Citation: Rehman, B., Nasir, R., Ikram, H., & Umar, A. (2023). Evolutionary study of the DGAT1 gene indicates the presence of positive selection in mammals. *Global Drug Design & Development Review, VIII*(II), 48-62. https://doi.org/10.31703/gdddr.2023(VIII-II).07

▪ **Pages:** 48 – 62 ▪ **Vol.** VIII, **No.** II (Spring 2023) ▪ **p- ISSN:** 2788-497X ▪ **e- ISSN:** 2788-4120

Corresponding Author: Bushra Rehman (Assistant Professor, Institute of Biotechnology and Microbiology, Bacha Khan University, Charsadda, KP, Pakistan. Email: dr.bushrabkuc.edu.pk)

Cite Us

▪ **DOI:** 10.31703/gdddr.2023(VIII-II).07 ▪ **URL:** http://dx.doi.org/10.31703/gdddr.2023(VIII-II).07

Evolutionary Study of DGAT1 Gene Indicates the Presence of Positive Selection in Mammals

Bushra Rehman [a] Rabiya Nasir [b]

Hira Ikram [c] Aiman Umar [d]

Abstract: *This study was focused on the evolving role of DGAT1 using phylogenetic analysis. DGAT1 was retrieved from available databases of 37 mammalian genomes to detect signatures of positive selection in individual codons. Evolutionary analyses likelihood (REL), fast unconstrained Bayesian approximation (FUBAR), and Mixed-effects model evolution (MEME). REL detected 3 positive selection sites (at 109, 195, and 211 positions), and MEME detected 20 sites. Among them, 211 sites were confirmed as a strong positive selection of the DGAT1 gene. The codon model evolution (CME) model clustered amino acid substitution rates at three class levels (0.06, 0.20, and 0.40). Multiple hit tests showed that 20 of the DGAT1 triplet codes for amino acid formation were changed into new codes. In which, the GAC code for aspartate and GTC for valine had the maximum counts of 14 and 16. Strong evidence of positive selection of the DGAT1 gene in mammals was found.*

Key Words: Evolutionary Study, DGAT1, Mammals, Positive Selection Variation

Introduction

Evolutionary biology gives an understanding of genetic improvement happening over species. It helps us to realize the gene functions and variation processes including their mutations, duplication, and genome-wide associations (Arbuckle, 2020; Ball and Balthazart, 2021). Wang et al. (2017) reported that the molecular evolution of some genes with different breeding histories can be used as a unique population structural genes. Which, can explain the genetic variation mechanisms and complex pathways. The primary enzyme that regulates the production of triacylglycerides (TGA) in the endoplasmic reticulum (ER) membrane is called diacylglycerol acyltransferase (DGAT) (Ying et al., 2017). Different organisms have been shown to harbour four DGAT subfamilies: DGAT1, DGAT2, DGAT3, and WAX-

[a] Assistant Professor, Institute of Biotechnology and Microbiology, Bacha Khan University, Charsadda, KP, Pakistan.
[b] Assistant Professor, Center of Excellence in Marine Biology, Karachi, Sindh, Pakistan.
[c] Lecturer, Institute of Biotechnology and Microbiology, Bacha Khan University, Charsadda, KP, Pakistan.
[d] Assistant Professor, Department of Food and Nutrition, Shaheed Benazir Bhutto Women University Peshawar, KP, Pakistan.

DGAT (Liu et al., 2020). whereas the main enzymes for TGA production and storage in animal tissues are DGAT1 and DGAT2 (Chitraju et al., 2019). Nevertheless, there is no evolutionary connection between the two enzymes. Consequently, a number of studies (Canon-Beltran et al., 2020; Hua et al., 2018; Karis et al., 2020) are attempting to determine how these two essentially distinct enzymes contribute to TGA storage in mammalian adipose tissues. As an ER multitopic membrane protein with a luminal side active site, DGAT1 is a member of the O-acyltransferase ER membrane gene family (Canon-Beltran et al., 2020). (Ma et al., 2018). This gene is a potential gene for milk fat and is found in the QTL of chromosome 14 (BTA14) areas (Bovenhuis et al., 2016). The evolution of DGAT1 has been shown to be driven by positive selection, however, the specific codons that are selected are unknown. DGAT1's changing function was investigated by phylogenetic analysis. In addition to diacylglycerol, this enzyme can transfer the fatty acyl moiety to a variety of acceptors, such as retinol or long-chain alcohols, to create retinyl esters or waxes, respectively (Chitraju et al., 2019). Additionally, Maraschin et al. (2019) described the Kennedy pathway and the monoacylglycerol pathway, citing the glycerol phosphate as one of the two primary channels for the synthesis of triacylglycerides. For example, Tabaran et al. (2015) found that the DGAT1 gene is essential for the synthesis of TGA in milk fat since it codes for the DGAT1 enzyme. As a result, the DGAT1 gene is regarded as one of the most significant milk fat candidate genes (Bovenhuis et al., 2016). Numerous bioinformatics models are widely employed to monitor the favourable choices of mammalian genes. Using the Markov Chain Monte Carlo (MCMC) procedure, the MEME model, for example, can combine the fixed effects to find instances of both positive selection and episodic diversifying selection at the individual branch site and rapid unconstrained Bayesian approximation (FUBAR) (Bakiu et al., 2015). According to Bartakova et al. (2021), the advantage of MEME is that it can enable each site to have its own selective history without requiring the branches that are being selected to be partitioned beforehand. Additionally, the evolutionary conservation of amino acid locations is crucial for preserving the structure and functionality of proteins. For instance, Onodera et al. (2019) reported that mammals with serine at 366 positions have more potential to control some cellular metabolism including biosynthesis and mitochondrial oxidation. Therefore, the detection of selected sites may enlighten the selection forces and detect the functionally significant sites for protein interaction. The Mechanistic-Empirical Combination (MEC) model can estimate the selection pressure at particular codons (Onodera et al., 2019). In order to distinguish the positive selection DGAT1 gene among various mammalian species, this study sought to examine selection model markers using the maximum likelihood probability approach. Additionally, it sought to provide information regarding the applicability of these markers assisted selection in the species diversity.

