Supplementary Material

Improving the expression of taxadiene synthase to enhance the titer of taxadiene in *Saccharomyces cerevisiae*

Chenglong Zhang[a,b], Jia Wang[a,b], Yi Shi[ab], Nan Wu[a,b], Xia Li [a,b],Ying Wang[a,b]， Bingzhi Li[a,b], Wenhai Xiao[a,b,c,d*], Mingdong Yao[a,b*] ,Yingjin Yuan[a,b]

a. Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering (Ministry of Education), School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China

b. Frontier Research Institute for Synthetic Biology, Tianjin University

c. School of Life Sciences, Faculty of Medicine, Tianjin University

d. Georgia Tech Shenzhen Institute, Tianjin University, Shenzhen 518071, China

***Correspondence:**

Mingdong Yao

Email: mingdong.yao@tju.edu.cn

Table S1. Plasmids used in this study.

Plasmid	Description	Source
pRS416	shuttle vector plasmid,Amp; Ura	This lab
pRS426	shuttle vector plasmid,Amp; Ura	This lab
pZCL096	pRS416-CYC1t-GAL1p-t60TS-RFP-GPM1t	This study
pZCL106	pRS416-CYC1t-GAL1p-t60TS-GPM1t	This study
pZCL116	pRS426-CYC1t-GAL1p-t60TS-GPM1t	This study

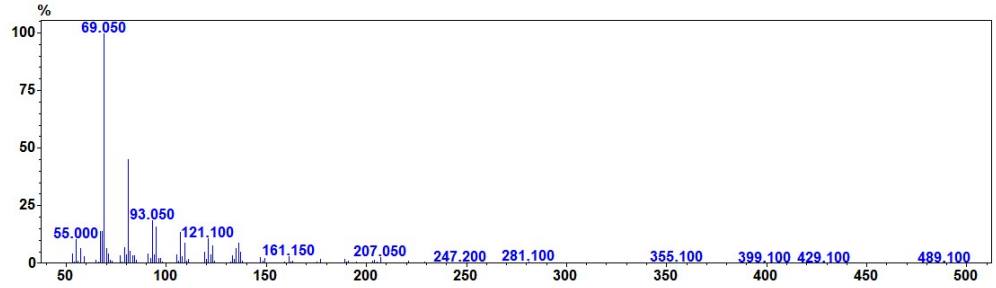
Table S2 Primer used in this study.

Primer name	Sequence (5′-3′)
PRB1-upF	CTCTATTTACACTCGGTGCCCTAGG
PRB1-upR	cttacgatacctgagtattcccacagttATTGTGCAGTGGACACGAGAGGA
PRB1-downF	aacggccttAacgacgtagtcgaAACTGGGGTAAGTGTGTCGACGTT
PRB1-downR	CAAGAATATCTCTCACTTGATCAAAGATTAAATCGGTC
PEP4-upF	GCTTGAAAGCATTATTGCCATTGGC
PEP4-upR	ctgagtattcccacagttAATGTAGGAAACAAGCCAAGGAACC
PEP4-downF	gaaacggccttAacgacgtagtcgaAAAGGATACTGAAAATGGCGGTGAAG
PEP4-downR	TCAAATTGCTTTGGCCAAACCAACC
CYM1-upF	ATGTTGCGGTTTCAGCGATTTGC
CYM1-upR	gtattcccacagttAGGTGGAGTCTAAATAAACACCTCTTAAATTAGC
CYM1-downF	acgacgtagtcgaAAATTCAGGTGTTGAACCAACTACAGCAG
CYM1-downR	CCTGTAACAGTATACCACGTTTAAAGATGTT
YAP3-upF	ATGACGCCTTCTAATATGGATGATAATACCAGC
YAP3-upR	tacctgagtattcccacagttAGGGACTCCTCGAACTTGGAGATATACC
YAP3-downF	cgacgtagtcgaAAAGAATTACAAGATAAATTGTTAGAAAGTGAAAGGAACC
YAP3-downR	TCATTCATTCAAAGTAGCTTCCTCCACTAGA
t60TS-RFP-F	CTAATACTATAACATACAATAATAAATGTCATCATCAACCGGTACTTCAAAGG
t60TS-RFP-R	CCTTTTGAAACGGGTCCGGAAACTTGGATAGGGTCAATATAGACC

Table S3 The Codon-optimized sequences of genes involved in this study.

