Genetic-metabolic coupling for targeted metabolic engineering

Stefano Cardinale [1, 2 *], Felipe Gonzalo Tueros [1] and Morten Otto Alexander Sommer [1 *]

[1] NNF-CFB, Technical University of Denmark, 2800 Kongens Lyngby, Denmark
[2]Lead Contact
[*]Corresponding authors: stefca@biosustain.dtu.dk and M.O.A.S: msom@bio.dtu.dk

SUMMARY

To produce chemicals, microbes typically employ potent biosynthetic enzymes that interact with native metabolism to affect cell fitness as well as chemical production. However, production optimization largely relies on data collected from wild type strains in the absence of metabolic perturbations, thus limiting their relevance to specific process scenarios. Here, we address this issue by coupling cell fitness to the production of thiamine diphosphate in *Escherichia coli* using a synthetic RNA biosensor.

We apply this system to interrogate a library of transposon mutants to elucidate the native gene network influencing both cell fitness and thiamine production. Specifically, we identify uncharacterized effectors of the OxyR-SoxR stress response that limit thiamine biosynthesis via alternative regulation of iron storage and Fe-S-cluster inclusion in enzymes.

Our study represents a new generalizable approach for the reliable high-throughput identification of genetic targets of both known and unknown function that are directly relevant to a specific biosynthetic process.

Keywords

RNA biosensors; thiamine biosynthesis; transposon mutagenesis; stress response; cofactors.

INTRODUCTION

A common problem in the field of biotechnology is the unpredictability inherent in engineering complex biological systems. The use of a whole-system approach can prevent some of the obstacles encountered in bioengineering, including toxicity due to metabolite over-production (Fletcher et al., 2016), metabolic bottlenecks (Lechner et al., 2016), and low product titers (Otero and Nielsen, 2013). Systems metabolic engineering relies on the generation and analysis of large datasets to identify key genetic components that contribute to high-yield production phenotypes. However, this methodology often assumes an unperturbed system, and our capacity to elucidate all metabolic and regulatory perturbations affecting the production of a given metabolite is limited.

The impact of engineering a cellular system to produce a primary metabolite can be substantial. The biosynthetic pathways of primary metabolites often utilize specialized enzymes of high catalytic rate and substrate affinity; these attributes support high flux through the pathway, even at low enzyme concentrations (Nam et al., 2012). When these enzymes are over-expressed, the supply of energy, cofactors, and carbon building blocks quickly become limiting and can lead to a strong metabolic burden on the cell. Indeed, metabolic burden can lead to low cell density and reduced titers during the production of recombinant vitamins, such as cobalamin (vitamin B_{12}) (Biedendieck et al., 2010) or riboflavin (vitamin B_2) (Lin et al., 2014). Furthermore, central cell metabolism, which provides building blocks for vitamin biosynthesis, is subject to complex regulation that can affect pathway optimization (Lin et al., 2014).

The effect of producing a metabolite is strongly dependent on the individual metabolite and how it affects cell metabolism and its regulation. Understanding the relevance of this context dependency at the genome level remains a challenge in the field of biotechnology. New approaches enabling high-throughput characterization of fitness and pathway yields are key to achieving this goal. Tagged transposon mutagenesis (Oh et al., 2010) and synthetic RNA-based biosensors (Townshend et al., 2015) are genetic tools that have been applied to study the condition-dependent contributions of genes to cell growth and synthetic pathway output.

Here, we present a methodology for characterizing the metabolic perturbations that are a direct result of the over-production of a primary metabolite in a bacterial cell. Genetic-metabolic coupling is based on the concurrent expression of a synthetic metabolic pathway and an end-product biosensor in a population of bar coded single-gene transposon insertion mutants. Mutant fitness data quantified by deep sequencing are used to identify which genes provide growth advantage exclusively upon selection for the end product. We use genetic-metabolic coupling to assess the influence of *Escherichia coli* cell functions during biosynthesis of a derivative of vitamin B_1 – thiamine diphosphate (TPP).

We found that the OxyR-Fur-mediated regulation of Fe-S cluster formation strongly affects TPP output. Ultimately, genetic-metabolic coupling allowed the identification of a small set of genes, several of which of uncharacterized or predicted function, with strong population-wide fitness phenotypes from over 2000 tested genes. When introducing specific knockouts in a clean genetic background, the identified genes resulted in the predicted changes to the TPP titers with different production plasmids and in different host strains. We finally demonstrate reverse tuning of several identified genetic targets with multi-copy plasmids further supporting the broad application of genetic-metabolic coupling for metabolic engineering.

RESULTS

Coupling population genetics to thiamine metabolic engineering

To create a TPP-overproducing strain for our analysis, we assembled the TPP biosynthetic genes on a plasmid. The genes were arranged in two operons that encoded the two different cellular branches required for the biosynthetic pathway. Four genes are involved in the synthesis of thiazole monophosphate (THZ-P) from L-cysteine, L-tyrosine and deoxy-D-xylulose-5-phosphate (DXP): *thi*FSGH, *thiC* for the synthesis of 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP-P) from 5-aminoimidazole riboside (AIR), *thiE* for coupling THZ-P to HMP-PP, and either *thiD*, the kinase that phosphorylates HMP-P to HMP-PP (plasmid pTHId), or *thiM* (plasmid pTHIm), a salvage enzyme that phosphorylates THZ. In the latter case, the strain relies on the native chromosomal *thiD* gene for HMP-P phosphorylation (details are provided in Sup. Fig. 1A-B). The synthesis of TPP relies on several cellular biosynthetic pathways for substrates and in particular purine nucleotides biosynthesis for AIR, cysteine biosynthesis, and S-adenosylmethionine (SAM) and NADPH as cofactors.

To quantitatively relay the intracellular TPP concentration at the single-cell level, we used a TPP biosensor (plasmid pTPP_Bios) based on an engineered TPP riboswitch (Genee et al., 2016). This riboswitch is used to activate the expression of two genes conferring resistance to the antibiotics chloramphenicol (CAM) and spectinomycin (SpeR). This dual selection design allows for the reliable selection of high TPP producers using both the pTHIm and pTHId plasmids with very low false-positive rate (Genee et al., 2016). The extracellular titers of *de novo*-produced TPP in *E. coli* MG1655::Δ*tbpA*Δ*thiM* (BS134), which carries deletions of the thiamine salvage and import pathways, were ~200 and ~800 μg/l for pTHIm and pTHId, respectively (Fig. 1A). The overexpression of thiamine biosynthetic genes reduced the growth rates by 10% and 35% for pTHIm and pTHId, respectively.

To understand the cellular components impacting the *E. coli* TPP biosynthetic capacity, we used whole-genome transposon (Tn) mutagenesis to construct libraries of *E. coli* MG1655 Tn5 mutant strains harboring the thiamine selection system (pTPP_Bios) and the plasmid carrying both *de novo* and salvage biosynthetic genes with a lower cell burden (pTHIm), or the plasmid without biosynthetic pathway, constructing the respective libraries lTpp and lMck (Figure 1B). Insertions were evenly distributed across the genome and through the lengths of open reading frames (Sup. Fig. 1C-D) (RefSeq. NC_000913.3). To characterize the impact of transposon insertions within various *E. coli* genes on the intracellular excess of TPP, we subjected the lMck and lTpp mutant libraries to a competitive fitness assay for 16 hours in shake flasks, obtaining robust fitness estimates for 1857 genes in the lMck library and 2054 genes in the lTpp library with high reproducibility across biological replicates (Sup. Fig. 1E). The relative gene fitness in each condition was calculated from strain bar code

abundance (log_2 fitness) using the relative abundance in the Tn libraries at time zero (Wetmore et al., 2015).

Iron and sulfur significantly impact TPP production

We applied a two-step methodology for identifying genes and biological functions with significant condition-specific fitness phenotype (top/bottom 5%, see also Sup. Methods), or a significantly reduced/increased relative abundance of TPP-producing Tn mutants compared to time zero or mutants carrying the backbone-only plasmid (lMck). First, extreme fitness values were identified by setting top and bottom thresholds for each condition. These values were used to build a factorized matrix with +1 or -1 values corresponding to significantly increased or decreased fitness, respectively, which was later used to establish gene-to-condition associations for functional enrichment (DAVID) (Huang et al., 2008) (Fig. 2A).

