

# Genomic Insight and Optimization of Astaxanthin Production from a new *Rhodotorula* sp. CP72-2

Engkarat Kingkaew<sup>1</sup>, Nisachon Tedsree<sup>2</sup>, Sukanya Phuengjayaem<sup>3</sup>, Pornchai Rojsitthisak<sup>4</sup>, Boonchoo Sritularak<sup>5</sup>, Worathat Thitikornpong<sup>4</sup>, Somphob Thompho<sup>6</sup>, Wuttichai Mhuantong<sup>7</sup>, and Somboon Tanasupawat<sup>1,\*</sup>

<sup>1</sup>Department of Biochemistry and Microbiology, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

<sup>2</sup>Faculty of Science and Arts, Burapha University, Chanthaburi Campus, Chanthaburi 22170, Thailand

<sup>3</sup>Department of Microbiology, Faculty of Science, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand

<sup>4</sup>Department of Food and Pharmaceutical Chemistry, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

<sup>5</sup>Department of Pharmacognosy and Pharmaceutical Botany, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

<sup>6</sup>Pharmaceutical Research Instrument Center, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok 10330, Thailand

<sup>7</sup>National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), Pathum Thani 12120, Thailand

\*Corresponding author: Somboon Tanasupawat; Tel +66-2-218-8376, Fax +66-2-254-5195 E-mail: Somboon.T@chula.ac.th

## Supplementary Table S1 Carbon assimilation of strain CP72-2

Substrate	CP72-2
D-Galactose	+
Cycloheximide (actidione)	-
Sucrose	+

Substrate	CP72-2
N-acetyl-glucosamine	-
Lactic acid	-
L-Arabinose	-
D-Cellobiose	-
D-Raffinose	+
D-Maltose	-
D-Trehalose	+
Potassium 2-ketogluconate	+
Methyl- $\alpha$ -D-Glucopyranoside	-
D-Mannitol	+
D-Lactose (bovine origin)	-
Inositol	-
No substrate	-
D-Sorbitol	+
D-Xylose	+
D-Ribose	-
Glycerol	+
L-Rhamnose	-
Palatinose	+
Erythritol	-
D-Melibiose	-
Sodium glucuronate	-
D-Melezitose	-
Potassium gluconate	+
Levulinic acid (levulinate)	-
D-Glucose	+
L-Sorbose	-
Glucosamine	-
Esculin ferric citrate	+

(-) negative test, (+) positive test

**Supplementary Table S2** ANI (%) values between the draft genomes of the *Rhodotorula* sp. CP72-2; *R. kratochvilovae* CBS 7436<sup>T</sup>; *R. kratochvilovae* LS11; *R. kratochvilovae* YM25235; *R. babjevae* CBS 7808<sup>T</sup>; *R. diobovata* UCD-FST 08-225; *R. glutinis* ZHK; *R. graminis* WP1; and *R. paludigena* CM33

**Genomic data:** 1, *Rhodotorula* sp. CP72-2 (PRJNA938253); 2, *R. kratochvilovae* CBS 7436<sup>T</sup> (JACAVT000000000); 3, *R. kratochvilovae* LS11 (PQDI000000000); 4, *R. kratochvilovae* YM25235 (JAHSUG000000000); 5, *R. babjevae* CBS 7808<sup>T</sup> (CAKKSZ000000000); 6, *R. diobovata* UCD-FST 08-225 (SOZI000000000); 7, *R. glutinis* ZHK (JAAGPT000000000); 8, *R. graminis* WP1 (JTAO000000000); 9, *R. paludigena* CM33 (SWEA000000000)

Query genome	Reference genome	ANI (%)
1	2	83.1646
1	3	83.1079
1	4	83.0901
1	5	80.3203
1	6	80.2167
1	7	80.1658
1	8	80.0865
1	9	79.3643

\*Data obtained from FASTANI

**Statistical supplementary Table S3** The effect of different carbon sources on astaxanthin production and growth of *Rhodotorula* sp. CP72-2

