/10.3389/fpls.2025.1452533.
28. Khan S, Korai Z, Yang L, Korai SK, Li S, Zulfiqar U, et al. Metabolic reprogramming in plant defense: linking signaling networks to metabolomics-driven insights. Plant Signal Behav. 2026;21:2672221. https://doi.org/10.1080/15592324.2026.2672221.
29. Marcelis LFM. Sink strength as a determinant of dry matter partitioning in the whole plant. J Exp Bot. 1996;47 Special Issue:1281–91.
30. Li Y-M, Forney C, Bondada B, Leng F, Xie Z-S. The Molecular Regulation of Carbon Sink Strength in Grapevine (Vitis vinifera L.). Front Plant Sci. 2021;11:606918. https://doi.org/10.3389/fpls.2020.606918.
31. Zhai Z, Keereetaweep J, Liu H, Xu C, Shanklin J. The Role of Sugar Signaling in Regulating Plant Fatty Acid Synthesis. Front Plant Sci. 2021;12:643843. https://doi.org/10.3389/fpls.2021.643843.

32. Braun DM, Wang L, Ruan Y-L. Understanding and manipulating sucrose phloem loading, unloading, metabolism, and signalling to enhance crop yield and food security. J Exp Bot. 2014;65:1713–35. https://doi.org/10.1093/jxb/ert416.
33. Knoblauch M, Knoblauch J, Mullendore DL, Savage JA, Babst BA, Beecher SD, et al. Testing the Münch hypothesis of long distance phloem transport in plants. eLife. 2016;5:e15341. https://doi.org/10.7554/eLife.15341.
34. Chen T, Zhang Z, Li B, Qin G, Tian S. Molecular basis for optimizing sugar metabolism and transport during fruit development. aBIOTECH. 2021;2:330–40. https://doi.org/10.1007/s42994-021-00061-2.
35. Falk KL, Behal RH, Xiang C, Oliver DJ. Metabolic Bypass of the Tricarboxylic Acid Cycle during Lipid Mobilization in Germinating Oilseeds1. Plant Physiol. 1998;117:473–81. https://doi.org/10.1104/pp.117.2.473.
36. Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. Nature. 2004;432:779–82. https://doi.org/10.1038/nature03145.
37. Gonçalves IL, Mielniczki-Pereira AA, Piovezan Borges AC, Valduga AT. Metabolic modeling and comparative biochemistry in glyoxylate cycle. Acta Sci Biol Sci. 2016;38:1. https://doi.org/10.4025/actascibiolsci.v38i1.24597.
38. Chew SY, Chee WJY, Than LTL. The glyoxylate cycle and alternative carbon metabolism as metabolic adaptation strategies of Candida glabrata: perspectives from Candida albicans and Saccharomyces cerevisiae. J Biomed Sci. 2019;26:52. https://doi.org/10.1186/s12929-019-0546-5.
39. Yang P, Liu W, Chen Y, Gong A-D. Engineering the glyoxylate cycle for chemical bioproduction. Front Bioeng Biotechnol. 2022;10:1066651. https://doi.org/10.3389/fbioe.2022.1066651.
40. Sharma N, Thakur M, Suryakumar P, Mukherjee P, Raza A, Prakash CS, et al. 'Breathing Out' under Heat Stress—Respiratory Control of Crop Yield under High Temperature. Agronomy. 2022;12:806. https://doi.org/10.3390/agronomy12040806.
41. Li G, Chen T, Feng B, Peng S, Tao L, Fu G. Respiration, not photosynthesis, determines yield loss caused by global warming. 2022. https://doi.org/10.22541/au.164864281.19004687/v1.
42. Wormit A, Butt SM, Chairam I, McKenna JF, Nunes-Nesi A, Kjaer L, et al. Osmosensitive Changes of Carbohydrate Metabolism in Response to Cellulose Biosynthesis Inhibition. Plant Physiol. 2012;159:105–17. https://doi.org/10.1104/pp.112.195198.