Materials and Methods

The nucleotide and amino acid sequence-coding DGAT1 gene

Ensembl (http://useast.ensembl.org/index.html), Uniprot (http://www.uniprot.org), and NCBI (www.ncbi.nlm.nih.gov/genbank) databases provide DGAT1 gene coding nucleotides and amino acid sequences. Clustal Omega was used in the MEGA 6.0 programme to align the DGAT1 protein sequences (Tamura et al., 2013). The species was identified using the accession number in addition to the mRNA and protein accession numbers shown in (Table 1). We used maximum likelihood methods in MEGA 6.0 to generate the phylogenetic tree of the DGAT1 gene. For taxonomic clustering, 1000 repeats were available using bootstrapping. Asif and Associates, 2017.

Molecular evolution and positive selection of DGAT1 codon

The molecular basis of evolution and the relevance of positive selection of DGAT1 was designed by

analysing the codon and sequence of DGAT1 and comparing the dN/dS ratio of ω for two maximum likelihood methodologies (Ahmad et al., 2018; Ahmad et al., 2019). Data Monkey (http://www.datamonkey.org/) and the HyPhy package were the software tools used. To discover the values globally of ω, we employed various likelihood programmes such as internal fixed effect likelihood (IFEL), rapid unconstrained Bayesian approximation (FUBAR), and random effect likelihood (REL). For positive site picks, the REL employed a 95% confidence interval and Bayes factor values greater than 20. According to Ahmad et al. (2017b), the other study used p values < 0.05 to evaluate significance. The estimated parameters were aligned using the multiple EM for motif elicitation (MEME) model and the equation ω = β/α. The alternative model includes four factors for each site: β , β+, q , and α calculating site-to-site substitution variability rates. The selective pressure was estimated using two parameters, β: β

α and β+. Based on the χ2 asymptotic distribution, values p < 0.05 were deemed significant (Pond & Muse, 2005).

Protein prediction by amino acids and nucleic acids

The online database Con Surf, accessible at http://consurftest.tau.ac.il, was utilised to forecast the proteins that have retained their amino and nucleic acid compositions over time. For sequence codon alignment of DGAT1, Yang et al. (2012) employed Selection version 2.2 (http://selecton.tau.ac.il/), which implements the mechanistic-empirical combination (MEC) model for predicting adaptive selection pressure at distinct codons. Ramachandran plot is used using http://vadar.wishartlab.com/ to predict the DGAT1. Furthermore, utilising a genetic algorithm and codon model evolution based on synonymous and nonsynonymous substitution rates, the evolutionary fingerprinting of the DGAT1 gene was deduced. The phylogenetic Markov model was utilised to explain the evolution of the codon model, which included character frequencies, replacement rates for amino acids, and amino acid substitution rate clustering.

Table 1

List of species and accession number of the NCBI gene bank database, which was used for the hypothesis testing

No	Species common name	Scientific name	Accession Number	Protein Accession Number
1	Human	*Homo sapiens*	NM_005465.4	NP_859029.1
2	House Mouse	*Mus musculus*	NM_001357390.1	XP_030109770.1
3	Norway rat	*Rattus norvegicus*	XM_006250322.3	XP_006250383.1
4	Chimpanzee	*Pan troglodytes*	XM_016934876.1	XP_016791361.1
5	White-tufted-ear marmoset	*Callithrix jacchus*	XM_017966707.1	XP_008983976.1
6	Feral Cattle	*Bos taurus*	NM_001191309.1	XP_024831736.1
7	Painted turtle	*Chrysemys picta*	XM_008170470.2	XP_008168692.1
8	Sheep	*Ovis aries*	XM_012187897.2	XP_027831384.1
9	Rhesus monkey	*Macaca mulatta*	NM_001266640.1	XP_028701012.1
10	Damara mole-rat	*Fukomys damarensis*	XM_010642598.2	XP_010640900.1
11	Chinese tree shrew	*Tupaia chinensis*	XM_014591111.1	XP_014446597.1
12	Water buffalo	*Bubalus bubalis*	XM_006045843.1	XP_006045905.1
13	Domestic ferret	*Mustela putorius*	XM_004756776.2	XP_012913885.1
14	Chinese hamster	*Cricetulus griseus*	XM_003508130.3	XP_016833997.1
15	Miniopterusnatalensis	*Miniopterus natalensis*	XM_016201398.1	XP_016056884.1
16	Egyptian fruit bat	*Rousettus aegyptiacus*	XM_016151253.1	XP_016006738.1

No	Species common name	Scientific name	Accession Number	Protein Accession Number
17	Sooty mangabey	*Cercocebus atys*	XM_012036298.1	XP_011891692.1
18	Chinese soft-shelled turtle	*Pelodiscus sinensis*	XM_006135202.2	XP_014435074.1
19	Cheetah	*Pelodiscus sinensis*	XM_015072881.1	XP_026902031.1
20	Domestic cat	*Felis catus*	XM_023247428.1	XP_023103194.1
21	Giant panda	*Ailuropoda melanoleuca*	XM_011223567.2	XP_011221869.1
22	Green monkey	*Chlorocebus sabaeus*	XM_007989961.1	XP_007988151.1
23	Gray short-tailed opossum	*Monodelphis domestica*	XM_016429466.1	XP_007481609.1
24	Long-tailed chinchilla	*Chinchilla lanigera*	XM_005374760.2	XP_005374816.1
25	Naked mole-rat	*Heterocephalus glaber*	XM_004853513.3	XP_004853570.1
26	Northern white-cheeked gibbon	*Nomascus leucogenys*	XM_012506588.1	XP_030668452.1
27	Przewalski horse	*Equus caballus*	XM_023632779.1	XP_008525153.1
28	Prairie vole	*Microtus ochrogaster*	XM_005369606.2	XP_013202436.1
29	Pacific walrus	*Odobenus rosmarus*	XM_012562385.1	XP_012417839.1
30	Pig-tailed macaque	*Macaca nemestrina*	XM_011729484.1	XP_011727791.1
31	Sumatran orangutan	*Pongo abelii*	XM_009248019.1	XP_024089808.1
32	Wild Bactrian camel	*Camelus ferus*	XM_014559995.1	XP_006185623.1
33	Western European hedgehog	*Erinaceus europaeus*	XM_007534664.1	XP_007534733.1
34	Weddell seal	*Leptonychotes weddellii*	XM_006740006.1	XP_006740067.1
35	American beaver	*Castor canadensis*	XM_020163460.1	XP_020019049.1
36	Australian saltwater crocodile	*Crocodylus porosus*	XM_019549778.1	XP_019405323.1
37	Koala	*Phascolarctos cinereus*	XM_020982080.1	XP_020837737.1
38	Anubis baboon	*Papio anubis*	XM_017958504.2	XP_017813935.1
39	Zebrafish	*Danio rerio*	NM_001197201.2	XP_001923454.3

Analysis of the protein-protein interaction network

Using the special linkage analysis of STRING (version 9.1, http://www.string-db.org/), the protein-protein interaction network with DGAT1 was predicted (Franceschini et al., 2013). Protein connections were identified using the online server data bank of biological interactions. According to Li et al. (2017), the pooled score < 0.4 was the cutoff standard value. The conserved motif analysis was conducted using the MEME online programme. The MEME website (http://meme-suite.org/tools/meme) provided an explanation of conserved motifs in the DGAT1.

