Supplemental Information

Engineering the Biocatalytic Selectivity of Iridoid Production in Saccharomyces

cerevisiae

John M. Billingsley^a, Anthony B. DeNicola^a, Joyann S. Barber^b, Man-Cheng Tang^a, Joe

Horecka^{c,d}, Angela Chu^{c,d}, Neil K. Garg^b, Yi Tang^{a,b,*}

^a Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, California 90095, United States

^b Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States

^c Stanford Genome Technology Center, Stanford University, Palo Alto, CA, USA

^d Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA

The Supplemental Information consists of 2 Supplementary Methods, 22 Supplementary Figures and 4 Supplementary Tables.

*Corresponding author. E-mail address: yitang@ucla.edu (Y. Tang)

Supplementary Method 1. Construction of JHY651

In order to improve the capacity of BY4742 (Brachmann et al., 1998) for expression of heterologous natural product pathways, a number of genomic changes were implemented. Six previously identified quantitative trait loci (QTLs) were repaired to restore sporulation (MKT1(30G) RME1(INS-308A) TAO3(1493Q)) (Deutschbauer and Davis, 2005) and mitochondrial genome stability (SAL1+ CAT5(91M) MIP1(661T)) (Dimitrov et al., 2009) in BY4742. Additionally, a Ty1 element which inactivates the HAP1 regulatory gene in S288c derived strains (Gaisne et al., 1999) was repaired. The resulting strain, DHY214 was further modified by deletion of PRB1 and PEP4, which mediate the degradation of heterologous proteins (Jones, 1991), resulting in the parent strain used in this study, JHY651 (S1). Strain modifications were performed using the previously described 50:50 method (Horecka and Davis, 2014).

Supplementary Method 2. Synthesis of 8-hydroxygeraniol



Synthesis of **10**: To a solution of selenium dioxide (0.68 g, 6.12 mmol, 0.4 equiv) in CH₂Cl₂ (77 mL) was added *tert*-butyl hydroperoxide (5.5 M in decane, 8.6 mL, 47.5 mmol, 3.1 equiv). The solution was then cooled to 0 °C and geranyl acetate (**9**, 3.3 mL, 15.3 mmol, 1.0 equiv) was added. After stirring for 7 h at 0 °C, EtOAc (200 mL) was added and the reaction was transferred to a separatory funnel. The layers were separated and the organic layer was washed successively with deionized water (2 x 75 mL), saturated aqueous NaHCO₃ (1 x 75 mL), deionized water (1 x 50 mL), and brine (1 x 80 mL). The organic layer was dried over MgSO₄, filtered, and concentrated under reduced pressure. The resulting crude oil was purified by flash chromatography (9:1 \rightarrow 1:1 hexanes:EtOAc) to afford allylic alcohol **10** (1.99 g, 61% yield) as a light yellow oil. Allylic alcohol **10**: R_f 0.44 (2:1 hexanes:EtOAc). Spectral data (Fig. S11) match those previously reported (Bogazkaya et al., 2014).



Synthesis of 8-hydroxygeraniol 1: To a solution of allylic alcohol **10** (1.99 g, 9.39 mmol, 1.0 equiv) in MeOH (60 mL) was added potassium hydroxide (1.56 g, 11.3 mmol, 1.2 equiv) as a solid in one portion. After stirring for 2.5 h at 23 °C, deionized water (40 mL) was added and the reaction was transferred to a separatory funnel. The layers were separated and the aqueous layer was extracted with Et₂O (3 x 80 mL). The combined organic layers were washed successively with 0.5 M HCl (1 x 50 mL), saturated aqueous NaHCO₃ (1 x 100 mL), brine (1 x 50 mL), and deionized water (1 x 50 mL). The organic layers were then dried over MgSO₄, filtered, and concentrated under reduced pressure. The resulting crude oil was purified by flash chromatography (1:1 hexanes:EtOAc) to afford 8-hydroxygeraniol (1, 1.29 g, 81% yield) as a light yellow oil. 1: R_f 0.14 (2:1 hexanes:EtOAc). Spectral data (Fig. S12) match those previously reported (Bogazkaya et al., 2014).



