



OPEN

## Transcriptome analysis of five different tissues of bitter gourd (*Momordica charantia* L.) fruit identifies full-length genes involved in seed oil biosynthesis

Kumar Ravichandiran & Madasamy Parani✉

The bitter gourd seed oil, rich in conjugated fatty acids, has therapeutic value to treat cancer, obesity, and aging. It also has an industrial application as a drying agent. Despite its significance, genomics studies are limited, and the genes for seed oil biosynthesis are not fully understood. In this study, we assembled the fruit transcriptome of bitter gourd using 254.5 million reads (Phred score > 30) from the green rind, white rind, pulp, immature seeds, and mature seeds. It consisted of 125,566 transcripts with N50 value 2,751 bp, mean length 960 bp, and 84% completeness. Transcript assembly was validated by RT-PCR and qRT-PCR analysis of a few selected transcripts. The transcripts were annotated against the NCBI non-redundant database using the BLASTX tool (E-value < 1E-05). In gene ontology terms, 99,443, 86,681, and 82,954 transcripts were classified under biological process, molecular function, and cellular component. From the fruit transcriptome, we identified 26, 3, and 10 full-length genes coding for all the enzymes required for synthesizing fatty acids, conjugated fatty acids, and triacylglycerol. The transcriptome, transcripts with tissue-specific expression patterns, and the full-length identified from this study will serve as an important genomics resource for this important medicinal plant.

Bitter gourd (*Momordica charantia* L.) is a vegetable crop widely cultivated in Asia and Africa. It is also called bitter melon, balsam-apple, bitter apple, balsam pear, balsam birne, African cucumber, bitter cucumber, bittergurke, karela, and carilla gourd<sup>1</sup>. It is a climber belonging to the Cucurbitaceae family and has about 60 annual and perennial species<sup>2,3</sup>. The bitter gourd genome is made of  $2n = 2x = 22$  chromosomes and 339 Mb. Bitter gourd fruits and leaves have a long history of use in traditional medicine for treating diabetes, infection, wounds, and osteoarthritis<sup>4</sup>. It can be used for birth control, abortion, and treating ailments like jaundice, leprosy, piles, psoriasis, and rheumatism<sup>5,6</sup>. In a randomized controlled trial, oral intake of dried bitter gourd fruit pulp 2000 mg/day exhibited hypoglycemic effect as indicated by a significant reduction in fructosamine levels<sup>7</sup>. In another clinical trial, daily consumption of 2.5 g of dried whole bitter gourd fruit powder (equivalent to 50 g fresh bitter gourd fruit) significantly lowered the blood glucose levels in individuals with prediabetes. The glucose lowering effect was higher among the participants with higher baseline fasting plasma glucose (FPG) levels<sup>8</sup>. A randomized controlled trial in type-2 diabetes patient without other complications showed that supplementation of bitter gourd fruit powder along with Metformin and Glibenclamide provided improved glycaemic control ( $p < 0.05$ )<sup>9</sup>.

Besides the fruits and leaves, bitter gourd seed oil (BGSO) also has therapeutic and industrial applications. The BGSO content in different bitter gourd varieties varies between 28 and 30%<sup>10,11</sup>. It is a rich source of polyunsaturated fatty acids and is considered a special oil due to the presence of 56–62%  $\alpha$ -eleostearic acid<sup>10,12</sup>. BGSO is used to manufacture paints, inks, and other coating materials because of its excellent drying property<sup>13</sup>. BGSO has significant health benefits because of its use in treating diabetes, cancer, obesity, and aging<sup>14–17</sup>. These properties were supported by mechanistic studies in the cell lines and animal models. Treatment with 10–100 mg/Kg body weight BGSO significantly reduced edema, epithelial disruption, mucosa erosions, and ulceration in

Department of Genetic Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Chengalpattu, Tamil Nadu 603203, India.  
✉email: paranim@srmist.edu.in

S. No	Sample name	Number of reads (million)	Number of bases (Gb)	Q30 >	
				Number of reads (million)	Number of bases (Gb)
1	Green rind	33.3	4.8	28.1	3.3
2	White rind	51.0	7.4	43.8	5.2
3	Seed	52.6	7.5	45.7	5.5
4	Pulp	74.9	10.8	64.0	7.6
5	Immature seeds	84.2	12.1	72.9	8.5
	Total	295.9	42.6	254.5	30.1

**Table 1.** Summary of raw reads and Q30 reads obtained from different tissues of bitter gourd.

Particulars	Number
Number of raw reads (in million)	295.9
Number of clean reads (in million)	254.5
No of bases (after processing in Gb)	30.1
Mean Phred Score	30
Total transcripts	125,566
Total length (bases in million)	196.86
Average length (bases)	1567.86
Median contig length	960
GC (%)	40.61
Contig N50 (bases)	2,751

**Table 2.** De novo assembly statistics of bitter gourd fruit transcriptome.

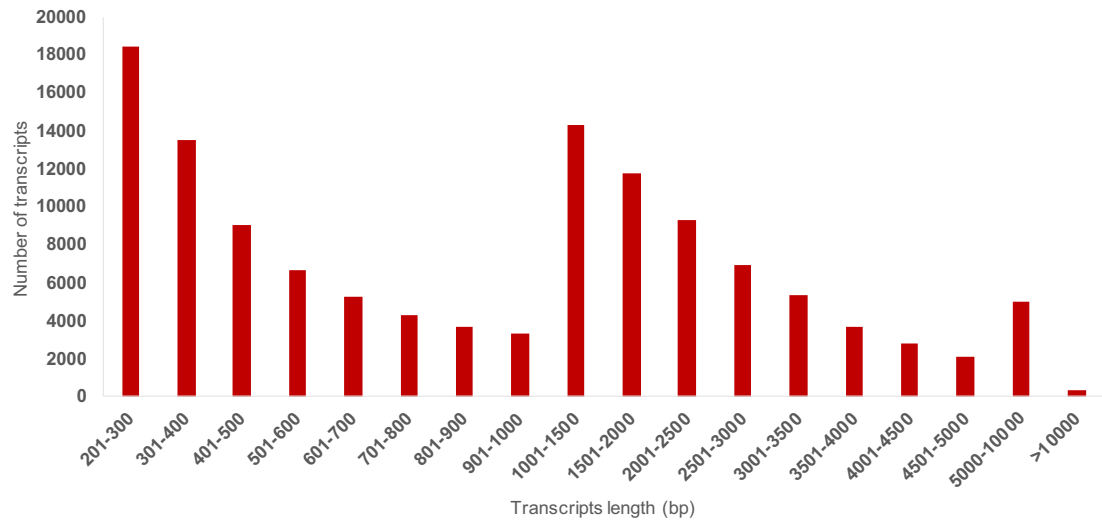
rats. The gastric juice of the treated rats showed elevated the pH without decreasing total acidity<sup>18</sup>. Accumulation of intracellular triglycerides was significantly lower in the hepatoma cells treated with BGSO than those treated with  $\alpha$ -linolenic acid. Anti-lipidemic property of BGSO was confirmed in vivo in mice, and it was attributed to an increase in NAD<sup>+</sup>/NADH ratio and activation of PPAR $\alpha$ , AMPK, and SIRT1 pathways<sup>19</sup>. BGSO showed anti-adiposity effect in mice, as indicated by the fat and plasma leptin content<sup>20</sup>. BGSO exhibited hypoglycemic potential by inhibiting  $\alpha$ -glucosidase and  $\alpha$ -amylase by 53% and 38%, respectively<sup>21</sup>. Oral diet supplemented with 0.5% BGSO nanoemulsion conferred protection against oxidative stress in the experimentally induced diabetic rats<sup>22</sup>.

Anti-cancer property of BGSO is supported by several experiments conducted in vitro and in vivo. BGSO triggered apoptosis in the colon cancer cells by upregulating GADD45 and p53 and downregulating Bcl-2 expression<sup>23</sup>. It was also reported that  $\alpha$ -eleostearic acid and its dihydroxy derivatives from BGSO inhibited the growth of leukemia and colon carcinoma cell lines<sup>24</sup>. In breast cancer cell lines,  $\alpha$ -eleostearic acid showed anti-cancer activity by arresting the cell cycle and decreasing HER2/HER3 levels<sup>25</sup>. It was observed that  $\alpha$ -eleostearic acid inhibited tumor formation and metastasis in the triple-negative breast cancer mice model<sup>26</sup>. Oral supplementation of hepatocellular carcinoma-induced rats with BGSO reduced hepatic dysplastic nodules and neoplastic lesions<sup>27</sup>.