Results

Phylogenetic relationship of DGAT1

To determine the origins of evolutionary alterations in the genes of 37 domesticated and wild mammalian species, phylogenetic tree analysis was conducted for the DGAT1 gene (Figure 1). which placed species with less significant relationships in different phylogenetic groups, such as the Western European hedgehog (Erinaceus europaeus) and the painted turtle

(Chrysemys picta), and placed species with evolutionarily close genes, such as the green monkey (Chlorocebus sabaeus) and Anubis baboon (Papio anubis), in the same groups. As a result, nine groupings were created from 37 species. We have determined the sites of the positive selection codons in these mammalian clades using phylogenetic tree analysis. We have determined the sites of the positive selection codons in these mammalian clades using phylogenetic tree analysis.

By estimating global ω values using random effect likelihood (REL), internal fixed effect likelihood (IFEL), and fast unconstrained Bayesian approximation (FUBAR) techniques, evolutionary evidence of positive selection was found. Three positive selection sites (at locations 109, 195, and 211) were found by the REL. Three positive selection sites were also found by IFEL (at codons 106, 254, and 319). In the meanwhile, 11 sites under positive selection were found by FUBAR analysis at locations 14, 21, 52, 153, 164, 221, 323, 325, 348, 349, and 480 (Table 2). which, at a 95% confidence interval, FUBAR found more positive locations than other analyses. Additionally, REL analysis, which found positive selection values >20, was based on the Bayes factor. When p-values were less than 0.05, every discovered site showed a significant difference (Table 2).

Table 2

Sites under positive selection in DGAT1 genes by using different approaches

IFEL	REL	FUBAR
Positive sites (p-value)	Positive sites (Bayes Factor)	Positive sites (Posterior Probability)
106(0.02), 254(0.04), 319(0.09)	109(45.07), 195(44.59), 211(152.35)	14(0.54), 21(0.54), 52(0.59), 153(0.92), 164(0.53), 221(0.66), 323(0.56), 325(0.55), 348(0.65), 349(0.56), 480(0.81)

Significant values (p < 0.05), Bayes factors >20.

Table 3 reports the distribution of synonymous (α) and non-synonymous (β) substitution rates over sites inferred by the MEME model, where the proportion of branches with β>α is significantly greater than 0. The *p*-value was derived using a mixture of χ2 distributions. In which, MEME analysis detected 20 sites undergone episodic diversifying selection (Table 3). Among these sites, 28, 29, 36, 42, 70, 74, 211, 244, 257, 282, and 542 were detected under episodic diversifying with *p*-values < 0.05. This model also estimated α and β substitution rates and the sites having values β > α were considered significant and determined these sites under diversifying selection (Ahmad et al., 2017a). For instance, 211 sites were inferred to have experienced pervasive nonsynonymous substitution throughout the evolutionary history with *a p*-value < 0.05. Moreover, this site evolved with β+ > α, and it was under positive selection in the other analyses. Therefore, it has been confirmed that the 211 site was considered as a strong positive selection of the *DGAT1* gene, whereas all other sites were conserved (Table 3).

Table 3

Mixed-effect model evolution (MEME) based on the episodic diversifying selection of DGAT1 genes

Codon	α	β^-	Pr[β=β^-]	β^+	Pr[β=β^+]	*p*-value
22	0.13	0.00	0.96	56.28	0.04	0.06
23	0.75	0.05	0.95	14.31	0.05	0.07
28	0.15	0.00	0.80	1.78	0.20	0.02
29	0.00	0.00	0.84	9.24	0.16	0.00

Codon	α	β^-	Pr[β=β^-]	β^+	Pr[β=β^+]	*p*-value
36	0.32	0.14	0.95	31.17	0.05	0.03
42	2.01	0.13	0.93	88.60	0.07	0.02
44	0.00	0.00	0.56	2.64	0.44	0.07
70	1.26	0.00	0.89	149.54	0.11	0.03
74	0.53	0.53	0.93	301.72	0.07	0.00
112	0.43	0.00	0.52	1.98	0.48	0.07
147	0.22	0.05	0.97	4.21	0.03	0.09
211	0.00	0.00	0.00	0.37	1.00	0.03
224	0.14	0.14	0.00	1.02	1.00	0.09
244	0.08	0.00	0.60	0.54	0.40	0.05
254	0.14	0.14	0.18	1.62	0.82	0.06
257	0.30	0.06	0.96	10.53	0.04	0.03
282	0.16	0.16	0.97	11.15	0.03	0.03
310	0.30	0.07	0.92	4.33	0.08	0.10
500	0.45	0.04	0.94	7.93	0.06	0.08
542	0.58	0.04	0.88	78.39	0.12	0.00

Codon Model Selection

The genetic algorithm was used to generate the codon model, which was used to identify the evolutionary fingerprint seen in the coding regions of the DGAT1 genes. wherein both synonymous and nonsynonymous substitution rates were used in the evolutionary fingerprinting method. Figure 2 shows the evolution of the codon model using the phylogenetic Markov model (CME). Character frequencies, substitution rates, and the clustering of amino acid substitution rates at three class levels (0.06, 0.20, and 0.40) are all included in this model. While DENQ had 50% substitution, FWY and HKR had less than 50% substitution, and the substitution pair ACGILMPSTV had 90% substitution. whereas the highest substitution rate for the various ratio classes was 0.40. Furthermore, among the different amino acid positions in the DGAT1 genes, the minimum was 0.06 (Figure 2).

In addition, the plot in Figure 3 used Gaussian-approximated variance to show the estimated distribution of α and β rate alignments. In which, the colored pixels show the density of the posterior sample distribution for the given rate. The diagonal line represented the idealized neutral evolution regime (ω = 1). The points above the line correspond to positive selection (ω > 1), and the points below the line were for negative selection (ω < 1). The graph showed the neutral evolution and only a few sites under the circle above the diagonal had positive evolution at the *DGAT1* gene. In which, the probability of site-to-site distribution ratio (ω = β/α) based on the likelihood log and Akaike information criterion (AIC) of five neutral evolution classes were identified at *DGAT1* gene (Figure 3). And, the likelihood log was -19009.06134 for five class rates using 37 parameters.