Supplementary Fig. 1. The monoterpene indole alkaloid biosynthetic pathway.

Supplementary Fig. 1 (continued). The monoterpene indole alkaloid biosynthetic pathway. 8-Hydroxygeraniol is converted to the dialdehyde by geraniol oxidoreductase (GOR). 8-oxogeranial undergoes reductive cyclization by iridoid synthase (ISY). The iterative P450 iridoid oxidase (IO) converts the C-4 methyl into a carboxy group. 7-deoxyloganitic acid glucosyltransferase (7-DLGT) stabilizes the hemiacetal. The P450 7-deoxyloganic acid hydroxylase (7-DLH) installs the hydroxyl, followed by O-methylation by loganic acid O-methyltransferase (LAMT). Secologanin synthase (SLS) then performs the oxidative cleavage of loganin. Meanwhile, tryptophan decarboxylase (TDC) produces the β -arylethylamine tryptamine. The Pictet-Spenglerase strictosidine synthase (STR) catalyzes the condensation of tryptamine with secologanin followed by ring closure to forge strictosidine.

> codon optimized ISY gene sequence

Supplementary Fig. 2. Gene sequences of GOR and ISY synthesized by Gen9.

> codon optimized GOR gene sequence



Supplementary Fig. 3. Monoterpene standard curve.



Supplementary Fig. 4. SDS-PAGE gel of purified enzymes used for *in vitro* assays.



Monoterpene profile in engineered strains

Supplementary Fig. 5. Monoterpene profile after bioconversion assay. (i) S1 with pJB031, (ii) S1 with pJB033, (iii) S4 with pJB033, (iv) S4 with pJB034.



Supplementary Fig. 6. GOR and OYE cofactor dependence, (i) 8-hydroxygeraniol standard, (ii) 8-hydroxytetrahydrogeraniol standard, (iii) 8-hydroxygeraniol + GOR + OYE2, without NAD⁺, (iv) 8-hydroxygeraniol + OYE2, without NADPH, (v) 8-hydroxygeraniol + OYE3, without NAD⁺, (vi) 8-hydroxygeraniol + GOR + OYE2, without NADPH.



Supplementary Fig. 7. ARI1, ADH6, ADH7 cofactor dependence. (i) 8hydroxygeraniol + ARI1 + NAD⁺, (ii) 8-hydroxygeraniol + ADH6 + NAD⁺, (iii) 8hydroxygeraniol + ADH7 + NAD⁺, (iv) 8-hydroxygeraniol + ARI1 + NADP⁺, (v) 8hydroxygeraniol + ADH6 + NADP⁺, (vi) 8-hydroxygeraniol + ADH7 + NADP⁺.



Supplementary Fig. 8. Mass spectra of compound 7 produced from 4 *in vitro* through the action of GOR and OYE.



Supplementary Fig. 9. Mass spectra of compound 8 produced from 4 *in vitro* through the action of ISY and ADH.



Supplementary Fig. 10. Additional fermentation data. Means and standard errors are reported for plasmid stability.

Supplementary Table 1. Primers used in this study.