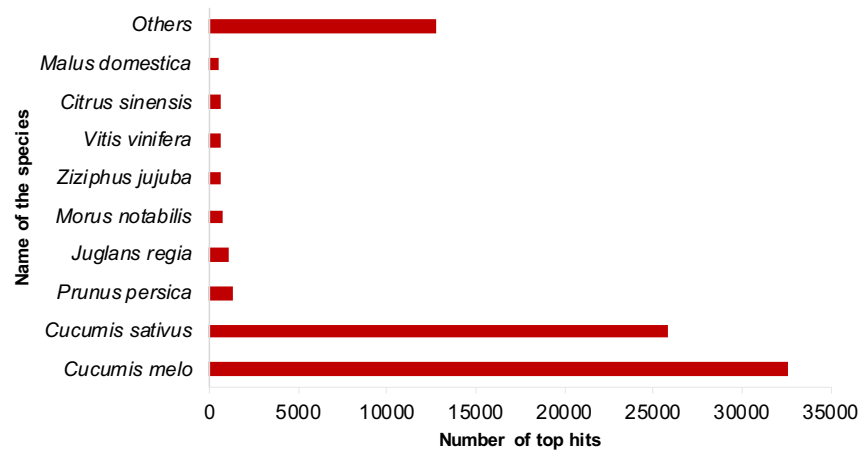
Despite the significance of bitter gourd in therapeutic and industrial applications, research on large-scale genomic studies and cloning of the genes involved in oil biosynthesis is limited. RNA-Sequencing is a cost-effective and high-throughput method to study genes and their expression patterns in qualitative and quantitative aspects. It also helps understand the functional elements of the genome and identify the genes involved in various pathways<sup>28,29</sup>. In the current study, we analyzed the bitter gourd fruit transcriptome in detail and identified full-length genes coding for all the enzymes involved in the biosynthesis of BGSO.

## Results and discussion

**Assembly of bitter gourd fruit transcriptome.** We generated 295.9 million 150 bp raw reads from the green rind, white rind, pulp, immature seeds, and mature seeds of bitter gourd. The reads from individual tissues varied between 33.3 million (green rind) and 84.2 million (immature seeds). Post quality filtering for low-quality reads (Phred score < 30) and adaptor removal, 254.5 million reads (86%) were retained for further processing (Table 1). About 14% of the reads were discarded after quality filtering. Even though the rejection rate is higher than the average percentage of data discarded after quality filtering, the total number of reads available for assembly is much higher than the requirement for a good assembly. Using trinity assembler, assembly of these reads generated 146,719 transcripts with 1,064 bp mean length and 2,855 bp N50. After clustering using CD-HIT, we obtained 125,566 transcripts with 960 bp mean length and 2,751 bp N50 (Table 2). The length distribution of the assembled transcripts is given in Fig. 1. The assembly parameters indicated that the transcripts assembled in this study could be reliably used for gene prediction and other downstream analyses.



**Figure 1.** Length distribution of the assembled transcripts from bitter gourd fruit transcriptome.

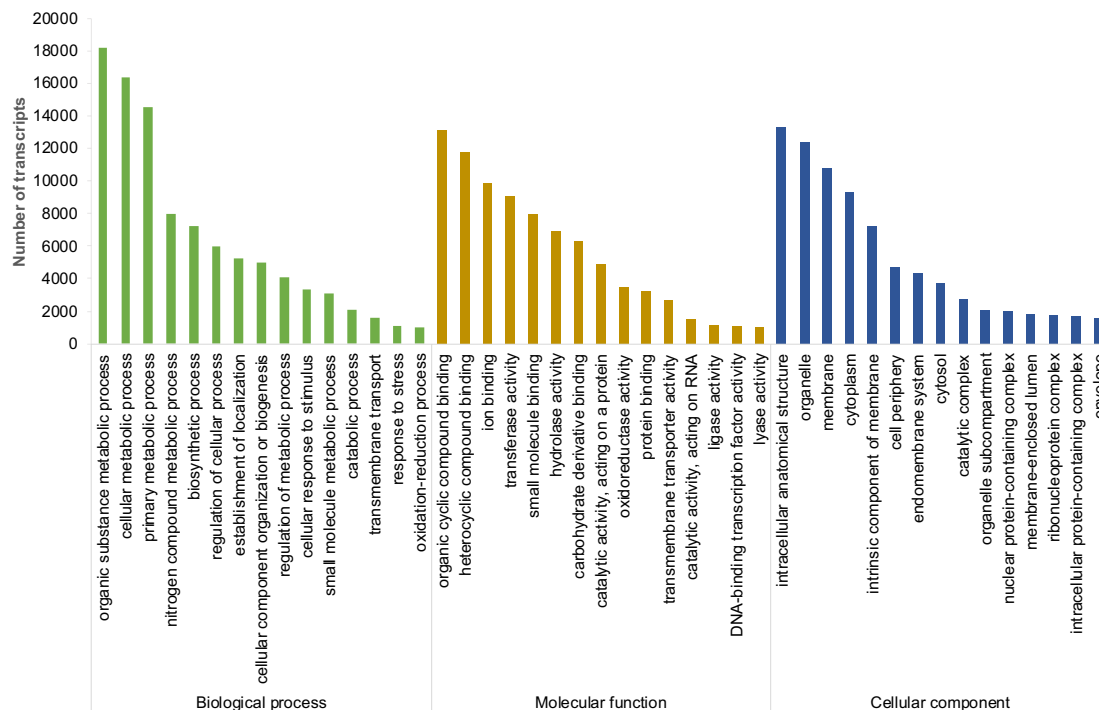


**Figure 2.** BLASTX top hit species distribution of bitter gourd transcripts against the nr database.

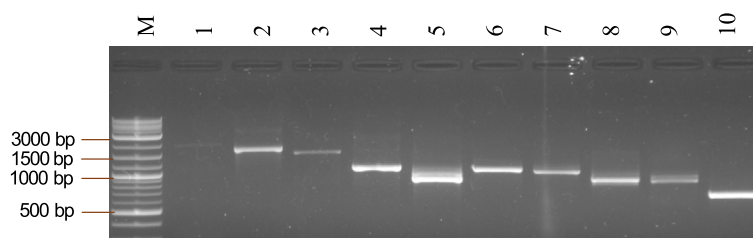
**Annotation of the transcripts.** Based on BLASTX similarity searches, we annotated 76,794 transcripts, and 48,771 transcripts did not show significant similarity with the existing genes. The BLASTX results were analyzed using BLAST2GO for further annotation. As shown in Fig. 2, the top-hit species for the bitter gourd transcripts was *Cucumis melo* followed by *Cucumis sativus*. According to the APG IV phylogenetic classification<sup>30</sup>, both the top-hit species are taxonomically closely related to bitter gourd.

**Functional classification of the transcripts.** The functional classification of the transcripts based on GO terms showed that 99,443 and 86,681, and 82,954 were classified under biological process, molecular function, and cellular component, respectively. Under the biological process, the highest number of the transcripts were related to the metabolic process (18,231 transcripts), followed by the cellular metabolic process (16,342 transcripts) and primary metabolic process (14,567 transcripts). In the case of molecular function, transcripts involved in organic cyclic compound binding were more abundant (13,119 transcripts), followed by heterocyclic compound binding (11,765 transcripts) and ion binding (9,878 transcripts). In the cellular component category, transcripts responsible for the intracellular anatomical structure were found in more numbers (13,265 transcripts) than organelle (12,412 transcripts), membrane (10,768 transcripts), and others (Fig. 3).