The modified Bayesian Information Criterion (mBIC) of the DGAT1 in 4032 logs (L) was used to determine the codon models (Table 4). It was believed that the model with -19302.3 L was the most useful for studying amino acid substitution. The greatest estimate of a single rate (dN/dS) substitution in the third class (Table 4) was 0.40/15, and the distribution of amino acids was classified into three classes based on these rates (Figure 2). Furthermore, 644.48 mBIC and 342.30 L improvements are provided by this approach, assuming that the rate of all non-synonymous modifications stays constant.

Table 4

Codon model selection is based on the modified Bayesian Information Criterion (mBIC) of the DGAT1 gene from different organisms.

Classes	Models	Credible	mBIC	ΔmBIC	dN/dS (Rates in class)		
1	1	0	40974.5		0.16/75		
2	2935	0	40387.1	587.41	0.07/50	0.32/25	
3	1096	132	40330.0	57.07	0.06/42	0.20/18	0.40/15

N: number of rate classes included in models; Models: genetic algorithm models; Credible: all the models evaluated by a genetic algorithm within 9.21 mBIC unit (the best model has credible values 0.01 or >1); mBIC: modified Bayesian Information Criterion; ΔmBIC: mBIC for N rate classes compared to N 1 rate classes; dN/dS: maximum likelihood estimates for each rate class.

The maximum likelihood analysis of the DGAT1 gene's codon by codon positive selection is shown in Table 5. Onodera et al. (2019) found that when dN-dS values were higher than 1, locations were significant and codons had experienced positive selection. For instance, among the twenty-one identified codons, only codons 55 and 185 had dN-dS values smaller than 1, suggesting that they were not under positive selection.

Table 5

Maximum likelihood analysis of DGAT1 gene for codon by codon positive selection.

Codon #	Codon Start	Triplet	Syn (s)	Nonsyn (n)	Syn sites (S)	Nonsyn sites (N)	dS	dN	dN-dS
7	670	GGG	2.67	14.33	0.87	2.08	3.06	6.88	3.82
8	673	GTC	3.00	10.00	0.99	2.01	3.03	4.98	1.94
25	724	GAG	2.25	11.75	0.67	2.22	3.37	5.30	1.92
33	748	GTG	0.33	10.67	0.90	2.08	0.37	5.12	4.75
34	751	CTG	3.00	8.00	1.27	1.72	2.37	4.66	2.29
43	778	TTC	0.00	4.00	0.73	2.27	0.00	1.76	1.76
55	814	TGG	0.00	2.00	0.15	2.17	0.00	0.92	0.92
56	817	TGC	0.00	3.00	0.70	2.16	0.00	1.39	1.39
93	1093	ATG	0.00	3.00	0.20	2.80	0.00	1.07	1.07
113	1153	CAG	0.33	4.67	0.56	2.08	0.59	2.25	1.66
114	1156	AAC	0.00	3.00	0.67	2.33	0.00	1.29	1.29
120	1174	AAG	1.00	7.00	0.63	2.26	1.58	3.10	1.52
122	1180	ATG	0.00	3.00	0.30	2.70	0.00	1.11	1.11
161	1297	ATG	0.50	6.50	0.36	2.64	1.39	2.46	1.07
185	1375	TGG	0.00	2.00	0.19	2.12	0.00	0.94	0.94
194	1402	ATC	3.00	12.00	0.93	2.07	3.21	5.81	2.60
199	1417	AAG	0.33	5.67	0.62	2.25	0.53	2.52	1.99
201	1423	ATG	0.00	8.00	0.31	2.68	0.00	2.99	2.99
206	1438	AGC	3.67	15.33	0.83	2.15	4.44	7.12	2.68
255	1621	CTG	2.50	9.50	1.14	1.86	2.19	5.12	2.93
258	1630	ATC	0.00	3.00	0.74	2.26	0.00	1.33	1.33

dN-dS values > 1 indicate significantly the codons have undergone positive selection and positions with their synonymous and non-synonymous substitution site.

Positive selection of amino acid positions

The Mechanistic-Empirical Combination (MEC) the model estimated the selection pressure at particular codons at various codons in DGAT1 under the positive selection (Figure 4). The indicators of selection masked codons confirmed the positive selection differences on 11, 21, 191, 221, and 231 codons.

A map of all the discovered codon locations' overall performance is presented in Figure 5. which shows the ambiguous, synonymous, and nonsynonymous codon changes with the evolutionary time. Both the synonym and nonsynonymous mutant codons performed consistently up until codon number 20. After that, when the number of codons climbed to 580, there was a discernible improvement in performance. The synonymous cumulative rate was lower at codon 471 than the non-synonymous cumulative rate. Furthermore, the performance of the ambiguous codon increased as the codon placements were initiated, and it eventually stabilised until codon 471.

Ramachandran plot

The energetically permitted regions of the backbone dihedral angle psi (ψ, x-axis) against phi (φ, y-axis) of the amino acid residues were displayed using a Ramachandran plot. The energetically permitted portions of the amino acid residues backbone dihedral angle psi (ψ, x-axis) versus phi (φ, y-axis) that appeared in the protein structure were displayed using a Ramachandran plot (Vimala et al., 2021). In Figure 6, the amino acids are represented by a black dot, while the permitted regions with α-helical and β-sheet conformations are shown by a red region. The protein exhibits more right-handed α-helix secondary structure than left-handed α-helix and β-sheet conformations, according to the cluster dots displayed on the map.

Thus, phi and psi, respectively, have established the relative rotational angle of the protein torsion. The relative rotational angle of the protein torsion by phi and psi, respectively, was verified by the cluster dots displayed in the Ramachandran plot. This could explain how certain amino acids and tight atom interactions play a crucial role in the functionality of produced proteins (Vimala et al., 2021).

Motif Compositions of DGAT1 protein

To gain a better understanding of the variations in the development and function of proteins, as well as to identify the distinct areas of various DGAT1 proteins, conserved motifs were predicted through the use of the online MEME tool (Mulaudzi-Masuku et al., 2019). The analysis of conserved regions in the DGAT1 gene from various animals is displayed in Figure 7. Using BLAST, ten motifs in all were predicted and their sequences were confirmed. DGAT1 protein from the same group generally had comparable motifs. Furthermore, Table 6 shows that the lengths of the motifs varied from 29 to 50 amino acid residues.

