Supplementary Information

Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control

Effendi Leonard, Parayil Kumaran Ajikumar, Kelly Thayer, Wen-Hai Xiao,

Jeffrey D. Mo, Bruce Tidor, Gregory Stephanopoulos,

Kristala L. J. Prather

Supplementary Table 1. Diterpenoid production (mg/L) from *E. coli* expressing wild-type GGPPS and LPS, with variable overexpression of genes in the MEP pathway as operons *dxs*, *idi*, *isp*D, and *isp*F (encoding for 1-deoxy-D-xylulose-5-phosphate synthase, isopentenyl diphosphate isomerase, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, and 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, respectively).

Upstream pathway overexpression	**Copy number of *dxs*, *idi*, *isp*D, *isp*F operon**	**Total diterpenoid titer (mg/L)**	**Fold improvement**
Wild type	0	0.15	1
Ch1TrcMEP	1	0.63	4.2
p5TrcMEP	~5	91.70	611.33
p10TrcMEP	~10	26.53	176.87

In addition to the native copy of the enzymes, additional copies of the genes were contained in the strains as operon *dxs*-*idi*-*isp*D-*isp*F under a Trc promoter. The single copy MEP strain was constructed by chromosomal localization of the MEP pathway and higher copy numbers were achieved through plasmid-based expression. Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Table 2. Diterpenoid production from precursor-amplified *E. coli* strains harboring GGPPS and LPS mutants (phylogenetically-based mutations)

LPS mutation	Titer (mg/L)
WT	26.5
M593I	98.6
C618N	4.0
L619F	12.7
A620T	33.9
L696Q	42.0
Y700H	11.5
K723S	48.0
A727S	0.0
A729G	0.6
V731L	2.3
N769A	0.0
E777A	0.0
N838E	58.5
G854T	36.8
I855L	17.7

WT, wild type LPS.

Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Table 3. Diterpenoid production from precursor-amplified *E. coli* strains harboring GGPPS and LPS mutants (saturation mutagenesis of Met593)

LPS Met593 mutation	Product Selectivity (%)			Titer (mg/L)
	1	2	3	
WT	87	11	2	26.5
Ala	65	26	9	16.2
Cys	65	28	7	48.4
Asp	51	15	34	10.3
Glu	30	7	63	9.7
Phe	ND	ND	100	3.8
Gly	67	23	10	1.3
His	ND	ND	ND	0.0
Ile	84	12	5	98.6
Lys	ND	ND	ND	0.0
Leu	80	13	7	55.2
Asn	75	11	14	40.4
Pro	49	8	43	12.6
Gln	43	TA	57	8.6
Arg	TA	ND	100	4.6
Ser	80	11	9	39.8
Thr	78	16	6	40.2
Val	71	22	6	18.2
Trp	ND	ND	ND	0.0
Tyr	ND	ND	ND	0.0

WT, wild type LPS; TA, trace amounts (< 0.1%); ND, not detected.

Levopimaradiene, 1; abietadiene, 2; sandaracopimaradiene, 3. Neoabietadiene is not included in the table because it was only produced in trace amounts in all strains.

Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Table 4. Diterpenoid production from precursor-amplified *E. coli* strains harboring GGPPS and LPS mutants (saturation mutagenesis of Tyr700)

LPS Tyr700 mutation	Product Selectivity (%)			Titer (mg/L)
	1	2	3	
WT	87	11	2	26.5
Ala	81	9	10	75.6
Cys	76	7	17	59.1
Asp	81	ND	19	13.3
Glu	64	7	29	46.3
Phe	79	16	5	133.5
Gly	79	6	15	36.1
His	74	ND	26	6.2
Ile	60	6	34	31.5
Lys	TA	ND	100	4.3
Leu	72	7	21	48.2
Met	80	13	7	132.8
Asn	70	9	21	60.2
Pro	56	ND	44	8.3
Gln	59	6	35	41.2
Arg	33	ND	67	5.0
Ser	78	7	15	84.9
Thr	72	7	21	65.4
Val	56	6	38	31.3
Trp	84	6	10	100.7

WT, wild type LPS; TA, trace amounts (< 0.1%); ND, not detected.

Levopimaradiene, 1; abietadiene, 2; sandaracopimaradiene, 3. Neoabietadiene is not included in the table because it was only produced in trace amounts in all strains.

Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Table 5. Diterpenoid production from precursor-amplified *E. coli* strains harboring GGPPS and LPS mutants (saturation mutagenesis of Ala620)

LPS Ala620 mutation	Product Selectivity (%)			Titer (mg/L)
	1	2	3	
WT	87	11	2	26.5
Cys	87	10	3	40.0
Asp	97	2	1	1.7
Glu	ND	ND	ND	0.0
Phe	ND	ND	ND	0.0
Gly	86	12	2	29.2
His	100	ND	ND	0.9
Ile	ND	ND	ND	0.0
Lys	ND	ND	ND	0.0
Leu	97	2	1	7.3
Met	ND	ND	ND	0.0
Asn	97	2	1	2.6
Pro	ND	ND	ND	0.0
Gln	ND	ND	ND	0.0
Arg	ND	ND	ND	0.0
Ser	92	8	TA	42.2
Thr	87	11	2	33.9
Val	92	8	TA	42.5
Trp	ND	ND	ND	0.0
Tyr	ND	ND	ND	0.0

WT, wild type LPS; TA, trace amounts (< 0.1%); ND, not detected.

Levopimaradiene, 1; abietadiene, 2; sandaracopimaradiene, 3. Neoabietadiene is not included in the table because it was only produced in trace amounts in all strains.

Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Table 6. Diterpenoid production from precursor-amplified *E. coli* strains harboring GGPPS and LPS M593I mutants (saturation mutagenesis of Tyr700)

LPS (M593I) Tyr700 mutation	Product Selectivity (%)			Titer (mg/L)
	1	2	3	
WT	87	11	2	26.5
Ala	91	2	7	167.8
Cys	97	1	2	155.7
Asp	85	TA	15	66.1
Glu	60	TA	40	92.4
Phe	85	9	6	273.8
Gly	80	TA	20	59.2
His	70	TA	30	63.8
Ile	65	TA	35	78.6
Lys	ND	ND	ND	0.0
Leu	73	2	25	97.4
Met	89	2	9	132.1
Asn	79	TA	21	93.7
Pro	0	TA	ND	0.0
Gln	60	TA	40	26.9
Arg	ND	ND	ND	0.0
Ser	84	TA	16	48.1
Thr	73	TA	27	22.7
Val	91	TA	9	83.5
Trp	63	TA	37	51.6

WT, wild-type LPS; TA, trace amounts (< 0.1%); ND, not detected.

Levopimaradiene, 1; abietadiene, 2; sandaracopimaradiene, 3. Neoabietadiene is not included in the table because it was only produced in trace amounts in all strains.

Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

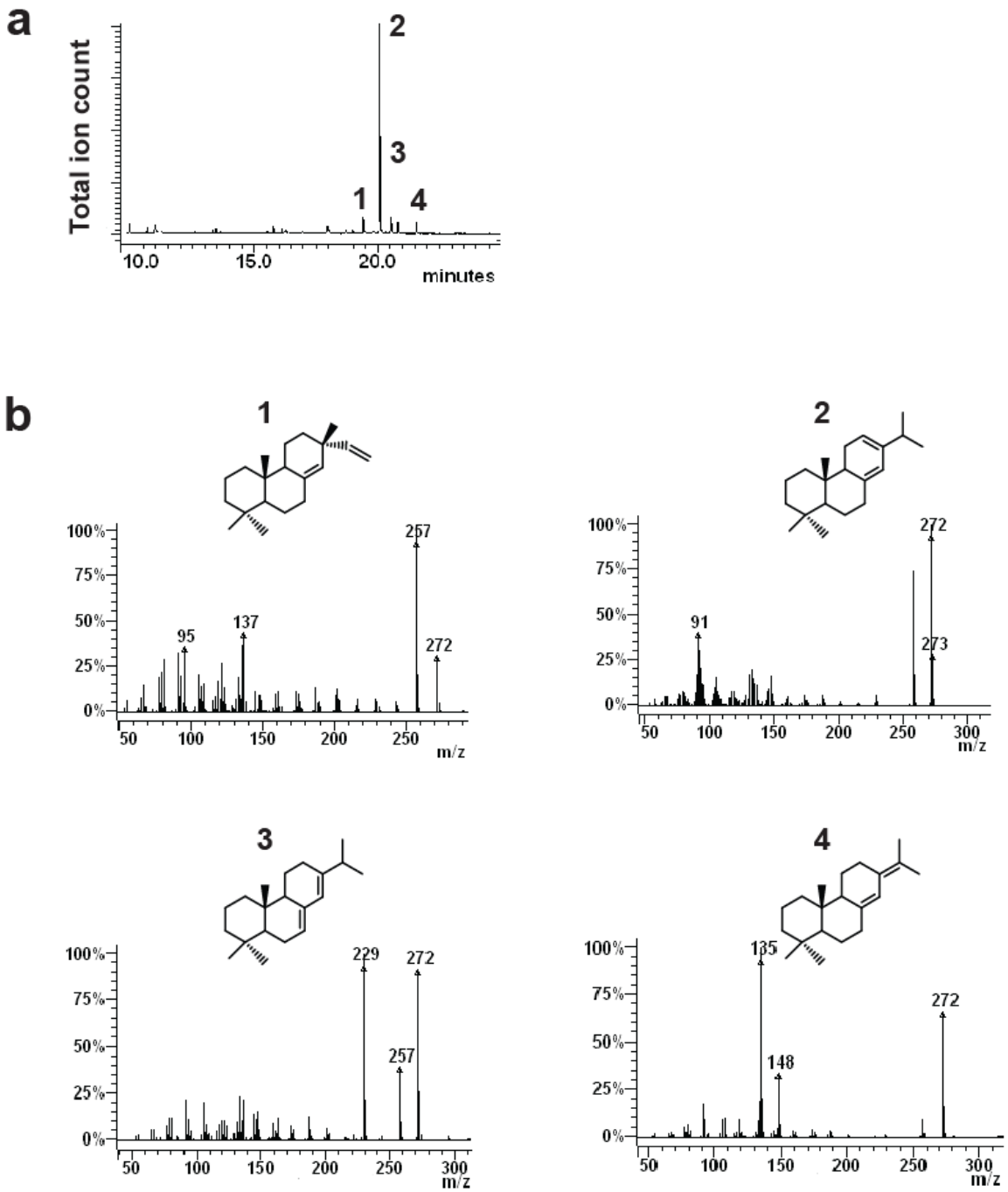
Supplementary Table 7. Diterpenoid production from precursor-amplified *E. coli* strains harboring isolated GGPPS mutants and the LPS M593I/Y700F variant

GGPPS mutation	Titer (mg/L)
WT	273.8
G1	261.3
G2	396.2
G3	343.2
G4	316.8
G5	257.1
G6	242.4
G7	380.9
G8	351.0
G9	350.7
G10	468.7
G11	211.8
G12	366.8
G13	411.1
G14	287.2
G15	406.7

WT, wild type GGPPS.