Genes	Sequences
Tbt60TS from Taxus brevifolia	atgtcatcatcaaccggtacttcaaaggtagtatcagagacttcttctaccatagtagatgatatcccaagattgtctgctaactaccacg gagatttgtggcatcacaacgtaatacaaacattggagacacctttagagagtcttctacataccaggagagggccgatgaattggt cgtcaagatcaaagatatgtttaacgcattaggtgacggtgatatctcaccttctgcctatgacacagcttgggttgctagattggccac aatttcttctgatggatctgaaaagcctaggtttcctcaggcattaaactgggtctttaacaatcaattgcaagacggttcatggggtata gagtctcacttttctttatgcgatagattattgaacactacaaactcagtaatcgctttatctgtctggaaaaccggacactcacaggttca acagggtgccgagttcatagccgaaaatttgaggttattaaacgaggaagatgagttatctcctgacttccaaatcatattcccagcttt gttacagaaagccaaggctttgggtattaatttgccttacgacttgccattcattaaatacttgtcaaccactagggaagctaggttaact gacgtatcagctgccgccgataatatccctgctaatatgttgaatgccttagagggattggaggaagtaatagactggaacaaaatca tgagatttcaatctaaggacggatcattcttgtcttcacctgcatcaactgcttgcgttttgatgaataccggagatgagaagtgctttaca ttcttaaacaacttattggataagttcggaggatgtgtcccttgtatgtattctatcgacttgttggagaggttgtctttagtcgacaacata gaacacttaggtattggtaggcactttaagcaggaaattaagggagctttggattatgtctacagacactggtcagaaagaggaatcg gttggggtagagactcattagtaccagatttaaacaccacagcattgggattgagaacattaaggatgcatggatataacgtttcttca gatgttttgaataatttcaaagacgagaacggtagattcttttcatctgccggtcagacacacgtcgaattaagatcagtagtaaacttgt ttagagcctcagatttggccttcccagacgaaagagccatggatgacgctaggaagttcgctgaaccttacttgagagaggcattgg caactaagatctctactaacacaaaattgttcaaggaaatcgaatacgtcgttgagtacccatggcatatgtcaataccaaggttggag gcaaggtcttacatagactcttatgacgataactatgtttggcagagaaagacattgtatagaatgccatctttgtctaactcaaagtgttt ggaattagccaaattggattttaacatcgtacaatcattgcaccaggaagaattaaagttattgactagatggtggaaggaatcaggtat ggccgacataaacttcacaagacatagggtagccgaggtttacttctcttcagctactttcgagccagagtattctgctacaaggattg ccttcaccaaaataggatgtttacaggtcttgttcgacgatatggcagacatttttgccactttggacgaattgaagtcattcaccgagg gtgttaagagatgggatacttctttgttgcacgaaataccagagtgtatgcaaacatgctttaaggtctggtttaagttaatggaagaggt caacaacgatgtagtcaaagtccaaggaagagacatgttagctcacattagaaaaccatgggagttatactttaactgctacgttcaa gaaagggagtggttagaggctggttacatccctacttttgaggaatatttgaaaacctacgctatttcagtcggtttaggaccttgtacat tgcagcctatcttattgatgggagagttagtaaaggacgacgttgtagaaaaggtccactacccttctaatatgttcgagttggtctcatt gtcatggagattgacaaacgacaccaagacttatcaggccgagaaggctagaggacaacaagcatctggaatcgcatgttacatga aggataacccaggtgccacagaagaggacgctatcaagcatatttgcagagtagttgacagggccttaaaagaagcctcattcgag tatttcaagccttcaaatgacatacctatgggttgtaaatcatttatcttcaacttaagattatgcgttcagatattttacaaattcatagatgg atacggtattgctaacgaagagataaaagactatataaggaaggtctatattgaccctatccaagtttaa

