Supplementary Material

Defining and enhancing the biosynthesis of astaxanthin and Docosahexaenoic acid in *Aurantiochytrium* sp. SK4

Jingrun Ye[a, b], Mengmeng Liu[a, b], Mingxia He[a], Ying Ye[a,b], Junchao Huang[a]*

[a]Key Laboratory of Economic Plants and Biotechnology, Yunnan Key Laboratory for Wild Plant Resources, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201

[b]University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China

***Corresponding author**

Email: huangjc@mail.kib.ac.cn.

Phone: +86-875-65228058

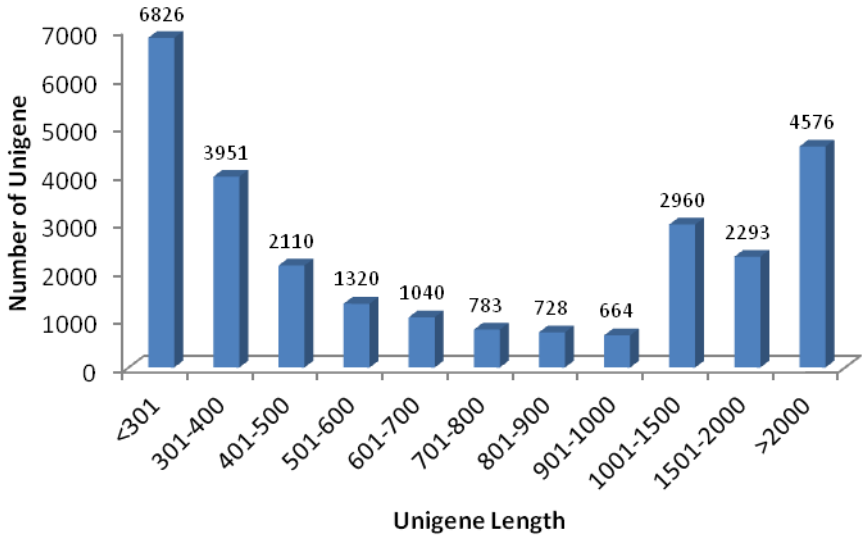

Figure S1. Uni-gene length distribution.

