Title

A systems approach discovers the role and characteristics of seven LysR type transcription factors in Escherichia coli

Journal

Scientific Reports, 12(1)

ISSN

2045-2322

Authors

Rodionova, Irina A

Gao, Ye

Monk, Jonathan

et al.

Publication Date

2022

DOI

10.1038/s41598-022-11134-7

scientific reports

Check for updates

OPEN

A systems approach discovers the role and characteristics of seven LysR type transcription factors in *Escherichia coli*

Irina A. Rodionova[1,2,6]✉, Ye Gao[1,2,6], Jonathan Monk[1], Ying Hefner[1], Nicholas Wong[2], Richard Szubin[1], Hyun Gyu Lim[1], Dmitry A. Rodionov[3], Zhongge Zhang[2], Milton H. Saier Jr.[2] & Bernhard O. Palsson[1,4,5]✉

Although *Escherichia coli* K-12 strains represent perhaps the best known model bacteria, we do not know the identity or functions of all of their transcription factors (TFs). It is now possible to systematically discover the physiological function of TFs in *E. coli* BW25113 using a set of synergistic methods; including ChIP-exo, growth phenotyping, conserved gene clustering, and transcriptome analysis. Among 47 LysR-type TFs (LTFs) found on the *E. coli* K-12 genome, many regulate nitrogen source utilization or amino acid metabolism. However, 19 LTFs remain unknown. In this study, we elucidated the regulation of seven of these 19 LTFs: YbdO, YbeF, YcaN, YbhD, YgfI, YiaU, YneJ. We show that: (1) YbdO (tentatively re-named CitR) regulation has an effect on bacterial growth at low pH with citrate supplementation. CitR is a repressor of the *ybdNM* operon and is implicated in the regulation of citrate lyase genes (*citCDEFG*); (2) YgfI (tentatively re-named DhfA) activates the *dhaKLM* operon that encodes the phosphotransferase system, DhfA is involved in formate, glycerol and dihydroxyacetone utilization; (3) YiaU (tentatively re-named LpsR) regulates the *yiaT* gene encoding an outer membrane protein, and *waaPSBOJYZU* operon is also important in determining cell density at the stationary phase and resistance to oxacillin microaerobically; (4) YneJ, re-named here as PtrR, directly regulates the expression of the succinate-semialdehyde dehydrogenase, Sad (also known as YneI), and is a predicted regulator of *fnrS* (a small RNA molecule). PtrR is important for bacterial growth in the presence of L-glutamate and putrescine as nitrogen/energy sources; and (5) YbhD and YcaN regulate adjacent y-genes on the genome. We have thus established the functions for four LTFs and identified the target genes for three LTFs.

Recently, the roles of uncharacterized LysR-type transcription factors (LTFs) have been identified via multiple approaches, including transcriptome analysis of uncharacterized TF (yTF)-deleted mutants (machine-learning-based), gene clustering, and the detection of DNA-binding sites[1]. The predicted yTF targets annotated as transporters and enzymes define the TF physiological role. A combination of the knowledge from EcoCyc, Fitness Browser, and iModulonDB with TF DNA-binding data provided hypotheses for the putative physiological functions of yTFs under different growth conditions[2]. Indeed, as an example, the function of one TF, PunR, was recently discovered to be an activator for *punC*, purine transporter important for adenosine utilization as a nitrogen source. Recently novel regulators for *csgD* were found by the promoter-specific screening of the 198 purified TFs including yTFs: YbiH, YdcI, YhjC, YiaU, YjgJ and YjiR. Two regulators for *csgD*: repressor YiaJ and activator YhjC (renamed RcdB) were found[3]. The SELEX method for the detection of the YbiH DNA-binding sites and antibiotics sensitivity for the yTF deletion mutant were used to predict the function for the CecR(YbiH)[4].

The RNA-seq data for the TF deletion mutants under specific growth conditions provide information about transcription affected by mutation (Fig. 1). In *Escherichia coli*, LTFs are regulators for amino acids (AAs), purine,

[1]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093-0116, USA. [2]Division of Biological Sciences, Department of Molecular Biology, University of California San Diego, La Jolla, CA 92093-0116, USA. [3]Sanford-Burnham-Prebys Medical Discovery Institute, La Jolla, CA 92037, USA. [4]Department of Pediatrics, University of California San Diego, La Jolla, CA 92093, USA. [5]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Lyngby, Denmark. [6]These authors contributed equally: Irina A. Rodionova and Ye Gao. ✉email: irodionova@ucsd.edu; palsson@ucsd.edu

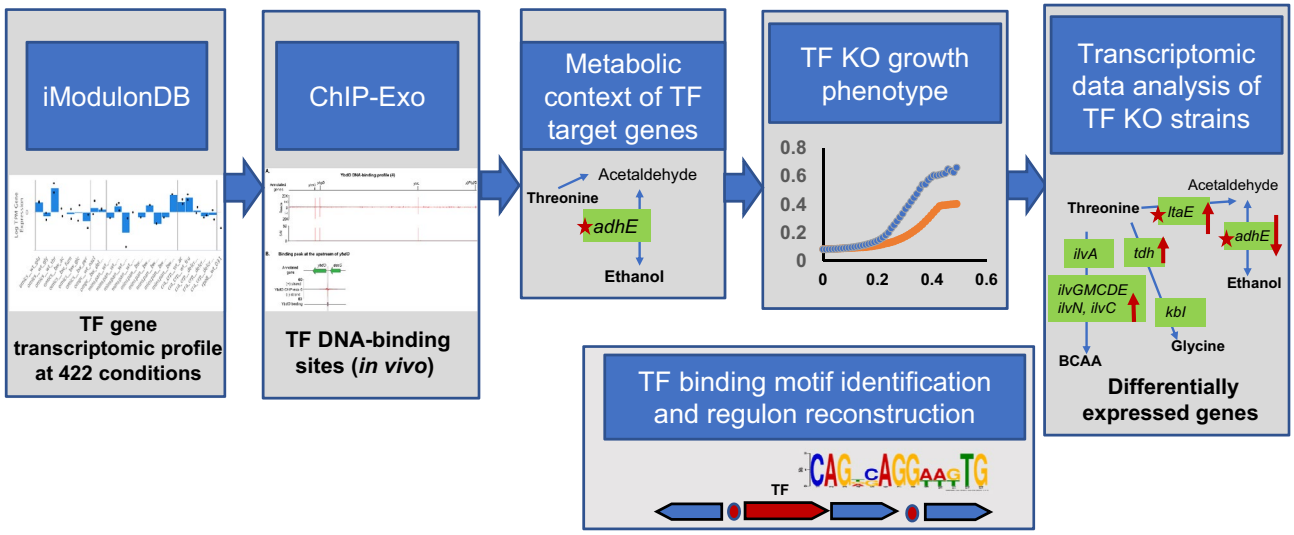

Figure 1. Systems approach for the prediction of transcription factor's function. A systematic approach to identify LysR family unknown transcriptional factors physiological function in *E. coli*.

and dicarboxylate metabolism, nitrogen assimilation (NAC), antibiotic resistance, and virulence. LTF regulatory proteins protect 50–60 bp regions with TA-rich regulatory binding sites and activation binding sites (ABS) and regulate different metabolic pathways. LTFs are known to regulate conserved gene clusters that are adjacent to the genes encoding the regulator[5], but additionally, LTF autoregulation is a common property. As one of the largest families of HTH-type regulators, LTFs contain an N-terminal helix-turn-helix DNA-binding domain and C-terminal co-inducer binding domain (Fig. S1). Given the broad conservation of LTFs, it is possible that they regulate a wide variety of target genes with diverse physiological functions using common regulatory features.

LTFs account for 16.7% (47 out of 280) of the total number of transcription factors in *Escherichia coli* K-12[6]. Out of the 47 LTFs, 26 (AbgR, AllS, ArgP, Cbl, CynR, CysB, Dan, DmlR, DsdC, GcvA, HcaR, HdfR, IlvY, LeuO, LrhA, LysR, MetR, Nac, NhaR, OxyR, PerR, PgrR, QseA, QseD, TdcA, XapR) have known functions and the majority were shown to regulate adjacent genes[7]. However, there are still many uncharacterized TFs belonging to the LysR-type family, which require further studies to determine their regulatory functions. This effort is also important for the reconstruction of the transcriptional regulatory network (TRN). The set of the six yTFs from LysR family (YbdO, YbeF, YcaN, YbhD, YgfI, YiaU) upregulated in the presence of L-threonine as supplement (iModulonDB, PRECISE) to minimal growth medium and YneJ, transcriptional analysis for the regulation of putrescine/L-glutamate utilization in minimal medium were characterized in our research.