yZCL080



GGOH Standard

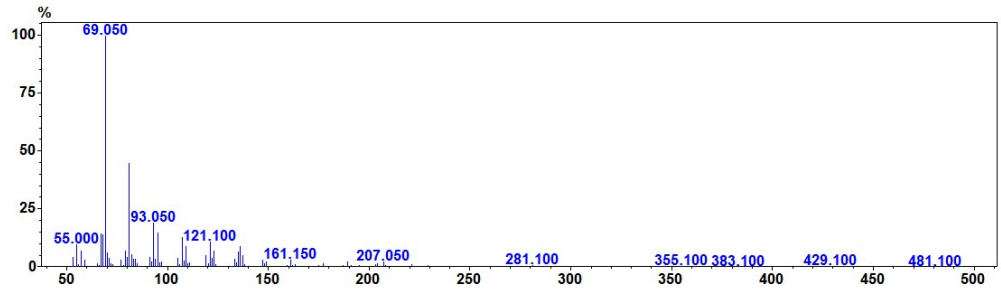


Fig. S1 The MS results for GGOH standard and the production of the strain yZCL080 were also showed.

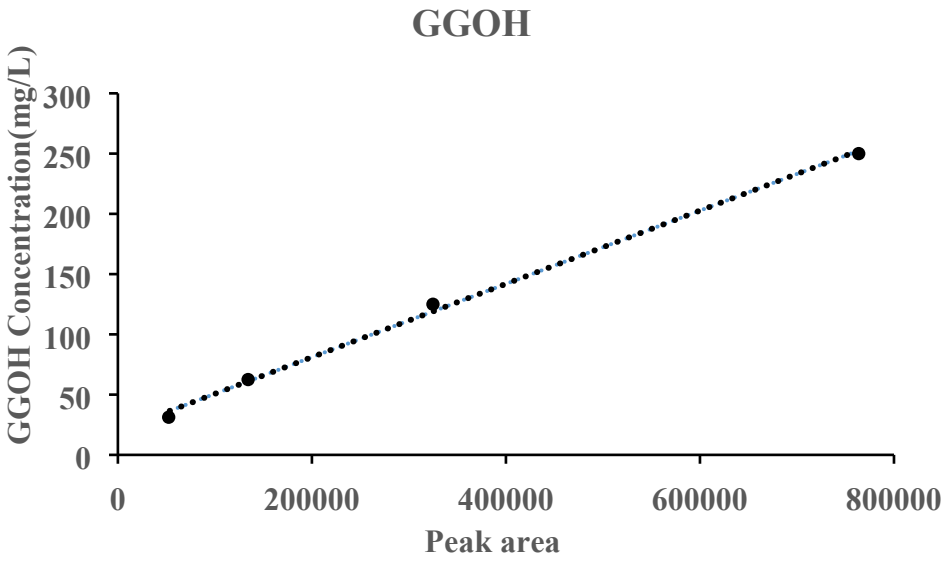


Fig. S2 Calibration curve of GGOH concentration.