**Quantification and completeness analysis of the transcripts.** The abundance of the transcripts in the transcriptome was calculated based on the FPKM value using RSEM. The data for all the transcripts from the respective tissues are given in Supplementary Table S3. The top 15 abundant transcripts in each tissue and overall abundance in the fruit transcriptome are given in Supplementary Table S4. The results showed that some genes are abundantly expressed in each tissue. The observed expression patterns of the transcripts can be used to isolate tissue-specific genes and promoters for applied research in bitter gourd. For example, the proteins from the bitter gourd fruit pulp showed antidiabetic properties in vitro and in vivo<sup>31,32</sup> but the corresponding genes are



**Figure 3.** Gene Ontology classification of bitter gourd transcripts under three major categories of GO terms.



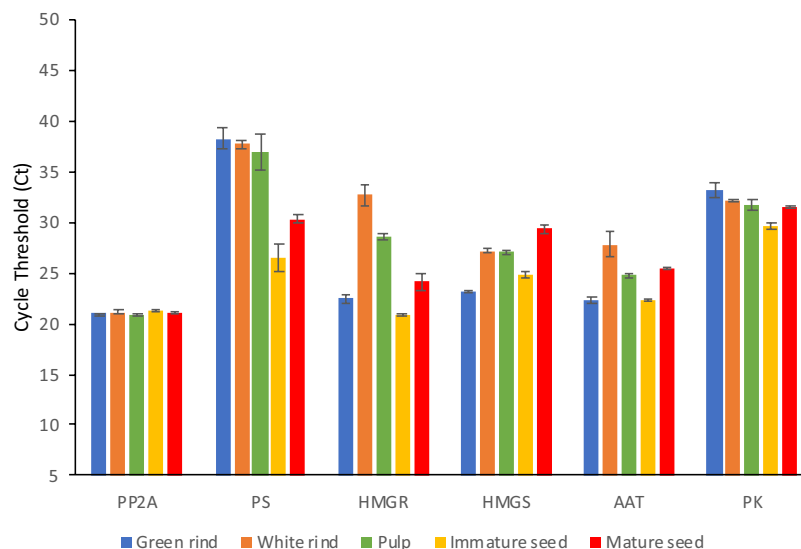
**Figure 4.** Reverse transcription PCR amplification of ten selected genes from bitter gourd. M is the 1 Kb DNA marker, followed by the RT-PCR product from Squalene synthase (lane 1), Hydroxymethylglutaryl-CoA synthase (lane 2), Acetyl-CoA acetyltransferase (lane 3), Methyltransferase-like protein 13 (lane 4), Phytoene synthase (lane 5), 3-hydroxy-3-methylglutaryl-coenzyme A reductase (lane 6), Chaperone protein dnaJ 10 (lane 7), Isopentenyl phosphate kinase (lane 8), Myb-related protein 306 (lane 9), and Late embryogenesis abundant protein M17 (lane 10).

yet to be cloned. Analysis of gene completeness showed 94% complete BUSCOs (Supplementary Fig. S1), which indicated that the transcript assembly is of good quality.

**Validation of transcript assembly.** The Assembly of the transcripts was validated by performing reverse transcription PCR (RT-PCR) amplification of selected genes using the RNA from the same tissues. We selected ten transcripts ranging in size between 609 bp and 2,567 bp, and their details are given in Supplementary Table S1. In agarose gel analysis of the RT-PCR products, we observed the amplification of cDNA fragments of approximately the same size as expected based on the transcript assembly (Fig. 4). This result indicated that the transcripts assembled in this study could be reliably used for gene identification.

**Validation of gene expression.** The expression of selected transcripts associated with terpenoid and steroid bio-synthesis pathways was validated using qRT-PCR. The genes chosen for this study included Phytoene synthase, Acetyl-CoA transferase, 3-Hydroxy-3-methylglutaryl-coenzyme A reductase 1, Phosphate kinase, and Hydroxymethylglutaryl-CoA synthase. The FPKM values for these genes are given in Supplementary Table S5. Expression of these genes, as assessed from the cycle threshold (ct) values, largely correlated with the expression levels in terms of FPKM (Fig. 5).

**Identification of simple sequence repeats (SSRs).** We identified 57,093 SSRs from 125,566 assembled transcripts. The majority of the SSRs were either di-nucleotide repeats (59.0%) or tri-nucleotide repeats (37.4%).



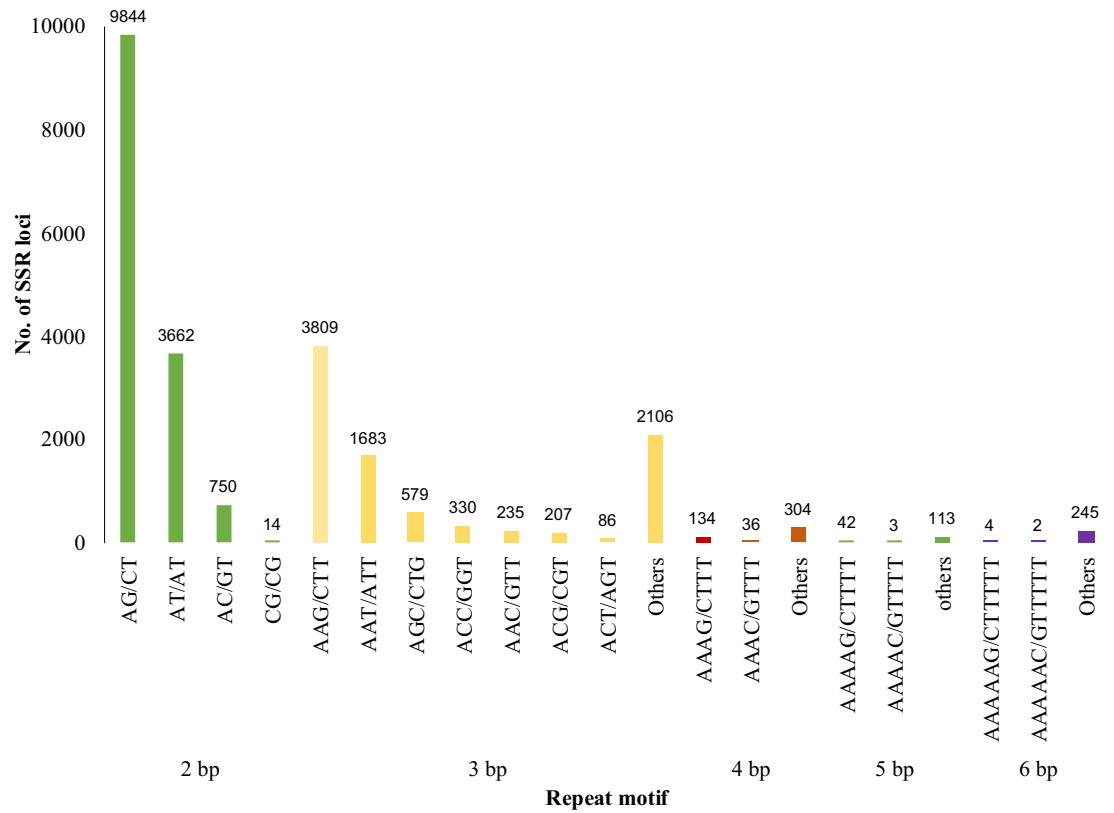
**Figure 5.** Quantitative reverse transcription PCR (qRT-PCR) analysis of Serine/Threonine-protein phosphatase (PP2A), Phytoene synthase (PS), 3-Hydroxy 3-Methylglutaryl CoA reductase (HMGR), Hydroxymethyl glutaryl CoA synthase (HMGS), Acetyl CoA acetyltransferase (AAT), and Phosphate kinase (PK).

Motif length	Repeats							Total	%
	5	6	7	8	9	10	>10		
Di	0	5374	2785	1855	1104	698	2454	14,270	59.00
Tri	4617	2145	1001	686	108	140	338	9035	37.35
Tetra	328	109	17	17	3	0	0	474	1.96
Penta	127	27	1	1	0	2	0	158	0.65
Hexa	143	71	15	11	6	5		251	1.04
Total	5215	7726	3819	2570	1221	845	2792	24,188	100
%	21.56	31.94	15.79	10.63	5.05	3.49	11.54		

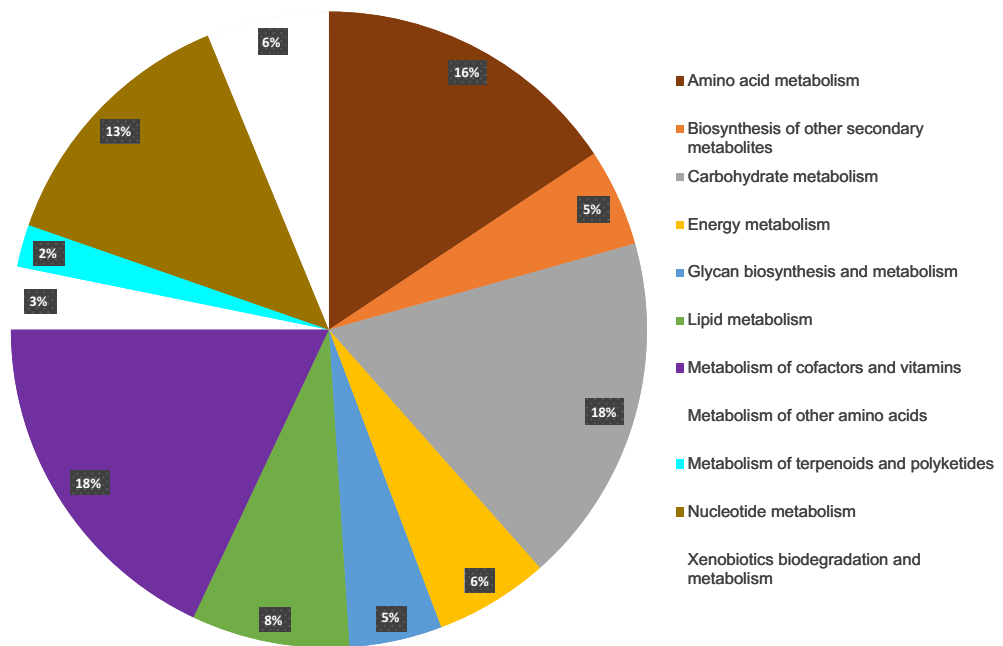
**Table 3.** Type of SSR repeats and their frequency identified from bitter gourd fruit transcriptome.

