Bioinformation Volume 5

open access

www.bioinformation.net Issue 3

Hypothesis

Analysis of beta-carotene hydroxylase gene cDNA isolated from the American oil-palm (*Elaeis oleifera*) mesocarp tissue cDNA library

Subhash J. Bhore[1, 2*], Amelia Kassim[1], Chye Ying Loh[2], Farida H. Shah[1]

[1]Molecular Biology Division, Melaka Institute of Biotechnology, Lot 7, Melaka International Trade Centre City, 75450 Ayer Keroh, Melaka, Malaysia; [2]Department of Biotechnology, Faculty of Applied Sciences, AIMST University, Bedong-Semeling Road, Bedong, 08100, Kedah Darul Aman, Malaysia; Subhash J Bhore- Email: subhashbhore@gmail.com; Phone: +60-4-429 8176; Fax: +60-4-429 8109, *Corresponding author

Received August 19, 2010; Accepted August 26, 2010; Published September 20, 2010

Abstract

It is well known that the nutritional quality of the American oil-palm (*Elaeis oleifera*) mesocarp oil is superior to that of African oil-palm (*Elaeis guineensis* Jacq. Tenera) mesocarp oil. Therefore, it is of important to identify the genetic features for its superior value. This could be achieved through the genome sequencing of the oil-palm. However, the genome sequence is not available in the public domain due to commercial secrecy. Hence, we constructed a cDNA library and generated expressed sequence tags (3,205) from the mesocarp tissue of the American oil-palm. We continued to annotate each of these cDNAs after submitting to GenBank/DDBJ/EMBL. A rough analysis turned our attention to the beta-carotene hydroxylase (*Chyb*) enzyme encoding cDNA. Then, we completed the full sequencing of cDNA clone for its both strands using M13 forward and reverse primers. The full nucleotide and protein sequence was further analyzed and annotated using various Bioinformatics tools. The analysis results showed the presence of fatty acid hydroxylase superfamily domain in the protein sequence. The multiple sequence alignment of selected *Chyb* amino acid sequences from other plant species and algal members with *E. oleifera Chyb* using ClustalW and its phylogenetic analysis suggest that *Chyb* from monocotyledonous plant species, *Lilium hubrid*, *Crocus sativus* and *Zea mays* are the most evolutionary related with *E. oleifera Chyb*. This study reports the annotation of *E. oleifera Chyb*.

Keywords: African oil-palm, American oil-palm, fatty acids, fatty acid hydroxylase, oleic acid, sterol desaturase, zeaxanthin

Abbreviations: ESTs, expressed sequence tags; *EoChyb, Elaeis oleifera* beta-carotene hydroxylase; MC, main cluster

Background:

Oil-palm is the major commodity worldwide. The second largest source of fats and oils to the world market of fats and oils is from *Elaeis guineensis* Jacq Tenera, which is commonly known as African oil-palm. Malaysia is the main producer and exporter of the palm oil. African oil-palm is cultivated on industrial scale due to high palm oil yield derived from its fruit mesocarp tissue. African oil-palm, *E. guineensis* Jacq have three different forms (also called as varieties), namely, 'Pisifera', 'Dura', and 'Tenera'. These three forms are distinguished on the basis of shell thickness of fruits. The commercially cultivated African oil-palm is a hybrid from 'Dura' (♀) and 'Pisifera' (♂) **[1]**. Another oil-palm species which is economically less important is *Elaeis oleifera*. This species is also called as American oil-palm.

Palmitic acid, ($C_{16:0}$) is the predominant fatty acid in palm oil derived from *E. guineensis* Jacq. Tenera fruit mesocarp tissue. Whereas oil derived from *E. oleifera* fruit mesocarp tissue is predominant with oleic acid ($C_{18:1}$) (68.6%), a fatty acid good for health **[2]**. The $C_{16:0}$ is the major (44%) saturated fatty acid in palm oil derived from *E. guineensis.* However, the $C_{16:0}$ content in *E. oleifera* is only 25%. In spite of the high content of the healthy $C_{18:1}$ in the *E. oleifera* oil, it is not preferred for commercial plantation due to its poor oil yield.

The single pass, partial sequencing of randomly isolated anonymous cDNA clones also called as ESTs has become a rapid and cost-effective means in gaining information about gene expression and their regulation **[3]**. In addition, generated ESTs data is useful in new and novel gene's discovery **[4]**, evaluation of the genome for gene content and its structure, and for *in silico* comparative expression analysis between different plant tissues **[5]**. Most importantly, ESTs serves as a valuable resource for high-throughput expression analysis using cDNA-microarray technology **[6]**. In higher plants, numerous genes have been identified by random nucleotide sequencing of cDNA clones **[7-11]**. Therefore, in order to study the gene expression and their patterns in *E. oleifera* fruit mesocarp tissue, ESTs generation project was initiated. So far, 3,205 ESTs are generated from 17 weeks old *E. oleifera* mesocarp tissue cDNA library (our unpublished work). Beta carotene hydroxylase (*Chyb*) is one of the isolated cDNA (ESTs) clones.

