

## Electronic Supplementary Material

# Improving lipid production by *Rhodotorula glutinis* for renewable fuel production based on machine learning

Lihe Zhang<sup>2,3</sup>, Changwei Zhang<sup>2,3</sup>, Xi Zhao<sup>2,3</sup>, Changliu He<sup>2,3</sup>, Xu Zhang (✉)<sup>1,2,3</sup>

1 Beijing Key Lab of Bioprocess, Beijing University of Chemical Technology, Beijing 10029, China

2 National Energy R&D Center for Biorefinery, Beijing University of Chemical Technology, Beijing 100029, China

3 Collage of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

E-mail: zhangxu@mail.buct.edu.cn

### 1. Analytical methods

Dry cell weight method was used to measure the biomass. Moreover, the concentration of glucose was measured by a glucose biosensor (SBA 40C, Biological Institute of Shandong Academy of Sciences). Lipid was extracted using the method reported by. The lipid components were analyzed as described in the existing study.

HPLC (Thermo Scientific, Waltham, MA, USA) was used to measure the concentration of sugars, organic acids and other organics.

### 2. Experimental results of oil fermentation

Table S1 Effects of main nitrogen sources and concentration on lipid fermentation

Glucose (g/L)	Yeast Extract (g/L)	Corn steep Liquor (g/L)	Urea (g/L)	Ammonium sulfate (g/L)	Pepton e(g/L)	Ammonia chloride(g/L)	Fermentatio n period (h)	Residual Sugar (g/L)	Biomass (g/L)	Lipid (g/L)
42.5	0	2	0	2	0	0	93	0.10	13.99	4.05
41.5	0	0	0	2	2	0	72	5.20	13.28	3.67
45	2	0	0	2	0	0	72	1.00	15.49	4.34
23	2	0	0	2	0	0	47	0.00	12.69	2.27
68	2	0	0	2	0	0	165	2.00	18.26	4.78
98.6	2	0	0	2	0	0	213	20.00	19.1	6.21
42	1.5	0	0	2	0	0	72	0.00	16.84	4.43
42	2	0	0	2	0	0	66	3.18	15.59	4.24
42	2.5	0	0	2	0	0	66	2.60	16.22	4.56
42.5	2	0	0	2	0	0	72	0.00	15.15	4.88
42.5	2	0	0	2	0	0	72	0.00	15.41	3.44
42.5	2	0	0	2	0	0	72	0.00	15.68	4.20
102.63	2	0	0	2	0	0	190	27.00	24	5.69
45.95	1	0.5	1	0	0	0	80	0.00	18.73	5.29
45.95	1	0.5	2	0	0	0	80	0.00	19.23	2.95
45.95	1	0.5	4	0	0	0	76	1.00	18.90	2.90
45.95	1	0.5	0	0	0	1	80	15.00	10.00	3.00
45.95	1	0.5	0	0	0	2	80	3.00	14.23	3.42
45.95	1	0.5	0	0	0	4	80	0.00	15.70	3.04

Other components(g/L): KH<sub>2</sub>PO<sub>4</sub> 7, Na<sub>2</sub>SO<sub>4</sub> 2, MgSO<sub>4</sub> 1.5