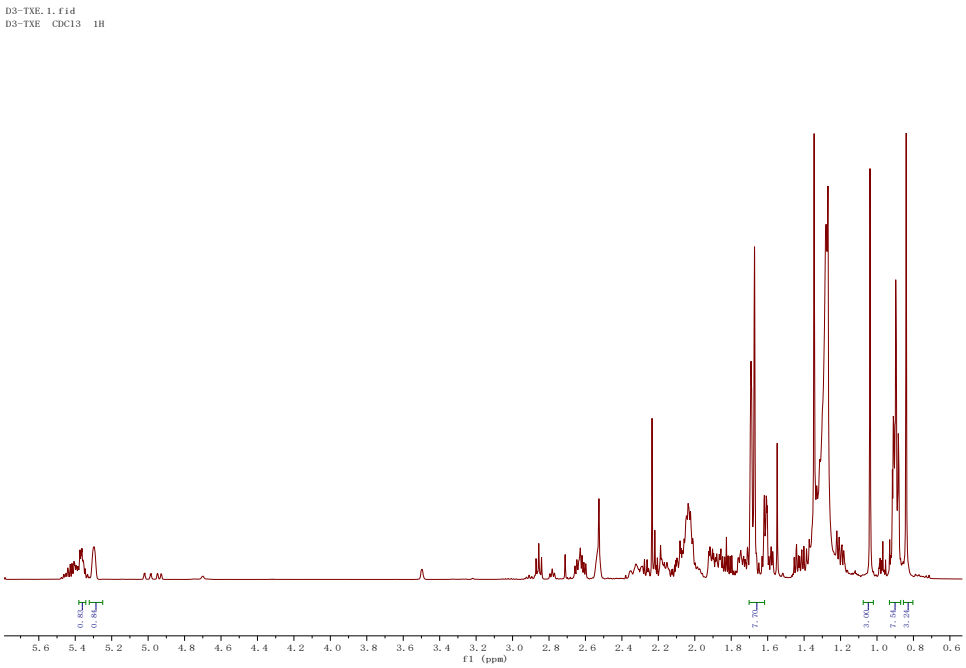


Fig. S3 ^{1}H NMR spectra of purified taxadiene samples

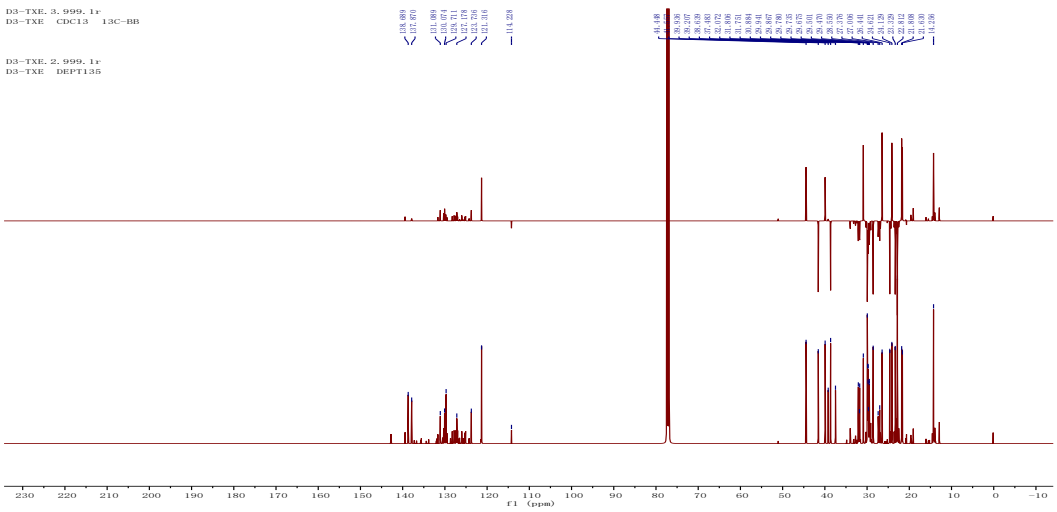


Fig. S4 ^{13}C NMR spectra of purified taxadiene samples

Table S4 The ^{13}C NMR data of purified samples were compared with the literature data[1]

No	Ref1.	Exp1.
1	44.2(d)	44.30(d)
2	28.4(t)	28.41(t)
3	39.8(d)	39.79(d)
4	138.1(s)	138.53(s)
5	121.3(d)	121.16(d)
6	22.7(t)	22.70(t)
7	38.5(t)	38.50(t)
8	37.3(s)	37.34(s)
9	41.4(t)	41.41(t)
10	24.5(t)	24.62(t)
11	137.5(s)	137.71(s)
12	129.3(s)	129.92(s)
13	29.6(t)	29.80(t)
14	23.3(t)	23.19(t)
15	38.9(s)	39.06(s)
16	30.6(q)	30.74(q)
17	26.1(q)	26.30(q)
18	21.2(q)	21.49(q)
19	21.5(q)	21.67(q)
20	23.9(q)	23.99(q)