Di-nucleotide repeats with six repeats and tri-nucleotide repeats with five repeats were found to be the most abundant ones. Only a small proportion of the SSRs belonged to the tetra-nucleotide (2.0%), hexa-nucleotide (1.0%), and penta-nucleotide repeats (0.65%) (Table 3). AG/CT was the most abundant type (68.98%) in terms of the actual sequence present in the repeats. AAG/CTT was the most abundant tri-nucleotide repeat (42.16%). The details of different types of repeats identified from the transcripts are given in Fig. 6. SSR marker-based study of genetic diversity in 211 accessions of bitter gourd showed clear segregation according to geographical origin<sup>33</sup>. Alhariri et al.<sup>34</sup> have analyzed 51 accessions using 61 primers and clustered the accession based on the polymorphism in SSR markers. Surprisingly, the 51 accessions formed three groups according to fruit size. These reports indicate that the SSR markers identified in this study will be highly useful for varietal identifications, mapping, and association studies in bitter gourd.

**Identification of transcripts related to secondary metabolites.** In KEGG pathway analysis, we mapped 25,497 transcripts to 147 pathways. About half of the mapped transcripts (51.5%) were related to the metabolism of cofactors and vitamins (18%), carbohydrates (17.8%), amino acids (15.7%) (Fig. 7). Antidiabetic properties of bitter gourd fruits have been demonstrated in several in vitro, in vivo, and clinical studies<sup>8,35,36</sup>. Most of the compounds responsible for its bitter taste and antidiabetic properties are triterpenoids and steroids<sup>37</sup>. The carbon skeletons of cucurbitacin triterpenoids are derived from 2,3-oxidosqualene by the action of oxidosqualene cyclase. Although cucurbitacins accumulated in high amounts in the fruits, cucurbitadienol synthase gene was highly expressed in the leaves. This indicated that cucurbitacins might be synthesized in the leaves and transported to the fruits<sup>38</sup>. Charantin also accumulated in high amounts in the fruits compared to the root, stem, leaves, and flowers<sup>39</sup>. The same study identified fifteen genes related to triterpenoid biosynthesis from the RNASeq library constructed from the bitter gourd seedlings. However, most of them encoded for the enzymes involved only up to the synthesis of squalene and 2,3-oxidosqualene. We identified 191 and 99 transcripts related to the terpenoid backbone and steroid biosynthesis from the bitter gourd fruit transcriptome. In our study also, most of the transcripts were mapped up to the synthesis of squalene and squalene-2,3-epoxide

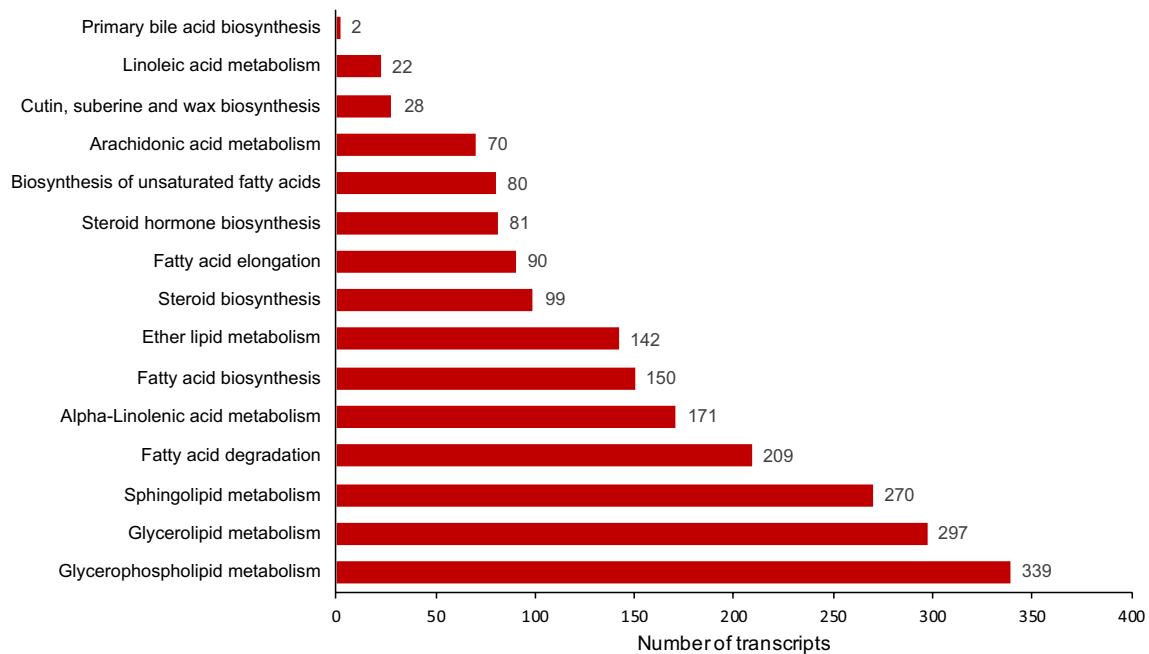


**Figure 6.** Frequency of different types of SSR motifs identified from bitter gourd fruit transcriptome.



**Figure 7.** KEGG pathway analysis of the bitter gourd fruit transcriptome.

(Supplementary Figures S2 & S3). We could not identify the genes involved in the synthesis of cucurbitacin and charantin. Together with previous reports, our results indicate that mature leaves may be a good source for cloning the genes related to the biosynthesis of the specific triterpenoid and steroid compounds from bitter gourd.



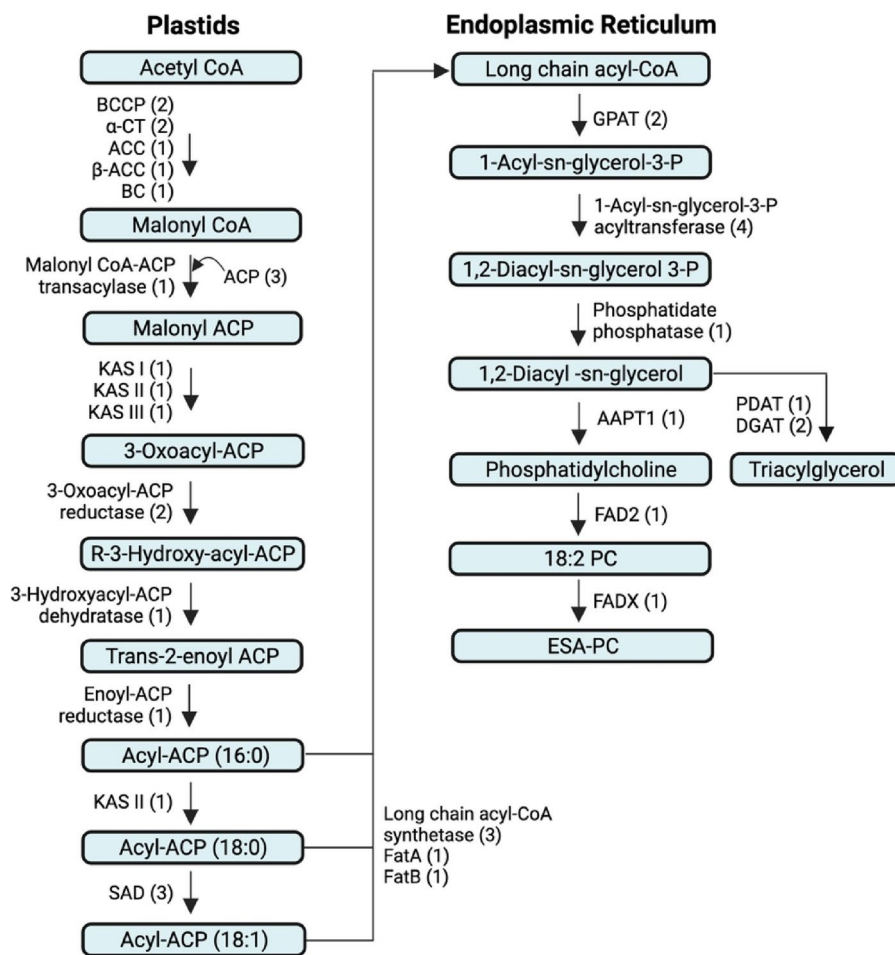
**Figure 8.** Details of the number of transcripts involved in lipid metabolism in the bitter gourd fruit transcriptome.