We focused our analysis on two effects: the effect of having biosynthetic genes for TPP compared to control backbone-only plasmid and the effect of selecting for intracellular TPP compared to no selection. In both cases we expected that the deletion of a gene encoding factors important for the biosynthesis of TPP, which was not provided in the cultivation media, or for replenishing cellular resources would lead to relatively reduced fitness compared to the full population.

In the presence of the TPP heterologous pathway but without selective pressure (fTpp), we identified 138 genes with significant fitness changes, of which 102 exhibited reduced and 36 exhibited increased fitness (Sup. Fig. 2A and Sup. Table 1). Knockouts of genes affecting cofactor or substrate supply to the TPP pathway showed fitness defects: *rtn* (ΔF: -1.0, ΔF: fitness difference to control) and *fucA* (ΔF: -1.3), implicated in ribose metabolism, *bfr* (ΔF: -1.1) and *yaaA* (ΔF: -0.8) in iron homeostasis, and *dxs*, involved in DXP biosynthesis. We also found enrichment for knockouts of genes involved in protein folding (Sup. Table 1, $P < 0.05$, with Bonferroni correction for multi-hypotheses testing). During selection for TPP without the TPP heterologous pathway ($fMck^{CS}$), several additional genes involved in cellular iron homeostasis (*fes*, *ftnA* and *cueO*) and the Gene Ontology (GO) class "water-soluble vitamin biosynthesis functions" (GO:0006767, $P < 0.05$, Sup. Fig. 2B and Sup. Table 1) showed reduced fitness. Overall, this result indicates that TPP biosynthesis relies strongly on the cell protein and vitamin synthesis apparatus, especially cellular iron availability. This result also highlights that our synthetic TPP biosensor is sensitive to the concentration of endogenous TPP.

We next studied the effect of selecting for TPP in the presence of the heterologous TPP pathway ($fTpp^{CS}$). We compiled a list of 45 genes with significant fitness changes, of which 27 were specific to this condition (e.g., with a neutral phenotype in the absence of selection or biosynthetic genes; Fig. 2B and Sup. Fig. 2C-D). This set of genes was functionally enriched for iron and glutathione ABC

transporters (Kyoto Encyclopedia of Genes and Genomes [KEGG] pathway, DAVID Bonferroni-corrected $P < 0.01$), involvement in the sulfur-compound biosynthetic process (GO:0044272 $P \leq 0.05$), and DNA restriction and modification (GO:0009307, $P < 0.01$). In addition, similar phenotypes were shared by functionally related genes: reduced Fe-S cluster availability (*csdE, cysE, ccmC* mutations) presented reduced fitness (Fig. 2B –bottom), while gene knockouts that were involved in sulfur assimilation, major redox regulators OxyR and SoxR, and HSD DNA restriction enzymes resulted in a positive fitness shift (Fig. 2B). Indeed, the pTHId pathway was more stable in a Δ*hsd*MR knockout compared to wild type (Sup. Fig. 3), possibly suggesting direct cleavage by this endonuclease.

SAM turnover during TPP biosynthesis requires extensive NADPH reductive power (Chen et al., 2015). $NADP^+$ can increase cellular oxidation, impair Fe-S cluster formation, and cause accumulation of hydrogen peroxide and reactive oxygen species (ROS) all tightly controlled by OxyR-SoxR regulons (Imlay, 2015). We tested the effect on TPP titer from limiting Fe-S clusters and the activation of OxyR-SoxR regulons. Sequestration of intracellular unbound iron by iron chelators o-phenanthroline and 2,2'-dipyridyl (Chupka et al., 1988) led to a small but significant reduction in TPP titer already at low concentrations (~25 μM) whereas thiourea showed near-normal TPP levels (Fig. 2C, $P < 0.02$, red bars). These effects were more significant at 5-fold higher compounds concentration, (Fig. 2C, cyan bars), indicating that free iron and Fe-S clusters are limiting during TPP production and the removal of ROS has an effect similar to the OxyR knockouts and regulon down-regulation.

Elucidation of the Fe-S cluster biogenesis network

We used a machine-learning algorithm to mine RefSeq gene descriptions (Tatusova et al., 2014) and quantified for each condition the effects of knocking out genes with biological functions related to iron, sulfur, cysteine, tyrosine and ribose, which are cofactors or substrates involved in TPP biosynthesis (Chatterjee et al., 2008; Kriek et al., 2007). We also looked at cellular redox regulation, which is believed to influence cellular vitamin metabolism (Dougherty and Downs, 2006). We identified 108 genes in the *E. coli* genome related to these functions, among which 57 genes (~53%) showed significant fitness shifts during selection for TPP production: 35 were associated with cofactor supply, and 25 were associated with cellular redox regulation ($fTpp^{CS}$ and fTPP, $P < 0.1$ compared to a random set by random sampling with replacement or bootstrapping, Fig. 3A). The presence of a significant number of genes involved in redox regulation suggests that oxidative stress can be relevant during TPP production.

To extend our analysis beyond a specific enzyme or transporter, we quantified the roles of the production phenotypes of parent GO classes (Keseler et al., 2009) associated with genes with extreme

fitness phenotypes. We found that mutants of 'Cellular Iron Ion Homeostasis' (GO: 006879) genes had a significant low-fitness phenotype during selection for TPP ($P < 0.05$, Sup. Table 2); however, mutations in genes involved in iron, siderophores and ferric-enterobactin transport (GO: 00015682, 0015684, 0015685, 015891) were associated with significantly higher fitness ($P < 0.1$, Fig. 3Bi and Sup. Fig. 4A-B).

In bacteria, sulfur is mainly assimilated in the cell in a reduced form, as hydrogen sulfide from sulfate, and is incorporated into the amino acids cysteine and methionine. We found that the average fitness of mutants in genes involved in 'Sulfate Assimilation' (GO: 0000103) had significantly increased ($P < 0.01$) during selection for TPP (Fig. 3Bii, Sup. Fig. 4C). We also found that impairment of genes related to 'Methionine Biosynthetic Process' (GO:0009086) but not cysteine biosynthesis (Δ*cysE*: -1.5 ΔF) led to increased fitness ($P < 0.1$) (Sup. Fig. 4D-E). In *E. coli*, the sulfur found in Fe-S clusters is derived from cytosolic L-cysteine mainly through the enzyme *IscS* (cysteine desulfurase) and the conversion of L-cysteine to L-alanine. We could not obtain fitness data for members of the ISC complex, likely because of its essentiality, but the deletion of the second sulfur-accepting complex – the CSD sulfur transfer system – showed reduced fitness upon selection for TPP (ΔF**,** Δ*csdA*: -0.9, Δ*csdE*: -3.4) (Fig. 3Bii). Surprisingly, we found that mutants in the SUF system for iron-sulfur cluster assembly that is employed under iron starvation or oxidative stress conditions (Outten et al., 2004) exhibited higher fitness (ΔF**,** Δ*sufS*: +0.4, Δ*sufA*: +0.7) (Fig. 3Bii).

The process of inclusion of Fe-S clusters in enzymes in *E. coli* is highly sensitive to the environment and its redox state (Crack et al., 2014; Ding et al., 2005). Transposon insertions in *soxR* and *oxyR* both showed higher fitness during selection for TPP (ΔF, respectively, +1.0 and +1.13). In addition to the *suf* operon, we found that insertions in other components of this response pathway had similar effects: the iron chelators *yaaA* and *dps* (ΔF: +1.4 and +0.6, respectively), *yhcN* (ΔF: +0.6), which is induced during peroxide stress (Lee et al., 2010), and the chaperone Hsp33 (ΔF: Δ*hslO* +0.4). During oxidative stress, both OxyR and SoxR upregulate the *fur* regulon that is regarded to play a key role in maintaining the stability of cellular iron levels (Seo et al., 2014). Ferritin and enterobactin have respectively key iron storage and iron uptake functions in *E. coli*. Bacterioferritin, whose expression is controlled by Fur via the small regulatory RNA RyhB, is important for sequestering iron thus limiting cellular oxidative damage caused by Fenton chemistry (Bou-Abdallah et al., 2002). We found that while Tn insertions in iron-storage ferritin and enterochelin esterase, which releases iron from ferric enterobactin, resulted in fitness defects (ΔF**:** Δ*ftnAB* ~ -2.0, Δ*fes* -3.3), the lack of bacterioferritin had a positive effect on fitness during selection for TPP (ΔF**:** Δ*bfr* +1.0) (Fig. 3Biv).