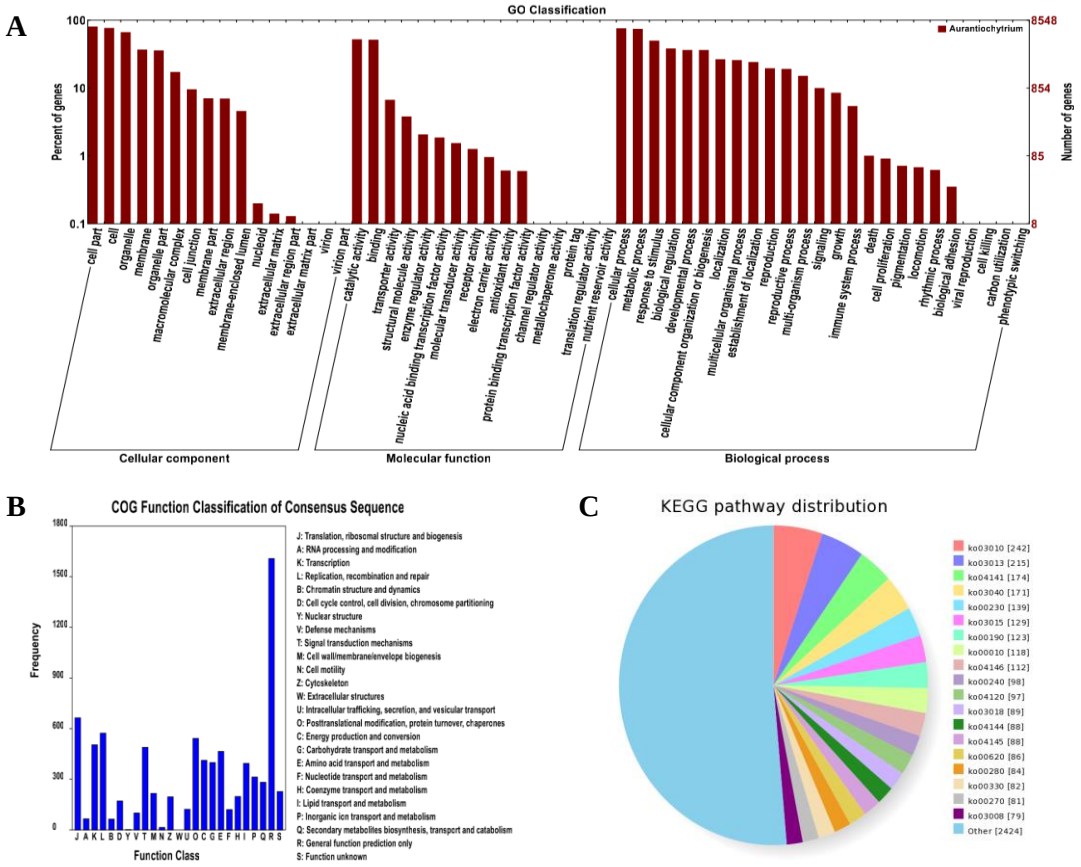

Figure S2. Functional annotation of assembled uni-genes in *Aurantiochytrium sp.* SK4. (A) GO classification, (B) COG classification, (C) KEGG pathway.

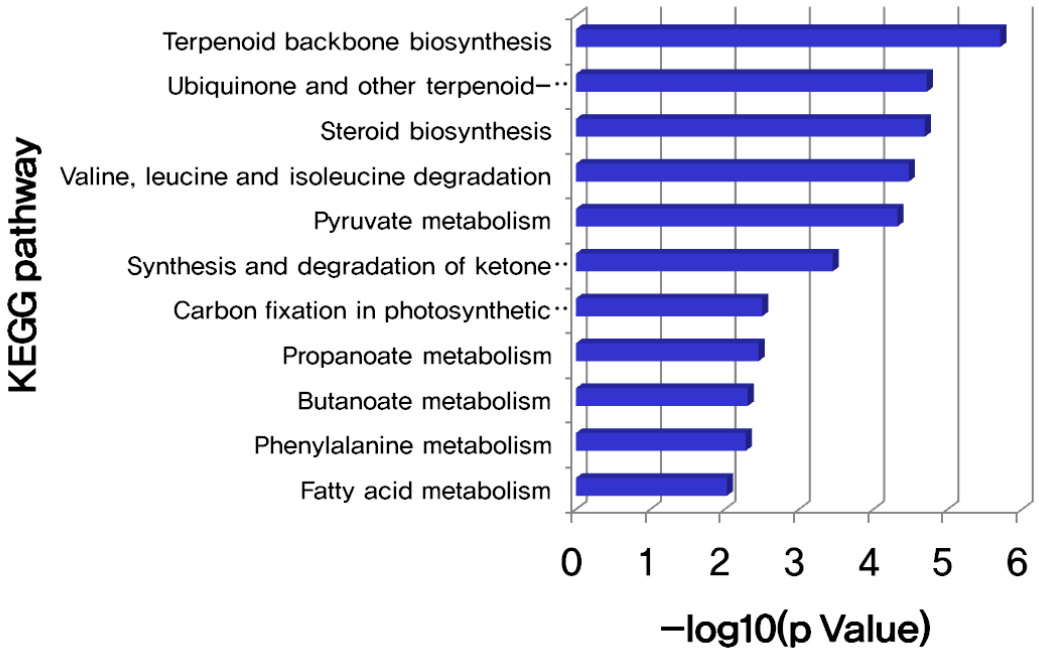

Figure S3. KEGG pathway enrichment analysis of differentially expressed genes (DEGs).