Escherichia coli is a representative of the commensal mammalian intestinal microbiota and is the best characterized model gram-negative bacterium. Nutrient starvation conditions are important for the gut microbiome bacterial community as they cause stress, activating different survival mechanisms[8], and TRNs rewire the metabolism under different conditions. iModulonDB is a collection of *E. coli* MG1655 transcriptomics data for different growth medium and stress conditions[9]. Machine-learning based ICA data analysis (iModulonDB), RNA-seq data for the TF knockout, and ChIP-exo data (in vivo DNA-binding sites) are useful resources for LTF characterization. Recently, the ChIP-exo results for verified uncharacterized TFs in *E. coli* MG1655 was published[1]. The gene expression profiling for LTFs under multiple growth conditions in iModulonDB provides important information for predicting the growth conditions to verification of function for the yTFs[9].

We performed systems analysis for the prediction of LTF functional role using data combined from previously published chromatin immunoprecipitation with exonuclease treatment (ChIP-exo) detected LTF DNA-binding sites, transcriptome data from iModulonDB (Fig. 1), RNA-seq analysis for LTF deletion mutants (experimental data), and genome cluster analysis using the bioinformatics tools (Fig. 1). We generated RNA-seq data for seven LTF mutants (*ybdO, ybeF, ycaN, ybhD, ygfI, yiaU,* and *yneJ*), analyzed the conserved genome clustering with LTF genes, and detected conserved genes that were differentially expressed in the LTF knockout. Accordingly, we generated a hypothesis about the possible regulatory targets of three LTFs and their physiological functions in response to the high L-threonine concentration in minimal medium (Fig. 1). The LTF regulation in response to the imbalance of Thr and catabolism for energy has been studied. The LTF deletion effect on the regulation of biochemical pathways of the Thr utilization and metabolism of glycerol, citrate, and formate (Fig. 1) had been investigated in *E. coli* BW25113.

Here, using a transcriptomic analysis systems approach, we identify the LTFs, YbdO, YgfI, YcaN, YbhD, important for Thr utilization pathway in minimal medium and regulating catabolism of citrate, pyruvate, formate or malate, and YiaU, YbeF important for liposaccharide modification and flagella biosynthesis and YneJ (PtrR), that directly controls *sad* gene expression in response to GABA, an intermediate of the Ptr utilization pathway.

For the detailed analysis, we investigated the effect of the YneJ transcriptional response in minimal medium under nitrogen starvation conditions. The *yneJ-sad* cluster is conserved in enterobacterial genomes, and Sad is involved in putrescine utilization as nitrogen source[10]. Many bacterial species, including *E. coli*, can simultaneously utilize L-glutamate (Glu) and the polyamine putrescine (Ptr) under carbon/nitrogen starvation conditions.

Glu is also essential for tetrahydrofolate polyglutamylation[11]. Ptr is important for bacterial growth and for efficient DNA replication, transcription, and translation[12,13] and plays an important role in maintaining compact conformations of negatively charged nucleic acids[14]. Ptr is also involved in multiple antibiotic resistance mechanisms under stress conditions[15]. The *puuA*, *puuD, puuE,* and *puuP* genes in *E. coli* are induced by Ptr and regulated at the transcriptional level by the Ptr-responsive repressor PuuR[16,17]. The expression of *sad* (*yneI*) is induced by the addition of Ptr to the medium[18]; however, it is not regulated by PuuR, and a transcriptional regulator for *sad* had not been described before this work.

The predicted PtrR binding site in the *sad* promoter region was confirmed via an in vitro binding assay with the purified PtrR protein and using the ChIP-exo assay. We further compared the whole genome transcriptional response of the *ptrR* knockout and wild type *E. coli* strains to carbon/nitrogen starvation in the presence of Ptr and L-glutamate using RNA-Seq analyses, and the PtrR DNA-binding site was predicted for *fnrS*, encoding small regulatory RNA. The physiological roles of PtrR regulation and antibiotic resistance are discussed.

Results

An integrated systems-approach uncovers LysR-type transcription factors in *E. coli* K-12. Previously, we generated a list of candidate transcriptional factors (TFs) from uncharacterized genes ("y-genes") using a homology-based algorithm[19]. Among these TFs, it was predicted that YeiE, YfeR, YidZ, YafC, YahB, YbdO, YbeF, YbhD, YcaN, YdcI, YdhB, YeeY, YfiE, YgfI, YhaJ, YhjC, YiaU, YneJ, and YnfL belong to the LTF based on Hidden Markov Model. YdhB (re-named PunR) function had been shown to be an activator for the purine transporter, PunC[1,2]. Recently YdcI (*Salmonella enterica*) function was shown related to biofilm formation[20]. We further chose 7 uncharacterized LysR-type TFs (yLTFs), which have transcriptional responses (increased mRNAs) to the presence of a L-threonine (Thr) in M9 medium (iModulonDB, PRECISE2). L-threonine is an important source for L-serine, L-glycine, branched-chain amino acid biosynthesis, and formate, which is important for anaerobic respiration. To elucidate roles of each yLTF and genome-wide target genes, we performed gene expression profiling via RNA-Seq and ChIP-exo detection of the LTF DNA-binding and growth phenotyping. The overall workflow is shown in Fig. 1.

YbdO (CitR) regulatory effects are involved in citrate utilization and YbeF is involved in the flagella biosynthesis. The physiological function for LTFs CitR and YbeF was previously unknown. The YbdO (CitR) ChIP-exo result detected the peak for DNA-binding upstream of *citR-ybdNM* (Fig. 2A,B), that further was confirmed by the transcriptional analysis. The other ChiP-exo detected peaks for possible suggested genes regulation were not found in DEGs for *citR* deletion mutant strain. The *citR* (regulated by HNS) and *ybdNM* operons are conserved adjacent genes in Enterobacteriaceae (Fig. 2D) and the *ybdN* promoter was predicted to be regulated by FNR (EcoCyc). Further, gene expression profiling showed that *citR* deletion leads to strong upregulation of the *ybdNM* operon and the adjacent genes *citCDEF.* Interestingly that genome context analysis evolved that *citR* is conserved with the citrate lyase encoding genes cluster *citCDEF.* (Fig. 2C). Therefore, this result confirmed autoregulation of *citR-ybdNM,* and regulation for citrate lyase which is involved in anaerobic metabolism. From PRECISE2, upregulation of *citC* was detected when *E. coli* MG1655 strain was grown in M9 medium supplemented with Thr, suggesting that citrate utilization is important for the catabolizm of Thr. The YbdM has been predicted as yTF (Uniprot) and probably related to *cit* operon regulation. The DNA-binding site upstream of the *citR* gene was predicted by the phylogenetic footprinting method (Fig. S2), suggesting CitR is an autoregulator and *ybdMN* regulation.

To test the regulation of the citrate lyase encoding gene cluster (*citEFG*), the phenotype for growth in the M9 medium supplemented with citrate at low pH (pH 6.5) showed an effect of *citR* deletion on citrate utilization, suggesting artificial de-repression of citrate lyase (Fig. 2E-F), but no difference for the growth in M9 medium supplemented with citrate was found at pH 7.5. Additionally, *citR* mutation decreases the *E. coli* BW25113 fitness phenotype for motility in LB medium and affects growth using glycolate as carbon source (Fitness Browser, fit.genomics.lbl.gov) and microaerobic utilization of formate as carbon source (Fig. 7).

CitR and YbeF are paralogs (identity 30%), presented in conserved gene clusters in Enterobacterial genomes with citrate lyase operon *cit* (*citCDEFXGT*, Figs. 2C, S3). Interestingly the transcriptomic analysis of the *citR* and *ybeF* deletion mutants showed that the expression level of *lrhA* was substantially decreased (− 4.3-fold and − 3.5 -fold, respectively, Table 1, Supplementary data), likely indicating that YbeF and CitR has affected *lrhA* regulation (CitR/YbeF DNA-binding was not detected/predicted). The FlhDC iModulon includes FlhDC regulated genes that were upregulated in the *ybeF* and *citR* deletion mutants (Fig. S4B). LrhA is an LTF repressor for *flhDC*, flagellar biosynthesis genes responsible for motility and chemotaxis[21]. The deletion of *ybeF* showed downregulation of *lrhA* and upregulation of *flhC.* The iModulon FlhDC (Fig. S4) was substantially upregulated in *ybeF* mutant strain, suggesting connection for YbeF transcriptional regulation and flagella biosynthesis.