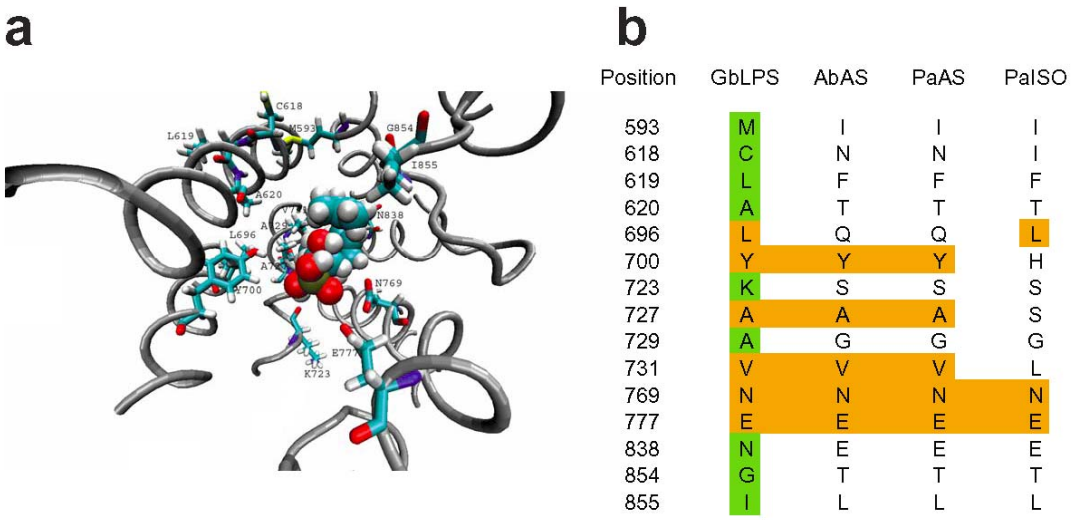
Quantification of production was determined based on sampling an average of three independent *E. coli* colonies harboring the mutated pathway. Product selectivity of all mutant pathways remained similar (within ± 1%) to the pathway harboring wild-type GGPPS and LPS(M593I/Y700F). Standard deviations were lower than 5%. Cell density (A_{600}) in all samples was 5 ± 0.5.

Supplementary Figure 1. GC-MS chromatogram of diterpenoid products from the engineered *E. coli* strain.



(a) Total ion chromatogram and retention time of the diterpenoids secreted in the culture media used to cultivate the pre-engineered *E. coli* strain expressing the wild type GGPPS and LPS. (**b**) GC-MS spectra of the product peaks corresponding to sandaracopimaradiene (**1**), levopimaradiene (**2**), abietadiene (**3**), and neoabietadiene (**4**) as previously reported in literatures(1, 2).

Supplementary Figure 2. Probing putative LPS binding pocket.



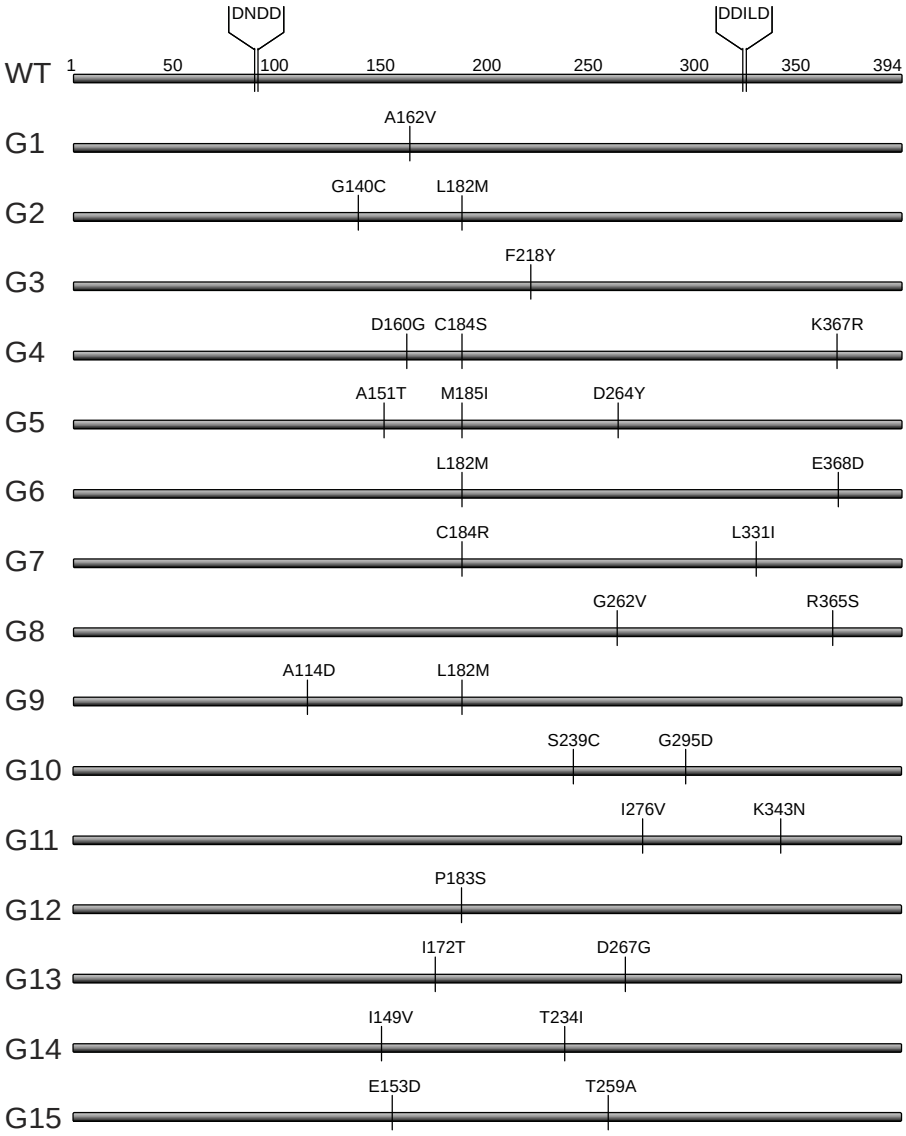
(**a**) The homology structural model for LPS second active site. Fifteen residues within 10 Å which were targeted for phylogenetic-based mutational analysis are shown. The substrate in the binding pocket is farnesyl hydroxyphosphonate. (**b**) Comparison of selected residues in *G. biloba* LPS (GbLPS) with those in *A. grandis* and *P. abies* abietadiene synthase (AbAS and PaAS), and *P. abies* isopimaradiene synthase (PaISO). Highlighted in green are residues in GbLPS that are different from all of the three paralogous enzymes. Highlighted in orange are residues in GbLPS found in any one of the paralogous enzymes.