	1001.9.10		
Category	tegory Primer Sequence		Amplicon
GOR	R P01 CAAAACGTAGGGGCAAACAAACG		ADH2p
expression	P02	GGTATTACGATATAGTTAATAGTTGATAGTTGATTG	. '
cassette	P03	AACTATCAACTATTAACTATATCGTAATACCATGACTAAA	GOR
assembly		ACTAATTCTCCAGCCCCATC	
	P04	TGTGCTAGTGTCTCCCGTCTTCTGTCTTAGAACTTAATAA	
	_	CAACTTTGACACAGTCTGGG	
	P05	GACAGAAGACGGGAGACACTAGCAC	PRM9t
	P06	GGCATTTTCAACATCGTATTTTCCGAAGC	
ISY	P07	CAATAGGAAAAAACCGAGCTTCCTTTC	PCK1p
expression	P08	GTTGTTATTTATTATGGAATAATTAGTTGCGTG	
cassette	P09	ACTCACGCAACTAATTATTCCATAATAAAATAACAACATG	ISY
assembly		TCCTGGTGGTGGAAAAGGTC	
,	P10	AATCTTTGACTATTCAATCATTGCGCTTATGGAATGAATC	
		TGTAGTCTCTCATCTTGTCG	
	P11	GCGCAATGATTGAATAGTCAAAGATT	CPS1t
	P12	ATTTGACACTTGATTTGACACTTCTTT	
pJB031	P13	GCGGCCGCAACGAAGCATCTGTGCTTCATTTGTAG	NotI-pJB031
assembly	P14	GCGGCCGCGTTTTTCCATAGGCTCCG	pJB031-Notl
p.IB033	P15	TGCTCGTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	ADH2n-
assembly	1.10	CAAAACGTAGGGGCAAACAAACG	GOR-PRM9t
accombry	P16	GTTCTACAAAATGAAGCACAGATGCTTCGTTGGCATTTT	
	1.10	CAACATCGTATTTTCCGAAGC	
	P17	GATCCGATGATAAGCTGTCAAACATGAG	LIRA3
	P18	GCGTTTTTCCATAGGCTCCG	
	P19		20
	P20	GAAGCATATTTGAGAAGATGCGGC	- - - P
n.IB034	P21		ADH2n-
assembly	121		GOR-PRM9t
assembly	P22	GGATGAAAGGAAGCTCGGTTTTTTCCTATTGGGCATTTT	
	1 22	CAACATCGTATTTTCCGAAGC	
	P23	TTGTTCTACAAAATGAAGCACAGATGCTTCGTTATTTGAC	PCK1p-ISY-
	1 20	ACTTGATTTGACACTTCTTT	CPS1t
	P24	GATCCGATGATAAGCTGTCAAACATGAG	URA3
	P25	GCGTTTTTCCATAGGCTCCG	
	P26		20
	P27	GAAGCATATTTGAGAAGATGCGGC	_ <u>~</u> µ
n IB097	P28		kanMX-I FU2
assembly	1 20		
assembly	P29		
	125	CGCGCCAACAAATATATTGC	
	P30		1 FU2-2u
	1.00	AAATAGCAAGTTAAAATAAGG	
	P31		
	101	GACGAAAGGGCCTCGTG	
	P32		iCas9
	P33		10030
n IB0/12	P3/		GOR
assembly	P35		
accombry	1 33	AGTCTGGG	
n IB0/13	P36		197
assembly	1 30	GTCG	
accombry	1		1