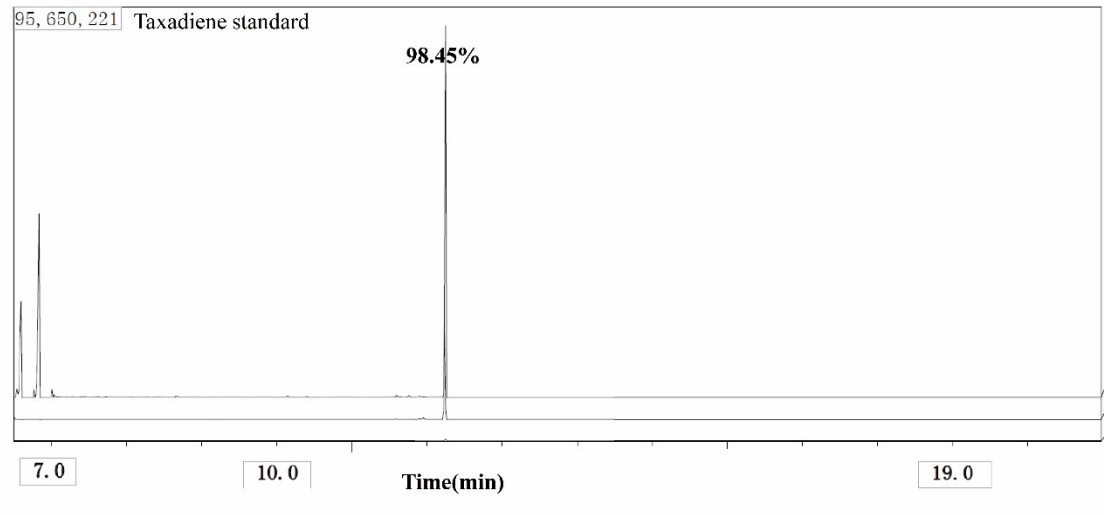


Fig. S5 Purity analysis of taxadiene standard by gas chromatography. The solvent peak appears around 7min.