Supplementary Figure 3. Amino acid alignment of *Sinapis alba* GGPPS (SaGGPPS) and *Taxus Canadensis* GGPPS (TcGGPPS).

```
SaGGPPS    (1) -----MASSVTPLGSWVLLHHHPSTILTQSRSRSPPS-------------
TcGGPPS    (1) MAYTAMAAGTQSLQLRTVASYQECNSMRSCFKLTPFKSFHGVNFNVPSLG

SaGGPPS   (33) ------LITLKPISLTPKRTVSSSS-----SSSLITKEDNNLKSSSSSFD
TcGGPPS   (51) AANCEIMGHLKLGSLPYKQCSVSSKSTKTMAQLVDLAETEKAEGKDIEFD

SaGGPPS   (72) FMSYIIRKADSVNKALDSAVPLREPLKIHEAMRYSLLAGGKRVRPVLCIA
TcGGPPS  (101) FNEYMKSKAVAVDAALDKAIPLEYPEKIHESMRYSLLAGGKRVRPALCIA

SaGGPPS  (122) ACELVGGEESLAMPARCAVEMIHTMSLIHDDLPCMDNDDLRRGKPTNHKV
TcGGPPS  (151) ACELVGGSQDLAMPTACAMEMIHTMSLIHDDLPCMDNDDFRRGKPTNHKV

SaGGPPS  (172) YGEDVAVLAGDALLSFAFEHLASATSSEVSPARVVRAVGELAKAIGTEGL
TcGGPPS  (201) FGEDTAVLAGDALLSFAFEHIAVATSKTVPSDRTLRVISELGKTIGSQGL

SaGGPPS  (222) VAGQVVDISSEGLDLNNVGLEHLKFIHLHKTAALLEASAVLGGIIGGGSD
TcGGPPS  (251) VGGQVVDITSEG--DANVDLKTLEWIHIHKTAVLLECSVVSGGILGGATE

SaGGPPS  (272) EEIERLRKFARCIGLLFQVVDDILDVTKSSQELGKTAGKDLIADKLTYPK
TcGGPPS  (299) DEIARIRRYARCVGLLFQVVDDILDVTKSSEELGKTAGKDLLTDKATYPK

SaGGPPS  (322) LMGLEKSREFAEKLNTEARDQLLGFDSDKVAPLLALANYIANRQN
TcGGPPS  (349) LMGLEKAKEFAAELATRAKEELSSFDQIKAAPLLGLADYIAFRQN
```

Highlighted in green are homologous residues.

Supplementary Figure 4. Location and identification of mutations carried by the isolated *ggpps* variants.



Supplementary Figure 5. Amino acid alignment of EAS(3) and LPS(4).

```
EAS    (1) -------------------------------------------------
LPS    (1) MAGVLFANLPCSLQLSPKVPFRQSTNILIPFHKRSSFGFNAQHCVRSHLR

EAS    (1) -------------------------------------------------
LPS   (51) LRWNCVGIHASAAETRPDQLPQEERFVSRLNADYHPAVWKDDFIDSLTSP

EAS    (1) -------------------------------------------------
LPS  (101) NSHATSKSSVDETINKRIQTLVKEIQCMFQSMGDGETNPSAYDTAWVARI

EAS    (1) -------------------------------------------------
LPS  (151) PSIDGSGAPQFPQTLQWILNNQLPDGSWGEECIFLAYDRVLNTLACLLTL

EAS    (1) -------------------------------------------------
LPS  (201) KIWNKGDIQVQKGVEFVRKHMEEMKDEADNHRPSGFEVVFPAMLDEAKSL

EAS    (1) -----MASAAVANYEEEIVRPVADFSPSLWGDQFLSFSIDNQVAEKYIYA
LPS  (251) GLDLPYHLPFISQIHQKRQKKLQKIPLNVLHNHQTALLYSLEGLQDVVDW

EAS   (46) QEIEALKEQTRSMLLATG----------RK--------------------
LPS  (301) QEITNLQSRDGSFLSSPASTACVFMHTQNKRCLHFLNFVLSKFGDYVPCH

EAS   (66) ----LADTLNLIDIIERLGISYHFEKEIDEILDQIYN---------QN--
LPS  (351) YPLDLFERLWAVDTVERLGIDRYFKKEIKESLDYVYRYWDAERGVGWARC

EAS  (101) SNCNDLCTSALQFRLLRQHGFNISPEIFSKFQDENG--K-FKESLASDVL
LPS  (401) NPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEKGDFFCFAGQTQIGVT

EAS  (148) GLLNLYEASHVRTHADDILEDALAFSTIHLESAAPH--------LKSPLR
LPS  (451) DNLNLYRCSQVCFPGEKIMEEAKTFTTNHLQNALAKNNAFDKWAVKKDLP

EAS  (190) EQVTHALEQCLHKGVPRVETRFFISSIYDKEQSK-----------NNVLL
LPS  (501) GEVEYAIKYPWHRSMPRLEARSYIEQFGSNDVWLGKTVYKMLYVSNEKYL

EAS  (229) RFAKLDFNLLQMLHKQELAQVSRWWKDLDFVTTLPYARDRVVECYFWALG
LPS  (551) ELAKLDFNMVQALHQKETQHIVSWWRESGFN-DLTFTRQRPVEMYFSVAV

EAS  (279) VYFEPQYSQARVMLVKTISMISIVDDTFDAYGTVKELEAYTDAIQRWDIN
LPS  (600) SMFEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDDLKLFSEAVRRWDIS

EAS  (329) EIDRLPD-YMKISYKAILDLYKDYEKELSSAGRSHIVCHAIERMKEVVRN
LPS  (650) VLDSVRDNQLKVCFLGLYNTVNGFGKDGLKEQGRDVLGYLRKVWEGLLAS

EAS  (378) YNVESTWFIEGYMPPVSEYLSNALATTTYYYLATTSYLG-MKSATEQDFE
LPS  (700) YTKEAEWSAAKYVPTFNEYVENAKVSIALATVVLNSIFFTGELLPDYILQ

EAS  (427) WLSKNPKILEASVIICRVIDDTATYEVEKSRGQIATGIECCMRDYG-IST
LPS  (750) QVDLRSKFLHLVSLTGRLINDTKTYQAERNRGELVSSVQCYMRENPECTE

EAS  (476) KEAMAKFQNMAETAWKDINEGLLRPTPVSTEFLTPILNLARIVEVTYIHN
LPS  (800) EEALSHVYGIIDNALKELNWELANPASNAPLCVRRLLFNTARVMQLFYMY

EAS  (526) LDGYTHPEKVLKPHIINLLVDSIKI
LPS  (850) RDGFGISDKEMKDHVSRTLFDPVA-
```