Table S2 Effect of main component concentration of medium on lipid fermentation

Glucose (g/L)	pH	Yeast Extract (g/L)	Corn Steep Liquor (g/L)	Sodium Citrate (g/L)	K <sub>2</sub> HPO <sub>4</sub> (g/L)	KH <sub>2</sub> PO <sub>4</sub> (g/L)	MgSO <sub>4</sub> (g/L)	Na <sub>2</sub> SO <sub>4</sub> (g/L)	Monosodium Glutamate (g/L)	Ammonium sulfate (g/L)	Fermentation period (h)	Residual Sugar (g/L)	Biomass (g/L)	Lipid (g/L)
49.1	5.5	0.5	1	0	1.5	1	1.5	2	0	2	84	2.73	15.48	3.97
49.1	5.5	0.5	1	0	1.5	1	1.5	2	0	1	101	14.4	11.08	3.05
49.1	5.5	0.5	1	0	1.5	1	1.5	2	0	4	84	0.2	16.91	3.11
49.1	5.5	0.5	1	0	0.5	1	1.5	2	0	2	101	0.1	15.01	3.51
49.1	5.5	0.5	1	0	5	1	1.5	2	0	2	101	0.1	16.27	4.09
49.1	5.5	0.5	1	0	1.5	1	1.5	1	0	2	84	4.1	14.35	3.66
49.1	5.5	0.5	1	0	1.5	1	1.5	4	0	2	101	0.1	14.92	4.01
48	5.5	0.5	1	0	1.5	1	1.5	2	2	2	84	0.3	16.76	4.35
48	5.5	0.5	1	0	1.5	1	1.5	2	4	2	84	0.1	19.46	4.56
48	5.5	0.5	1	0	1.5	1	0.5	2	0	2	84	2.8	14.90	3.92
48	5.5	0.5	1	0	1.5	1	4	2	0	2	84	2.3	14.99	3.85
48	5.5	0.5	1	0	1.5	0.5	1.5	2	0	2	101	0.1	14.51	4.28
48	5.5	0.5	1	0	1.5	3	1.5	2	0	2	101	0.1	15.09	4.13
47.3	6	0.5	1	0	1.5	1	1.5	2	0	2	101	0.1	14.4	3.86
47.3	7	0.5	1	0	1.5	1	1.5	2	0	2	101	0.1	15.09	3.68
47.3	5.5	0.5	0.5	0	1.5	1	1.5	2	0	2	101	0.1	14.87	3.56
47.3	5.5	0.5	3	0	1.5	1	1.5	2	0	2	72	3.7	16.08	3.96
47.3	5.5	1	1.02	0	1.5	1	1.5	2	0	2	84	0.6	15.69	3.83
47.3	5.5	3	1.02	0	1.5	1	1.5	2	0	2	84	0.1	16.97	4.55
45.95	5.5	1	0.5	0	1.5	1	1.5	2	0	0	76	39	4.55	1.54
45.95	5.5	1	0.5	0	1.5	1	1.5	2	0	1	76	16.5	9.8	2.45

---

45.95	5.5	1	0.5	0	1.5	1	1.5	2	0	2	80	2.3	13.78	3.80
45.95	5.5	1	0.5	0	1.5	1	1.5	2	0	4	80	0.2	16.03	2.55
45.95	5.5	1	0.5	0	1.5	1	1.5	2	2	2	80	0	15.94	4.03
45.95	5.5	1	0.5	0	1.5	1	1.5	2	8	2	80	0	21.07	4.61
45.95	5.5	1	0.5	0	1.5	1	1.5	2	0	2	80	0.2	13.77	3.73
45.95	5.5	1	0.5	1	1.5	1	1.5	2	0	2	80	0.2	14.46	3.81
45.95	5.5	1	0.5	2	1.5	1	1.5	2	0	2	80	0.2	15.67	4.50
45.95	5.5	1	0.5	4	1.5	1	1.5	2	0	2	80	0.2	17	4.98

---

Table S3 Effects of different concentrations of carbon and nitrogen sources on lipid fermentation

Glucose (g/L)	Yeast Extract (g/L)	Corn Steep Liquor (g/L)	Urea (g/L)	Sodium Citrate (g/L)	K <sub>2</sub> HPO <sub>4</sub> (g/L)	KH <sub>2</sub> PO <sub>4</sub> (g/L)	MgSO <sub>4</sub> (g/L)	Na <sub>2</sub> SO <sub>4</sub> (g/L)	Fermentation period (h)	Residual Sugar (g/L)	Biomass (g/L)	Lipid (g/L)
42	1	0.5	1	4	2	1	2	2	68	1.00	17.70	4.72
41	2	1	2	4	2	1	2	2	68	0.00	18.20	3.90
41.1	4	2	4	4	2	1	2	2	60	1.00	20.00	3.69
65	1	0.5	1	4	2	1	2	2	144	8.00	23.00	8.21
65.6	2	1	2	4	2	1	2	2	95	1.00	27.80	8.40
64	4	2	4	4	2	1	2	2	95	0.00	28.80	5.80
64.1	6	3	4	4	2	1	2	2	84	1.00	29.10	4.91
93	1	0.5	1	4	2	1	2	2	300	17.70	26.00	9.20
92.2	2	1	2	4	2	1	2	2	160	0.00	32.40	12.10
77.8	4	2	4	4	2	1	2	2	140	0.00	32.40	6.40
90	6	4	4	4	2	1	2	2	140	1.00	35.90	7.30
123	1	0.5	1	4	2	1	2	2	156	68.00	22.00	7.40
141	2	1	2	4	2	1	2	2	156	65.00	32.00	9.80
124.9	4	2	4	4	2	1	2	2	214	0.00	40.00	12.00
124.9	6	4	4	4	2	1	2	2	156	18.00	42.00	10.60
124.9	6	4	4	4	4	1	2	2	214	0.00	39.00	8.80

Table S4 Effect of carbon source concentration and component concentration on lipid fermentation

