

Supplementary Material

Expression of phosphotransacetylase in *Rhodospiridium toruloides* leading to improved cell growth and lipid production

Xiaobing Yang^{a,b}, Wenyi Sun^a, Hongwei Shen^a, Sufang Zhang^a, Xiang Jiao^a, Zongbao K. Zhao^{a,c,*}

a. Division of Biotechnology, Dalian Institute of Chemical Physics, CAS, Dalian 116023, PR China

b. College of Enology, Northwest A&F University, Yangling, Shaanxi 712100, PR China

c. Dalian National Laboratory for Clean Energy, Dalian Institute of Chemical Physics, CAS, Dalian 116023, PR China

*Correspondence:

Associate Prof. Xiaobing Yang Mobile: +86-187-2952-0029, E-mail: yangxb@nwafu.edu.cn

Table S1. Primers used in this study.

Name	Sequence	Function
Pta-EcoR V-F	CCATGGGATATCATGCACCACCACCACCAC GCTGACCTCTTCTCGAC	Gene cloning and colony PCR
Pta-Spe I-R	GGACTAGTCTAGAGAGCCTGAGCCGCC	
Pta-RtPCR-F	ACGAGAACTACTTTCGGCAC	RT-PCR detection
Pta-RtPCR-R	CTAGAGAGCCTGAGCCG	
Pta- RB764	CGGAGATCAAGGGCGACGCTAA	
Pta- RB826	TACAAGATCGCTCAGCGCCTCG	
Pta-RB947	TCATCACGGCGGCTCAGGCTCT	
LAD-1	ACGATGGACTCCAGAGCGGCCGCVNVNNGGAA	Primer used for TAIL-PCR
LAD-2	ACGATGGACTCCAGAGCGGCCGCBNBNNNGGTT	
LAD-3	ACGATGGACTCCAGAGCGGCCGCVNVNNNCCA A	
LAD-4	ACGATGGACTCCAGAGCGGCCGCBDBNNNCGG T	
GPD-Xba I-F	CTAGTCTAGATCCATGCTGCTGCGATCTGGGAGT G	Primer used for Glyceraldehyde- 3-phosphate dehydrogenase Promoter (<i>Pgpd</i>)
GPD-EcoR V-R	CCATGGGATATCTGTGACTGATCTGGTGTGTTTC TGA	

Table S2. The T-DNA insertion information of the mutants.