YcaN and YbhD regulators mutant characterization. The *ybhD* and *ycaN* mutant strains and WT were collected at the late-exponential phase after growth in the M9 medium supplemented with Thr. The conserved gene cluster *ycaN-ycaK-ycaM* with adjacent *ycaC* and *ycaD* genes was detected in the *Escherichia coli* K12 and *Shigella boydii* genomes. We noticed that in the *ycaN* deletion mutant strain differentially expressed genes *ycaC* (downregulated) and *ycaD, focA* (encoding formate channel) (upregulated) are genes adjacent to *ycaN.* It is interesting that *ycaK, ycaC,* and *ycaD* are predicted to be regulated by Nac (belonging to LTF), as Nac function is nitrogen assimilation, and *ycaC* and *ycaD* are probably nitrogen assimilation function related (EcoCyc). In *ycaN* deletion mutant highly DEGs are the *tnaAB* genes, encoding L-tryptophanase and a tryptophan transporter, L-arginine degradation, *astCADBE*, an autoinducer-2 transport system (*lsrABCD, lsrR*), and the HTH-type transcriptional regulator, *galS,* the genes were strongly downregulated (Supplementary Data 1). The genes encod-

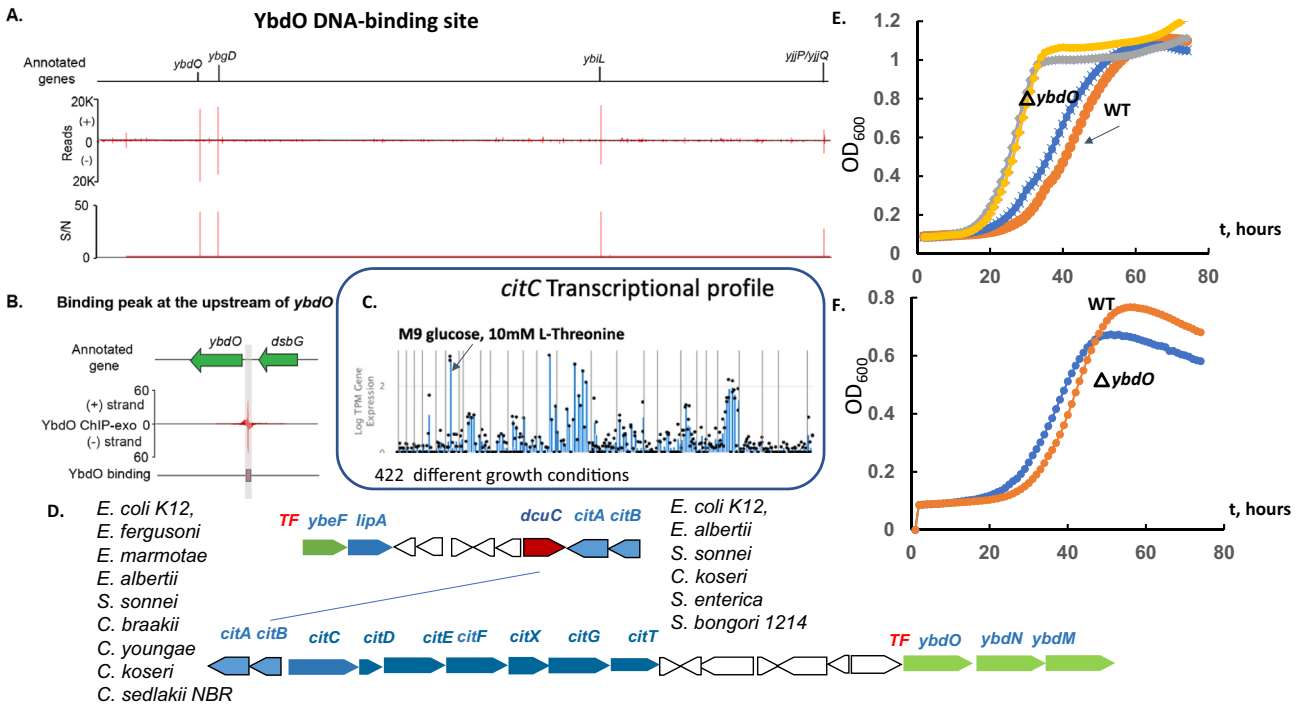

Figure 2. The systems approach for YbdO and YbeF transcriptional factors function prediction. (**A**) The genome-wide binding of YbdO. (**B**) The zoom-in of the binding site at the promoter of the gene *ybdO*. (**C**) Transcriptomic data for *citC* (citrate transporter) (**D**) *citCDEFXGT* (citrate lyase encoding) genes clustering with *ybdO-ybdMN,* and *ybeF,* the analysis across the closely related bacteria. *dcuC*-anaerobic dicarboxylate transporter, *lipA*- lipoyl synthase (**E**) The growth measurement of the *ybdO* mutant compared to the wild type BW25113 strain of *E. coli* in 96-well plates microaerobically in M9 glucose supplemented by 30 mM citrate at pH 7.5 (orange and grey line) or glycerol as carbon source (blue and yellow line). (**F**) Growth was measured in 96-well plates (M9 with glycerol) as the carbon source, supplemented by 30 mM citrate at pH 6.5 (WT-orange and *ybdO* mutant strain-blue line).

ing amino acid metabolism (L-valine biosynthesis (IlvB, IlvN), threonine dehydrogenase (Tdh), transcriptional activator (TdcA), and glycolate utilization pathway (GlcC, GlcD, GlcF)) were additionally downregulated in the *ycaN* mutant. The *ycaN* deletion mutant phenotype microarray applied for 95 carbon sources had evolved the negative phenotype for formate utilization specifically suggesting YcaN dependent *focA* regulation involved formate utilization (Fig. 7).

The *ybhD* gene is divergently oriented with respect to the conserved Enterobacterial gene cluster *ybhHI* and the putative hydrolase gene, *ybhJ* (Fig. S5A). YbhH is a 4-oxalomesaconate tautomerase homologous protein, and YbhI is a putative tricarboxylate transporter, homologous to 2-oxoglutarate/malate translocator, (id. 35%) (iModulonDB) (Fig. S5B). The YbhH encoding gene was strongly upregulated in a *ybhD* deletion mutant, as shown by transcriptomic analysis. *ybhH* and *ybhI* transcription is regulated by Nac (EcoCyc, iModulonDB) and the functional relation to nitrogen assimilation is suggested. The *ybhD* deletion strain growth phenotype in the M9 minimal medium with glycerol (carbon source), supplemented by L-malate was detected (Fig. S6).

YgfI (DhfA) regulation and glycerol and formate utilization. ChIP-exo assays previously detected DhfA binding upstream of the *dhaKLM* operon encoding the dihydroxyacetone phosphotransferase (DHAK) (Fig. 3A-B). The transcriptional activation of DhaKLM is important for glycerol utilization and M9 supplemented by Thr or L-tryptophan (Fig. 3B) (iModulonDB). DHAK in the *E. coli* MG1655 strain is involved in glycerol utilization (Uniprot). Accordingly, we decided to test the effect of the *dhfA* deletion on the growth on glucose (Fig. 3C-D) and glycerol (Fig. 3E). The resulting deficiency in growth on glycerol is potentially explained by *dhaKLM* as well as *pflB, hycBCD, hycEFG,* and *adhB* transcriptional DhfA activation, as shown by RNA-seq (Table 2), and the DhfA binding site was predicted upstream of those genes (Table S1). The DEGs detected by RNA-seq showed substantial downregulation of formate fermentation related genes (Fig. 3F) encoding pyruvate-formate lyase (*pflB*), fumarate reductase (*frdABCD*), formate hydrogenlyase (*hycBCD, hycEFG),* and the regulator (*hycA*), as well as hydrogenase encoding genes (*hybABC, hybEF,* and *hybO*) and the gene encoding protein involved in maturation of all hydrogenases isozymes (*hypB*) (Fig. 3F) and the *dhaKLM* operon (Table 2, Supplementary Data). *dhfA* deletion had little effect on growth using glucose as the carbon source (Fig. 3C) and with Thr supplement (Fig. 3D) aerobically, but the *dhfA* strain had a low growth rate microaerobically without Thr supplement in minimal medium (Fig. S7).