pJB043	P37	GGTGCTCGAGTTGGAATGAATCTGTAGTCTCTCATCTTG	
assembly		TCG	
pJB044	P38	TAATGCTAGCATGCCATTTGTTAAGGACTTTAAGCC	OYE2
assembly	bly P39 TAATGAGCTCTTAATTTTTGTCCCAACCGAGTTTTAGAGC		
pJB045	P40	TAATGCTAGCATGCCATTTGTAAAAGGTTTTGAGCC	OYE3
assembly	P41	TAATGAGCTCTCAGTTCTTGTTCCAACCTAAATCTACTGC	
pJB072	P42	TAATGCTAGCATGTCTTATCCTGAGAAATTTGAAGGTATC	ADH6
assembly		G	
	P43	TAATGAGCTCCTAGTCTGAAAATTCTTTGTCGTAGCC	
pJB071	P44	TAATGGTCTCGCTAGCATGCTTTACCCAGAAAAATTTCA	ADH7
assembly		GG	
	P45	TAATGAGCTCCTATTTATGGAATTTCTTATCATAATCGAC	
		CAAAG	
pJB095	P46	TAATGCTAGCATGACTACTGATACCACTGTTTTCGT	ARI1
assembly	P47	GCATACAATCAACTATCAACTATTAACTATATCGTAATAC	
		CATGGGCAGCAGCCATCATC	
	P48	TGTGTGCTAGTGTCTCCCGTCTTCTGTCCTTAGGCTTCA	
		TTTTGAACTTCTAACATTTGC	
S6		AGTAATTGTGCATTGTACAACTGTGCTAAACAGACTTAAA	
construction	P49	AAAGTAATAATTTAACCATTATTTTTTCCTCAACATAACG	ARI1p-LEU2
		AGAACACAC	
	D50	CTATAGATTTGCCTATTGGAGTGATCAAAAAAAACTTCAA	
	FJU	TTAGCGTGATACGGATTTTCTTAACTTCTTCGGCGACAG	LL02-ANT
S8		ATATTATTGAAATAAAAAAAAAAAAAAAAAAAAAAAAAA	
construction	P51	TACAAAAAAAAAAAACCATTATTTTTTCCTCAACATAACG	ADH7p-LEU2
		AGAACACAC	
	P52	TTCCGAGATTTGACATGCATTTTAAGAGATTCTGAAAAAT	
	T JZ	ATTACGTATATAGGATTTTCTTAACTTCTTCGGCGACAG	

Supplementary Table 2. gBlocks used for HICRISPR

gBlock	Sequence
∆oye2	CCTTTGGTCTCACCAAAACATAGCCTTGAAAATCTTGGTCCATTCTTTAATTTGTTCTTCGG
	ACCAGATCATTGTCGTAACCCCCAGATTGTGGAGAGGGAAAGGTACCTTCAGTGATATGT
	TCTTCGGACCAGATACCGTTTTAGAGAGAGACCTTTC
∆oye2	CCTTTGGTCTCACCAAAACAGACCTGGTACCATGATCATCACGGAAGGTACGTTTATTTCC
	CCTCAAGCATGACAACGCCCCTGGGATTTGGTCTGATGAGCAGGTCGCTGAGTGGAAGT
	TTATTTCCCCTCAAGCCGGGTTTTAGAGAGAGACCTTTC
∆gre2	CTTTGGTCTCACCAAAACTCCTTGCCGTGCTTTTGGAAAACATGGTCAAATGCGTCCAGCT
	TAGATATGGCCAGAGGTATAGACATAGCCAACTTCCATGGAGAATTTTGGGTTGTTACCAA
	AGGCCTCCGTTAAATTCGCGTCCAGCTTAGATATGTCGTTTTAGAGAGAG
∆adh6	CTTTGGTCTCACCAAAACAAGTTTGTTACCACATACAGTCAGCCTTATGAAGACGGCTATG
	TGTCGCAGACGTCTCACGGATCGTATATATGCAAACTACGTCAGAGTTCATGAACATTTTG
	TGGTGCCTATCCCAGAGGACGGCTATGTGTCGCAGGGGTTTTAGAGAGAG



Supplementary Fig. 11 1H NMR of 10.



Supplementary Fig. 12 1H NMR of 1.