By understating potential applications of *Chyb* in genetic engineering of oil-palm and or in other plants, clone was fully sequenced. The *Chyb* is involved in zeaxanthin biosynthesis by hydroxylating beta-carotene, but the enzyme may be involved in other pathways **[12-13]**. The products of this enzyme are zeaxanthin and beta-cryptoxanthin **[14]**. Beta-carotene, zeaxanthin and beta-cryptoxanthin are categorized under carotenoids which have many industrial applications as food and feed additives, and are used in cosmetics and as nutraceutical. Currently, vitamin A deficiency is a global health burden which could be alleviated through provitamin A carotenoid biofortification in suitable plants **[15]**. If we understand in depth the regulation of carotenoid biosynthesis, enhancement of beta-carotene could be done by limiting beta-carotene hydroxylation. Therefore in order to understand more about *EoChyb*, its cDNA clone is analyzed using computational tools. The nucleotide and protein sequence of *EoChyb* cDNA is analyzed and annotated in this study to find out their features. The *EoChyb* annotation and features are reported in this paper.

Bioinformation Volume 5
open access
www.bioinformation.net Issue 3 Hypothesis

Methodology:

cDNA library and *EoChyb* clone isolation

Elaeis oleifera, seventeen week's old [weeks after anthesis (WAA)] mesocarp tissue cDNA library constructed using the 'CloneMiner cDNA library construction kit' (Invitrogen Corporation) for ESTs generation (our unpublished work) was used to isolate *EoChyb*. The *EoChyb* cDNA clone was isolated by random method of cDNA isolation **[3]**.

The cDNA nucleotide sequence reported in this paper has been submitted to GenBank/DDBJ/EMBL under accession number EU057623.

Nucleotide sequencing

The *E. coli* DH5α cells harboring *EoChyb* cDNA were cultivated in 10 ml LB medium (supplemented with Kanamycin) overnight in dark at 37 ºC, 160 rpm. The cultivated *E. coli* DH5α cells were harvested from broth and plasmid DNA was isolated by using a commercial kit, Wizard® *Plus* SV Minipreps DNA purification system (Promega). Sequencing reactions were carried out for both strands using M13 (Forward); 5'-GTAAAACGACGGCCAG-3' and M13 (Reverse); 5'-GGATAACAATTTCACACAGG-3' primers.

cDNA and protein sequence analysis

The cDNA sequence was edited manually to eliminate vector and adaptor sequences from 5' and 3' ends. Finalized cDNA sequence analysis was performed using online free bioinformatics tools. The similarity searches were performed using blast programs (BlastN, BlastX, and BlastP) against the databases available at NCBI **[16]**. EMBOSS pairwise alignment algorithm [http://www.ebi.ac.uk/Tools/emboss/align/] was used to compare 2 cDNA sequences to find out homology %. To find out general features of the cDNA sequence including amino acid composition and isoelectric point analysis was performed online using bioinformatics tools available at JustBio **[17]**. Guanine and cytosine content (GC %) calculation was carried out by using 'DNA/RNA base composition calculator', a free online bioinformatics tool that calculates the molecular mass, elemental composition, base composition, and percent AT and GC content for cDNA/DNA/RNA sequences **[18]**. Alignment of amino acid sequences and dendrogram construction was carried out using multiple sequence alignment by ClustalW **[19]** program. Whereas, to find out fully conserved residues; residues with conserved strong groups and residues with conserved weak groups in *EoChyb*, amino acid sequences were aligned by using clustal 2.0.11 multiple sequence alignment program.

Results:

***EoChyb* clone isolation**

By random method of cDNA clone isolation, *EoChyb* clone was isolated from 17 day old mesocarp tissue cDNA library. The serial number of this randomly isolated cDNA clone in ESTs generation project was 2962; hence clone identity 'EoEST-2962' was given to *EoChyb* cDNA clone.

Nucleotide sequencing

The +ve and –ve strand of *EoChyb* cDNA clone were sequenced using M13 forward and M13 reverse primers, respectively. The *EoChyb* +ve and –ve strand cDNA sequence after removal of vector and adaptor sequence was compared and searched for overlaps using blast (bl2seq) program **[16]**. The results produced shows that *EoChyb* cDNA sequence is 1414 bp in length. By analyzing the finalized cDNA sequence and its deduced amino acid sequence, *Chyb* identity was given to the cDNA clone.