Of the 37 species that were the subject of the study, 29 species had 10 motifs in each DGAT1 protein, whereas the remaining 8 species had 9 motifs. In eight species of DGAT1 proteins, motifs in the colours yellow, orange, and light blue were also absent. For example, neither the Chinese soft-shell turtle (Pelodiscus sinensis) nor the feral cattle (Bos taurus) have the light blue motif that is part of the membrane-bound O-acyltransferase family MBOAT and possesses amino acid residue sequence (RLLEMLFFTQLQVGLIQQWMVPTIQNSMK PFKDMDYSRIIE). Additionally, Table 6 shows that the Egyptian fruit bat (Rousettus aegyptiacus) and the Canadian beaver (Castor canadensis) lack the yellow motif, which is an amino acid residue sequence (APDKDGDVGSGHWELRCHRLQDSLFSS) with no description.

Table 6

The length of motifs ranged and amino acid residues.

No.	Motif	Size	Description
1	MQFGDREFYRDWWNSESVTYFWQNWNIPVHKWCJRHFY KPM	41	MBOAT, membrane-bound O-acyltransferase family
2	RLLEMLFFTQLQVGLIQQWMVPTIQNSMKPFKDMDYSRII E	41	MBOAT, membrane-bound O-acyltransferase family

No.	Motif	Size	Description
3	SNYRGILNWCVVMLILSNARLFLENLIKYGILVDPIQVVSLF LKDPYSWP	50	No Description
4	WAFTGMMAQIPLAWIVGRFFQGNYGNAAVWLTLIIGQPV AVLMYVHDYYV	50	MBOAT, membrane-bound O-acyltransferase family
5	TVSYPDNLTYRDLYYFLFAPTLCYELNFPRSPRIRKRFLLR	41	No Description
6	VAAFQVEKRLAVGALTEQAGLLLHVANLATILCFPAAVAL LVESITPVGS	50	No Description
7	LLKLAVPNHLIWLIFFYWLFHSCLNAVAE	29	No Description
8	SKWMARTGVFLASAFFHEYLVSIPLRMFR	29	MBOAT, membrane-bound O-acyltransferase family
9	ALMVYTILFLKLFSYRDVNLWCRZRRAKA	29	Cry35Ab1 HTH C-terminal domain
10	APDKDGDADVGSGHWELRCHRLQDSLFSS	29	No Description

Multi Hits Likelihood Test

A circus network was built to investigate the shift in triplet codes on the amino acids, and the multiple hits test was carried out using the Datamonkey web tool (Figure 8). The findings demonstrated that while the triplet code has changed, the amino acids in DGAT1 have remained conserved (Table 7). It led to the alteration of the amino acids and polypeptide synthesis by changing 20 triplet codes into new codes. Furthermore, aspartate's GAC code was the most frequently altered target code to become other amino acids (Table 7). Moreover, GTC for valine had the highest count of 16 to be changed into AAG and the highest count of 14 for GTC to be transformed into CAG code. Therefore, it could be concluded that in DGAT1 protein, the aspartic acid and valine were the most sourced amino acids for different other targeted amino acids.

Table 7

The list of DGAT1 amino acids remains conserved and also changed due to changes in the triplet code.

No	Source	Target	Source Amino Acid	Target Amino Acid	Count
1	GAC	CCA	D	P	2
2	GAC	CCG	D	P	2
3	GAC	CCT	D	P	2
4	GAC	CTG	D	L	4
5	GAC	ATG	D	M	2
6	GAC	CGG	D	R	2
7	GAC	ACA	D	T	2
8	GAC	TTG	D	L	2
9	GAC	YCT	D	S	2
10	GAC	TGG	D	W	1
11	GTC	CGA	V	R	2
12	GTC	AAG	V	K	16
13	GTC	ACT	v	T	4
14	GTC	AGA	v	R	5
15	GTC	AAA	v	K	3
16	GTC	CCG	v	P	3
17	GTC	CAG	v	Q	14
18	GTC	YGG	v	W	8
19	GTC	AAT	v	N	5
20	GTC	TAT	v	Y	5

The network of Protein-Protein Interaction (PPI)

Further information regarding the molecular function of DGAT1 could be obtained through the study of the protein-protein interaction network (Rao et al., 2014). The PPI linkage in this study consisted of 32 edges, or line networks connecting the nodes, and 11 nodes, which represent proteins encoded with DGAT1. (Figure 9). This suggests that DGAT1 can be connected to multiple proteins via the protein interaction network. It showed that DGAT1 was networked with the other ten essential genes that are co-expressed from the PPI network. PPAP2A, AGPAT2, PLD1, SLC27A2, CD36, RPE65, DGAT2, MOGAT3, MOGAT1, and MOGAT2 are some of these genes (Figure 9).

Phosphatidic acid phosphatase 2a (PPAP2A) is one of these genes that catalyses the hydrolysis and uptake of lipids from extracellular space by converting phosphatidylcholine (PC) to phosphatidic acid and diacylglycerol (Goto et al., 2021). Additionally, the acylglycerol O-acyltransferase family (DGAT2/MOGAT) includes the genes MOGAT1, MOGAT2, and the related MOGAT3 genes. These genes are involved in the synthesis of diacylglycerol (DAG) and triacylglycerol (TAG) from monoacylglycerol (MAG), and their related pathways include metabolism and glycerolipid metabolism (Agarwal et al., 2019). In biological signalling pathways, all of these genes may be involved in the overexpression of the DGAT1 gene.

Discussion

Given that they contain proteins with biological functions intended to protect the host, mammalian genes are among those that are developing the fastest (Madende & Osthoff, 2019). An attempt was undertaken to determine whether the places under positive selection had any specific functions that contributed to their evolution since these positions are probably correlated with sites of significant activity. For contemporary evolutionary research, identifying the selection in the mammalian genome provides a crucial research platform.

The DGAT1 gene's evolutionary changes among 37 mammalian species were the main focus of this investigation. which separated these 37 species into nine groups according to how similar they were. For example, the phylogenetic tree revealed that the genes of some species, such as P. anubis and C. sabaeus, had evolved closely together. Thus, in order to determine the sites of the positive selection codons, we performed an initial analysis of the phylogenetic tree. Farmanullah et al. (2020) identified the positive selection codon positions in the mammalian AKT3 gene using a phylogenetic tree; REL confirmed that these places comprise 32 legitimate selection sites. The DGAT1 gene in animals has undergone considerable evolutionary positive selection, according to the findings. This has three DGAT1 positive selection sites identified by REL and IFEL.

