Supporting Information

Santos et al. 10.1073/pnas.1206346109

SI Text

Methods. ***Cloning of pZE-kan^{FRT}-tyrA^{fbr}aroG^{fbr} plasmids.*** Plasmid pCL1920 :: *tyrA*^{fbr}*aroG*^{fbr}(1) was used as a template for the amplification of *tyrA*^{fbr}-*aroG*^{fbr} with *Pfu* Turbo DNA polymerase (Stratagene) and the following primers: CS114 tyrAfbr sense KpnI and CS278 aroGfbr anti MluI. This PCR product was then digested with KpnI and MluI and ligated into five KpnI/MluI digested pZE-*gfp* plasmids taken from a synthetic constitutive promoter library (2). The resulting plasmid pZE-*tyrA*^{fbr}*aroG*^{fbr} consists of five different versions, each corresponding to different promoter strengths (R, Y, W, B, and P_L) (Table S1). To construct pZE-*kan*^{FRT}-*tyrA*^{fbr}*aroG*^{fbr}, primers CS279 pKD13 sense SacI and CS280 pKD13 anti AatII were used to amplify an FRT-flanked kanamycin resistance gene (*kan*) on the plasmid pKD13 (3). After digestion with SacI and AatII, this product was ligated to SacI/AatII-digested pZE-*tyrA*^{fbr}*aroG*^{fbr} and transformed into chemically competent *E. coli* DH5α cells as described in the protocol. Following transformation, all plasmid constructs were isolated, verified by sequencing, and transformed into *E. coli* K-12 Δ*pheA* Δ*tyrR*. All enzymes used in the cloning procedure were purchased from New England Biolabs. The names and sequences of all primers used for gene amplification are listed in Table S3.

Chromosomal integration of tyrA^{fbr}-aroG^{fbr} cassette. The *kan*^{FRT}-*tyrA*^{fbr}-*aroG*^{fbr} cassette was integrated into the *lacZ* locus of *E. coli* K-12 Δ*pheA* Δ*tyrR* using a lambda-red recombination-based method (3). Briefly, *kan*^{FRT}-*tyrA*^{fbr}-*aroG*^{fbr} was amplified from pZE-*kan*^{FRT}-*tyrA*^{fbr}*aroG*^{fbr} with primers CS173 aroG-lacZ anti and CS281 pKD13-lacZ sense and Platinum *Pfx* DNA polymerase (Invitrogen). Both primers incorporated 75–77 bp of homology with the ends of the *lacZ* gene to facilitate integration into the proper locus. Following transformation of the cassette into *E. coli* K-12 Δ*pheA* Δ*tyrR* pJM12 (a pKD46 derivative), colonies were verified by colony PCR and sequencing. Excision of FRT-flanked *kan* was mediated by transformation with FLP recombinase-expressing pCP20 as described in the literature (3). Five versions of *E. coli* K-12 Δ*pheA* Δ*tyrR lacZ* :: *tyrA*^{fbr}*aroG*^{fbr} were constructed by this process with each representing different promoter strengths (R, Y, W, B, P_L).

To integrate a second *tyrA*^{fbr}-*aroG*^{fbr} cassette within the *tyrR* locus, *kan*^{FRT}-*tyrA*^{fbr}-*aroG*^{fbr} was amplified from pZE-*kan*^{FRT}-*tyrA*^{fbr}*aroG*^{fbr} (using promoters W, B, and P_L only) with primers CS286 pKD13-tyrR sense and CS287 aroG-tyrR anti. These primers incorporated 77–78 bp of homology with the ends of the *tyrR* gene to facilitate integration into the proper locus. Following transformation into *E. coli* K-12 Δ*pheA* Δ*tyrR* pJM12, verification and excision of *kan* were performed as described earlier. Integration of a second *tyrA*^{fbr}-*aroG*^{fbr} cassette into these strains was mediated by P1 transduction (4) of *lacZ* :: *kan*^{FRT}-$P_{LtetO-1}$-*tyrA*^{fbr}*aroG*^{fbr} from the previously constructed strain *E. coli* K-12 Δ*pheA* Δ*tyrR lacZ* :: *kan*^{FRT}-$P_{LtetO-1}$-*tyrA*^{fbr}*aroG*^{fbr}. Selection on kanamycin, verification, and subsequent *kan* excision resulted in three separate versions of *E. coli* K-12 Δ*pheA* Δ*tyrR lacZ* :: *tyrA*^{fbr}*aroG*^{fbr}*tyrR* :: *tyrA*^{fbr}*aroG*^{fbr}. All three strains contain a P_L promoter at the *lacZ* locus and the promoters R, B, or P_L at the *tyrR* site. The strain *E. coli* K-12 Δ*pheA* Δ*tyrR lacZ* :: $P_{LtetO-1}$-*tyrA*^{fbr}*aroG*^{fbr}*tyrR* :: $P_{LtetO-1}$-*tyrA*^{fbr}*aroG*^{fbr} (which makes use of the highest strength promoter, P_L, for both cassettes) will henceforth be referred to as parental strain P2.