**Identification of full-length genes related to bitter gourd seed oil biosynthesis.** We mapped 2,050 transcripts to lipid metabolism from the assembled transcripts, including 15 pathways related to the biosynthesis and metabolism of various fatty acids and lipids (Fig. 8). Among them, 150, 80, 297, and 339 transcripts were related to the biosynthesis of fatty acids, unsaturated fatty acids, glycerolipids, and glycerophospho-lipids. Plants synthesize various fatty acids and lipids, which have medicinal, cosmetic, and industrial applications<sup>40–42</sup>. Fatty acid biosynthesis occurs in plastids using acetyl Co-A as the major precursor and involves fatty acid synthases and carrier proteins. Fatty acids are synthesized up to 16:0 long-chain acyl carrier protein (ACP) during this process. The 16:0 fatty acids are elongated to 18:0 ACP and desaturated to 18:1 ACP by ketoacyl synthase II (KAS II) and stearoyl-ACP desaturase, respectively. All the long-chain ACPs from the chloroplast contribute to the long-chain acyl-CoA pool, which is used to synthesize oils and lipids in the endoplasmic reticulum. Fatty acids are incorporated into glycerol-3P to obtain diacylglycerol (DAG). The DAG pool is used to make triacylglycerol (oil) or conjugated fatty acids through distinct pathways. A schematic diagram representing the biochemical pathways of fatty acids, conjugated fatty acids, and triacylglycerol is shown in Fig. 9.

Yang et al. 2010 studied the transcriptome of bitter gourd seeds by 454 sequencing and identified several transcripts related to oil biosynthesis genes. However, full-length genes were analyzed only for diacylglycerol acyltransferase 1 and 2<sup>43</sup>. In our study, we identified 2,050 transcripts related to lipid metabolism. The transcripts related to fatty acid biosynthesis, glycerolipid metabolism, and glycerophospholipid metabolism were analyzed to determine full-length genes coding for the enzymes and proteins involved in the biosynthesis of fatty acids conjugated fatty acids, and triacylglycerol. We identified 39 full-length genes (GenBank Accession No: ON175892 to ON175930), which code for 23 enzymes and two carrier proteins required for the biosynthesis of fatty acids, conjugated fatty acids, and triacylglycerol (Fig. 9, Table 4). This included ten full-length genes coding for the five enzymes needed for converting long-chain acyl-CoA to triacylglycerol. The highest number of four genes were identified for 1-acyl-sn-glycerol-3-phosphate acyltransferase (AGPAT). We also found two genes for diacylglycerol acyltransferase (DGAT) and one for phospholipid: diacylglycerol acyltransferase (PDAT), which are essential to convert diacyl glycerol pool to triacylglycerol. These oil biosynthesis genes can be used to increase the oil content through genetic engineering approaches. This is evident from the studies, which showed that overexpression of heterologous genes could achieve as high as a ten-fold increase in seed oil content<sup>44,45</sup>.

Bitter gourd seed oil's medicinal and industrial applications are attributed to the conjugated fatty acid, mainly  $\alpha$ -eleostearic acid. Plants with a negligible amount of conjugated fatty acids can also be metabolically engineered to produce a higher amount of these fatty acids by transferring the relevant genes. In Soybean and Arabidopsis, conjugated fatty acid content increased by 15 to 20% when delta-9 and delta-12 conjugase genes were expressed under seed-specific promoters<sup>46</sup>. In Arabidopsis, the accumulation of conjugated fatty acids was further improved by expressing both fatty acid desaturase and delta-9 and delta-12 conjugase<sup>47</sup>. Developing bitter gourd and other plant varieties with a high amount of conjugated fatty acids will require identifying the genes necessary to divert diacylglycerol from triacylglycerol synthesis. Three enzymes, CDP-choline diacyl glycerol choline phosphotransferase (AAPT), fatty acid desaturase (FAD), and fatty acid conjugase (FADX), are required to synthesize  $\alpha$ -eleostearic acid from diacylglycerol<sup>48</sup>. Several transcripts related to the biosynthesis of conjugated fatty acids were identified by analyzing the normalized cDNA library of seed tissues<sup>43</sup>. However, the





**Figure 9.** Schematic representation of the biochemical pathway for synthesizing fatty acids, conjugated fatty acids, and triacylglycerol. The numbers given in the bracket indicate the number of full-length genes identified for the respective enzymes and carrier proteins from the bitter melon fruit transcriptome. *BCCP* biotin carboxyl carrier protein of acetyl-CoA carboxylase; *α-CT* acetyl-CoA carboxylase carboxyl transferase subunit alpha; *ACC* acetyl-CoA carboxylase; *β-ACC* acetyl-CoA carboxylase beta subunit; *BC* biotin carboxylase; *ACP* acyl carrier protein; *KAS I* 3-oxoacyl-ACP synthase; *KAS II* 3-oxoacyl-ACP synthase II; *KAS III* 3-oxoacyl-ACP synthase III; *SAD* stearoyl-ACP-desaturase; *FatA* oleoyl-ACP thioesterase; *FatB* acyl-ACP thioesterase ATL3; *GPAT* glycerol-3-phosphate acyltransferase; *AAPT* phosphatidylcholine:diacylglycerol cholinephosphotransferase; *FAD2* delta(12)-fatty acid desaturase 2; *FADX* delta(12)-fatty acid conjugase; *PDAT* phospholipid:diacylglycerol acyltransferase; *DGAT* diacylglycerol O-acyltransferase; *PC* phosphatidylcholine; *ESA* eleostearic acid.

identification of full-length genes was limited to diacylglycerol acyltransferase 1 and 2, which do not directly affect the production of conjugated fatty acid. In the current study, we identified full-length genes for all three enzymes, all of which were present as single-copy genes (Fig. 9, Table 4). These genes will help to improve the conjugated fatty acid content in bitter melon and other high-yielding oilseed crops like brassica, sunflower, groundnut, and soybean. Though this has been demonstrated in *Arabidopsis* by co-expressing *FADX* and *DGAT* genes from tung<sup>49</sup>, further research in important oilseed crops is required to achieve the scale of production needed for medicinal and industrial applications.

## Materials and methods

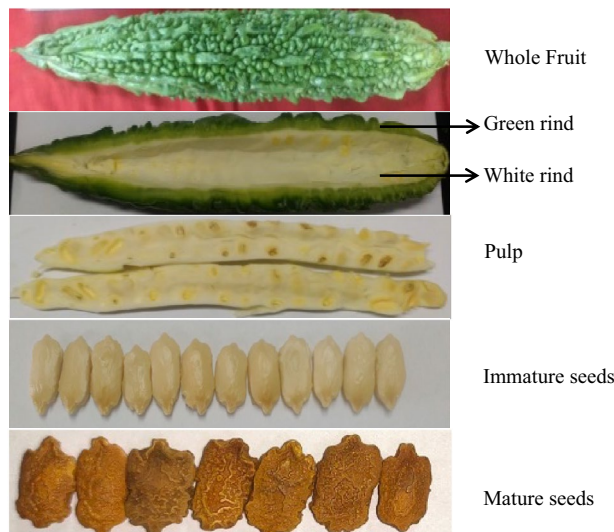
**Plant material.** Seeds of *Momordica charantia* L. variety Co1 were obtained from Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India. It is a cultivated variety and its fruits are usually long and dark green weighing upto 100–120 g. It was released for cultivation in 1978, and can yield up to 14 tonnes/ha. The plants were raised from the seeds in a greenhouse (23–25 °C temperature, 16 h light, and 50% relative humidity). Green unripe fruits were cleaned with sterile distilled water, and dissected to separate green rind, white rind, pulp, and immature seeds. Mature seeds were collected from ripe fruits (Fig. 10). The separated tissues were flash-frozen under liquid nitrogen before storing in a freezer (-80 °C) until use. The plant studies were carried out in accordance with relevant institutional, national, and international guidelines and legislation.