cDNA and protein sequence analysis

The general features of the cDNA nucleotide and protein sequence as revealed by bioinformatics tools available at JustBio are summarized in **Table 1 (see supplementary material)**. The annotated nucleotide sequence was submitted to GenBank/DDBJ/EMBL under accession number EU057623. Whereas the comparison of the *EoChyb* at nucleotide and protein level with *Chyb* from other plant species and algal members is depicted in **Table 2 (see supplementary material)**. The nucleotide and deduced amino acid sequences of *EoChyb* cDNA, an open reading frame (ORF), 5' and 3' untranslated regions (UTR), initiation and termination codon, and the amino acid residues of beta-carotene hydroxylase and sterol desaturase conserved domains are shown in **Figure 1**. The *EoChyb*, and *Chyb* from other selected plants and algal members *Chyb* amino acid sequence alignment based phylogenetic analysis was carried out and the constructed dendrogram is shown in **Figure 2**. Clustal 2.0.11 multiple sequence alignment program produced results of *EoChyb* amino acid sequence alignment with amino acid sequences of *Chyb* from other organisms and showed single fully conserved residues, residues with conserved strong groups and residues with conserved weak groups **(Figure 3)**.

Discussion:

The availability of *EoChyb* cDNA clone in 17 day old mesocarp tissue cDNA library indicates that *Chyb* is expressed in developing *E. oleifera* fruit mesocarp tissue. However, the level of its expression, pattern of expression and tissue specificity is not known. The analysis of developing fruit mesocarp by Northern blot technique could shade the light on level and pattern of its expression **[20]**. The GC content in *EoChyb* cDNA is 56%. This GC % is close to the predicted GC content in coding sequences of *Elaeis oleifera* and *Elaeis guineensis* **[21-22]**. Homology analysis using BlastP indicates that *EoChyb* protein shows 82-98% homology with *Chyb* from other monocot plant species. Comparison of *EoChyb* with *Chyb* from 13 dicot plant species shows range of homology from 71-98%. However, *Chyb* from algal member shows only 78-84% homology with *EoChyb* **(Table 2, see supplementary material)**. The relatively low level of *EoChyb* homology with *Chyb* from algal member is in line with the evolution in plant species. Monocots are highly evolved in comparison with the dicots and algae. Conserved domain search in *EoChyb* indicates that amino acid residue 46-324 are part of the beta-carotene hydroxylase (fatty acid hydroxylase super family) conserved domain; and amino acid residues 184-309 belongs to the sterol desaturase [lipid metabolism] domain **(Figure 1)**. However, latter's role in sterol desaturation is not clear **[23]**. The *EoChyb* amino acid composition analysis revealed that it is rich in alanine (A) amino acid **(**See **Figure 1** and **Table 3** in Supplementary material**)**.

The dendrogram constructed to study the phylogenetic relationship between *EoChyb* and *Chyb* protein sequences from 13 dicots, 3 monocots and 3 algal members shows 3 MC. The *Chyb* proteins from dicots, monocots and algal members were precisely grouped in MC1, MC2, and MC3 respectively **(Figure 2)**. All the *Chyb* proteins from dicotyledonous plants in MC1 were sub-grouped into 3 sub-clusters. The sub-cluster-1 is of *Chyb* from *Citrus sinensis* and *Glycine max*. The sub-cluster-2 is of *Chyb* from *Coffea arabica*, *Solanum lycopersicum*, *Gentiana lutea*, *Diospyros kaki* and *Capsicum annuum*. Whereas, *Vitis vinifera*, *Adonis aestivalis*, *Chrysanthemum x morifolium*, *Brassica napus*, *Arabidopsis thaliana* and *Daucus carota* were grouped under sub-cluster-3 of MC1. All the *Chyb* protein sequences from monocots, *Crocus sativus*, *Elaeis oleifera*, *Lilium hybrid* and *Zea mays* used in phylogenetic analysis were grouped under MC2; and three algal *Chyb* protein sequences from *Chlamydomonas reinhardtii*, *Muriella zofingiensis* and *Haematococcus pluvialis* were grouped under MC3 **(Figure 2)**. This classification of *Chyb* is in line with the phylogenetic lineages that have molecular data in the NCBI databases **[24]**. The dendrogram analysis also reflects that *EoChyb* is phylogenetically closer to the *Chyb* from *Lilium hybrid* and other monocots used in the study.

The multiple sequence alignment of *Chyb* amino acid sequences of *EoChyb* and *Chyb* protein sequences from 13 dicots, 3 monocots and 3 algal members **(Table 2 see supplementary material)** clearly shows the well conserved catalytic residues in *Chyb*. The analysis of the multiple sequence alignment of *Chyb* amino acid sequences clearly shows that in total 51 amino acid residues are fully conserved in *Chyb* of the all plant species and algal members **(Figure 3)**. It was also evident that across the *Chyb* amino acid sequence, in total 28 amino acid residues were conserved strongly **(Figure 3)**.