Creation of pZE overexpression plasmids. *E. coli* K-12 genomic DNA was prepared using the Wizard Genomic DNA Purification Kit (Promega) and used as a template for the synthesis of genes *gadE*, *ydeO*, *evgA*, and a truncated *relA*(5)with Phusion DNA polymerase (New England Biolabs). All PCR fragments were digested with KpnI and BsaI or MluI and subsequently ligated to similarly digested pZE-*gfp* plasmids. To enable testing of both low- and high-level expression, two promoters from a synthetic library were used in the creation of each plasmid construct (II, AA, or Y for low expression; P_L for high expression) (Table S1) (2). Following transformation into chemically competent *E. coli* DH5α (Invitrogen), plasmids were isolated and verified by both PCR and sequencing. All enzymes used in the cloning procedure were purchased from New England Biolabs. The names and sequences of all primers used for gene amplification are listed in Table S3.

Reengineering of SNPs into the bacterial chromosome. We utilized a two-step lambda-red recombination based method (3) to delete the relevant locus in a P2 background and subsequently replace it with a SNP-substituted variant. Two knockout cassettes for both the *purF* and *hisH* genes were generated by amplification of pKD13's *kan*^{FRT} region using Taq DNA polymerase (New England Biolabs). Primers CS707 pKD13 kan-purF sense and CS708 pKD13 kan-purF anti were used for the generation of *purF*::*kan* integration cassettes, and CS709 pKD13 kan-hisH sense and CS710 pKD13 kan-hisH anti were used to create the *hisH*::*kan* fragment (Table S3). Cassettes were transformed into P2 pJM12 (a pKD46 variant utilizing a trc promoter), and colonies were verified by colony PCR and sequencing. Cassettes for the second round of integration were generated through the amplification of SNP-containing *purF* and *hisH* using primers CS699 purF sense, CS700 purF anti, CS701 hisH sense, and CS702 hisH anti (Table S3) and genomic DNA preparations from *rpoA14*, *rpoA27*, and *rpoD3* as templates. SNP variants were then transformed into either P2 *purF* :: *kan*^{FRT} pJM12 or P2 *hisH* :: *kan*^{FRT} pJM12 and grown overnight in 5 ml M9 minimal medium with 5 g/l glucose. Correct transformants were selected by this alternate method due to the absence of an antibiotic marker for the second integration event. However, because *purF* and *hisH* deletion strains exhibit much slower growth rates in minimal medium than their SNP-containing counterparts, adequate selection was achieved after just a single round of overnight growth in minimal media. Individual colonies were tested for the loss of kanamycin resistance (indicating replacement of *kan*^{FRT} with a mutated *purF* or *hisH*) and validated by colony PCR and sequencing.

Mutational analysis of strains. Mutation frequency analysis protocols were adapted from a previously published report on the use of *rpoB* to analyze the specificity of base substitutions in *E. coli* 6. Briefly, cells were inoculated at a starting OD_{600} of 0.01 and grown for 48 h in 5 ml LB cultures. After appropriate dilution, cultures were plated on LB-agar with and without 100 μg/ml rifampicin (Sigma). (Dilution is necessary to recover distinct colonies as opposed to lawn growth, with an approximate target of less than 200 colonies per plate). The number of colony forming units (cfus) on each plate was quantified after 16–20 h, and average mutation frequencies were determined by dividing the average number of rifampicin-resistant colonies by the average number of cfus growing on unselective media (LB). At least ten culture and plate replicates were used in this calculation for each strain tested.