We found that the control of OxyR-SoxR activation also had a significant effect. OxyR is activated by oxidation, likely by oxidized glutathione, for which just a Cys^{199} – Cys^{208} disulfide bond appears to be sufficient (Georgiou, 2002). We found that the impairment of glutathione import to the cytosol and *E. coli* disulfide bond system enzymes all result in negative fitness (ΔF**,** Δ*gsiA*: -1.4, Δ*dsbA*: -2.0, Δ*dsbC*: -1.7). In contrast, the lack of *soxR* inactivation by the RSX Membrane Reducing System (Koo et al., 2003) showed positive growth phenotypes (ΔF**,** Δ*rsxA*: +1, Δ*rsxD*: +0.8, Δ*rsxE*: +0.6) (Fig. 3Biii).

Overall, our data suggest that TPP overproduction likely causes oxidative stress as previously reported (Kriek et al., 2007) and probably iron deprivation, resulting in an activation of the OxyR-SoxR and possibly Fur regulons, which could be responsible for the distinct effects on iron uptake and sequestration (Fig. 3Biv and i) due to negative feedback (Seo et al., 2014).

Population dynamics effectively identify metabolic targets

We set out to examine whether population-based screening could be used to reliably identify metabolic targets by coupling genetic modification to metabolism. We measured the extracellular titers of TPP produced *de novo* from the plasmid pTHId in 24 genomic deletions of *E. coli* MG1655 BS134 strains selected from the characterized Fe-S biogenesis network (Fig. 3B). For comparison with the newly identified genetic targets we selected the following five loci previously shown to affect thiamine biosynthesis: *mrp* (Boyd et al., 2008), *purF* (Frodyma et al., 2000), *iscAU* (Agar et al., 2000), *iscR* (Giel et al., 2013) and *yggX* (Gralnick and Downs, 2003). We expected to see reductions in biosynthetic output for Δ*mrp*, Δ*purF* and Δ*iscAUS*, and an increase for Δ*iscR*, a regulator that inhibits transcription from the *iscAUS* operon (Giel et al., 2013). We created 14 single-gene knockouts and 10 multi-gene locus lesions (Sup. Table 3). For the latter, we expected to observe an additive effect of the genes on TPP output because their products did not appear to be part of enzyme cascades (Sup. Table 4).

We observed strong correlation between the shift in fitness of a gene knockout in the multiplexed competitive growth assays and the extracellular TPP titer of the isolated gene knockout (correlation: 0.89 excluding the five test loci). Iron supply and availability strongly contributed to TPP output. Disruption of iron storage in ferritin and mobilization from enterobactin completely disrupted the TPP titer (Δ*ftnB* and Δ*fes-entF*, Fig. 4A). In contrast, knockouts of two genes induced by Fur-OxyR regulons – *bfr* and *yaaA* – whose mutants showed fitness improvements in competitive assays (Fig. 3Bi), yielded higher extracellular TPP titers (Δ*bfr*: +0.3 ±0.2, Δ*yaaA*: +0.4 ±0.1, fold difference from mean titer, see also Sup. Fig. 5) compared to the average titer across all knockouts and wild type.

We then investigated pathways involved in S-incorporation into Fe-S in *E. coli*, including free cytosolic cysteine. TPP output decreased as expected from previous reports (Agar et al., 2000; Boyd et

al., 2008; Giel et al., 2013) in the two test loci Δ*mrp* and Δ*iscAUS* (respectively, -0.8 ±0.04 and undetected, fold difference). In contrast, knockout of the repressor *iscR* resulted in a higher TPP titer (+0.7 ±0.06, Fig. 4B). We found that disrupting cysteine formation and the mobilization of its sulfur by the cellular sulfur transfer system CSD significantly decreased the TPP titer compared to the mean knockout effect (Δ*cysE:* undetected, Δ*csdA-ygdK:* -2.7 ±0.09), which is in agreement with the reduced fitness of mutants in its components (Fig. 3Bii). We previously found that mutants in the *E. coli* SUF sulfur transfer system and in various proteins involved in sulfate assimilation and export exhibited improved growth during TPP selection (Fig. 3Bii). The TPP output of single-locus knockouts agreed with these observations: the average extracellular TPP titers were higher in these knockouts compared to other knockouts and wild type (Δ*sufAS:* +0.4 ±0.2, Δ*cysN:* +0.6 ±0.1, Δ*cysP:* +0.5 ±0.04, Δ*eamA:* +0.7 ±0.2, fold difference) (Fig. 4B).

The remaining tested loci encoded proteins involved in redox control of a number of cellular functions, particularly iron homeostasis, which is tightly regulated by the SoxR-OxyR-Fur regulons (Jang and Imlay, 2011; Seo et al., 2014). The OxyR regulon is activated by an environment that favors disulfide bonds, in which glutathione plays a key role (Georgiou, 2002), and by nitric oxide via S-nitrosylation (Seth et al., 2012). Deletion of the *nar*VYZ locus, for which a mild fitness effect was calculated (**Δ**F: +0.4), showed TPP titers close to the knockouts average (Δ*narVYZ*; 0.10 ±0.3) (Fig. 4C). In contrast, we found that knockouts of the three predicted activators of the OxyR regulon exhibited significantly reduced TPP titers: *Δdsb*A, which is involved in disulfide bond formation (Depuydt et al., 2007) (-1.8 ±0.2), Δ*yliAC* (-1.7 ±0.3) and Δ*napCG* (undetected). Mutants in these loci showed reduced fitness during TPP production, whereas those in the SoxR-OxyR regulons had higher fitness. These observations suggest a role for the SoxR-OxyR regulon in diminishing thiamine biosynthesis. Indeed, knockouts of three downstream genes activated by OxyR – the peroxide-response genes *yhcN* (Lee et al., 2010) and *yggX* (Gralnick and Downs, 2003) – had higher extracellular TPP titers (Δ*yggX*: +0.6 ±0.1, Δ*argR-yhcN*: +1.3 ±0.1, fold difference) (Fig. 4C), in agreement with the shifts in fitness of their pooled mutants (Fig. 3Biii and Sup. Table 3).

Tuning of central Fe-S control

We examined the role of specific regulators of the complex *E. coli* SoxR-OxyR-Fur regulons during TPP production. We quantified the extracellular TPP titer obtained with Δ*soxR*, Δ*oxyR* and Δ*fur* single-gene knockouts introduced in *E. coli* strain TOP10, which lacks DNA restriction systems, and the production strain BS134. We found that compared to the background strains, deletion of SoxR did not lead to increased TPP titers, but the deletion of both OxyR and Fur dramatically increased pathway output in both TOP10 and BS134 (Figure 5A-B, respectively). This result pinpointed the OxyR-Fur

regulon, which directs iron homeostasis (Seo et al., 2014), as central cellular response to TPP production.

In order to determine whether this response and Fe-S cluster formation could be tuned via the identified genetic targets, we over-expressed the following genes as GFP-fusions (Kitagawa et al., 2005): *iscS, dsbA, csdA, oxyR, yhcN, yncD* and *tonB,* in the same strain BS134 used for knockouts. Compared to knockouts, TPP titers resulting from gene over-expression showed an opposite shift. In particular, the titer increased during over-expression of *iscS* and *dsbA* while decreased during over-expression of *oxyR, yhcN* and *yncD* (Figure 5C).