**ANOVA**

		Sum of Squares	df	Mean Square	F	Sig.
Astaxanthin content (µg/g DCW)	Between Groups	8207.728	3	2735.909	76.788	.000
	Within Groups	285.035	8	35.629		
	Total	8492.763	11			
Astaxanthin yield (mg/L)	Between Groups	1.514	3	.505	119.935	.000
	Within Groups	.034	8	.004		
	Total	1.548	11			
Biomass (g/L)	Between Groups	20.032	3	6.677	146.728	.000
	Within Groups	.364	8	.046		
	Total	20.396	11			

**Astaxanthin content (ug/g DCW)**

Carbon source		N	Subset for alpha = 0.05	
			1	2
Tukey HSD(a)	fructose	3	217.88000	281.60000 1.000
	maltose	3	221.23000	
	sucrose	3	225.45333	
	glucose	3		
	Sig.		.453	
Duncan(a)	fructose	3	217.88000	281.60000 1.000
	maltose	3	221.23000	
	sucrose	3	225.45333	
	glucose	3		
	Sig.		.174	

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Astaxanthin yield (mg/L)**

Carbon source		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	maltose	3	.77483		
	fructose	3		1.40563	
	sucrose	3		1.52070	
	glucose	3			1.72663
	Sig.		1.000	.210	1.000

Duncan(a)	maltose	3	.77483		
	fructose	3		1.40563	
	sucrose	3		1.52070	
	glucose	3			1.72663
	Sig.		1.000	.062	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

#### Biomass (g/L)

Carbon source		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	maltose	3	3.50133		
	glucose	3		6.13133	
	fructose	3		6.45133	6.45133
	sucrose	3			6.74533
	Sig.		1.000	.324	.389
Duncan(a)	maltose	3	3.50133		
	glucose	3		6.13133	
	fructose	3		6.45133	6.45133
	sucrose	3			6.74533
	Sig.		1.000	.103	.130

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Statistical supplementary Table S4** The effect of glucose concentration on astaxanthin production and growth of *Rhodotorula* sp. CP72-2

**ANOVA**

		Sum of Squares	df	Mean Square	F	Sig.
Astaxanthin content (µg/g DCW)	Between Groups	13919.437	4	3479.859	76.182	.000
	Within Groups	456.783	10	45.678		
	Total	14376.221	14			
Astaxanthin yield (mg/L)	Between Groups	11.061	4	2.765	178.146	.000
	Within Groups	.155	10	.016		
	Total	11.216	14			
Biomass (g/L)	Between Groups	216.540	4	54.135	1759.581	.000
	Within Groups	.308	10	.031		
	Total	216.848	14			

**Astaxanthin content (µg/g DCW)**

Glucose concentra- tion		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	50 g/L	3	250.87667		
	40 g/L	3	259.27667	259.27667	
	10 g/L	3		271.92667	
	30 g/L	3		276.39000	
	20 g/L	3			337.33667
	Sig.		.572	.067	1.000
Duncan(a)	50 g/L	3	250.87667		
	40 g/L	3	259.27667		
	10 g/L	3		271.92667	
	30 g/L	3		276.39000	
	20 g/L	3			337.33667
	Sig.		.159	.437	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Astaxanthin yield (mg/L)**

Glucose concentration		N	Subset for alpha = 0.05				
			1	2	3	4	5
Tukey	10 g/L	3	1.59803				
HSD(a)	20 g/L	3		2.83603			

	30 g/L	3		3.11533			
	40 g/L	3			3.63503		
	50 g/L	3				4.13980	
	Sig.		1.000	.116	1.000	1.000	
Dun-	10 g/L	3	1.59803				
can(a)	20 g/L	3		2.83603			
	30 g/L	3			3.11533		
	40 g/L	3				3.63503	
	50 g/L	3					4.13980
	Sig.		1.000	1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