Methods

Cloning and pathway construction. The sequences of *ggpps (5)* and *lps(4)* were obtained from *Taxus canadensis* and *Gingko biloba*, respectively (Genbank accession codes: AF081514 and AF331704). We chose *T. canadensis* GGPPS as the upstream enzyme of LPS because this enzyme has high specificity toward FPP to synthesize GGPP, the diterpenoid substrate (5). Genes were custom-synthesized (DNA 2.0) to incorporate *E. coli* codon bias, remove restriction sites for cloning purposes, and establish a ~50% GC-content. Nucleotides corresponding to the 98 N-terminal amino acids of GGPPS (plastid transit peptide) were removed by designing custom oligonucleotides to generate mature proteins as previously described (5). In the case of LPS, truncation of 40 N-terminal amino acids was chosen because its incorporation into the levopimaradiene pathway gave rise to the most stable diterpenoid production in comparison to 60- and 80-amino acid truncations. In all cases, a start codon was introduced in the truncated gene fragments. For creating mutagenesis templates and sequencing purposes, *ggpps* and *lps* were individually cloned into pTrc99A (GE Healthcare) into the HindIII–EcoRI and EcoRI–SalI restriction sites, respectively.

The levopimaradiene pathways (wild type and mutants) were constructed by cloning PCR fragments of *ggpps* and *lps* into the HindIII–EcoRI and EcoRI–SalI sites of pTrcMod (6) to create pTrcGGPPS-LPS. To allow high throughput screening of GGPPS mutants, the biosynthetic gene cluster consisting of *crt*B and *crt*I derived from plasmid pAC-LYC (7) were cloned into the EcoRI – SalI

sites of pTrcMod to yield pTrcCRT. The mutant *ggpps* library was subsequently cloned into pTrcCRT in between the HindIII and EcoRI sites to create pTrcGGPPS*-CRT. In all cases, *E. coli* MG1655 Δ (*end*A, *rec*A) overexpressing the MEP pathway was used as the expression strain of the various pathways (wild-type and mutant levopimaradiene pathways, wild-type and mutant lycopene pathways).

Construction of the MEP pathway (*dxs-idi-ispDF* operon). The *dxs-idi-ispDF* operon was initially constructed by cloning each of the genes from the genome of *E. coli* K12 MG1655 into the pET21C+ plasmid with a T7 promoter (p20T7MEP). Using the primers dxsidiispDFNcoI(s) and dxsidiispDFKpnI(a) (SI Table 8), the *dxs-idi-ispDF* operon was sub-cloned into the pTrcMod plasmid after being digested with NcoI and KpnI restriction enzymes for the construction of p20TrcMEP. The p20TrcMEP plasmid was digested with MluI and PmeI and cloned into MluI and PmeI digested pACYC184-melA(P2A) plasmid to construct p10TrcMEP plasmid. For the construction of pACYC184-melA(P2A) plasmid, lacI-Ptrc-melA was amplified from pTrcmelAmut1(8) using the primers CS173blacIqEagI(s) and CS174brrnBBamHI(a). Then the PCR product and pACYC184 were digested with EagI and BamHI and ligated together. The p20TrcMEP plasmid was digested with BstZ17I and ScaI and cloned into PvuII digested pCL1920 plasmid to construct p5TrcMEP plasmid (9, 10).

To construct plasmids with the FRP-Km-FRP cassette used for amplifying the subsequent sequence for chromosomal localization of the pathway, the FRP-Km-FRP cassette was amplified from the pkD13 plasmid using the

KmFRPSacI(s) and KmFRPScaI(a) primers. This fragments was inserted into the p20TrcMEP plasmid was digested with SacI/ScaI (p20TrcMEPKm-FRP).

The MEP pathways constructed under the promoter Trc were localized to the arabinose operon region in the *E. coli* chromosome with the Kan marker. The PCR fragments were amplified from p20T7MEPKmFRP, using the primers IntTrc(s) and Int(a) and then electroporated into *E. coli* MG1655 Δ(*end*A, *rec*A) for chromosomal integration through the λ-Red recombination technique(11). The site specific localization was confirmed by PCR after successful gene integration.

Cultivation of engineered strains in 15 mL culture tube and 3-L bioreactor. Single transformants of pre-engineered *E. coli* strains expressing pTrcGGPPS-LPS or their mutant variants were cultivated for 18 h at 30°C in LB medium. For library characterization, these preinnocula were used to seed fresh 2-mL cultures in 15-mL vented tubes at a starting A_{600} of 0.1. The medium was composed of yeast extract, 5 g/L; Trypton, 10 g/L; glycerol, 15 g/L; NaCl, 10 g/L; HEPES, 100 mM; 3 ml 50% antifoam B. pH was adjusted to 7.6. Cultures were grown for 120 h at 22°C prior to diterpenoid analysis.

Scale-up experiments (1-L cultures) were done in 3-L bioreactors using the same culture medium except that glycerol was initially supplied at 10 g/L. During the fermentation, 3 g/L glycerol was added to the bioreactor when the concentration was observed to fall to 0.5 – 1 g/L. The aeration level was set to 0.5 vvm, and dissolved oxygen level was controlled at more than 20% during the course of fermentation by increasing agitation speed. All cultures were supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. To

minimize the loss of diterpenoids due to air-stripping, 20% dodecane was added into the culture.