DISCUSSION

The study shows that by coupling the growth of genome-wide bar coded transposon mutants to a biosynthetic product it is possible to identify product-specific metabolic and regulatory bottlenecks. We show that disruptions in the assimilation of certain forms of iron and sulfur led unexpectedly to improved fitness upon selection for TPP production (Fig. 3Bi and 3Bii), a result that was supported by measuring the extracellular TPP titer from individual gene knockouts (Fig. 4). Further, over-expression of *tonB,* involved in the uptake of iron-siderophore chelates, led to a significant decrease in TPP titer (Fig. 5C). We included in our validation genes that are poorly characterized or of only predicted function, which constitute the so-called Y-genes within the *E. coli* genome, and have strong fitness phenotypes with regard to TPP production. Two genes in particular have predicted function in siderophore uptake (*yncD*, a TonB-dependent receptor) and aryl-sulfate assimilation (*yidJ*, a predicted sulfatase). Both *yncD* and *yidJ* knockouts exhibited TPP titers 20-25% higher than that of wild type cells (Fig. 4), further confirming that certain assimilatory activities may play a role in inhibiting TPP biosynthesis possibly through the regulator Fur.

Applied to TPP biosynthesis, genetic-metabolic coupling uncovered that the regulation of Fe-S homeostasis mediated by OxyR-Fur regulons diminishes biosynthetic output by depleting or sequestering cellular Fe-S clusters. In fact, Fe-S clusters are critical for the function of TPP pathway enzymes ThiC and ThiH. The effects of changing global regulators are notoriously idiosyncratic. Nevertheless, TPP-coupled fitness data of mutants in the regulons could be reproduced in different *E. coli* strains and production plasmids (Fig. 5A-B). Although mutants were not screened for fitness against wild type cells at the population level, 9 of 10 selected high-fitness mutants when introduced in wild type exhibited an extracellular TPP titer 0.2- to 2-fold higher (with the exception of *nar*VYZ), while excluding *yidJ* all other low-fitness mutants exhibited significantly lower titers (Fig. 4). We also

demonstrate that population-wide mutant fitness data could serve as basis for tuning a heterologous pathway through over-expression of genetic targets from multi-copy plasmids (Fig. 5C).

The characterization of how important parts of the *E. coli* stress response and nutrient assimilation affect a complex heterologous pathway and the identification of poorly characterized genes, which today still represent ~30% of the *E. coli* genome, contributing to this response demonstrate the power of genetic-metabolic coupling for the identification of target genes for the metabolic engineering of a specific biosynthetic product.

EXPERIMENTAL PROCEDURES

Strains, plasmids and cultivation media

Plasmid DNA, PCR products and gel extractions were prepared with appropriate kits (Qiagen ©). Synthetic oligonucleotides were purchased from Integrated DNA Technologies (IDT, Inc). In-frame GFP fusions of selected genes for expression from multi-copy plasmid were obtained from the ASKA collection (Kitagawa et al., 2005). Single-locus deletions were constructed by employing lambda Red recombineering. The thiamine selection system (pTPP_Bios) was described in a previous study from our laboratory (Genee et al., 2016). The TPP biosynthetic pathway was assembled from native *E. coli* MG1655 genes via PCR amplification. A detailed description of the molecular tools is provided in Sup. Info.

Analytical quantification of TPP

Extracellular TPP concentration was measured using a modified thiochrome-high-performance liquid chromatography (HPLC) assay procedure. Briefly, fresh colonies carrying either pTHId or pTHIm were inoculated into 1 mL MOPS rich media lacking thiamine (see above) and containing 30 µg/mL chloramphenicol and 50 µg/mL spectinomycin in deep-well culture plates and then cultivated twice to saturation in a 48-hour period (Sup. Info.).

Construction of Tn-mutants, TnSeq and BarSeq protocols

For construction of tagged transposon mutants of *E. coli* MG1655 and bioinformatic analysis of NGS data we relied on reported methodology and tools (Wetmore et al., 2015). BarSeq reads are converted into a table reporting the number of times each bar code is observed using a custom Perl script (MultiCodes.pl). Given a table of bar codes, where they map in the genome, and their counts in each sample, strain and gene fitness values are estimated with a custom R script (FEBA.R) (Sup. Info.). The bioinformatic tools can be accessed at https://bitbucket.org/berkeleylab/feba.

Calculation of gene fitness

The gene fitness is the weighted average of Tn mutant strains fitness within the gene, that is strains with more reads have less noisy fitness estimates and are weighted more highly. In more detail, we first selected a subset of strains and genes that have adequate coverage in the time-zero samples (2 reads per strain and 20 reads per gene are considered adequate). Only strains that lie within the central 10 to 90% of a gene coding region are considered. Then, for each sample:

Strain fitness = $Log_2(n_{After} + \sqrt{\psi}) - Log_2(n_0 + \sqrt{\frac{1}{\psi}})$

where ψ is a "pseudocount" to control for very noisy estimates when read counts are very low. (Further details can be found in Sup. Info. and Wetmore et al., 2015).

ABBREVIATIONS

KO: knockout, Tn: transposon, aa: amino acid, DNA: deoxyribonucleic acid, RNA: ribonucleic acid, RBS: ribosome binding site.

SUPPLEMENTAL INFORMATION

Refer to the web version on PubMed Central for supplementary materials.

AUTHOR CONTRIBUTIONS

Conceptualization, S.C. and M.O.A.S.; Methodology, S.C. and M.O.A.S.; Investigation, S.C. and F.G.T.; Writing – Original Draft, S.C.; Writing – Review & Editing, S.C. and M.O.A.S.; Funding Acquisition, M.O.A.S.

Competing financial interests

The authors have no competing financial interests to declare.

ACKNOWLEDGMENTS

This research was funded by the European Union Seventh Framework Programme (FP7-KBBE-2013-7-single-stage) under grant no. 613745 "Promys" and the Novo Nordisk Foundation (NNF) grant no. 11355-444 "Biobase". F.G.T acknowledges NNF Ph. D. fellowship grant no. NNF16CC0020908. M.O.A.S acknowledges additional funding from the NNF.

We thank Dr. Adam Deutschbauer (Lawrence Berkeley National Laboratory – Berkeley, USA) for kindly providing strains and plasmids utilized here for transposon mutagenesis. We also thank H. Genee and L. Gronenberg (Biosyntia ApS) for donating the source molecular parts for producing and selecting thiamine derivatives and for establishing analytical protocols. We also thank Dr. Lei Yang and Pernille Smith (iLoop – DTU Biosustain) for constructing strain BS134, and Dr. Anna Koza for technical support with Illumina sequencing.

REFERENCES

Agar, J.N., Krebs, C., Frazzon, J., Huynh, B.H., Dean, D.R., Johnson, M.K., April, R. V, Re, V., Recei, M., and May, V. (2000). IscU as a Scaffold for Iron - Sulfur Cluster Biosynthesis□: Sequential Assembly of [2Fe-2S] and [4Fe-4S] Clusters in IscU †. 7856–7862.

Biedendieck, R., Malten, M., Barg, H., Bunk, B., Martens, J.H., Deery, E., Leech, H., Warren, M.J., and Jahn, D. (2010). Metabolic engineering of cobalamin (vitamin B12) production in Bacillus megaterium. Microb. Biotechnol. *3*, 24–37.

Bou-Abdallah, F., Lewin, A.C., Le Brun, N.E., Moore, G.R., and Dennis Chasteen, N. (2002). Iron detoxification properties of Escherichia coli bacterioferritin. Attenuation of oxyradical chemistry. J. Biol. Chem. *277*, 37064–37069.

Boyd, J.M., Pierik, A.J., Netz, D.J.A., Lill, R., and Downs, D.M. (2008). Bacterial ApbC can bind and effectively transfer iron-sulfur clusters. Biochemistry *47*, 8195–8202.

Chatterjee, A., Li, Y., Zhang, Y., Grove, T.L., Lee, M., Krebs, C., Booker, S.J., Begley, T.P., and Ealick, S.E. (2008). Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. Nat. Chem. Biol. *4*, 1–8.

Chen, Y., Xu, D., Fan, L., Zhang, X., and Tan, T. (2015). Manipulating multi-system of NADPH regulation in Escherichia coli for enhanced S - adenosylmethionine production. RSC Adv. *5*, 41103–41111.