1. Santos CNS, Stephanopoulos G (2008) Melanin-based high-throughput screen forL-tyrosine production in *Escherichia coli*. *Appl Environ Microbiol* 74:1190–1197.
2. Alper H, Fischer C, Nevoigt E, Stephanopoulos G (2005) Tuning genetic control through promoter engineering. *Proc Natl Acad Sci USA* 102:12678–12683.
3. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* 97:6640–6645.
4. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed.
5. Schreiber G, et al. (1991) Overexpression of the *relA* gene in *Escherichia coli*. *J Biol Chem* 266:3760–3767.
6. Garibyan L, et al. (2003) Use of the *rpoB* gene to determine the specificity of base substitution mutations on the *Escherichia coli* chromosome. *DNA Repair (Amst)* 2:593–608.

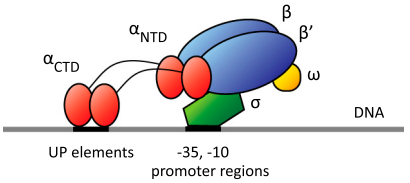

Fig. S1. Structure and DNA interactions of RNA polymerase. The RNA polymerase holoenzyme is composed of the following subunits: $\alpha_2\beta\beta'\omega$. The C-terminal domain of the α subunit (αCTD) is capable of influencing promoter specificity through its interactions with upstream promoter (UP) elements and other activator/repressor proteins. *E. coli*'s seven sigma factors, which include the principal sigma factor σ^{70}, also control transcription through their interactions with the –35 and –10 regions of promoters.

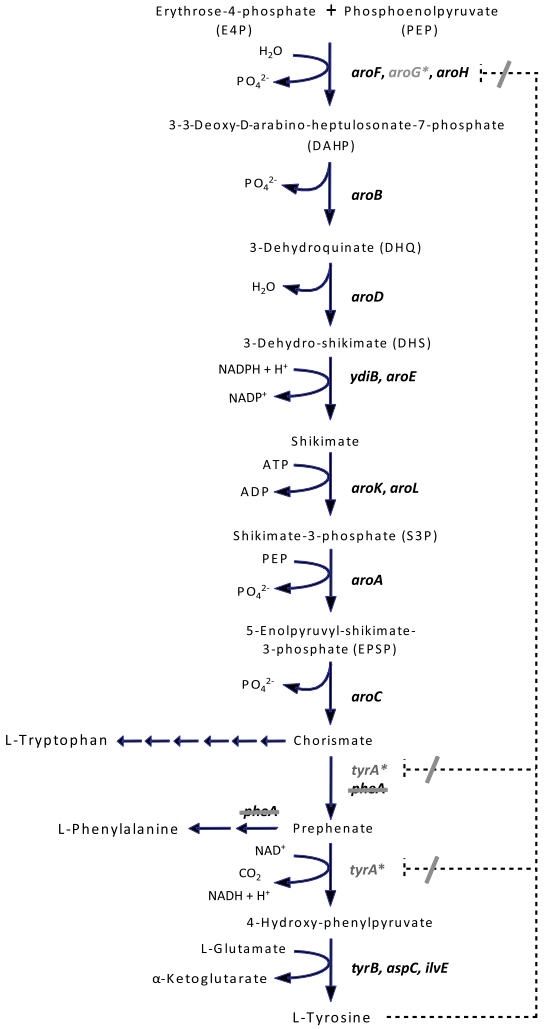

Fig. S2. Aromatic amino acid biosynthetic pathway. Dashed line indicates allosteric regulation by L-tyrosine. Asterisks indicate overexpression of a feedback-resistant (fbr) form of the enzyme. Gray strikeouts or lines indicate deletion or inactivation of the gene or feedback loop.