43. Verbančič J, Lunn JE, Stitt M, Persson S. Carbon Supply and the Regulation of Cell Wall Synthesis. Mol Plant. 2018;11:75–94. https://doi.org/10.1016/j.molp.2017.10.004.
44. Gutschker S, Ruescher D, Rabbi IY, Rosado-Souza L, Pommerrenig B, Pauly M, et al. Carbon usage in yellow-fleshed *Manihot esculenta* storage roots shifts from starch biosynthesis to cell wall and raffinose biosynthesis via the *myo* -inositol pathway. Plant J. 2024;119:2045–62. https://doi.org/10.1111/tpj.16909.
45. Wong YC, Teh HF, Mebus K, Ooi TEK, Kwong QB, Koo KL, et al. Differential gene expression at different stages of mesocarp development in high- and low-yielding oil palm. BMC Genomics. 2017;18:470. https://doi.org/10.1186/s12864-017-3855-7.

46. Bates PD, Stymne S, Ohlrogge J. Biochemical pathways in seed oil synthesis. Curr Opin Plant Biol. 2013;16:358–64. https://doi.org/10.1016/j.pbi.2013.02.015.
47. Ischebeck T, Krawczyk HE, Mullen RT, Dyer JM, Chapman KD. Lipid droplets in plants and algae: Distribution, formation, turnover and function. Semin Cell Dev Biol. 2020;108:82–93. https://doi.org/10.1016/j.semcdb.2020.02.014.
48. Cai Y, Horn PJ. Packaging "vegetable oils": Insights into plant lipid droplet proteins. Plant Physiol. 2025;197:kiae533. https://doi.org/10.1093/plphys/kiae533.
49. Reitzer L. Biosynthesis of Glutamate, Aspartate, Asparagine, L -Alanine, and D -Alanine. EcoSal Plus. 2004;1:10.1128/ecosalplus.3.6.1.3. https://doi.org/10.1128/ecosalplus.3.6.1.3.
50. Fortunato S, Nigro D, Lasorella C, Marcotuli I, Gadaleta A, De Pinto MC. The Role of Glutamine Synthetase (GS) and Glutamate Synthase (GOGAT) in the Improvement of Nitrogen Use Efficiency in Cereals. Biomolecules. 2023;13:1771. https://doi.org/10.3390/biom13121771.
51. Li S, Ahmed W, Jiang T, Yang D, Yang L, Hu X, et al. Amino acid metabolism pathways as key regulators of nitrogen distribution in tobacco: insights from transcriptome and WGCNA analyses. BMC Plant Biol. 2025;25:393. https://doi.org/10.1186/s12870-025-06390-4.
52. Farrow SC, Facchini PJ. Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. Front Plant Sci. 2014;5. https://doi.org/10.3389/fpls.2014.00524.
53. Ostojic SM. Tackling guanidinoacetic acid for advanced cellular bioenergetics. Nutrition. 2017;34:55–7. https://doi.org/10.1016/j.nut.2016.09.010.
54. Ostojic SM. Guanidinoacetic acid as a performance-enhancing agent. Amino Acids. 2016;48:1867–75. https://doi.org/10.1007/s00726-015-2106-y.
55. Funck D, Sinn M, Forlani G, Hartig JS. Guanidine production by plant homoarginine-6-hydroxylases. eLife. 2024;12:RP91458. https://doi.org/10.7554/eLife.91458.
56. Yi S, Hu S, Wang J, Abudukelimu A, Wang Y, Li X, et al. Effect of Guanidinoacetic Acid Supplementation on Growth Performance, Rumen Fermentation, Blood Indices, Nutrient Digestion, and Nitrogen Metabolism in Angus Steers. Animals. 2024;14:401. https://doi.org/10.3390/ani14030401.