Glucose (g/L)	Yeast Extract (g/L)	Corn Steep Liquor (g/L)	Urea (g/L)	Sodium Citrate (g/L)	K <sub>2</sub> HPO <sub>4</sub> (g/L)	KH <sub>2</sub> PO <sub>4</sub> (g/L)	MgSO <sub>4</sub> (g/L)	Na <sub>2</sub> SO <sub>4</sub> (g/L)	Monosodium Glutamate (g/L)	Fermentation period (h)	Residual Sugar (g/L)	Biomass (g/L)	Lipid (g/L)
75.95	1	0.5	1	3	1	1	1	1	0	144	27.00	13.95	5.13
78.06	2	1	2	6	1	1	1	1	2	144	1.00	28.50	8.70
75.95	3	2	3	9	1	1	1	1	4	125	0.70	32.00	7.30
72	1	0.5	1	3	1	1	1	1	2	125	25.30	17.10	6.20
78.5	2	1	2	6	1	1	1	1	4	125	1.50	31.00	10.35
74.74	3	2	3	9	1	1	1	1	0	125	0.50	32.00	8.55
83	1	1	3	3	2	2	1	2	0	125	4.50	31.40	8.99
76.33	2	2	1	6	2	2	1	2	2	144	8.20	30.10	9.85
80.1	3	0.5	2	9	2	2	1	2	4	144	1.20	33.20	11.00
78.8	1	2	2	3	2	2	2	1	0	144	1.00	30.10	10.50
77.21	2	0.5	3	6	2	2	2	1	2	125	5.00	29.90	9.94
70.82	3	1	1	9	2	2	2	1	4	144	6.20	31.00	10.70
79.27	2	2	1	3	1	2	2	2	4	144	5.70	31.00	11.30
82.35	3	0.5	2	6	1	2	2	2	0	125	7.00	28.90	10.41
74.36	1	1	3	9	1	2	2	2	2	125	2.00	30.00	8.50
74	2	2	2	3	2	1	2	2	0	125	1.00	28.00	9.65
75.04	3	0.5	3	6	2	1	2	2	2	125	0.60	28.00	8.55

74	1	1	1	9	2	1	2	2	4	144	8.40	28.00	9.10
85.64	2	0.5	3	3	2	1	1	2	4	145	4.00	31.00	9.51
93.52	3	1	1	6	2	1	1	2	0	192	25.00	23.50	6.74
87.32	1	2	2	9	2	1	1	2	2	168	9.40	29.90	9.80
89.24	2	1	3	3	1	2	1	2	4	145	2.30	33.10	10.11
94.5	3	2	1	6	1	2	1	2	0	145	25.00	25.30	6.70
86.46	1	0.5	2	9	1	2	1	2	2	145	7.50	27.00	9.54
86.46	1	0.5	2	9	1	2	1	2	2	168	0.70	27.50	9.82
106.7	3	1	1	3	2	2	1	1	2	192	14.90	30.00	8.82
92.32	1	2	2	6	2	2	1	1	4	168	0.60	30.85	10.99
90.68	2	0.5	3	9	2	2	1	1	0	145	1.60	31.80	8.93
94.25	3	1	2	3	1	1	2	2	2	168	4.00	30.90	10.11
92.93	1	2	3	6	1	1	2	2	4	168	0.00	33.9	9.62
89.99	2	0.5	1	9	1	1	2	2	0	192	21.40	27.00	6.56
89.95	3	2	3	3	2	1	2	1	2	145	1.00	33.50	9.42
101	1	0.5	1	6	2	1	2	1	4	192	17.40	28.80	8.40
85.27	2	1	2	9	2	1	2	1	0	168	6.00	30.10	9.36
92.34	3	0.5	2	3	1	2	2	1	4	192	3.00	29.80	9.82
104	1	1	3	6	1	2	2	1	0	168	1.80	30.80	9.33
102	2	2	1	9	1	2	2	1	2	192	19.60	29.00	8.54

---

### 3.1. Evaluation of biomass prediction model based on BP-ANN

Table S5 Evaluation of biomass prediction model based on BP-ANN

BP-ANN-Biomass				
Number of neurons	Training data		Test data	
	R <sup>2</sup>	MSE	R <sup>2</sup>	MSE
5	0.98909	0.0043128	0.89538	0.073539
6	0.99121	0.0034767	0.91395	0.05301
7	0.99236	0.0030256	0.91339	0.065231
8	0.99326	0.0026695	0.92792	0.061748
9	0.99413	0.0023257	0.79528	0.25557
10	0.99283	0.0028403	0.91274	0.052726
11	0.99563	0.0017336	0.8844	0.074111
12	0.99554	0.0017704	0.86052	0.073681
13	0.99601	0.0015834	0.876345	0.082721
14	0.9964	0.0014301	0.53251	0.32969
15	0.99748	0.0009998	0.75085	0.17524
16	0.99749	0.00099756	0.43495	0.88349
17	0.99748	0.00099869	0.41818	1.5245
18	0.99749	0.00099633	0.69928	0.24139
19	0.99749	0.00099643	0.35098	1.493
20	0.9975	0.00099437	0.84326	0.24782
21	0.99749	0.00099639	0.47755	1.1923
22	0.99749	0.00099789	0.70359	0.27107
23	0.9975	0.00099148	0.46745	1.1923
24	0.99749	0.00099759	0.71287	0.71261
25	0.99749	0.00099547	0.81468	0.2214