Additionally, REL analysis, which found positive selection values >20, was based on the Bayes factor. The amino acid substitutions under Bayes Empirical Bayes-based selection analysis demonstrated a positive selection of different amino acid positions in the MKRN3 gene by IFEL and REL models, according to Ahmad et al. (2019). Under positive selection, they found several sites with a Bayes factor > 20 and a $p < 0.05$. Consequently, REL may be a better model to identify the codon locations of DGAT1 that are subject to positive or negative selection.

Twenty locations that experienced episodic diversifying selection were identified by the MEME model based on the distribution of α and β. Of these, p-values < 0.05 were found for 28, 29, 36, 42, 70, 74, 211, 244, 257, 282, and 542 under episodic diversifying. wherein the DGAT1 gene's substantial positive selection at the 211 site was verified. According to Bartakova et al. (2021), the advantage of MEME is that it can enable each site to have its own selective history without requiring the branches that are being selected to be partitioned beforehand. Furthermore, Farmanullah

et al. (2020) discovered 20 MEME-based positive selection sites in the AKT3 gene.

The mBIC identified 19 codons that had positive selection with dN-dS values > 1. In which, the positions of amino acids' evolutionary conservation is important for maintaining the protein structure and function. For instance, Onodera et al. (2019) reported that mammals with serine at 366 positions have more potential to control some cellular metabolism including biosynthesis and mitochondrial oxidation. Therefore, the detection of selected sites may enlighten the selection forces and detect the functionally significant sites for protein interaction. Also, MEC confirmed the positive selection differences on 11, 21, 191, 221, and 231 codons. Moreover, based on the similarities between sequences on the phylogenic relationship, there were many conserved amino acids even with the positive selection presence (da et al., 2021). The collective performance showed different stability after codon number 20 with a significant increase till codon 580. Barbour et al. (2013) mentioned that the differences in the cumulative behaviour between the nonsynonymous and synonymous distribution could indicate the differences in positive selection substitution among the mammalian species.

Furthermore, based on the similarities between sequences on the phylogenic relationship, there were highly conserved amino acids even in the presence of positive selection (da et al., 2021). The overall performance showed a discernible improvement from codon 20 and persisted until codon 580. Variations in synonymous and nonsynonymous distributions' cumulative behaviour could indicate differences in positive selection replacement among mammalian species, according to Barbour et al. (2013).

Additionally, the DGAT1 PPI linkage contained 11 nodes, indicating that other proteins are able to link with DGAT1 via the protein interaction network. In the human BMP15 gene, Auclair et al. (2013) found positive selection signals across 24 mammalian species. Because DGAT1 is upregulated, these genes might be connected to biological signalling networks. AGPAT2 and PLD1, for example, are important genes in the route leading to phospholipid production. The class B scavenger receptor family of cell surface proteins includes the integral membrane protein CD36 antigen, which is involved in the import of fatty acids within cells (Jay et al., 2020).

Furthermore, during phototransduction, RPE65, which is expressed in the retinal pigment epithelium, is in charge of converting all-trans-retinyl esters to 11-cis-retinol (Jakobiec et al., 2021). As a result, DGAT1 is an enzyme that accelerates the last stage of triglyceride synthesis in adipose tissue.

Conclusion

The evolutionary investigation of the DGAT1 gene in mammals has a significant potential to control the expression of the related gene in the lipid mechanisms of mammal cells. This can open the scientific field for biological studies and their relations in the medical field. The present work included different bioinformatics methods as a point of clarification and comparison. The major benefit of the used models is that they can identify which specific codons and codes that should focus on in studying the DGAT1 gene for its translation, esterification, or metabolic action. This study emphasized the great importance of the structural differences of DGAT1 between the mammal species on their different functional responses against immune diseases.

Conflict of Interests

Regarding the release of this paper, the authors affirm that they have no conflicts of interest.

Acknowledgements

The authors would like to thank the Guangxi Key Laboratory of Buffalo Genetics, Reproduction, and Breeding project, the Ministry of Agriculture's Research Platform Construction project, and the International Exchange and Cooperation in Science and Technology project (SNKF-2016-01) for their financial support.

References

Agarwal, A.K., Tunison, K., Mitsche, M.A., McDonald, J.G., & Garg, A. (2019). Insights into lipid accumulation in skeletal muscle in dysferlin-deficient mice. *Journal of lipid research 60*, 2057-2073. https://doi.org/10.1194/jlr.RA119000399

Ahmad, H.I., Ahmad, M.J., Adeel, M.M., Asif, A.R., & Du, X. (2018). Positive selection drives the evolution of endocrine regulatory bone morphogenetic protein system in mammals. *Oncotarget 9,* 18435-18445. https://doi.org/10.18632/oncotarget.24240

Ahmad, H.I., Liu, G., Jiang, X., Liu, C., Chong, Y., & Huarong, H. (2017). Adaptive molecular evolution of MC1R gene reveals the evidence for positive diversifying selection in indigenous goat populations. *Ecology and evolution 7*, 5170-5180. https://doi.org/10.1002/ece3.2919

Ahmad, H.I., Liu, G.Q., Jiang, X.P., Liu, C.H., Chong, Y.Q., & Huang, H.R. (2017). Adaptive molecular evolution of MC1R gene reveals the evidence for positive diversifying selection in indigenous goat populations. *Ecology and evolution 7,* 5170-5180. https://doi.org/10.1002/ece3.2919

Ahmad, M.J., Ahmad, H.I., Adeel, M.M., Liang, A., Hua, G., Murtaza, S., Mirza, R.H., Elokil, A., Ullah, F., & Yang, L. (2019). Evolutionary Analysis of Makorin Ring Finger Protein 3 Reveals Positive Selection in Mammals. *Evolutionary bioinformatics online 15*, 1176934319834612. https://doi.org/10.1177/1176934319834612

Arbuckle, K. (2020). From molecules to macroevolution: Venom as a model system for evolutionary biology across levels of life. *Toxicon: X 6*, 100034. https://doi.org/10.1016/j.toxcx.2020.100034

Asif, A.R., Awais, M., Qadri, S., Ahmad, H.I., & Du, X. (2017). Positive selection of IL-33 in adaptive immunity of domestic Chinese goats. *Ecology and evolution 7,* 1954-1963. https://doi.org/10.1002/ece3.2813