Strain	T-DNA position	inserted	Functional annotation	Insertion region	Source
NP-Pta-4	XM_016413716.1		Hypothetical protein	Scaffold 1, 74977	This study
NP-Pta-6	RHTO_01356- RHTO_01357		Oligopeptide transporter and RNA polymerase II associated protein 3	Scaffold 13, 110670	This study
NP-Pta-15	XM_016413716.1		Hypothetical protein	Scaffold 1, 75043	This study
NP-Pta-20	XM_016413716.1		Hypothetical protein	Scaffold 1, 74977	This study

```

      10      20      30      40      50      60      70      80      90      100     110
Pta GTG----ATG.seq ATGGCAGATTATTTTCACAGTGCAGAAAAGTAGCTGGAAGAGACGTTAAATTTGATTTTCCTGAAAGGCTTAGACGAGCGTATTTTAGAAGCGGTCAAGCTTCAGGA 114
PTA redressed.seq  ATGGCTGACCTCTTCTCGACCGTCCAGGAGAAGGTCGCCGGCAAGGACGTC AAGATCGTCTCCAGAGGGCCCTCGACGAGCGCATCCTCGAGGCTGTCTCGAAGCTCGCCGCG 114

      120     130     140     150     160     170     180     190     200     210     220
Pta GTG----ATG.seq AACAAAGTGTGTAATCCGATTTGTGATCGGCAATGAAATGAGATCCAAAGCAAAAGCAAAAGAAITGAACCTTACGCTGGCGCGCTTAAGATTATGATCCTCATACATATGAA 228
PTA redressed.seq  AACAAAGTCTCAACCTATCGTCAATCGGCAAGGAGAACGAGATCCAGGCTAAGGCTAAGGAGTCAACCTCACCTCGCGCGCTCAAGATCTACGACCTCAACCTACCTACGAG 228

      230     240     250     260     270     280     290     300     310     320     330     340
Pta GTG----ATG.seq GGCATGGAAGACCTTGTCAAGCATTCGTAGAACGCCGCAAGGGCAAGGCACTGAGAAACAGGCTCGTAAGCAITATTAGACGAGAAGCTACTTCGGTACATGCTGGTGTAT 342
PTA redressed.seq  GGCATGGAAGACCTGTCCAGGCTTTCGTGAGCCGCGCAAGGGCAAGGCTACCGAGGAGCAGGCTCGCAAGGCTCTCCAGGCTCTCCAGCAGAGAAGCTACTTCGGCACCAATGCTGTCTAC 342

      350     360     370     380     390     400     410     420     430     440     450
Pta GTG----ATG.seq AAGAGGCTTGGACGGGCTTGTAGCGGAGCTGCTCACTCACTGCTGACACTGCTCCGCCCGGCTCTTCAATCATCAAAACCAAGGAGGCGGTGAAAAGACTCTCGGGCGT 456
PTA redressed.seq  AAGGGCTCGCCAGCGGCTGTCTCGGGCGTCTCACTCGACCGCTGACACTGCTCCGCCCGCTCTCCAGATCATCAAGACCAAGGAGGCGGTCAAGAAGACTCGGGCGT 456

      460     470     480     490     500     510     520     530     540     550     560     570
Pta GTG----ATG.seq TTCATCATGGCTCGCGGAGAGAGCAATACGTAATTCGCAAGATTCCGCATCAACATTGCACCTGACAGCCAGATCTTGGCCAGATTGCGGAGATTGCATCGAAGGTGCCAATACGCCAAA 570
PTA redressed.seq  TTCATCATGGCTCGCGGAGAGAGCAATACGTAATTCGCAAGATTCCGCATCAACATTGCCTCCAGACTCGCAGGACTCGCCGAGATCGCTATCGAGTCGGTAAACACTGCTAAG 570

      580     590     600     610     620     630     640     650     660     670     680
Pta GTG----ATG.seq ATGTTGACATTGAGCCTCGCGTGGCAATGCTCAGCTTCTCTACAAAAGGCTCAGCAAAATCTGATGAAACAGAAAAGTAGCCGGATGCAGTGAATTCGCAAGGAAAAGGC 684
PTA redressed.seq  ATGTTGACATTGAGCCTCGCGTGGCAATGCTCAGCTTCTCTACAAAAGGCTCAGCAAAATCTGATGAAACAGAAAAGTAGCCGGATGCAGTGAATTCGCAAGGAAAAGGC 684

      690     700     710     720     730     740     750     760     770     780     790
Pta GTG----ATG.seq CCTGAACTGACACTTGCAGCGCAATTCGAATTTGATGCTGCAATTTGTTCCATCTGTAGCTGAGAAAAGGCGCCGATTCCGAGATCAAAAGGGGACGCTAACGTAATTCGTAATC 798
PTA redressed.seq  CCAGAGCTCACCTCGACGGCGAGTCCAGTTGACGCCGCTTTGTCTCTCCAGGCTGAGAAAGAGGCTCCAGACTCGGAGATCAAGGGGACGCTAACGTAATTCGTAATC 798

      800     810     820     830     840     850     860     870     880     890     900     910
Pta GTG----ATG.seq CCAAGTCTTGAAGCAGGAACATCGGCTATAAATCGCTCAGCGTTTGGCAACTTTGAAAGCGTAGGACCAATCCTGCAAGGTTTAAATATGCGCTGTAACGACCTTTCAGA 912
PTA redressed.seq  CCTTCCCTCGAGGCCGGCAACATCGGCTAACAGATCGCTCAGCGCTCGGCAACTTGGAGGCTGTCCGCCCTATCTCCAGGGCTCAAGATGCCAGTCAACGACCTTTCAGC 912

      920     930     940     950     960     970
Pta GTG----ATG.seq GGAATGTAACGCTGAAAGATGTTTACAATCTCGCAATTAATTACAGCGGCGCAAGCACTGTAA 972
PTA redressed.seq  GGCTGCAACGCTGAGGAGCTTACAACCTCGCTTCATCACGCGGCTCAGGCTCTAG 972

```

Figure S1. Alignment of Pta sequence between the native (Pta GTG--ATG) and the de novo synthesized (PTA redressed). The initial codon GTG from *B. subtilis* was changed to ATG when the alignment performed.