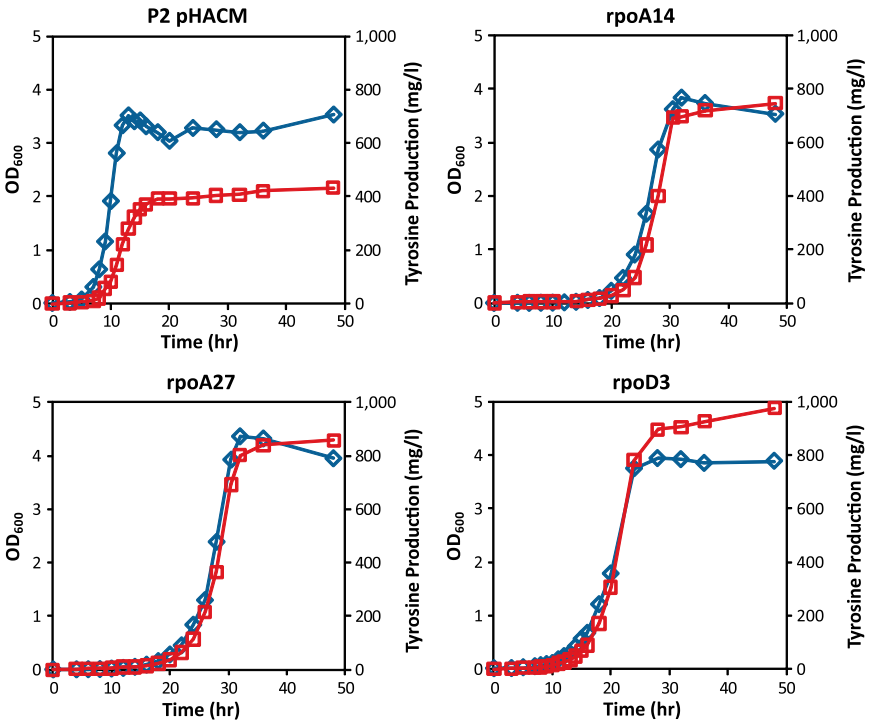

Fig. S3. Representative growth (blue diamonds) and L-tyrosine production (red squares) profiles for parental P2 with pHACM (empty plasmid) and three global transcription machinery engineering (gTME) mutants. Specific growth rates for these strains are as follows: 0.617 h^{-1} (P2 pHACM), 0.296 h^{-1} (*rpoA14*), 0.249 h^{-1} (*rpoA27*), and 0.290 h^{-1} (*rpoD3*).

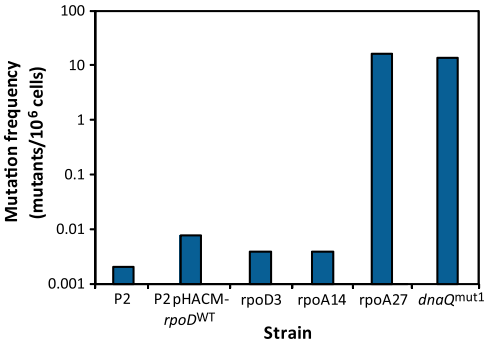

Fig. S4. Mutation frequencies of select strains. Mutation frequencies were measured after 48-h growth in 5 ml LB. Mutation data for *dnaQ*^{mut1} (K-12 Δ*pheA* Δ*tyrR* with a truncated *dnaQ* gene) was included as a positive control (1).

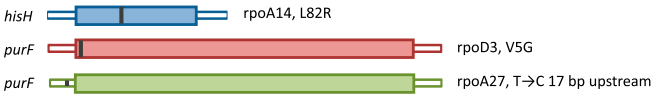

Fig. S5. Locations of validated SNPs within the *hisH* and *purF* loci.

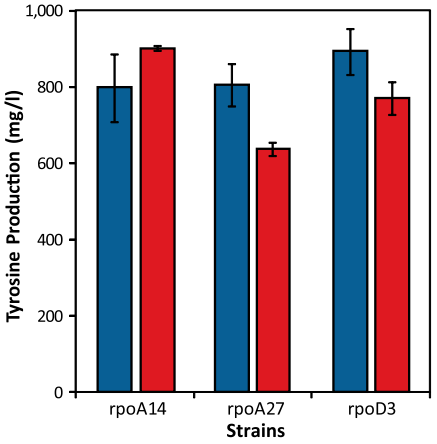

Fig. S6. Comparison of original and reconstructed gTME strains. L-tyrosine titers for the original isolates are shown in blue. The performance of reconstructed strains containing an introduced SNP and their corresponding *rpoA* or *rpoD* plasmids is shown in red. All L-tyrosine concentrations were measured after 48 h.

Table S1. Relative strengths of synthetic constitutive promoters

Promoter	Average promoter strength metric (0–1)*
II	0.04
R	0.14
AA	0.22
Y	0.31
W	0.54
B	0.82
P_L	0.87

*Data provided by C. Fischer.