Auclair, S., Rossetti, R., Meslin, C., Monestier, O., Di Pasquale, E., Pascal, G., Persani, L., & Fabre, S. (2013). Positive selection in bone morphogenetic protein 15 targets a natural mutation associated with primary ovarian insufficiency in human. *PloS one 8*, e78199. https://doi.org/10.1371/journal.pone.0078199

Bakiu, R., Korro, K., & Santovito, G. (2015). Positive selection effects on the biochemical properties of mammal pyroglutamylated RFamide peptide receptor (QRFPR). *Italian Journal of Zoology 82*, 309-326. https://doi.org/10.1080/11250003.2015.1018352

Ball, G.F., & Balthazart, J. (2021). Evolutionary neuroscience: Are the brains of birds and mammals really so different? *Current biology: CB 31*, R840-R842. https://doi.org/10.1016/j.cub.2021.05.004

Barbour, K.W., Xing, Y.Y., Pena, E.A., & Berger, F.G. (2013). Characterization of the bipartite degron that regulates ubiquitin-independent degradation of thymidylate synthase. *Bioscience Reports 33*, 165-173. https://doi.org/10.1042/BSR20120112

Bartakova, V., Bryjova, A., Nicolas, V., Lavrenchenko, L.A., & Bryja, J. (2021). Mitogenomics of the endemic Ethiopian rats: looking for footprints of adaptive evolution in sky islands. *Mitochondrion 57*, 182-191. https://doi.org/10.1016/j.mito.2020.12.015

Bovenhuis, H., Visker, M., Poulsen, N.A., Sehested, J., van Valenberg, H., van Arendonk, J., Larsen, L.B., & Buitenhuis, A.J. (2016). Effects of the diacylglycerol o-acyltransferase 1 (DGAT1) K232A polymorphism on fatty acid, protein, and mineral composition of dairy cattle milk. *Journal of Dairy Science 99*, 3113-3123. https://doi.org/10.3168/jds.2015-10462

Canon-Beltran, K., Giraldo-Giraldo, J., Cajas, Y.N., Beltran-Brena, P., Hidalgo, C.O., Vasquez, N., Leal, C.L.V., Gutierrez-Adan, A., Gonzalez, E.M., & Rizos, D. (2020). Inhibiting diacylglycerol acyltransferase-1

reduces lipid biosynthesis in bovine blastocysts produced in vitro. *Theriogenology 158*, 267-276. https://doi.org/10.1016/j.theriogenology.2020.09.014

Chitraju, C., Walther, T.C., Farese, & R.V., Jr. (2019). The triglyceride synthesis enzymes DGAT1 and DGAT2 have distinct and overlapping functions in adipocytes. *Journal of lipid research 60*, 1112-1120. https://doi.org/10.1194/jlr.M093112

Da, S.V.D., Polveiro, R.C., Butler, T.J., Hackett, T.A., Braga, C.P., Puniya, B.L., Teixeira, W.F.P., de, M.P.P., Adamec, J., & Feitosa, F.L.F. (2021). An in silico, structural, and biological analysis of lactoferrin of different mammals. *International journal of biological macromolecules 187*, 119-126. https://doi.org/10.1016/j.ijbiomac.2021.07.102

Farmanullah, Hosseini, S.M., Liang, A., Hua, G., Rehman, Z.U., Talpur, H.S., Salim, M., Ahmad, S., Abulaiti, A., Khan, M., Safdar, M., Kakar, I.U., Ahmad, Z., Ahmad, M.Z., Tingzhu, Y., Schreurs, N.M., Bano, I., & Yang, L. (2020). Adaptive Molecular Evolution of AKT3 Gene for Positive Diversifying Selection in Mammals. *BioMed research international 2020*, 2584627. https://doi.org/10.1155/2020/2584627

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., von Mering, C., & Jensen, L.J. (2013). STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research 41*, D808-815. https://doi.org/10.1093/nar/gks1094

Goto, H., Miyamoto, M., & Kihara, A. (2021). Direct uptake of sphingosine-1-phosphate independent of phospholipid phosphatases. *The Journal of Biological Chemistry 296*, 100605. https://doi.org/10.1016/j.jbc.2021.100605

Hua, C., Geng, Y., Niu, L., Chen, Q., Cai, L., Tao, S., Ni, Y., & Zhao, R. (2018). Stimulating lipolysis in subcutaneous adipose tissues by chronic dexamethasone administration in goats. *Livestock Science 214*, 62-67. https://doi.org/10.1016/j.livsci.2018.05.020

Jakobiec, F.A., Barrantes, P.C., Yonekawa, Y., Lad, E.M., & Proia, A.D. (2021). Subretinal Mononuclear Cells in Coats' Disease Studied with RPE65 and CD163: Evidence for Histiocytoid Pigment Epithelial Cells. *American journal of ophthalmology 222*, 388-396. https://doi.org/10.1016/j.ajo.2020.09.020

Jay, A.G., Simard, J.R., Huang, N., & Hamilton, J.A. (2020). SSO and other putative inhibitors of FA transport across membranes by CD36 disrupt intracellular metabolism, but do not affect FA translocation. *Journal of lipid research 61*, 790-807. https://doi.org/10.1194/jlr.RA120000648

Karis, P., Jaakson, H., Ling, K., Bruckmaier, R.M., Gross, J.J., Parn, P., Kaart, T., & Ots, M. (2020). Body condition and insulin resistance interactions with periparturient gene expression in adipose tissue and lipid metabolism in dairy cows. *Journal of Dairy Science 103*, 3708-3718. https://doi.org/10.3168/jds.2019-17373

Li, H., Zhao, X., Wang, J., Zong, M., & Yang, H. (2017). Bioinformatics analysis of gene expression profile data to screen key genes involved in pulmonary sarcoidosis. *Gene 596*, 98-104. https://doi.org/10.1016/j.gene.2016.09.037

Liu, J., Wang, Z., Li, J., Li, H., & Yang, L. (2020). Genome-wide identification of Diacylglycerol Acyltransferases (DGAT) family genes influencing Milk production in Buffalo. *BMC genetics 21*, 26. https://doi.org/10.1186/s12863-020-0832-y

Ma, D., Wang, Z., Merrikh, C.N., Lang, K.S., Lu, P., Li, X., Merrikh, H., Rao, Z., & Xu, W. (2018). Crystal structure of a membrane-bound O-acyltransferase. *Nature 562*, 286-290. https://doi.org/10.1038/s41586-018-0568-2