Bioinformation Volume 5

open access

www.bioinformation.net Issue 3

Hypothesis

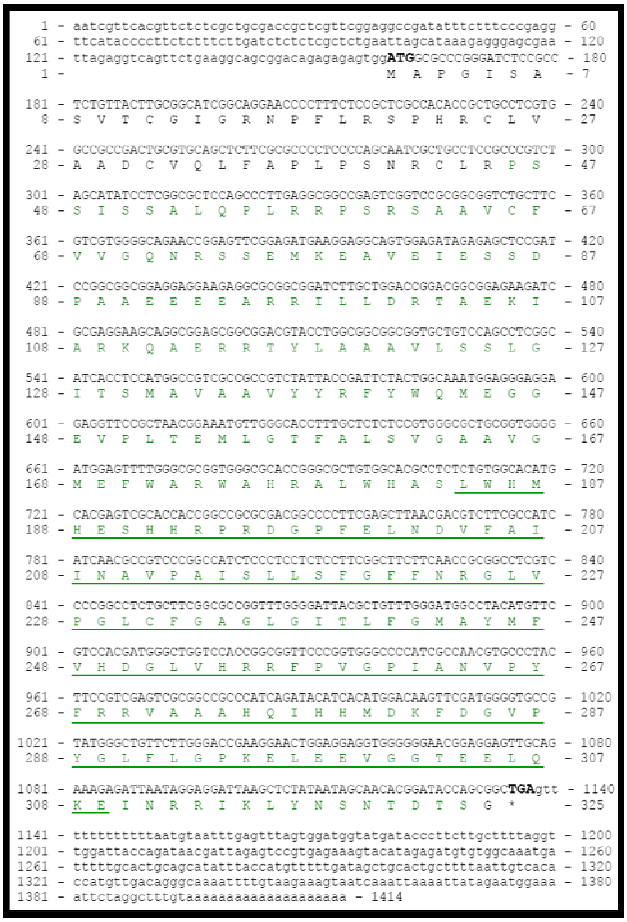

Figure 1: Nucleotide and deduced amino acid sequences of *E. oleifera* beta-carotene hydroxylase cDNA clone. An open reading frame and noncoding regions are shown in capital and small letters, respectively. The deduced amino acid sequence is given below the nucleotide sequence, which is numbered at the both ends of each sequence line. The open reading frame encodes for a protein of 325 amino acid residues. Amino acid residues are numbered beginning with the initial methionine till last glycine (G) residue. Initiation and termination codons are shown in bold. The beta-carotene hydroxylase (fatty acid hydroxylase super family) conserved domain (PLN02601) **[25]** residues are shown in green colour. The sterol desaturase [Lipid metabolism] domain (ERG3) **[26]** residues are underlined. *represent the termination codon. This cDNA was isolated by random method of gene isolation from *E. oleifera* 17 week old mesocarp tissue cDNA library.

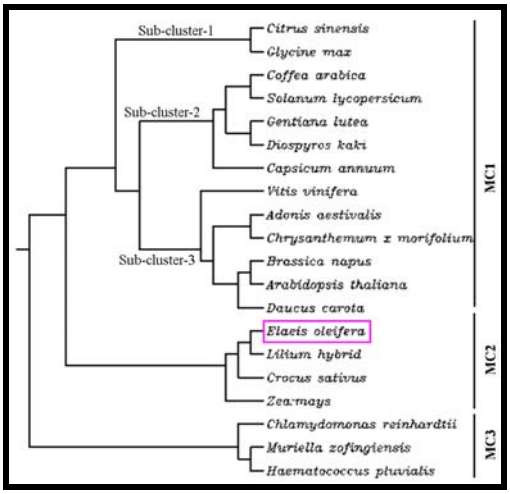

Figure 2: Rooted dendrogram showing clustering of beta-carotene hydroxylase (*Chyb*) from *E. oleifera* and other organisms. Amino acid sequences for different organisms were obtained from NCBI database. Alignment of amino acid sequences and dendrogram construction was carried out using multiple sequence alignment by ClustalW **[19]** program using default parameters. Location of *E. oleifera Chyb* in phylogenetic tree is shown in pink box. The ID of *Chyb* proteins used in the study is given in **Table 2 (see supplementary material)**. MC stands for main cluster.