GC-MS analysis of diterpenoids. For analysis of small-scale cultivations (libraries), 1 mL hexane was added into 1.5 mL culture aliquots and vortexed for 30 min. The mixture was centrifuged to separate the organic layer. For analysis of bioreactor cultivations, 1 μL of the dodecane layer was diluted to 200 μL with hexane. In both cases, 1 μL of hexane (containing the analytes) was analyzed by GC-MS (Varian Saturn 3800 GC attached to a Varian 2000 MS). The sample was injected into a HP5ms column 30m x 250 μM x 0.25 μM thickness (Agilent). Helium (ultra purity) at a flow rate 1.0 ml/min was used as a carrier gas. The oven temperature was first kept constant at 50° C for 1 min, and then increased to 220° C at the increment of 10° C/min, and finally held at this temperature for 10 min. The injector and transfer line temperatures were set at 200° C and 250°C, respectively. Because levopimaradiene, abietadiene, and sandaracopimaradiene are not commercially available, taxadiene, a diterpenoid possessing the same molecular mass as levopimaradiene, abietadiene, sandaracopimaradiene was used to construct a calibration curve for the peak areas obtained from the GC-MS.

Supplementary Table 8. Custom oligonucleotides used for MEP pathway construction

1	dxsNdeI(s)	CGGCATATGAGTTTTGATATTGCCAAATACCCG
2	dxsNheI(a)	CGGCTAGCTTATGCCAGCCAGGCCTTGATTTTG
3	idiNheI (s)	CGCGGCTAGCGAAGGAGATATACATATGCAAACGGAACACGTCATTTTATTG
4	idiEcoRI(a)	CGGAATTCGCTCACAACCCCGGCAAATGTCGG
5	ispDFEcoRI(s)	GCGAATTCGAAGGAGATATACATATGGCAACCACTCATTTGGATGTTTG
6	ispDFXhoI(a)	GCGCTCGAGTCATTTTGTTGCCTTAATGAGTAGCGCC
7	dxsidiispDFNcoI(s)	TAAACCATGGGTTTTGATATTGCCAAATACCCG
8	dxsidiispDFKpnI(a)	CGGGGTACCTCATTTTGTTGCCTTAATGAGTAGCGC
10	KmFRPScaI(a)	GACGAGTACTGAACGTCGGAATTGATCCGTCGAC
11	KmFRPSacI(s)	GACGGAGCTCGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTTGTGTAGGCTGGAGCTGCTTCG
12	IntTrc(s)	ATGACGATTTTTGATAATTATGAAGTGTGGTTTGTCATTGGCATCCGCTTACAGACAAGCTGTG
13	Int(a)	TTAGCGACGAAACCCGTAATACACTTCGTTCCAGCGCAGCCGACGTCGGAATTGATCCGTCGAC
14	CS173blacIqEagI(s)	GTTCGGCCG AGCTGTGACCGTCTCCGGGAGC
15	CS174brrnBBamHI(a)	GTTCACCGACAAACAACAGATAAAACGAAA

Supplementary Table 9. Custom oligonucleotides used for LPS mutagenesis

Mutation	5’-3’ Sequence
F-C618N	CGTACGCAAAAACCTCTAACCTGGCCGTAATCCTGG
R-C618N	CCAGGATTACGGCCAGGTTAGAGGTTTTTGCGTACG
F-L619F	GCAAAAACCTCTTGCTTCGCCGTAATCCTGGACGATC
R-L619F	GATCGTCCAGGATTACGGCGAAGCAAGAGGTTTTTGC
F-L696Q	GTAAAGTTTGGGAGGGCCAGCTGGCCTCCTATAC
R-L696Q	GTATAGGAGGCCAGCTGGCCCTCCCAAACTTTAC
F-K723S	GTATGTCGAGAACGCTAGTGTTAGCATCGCGCTGG
R-K723S	CCAGCGCGATGCTAACACTAGCGTTCTCGACATAC
F-A727S	CTAAAGTTAGCATCTCGCTGGCGACCGTTGTTCTG
R-A727S	CAGAACAACGGTCGCCAGCGAGATGCTAACTTTAG
F-A729G	CTAAAGTTAGCATCGCGCTGGGGACCGTTGTTCTG
R-A729G	CAGAACAACGGTCCCCAGCGCGATGCTAACTTTAG
F-V731L	CATCGCGCTGGCGACCCTTGTTCTGAACTC
R-V731L	GAGTTCAGAACAAGGGTCGCCAGCGCGATG
F-N769A	CCGGCCGTCTGATTGCCGACACCAAAACCTATCAG
R-N769A	CTGATAGGTTTTGGTGTCGGCAATCAGACGGCCGG
F-E777A	CCAAAACCTATCAGGCTGCACGTAACCGTGG
R-E777A	CCACGGTTACGTGCAGCCTGATAGGTTTTGG
F-N838E	CGTCGTCTGCTGTTCGAGACCGCGCGTGTAATGC
R-N838E	GCATTACACGCGCGGTCTCGAACAGCAGACGACG
F-G854T	GTACCGCGATGGCTTCACCATCAGCGATAAAGAAATG
R-G854T	CATTTCTTTATCGCTGATGGTGAAGCCATCGCGGTAC
F-I855L	CCGCGATGGCTTCGGCCTCAGCGATAAAG
R-I855L	CTTTATCGCTGAGGCCGAAGCCATCGCGG
F-M593A	GTCAGCGCCCGGTTGAAGCGTACTTTTCTGTTGCAG
R-M593A	CTGCAACAGAAAAGTACGCTTCAACCGGGCGCTGAC
F-M593C	GTCAGCGCCCGGTTGAATGTTACTTTTCTGTTGCAG
R-M593C	CTGCAACAGAAAAGTAACATTCAACCGGGCGCTGAC
F-M593D	GTCAGCGCCCGGTTGAAGACTACTTTTCTGTTGCAG
R-M593D	CTGCAACAGAAAAGTAGTCTTCAACCGGGCGCTGAC
F-M593E	GTCAGCGCCCGGTTGAAGAGTACTTTTCTGTTGCAG
R-M593E	CTGCAACAGAAAAGTACTCTTCAACCGGGCGCTGAC
F-M593F	GTCAGCGCCCGGTTGAATTTTACTTTTCTGTTGCAG
R-M593F	CTGCAACAGAAAAGTAAAATTCAACCGGGCGCTGAC
F-M593G	GTCAGCGCCCGGTTGAAGGGTACTTTTCTGTTGCAG
R-M593G	CTGCAACAGAAAAGTACCCTTCAACCGGGCGCTGAC
F-M593H	FGTCAGCGCCCGGTTGAACACTACTTTTCTGTTGCAG
R-M593H	CTGCAACAGAAAAGTAGTGTTCAACCGGGCGCTGAC
F-M593I	GTCAGCGCCCGGTTGAAATCTACTTTTCTGTTGCAG
R-M593I	CTGCAACAGAAAAGTAGATTTCAACCGGGCGCTGAC