Chupka, W.A., Jivery, W.T., and Introduction, I. (1988). Toxic DNA Damage by Hydrogen Peroxide Through the Fenton Reaction in Vivo and in Vitro. *240*.

Crack, J.C., Thomson, A.J., and Brun, N.E. Le (2014). Iron − Sulfur Clusters as Biological Sensors□: The Chemistry of Reactions with Molecular Oxygen and Nitric Oxide. Acc. Chem. Res.

Depuydt, M., Leonard, S.E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., Messens, J., Carroll, K.S., and Collet, J. (2007). A Periplasmic Reducing System Protects Single Cysteine Residues from Oxidation. Science (80-.). 8–11.

Ding, H., Harrison, K., and Lu, J. (2005). Thioredoxin Reductase System Mediates Iron Binding in IscA and Iron Delivery for the Iron-Sulfur Cluster Assembly in IscU *. J. Biol. Chem. *280*, 30432–30437.

Dougherty, M.J., and Downs, D.M. (2006). A connection between iron – sulfur cluster metabolism and the biosynthesis of 4-amino-5- hydroxymethyl-2-methylpyrimidine pyrophosphate in Salmonella enterica. Microbiology 2345–2353.

Fletcher, E., Krivoruchko, A., and Nielsen, J. (2016). Industrial Systems Biology and Its Impact on Synthetic Biology of Yeast Cell Factories. Biotechnol. Bioeng. *113*, 1164–1170.

Frodyma, M., Rubio, A., and Downs, D.M. (2000). Reduced Flux through the Purine Biosynthetic Pathway Results in an Increased Requirement for Coenzyme A in Thiamine Synthesis in Salmonella enterica Serovar Typhimurium. J. Bacteriol. *182*, 236–240.

Genee, H.J., Bali, A.P., Petersen, S.D., Siedler, S., Bonde, M.T., Gronenberg, L.S., Kristensen, M., Harrison, S.J., and Sommer, M.O.A. (2016). Functional mining of transporters using synthetic selections. Nat. Chem. Biol. *12*, 1015–1022.

Georgiou, G. (2002). How to Flip the (Redox) Switch Minireview. Cell *111*, 607–610.

Giel, J.L., Nesbit, A.D., Mettert, E.L., Fleischhacker, A.S., Wanta, B.T., and Kiley, P.J. (2013). Regulation of iron-sulfur cluster homeostasis through transcriptional control of the Isc pathway by [2Fe-2S]-IscR in Escherichia coli. Mol. Microbiol. *87*, 478–492.

Gralnick, J.A., and Downs, D.M. (2003). The YggX Protein of Salmonella enterica Is Involved in Fe (II) Trafficking and Minimizes the DNA Damage Caused by. J. Biol. Chem. *278*, 20708–20715.

418 Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists
419 using DAVID bioinformatics resources. Nat. Protoc.

420 Imlay, J.A. (2015). Transcription Factors That Defend Bacteria Against Reactive Oxygen Species. Annu. Rev.
421 Microbiol. *69*.

422 Jang, S., and Imlay, J.A. (2011). Hydrogen peroxide inactivates the Escherichia coli Isc iron-sulfur assembly system,
423 and OxyR induces the Suf system to compensate. Mol. Microbiol. *78*, 1448–1467.

424 Keseler, I.M., Bonavides-Martínez, C., Collado-Vides, J., Gama-Castro, S., Gunsalus, R.P., Johnson, D.A.,
425 Krummenacker, M., Nolan, L.M., Paley, S., Paulsen, I.T., et al. (2009). EcoCyc: a comprehensive view of
426 Escherichia coli biology. Nucleic Acids Res. *37*, D464–D470.

427 Kitagawa, M., Ara, T., Arifuzzaman, M., and Ioka-nakamichi, T. (2005). Complete set of ORF clones of Escherichia
428 coli ASKA library (A Complete S et of E . coli K -12 ORF A rchive): Unique Resources for Biological Research.
429 DNA Res. *299*, 291–299.

430 Koo, M., Lee, J., Rah, S., Yeo, W., Lee, J., Lee, K., Koh, Y., Kang, S., and Roe, J. (2003). A reducing system of the
431 superoxide sensor SoxR in Escherichia coli. EMBO J. *22*.

432 Kriek, M., Martins, F., Leonardi, R., Fairhurst, S.A., Lowe, D.J., and Roach, P.L. (2007). Thiazole Synthase from
433 Escherichia coli AN INVESTIGATION OF THE SUBSTRATES AND PURIFIED PROTEINS REQUIRED FOR.
434 J. Biol. Chem. *282*, 17413–17423.

435 Lechner, A., Brunk, E., and Keasling, J. (2016). The Need for Integrated Approaches in Metabolic Engineering.
436 Cold Spring Harb Pespect Biol.

437 Lee, J., Hiibel, S.R., Reardon, K.F., and Wood, T.K. (2010). Identification of stress-related proteins in Escherichia
438 coli using the pollutant cis -dichloroethylene. J. Appl. Microbiol. *108*, 2088–2102.

439 Lin, Z., Xu, Z., Li, Y., Wang, Z., Chen, T., and Zhao, X. (2014). Metabolic engineering of Escherichia coli for the
440 production of riboflavin. Microb. Cell Fact. *13*.

441 Nam, H., Lewis, N.E., Lerman, J.A., Lee, D., Chang, R.L., Kim, D., and Palsson, B.O. (2012). Network Context and
442 Selection in the Evolution to Enzyme Specificity. Science (80-.). *6185*.

443 Oh, J., Fung, E., Price, M.N., Dehal, P.S., Davis, R.W., Giaever, G., Nislow, C., Arkin, A.P., and Deutschbauer, A.
444 (2010). A universal TagModule collection for parallel genetic analysis of microorganisms. Nucleic Acids Res. *38*,
445 e146.

446 Otero, M., and Nielsen, J. (2013). Genome-scale modeling enables metabolic engineering of Saccharomyces
447 cerevisiae for succinic acid production. J. Microbiol. Biotechnol. 735–747.

448 Outten, F.W., Djaman, O., and Storz, G. (2004). A suf operon requirement for Fe – S cluster assembly during iron
449 starvation in Escherichia coli. Mol. Microbiol. *2004*, 861–872.

450 Seo, S.W., Kim, D., Latif, H., O'Brien, E.J., Szubin, R., and Palsson, B.O. (2014). Deciphering Fur transcriptional
451 regulatory network highlights its complex role beyond iron metabolism in Escherichia coli. Nat. Commun. *5*, 4910.

452 Seth, D., Hausladen, A., Wang, Y., and Stamler, J.S. (2012). Endogenous Protein S-Nitrosylation in E . coli:
453 Regulation by OxyR. Science (80-.).

454 Tatusova, T., Ciufo, S., Fedorov, B., O'Neill, K., and Tolstoy, I. (2014). RefSeq microbial genomes database: New
455 representation and annotation strategy. Nucleic Acids Res. *42*, 553–559.

456 Townshend, B., Kennedy, A.B., Xiang, J.S., and Smolke, C.D. (2015). High-throughput cellular RNA device
457 engineering. Nat. Methods *12*.

458 Wetmore, K.M., Price, M.N., Waters, R.J., Lamson, J.S., He, J., Hoover, C. a, Blow, M.J., Bristow, J., Butland, G.,
459 and Arkin, A.P. (2015). Rapid Quantification of Mutant Fitness in Diverse Bacteria by Sequencing Randomly Bar-
460 Coded Transposons. MBio *6*, 1–15.

461

FIGURE LEGENDS

Figure 1. Genetic-metabolic coupling design.
A) Quantification of extracellular TPP titer with pTHId and pTHIm (cyan and red, respectively) compared to a control (no plasmid, grey) in *E. coli* MG1655 and BS134 strains (n = 4-6) **B)** Graphical scheme of the experimental design. A wild type *E. coli* strain carrying the dual riboswitch (pTPP_Bios) is subject to mutagenesis via a tagged Tn5 transposon. Following library characterization, the TPP production plasmid (pTHIm) or the backbone plasmid (pMck) is electroporated in the pool of mutants to obtain respectively the fTpp and fMck libraries. These libraries are finally used for fitness assays in absence or presence (^{CS}) of selection for TPP production.