The detected DhfA binding upstream of *dhaKLM* and the defect for the growth on glycerol (Fig. 3E) for the *dhfA* mutant confirm the DhfA-dependent transcriptional activation of DHAK, as DHAK is involved in glycerol

	Gene name	Gene function	*p* value	Base Mean	$\log^2$ Fold Change
b2289	*lrhA*	DNA-binding transcriptional dual regulator LrhA	3.89E−26	145	−4.3
b1496	*yddA*	ABC transporter family protein YddA	0.000511	13	−3.2
b4462	*ygaQ*	Putative uncharacterized protein YgaQ	4.66E−05	15	−3.0
b2368	*emrK*	Tripartite efflux pump membrane fusion protein EmrK	0.000211	16	−2.8
b2273	*yfbN*	Uncharacterized protein YfbN	0.000348	16	−2.5
b2845	*yqeG*	Putative transporter YqeG	2.13E−05	24	−2.5
b2373	*oxc*	Oxalyl-CoA decarboxylase	0.00095	11	−2.1
b4660_1	*yhiL*	Uncharacterized protein	0.000401	28	−2.0
b2310	*argT*	Lysine/arginine/ornithine ABC transporter periplasmic binding protein	0.000795	761	−1.8
b2349	*intS*	CPS-53 (KpLE1) prophage prophage CPS-53 integrase	0.000511	101	−1.6
b3043	*ygiL*	Putative fimbrial protein YgiL	0.00100	10	−1.6
b2309	*hisJ*	Histidine ABC transporter periplasmic binding protein	0.000348	216	−1.4
b1025	*dgcT*	Putative diguanylate cyclase DgcT	0.001038	122	−1.4
b0854	*potF*	Putrescine ABC transporter periplasmic binding protein	0.000346	400	−1.4
b2306	*hisP*	Lysine/arginine/ornithine ABC transporter/histidine ABC transporter, ATP binding subunit	4.90E−05	264	−1.4
b0287	*yagU*	Inner membrane protein that contributes to acid resistance	0.000166	870	−1.2
b0641	*lptE*	Lipopolysaccharide assembly protein LptE	0.000226	146	1.2
b4316	*fimC*	Type 1 fimbriae periplasmic chaperone	2.37E−05	168	1.2
b0605	*ahpC*	Alkyl hydroperoxide reductase AhpC component	5.59E−07	585	1.2
b4315	*fimI*	Putative fimbrial protein FimI	0.000279	291	1.3
b0622	*pagP*	Lipid IVA palmitoyltransferase	7.10E−05	45	1.6
b2013	*yeeE*	Inner membrane protein YeeE	0.000163	315	1.7
b1729	*tcyP*	Cystine/sulfocysteine:cation symporter	0.000397	655	1.8
b2423	*cysW*	Sulfate/thiosulfate ABC transporter inner membrane subunit CysW	5.00E−05	110	2.0
b2424	*cysU*	Sulfate/thiosulfate ABC transporter inner membrane subunit CysU	1.29E−05	245	2.0
b1950	*fliR*	Flagellar biosynthesis protein FliR	0.000674	20	2.3
b2751	*cysN*	Sulfate adenylyltransferase subunit 1	2.77E−08	297	2.4
b1879	*flhA*	Flagellar biosynthesis protein FlhA	1.37E−09	190	2.5
b2752	*cysD*	Sulfate adenylyltransferase subunit 2	2.03E−08	249	2.5
b0601	*ybdM*	ParB-like nuclease domain-containing protein YbdM	0.000248	20	2.5
b2764	*cysJ*	Sulfite reductase flavoprotein subunit	1.62E−08	381	2.6
b2750	*cysC*	Adenylyl-sulfate kinase	1.34E−08	162.1	2.6
b4110	*yjcZ*	Uncharacterized protein YjcZ	2.81E−08	351.3	2.7
b2422	*cysA*	Sulfate/thiosulfate ABC transporter ATP binding subunit	4.46E−08	296	2.7
b2763	*cysI*	Sulfite reductase hemoprotein subunit	3.02E−09	567	2.8
b1070	*flgN*	Flagellar biosynthesis protein FlgN	2.32E−06	338	2.9
b1880	*flhB*	Flagellar biosynthesis protein FlhB	4.53E−10	118	3.0
b1071	*flgM*	Anti-sigma factor for FliA (sigma(28))	7.52E−05	86	3.0
b1946	*fliN*	Flagellar motor switch protein FliN	1.42E−05	30	3.1
b1566	*flxA*	Qin prophage protein FlxA	8.91E−06	402	3.1
b2762	*cysH*	Phosphoadenosine phosphosulfate reductase	3.14E−09	145	3.2
b3525	*pdeH*	c-di-GMP phosphodiesterase PdeH	6.75E−11	176	3.3
b1072	*flgA*	Flagellar basal body P-ring formation protein FlgA	4.65E−13	225	3.3
b0615	*citF*	Citrate lyase alpha subunit	3.40E−08	20	3.3
b1941	*fliI*	Flagellum-specific ATP synthase FliI	1.66E−11	165	3.4
b1948	*fliP*	Flagellar biosynthesis protein FliP	2.90E−10	40	3.5
b1081	*flgJ*	Putative peptidoglycan hydrolase FlgJ	8.01E−15	126	3.5
b1943	*fliK*	Flagellar hook-length control protein	7.25E−10	87.3	3.7
b1083	*flgL*	Flagellar hook-filament junction protein 2	5.67E−16	3080	3.7
b1080	*flgI*	Flagellar P-ring protein	3.67E−17	263	3.7
b1942	*fliJ*	Flagellar biosynthesis protein FliJ	1.46E−05	27	3.7
b1939	*fliG*	Flagellar motor switch protein FliG	3.18E−13	213	3.8
b1940	*fliH*	Flagellar biosynthesis protein FliH	1.09E−12	73	3.8
b1938	*fliF*	Flagellar basal-body MS-ring and collar protein	1.44E−24	645	3.8
Continued					

	Gene name	Gene function	*p* value	Base Mean	log^2 Fold Change
b1945	*fliM*	Flagellar motor switch protein FliM	3.39E−18	174	3.9
b1078	*flgG*	Flagellar basal-body rod protein FlgG	2.24E−17	209	3.9
b1882	*cheY*	Chemotaxis protein CheY	2.68E−11	232	4.0
b0618	*citC*	Citrate lyase synthetase	3.64E−07	12	4.0
b1922	*fliA*	RNA polymerase sigma 28 (sigma F) factor	8.52E−18	3038	4.0
b1074	*flgC*	Flagellar basal-body rod protein FlgC	1.73E−19	130	4.1
b1887	*cheW*	Chemotaxis protein CheW	1.74E−16	345	4.1
b1944	*fliL*	Flagellar protein FliL	2.57E−11	46	4.1
b1079	*flgH*	Flagellar L-ring protein	1.32E−16	115	4.2
b1925	*fliS*	Flagellar biosynthesis protein FliS	4.36E−07	134	4.2
b4355	*tsr*	Methyl-accepting chemotaxis protein Tsr	2.59E−21	2250	4.2
b1923	*fliC*	Flagellar filament structural protein	3.74E−19	18,478	4.2
b1073	*flgB*	Flagellar basal-body rod protein FlgB	2.75E−20	268	4.3
b1077	*flgF*	Flagellar basal-body rod protein FlgF	2.31E−23	532	4.3
b1881	*cheZ*	Chemotaxis protein CheZ	1.19E−26	381	4.3
b1082	*flgK*	Flagellar hook-filament junction protein 1	1.75E−21	1771	4.3
b1924	*fliD*	flagellar filament capping protein	1.63E−16	623	4.3
b1947	*fliO*	Flagellar biosynthesis protein FliO	7.05E−16	68	4.4
b0616	*citE*	Citrate lyase beta subunit	7.69E−07	9	4.4
b1884	*cheR*	Chemotaxis protein methyltransferase	1.17E−11	158	4.4
b1075	*flgD*	Flagellar biosynthesis initiation of hook assembly	1.13E−21	400	4.5
b1883	*cheB*	Protein-glutamate methylesterase/protein glutamine deamidase	1.04E−20	415	4.5
b1921	*fliZ*	DNA-binding transcriptional regulator FliZ	7.56E−15	155	4.6
b1076	*flgE*	Flagellar hook protein FlgE	1.14E−22	677	4.6
b1890	*motA*	Motility protein A	4.60E−21	267	4.6
b1885	*tap*	Methyl-accepting chemotaxis protein Tap	9.03E−23	1642	4.6
b1888	*cheA*	chemotaxis protein CheA	1.33E−27	679	4.8
b1889	*motB*	Motility protein B	5.67E−26	410	4.8
b1886	*tar*	Methyl-accepting chemotaxis protein Tar	7.76E−31	1507	4.9
b1926	*fliT*	Flagellar biosynthesis protein FliT	0.000596	21	5.3
b0602	*ybdN*	Putative PAPS reductase/DUF3440 domain-containing protein YbdN	1.86E−22	36	5.5