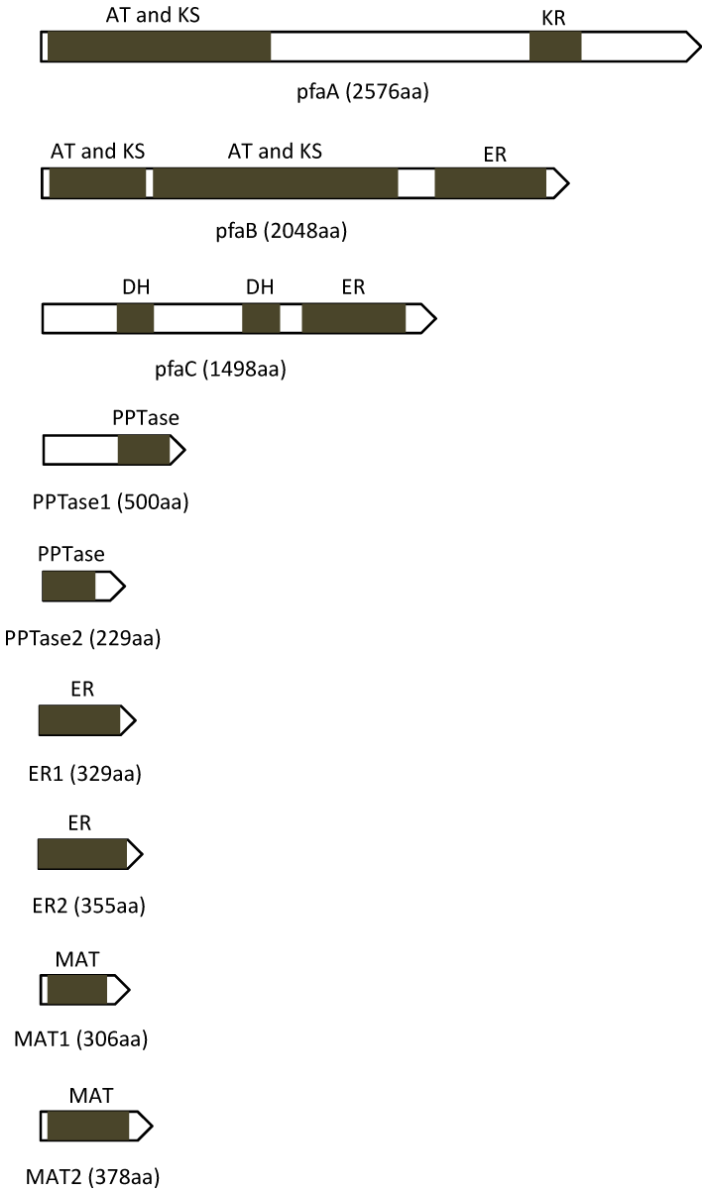

Figure S4. Genes encoding enzymes of the polyketide synthase (PKS pathway) in *Aurantiochytrium* sp.SK4. Dark gray areas indicate proposed enzymatic domains. KS, 3-ketoacyl synthase; KR, 3-ketoacyl-ACP reductase; ER, enoyl reducatase; DH, dehydrase/isomerase; PPTase, phosphopantetheine transferase; MAT, malonyl-CoA:ACP acyltransferase. aa, amino acid.