Table S2. Validated SNPs in gTME-derived mutants

	rpoD3		*rpoA14*		*rpoA27*		
Position	R	S	R	S	R	S	Annotation
2092803			T	G			*hisH*; imidazole glycerol phosphate synthase subunit, glutamine amidotransferase (histidine biosynthesis)
2428247	A	C					*purF*; amidophosphoribosyl transferase (*de novo* purine biosynthesis)
2428277					T	C	

R = reference (P2) sequence; S = substituted base pair

Table S3. Primers used in this study

Primer name	Primer sequence (5′ → 3′)
CS114 tyrAfbr sense KpnI	GCT CGG TAC CAT GGT TGC TGA ATT GAC CGC ATT ACG
CS173 aroG-lacZ anti	TTC CGG CAC CAG AAG CGG TGC CGG AAA GCT GGC TGG AGT GCG ATC TTC CTG AGG CCG ATA CTG TCG TCG TCC CCT TTA CCC GCG ACG C
CS278 aroGfbr anti MluI	CGA CGC GTT TAC CCG CGA CGC GCT TTT ACT G
CS279 pKD13 sense SacI	TCC GAG CTC TTG TGT AGG CTG GAG CTG CTT CGA
CS280 pKD13 anti AatII	TCT TAG ACG TCG GAA TTG ATC CGT CGA CCT GCA GTT CGA A
CS281 pKD13-lacZ sense	CGC GTG CAG CAG ATG GCG ATG GCT GGT TTC CAT CAG TTG CTG TTG ACT GTA GCG GCT GAT GTT GAA CTG GAA GTC GTG TAG GCT GGA GCT GCT TCG A
CS286 pKD13-tyrR sense	TGC AAT ATC GGG TGC TGA CCG GAT ATC TTT ACG CCG AAG TGC CCG TTT TTC CGT CTT TGT GTC AAT GAT TGT TGA CAG GTG TAG GCT GGA GCT GCT TCG A
CS287 aroG-tyrR anti	TAA TTT AAT ATG CCT GAT GGT GTT GCA CCA TCA GGC ATA TTC GCG CTT ACT CTT CGT TCT TCT TCT GAC TCA GAC CAT TAC CCG CGA CGC GCT TTT ACT G
CS555 evgA sense KpnI	CTC GGT ACC ATG AAC GCA ATA ATT ATT GAT GAC CAT CC
CS556 evgA anti MluI	CGA CGC GT T TAG CCG ATT TTG TTA CGT TGT GCG
CS557 ydeO sense KpnI	CTC GGT ACC ATG TCG CTC GTT TGT TCT GTT ATA TTT ATT C
CS558 ydeO anti BsaI	GGT CTC TCT TTT CAA ATA GCT AAA GCA TTC ATC GTG TTG C

Primer name	Primer sequence (5′ → 3′)
CS559 gadE sense KpnI	CTC GGT ACC ATG ATT TTT CTC ATG ACG AAA GAT TCT TTT C
CS560 gadE anti MluI	CGA CGC GTC TAA AAA TAA GAT GTG ATA CCC AGG GTG ACG
CS582 relA sense KpnI	CTC GGT ACC ATG GTT GCG GTA AGA AGT GCA CAT ATC A
CS584 relA trunc anti MluI	CGA CGC GTT TAC AGC TGG TAG GTG AAC GGC ACA AT
CS699 purF sense	GCA GCA ATG GCA GCG AAA ATA TTG
CS700 purF anti	CAG TCT GGT TTA CGG GCT TTG AAG AC
CS701 hisH sense	TCT CAG CAC CGA AAT GAT CGA GCA
CS702 hisH anti	CCG GAA TAA TCA TCA CAT CTC CAG GA
CS707 pKD13 kan-purF sense	TAA CGC ACA TGA CCA ATG CCC ATA TTG CCC TGC AAA CGC TGC ATA TGG CGA GCG TGT AGG CTG GAG CTG CTT C
CS708 pKD13 kan-purF anti	CGG TAC TGT TTA TCG CTA CCC TGA TCG TTG GTG CTA TCG TGA ACT TCG TGA TCC GTC GAC CTG CAG TTC GA
CS709 pKD13 kan-hisH sense	CGT GAC CCG GAC GTC GTG TTG CTG GCC GAT AAA CTG TTT TTA CCC GGC GTT GGC ACT GAT CCG TCG ACC TGC AGT TCG A
CS710 pKD13 kan-hisH anti	CGG CAT TGC GTA GCT GTG AAC AAA GTA AAA GTA CGC GCC GTC TTC AAT CCC CTG TGT AGG CTG GAG CTG CTT C