Figure 2. Gene knockouts with significant fitness shift during TPP production.
A) Data analysis pipeline: significant fitness values are identified and used to build a phenotypic binary map. Functional enrichment in gene clusters is examined by DAVID and by mining RefSeq gene descriptions for functional annotations. **B)** Hierarchical clustering of set of genes with significant fitness shift exclusively during TPP production and selection (right-most column $fTpp^{CS}$, +/- or red/cyan indicate significantly increased/decreased fitness value) (see also Sup. Fig. 2C). **C)** Extracellular TPP concentration from *E. coli* either untreated (-) or grown in presence of low (red bars, phe=20 µM, dpd=30 µM, thiU=100 µM) or high (5x) (cyan bars) concentrations of o-phenanthroline (phe and phe^{+}), 2,2'-dipyridyl (dpd and dpd^{+}) or thiourea (thiU or $thiU^{+}$) (n = 4-5) (p: two-tailed Student's t-Test).

Figure 3. Diverse effects of Fe-S clusters metabolism
A) Venn diagram of sets of genes with significant fitness for each cultivation condition (p: the P-value of Fe-S metabolism enrichment for each condition found via mining gene annotations, *: significant, bootstrap n = 10000; null hypothesis: enrichments can occur with the same probability in a random set of genes of equal size). **B)** A map of the Fe-S biogenesis and regulatory network. Four sub-networks are represented: iron supply (*i*), sulfur supply (*ii*), redox regulation of oxyR-soxR regulons (*iii*) and Fe-S metabolism (*iv*) (orange/cyan arrowheads: improved/reduced fitness, small-size: 1-2 St. Dev. from mean, big-size: > 2 St. Dev. from mean) (see also Sup. Methods and Table 2).

Figure 4. TPP quantification in a set of genetic targets.
Extracellular titers for individual locus deletions compared to the average titers across all strains (fold difference, for absolute values see Sup. Fig. 5). Strains are grouped in three different sub-networks involved in Fe-S biogenesis: iron supply (**A**), sulfur supply (**B**) and oxidative regulation (**C**). (*: Test KO, fitness error bars are reported for multi-gene loci) (Error bars: St. Dev. n = 3-5).

Figure 5. Reverse-engineering OxyR and Fe-S regulation
Extracellular titers obtained by engineering central Fe-S cluster formation and the cellular stress response. (**A-B**) TPP titer measured in knockouts of central oxidative stress response in *E. coli* TOP10 (A) and BS134 (B) strains bearing pTHIm and pTHId production plasmids (respectively cyan and orange, pink dashed line: background strain) (n = 6-8). (**C**). Extracellular TPP titer quantified in *E. coli* BS134 over-expressing several identified genetic targets in Fe-S cluster metabolism and oxidative response (cyan bars) compared to their deletion (knockouts, orange bars) (Pink dashed line: average production) (n = 6-8).

Figure 1

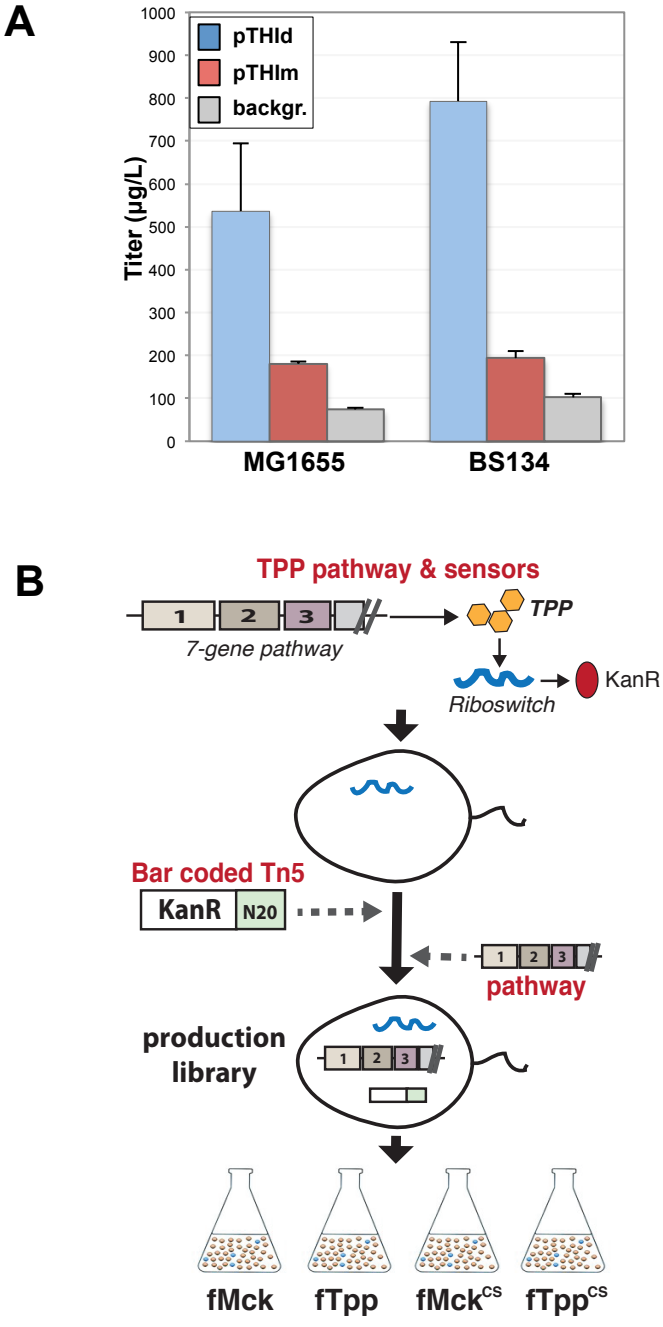
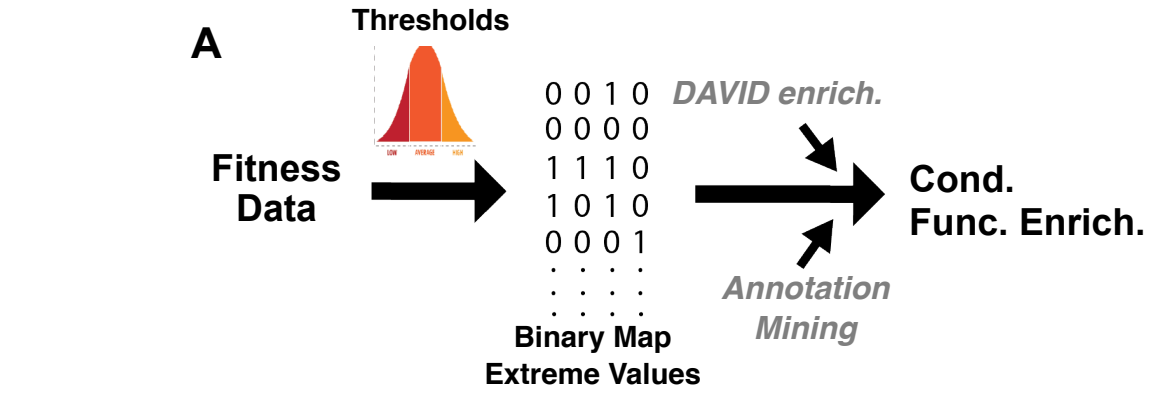
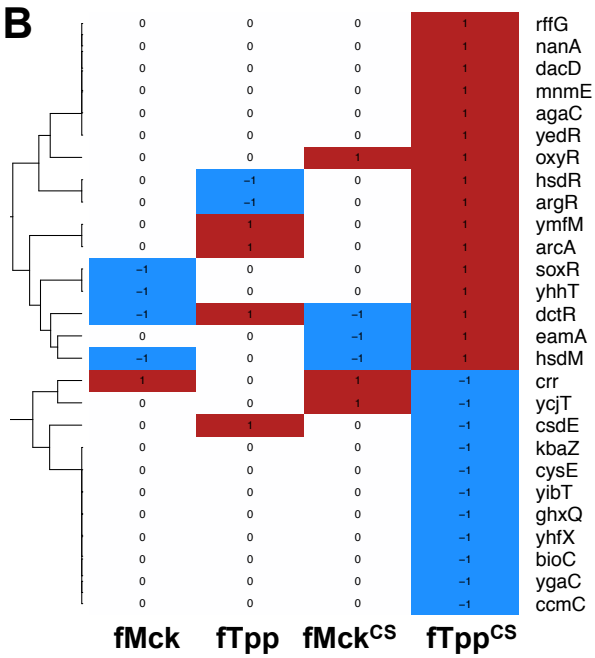