Bioinformation Volume 5

open access

www.bioinformation.net Issue 3

Hypothesis

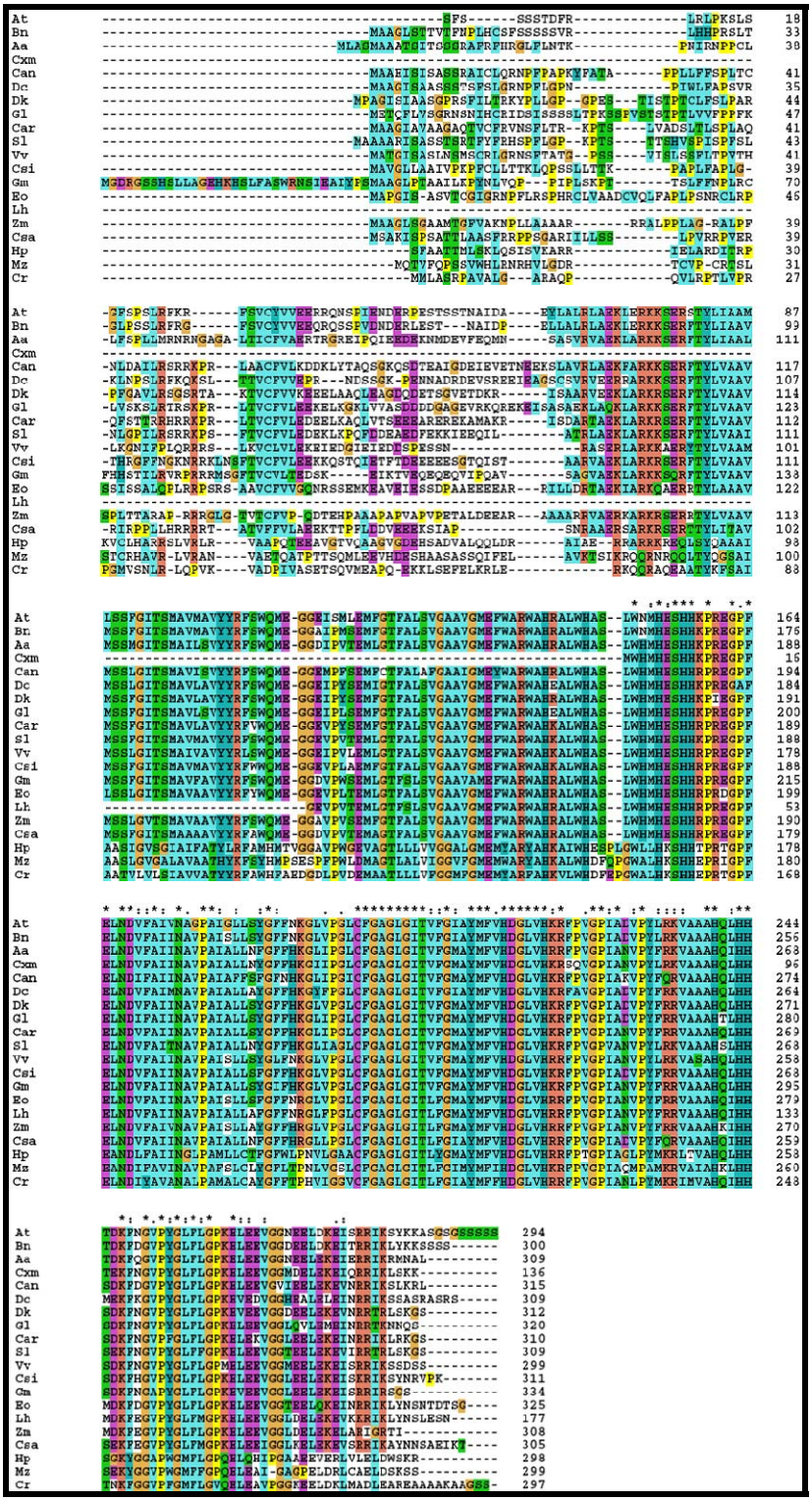

Figure 3: Similarity comparison of amino acid sequences of the *E. oleifera* beta-carotene hydroxylase (*Chyb*) protein and *Chyb* amino acid sequences from other organisms. Amino acid sequences are numbered at the end of each sequence row. (*), (:) and (.) denote single fully conserved residues, residues with conserved strong groups and residues with conserved weak groups in *Chyb*, respectively. Aa, *Adonis aestivalis*; At, *Arabidopsis thaliana*; Bn, *Brassica napus*; Can, *Capsicum annuum*; Car, *Coffea arabica*; Cr, *Chlamydomonas reinhardtii*; Csi, *Citrus sinensis*; Csa, *Crocus sativus*; Cxm, *Chrysanthemum* x *morifolium* hybrid; Dc, *Daucus carota*; Dk, *Diospyros kaki*; Eo, *Elaeis oleifera*; Gl, *Gentiana lutea*; Gm, *Glycine max*; Hp, *Haematococcus pluvialis*; Lh, *Lilium hybrid*; Mz, *Muriella zofingiensis*; Sl, *Solanum lycopersicum*; Vv, *Vitis vinifera*; Zm, *Zea mays.* This alignment is produced by clustal 2.0.11 multiple sequence alignment program.

Conclusion:
This study has annotated the salient features of randomly isolated *EoChyb* cDNA clone using Bioinformatics tools. Bioinformatics analyses revealed that *EoChyb* protein is carrying conserved domains for the beta-carotene hydroxylase and the sterol desaturase. Furthermore, the study also shows the 51 fully conserved amino acid residues in *Chyb* from the flowering (monocot and dicot) plants and algal members, which could be of relevance to gain insights into the evolution of *Chyb*.

Acknowledgements
Authors are grateful to the Ministry of Science, Technology and Innovation (MOSTI) of Malaysian Government for research funding [Grant Code: IRPA: 01-02-02-0014/PR0015/07-07], and to the United Plantation Berhad, Perak, Malaysia for supplying fruit samples of *Elaeis oleifera* for this study.