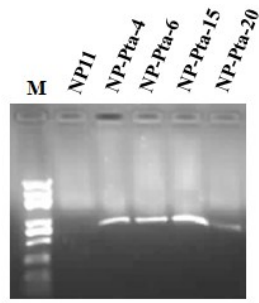


Figure S2. RT-PCR results of the engineered strains. The RT-PCR was performed using the PrimeScript™ II High Fidelity RT-PCR Kit with the Primer pair Pta-RtPCR-F/Pta-RtPCR-R.

```

>Rhodosporidium
  toruloides|RHTO_04463.t1p|RHTO_04463.t1|RHTO_04463|phosp
  hoketolase|112|866
  Length = 865

  Score = 473 bits (1218), Expect = e-134
  Identities = 279/774 (36%), Positives = 420/774 (54%), Gaps = 47/774 (6%)

Query: 22 IEGVDKTYRAANYLSIGQIVLRSNPLMKFPFTREDVQKR——LVGHWGTTPLGNFLIG 76
      ++ + WR + YL  QI+L+SN ++  T++DVK R   L G   GL+
Sbjct: 46 LDAIVANWRSLCYLGASQIFLQSNAILSRLTKDDVKPRAAMTLAGD—IQGGLSLAYV 103

Query: 77 HINRLIADH——QQNTVIDMGFGHGGPAGTAQSYLDGTYTEYFPNITKDEAGLQKFF 130
      H  LI          + + + GPGHG PA + Y++G +++++P  +E GL+KF
Sbjct: 104 HTQALIRRRGDEEGAEPKMFVVTGPGHGAPAILSPLYIEGAISKFPYQVPLNEQGLEKRV 163

Query: 131 RQFSYPGGIFSHYAFETPGSIMEGGELGYALSHAYGAVMNNPSLFPVAIVGDGEAETGPL 190
      + FS+PGG PSH  ETPG IMEGGELGYAL  AYG+VM+ P      +VGDGE+ETGP
Sbjct: 164 KYFSWPGGPFPSHVAETPGCIHEGGELGYALGVAYGSMRDRPEQISVVVVGDGESETGPT 223

Query: 191 ATGWQSNKLNPRTDGIVLPFILHLNGYKIANPTILSRISDEELHEFFHGMGYEPEYFVAG 250
      AT W S+K ++P  G VLPILH+NG+KI+  TI  + + EL  + G GY+
Sbjct: 224 ATAWSHKQWLDPAESGAVLPFILHVGFKISERTIPGTMNVELSLLYSGVGVQVRFVEYK 283

Query: 251 FDNEDHLS———IHREFAELFETVDFEICDIKAAAQT—DIMTRPFYPMIIFRTPK 300
      E H+          +H  A  + + EI  I+ AA++  + +P +PMII R+PK
Sbjct: 284 AQGEAHMGGNDPADRVLHEDMAASLDWAYGEIRKIQKAARSGGKPIDKPRWPMIILRSPK 343

Query: 301 GWTCPKFDIDGKTEGSRKSHQVPLASARDTEAHFEVLKNWLESYKPEELFDENGAVKPEV 360
      GWT P  GK+  ++ SHQVPL  A+  +  L+ WL+SY+ ++LFD  +
Sbjct: 344 GWTGFSSEHGKQLLNNFASHQVPLPDAKTDDANAYLERWLKSYEADKLFDFSDNLKRG 403

Query: 361 TAFMPTGELRIGENFNANGGRIREELKLPKLEDYEVKEVAEYGHGWGQ—LEATRRLGVYT 419
      T F          + ++  G ++E  + E+ + +YG  G+ +  + + Y
Sbjct: 404 TIFDQLLYEALPKDMERRLGFVKETYN—GYKPLEDDWKKYGFKKGEDVSCMKALAGYL 461

Query: 420 RDIKKNPDSFRIFGPDETASNRLQAAYDVT—NEQWVAGYLSAQVDEHMAVTVGQVTEQL 477
      D+IK NP  FRIF  PDE A N+L  + VT  N QW+          E  G+VTE L
Sbjct: 462 TDVIRKRPKEFRIFSPDELALNKLDGVFSVTERNMQWD———PETAHKGGRVTEML 513

Query: 478 SEHQMEGFLEGYLLTGRHGIWSSYESFVHVIDSMLNQHAKWLEATVREIPWRKPISSMNL 537
      SEH ++ +L+GY LTGRHG++ SYE+F+ ++ +M  Q+ K+++  + E  WR P +S+
Sbjct: 514 SEHSLQAWLQGYTLTGRHGVFVSYEAFLGIVATMTVQYTKFMKML—ETNWRGPTASLTY 572

Query: 538 LVSSHVWRQDHNGFSSHQDPGVTSVLLNKCFFNNDHVIGIYFPVDSNMLLAVAEEKCYSTNK 597
      + +S  RQ+HNG+SHQ+PG  S +L+          + +YFP D+N  ++V  C +S N
Sbjct: 573 IETSTWTRQEHNGYSHQNPGFVSTVLSL——PSQLARVYFPPSDANTSVSVAIHLRSLKNV 629

Query: 598 INAIAGKQPAATWLTLDDEARAELEKGAEEWKNVKSNDQAQIVLAATGDVPTQEDMA 657
      IN I+  K P  +L+++EA  GA+ W+  S V  + +VL  G  T+E++
Sbjct: 630 INLIVGTAKPTPVYLSVEEAERHCIAGASVWENYS—VDKGVDPDVLVIGIGYELTEEVH 688

Query: 658 AADKLD——AMGKFKVNVVVDLVLKLSAKENNEALSDEEFAELFTEDKPVLFAYHSYAR 714
      AA  L          ++ +VVNVVDL+ L  ++  AL + F  LF  P++F YH YA
Sbjct: 689 AAALLRKDFGTELVRVNVVDLLVLAPKGDHPHALDEAGFNSLFPFGVPIIFVNYHYAG 748

Query: 715 DVRGLIYDRFN——HDNFNVHGYEEQGSTTTPYDMVRVNNIDRYELQAEALRMI 765
      + L++DR +          + Y  EQG+TTP+ M+  NN DR+ L AEAL M+
Sbjct: 749 QLASLLFDRIGHSVGRSRMRIFAYSEQGTTTTFPAMMCCNNTDRFNLAEEALEMV 802

```

Figure S3. Alignment of the Pk genes of *Bifidobacterium adolescentis* and *R. toruloides*.

Query: Pk of *B. adolescentis* GenBank: KLE27828.1; Sbjct: PK of NP11 GenBank: EMS19462.1or RHTO_04463.