57. Shi K, Gao Z, Shi T-Q, Song P, Ren L-J, Huang H, et al. Reactive Oxygen Species-Mediated Cellular Stress Response and Lipid Accumulation in Oleaginous Microorganisms: The State of the Art and Future Perspectives. Front Microbiol. 2017;8:793. https://doi.org/10.3389/fmicb.2017.00793.
58. Xi Y, Cai J, Li G, Huang H, Peng X, Zhu G. High CO_2 facilitates fatty acid biosynthesis and mitigates cellular oxidative stress caused by CAC2 dysfunction in *Arabidopsis*. Plant J. 2023;115:1316–30. https://doi.org/10.1111/tpj.16321.
59. Wang L, Yang T, Pan Y, Shi L, Jin Y, Huang X. The Metabolism of Reactive Oxygen Species and Their Effects on Lipid Biosynthesis of Microalgae. Int J Mol Sci. 2023;24:11041. https://doi.org/10.3390/ijms241311041.
60. Singh R, Singh S, Parihar P, Mishra RK, Tripathi DK, Singh VP, et al. Reactive Oxygen Species (ROS): Beneficial Companions of Plants' Developmental Processes. Front Plant Sci. 2016;7. https://doi.org/10.3389/fpls.2016.01299.

673 61. Alam P, Faizan M, Arif Y, Azzam MM, Hayat S, Afzal S, et al. Reactive oxygen species:
674 balancing agents in plants. Front Plant Sci. 2025;16:1713590.
675 https://doi.org/10.3389/fpls.2025.1713590.
676 62. Slocum RD. Genes, enzymes and regulation of arginine biosynthesis in plants. Plant
677 Physiol Biochem. 2005;43:729–45. https://doi.org/10.1016/j.plaphy.2005.06.007.
678 63. Shi H, Ye T, Chen F, Cheng Z, Wang Y, Yang P, et al. Manipulation of arginase
679 expression modulates abiotic stress tolerance in Arabidopsis: effect on arginine
680 metabolism and ROS accumulation. J Exp Bot. 2013;64:1367–79.
681 https://doi.org/10.1093/jxb/ers400.
682 64. Hildebrandt TM. Synthesis versus degradation: directions of amino acid metabolism
683 during Arabidopsis abiotic stress response. Plant Mol Biol. 2018;98:121–35.
684 https://doi.org/10.1007/s11103-018-0767-0.
685 65. Loei H, Lim J, Tan M, Lim TK, Lin QS, Chew FT, et al. Proteomic Analysis of the Oil
686 Palm Fruit Mesocarp Reveals Elevated Oxidative Phosphorylation Activity is Critical for
687 Increased Storage Oil Production. J Proteome Res. 2013;12:5096–109.
688 https://doi.org/10.1021/pr400606h.
689 66. Ghifari AS, Saha S, Murcha MW. The biogenesis and regulation of the plant oxidative
690 phosphorylation system. Plant Physiol. 2023;192:728–47.
691 https://doi.org/10.1093/plphys/kiad108.
692 67. Guerin C, Joët T, Serret J, Lashermes P, Vaissayre V, Agbessi MDT, et al. Gene
693 coexpression network analysis of oil biosynthesis in an interspecific backcross of oil
694 palm. Plant J. 2016;87:423–41. https://doi.org/10.1111/tpj.13208.
695 68. Huang W, Han S, Wang L, Li W. Carbon and nitrogen metabolic regulation in freshwater
696 plant Ottelia alismoides in response to carbon limitation: A metabolite perspective. Front
697 Plant Sci. 2022;13:962622. https://doi.org/10.3389/fpls.2022.962622.
698 69. Fernie AR, Bachem CWB, Helariutta Y, Neuhaus HE, Prat S, Ruan Y-L, et al.
699 Synchronization of developmental, molecular and metabolic aspects of source–sink
700 interactions. Nat Plants. 2020;6:55–66. https://doi.org/10.1038/s41477-020-0590-x.