Table 1. Differentially expressed genes revealed by RNA-Seq *ybdO* deletion mutant strain and wild type *E. coli* strains during growth in M9 medium with glucose as the primary carbon source and 7 mM L-threonine as supplement.

utilization[22]. We suggest re-naming YgfI to DhfA—dihydroxyacteone/formate utilization activator. The fermentation products during microaerobic growth on M9 glucose were detected in the Thr supplemented M9 medium. The analysis evolved the higher efflux of formate and acetate for the *dhfA*mutant strain (Fig. 3G), suggesting that DhfA is important for formate utilization. The DhfA dependent activation of *hyc* and *hyb* operons was important the formate utilization microaerobically as was shown by the phenotype microarray analysis (Fig. 7).

YiaU (LpsR) regulatory network and yiaU mutant growth phenotype. ChIP-exo results show the LpsR binding for the regulation of *waaP* (encoding lipopolysaccharide (LPS) biosynthesis glycero-D-manno-heptose kinase) (Fig. 4A). The results suggest that LpsR is important for LPS biosynthesis at specific conditions and the majority of DEGs encode the proteins involved in cell wall/membrane/envelope biogenesis, carbohydrate transport, energy production, and amino acid metabolism (Fig. 4B). Accordingly, gene expression for the *lpsR* mutant was lower for the genes from the operon *waaPSBOJYZU,* suggesting that LpsR is the LPS biosynthesis operon activator (Table 3); previously the regulator for *waa* operon was not known (ecocyc.org). Additionally, ChIP-exo detected binding for the genes encoding adenine transporter, *adeP*, suggests transcriptional regulation, and they were detected as DEGs for the *yiaU* mutant (Fig. 4D). We suggest re-name YiaU to LpsR—LPS biosynthesis regulator. The Biolog plates with the antibiotics were tested microaerobically in RPMI10LB medium for the possible LTF *waa* operon activation in response to the stress. LpsR was found essential for survival at high concentration of oxacillin microaerobically (Fig. 4C) in RPMI_10LB medium. WaaZ and WaaY were shown to be essential for survival with nafcillin in *E. coli*[23]. The difference for the growth in exponential/log phase for the *lpsR* mutant strain had not been detected (Fig. 4D), although LpsR regulation had affected final OD_{600} (stationary phase) in the M9 glucose medium. The phenotype (log phase) for the mutant in M9 with 0.3 M NaCl added was additionally detected (Fig. 4D).

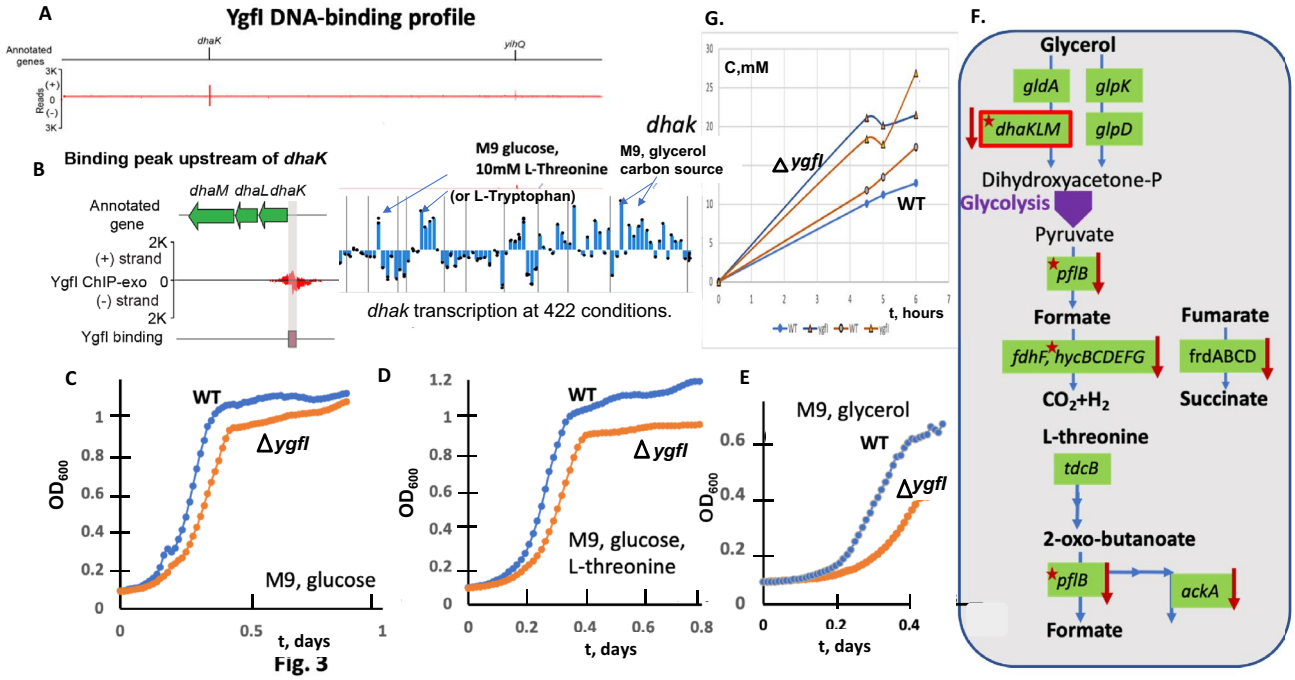

Figure 3. The systems approach for YgfI transcriptional factor function prediction. (**A**-**B**) The genome-wide binding of YgfI and the zoom-in for the *dhaK* binding site. Transcriptomic profile for the YgfI regulated *dhaK* gene. (**D**-**E**) The YgfI function was assessed based on the growth phenotype under different conditions. The growth measurement of the *ygfI* mutant (orange line) compared to the wild type BW25113 strain of *E. coli* was measured in 96-well plates in M9 medium or the same medium supplemented with 7 mM L-threonine (**C**) or M9 medium with glycerol as the carbon source (**F**) The glycerol utilization pathway – glycerol kinase (GK, *glpK*), glycerol-3phosphate dehydrogenase (G3PDH, *glpD*), glycerol dehydrogenase (GlyDH, *gldA*), dihydroacetone kinase (DHAK, *dhaKLM*). RNA-seq differentially expressed genes are marked by a red arrow. Transcriptomic data analysis for the *ygfI* mutant compared to the wild type BW25113 strain of *E. coli* shows the glycerol and L-threonine utilization pathway genes (the DEGs are marked by arrows, predicted YgfI binding site marked by red stars) regulation. (**G**) The concentration (mM) for the formate (orange line) and acetate (blue line) produced by *E. coli* BW25113 WT (circles) and *ygfI* deletion mutant (triangles) strains in M9 glucose medium supplemented by 7 mM Thr (microaerobic conditions).

The LysR family representatives are known to regulate adjacent genes, and *lpsR-yiaT* are conserved in the bacterial genomes (Fig. 4E). The RNA-seq results for the *lpsR* deletion mutant strain showed upregulation of the *yiaT* gene, encoding a predicted outer membrane protein membrane anchor for the surface display for the proteins, homologue of MipA. MipA is an MltA (murein-degrading enzyme) interacting protein.