S. no	Name of the gene	mRNA (nt)	CDS (bp)	ORF (aa)	5'UTR (bp)	3'UTR (bp)	Genbank accession no
1	Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP), chloroplast	1484	867	288	398	219	ON175892
2	Biotin carboxyl carrier protein of acetyl-CoA carboxylase 1, (BCCP) chloroplast	1419	852	283	119	448	ON175893
3	Acetyl-CoA carboxylase carboxyl transferase subunit alpha ( $\alpha$ -CT), chloroplast isoform 1	2641	2205	734	277	159	ON175894
4	Acetyl-CoA carboxylase carboxyl transferase subunit alpha ( $\alpha$ -CT), chloroplast isoform 2	3508	2367	788	350	791	ON175895
5	Acetyl-CoA carboxylase 1 (ACC)	8184	6798	2265	957	429	ON175896
6	Acetyl-CoA carboxylase beta subunit ( $\beta$ -ACC)	3965	1578	525	216	2171	ON175897
7	Biotin carboxylase 1 (BC 1), chloroplast	2070	1590	529	143	337	ON175898
8	Malonyl CoA-ACP transacylase (MCAT)	2000	1137	378	449	414	ON175899
9	Acyl carrier protein (ACP)	803	462	153	168	173	ON175900
10	Acyl carrier protein 1 (ACP 1), chloroplast	765	420	139	121	224	ON175901
11	Acyl carrier protein 4 (ACP 4), chloroplast	1614	432	143	305	877	ON175902
12	3-oxoacyl-ACP synthase I (KAS I), chloroplast	2172	1413	470	331	428	ON175903
13	3-oxoacyl-ACP synthase II (KAS II), chloroplast isoform 1	2623	1695	564	402	526	ON175904
14	3-oxoacyl-ACP synthase III (KAS III), chloroplast	1917	1215	404	457	245	ON175905
15	3-oxoacyl-ACP reductase 1 (FabG)	2151	867	288	312	972	ON175906
16	3-oxoacyl-ACP reductase 2 (FabG)	1672	909	302	485	278	ON175907
17	3-hydroxyacyl-ACP dehydratase (FabZ)	1157	675	224	265	217	ON175908
18	Enoyl-ACP reductase, chloroplast (FabI)	1720	1176	391	252	292	ON175909
19	Stearoyl-ACP 9-desaturase 1 (SAD1), chloroplast	1882	1191	396	225	466	ON175910
20	Stearoyl-ACP 9-desaturase 5 (SAD5), chloroplast	1956	1170	389	206	580	ON175911
21	Stearoyl-ACP 9-desaturase 6 (SAD6), chloroplast	1425	1197	398	56	172	ON175912
22	Long chain acyl-CoA synthetase 1 (LACS1)	2854	2016	671	307	531	ON175913
23	Long chain acyl-CoA synthetase 2 (LACS2)	2570	1986	661	349	235	ON175914
24	Long chain acyl-CoA synthetase 9, (LACS9) chloroplast	2969	2085	694	606	278	ON175915
25	Oleoyl-ACP thioesterase 1, chloroplast (FatA)	3544	1119	372	884	1541	ON175916
26	Acyl-ACP thioesterase ATL3, chloroplast (FatB)	1070	615	204	112	343	ON175917
27	Glycerol-3-phosphate acyltransferase 3 (GPAT)	1949	1128	375	339	482	ON175918
28	Glycerol-3-phosphate acyltransferase, chloroplast (GPAT)	1823	1353	450	270	200	ON175919
29	1-acyl-sn-glycerol-3-phosphate acyltransferase (AGPAT)	2033	897	298	2	1134	ON175920
30	1-acyl-sn-glycerol-3-phosphate acyltransferase 1, chloroplast (AGPAT1)	1888	1083	360	147	658	ON175921
31	1-acyl-sn-glycerol-3-phosphate acyltransferase 2 (AGPAT2)	1788	1149	382	297	342	ON175922
32	1-acyl-sn-glycerol-3-phosphate acyltransferase 3 (AGPAT3)	1807	1143	380	200	464	ON175923
33	Phosphatidate phosphatase 1 (PAP1)	3757	2676	891	614	467	ON175924
34	Phosphatidylcholine:diacylglycerol cholinephosphotransferase 1 (AAP1)	1402	813	270	239	350	ON175925
35	Delta(12)-fatty acid desaturase 2 (FAD2)	1827	1155	384	259	413	ON175926
36	Delta(12)-fatty acid conjugase (FADX)	1674	1200	399	133	341	ON175927
37	Phospholipid:diacylglycerol acyltransferase 1 (PDAT1)	2934	2004	667	558	372	ON175928
38	Diacylglycerol O-acyltransferase 1A (DGAT1A)	2191	1596	531	320	275	ON175929
39	Diacylglycerol O-acyltransferase 2 (DGAT2)	2625	966	321	199	1460	ON175930

**Table 4.** Details of the full-length genes identified from the bitter melon fruit transcriptome.

**RNA isolation.** RNA isolation from the different samples of *M.charantia* fruits was done using the Trizol reagent (Invitrogen). Frozen tissue sample of 100 mg was ground to fine powder under liquid nitrogen, 1 ml of Trizol was added, and the tissue suspension was transferred to a 1.5 ml centrifuge tube. Chloroform (0.2 volume) was added, mixed vigorously, and centrifuged at 12,000 rpm for 15 min at 4 °C. The aqueous phase was transferred to a fresh 1.5 ml centrifuge tube, and 0.7 volume of isopropanol was added. The content was mixed gently, centrifuged at 12,000 rpm for 10 min at 4 °C, and the supernatant was discarded. The pellet was washed by adding 500  $\mu$ l of 70% ethanol, and centrifuging at 10,000 rpm for 5 min at room temperature. The supernatant was discarded, pellet was air-dried, and dissolved in nuclease-free water. The total RNA was checked in the gel, and treated DNase I (Qiagen, Germany) followed by column purification using RNeasy MinElute Cleanup Kit (Qiagen, Germany). The purified RNA was assessed using a spectrophotometer (Eppendorf, Germany), Qubit 2.0 fluorimeter (Invitrogen, USA), and Bioanalyzer (Agilent Technologies, USA). The RNA integrity number



**Figure 10.** The five different tissues of bitter melon fruits used for the transcriptome sequencing.

(RIN value) of RNA samples ranged from 7.3 to 9.3. The RNA samples with RIN values above eight were used for complementary DNA synthesis.

**Library preparation and sequencing.** The sequencing library was prepared using one  $\mu\text{g}$  total RNA as a template using TruSeq mRNA Library Preparation Kit V2 (Illumina Inc, USA). Poly(A)-RNA was purified from total RNA using oligo-dT magnetic beads and fragmented before cDNA synthesis. The fragmented cDNAs were used as template for cDNA synthesis using SuperScript II reverse transcriptase (Invitrogen, USA). The first strand cDNA was converted to double strand cDNA and purified using AMPure XP beads. The purified cDNAs were end-repaired to generate blunt-ended cDNAs and single 'A' nucleotide tail was added to the 3' ends. These 'A' tails were complementary to the 'T' overhang in the adapters. Adapters were ligated to the cDNA ends for indexing and size selected for 420 bp fragments using AMPure XP beads. The size-selected cDNAs were enriched by PCR. The PCR-amplified library was purified, and analyzed in Bioanalyzer 2100 (Agilent Technologies, USA) using High Sensitivity DNA Chips to determine the insert size and yield. The DNA yield was normalized to 10 nM, and pooled for cluster generation in the flow cell. The libraries were subjected to  $2 \times 150$  bp paired-end sequencing in the Illumina platform (Illumina, USA).

**De novo assembly and clustering.** FastQC v0.11.2<sup>27</sup> was used for the quality analysis of the reads. Removal of adaptor sequences and trimming was performed using Cut adapt v1.7.1<sup>50</sup> and Sickle v1.33<sup>51</sup>. The high-quality (Q30>) paired-end reads were assembled using the Trinity assembler with three modules<sup>52</sup>. Inchworm assembled unique transcripts using k-mer values. Chrysalis clustered related contigs and constructed de Bruijn graphs. Butterfly reconstructed full-length transcripts and distinct transcripts for splice isoforms. The redundancy among the assembled transcripts was reduced by clustering using CD-HIT v4.0<sup>53</sup>.

**Annotation of the transcripts.** To perform functional annotation, de novo assembled transcripts of bitter melon were analyzed using BLASTX algorithm plant non-redundant protein database at National Center for Biotechnology Information (NCBI). The annotations of the best hits were saved and subjected to gene ontology (GO) analysis using Blast2GO. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used, and the transcripts were mapped to different pathways<sup>54,55</sup>.