A
1000
900
800
700
600
500
400
300
200
100
0
Titer (µg/L)
pTHId
pTHIm
backgr.
MG1655
BS134
B
TPP pathway & sensors
1
2
3
7-gene pathway
TPP
Riboswitch
KanR
Bar coded Tn5
KanR
N20
pathway
production
library
fMck
fTpp
fMck^CS
fTpp^CS

Figure 2


A
Thresholds
Fitness Data
0 0 1 0
0 0 0 0
1 1 1 0
1 0 1 0
0 0 0 1
Binary Map
Extreme Values
DAVID enrich.
Annotation Mining
Cond.
Func. Enrich.


B
rffG
nanA
dacD
mnmE
agaC
yedR
oxyR
hsdR
argR
ymfM
arcA
soxR
yhhT
dctR
eamA
hsdM
crr
ycjT
csdE
kbaZ
cysE
yibT
ghxQ
yhfX
bioC
ygaC
ccmC
fMck
fTpp
fMck^CS
fTpp^CS

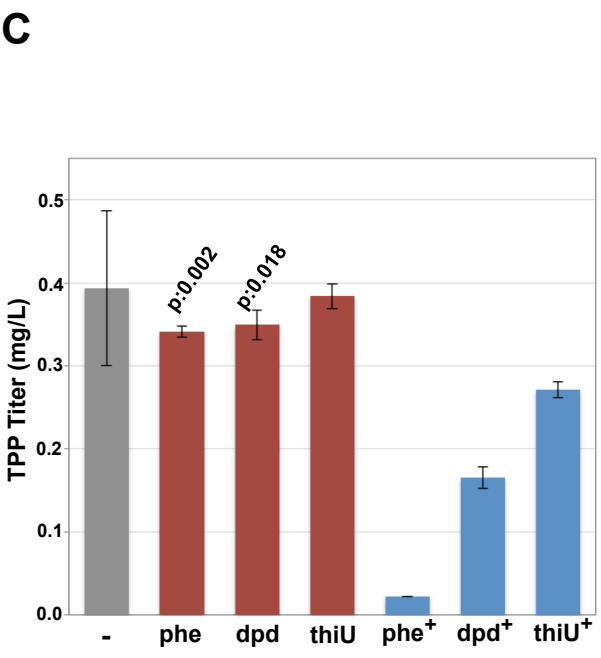

C
TPP Titer (mg/L)
0.0
0.1
0.2
0.3
0.4
0.5
p:0.002
p:0.018
-
phe
dpd
thiU
phe+
dpd+
thiU+

Figure 3

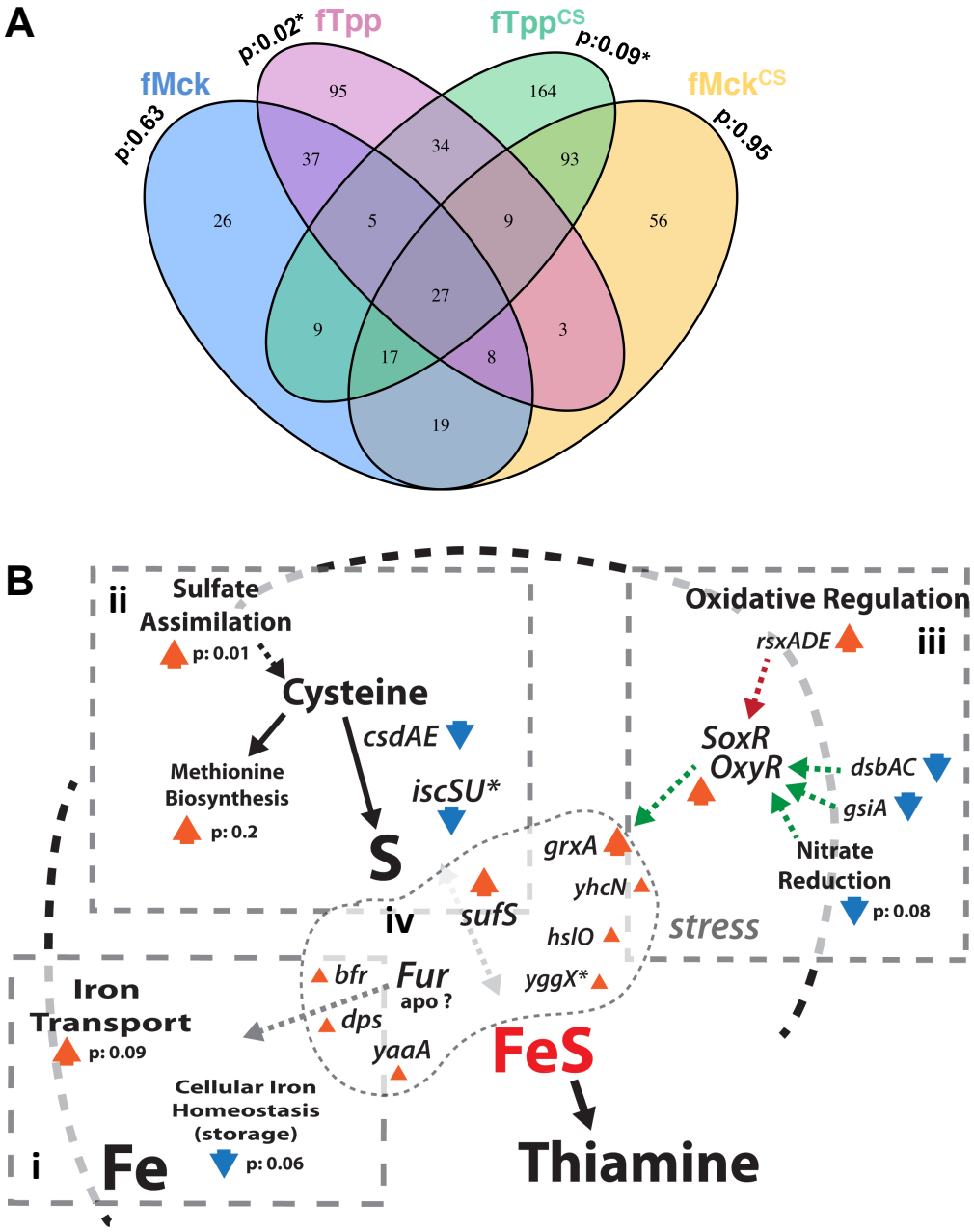

A
fMck
fTpp
fTpp^CS
fMck^CS
p:0.63
p:0.02*
p:0.09*
p:0.95
26
95
164
56
37
34
93
5
9
27
9
3
17
8
19
B
ii
Sulfate Assimilation
p: 0.01
Cysteine
csdAE
Methionine Biosynthesis
p: 0.2
iscSU*
S
iv
sufS
Oxidative Regulation
iii
rsxADE
SoxR
OxyR
dsbAC
gsiA
grxA
yhcN
hslO
yggX*
stress
Nitrate Reduction
p: 0.08
bfr
Fur
apo ?
dps
yaaA
Iron Transport
p: 0.09
Cellular Iron Homeostasis (storage)
p: 0.06
i
Fe
FeS
Thiamine