YneJ (PtrR) regulatory effects for sad and fnrS transcription and putrescine utilization. Two distinct Ptr utilization pathways are known for *E. coli* (Fig. 5A). The first is catalyzed by the PuuA, PuuB, PuuC, and PuuD enzymes encoded by the *puuP, puuA, puuDR, puuCB, puuE* gene cluster and involves degradation of Ptr to γ-aminobutyric acid (GABA) via γ-glutamylated intermediates. The alternative pathway of Ptr degradation to GABA consists of PatA (Ptr aminotransferase) and PatD (γ-aminobutyraldehyde dehydrogenase)[24]. The PuuABCDE pathway is essential for Ptr utilization in *E. coli* using PuuP as the major Ptr transporter[25]. GABA is further utilized by two alternative 4-aminobutyrate aminotransferases (GABA-AT) encoded by *gabT* and *puuE*, and also two succinate semialdehyde dehydrogenases (SSADH) encoded by *gabD* and *sad*[26,27].

We decided to characterize in detail the YneJ (re-named PtrR) by analyzing the PtrR ChIP-exo detected binding sites[1]. The *ptrR* gene is located in a conserved gene cluster with the divergently transcribed *sad* (*yneI)* gene, which encodes succinate semialdehyde dehydrogenase and *yneH* (glsB) glutaminase (Fig. 6A and 6C), upregulated in the evolved *yneJ* mutant strain (iModulonDB)[28]. To identify and characterize DNA binding sites of PtrR in the *E. coli* genome we utilized a combined bioinformatics and experimental approach. First, we applied a comparative genomic approach of phylogenetic footprinting[27] to predict putative PtrR-binding sites in the common intergenic region of the *ptrR* and *sad* genes (Fig. 6C, Fig. S8). The *ptrR*/*sad* genes are conserved in several taxonomic groups including *Escherichia*/*Salmonella*/*Shigella*, *Citrobacter*, *Enterobacter*, and *Klebsiella*, as well as in *Pseudomonas* spp. In *E. coli* and closely related enterobacteria the *sad* gene belongs to the putative *sad-yneH* gene cluster, while in *Enterobacter* and *Citrobacter* the orthologous genes include an additional gene encoding the methyl-accepting chemotaxis protein I (serine chemoreceptor protein, Mcp) (Fig. 6C). The multiple sequence alignment of *ptrR*/*sad* upstream regions from *E. coli* and closely related enterobacteria (termed Group 1 species) contains a conserved 15-bp palindromic sequence with consensus TTCACnAATnGAGAA downstream

	Gene name	Gene function	Base mean	log^2 fold change
b2724	*hycB*	Formate hydrogenlyase subunit HycB	29.3	− 9.0
b1557	*cspB*	Qin prophage cold shock-like protein CspB	48.8	− 8.3
b1937	*fliE*	Flagellar basal-body protein FliE	66.8	7.7
b0990	*cspG*	cold shock protein CspG	10.2	− 7.6
b2727	*hypB*	Hydrogenase isoenzymes nickel incorporation protein HypB	53.6	− 7.5
b2921	*ygfI*	Putative LysR-type transcriptional regulator	21.6	− 7.3
b1922	*fliA*	RNA polymerase sigma 28 (sigma F) factor	3038.3	7.2
b3556	*cspA*	cold shock protein CspA	247.8	− 7.2
b0572	*cusC*	Copper/silver export system outer membrane channel	1017.0	7.1
b2720	*hycF*	Formate hydrogenlyase subunit HycF	37.7	− 7.1
b4335	*yjiM*	Putative dehydratase subunit	228.8	− 6.9
b1904	*yecR*	Lipoprotein YecR	6.0	6.6
b1938	*fliF*	flagellar basal-body MS-ring and collar protein	645.3	6.5
b1939	*fliG*	flagellar motor switch protein FliG	213.0	6.5
b2721	*hycE*	Formate hydrogenlyase subunit HycE	154.6	− 6.4
b1729	*tcyP*	Cystine/sulfocysteine:cation symporter	654.8	6.3
b4037	*malM*	Maltose regulon periplasmic protein	110.6	− 6.2
b1409	*ynbB*	Putative CDP-diglyceride synthase	2.4	− 6.1
b1674	*ydhY*	Putative 4Fe-4S ferredoxin-type protein	80.9	− 6.1
b2728	*hypC*	Hydrogenase 3 maturation protein HypC	5.7	− 6.1
b1072	*flgA*	Flagellar basal body P-ring formation protein FlgA	225.3	6.0
b1886	*tar*	Methyl-accepting chemotaxis protein Tar	1507.2	6.0
b1258	*yciF*	DUF892 domain-containing protein YciF	27.2	6.0
b1566	*flxA*	Qin prophage protein FlxA	402.2	6.0
b1887	*cheW*	Chemotaxis protein CheW	345.4	6.0
b2378	*lpxP*	Palmitoleoyl acyltransferase	422.4	− 5.9
b1073	*flgB*	Flagellar basal-body rod protein FlgB	267.6	5.9
b1375	*ynaE*	Rac prophage uncharacterized protein YnaE	4.5	− 5.9
b1923	*fliC*	Flagellar filament structural protein	18,478.4	5.8
b1880	*flhB*	Flagellar biosynthesis protein FlhB	118.3	5.8
b1942	*fliJ*	Flagellar biosynthesis protein FliJ	27.2	5.8
b2971	*yghG*	Lipoprotein YghG	3.4	5.8
b4380	*yjjI*	DUF3029 domain-containing protein YjjI	128.5	− 5.7
b4034	*malE*	Maltose ABC transporter periplasmic binding protein	402.0	− 5.7
b2997	*hybO*	Hydrogenase 2 small subunit	245.6	− 5.7
b1241	*adhE*	aldehyde-alcohol dehydrogenase	8441.5	− 5.7
b1925	*fliS*	Flagellar biosynthesis protein FliS	134.0	5.6
b1587	*ynfE*	Putative selenate reductase YnfE	250.8	− 5.6
b1083	*flgL*	Flagellar hook-filament junction protein 2	3080.4	5.6
b1552	*cspI*	Qin prophage cold shock protein CspI	6.3	− 5.6
b3370	*frlA*	Fructoselysine/psicoselysine transporter	2.1	− 5.5
b1589	*ynfG*	Putative oxidoreductase YnfG	87.0	− 5.5
b1890	*motA*	Motility protein A	267.0	5.4
b1940	*fliH*	Flagellar biosynthesis protein FliH	73.1	5.4
b0903	*pflB*	Pyruvate formate-lyase	23,055.8	− 5.4
b4355	*tsr*	Methyl-accepting chemotaxis protein Tsr	2250.0	5.4
b0849	*grxA*	Reduced glutaredoxin 1	8.0	5.4
b2995	*hybB*	Hydrogenase 2 membrane subunit	157.1	− 5.4
b2723	*hycC*	Formate hydrogenlyase subunit HycC	118.9	− 5.4
b0894	*dmsA*	Dimethyl sulfoxide reductase subunit A	454.5	− 5.4
b2722	*hycD*	Formate hydrogenlyase subunit HycD	50.9	− 5.4
b1944	*fliL*	Flagellar protein FliL	46.2	5.4
b4334	*yjiL*	Putative ATPase activator of (R)-hydroxyglutaryl-CoA dehdratase	50.6	− 5.3
b1946	*fliN*	Flagellar motor switch protein FliN	30.4	5.3
b4036	*lamB*	Maltose outer membrane channel/phage lambda receptor protein	191.8	− 5.3
b1757	*ynjE*	Molybdopterin synthase sulfurtransferase	335.9	− 5.3
Continued				

	Gene name	Gene function	Base mean	log^2 fold change
b1531	*marA*	DNA-binding transcriptional dual regulator MarA	33.0	5.3
b2021	*hisC*	Histidinol-phosphate aminotransferase	960.3	− 5.3
b1751	*ydjY*	4Fe-4S ferredoxin-type domain-containing protein YdjY	133.1	− 5.2
b1926	*fliT*	Flagellar biosynthesis protein FliT	21.2	5.2
b4307	*yjhQ*	KpLE2 phage-like element putative acetyltransferase TopAI antitoxin YjhQ	2.1	− 5.2
b2024	*hisA*	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino] imidazole-4-carboxamide isomerase	608.7	− 5.2
b1889	*motB*	Motility protein B	409.8	5.2
b4154	*frdA*	Fumarate reductase flavoprotein subunit	2376.0	− 5.2
b1112	*bhsA*	DUF1471 domain-containing multiple stress resistance outer membrane protein BhsA	3.5	5.2
b0621	*dcuC*	Anaerobic C4-dicarboxylate transporter DcuC	120.9	− 5.1
b1878	*flhE*	Flagellar protein	27.4	5.1
b2020	*hisD*	Histidinal/histidinol dehydrogenase	1390.9	− 5.1
b1945	*fliM*	Flagellar motor switch protein FliM	174.1	5.1
b1885	*tap*	Methyl-accepting chemotaxis protein Tap	1641.7	5.0
b2022	*hisB*	Imidazoleglycerol-phosphate dehydratase/histidinol-phosphatase	950.7	− 5.0
b0297	*eaeH*	Putative porin domain-containing protein EaeH	1.9	− 5.0
b3476	*nikA*	$Ni^{2(+)}$ ABC transporter periplasmic binding protein	126.7	− 5.0
b1080	*flgI*	Flagellar P-ring protein	263.1	5.0
b1742	*ves*	HutD family protein Ves	111.3	4.9
b1200	*dhaK*	Dihydroxyacetone kinase subunit K	127.4	− 1.4
b3634	*coaD*	Pantetheine-phosphate adenylyltransferase	40.9	2.6
b2025	*hisF*	Imidazole glycerol phosphate synthase subunit HisF	1087.6	− 4.9