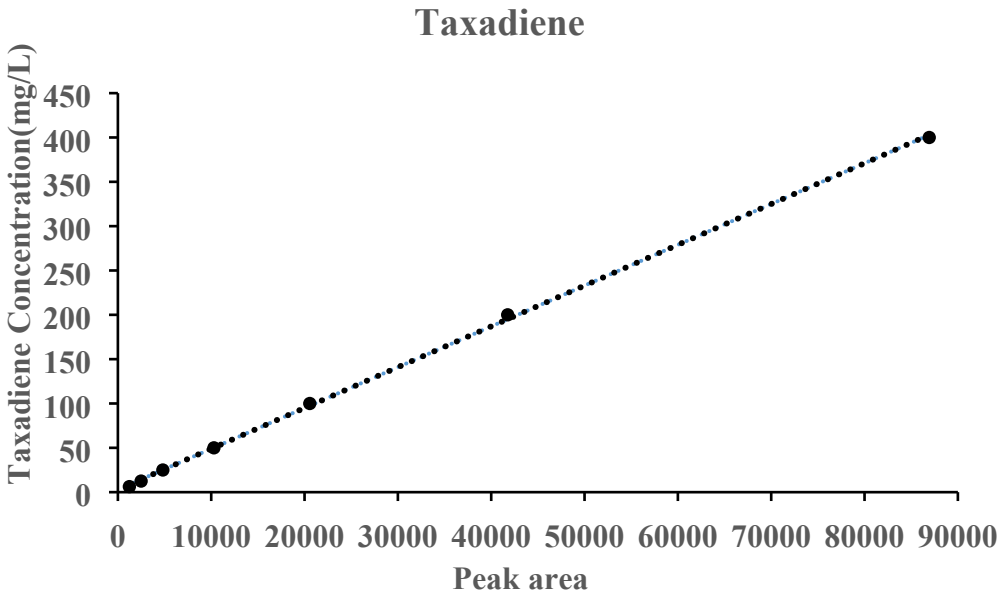


Fig. S6 Calibration curve of taxadiene concentration

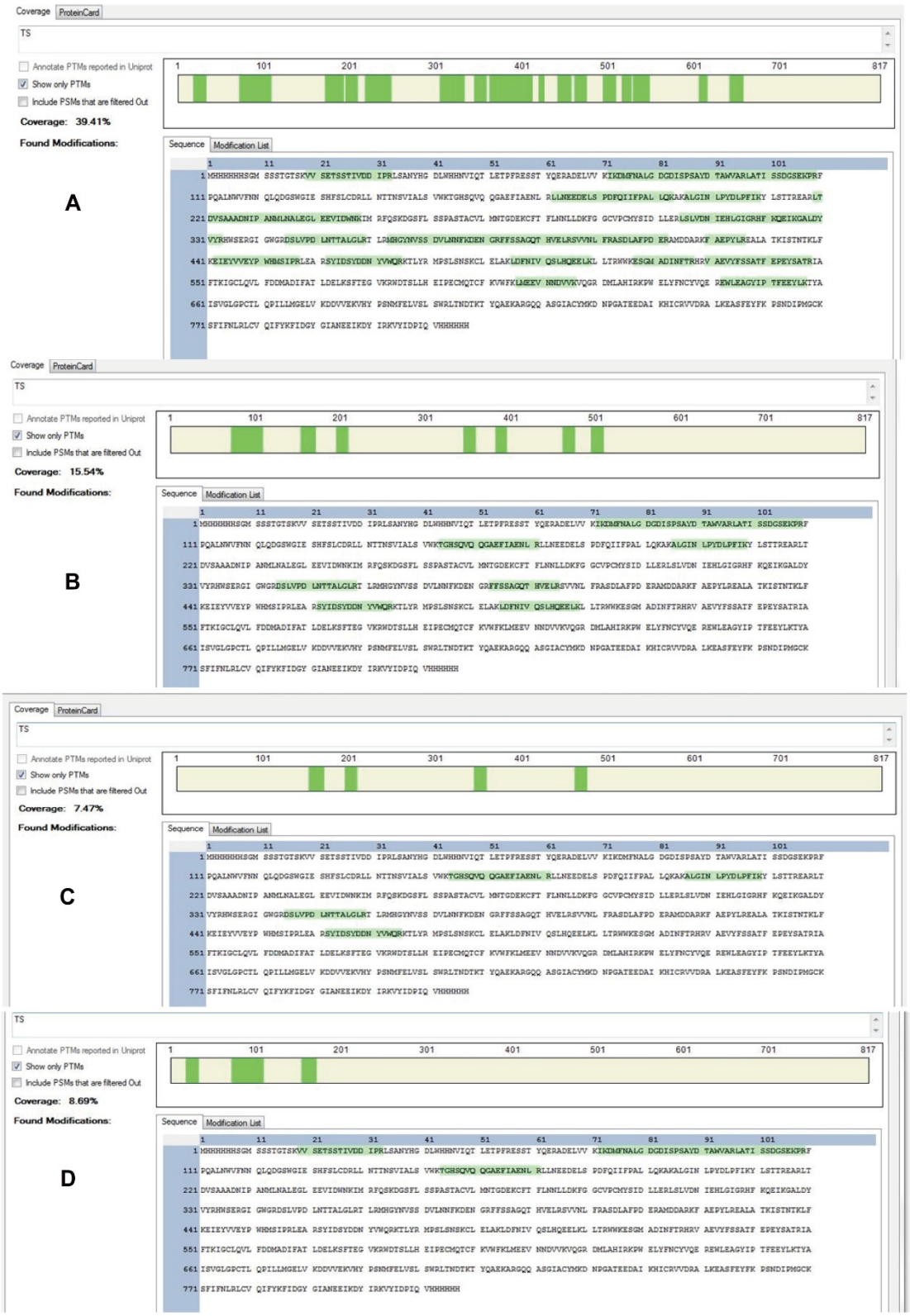

Coverage
ProteinCard
TS
Annotate PTMs reported in Uniprot
Show only PTMs
Include PSMs that are filtered Out
Coverage: 39.41%
Found Modifications:
Sequence
Modification List
A
Coverage: 15.54%
B
Coverage: 7.47%
C
Coverage: 8.69%
D