### 3.2. Evaluation of lipid concentration prediction model based on BP-ANN

Table S6 Evaluation of lipid concentration prediction model based on BP-ANN

BP-ANN- lipid concentration				
Number of neurons	Training data		Test data	
	R <sup>2</sup>	MSE	R <sup>2</sup>	MSE
5	0.98069	0.008319	0.95092	0.11048
6	0.988832	0.0050529	0.94154	0.063246
7	0.98817	0.0051148	0.25754	2.5302
8	0.99054	0.0040941	0.95846	0.062235
9	0.9935	0.0028164	0.95109	0.036979
10	0.99354	0.0027992	0.91337	0.082682
11	0.9945	0.0023864	0.87887	0.087868
12	0.9946	0.0020909	0.92202	0.14801

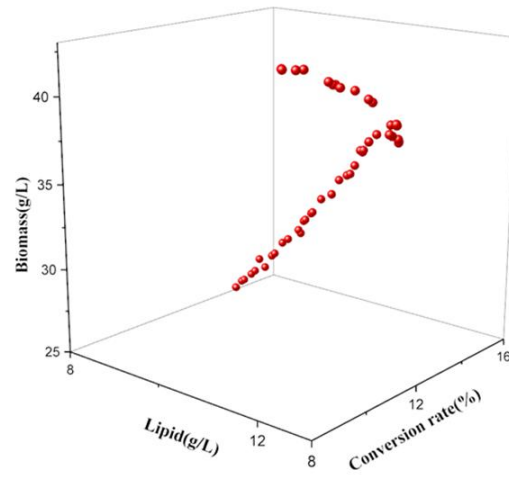
13	0.99547	0.0019655	0.8377	0.23726
14	0.9965	0.0015191	0.86095	0.19266
15	0.99629	0.0016095	0.84237	0.15279
16	0.99682	0.0013802	0.69792	0.28464
17	0.99697	0.0013149	0.80397	0.50194
18	0.9977	0.0009998	0.86966	0.091377
19	0.9977	0.00099972	0.38595	1.4633
20	0.9977	0.00099877	0.74298	0.30582
21	0.99771	0.00099629	0.85695	0.44763
22	0.9977	0.0009985	0.66222	0.69615
23	0.9977	0.00099755	0.70114	0.51976
24	0.9977	0.00099727	0.74534	0.63125
25	0.99777	0.00097074	0.77915	0.50392

### 3.3. Evaluation of consumption of glucose prediction model based on BP-ANN

Table S7 Evaluation of consumption of glucose prediction model based on BP-ANN

BP-ANN- consumption of glucose				
Number of neurons	Training data		Test data	
	R <sup>2</sup>	MSE	R <sup>2</sup>	MSE
5	0.99511	0.00177	0.96966	0.014209
6	0.99635	0.001293	0.95713	0.023219
7	0.9955	0.001594	0.95963	0.023184
8	0.99716	0.001009	0.96105	0.025822
9	0.99719	0.000995	0.87517	0.099974
10	0.99719	0.000997	0.90248	0.052723
11	0.99718	0.000999	0.96051	0.022268
12	0.99718	0.000999	0.93951	0.034585
13	0.99721	0.000989	0.95211	0.022923
14	0.99721	0.000993	0.88806	0.061252
15	0.99722	0.000987	0.95671	0.029942
16	0.99722	0.000986	0.94781	0.03127
17	0.99724	0.000983	0.95631	0.03025
18	0.99725	0.000977	0.97011	0.014423
19	0.99728	0.000966	0.96195	0.022222
20	0.99728	0.000968	0.93412	0.030536
21	0.9972	0.000995	0.96683	0.0252
22	0.99725	0.000977	0.93052	0.037956
23	0.99721	0.000991	0.95092	0.046684
24	0.99724	0.000982	0.9647	0.020192
25	0.99768	0.000904	0.93684	0.035301

### 3.4. Three objective optimization results of lipid fermentation



**Figure S1** Three-objective Pareto optimal solution distribution curve based on genetic algorithm