References:
[1] F Muniran *et al*. *Indian J Exp Biol*. **46**: 79 (2008) [PMID: 18697576]
[2] O Asemota & FH Shah, *Afr J Biotechnol*. **11**: 595 (2004)
[3] BM Lange *et al*. *PNAS*. **97**:2934-2939 (2000) [PMID: 10717007]
[4] J Ohlrogge, C Benning, *Curr Opin Plant Biol.* **3**: 224 (2000) [PMID: 10837264]
[5] CM Ronning *et al*. *Plant Physiol*. **131**:419 (2003) [PMID: 12586867]
[6] N Sreenivasulu *et al*. *Plant J*. **37**:539 (2004) [PMID: 14756762]
[7] J Liu *et al*. *Plant Mol Biol*. **29**: 685 (1995) [PMID: 8541495]
[8] CS Keith *et al*. *Plant Physiol* **101**: 329 (1993) [PMID: 8278499]
[9] YS Park *et al*. *Plant Physiol* **103**: 359 (1993) [PMID: 8029332]
[10] H Hofte *et al*. *Plant J*. **4**: 1051 (1993) [PMID: 8281187]
[11] M Umeda & H Uchimiya, *Plant Physiol* **106**:1015(1994) [PMID: 12232382]
[12] A Marchler-Bauer *et al*. *Nucleic Acids Res*. **37**: (D)205 (2009) [PMID: 18984618]
[13] JF Martin at al. *Microb Cell Fact*. **7**: 3 (2008) [PMID: 18289382]
[14] I Baroli *et al*. *Plant Cell*. **15**: 992 (2003) [PMID: 12671093]
[15] R Vallabhaneni *et al*. *Plant Physiol* **151**: 1635 (2009) [PMID: 19767386]
[16] http://www.ncbi.nlm.nih.gov/
[17] http://www.justbio.com/
[18] http://www.currentprotocols.com/tools/dnarna-base-composition-calculator
[19] http://align.genome.jp/
[20] FH Shah, TS Cha, *Plant Sci.* **154**: 153 (2000) [PMID: 10729614]
[21] http://www.kazusa.or.jp/codon/cgi-bin/spsearch.cgi?species=Elaeis+&c=s
[22] R Singh *et al*. *BMC Plant Biol*. **9**:114 (2009) [PMID: 19706196]
[23] BA Arthington, *Gene*. **102**: 39 (1991) [PMID: 1864507]
[24] http://www.ncbi.nlm.nih.gov/taxonomy
[25] http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?ascbin=8&maxaln=10& seltype=2&uid=166242
[26] http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?ascbin=8&maxaln=10&seltype=2&uid=32818

Edited by P. Kangueane
Citation: Bhore *et al.* Bioinformation 5(3): 104-112 (2010)

Bioinformation Volume 5

open access

www.bioinformation.net Issue 3 **Hypothesis**

Supplementary material:

Supplementary Figure 1: Individual amino acids (%) in the *E. oleifera* beta-carotene hydroxylase (*Chyb*) protein. On X axis, the amino acids are represented by a single letter code. The amino acid composition was calculated by ProtCalc program **[17]**.