Table 2. Differentially expressed genes revealed by RNA-Seq *ygfI* deletion mutant strain and wild type *E. coli* strains during growth in M9 medium with glucose as the primary carbon source and 7 mM L-threonine as supplement.

predicted sigma-E dependent promoter (Fig. 6A). We also analyzed upstream regions of *ptrR* orthologs in other enterobacterial genomes (Group 2 species), where the *sad* gene ortholog is absent and *ptrR* is co-localized with an uncharacterized MFS-family transporter gene. Further, we predicted two conserved DNA sites with similar consensus sequences located in their common intergenic region (Fig. 6C).

We further confirmed the identified putative PtrR-binding site upstream of the *sad* genes in *E. coli* and conducted genome-wide mapping of other PtrR-binding sites using the ChIP-exo method. To identify in vivo PtrR binding sites, *E. coli* was grown under glucose as the carbon source in the M9 minimal media. A total of nine PtrR-binding sites were detected in these experiments. PtrR binds in the promoter regions of the *fhuC, moeB, dhaK, fnrS, gltS/xanP,* and *ptrR/sad* genes. The experimentally identified 50-bp PtrR-binding region at *sad/ptrR* genes contains the conserved palindromic DNA motif identified via phylogenetic footprinting (Fig. S8). Comparison of this DNA motif with eight other regions containing experimentally mapped PtrR-binding regions did not reveal significant sequence similarity except for the PtrR-binding area at *fnrS*, which shares a common consensus with the identified DNA motif at *sad/ptrR*. We created multiple alignments of the upstream DNA sequences of closely related species with the beginning of the *E. coli* gene for *fnrS* and these binding sites corresponded to the ChIP-exo protected areas. The binding sites TTCACGAATCGaGAA, TTCtCGATTCGTGAA, and TgaAtGcAaCGTcAA were predicted for *ptrR*, *sad(yneI)*, and *fnrS*, respectively.

Experimental assessment of the computationally predicted PtrR DNA-binding site TTCtCGATTCGTGAA in the *sad* promoter region has been facilitated using a PtrR-binding fluorescent polarization (FP) assay (Fig. 6B). The recombinant overproduced PtrR was obtained using a strain from the ASKA collection. The PtrR purification procedure is described in Supplemental materials. The binding of the purified refolded PtrR protein to synthetic DNA fragments containing the predicted PtrR-binding site was assessed using FP in the assay mixture containing 10 mM urea (Fig. 6B). Specific binding of PtrR to the DNA (5′-GGGTTCTCGATTCGTGAAGGG-3′) was detected at 0.6 uM of PtrR in contrast to the negative control (PhrR)[29]. The fluorescent polarization assay showed binding for the PtrR to the predicted *sad* binding site and the addition of GABA leads to dissociation of PtrR from the fluorescently labeled DNA (Fig. 6B), suggesting a regulatory function for Ptr/GABA catabolism for energy. If the Ptr utilization pathway intermediate, GABA, accumulates, PtrR de-repress *sad* and *fnrS*. The upregulation for *sad* and *glsB* had been detected for the *yneJ* mutant strain for adapted *E. coli* MG1655 mutant (deletion *menF-entC-ubiC*) (Fig. 5D)[28].

An *E. coli* BW25113 (WT) and *ptrR* mutant strain growth phenotype on glutamate as the nitrogen source in minimal medium (glycerol as the carbon source) has been detected for the growth. Cells showed a growth phenotype when the *ptrR* gene was deleted under the starvation conditions (Fig. S9); a decrease in the growth

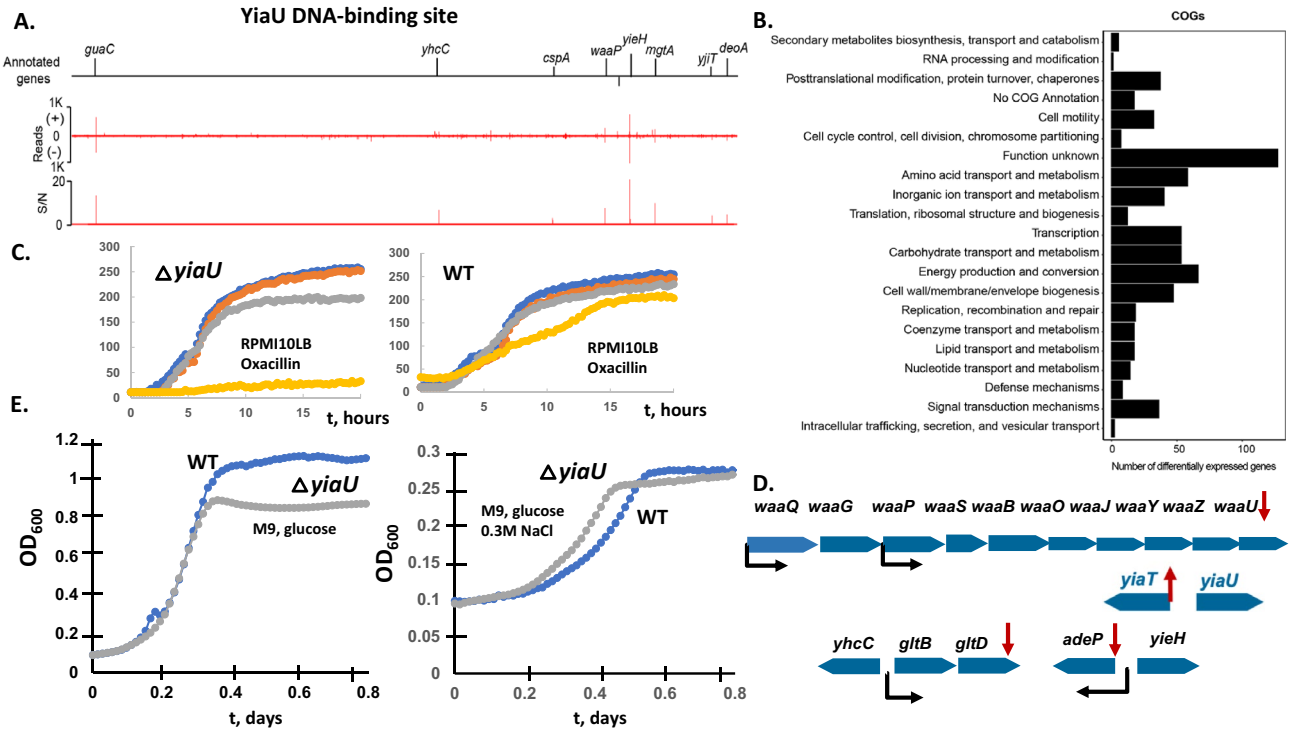

Figure 4. The systems approach for the function of the transcription factor YiaU. (**A**). The genome-wide binding of YiaU across the genome. (**B**). Clusters of Orthologous Groups (COGs) were enriched among the differentially expressed genes between the wild type BW25113 and *yiaU* mutant strains. (**C**). The phenotypes of the *E. coli* BW25113 and *yiaU* knockout strains in the Biolog plate PM12B, measured in RPMI_10LB medium for oxacillin at 4 different doubling concentrations. *yiaU* mutant strain (left panel) had a negative phenotype at the highest oxacillin concentration (8X) PM12B microarray (yellow line). The other antibiotics tested at 4 different concentrations are penicillin G, tetracycline, carbenicillin, penimepicycline, polymyxin B, paromomycin, vancomycin, D,L-serine hydroxamate, sisomycin, sulfamethazine, novobiocin, 2,4-diamino-6,7-dilsopropyl-pteridine, sulfadiazine, benzethonium chloride, tobramycin, 5-fluoroorotate, spectinomycin, L-aspartic-b-hydroxamate, spiramycin, rifampicin, dodecyl trimethyl ammonium bromide. (**D**). The growth measurement of the wild type and *yiaU* mutant strains under different conditions. Left panel: M9 glucose with L-threonine. Right panel: M9 glucose with L-threonine and 0.3 M NaCl. (**E**) Predicted structure and YiaU regulation of *waa* operon, *yiaT* and *adeP* in *E.coli* BW25113. The differentially expressed genes are shown by a red arrow. The *waa* and *gltBD* operon promoters are shown by a black arrow.