701 70. Hamany Djande CY, Pretorius C, Tugizimana F, Piater LA, Dubery IA. Metabolomics: A
702 Tool for Cultivar Phenotyping and Investigation of Grain Crops. Agronomy.
703 2020;10:831. https://doi.org/10.3390/agronomy10060831.
704 71. Shi T, Zhu A, Jia J, Hu X, Chen J, Liu W, et al. Metabolomics analysis and metabolite-
705 agronomic trait associations using kernels of wheat (*Triticum aestivum*) recombinant
706 inbred lines. Plant J. 2020;103:279–92. https://doi.org/10.1111/tpj.14727.
707 72. Song X, Fan B-L, Hong X, Su P, Sun M. Unveiling Plant Metabolic Diversity:
708 Integrating Metabolomics with Multi-Omics Approaches for Crop Improvement. Plants.
709 2026;15:846. https://doi.org/10.3390/plants15050846.
710 73. Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, et al. MetaboAnalyst 6.0: towards a unified
711 platform for metabolomics data processing, analysis and interpretation. Nucleic Acids
712 Res. 2024;52:W398–406. https://doi.org/10.1093/nar/gkae253.
713 74. Krumsiek J, Stückler F, Suhre K, Gieger C, Spector TD, Soranzo N, et al. Network-based
714 metabolite ratios for an improved functional characterization of genome-wide association
715 study results. 2016. https://doi.org/10.1101/048512.

75. Dan Z, Chen Y, Li H, Zeng Y, Xu W, Zhao W, et al. The metabolomic landscape of rice heterosis highlights pathway biomarkers for predicting complex phenotypes. Plant Physiol. 2021;187:1011–25. https://doi.org/10.1093/plphys/kiab273.
76. Deutsch JC. Ascorbic Acid Oxidation by Hydrogen Peroxide. Anal Biochem. 1998;255:1–7. https://doi.org/10.1006/abio.1997.2293.
77. Truffault V, Fry SC, Stevens RG, Gautier H. Ascorbate degradation in tomato leads to accumulation of oxalate, threonate and oxalyl threonate. Plant J. 2017;89:996–1008. https://doi.org/10.1111/tpj.13439.
78. Dewhirst RA, Fry SC. The oxidation of dehydroascorbic acid and 2,3-diketogulonate by distinct reactive oxygen species. Biochem J. 2018;475:3451–70. https://doi.org/10.1042/BCJ20180688.
79. Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry–based metabolite profiling in plants. Nat Protoc. 2006;1:387–96. https://doi.org/10.1038/nprot.2006.59.
80. Zanatta AC, Vilegas W, Edrada-Ebel R. UHPLC-(ESI)-HRMS and NMR-Based Metabolomics Approach to Access the Seasonality of Byrsonima intermedia and Serjania marginata From Brazilian Cerrado Flora Diversity. Front Chem. 2021;9:710025. https://doi.org/10.3389/fchem.2021.710025.
81. Nugroho S, Rahmadi HY, Simamora AN, Purba AR. 1H NMR metabolomic profiling of resistant and susceptible oil palm root tissues in response to Ganoderma boninense at the nursery stage. Sci Rep. 2025;15:16784. https://doi.org/10.1038/s41598-025-01691-y.
82. Lefort G, Liaubet L, Canlet C, Tardivel P, Père M-C, Quesnel H, et al. ASICS: an R package for a whole analysis workflow of 1D 1H NMR spectra. Bioinformatics. 2019;35:4356–63. https://doi.org/10.1093/bioinformatics/btz248.
83. Brans J-P, Mareschal B. Promethee Methods. In: Multiple Criteria Decision Analysis: State of the Art Surveys. New York: Springer-Verlag; 2005. p. 163–86. https://doi.org/10.1007/0-387-23081-5_5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable1MetaboliteprofileofprogenyRB03SComplete.xlsx