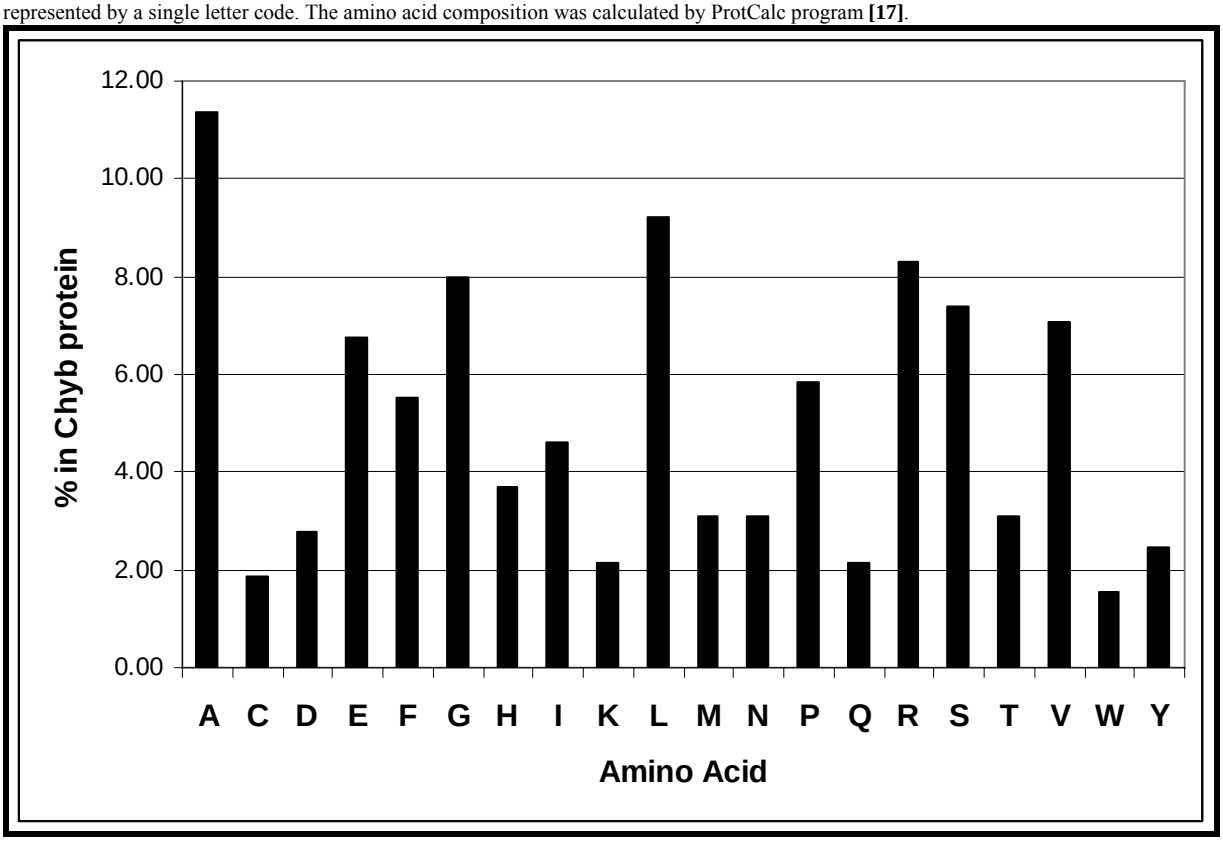

Table 1: The general features of *Elaeis oleifera* beta-carotene hydroxylase (*EoChyb*) cDNA and protein sequence

General Features	
cDNA sequence	
Size, bp	1414
Molecular Weight, Daltons	437125
5' UTR, bp	159
Coding Sequence	978
3' UTR, bp	277
Stop Codon	TGA (UGA)
G+C content, %	56
Protein sequence	
Length, amino acids	325
Molecular weight (Dalton)	35940.42
Isoelectric point	8.54

Table 2: Comparison of the *Elaeis oleifera* beta-carotene hydroxylase (*Chyb*) with *Chyb* from monocots, dicots and algal members

Species	GenBank Accession No.	Length compared		Homology (%)	
		nucleotide (bp)	amino acid	nucleotide level	amino acid level
Monocots					
Lilium hybrid	AB445122	722	177	85%	97%
Crocus sativus	AJ416711	918	305	81%	82%
Zea mays	AY844956	1392	308	87%	98%
Dicots					
Daucus carota	DQ192193	1231	309	71%	93%
Arabidopsis thaliana	U58919	956	294	73%	85%
Brassica napus	EF026098	903	300	52%	71%
Chrysanthemum x morifolium	AB205042	1227	136	74%	97%
Adonis aestivalis	EF120636	1187	309	74%	82%
Vitis vinifera	AF499108	900	299	78%	98%
Coffea arabica	DQ157169	933	310	75%	92%
Solanum lycopersicum	Y14809	1125	309	73%	87%
Gentiana lutea	AB027187	1519	320	73%	82%
Diospyros kaki	FJ790215	1365	312	79%	90%
Capsicuum annuum	Y09225	1112	315	71%	96%
Citrus sinensis	DQ228870	936	311	74%	90%
Glycine max	AY575953	1164	334	74%	96%
Algal members					
Haematococcus pluvialis	DQ257289	906	298	34.8%	81%
Muriella zofingiensis	EU016205	1545	299	67%	78%
Chlamydomonas reinhardtii	XM_001698646	1757	297	72%	84%

Bioinformation Volume 5

open access

www.bioinformation.net Issue 3 **Hypothesis**

Table 3: Amino acid composition of *E. oleifera* beta-carotene hydroxylase (*Chyb*) and in *Chyb* from other plant species and algal members