Madende, M., & Osthoff, G. (2019). Comparative genomics of casein genes. *The Journal of Dairy Research 86*, 323-330.

https://doi.org/10.1017/S0022029919000414

Maraschin, F.D.S., Kulcheski, F.R., Segatto, A.L.A., Trenz, T.S., Barrientos-Diaz, O., Margis-Pinheiro, M., Margis, R., & Turchetto-Zolet, A.C. (2019). Enzymes of glycerol-3-phosphate pathway in triacylglycerol synthesis in plants: Function, biotechnological application and evolution. *Progress in lipid research 73*, 46-64. https://doi.org/10.1016/j.plipres.2018.12.001

Mulaudzi-Masuku, T., Ikebudu, V., Muthevhuli, M., Faro, A., Gehring, C.A., & Iwuoha, E. (2019). Characterization and Expression Analysis of Heme Oxygenase Genes from Sorghum bicolor. Bioinformatics and biology insights 13, 1177932219860813. https://doi.org/10.1177/1177932219860813

Onodera, W., Asahi, T., & Sawamura, N. (2019). Positive selection of cereblon modified function including its E3 ubiquitin ligase activity and binding efficiency with AMPK. *Molecular phylogenetics and evolution 135*, 78-85. https://doi.org/10.1016/j.ympev.2019.03.001

Pasquali, C.C., Islam, Z., Adamoski, D., Ferreira, I.M., Righeto, R.D., Bettini, J., Portugal, R.V., Yue, W.W.-y., Gonzalez, A., Dias, S.M.G., & Ambrosio, A.L.B. (2017). The origin and evolution of human glutaminases and their atypical C-terminal ankyrin repeats. *Journal of Biological Chemistry 292*, 11572-11585. https://doi.org/10.1074/jbc.M117.787291

Pond, S.K., & Muse, S.V. (2005). Site-to-site variation of synonymous substitution rates. *Molecular biology and evolution 22*, 2375-2385. https://doi.org/10.1093/molbev/msi232

Rao, V.S., Srinivas, K., Sujini, G.N., & Kumar, G.N. (2014). Protein-protein interaction detection: methods and analysis. *International journal of proteomics 2014*, 147648. https://doi.org/10.1155/2014/147648

Tabaran, A., Balteanu, V.A., Gal, E., Pusta, D., Mihaiu, R., Dan, S.D., Tabaran, A.F., & Mihaiu, M. (2015). Influence of DGAT1 K232A polymorphism on milk fat percentage and fatty acid profiles in Romanian Holstein cattle. Animal biotechnology 26, 105-111. https://doi.org/10.1080/10495398.2014.933740

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular biology and evolution 30*, 2725-2729. https://doi.org/10.1093/molbev/mst197

Vimala, M., Stella Mary, S., Ramalakshmi, R., Muthu, S., Niranjana Devi, R., & Irfan, A. (2021). Quantum computational studies on optimization, donor-acceptor analysis and solvent effect on reactive sites, global descriptors, non-linear optical parameters of Methyl N-Boc-piperidine-3-carboxylate. *Journal of Molecular Liquids 343*, 117608. https://doi.org/10.1016/j.molliq.2021.117608

Wang, M.S., Otecko, N.O., Wang, S., Wu, D.D., Yang, M.M., Xu, Y.L., Murphy, R.W., Peng, M.S., & Zhang, Y.P. (2017). An Evolutionary Genomic Perspective on the Breeding of Dwarf Chickens. *Molecular biology and evolution 34*, 3081-3088. https://doi.org/10.1093/molbev/msx227

Yang, J.R., Liao, B.Y., Zhuang, S.M., & Zhang, J. (2012). Protein misinteraction avoidance causes highly expressed proteins to evolve slowly. *Proceedings of the National Academy of Sciences of the United States of America 109*, E831-840. https://doi.org/10.1073/pnas.1117408109

Ying, F., Gu, H., Xiong, Y., & Zuo, B. (2017). Analysis of Differentially Expressed Genes in Gastrocnemius Muscle between DGAT1 Transgenic Mice and Wild-Type Mice. *BioMed research international 2017*, 5404682. https://doi.org/10.1155/2017/5404682

Figures Captions

Figure 1: Phylogenetic tree of DGAT1 gene of 37 different species. The phylogenetic tree was constructed using the Maximum Likelihood (ML) method. Bootstrap values in percentage (1000 replicates) are indicated on the nodes. Different groups are shaded in different colours.

Figure 2: Evolutionary rate clustering in structured genetic algorithm models. The models were inferred from the DGAT gene alignment of different species. Each cluster was labelled with the Maximum Likelihood estimate of its inferred rate. The nodes (residues) are annotated by their biochemical properties and Stanfel class, and the rates (edges) are labelled with the model-averaged rate estimates.

Figure 3: Evolutionary fingerprints of the DGAT1 gene inferred from alignments on the log scale, with the diagonal line corresponding to the values $\alpha = \beta$ for neutral evolution. Dots in the circle indicating the ratio β/α and the area of the circle represent the weight of rate classes. The points above the diagonal correspond to positive selection and below the diagonal negative selection.

Figure 4: Selection pressures among goat DGAT1 gene sequences using the mechanistic-empirical combination (MEC) model of selection online tool. Yellow and brown highlights represent positive selection, grey and white highlights represent neutral selection, and purple highlight represents negative selection on codons.

Figure 5: Cumulative behaviour of synonymous, non-synonymous, and ambiguous codon mutation changes per site compared to the mutation changes estimates between the number of synonymous and non-synonymous mutations (ambiguous codons) per site. Red plot indicating non-synonymous, green synonymous, and blue for ambiguous codon positions in each group.

Figure 6: Ramachandran plot prediction to examine the structure of a protein, the conformation of the amino acid present in the protein, and close associates between the atoms.

Figure 7: Conserved protein motif analysis of PtPL1s by MEME. Conserved motifs in the DGAT1 were elucidated by the MEME website (http://meme-suite.org/tools/meme). Up to 10 motifs were shown by different colours from different species.

Figure 8: Multi Hits Likelihood Test by circus graph to show the multihits of each codon to its target codon.

Figure 9: The protein-protein interaction (PPI) network built by the STRING database for the DGAT1 gene. Grey and Red circles characterize downregulated and upregulated genes, respectively. Line thickness indicates the strength of the interaction. Dash and solid edge mean negative and positive correlation coefficients. Network nodes denote proteins' post-transcriptional modifications or splice isoforms, and each node represents all the proteins produced by a single, protein-coding gene locus.