### Biomass (g/L)

Glucose concentration		N	Subset for alpha = 0.05				
			1	2	3	4	5
Tukey HSD(a)	10 g/L	3	5.87667				
	20 g/L	3		8.40333			
	30 g/L	3			11.27200		
	40 g/L	3				14.01667	
	50 g/L	3					16.49933
	Sig.		1.000	1.000	1.000	1.000	1.000
Dun- can(a)	10 g/L	3	5.87667				
	20 g/L	3		8.40333			
	30 g/L	3			11.27200		
	40 g/L	3				14.01667	
	50 g/L	3					16.49933
	Sig.		1.000	1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Statistical supplementary Table S5** The effect of pH on astaxanthin production and growth of *Rhodotorula* sp. CP72-2

### ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Astaxanthin content (µg/g DCW)	Between Groups	9594.350	4	2398.587	120.526	.000
	Within Groups	199.010	10	19.901		
	Total	9793.360	14			
Astaxanthin yield (mg/L)	Between Groups	2.584	4	.646	31.363	.000

Biomass (g/L)	Within Groups	.206	10	.021		
	Total	2.790	14			
	Between Groups	1.144	4	.286	2.364	.123
	Within Groups	1.210	10	.121		
	Total	2.354	14			

#### Astaxanthin content (µg/g DCW)

pH		N	Subset for alpha = 0.05				
			1	2	3	4	5
Tukey HSD(a)	pH 8.5	3	202.37667				
	pH 7.5	3		234.18667			
	pH 6.5	3			248.93000		
	pH 5.5	3			257.91667		
	pH 4.5	3				277.93667	
	Sig.		1.000	1.000	.175	1.000	
Duncan(a)	pH 8.5	3	202.37667				
	pH 7.5	3		234.18667			
	pH 6.5	3			248.93000		
	pH 5.5	3				257.91667	
	pH 4.5	3					277.93667
	Sig.		1.000	1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

#### Astaxanthin yield (mg/L)

pH		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	pH 8.5	3	3.53050		
	pH 7.5	3		4.21920	
	pH 6.5	3		4.36653	
	pH 5.5	3		4.48497	4.48497
	pH 4.5	3			4.77643
	Sig.		1.000	.232	.170
Duncan(a)	pH 8.5	3	3.53050		
	pH 7.5	3		4.21920	
	pH 6.5	3		4.36653	
	pH 5.5	3		4.48497	
	pH 4.5	3			4.77643

Sig.		1.000	.055	1.000
------	--	-------	------	-------

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

### Biomass (g/L)

pH		N	Subset for alpha = 0.05	
			1	2
Tukey HSD(a)	pH 4.5	3	17.1827	
	pH 5.5	3	17.3893	
	pH 8.5	3	17.4480	
	pH 6.5	3	17.5313	
	pH 7.5	3	18.0153	
	Sig.		.087	
Dun- can(a)	pH 4.5	3	17.1827	
	pH 5.5	3	17.3893	17.3893
	pH 8.5	3	17.4480	17.4480
	pH 6.5	3	17.5313	17.5313
	pH 7.5	3	18.0153	18.0153
	Sig.		.279	.067

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Statistical supplementary Table S6** The effect of temperature on astaxanthin production and growth of *Rhodotorula* sp. CP72-2

### ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Astaxanthin content (µg/g DCW)	Between Groups	94195.806	4	23548.952	270.966	.000
	Within Groups	869.074	10	86.907		
	Total	95064.880	14			
Astaxanthin yield (mg/L)	Between Groups	31.345	4	7.836	384.167	.000
	Within Groups	.204	10	.020		
	Total	31.549	14			
Biomass (g/L)	Between Groups	90.376	4	22.594	205.679	.000
	Within Groups	1.099	10	.110		
	Total	91.475	14			

### Astaxanthin content (µg/g DCW)



Temperature		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	35 °C	3	116.62667		
	30 °C	3		235.62667	
	25 °C	3			311.98667
	20 °C	3			314.00000
	15 °C	3			328.10333
	Sig.		1.000	1.000	.284
Duncan(a)	35 °C	3	116.62667		
	30 °C	3		235.62667	
	25 °C	3			311.98667
	20 °C	3			314.00000
	15 °C	3			328.10333
	Sig.		1.000	1.000	.070