Figure 4

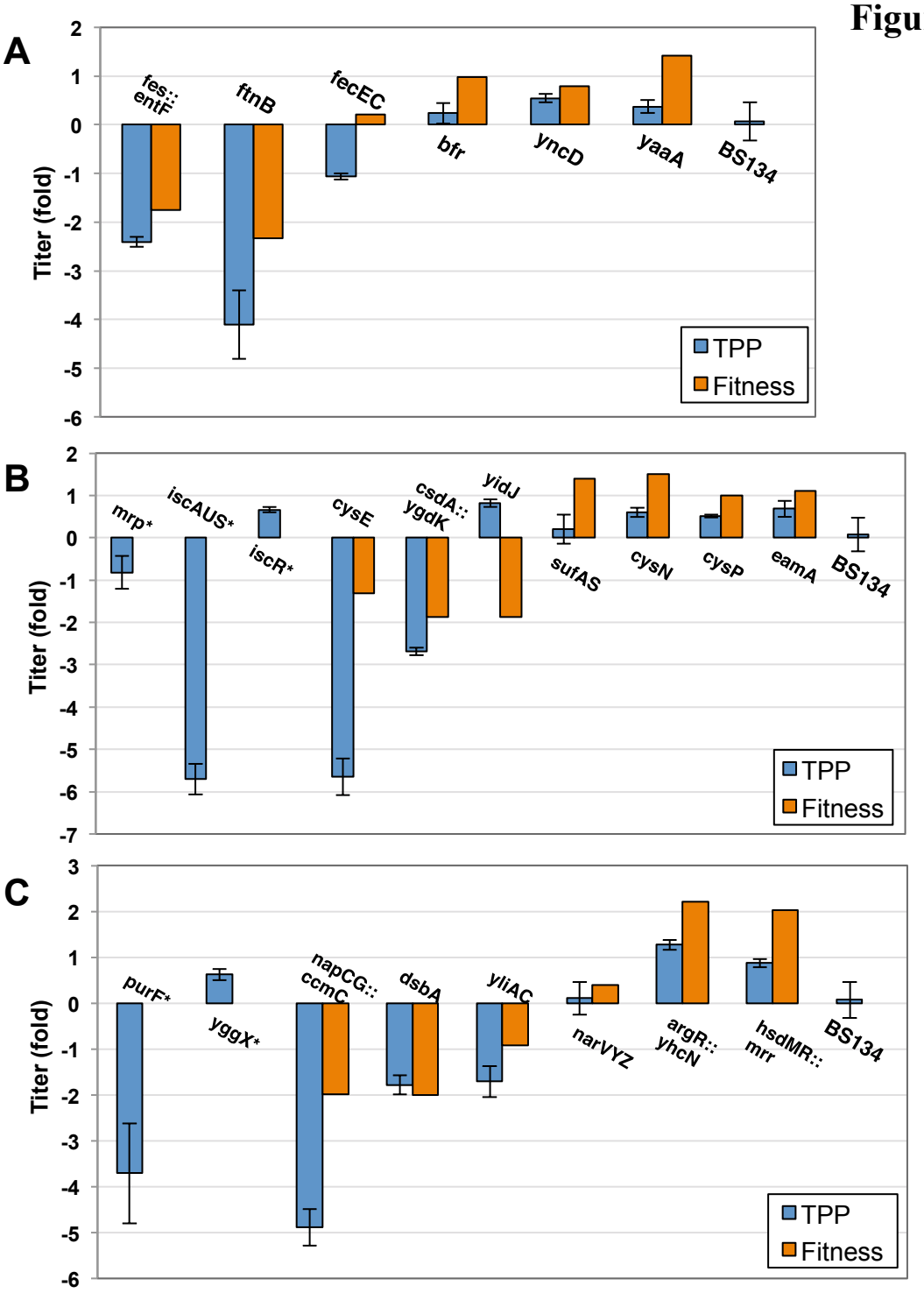
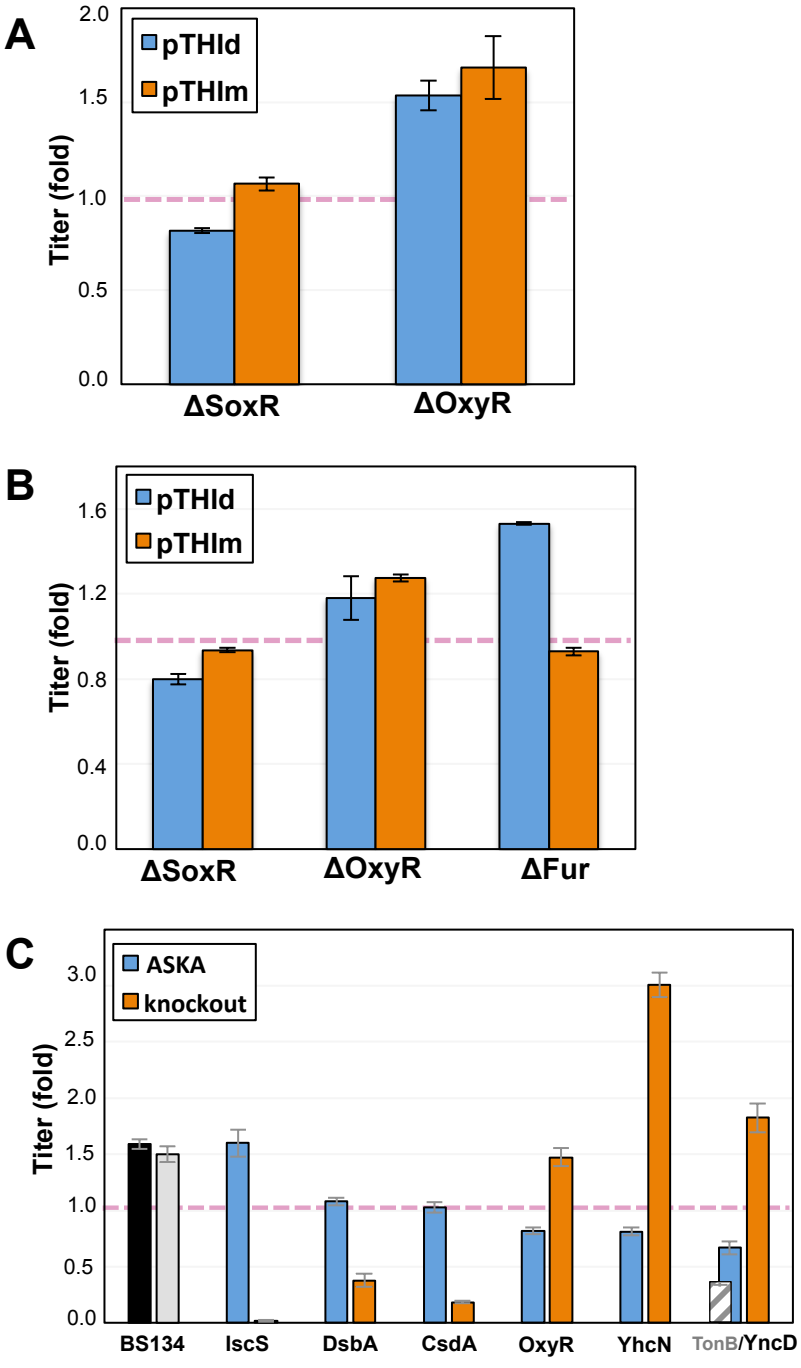

A
Titer (fold)
fes::
entF
ftnB
fecEC
bfr
yncD
yaaA
BS134
TPP
Fitness
B
Titer (fold)
mrp*
iscAUS*
iscR*
cysE
csdA::
ygdK
yidJ
sufAS
cysN
cysP
eamA
BS134
TPP
Fitness
C
Titer (fold)
purF*
yggX*
napCG::
ccmC
dsbA
yliAC
narVYZ
argR::
yhcN
hsdMR::
mrr
BS134
TPP
Fitness

Figure 5


A
pTHId
pTHIm
Titer (fold)
2.0
1.5
1.0
0.5
0.0
ΔSoxR
ΔOxyR
B
pTHId
pTHIm
Titer (fold)
1.6
1.2
0.8
0.4
0.0
ΔSoxR
ΔOxyR
ΔFur
C
ASKA
knockout
Titer (fold)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
BS134
IscS
DsbA
CsdA
OxyR
YhcN
TonB/YncD