F-M593K	GTCAGCGCCCGGTTGAAAAATACTTTTCTGTTGCAG
R-M593K	CTGCAACAGAAAAGTATTTTTCAACCGGGCGCTGAC
F-M593L	GTCAGCGCCCGGTTGAATTGTACTTTTCTGTTGCAG
R-M593L	CTGCAACAGAAAAGTACAATTCAACCGGGCGCTGAC
F-M593N	GTCAGCGCCCGGTTGAAAACTACTTTTCTGTTGCAG
R-M593N	CTGCAACAGAAAAGTAGTTTTCAACCGGGCGCTGAC
F-M593Q	GTCAGCGCCCGGTTGAACAGTACTTTTCTGTTGCAG
R-M593Q	CTGCAACAGAAAAGTACTGTTCAACCGGGCGCTGAC
F-M593P	GTCAGCGCCCGGTTGAACCGTACTTTTCTGTTGCAG
R-M593P	CTGCAACAGAAAAGTACGGTTCAACCGGGCGCTGAC
F-M593R	GTCAGCGCCCGGTTGAAAGGTACTTTTCTGTTGCAG
R-M593R	CTGCAACAGAAAAGTACCTTTCAACCGGGCGCTGAC
F-M593S	GTCAGCGCCCGGTTGAATCGTACTTTTCTGTTGCAG
R-M593S	CTGCAACAGAAAAGTACGATTCAACCGGGCGCTGAC
F-M593T	GTCAGCGCCCGGTTGAAACGTACTTTTCTGTTGCAG
R-M593T	CTGCAACAGAAAAGTACGTTTCAACCGGGCGCTGAC
F-M593V	GTCAGCGCCCGGTTGAAGTGTACTTTTCTGTTGCAG
R-M593V	CTGCAACAGAAAAGTACACTTCAACCGGGCGCTGAC
F-M593W	GTCAGCGCCCGGTTGAATGGTACTTTTCTGTTGCAG
R-M593W	CTGCAACAGAAAAGTACCATTCAACCGGGCGCTGAC
F-M593Y	GTCAGCGCCCGGTTGAATATTACTTTTCTGTTGCAG
R-M593Y	CTGCAACAGAAAAGTAATATTCAACCGGGCGCTGAC
F-A620C	CCTCTTGCCTGTGCGTAATCCTGGACG
R-A620C	CGTCCAGGATTACGCACAGGCAAGAGG
F-A620D	CCTCTTGCCTGGACGTAATCCTGGACG
R-A620D	CGTCCAGGATTACGTCCAGGCAAGAGG
F-A620E	CCTCTTGCCTGGAAGTAATCCTGGACG
R-A620E	CGTCCAGGATTACTTCCAGGCAAGAGG
F-A620F	CCTCTTGCCTGTTCGTAATCCTGGACG
R-A620F	CGTCCAGGATTACGAACAGGCAAGAGG
F-A620G	CCTCTTGCCTGGGCGTAATCCTGGACG
R-A620G	CGTCCAGGATTACGCCCAGGCAAGAGG
F-A620H	CCTCTTGCCTGCACGTAATCCTGGACG
R-A620H	CGTCCAGGATTACGTGCAGGCAAGAGG
F-A620I	CCTCTTGCCTGATCGTAATCCTGGACG
R-A620I	CGTCCAGGATTACGATCAGGCAAGAGG
F-A620K	CCTCTTGCCTGAAAGTAATCCTGGACG
R-A620K	CGTCCAGGATTACTTTCAGGCAAGAGG
F-A620L	CCTCTTGCCTGCTCGTAATCCTGGACG
R-A620L	CGTCCAGGATTACGAGCAGGCAAGAGG
F- A620M	CCTCTTGCCTGATGGTAATCCTGGACG
R-A620M	CGTCCAGGATTACCATCAGGCAAGAGG
F-A620N	CCTCTTGCCTGAACGTAATCCTGGACG
R-A620N	CGTCCAGGATTACGTTCAGGCAAGAGG
F-A620P	CCTCTTGCCTGCCCGTAATCCTGGACG
R-A620P	CGTCCAGGATTACGGGCAGGCAAGAGG