Fig. S7 Comparative analysis of mass spectrometry results of TS protein. The four protein bands in the SC-TS sample are numbered 1, 2, 3, and 4 from top to bottom, and the protein spectrum segment of each protein band has a corresponding coverage range. **(A)** The first protein band (about 93KDa in size) in SDS-PAGE was detected by TS full-length mass spectrometry. **(B)** The second protein band (about 53KDa in size) in SDS-PAGE was detected by mass spectrometry. **(C)** The third protein strip (about 35KDa in size) in SDS-PAGE was used to detect the matching results of peptide by mass spectrometry. **(D)** The fourth protein band (about 15KDa in size) in SDS-PAGE was detected by mass spectrometry.

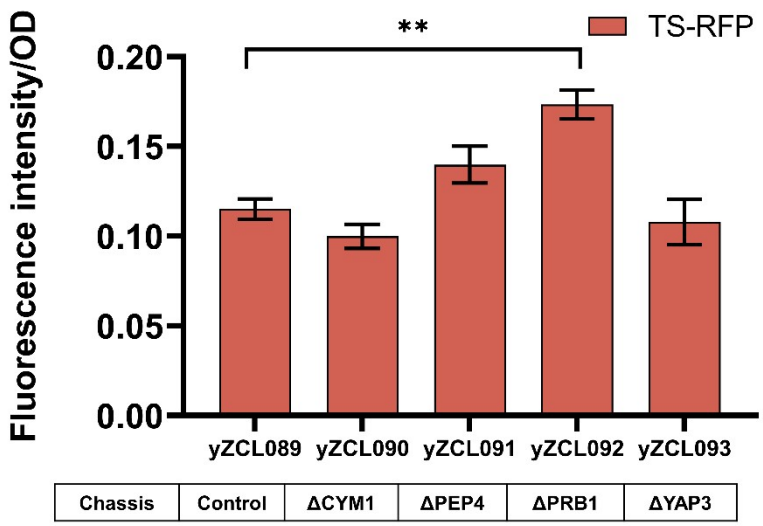


Fig. S8 Effect of endogenous protease knockout on TS expression. The single-copy pZCL096 plasmid carrying the TS-RFP fusion expression gene was introduced into control strains, and endogenous proteases PEP4, CYM1, YAP3, and PRB1 single knockout strains and yZCL089, yZCL090, yZCL091, yZCL092, and yZCL093 were successfully constructed.