Supplementary Table 3. NMR data for compound **4**. ¹H NMR spectrum (500 MHz), ¹³C NMR spectrum (125 MHz), CDCI₃

OH но′

compound 4

ОН HO

Key HMBC Correlations

Position	δ _H , mult (<i>J</i> in Hz)	δc	COSY
1	3.69, m	61.33	H2
2	1.37, m, overlap	40.09	
	1.59, m, overlap		
3	1.56, m, overlap	29.46	
4	1.14, m, overlap	37.37	
	1.29, m, overlap		
5	1.32, m, overlap	24.38	
6	1.08, m, overlap	33.38	
	1.37, m, overlap		
7	1.61, m, overlap	35.85	
8	3.43, m	68.53	H7
	3.50, m		
9	0.89, d (6.5)	19.76	H3
10	0.91, d (6.7)	16.70	H7

Supplementary Table 4. NMR data for the isolated saturated acid. ¹H NMR spectrum (500 MHz), ¹³C NMR spectrum (125 MHz), CDCl₃



8-hydroxy-2,6-dimethyloctanoic acid

но ĠН

Key HMBC Correlations

Position	δ_{H} , mult (J in Hz)	δ _C	COSY
1	3.68, m	61.28	H2
2	1.35, m, overlap	39.91	
	1.58, m, overlap		
3	1.56, m, overlap	29.39	
4	1.15, m, overlap	36.93	
	1.32, m, overlap		
5	1.32, m, overlap	24.62	
6	1.39, m, overlap	33.83	
	1.65, m, overlap		
7	2.47, m	39.27	H6, H10
8		181.53	
9	0.89, d (6.4)	19.74	H3
10	1.18, d (7.0)	17.06	H7

Position	δ_H , mult (J in Hz)	δc	COSY
1		177.89	
2	2.16, m	41.48	H3
	2.34, m		
3	1.97, m	30.19	H2, H9
4	1.15, m, overlap	37.01	
	1.32, m, overlap		
5	1.32, m, overlap	24.34	
6	1.09, m, overlap	33.18	
	1.39, m, overlap		
7	1.60, m, overlap	35.78	
8	3.44, m	68.46	H7
	3.50, m		
9	0.97, d (6.7)	19.87	H3
10	0.91, d (6.9)	16.67	H7

.OH но

8-hydroxy-3,7-dimethyloctanoic acid

ОН

Key HMBC Correlations







Supplementary Fig. 15. ¹H-¹H COSY Spectrum of Compound 4 in CDCl₃.









Supplementary Fig. 18. ¹H NMR Spectrum of the isolated saturated acid in CDCl₃.



Supplementary Fig. 21. ¹³C NMR Spectrum of the isolated saturated acid in CDCl₃.



Supplementary Fig. 20. ¹H-¹H COSY Spectrum of the isolated saturated acid in CDCl₃.



Supplementary Fig. 21. HSQC Spectrum of the isolated saturated acid in CDCl3.



Supplementary Fig. 22. HMBC Spectrum of the isolated saturated acid in CDCl₃.

Supplemental Information References

- Bogazkaya, A. M., von Buhler, C. J., Kriening, S., Busch, A., Seifert, A., Pleiss, J., Laschat, S., Urlacher, V. B., 2014. Selective allylic hydroxylation of acyclic terpenoids by CYP154E1 from *Thermobifida fusca* YX. Beilstein journal of organic chemistry. 10, 1347-1353.
- Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., Boeke, J. D., 1998. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast. 14, 115-32.
- Deutschbauer, A. M., Davis, R. W., 2005. Quantitative trait loci mapped to singlenucleotide resolution in yeast. Nature genetics. 37, 1333-40.
- Dimitrov, L. N., Brem, R. B., Kruglyak, L., Gottschling, D. E., 2009. Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of *Saccharomyces cerevisiae* S288C strains. Genetics. 183, 365-83.
- Gaisne, M., Becam, A. M., Verdiere, J., Herbert, C. J., 1999. A 'natural' mutation in *Saccharomyces cerevisiae* strains derived from S288c affects the complex regulatory gene HAP1 (CYP1). Current genetics. 36, 195-200.
- Horecka, J., Davis, R. W., 2014. The 50:50 method for PCR-based seamless genome editing in yeast. Yeast. 31, 103-12.
- Jones, E. W., 1991. Tackling the protease problem in *Saccharomyces cerevisiae*. Methods in enzymology. 194, 428-53.