**Transcript quantification.** The abundance of the transcripts in the bitter melon fruit tissues was measured using RNA-Seq by expectation maximization (RSEM) tool. In RSEM, the RNA-Seq reads were aligned to the reference transcripts to estimate the transcript abundance<sup>56</sup>. The fragments per kilobase per million (FPKM) and transcripts per million (TPM) were calculated to estimate the transcript abundance.

**Identification of simple sequence repeats.** Simple sequence repeats (SSRs) from bitter melon were identified using the MISA tool<sup>57</sup>. The search parameters identified different repeat types and number repeat for each kind.

**Assessment of gene completeness.** Completeness of the transcripts was estimated using Benchmarking Universal Single Copy Orthologs (BUSCO) analysis against a database of single-copy orthologous genes for plants (eudicotyledons\_odb10). The BUSCO analysis identifies the conserved orthologs and helps to assess the completeness of transcripts<sup>58</sup>.

**Validation of transcript assembly.** Ten assembled transcripts were selected at random and validated for the correctness of de novo assembly by reverse transcription PCR (RT-PCR) amplification of the transcripts. Transcript-specific primers were designed based on the sequence of the de novo assembled transcripts (Supplementary Table S1). cDNA was synthesized using oligo-dT primers, and individual transcripts were amplified using the respective transcript-specific primers. PCR amplification included initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 95 °C for 30 s, annealing (55–59 °C) for 30 s, and extension at 72 °C for 2 min, final extension at 72 °C for 5 min, and storage at 4 °C. The amplified products were analyzed by agarose gel electrophoresis.

**Analysis of gene expression.** Five transcripts were selected at random and analyzed for gene expression. Transcript-specific primers suitable for qRT-PCR were designed based on the sequence of the de novo assembled transcripts. Analysis of gene expression was performed using TB Green Premix Ex Taq II (Takara, Japan) and QuantStudio 5 (ThermoFisher Scientific, USA). As reported before, the serine/threonine phosphatase gene was used as an internal reference<sup>59</sup>. A negative control reaction without a template was included in all experiments done. The sequences of the primers designed for qRT-PCR analysis are given in Supplementary Table S2. Gene expression was analyzed using the individual efficiency corrected calculation method<sup>60</sup>.

## Conclusions

In the current study, we investigated the fruit transcriptome of bitter melon, focusing on identifying full-length genes involved in seed oil biosynthesis. We identified 39 full-length genes coding for 23 enzymes and two carrier proteins involved in the biosynthesis of fatty acids, conjugated fatty acids, and triacylglycerol. We also identified several transcripts related to the biosynthesis of the compounds responsible for the bitterness and antidiabetic properties. The tissue-specific abundance of the transcripts reported here will form the basis for identifying tissue-specific promoters. The full-length genes, gene expression data, and SSR markers reported from this study are vital genomic resources for further research.

## Data availability

Datasets used in this study were deposited in the SRA (<https://www.ncbi.nlm.nih.gov/sra>) under the SRA accession numbers SRR18210038 (immature seeds), SRR18210039 (pulp), SRR18210040 (mature seeds), SRR18210041 (white rind), and SRR18210042 (green rind). The full length oil biosynthesis genes were deposited in NCBI GenBank (Accession No: ON175892–ON175930).

Received: 13 May 2022; Accepted: 1 September 2022

Published online: 13 September 2022

## References

1. Basch, E., Gabardi, S. & Ulbricht, C. Bitter melon (*Momordica charantia*): A review of efficacy and safety. *Am. J. Health Syst. Pharm.* **60**, 356–359 (2003).
2. Bharathi, L. K. *et al.* Cytotaxonomical analysis of *Momordica* L. (Cucurbitaceae) species of Indian occurrence. *J. Genet.* **90**, 21–30 (2011).
3. Jeffrey, C. A review of the *Cucurbitaceae*. *Bot. J. Linn. Soc.* **81**, 233–247 (1980).
4. Bortolotti, M., Mercatelli, D. & Polito, L. *Momordica charantia*, a nutraceutical approach for inflammatory related diseases. *Front. Pharmacol.* **10**, 486 (2019).
5. May, L.S., Sanip, Z., Shokri, A.A., Abdul Kadir, A. & Lazin, M.R. The effects of *Momordica charantia* (bitter melon) supplementation in patients with primary knee osteoarthritis: A single-blinded, randomized controlled trial. *Complement. Ther. Clin. Pract.* **32**, 181–186 (2018).
6. Poolperm, S. & Jiraungkoorskul, W. An update review on the anthelmintic activity of bitter melon, *Momordica charantia*. *Pharmacogn. Rev.* **11**, 31–34 (2017).
7. Fuangchan, A., Sonthisombat, P., Seubnukarn, T., Chanouan, R., Chotchaisuwat, P., Sirigulsatien, V., Ingkaninan, K., Plianbangchang, P., Haines, S.T. Hypoglycemic effect of bitter melon compared with metformin in newly diagnosed type 2 diabetes patients. *J. Ethnopharmacol.* **24**, 134(2), 422–428 (2011).
8. Krawinkel, M. B. *et al.* Bitter melon reduces elevated fasting plasma glucose levels in an intervention study among prediabetics in Tanzania. *J. Ethnopharmacol.* **24**(216), 1–7 (2018).
9. Kumari, S., Dash, I. & Behera, K. K. Therapeutic effect of *Momordica charantia* on blood glucose, lipid profile and oxidative stress in type 2 diabetes mellitus patients: A randomised controlled trial. *J. Clin. Diagn. Res.* **12**, 21–25 (2018).
10. Nyam, K. L., Tan, C. P., Lai, O. M., Long, K. & Man, Y. B. C. Physicochemical properties and bioactive compounds of selected seed oils. *LWT Food Sci Technol.* **42**, 1396–1403 (2009).
11. Lee, S. T., Radu, S., Ariffin, A. & Ghazali, H. M. Physico-chemical characterization of oils extracted from noni, spinach, lady's finger, bitter melon, mustard seeds, and copra. *Int. J. Food Prop.* **18**, 2508–2527 (2014).
12. Yoshime, L.T., de Melo, I.L.P., Sattler, J.A.G., de Carvalho, E.B.D. & Mancini-Filho, J. Bitter melon (*Momordica charantia* L.) seed oil as a naturally rich source of bioactive compounds for nutraceutical purposes. *Nutrire.* **41**, 12 (2016).
13. Cahoon, E. B. *et al.* Biosynthetic origin of conjugated double bonds: production of fatty acid components of high-value drying oils in transgenic soybean embryos. *Proc. Natl. Acad. Sci. USA* **96**, 12935–12940 (1999).
14. Dahiya, D. K. & Puniya, A. K. Conjugated linoleic acid enriched skim milk prepared with *Lactobacillus fermentum* DDH127 endorsed antiobesity in mice. *Future Microbiol.* **13**, 1007–1020 (2018).
15. Koronowicz, A. A. & Banks, P. Antitumor properties of CLA-enriched food products. *Nutr. Cancer.* **70**, 529–545 (2018).
16. Wang, X. & Wu, J. Modulating effect of fatty acids and sterols on skin aging. *J. Funct. Foods.* **57**, 135–140 (2019).
17. Dhar Dubey, K. K., Sharma, G. & Kumar, A. Conjugated linolenic acids: implication in cancer. *J. Agric. Food Chem.* **67**, 6091–6101 (2019).
18. Abu Bakar, N., Hakim Abdullah, M. N., Lim, V. & Yong, Y. K. Essential oils derived from *Momordica charantia* seeds exhibited anticancer activity against hydrogen chloride/ethanol and indomethacin. *Evid. Based Complement Altern. Med.* **22**, 5525584 (2021).
19. Chen, G. C. *et al.* A conjugated fatty acid present at high levels in bitter melon seed favorably affects lipid metabolism in hepatocytes by increasing NAD(+)/NADH ratio and activating PPAR $\alpha$ , AMPK and SIRT1 signaling pathway. *J. Nutr. Biochem.* **33**, 28–35 (2016).