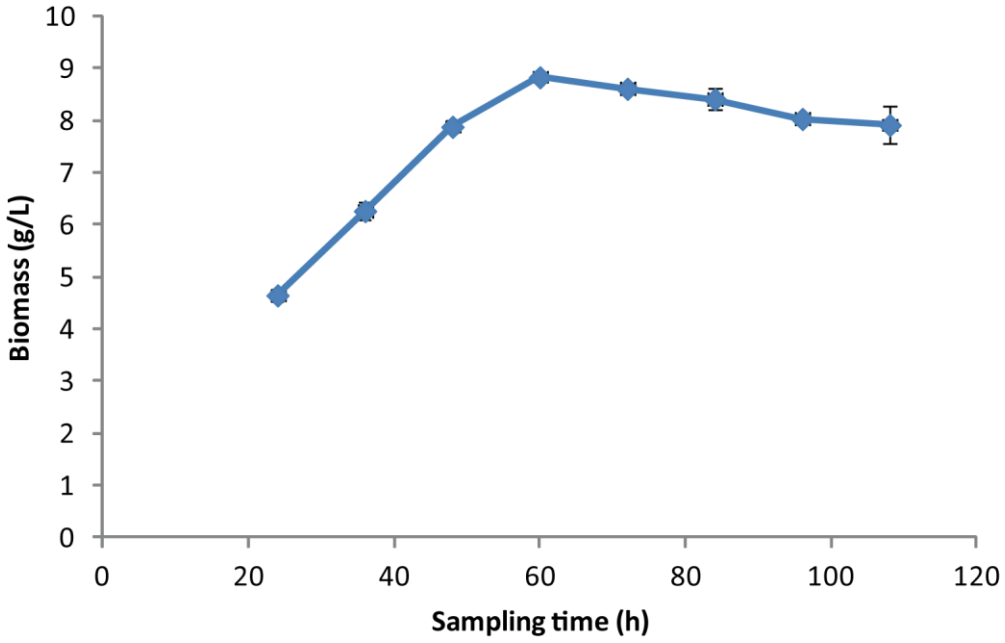

Figure S5. The growth curve of *Aurantiochytrium* sp.SK4 of Figure 1 and Figure 2.

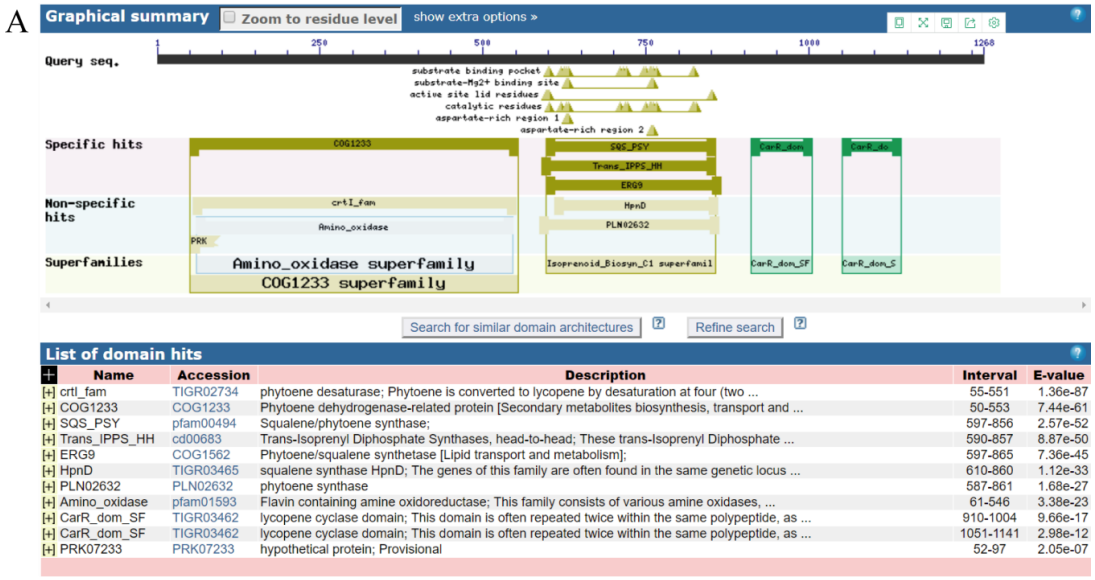

Figure S6. The conserved domains of CrtIBY and alignment of amino acid sequences of different CrtIBY. The conserved domains of CrtIBY from *Aurantiochytrium* sp.SK4 were analyzed by CDD/SPARCLE **(A)** and alignment of amino acid sequences of possible trifunctional β-carotene synthases, CrtIBY **(B)**. Sequences of CrtIBY of *Aurantiochytrium* sp. SK4 are compared with those of *Aurantiochytrium* sp. FCC1311, *Aurantiochytrium* sp. KH105, *Schizochytrium* sp. CCTCC M209059, *Aurantiochytrium* sp. T66, and *Thraustochytrium* sp. ATCC 26185.

Table S1

Comparison of *Aurantiochytrium* sp. SK4 genome statistics to other five algae and the *Arabidopsis thaliana* genome.