Amino acid (single letter code)	Source of *Chyb* and Their Amino Acid Composition (%)¶																		
	Dicotyledonous plants													Monocotyledonous plants				Algal Members	
	Csi	Gm	Car	Sl	Gl	Dk	Can	Vv	Aa	Cxm	Bn	At	Dc	Eo	Lh	Csa	Zm	Cr	Mz	Hp
Alanine(A)	9.97	9.88	11.61	10.03	8.12	10.58	11.43	8.36	10.36	7.35	8.67	8.50	11.33	11.38	9.60	12.13	16.56	14.14	11.04	13.76
Cysteine(C)	0.96	0.90	0.97	0.65	0.94	0.96	1.59	1.00	0.97	0.74	1.00	0.68	0.97	1.85	0.56	0.33	0.65	0.67	2.01	1.01
Aspartic acid (D)	1.61	1.80	1.94	1.94	2.81	2.56	2.86	2.34	1.94	2.21	3.33	2.72	2.27	2.77	2.26	1.97	1.95	3.03	2.68	2.35
Glutamic acid (E)	7.72	6.29	7.42	7.77	6.88	7.69	7.30	7.36	7.44	7.35	6.67	7.48	7.44	6.77	7.91	7.54	6.17	6.40	4.68	5.03
Phenylalanine (F)	8.04	6.29	7.10	7.77	6.56	6.73	7.94	4.68	7.12	6.62	6.67	6.80	6.80	5.54	8.47	6.56	5.84	6.40	5.35	3.69
Glycine (G)	8.36	8.08	7.74	7.77	7.81	9.62	6.67	8.70	8.09	11.03	8.00	8.84	8.41	8.00	10.73	7.21	9.09	8.08	8.03	10.07
Histidine(H)	4.50	5.09	4.19	4.53	4.06	3.85	3.81	4.01	4.21	7.35	4.33	3.40	4.21	3.69	6.21	4.26	4.22	4.04	5.69	4.70
Isoleucine(I)	3.54	4.79	3.55	3.88	4.38	4.17	6.03	5.35	5.83	6.62	4.00	4.08	4.21	4.62	3.95	5.57	2.27	3.37	4.35	4.70
Lysine(K)	5.79	3.89	5.16	5.83	5.94	4.17	5.40	3.68	3.88	7.35	4.33	5.10	4.53	2.15	3.39	3.28	2.27	4.38	3.34	3.02
Leucine(L)	9.65	8.68	9.35	8.74	9.38	8.33	8.57	10.70	9.06	9.56	10.00	9.18	7.44	9.23	10.17	8.20	9.09	9.76	9.03	11.41
Methionine(M)	3.22	3.29	3.23	3.24	3.12	2.88	3.17	4.35	4.85	3.68	3.33	3.06	3.24	3.08	3.95	2.95	3.57	4.71	3.68	3.36
Asparagine(N)	1.93	2.10	1.94	1.94	2.50	1.28	2.22	3.01	4.21	3.68	2.33	3.06	2.91	3.08	3.39	1.97	1.3	1.68	2.01	1.01
Proline(P)	5.47	5.99	4.19	5.50	5.00	5.77	5.40	4.68	4.86	5.88	5.00	4.42	4.85	5.85	5.65	6.23	7.14	6.40	5.69	4.70
Glutamine(Q)	1.93	2.40	1.94	0.97	1.88	1.28	1.90	1.00	1.62	2.21	1.33	1.02	0.97	2.15	0.56	0.98	0.65	3.37	4.35	3.36
Arginine(R)	5.14	6.29	7.74	5.50	5.31	6.09	6.03	6.02	7.12	3.68	6.67	6.12	6.47	8.31	5.65	9.18	9.09	5.05	5.02	7.38
Serine(S)	5.47	8.38	5.81	7.77	9.38	7.37	6.67	10.37	5.50	2.21	9.67	11.56	9.06	7.38	3.39	6.56	3.9	3.37	7.36	4.36
Threonine(T)	4.82	3.29	4.19	4.85	4.06	4.17	2.86	3.01	3.56	1.47	3.67	2.72	2.27	3.08	1.69	4.59	3.9	3.37	4.68	4.70
Valine(V)	7.72	7.78	8.06	7.12	8.12	8.01	5.71	7.02	5.83	7.35	7.33	6.12	7.77	7.08	7.91	6.56	8.44	8.08	6.69	6.71
Tryptophan(W)	1.93	2.10	1.61	1.62	1.56	1.60	1.59	1.67	1.62	0.74	1.67	1.70	1.94	1.54	2.26	1.64	1.62	1.01	2.01	2.01
Tyrosine(Y)	2.25	2.69	2.26	2.59	2.19	2.88	2.86	2.68	1.94	2.94	3.00	3.40	2.91	2.46	2.26	2.30	2.27	2.69	2.34	2.68

Bioinformation Volume 5

open access

www.bioinformation.net Issue 3 **Hypothesis**

¶Source of *Chyb*: Aa, *Adonis aestivalis*; At, *Arabidopsis thaliana*; Bn, *Brassica napus*; Can, *Capsicum annuum*; Car, *Coffea arabica*; Cr, *Chlamydomonas reinhardtii*; Csi, *Citrus sinensis*; Csa, *Crocus sativus*; Cxm, *Chrysanthemum* x *morifolium* hybrid; Dc, *Daucus carota*; Dk, *Diospyros kaki*; Eo, *Elaeis oleifera*; Gl, *Gentiana lutea*; Gm, *Glycine max*; Hp, *Haematococcus pluvialis*; Lh, *Lilium hybrid*; Mz, *Muriella zofingiensis*; Sl, *Solanum lycopersicum*; Vv, *Vitis vinifera*; Zm, *Zea mays*. The amino acid composition was calculated by ProtCalc program **[17]**.