20. Chen, G. C., Chen, W. H., Tseng, K. T., & Chao, P. M. The anti-adiposity effect of bitter melon seed oil is solely attributed to its fatty acid components. *Lipids Health Dis.* **29**;16(1):186 (2017).
21. Ahmad, Z., Zamhuri, K. F., Yaacob, A., Siong, C. H., Selvarajah, M., Ismail, A. & Nazrul Hakim, M. In vitro anti-diabetic activities and chemical analysis of polypeptide-k and oil isolated from seeds of *Momordica charantia* (bitter gourd). *Molecules.* **17**, 9631–9640 (2012).
22. Paul, D., Dey, T. K., Mukherjee, S., Ghosh, M. & Dhar, P. Comparative prophylactic effects of  $\alpha$ -eleostearic acid rich nano and conventional emulsions in induced diabetic rats. *J. Food Sci. Technol.* **51**, 1724–1736 (2014).
23. Yasui, Y. *et al.* Bitter gourd seed fatty acid rich in 9c,11t,13t-conjugated linolenic acid induces apoptosis and up-regulates the GADD45, p53 and PPAR gamma in human colon cancer Caco-2 cells. *Prostaglandins Leukot Essent Fatty Acids.* **73**, 113–119 (2005).
24. Kobori, M., Ohnishi-Kameyama, M., Akimoto, Y., Yukizaki, C. & Yoshida, M. Alpha-eleostearic acid and its dihydroxy derivative are major apoptosis-inducing components of bitter gourd. *J. Agric. Food Chem.* **56**, 10515–10520 (2008).
25. Zhuo, R. J. *et al.* A-eleostearic acid inhibits growth and induces apoptosis in breast cancer cells via HER2/HER3 signaling. *Mol. Med. Rep.* **9**, 993–998 (2014).
26. Beatty, A., Singh, T., Tyurina, Y. Y., Tyurin, V. A., Samovich, S., Nicolas, E., Maslar, K., Zhou, Y., Cai, K. Q., Tan, Y., Doll, S., Conrad, M., Subramanian, A., Bayir, H., Kagan, V. E., Rennefahrt, U. & Peterson, J. R. Ferroptotic cell death triggered by conjugated linolenic acids is mediated by ACSL1. *Nat Commun.* **14**(12), 2244 (2021).
27. Ranasinghe, K. N. K. *et al.* In vivo anticancer effects of *Momordica charantia* seed fat on hepatocellular carcinoma in a rat model. *J. Ayurveda Integr. Med.* **12**, 435–442 (2021).
28. Ozsolak, F. & Milos, P.M. RNA sequencing: Advances, challenges and opportunities. *Nat. Rev. Genet.* **12**, 87–98 (2011).
29. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
30. Chase, M. W. *et al.* An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **161**, 105–121 (2009).
31. Poovitha, S. & Parani, M. In vitro and in vivo  $\alpha$ -amylase and  $\alpha$ -glucosidase inhibiting activities of the protein extracts from two varieties of bitter gourd (*Momordica charantia* L.). *BMC Complement Altern. Med.* **18**(16), 185 (2016).
32. Poovitha, S. & Parani, M. Protein extracts from *Momordica charantia* var. *charantia* and *M. charantia* var. *muricata* show anti-lipidemic and antioxidant properties in experimental type 2 diabetic rats. *J. Food Biochem.* **44**(9), e13370 (2020).
33. Cui, J. *et al.* Genome-wide analysis of simple sequence repeats in bitter gourd (*Momordica charantia*). *Front Plant Sci.* **22**, 1103 (2017).
34. Alhariri, A., Behera, T. K., Jat, G. S., Devi, M. B., Boopalakrishnan, G., Hemeda, N. F., Teleb, A. A., Ismail, E. & Elkordy, A. Analysis of genetic diversity and population structure in bitter gourd (*Momordica charantia* L.) using morphological and SSR markers. *Plants (Basel).* **10**(9), 1860 (2021).
35. Chang, C. I., Cheng, S. Y., Nurlatifah, A. O., Sung, W. W., Tu, J. H., Lee, L. L. & Cheng, H. L. Bitter melon extract yields multiple effects on intestinal epithelial cells and likely contributes to anti-diabetic functions. *Int. J. Med. Sci.* **18**(8), 1848–1856 (2021).
36. Mahmoud, M. F., El Ashry, F. E., El Maraghy, N. N. & Fahmy, A. Studies on the antidiabetic activities of *Momordica charantia* fruit juice in streptozotocin-induced diabetic rats. *Pharm. Biol.* **55**(1), 758–765 (2017).
37. Jia, S., Shen, M., Zhang, F. & Xie, J. Recent advances in *momordica charantia*: functional components and biological activities. *Int J Mol Sci.* **18**(12), 2555 (2017).
38. Takase, S. *et al.* Identification of triterpene biosynthetic genes from *Momordica charantia* using RNA-seq analysis. *Biosci. Biotechnol. Biochem.* **83**, 251–261 (2019).
39. Cuong, D. M. *et al.* Accumulation of charantin and expression of triterpenoid biosynthesis genes in bitter melon (*Momordica charantia*). *J. Agric. Food Chem.* **23**(65), 7240–7249 (2017).
40. Bach, L. & Faure, J. D. Role of very-long-chain fatty acids in plant development, when chain length does matter. *C R Biol.* **333**(4), 361–370 (2010).
41. He, M. & Ding, N. Z. Plant unsaturated fatty acids: Multiple roles in stress response. *Front Plant. Sci.* **11**, 562785 (2020).
42. Harwood, J. L., Woodfield, H. K., Chen, G. & Weselake, R. J. Modification of oil crops to produce fatty acids for industrial applications. In *Fatty Acids: Chemistry, Synthesis and Applications* (ed. Ahmad, M. U.) 187–236 (Elsevier, 2017).
43. Yang, P., Li, X., Shipp, M. J., Shockey, J. M. & Cahoon, E. B. Mining the bitter melon (*Momordica Charantia* L.) seed transcriptome by 454 analysis of non-normalized and normalized cDNA populations for conjugated fatty acid metabolism-related genes. *BMC Plant Biol.* **10**, 250 (2010).
44. Alameldin, H., Izadi-Darbandi, A., Smith, S. A., Balan, V., Jones, A. D. & Sticklen, M. Production of seed-like storage lipids and increase in oil bodies in corn (Maize; *Zea mays* L.) vegetative biomass. *Ind. Crops Prod.* **108**, 526–534 (2017).
45. Savadi, S., Lambani, N., Kashyap, P. L. & Bisht, D. S. Genetic engineering approaches to enhance oil content in oilseed crops. *Plant Growth Regul.* **83**, 207–222 (2017).
46. Cahoon, E. B. *et al.* Conjugated fatty acids accumulate to high levels in phospholipids of metabolically engineered soybean and *Arabidopsis* seeds. *Phytochemistry* **67**(12), 1166–1176 (2006).
47. Mietkiewska, E. *et al.* Combined transgenic expression of *Punica granatum* conjugase (FADX) and FAD2 desaturase in high linoleic acid *Arabidopsis thaliana* mutant leads to increased accumulation of punicic acid. *Planta* **240**, 575–583 (2014).
48. Dyer, J. M., Chapital, D. C., Kuan, J. C., Mullen, R. T., Turner, C., McKeon, T. A. & Pepperman, A. B. Molecular analysis of a bifunctional fatty acid conjugase/desaturase from tung. Implications for the evolution of plant fatty acid diversity. *Plant Physiol.* **130**(4), 2027–2038 (2002).
49. Yurchenko, O. *et al.* Engineering the production of conjugated fatty acids in *Arabidopsis thaliana* leaves. *Plant Biotechnol. J.* **15**(8), 1010–1023 (2017).
50. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
51. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (2011).
52. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
53. Li, W. & Godzik, A. CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
54. Conesa, A. *et al.* Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).
55. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
56. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* **12**, 323 (2011).
57. Beier, S., Thiel, T., Munch, T., Scholz, U. & Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics.* **33**(16), 2583–2585 (2017).
58. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* **31**(19), 3210–3212 (2015).
59. Wang, Z. *et al.* Selection and validation of appropriate reference genes for real-time quantitative PCR analysis in *Momordica charantia*. *Phytochemistry* **164**, 1–11 (2019).
60. Rao, X., Huang, X., Zhou, Z. & Lin, X. An improvement of the 2<sup>-</sup>(-delta delta CT) method for quantitative real-time polymerase chain reaction data analysis. *Biostat Bioinforma Biomath.* **3**(3), 71–85 (2013).

61. Andrews, S. FastQC: A quality control tool for high throughput sequence data (2010).

### Author contributions

M.P. designed the experiments and guided the data analysis and reviewed the manuscript; K.R. collected the sample, performed the experiment and analysed the data; K.R. prepared the figures and prepared the manuscript.

### Funding

We received the grant from the SRM-DBT Partnership Platform for Contemporary Re-search Services and Skill Development in Advanced Life Sciences Technologies.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-19686-4>.

**Correspondence** and requests for materials should be addressed to M.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022