Organism	***Aurantiochytrium* sp. SK4**	***Arabidopsis thaliana***	***Chlamydomonas reinhardtii***	***Chlorella* sp.NC64A**	***Chromochloris zofingiensis***	***Coccomyxa subellipsoidea* C-169**	***Monoraphidium neglectum***
Sequenced genome size	49Mbp	119Mbp	107Mbp	42Mbp	58Mbp	49Mbp	67Mbp
Percent G+C in sequenced genome	56.7%	36%	64%	67%	51%	53%	65%
Coding sequence in sequenced genome	63.0%	28%	37%	32%	39%	25%	26%
Percent G+C in coding sequence	56.9%	44%	70%	69%	53%	61%	70%
Average number of exons	2.4	5.2	8.5	8.3	5.0	8.1	5.0
Average exon length	903nt	237nt	261nt	166nt	291nt	159nt	207nt
Percentage transcript with at least one intron	55.5%	76%	92%	98%	82%	94%	82%

Table S2

The expression of genes associated with carotenoids and fatty acid biosynthesis in transcriptome.

Pathway	Gene	RPKM	
		24h	96h
Astaxanthin pathway	*HMGS*	293.68	2.67
	HMGR	141.10	1.42
	MK	194.72	10.42
	PMK	21.56	12.61
	PPMD	60.02	1.11
	IDI	29.57	1.29
	CrtIBY	9.10	19.95
	CrtZ	3.87	12.66
	CrtO	14.94	28.95
FAS pathway	*Type 1 fatty acid synthase*	613.72	247.47
	Δ12 desaturase	377.4	2.0
	Δ5 desaturase	18.2	29.1
	ω-3 desaturase	0	0.6
	Δ4 desaturase	114.4	19.7
	Δ6 desaturase	0.7	0
	Δ9 desaturase	1.08	0
PKS pathway	*PKS pfaA*	754.56	1.95
	PKS pfaB	193.79	9.46
	PKS pfaC	381.54	4.38

Table S3

Contents of squalene in wild-type SK4 and the transformant AT26 at different stages.

	Sampling time	SK4	AT26
Squalene **content (mg × g^{-1} DW)**	48h	10.98 ± 0.13	0.34 ± 0.02
	72h	21.08 ± 0.06	0.41 ± 0.06
	96h	13.18 ± 0.05	0.64 ± 0.07

Data are shown as mean ± SD, n = 3.

Table S4

Primers Used for qRT-PCR and the detection of the p−−*VHb*−*ble*-2A-*IDI*-2A-*GPS* (VBIG).

Primer	Sequence
qRT-pcr-Actin-F	G A G G C C A T G T T T C A G A C C A T
qRT-pcr-Actin-R	A C G A G A G C C G T C A T T T C T G T
qRT-pcr-HMGS-F	C G C C G G C G T C G A C A G C A T
qRT-pcr-HMGS-R	G G G C A C G G C G G G C A A G A C
qRT-pcr-HMGR-F	C C G G C G C A A A A T G T C G A G T C T
qRT-pcr-HMGR-R	C C G C C G A C A G T G C C A A C C T C
qRT-pcr-MK-F	C C G C A A C C A C G A A A T C C T C C A A A A
qRT-pcr-MK-R	C G A G A G C G C C G G C A G A C T T G
qRT-pcr-PMK-F	G C C G T C T T T G C A G T T G T T G T T G A T T C
qRT-pcr-PMK-R	G C C G C C G A T C T T C A C T C A G C A A
qRT-pcr-crtIBY-F	T G G T G A C C T C G A T C A T G T G T
qRT-pcr-crtIBY-R	C G G C T C T A C A G G T A A T G A G T
qRT-pcr-FAS Ⅰ-R	G A G A A C G T C A G C A C C T T T G C
qRT-pcr-FAS Ⅰ-F	A G G C T C G A G A G A G C C T T G A C
qRT-pcr-pfaA-F	T G A T C C C T T C G T G A A T G A C C
qRT-pcr-pfaA-R	G C T C G T T G T G G A A C T G A A G G
qRT-pcr-pfaB-F	G T C A T T C T G C C C C T C A T C A T C A A C C
qRT-pcr-pfaB-R	G A C T G C T T G G C G A C C T G G T T C A C
qRT-pcr-pfaC-F	C C G C C C C A T C C A C G T C A T C C T C
qRT-pcr-pfaC-R	C C G G A C T G C T T G G C G A C C T G G T T
VBIGF	G G C T T T G G C G A T G A C G G T A T T G
VBIGR	C C C C T C C T C A T C T C G T C C C T G T