F-A620Q	CCTCTTGCCTGCAAGTAATCCTGGACG
R-A620Q	CGTCCAGGATTACTTGCAGGCAAGAGG
F-A620R	CCTCTTGCCTGCGCGTAATCCTGGACG
R-A620R	CGTCCAGGATTACGCGCAGGCAAGAGG
F-A620S	CCTCTTGCCTGTCCGTAATCCTGGACG
R-A620S	CGTCCAGGATTACGGACAGGCAAGAGG
F-A620T	CCTCTTGCCTGACCGTAATCCTGGACG
R-A620T	CGTCCAGGATTACGGTCAGGCAAGAGG
F-A620V	CCTCTTGCCTGGTCGTAATCCTGGACG
R-A620V	CGTCCAGGATTACGACCAGGCAAGAGG
F-A620W	CCTCTTGCCTGTGGGTAATCCTGGACG
R-A620W	CGTCCAGGATTACCCACAGGCAAGAGG
F-A620Y	CCTCTTGCCTGTACGTAATCCTGGACG
R-A620Y	CGTCCAGGATTACGTACAGGCAAGAGG
F-Y700A	GGCCTGCTGGCCTCCGCTACCAAGGAAGCG
R-Y700A	CGCTTCCTTGGTAGCGGAGGCCAGCAGGCC
F-Y700C	GGCCTGCTGGCCTCCTGTACCAAGGAAGCG
R-Y700C	CGCTTCCTTGGTACAGGAGGCCAGCAGGCC
F-Y700D	GGCCTGCTGGCCTCCGATACCAAGGAAGCG
R-Y700D	CGCTTCCTTGGTATCGGAGGCCAGCAGGCC
F-Y700E	GGCCTGCTGGCCTCCGAAACCAAGGAAGCG
R-Y700E	CGCTTCCTTGGTTTCGGAGGCCAGCAGGCC
F-Y700F	GGCCTGCTGGCCTCCTTTACCAAGGAAGCG
R-Y700F	CGCTTCCTTGGTAAAGGAGGCCAGCAGGCC
F-Y700G	GGCCTGCTGGCCTCCGGTACCAAGGAAGCG
R-Y700G	CGCTTCCTTGGTACCGGAGGCCAGCAGGCC
F-Y700H	GGCCTGCTGGCCTCCCATACCAAGGAAGCG
R-Y700H	CGCTTCCTTGGTATGGGAGGCCAGCAGGCC
F-Y700I	GGCCTGCTGGCCTCCATTACCAAGGAAGCG
R-Y700I	CGCTTCCTTGGTAATGGAGGCCAGCAGGCC
F-Y700K	GGCCTGCTGGCCTCCAAAACCAAGGAAGCG
R-Y700K	CGCTTCCTTGGTTTTGGAGGCCAGCAGGCC
F-Y700L	GGCCTGCTGGCCTCCTTAACCAAGGAAGCG
R-Y700L	CGCTTCCTTGGTTAAGGAGGCCAGCAGGCC
F-Y700M	GGCCTGCTGGCCTCCATGACCAAGGAAGCG
R-Y700M	CGCTTCCTTGGTCATGGAGGCCAGCAGGCC
F-Y700N	GGCCTGCTGGCCTCCAATACCAAGGAAGCG
R-Y700N	CGCTTCCTTGGTATTGGAGGCCAGCAGGCC
F-Y700P	GGCCTGCTGGCCTCCCCTACCAAGGAAGCG
R-Y700P	CGCTTCCTTGGTAGGGGAGGCCAGCAGGCC
F-Y700Q	GGCCTGCTGGCCTCCGAAACCAAGGAAGCG
R-Y700Q	CGCTTCCTTGGTTTCGGAGGCCAGCAGGCC
F-Y700R	GGCCTGCTGGCCTCCCGTACCAAGGAAGCG
RY700R	CGCTTCCTTGGTACGGGAGGCCAGCAGGCC
F-Y700S	GGCCTGCTGGCCTCCTCTACCAAGGAAGCG
R-Y700S	CGCTTCCTTGGTAGAGGAGGCCAGCAGGCC

F-Y700T	GGCCTGCTGGCCTCCACTACCAAGGAAGCG
R-Y700T	CGCTTCCTTGGTAGTGGAGGCCAGCAGGCC
F-Y700V	GGCCTGCTGGCCTCCGTTACCAAGGAAGCG
R-Y700V	CGCTTCCTTGGTAACGGAGGCCAGCAGGCC
F-Y700W	GGCCTGCTGGCCTCCTGGACCAAGGAAGCG
R-Y700W	CGCTTCCTTGGTCCAGGAGGCCAGCAGGCC

The letter F and R in the beginning of each mutagenic oligonucleotide indicates ‘forward’ and ‘reverse’ sequence, respectively.

REFERENCES

1. Cyr A, Wilderman PR, Determan M, & Peters RJ (2007) A modular approach for facile biosynthesis of labdane-related diterpenes. *J Am Chem Soc* 129(21):6684-6685.
2. Martin DM, Faldt J, & Bohlmann J (2004) Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiol* 135(4):1908-1927.
3. Facchini PJ & Chappell J (1992) Gene family for an elicitor-induced sesquiterpene cyclase in tobacco. *Proc Natl Acad Sci U S A* 89(22):11088-11092.
4. Schepmann HG, Pang J, & Matsuda SP (2001) Cloning and characterization of Ginkgo biloba levopimaradiene synthase which catalyzes the first committed step in ginkgolide biosynthesis. *Arch Biochem Biophys* 392(2):263-269.
5. Hefner J, Ketchum REB, & Croteau R (1998) Cloning and functional expression of a cDNA encoding geranylgeranyl diphosphate synthase from Taxus canadensis and assessment of the role of this prenyltransferase in cells induced for Taxol production. (Translated from English) *Archives of Biochemistry and Biophysics* 360(1):62-74 (in English).
6. Leonard E & Koffas MA (2007) Engineering of artificial plant cytochrome P450 enzymes for synthesis of isoflavones by Escherichia coli. *Appl Environ Microbiol* 73(22):7246-7251.
7. Cunningham FX, Jr., Sun Z, Chamovitz D, Hirschberg J, & Gantt E (1994) Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium Synechococcus sp strain PCC7942. *Plant Cell* 6(8):1107-1121.
8. Santos CNS & Stephanopoulos G (2008) Melanin-based high-throughput screen for L-tyrosine production in Escherichia coli. *Applied and environmental microbiology* 74(4):1190.
9. Lerner CG & Inouye M (1990) Low copy number plasmids for regulated low-level expression of cloned genes in Escherichia coli with blue/white insert screening capability. *Nucleic Acids Res* 18(15):4631.
10. Tyo KE, Ajikumar PK, & Stephanopoulos G (2009) Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat Biotechnol* 27(8):760-765.
11. Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* 97(12):6640-6645.