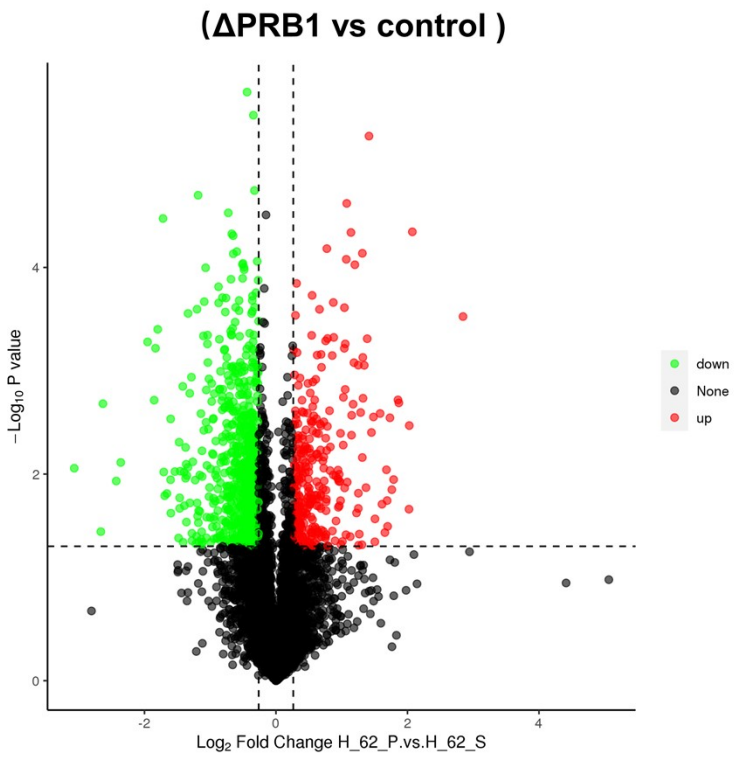


Fig. S9 Differential protein volcano map between ΔPRB1 and control group. The horizontal coordinate indicates the difference multiple of the differential protein (log2 value), the vertical axis indicates the *P*-value (-log10 value), the black represents the non-significantly different protein, the red represents the up-regulated protein, and the green represents the down-regulated protein.

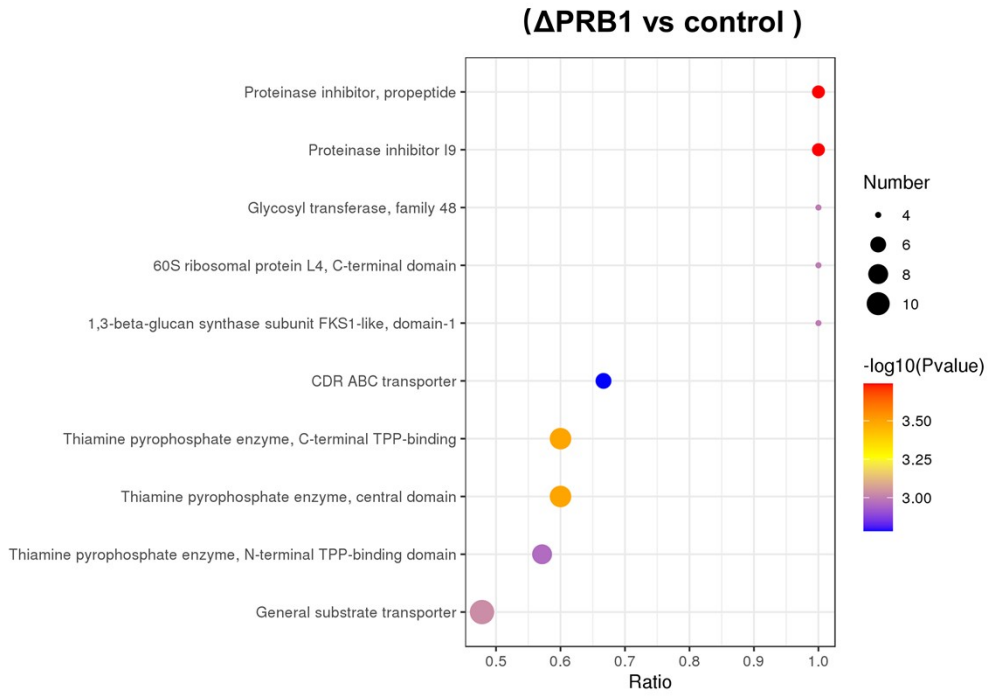


Fig. S10 Bubble map of IPR (domain) enrichment analysis. The horizontal coordinate is the ratio of the number of different proteins in the corresponding domain to the total number of proteins identified in the domain. The color of the dot represents the *p*-value of the hypergeometric test. The redder the color, the smaller the *p*-value, indicating the greater the reliability and statistical significance of the test. The size of the dot represents the number of different proteins in the corresponding domain, and the larger the dot, the more different proteins in that domain.

Ref 1. Q. Jin, D. C. Williams, M. Hezari, R. Croteau and R. M. Coates, J. Org. Chem., 2005, 70, 4667–4675.