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

#### Astaxanthin yield (mg/L)

Tempera- ture		N	Subset for alpha = 0.05				
			1	2	3	4	5
Tukey HSD(a)	35 °C	3	1.26133				
	30 °C	3		3.98643			
	15 °C	3		4.30807			
	20 °C	3			4.95633		
	25 °C	3				5.39293	
Duncan(a)	Sig.		1.000	.113	1.000	1.000	
	35 °C	3	1.26133				
	30 °C	3		3.98643			
	15 °C	3			4.30807		
	20 °C	3				4.95633	
	25 °C	3					5.39293
	Sig.		1.000	1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

#### Biomass (g/L)

Temperature		N	Subset for alpha = 0.05			
			1	2	3	4
Tukey HSD(a)	35 °C	3	10.83600			
	15 °C	3		13.13067		
	20 °C	3			15.78667	
	30 °C	3				16.91867
	25 °C	3				17.28267
	Sig.		1.000	1.000	1.000	.672
Duncan(a)	35 °C	3	10.83600			
	15 °C	3		13.13067		
	20 °C	3			15.78667	

30 °C	3				16.91867
25 °C	3				17.28267
Sig.		1.000	1.000	1.000	.208

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

**Statistical supplementary Table S7** The effect of incubation period on astaxanthin production and growth of *Rhodotorula* sp. CP72-2

#### ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Astaxanthin content (µg/g DCW)	Between Groups	13560.368	3	4520.123	110.617	.000
	Within Groups	326.903	8	40.863		
	Total	13887.271	11			
Astaxanthin yield (mg/L)	Between Groups	20.625	3	6.875	572.792	.000
	Within Groups	.096	8	.012		
	Total	20.721	11			
Productivity (mg/L/day)	Between Groups	1.168	3	.389	196.258	.000
	Within Groups	.016	8	.002		
	Total	1.184	11			
Biomass (g/L)	Between Groups	210.196	3	70.065	1107.446	.000
	Within Groups	.506	8	.063		
	Total	210.702	11			

#### Astaxanthin content (µg/g DCW)

Time		N	Subset for alpha = 0.05		
			1	2	3
Tukey HSD(a)	1 day	3	198.02667	240.63667	274.94000
	3 days	3			
	5 days	3			
	7 days	3			
	Sig.		1.000	1.000	.425
Duncan(a)	1 day	3	198.02667	240.63667	274.94000
	3 days	3			
	5 days	3			
	7 days	3			
	Sig.		1.000	1.000	.146

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

### Astaxanthin yield (mg/L)

		N	Subset for alpha = 0.05		
Time			1	2	3
Tukey	1 day	3	1.34997		
HSD(a)	3 days	3		4.13250	
	7 days	3			4.46757
	5 days	3			4.48120
	Sig.		1.000	1.000	.999
Dun-	1 day	3	1.34997		
can(a)	3 days	3		4.13250	
	7 days	3			4.46757
	5 days	3			4.48120
	Sig.		1.000	1.000	.883

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

### Productivity (mg/L/day)

		N	Subset for alpha = 0.05		
Time			1	2	3
Tukey	7 days	3	.63823		
HSD(a)	5 days	3		.89624	
	1 day	3			1.34994
	3 days	3			1.37750
	Sig.		1.000	1.000	.871
Dun-can(a)	7 days	3	.63823		
	5 days	3		.89624	
	1 day	3			1.34994
	3 days	3			1.37750
	Sig.		1.000	1.000	.470

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.

### Biomass (g/L)

Time		N	Subset for alpha = 0.05			
			1	2	3	4
Tukey HSD(a)	1 day	3	6.81667			
	7 days	3		15.76933		
	5 days	3		16.29733		

Duncan(a)	3 days	3			17.17133	
	Sig.		1.000	.122	1.000	
	1 day	3	6.81667			
	7 days	3		15.76933		
	5 days	3			16.29733	
	3 days	3				17.17133
	Sig.		1.000	1.000	1.000	1.000

Means for groups in homogeneous subsets are displayed.

a Uses Harmonic Mean Sample Size = 3.000.