Locus tag	Gene name	*P* value	Base mean	$\log^2$ fold change
b3584	*yiaT*	9.32E−07	26	2.8
b1787	*yeaK*	6.57E−10	57	− 4.4
b2036	*glf*	1.01E−09	2378	− 2.6
b2035	*wbbH*	1.63E−07	694	− 1.9
b3714	*adeP*	7.68E−05	357	− 1.32
b3213	*gltD*	0.000187	3158	− 1.27
b2028	*ugd*	0.005533	74	− 1.27
b3624	*waaZ*	5.06E−05	121	− 1.79
b3625	*waaY*	2.08E−05	178	− 1.77
b3622	*waaL*	3.53E−08	352	− 2.81
b3623	*waaU*	3.98E−05	129	− 1.72
b3629	*waaS*	1.84E−08	263	− 2.57
b3631	*waaG*	8.90E−05	140	− 1.61
b3630	*waaP*	0.001681	99	− 1.44
b3627	*waaO*	0.003191	199	− 1.22
b3628	*waaB*	3.78E−05	168	− 1.76

Table 3. Differentially expressed genes revealed by RNA-Seq of a *yiaU* knockout and wild type *E. coli* strains during growth in M9 medium supplemented with 7 mM L-threonine.

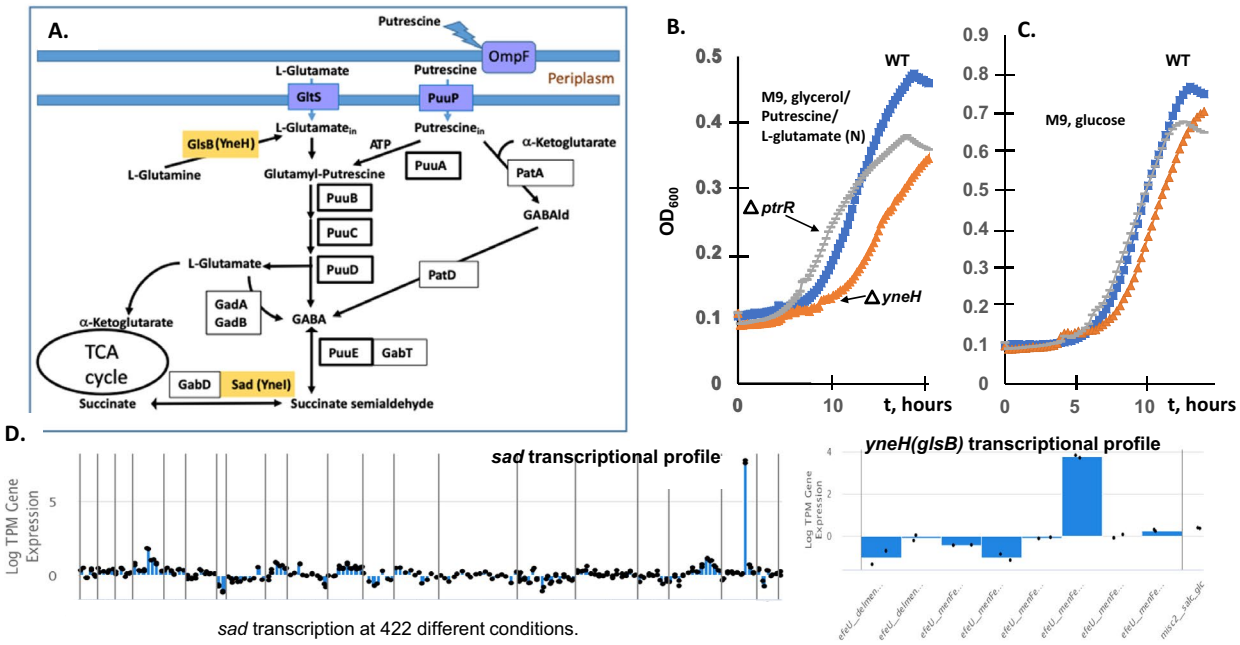

Figure 5. The experimental validation of the transcription factor YneJ (PtrR). (**A**) Overview of two alternative pathways of putrescine utilization in *E. coli*. PtrR-regulated genes are shown in yellow boxes. Transporters are shown in blue. Abbreviations: GABAld – gamma-aminobutyraldehyde, PatA—putrescine aminotransferase, PatD—gamma-aminobutyraldehyde dehydrogenase, YneI (Sad)—succinate-semialdehyde dehydrogenase, YneH (GlsB)—glutaminase. GadAB- two glutamate decarboxylase isoforms, GabD, Sad (YneI)-succinate-semialdehyde dehydrogenase, GlsB (YneH), glutaminase. (**B**) The growth measurement of the *ptrR*, *yneH* mutants compared to the wild type BW25113 strain of *E. coli*. The cell cultures were grown in M9 medium with 10 mM Glu/10 mM Ptr, as nitrogen sources and 0.4% glycerol (v/v) as the primary carbon source or (**C**) M9 glucose medium. (**D**) Transcriptomic data for sad (yneJ) and glsB (yneH) at different growth conditions for *E. coli* MG1655 strain and adapted MG1655 derivatives (iModulonDB, PRECISE2). The activation of the sad and glsB promoter at the ptrR mutant strain was detected.

rate was observed for the *ptrR* mutant. The mRNA level for *sad* was higher in the *ptrR* mutant at these conditions (Table S2, Table 4). The *ptrR* mutation led to the upregulation of 121 genes.

The phenotype microarray using Biolog PM2A was tested under the microaerobic conditions for the *ptrR* mutant phenotype and the *E. coli* WT BW25113 strain. The phenotype using Glu as the energy/nitrogen source was minimal when D-glucosamine or dihydroxyacetone was the carbon source. The *ptrR* mutant defect in growth/respiration with glycine, L-ornithine, or gamma-hydroxybutyrate was observed using M9 medium with L-glutamate as the sole nitrogen source. Phenotypes for the *ptrR* mutant with D-tagatose, oxalomalic acid, gamma-hydroxybutyrate, glycine, and L-alaninamide were observed under the same conditions with Glu as a supplement (Fig. S10). The regulatory effect of PtrR during aerobic growth with putrescine/Glu as the nitrogen source was detected for M9 medium with glycerol as the primary carbon source. The growth of BW25113 (WT) as well as *ptrR, yneH (glsB)* null mutant strains are shown in Fig. 5B-C. The *E. coli* WT strain had a longer lag-phase compared to the *ptrR* mutant. A growth deficiency for a *glsB* mutant was observed under the same conditions, suggesting a functional relationship between GlsB (YneH) and Sad, encoding genes conserved in genome clusters with *ptrR*. The effect of a *yneH* deletion was substantial as cells approached the stationary phase.

PtrR-dependent regulation during growth with Ptr/Glu as nitrogen sources and antibiotic resistance. The *E. coli* WT and *ptrR* mutant were grown aerobically in M9 medium with 20 mM Glu and Ptr as nitrogen sources and 0.2% glycerol. To determine the effect of the *ptrR* deletion mutation, the cells were collected at the log-phase, and total mRNA was purified (see Materials and methods). PuuR and PuuADE, SodB, and two copper related transport systems' mRNA levels increased in the *ptrR* mutant strain (Table 5). SodB (superoxide dismutase) mRNA was increased in *ptrR* mutant and SodB produced H_2O_2. The *copA* and *cus* operons are regulated by the CusSR and HprRS system. H_2O_2 is the effector for HprRS and likely has a transcriptional effect for the *copA* and *cus* system.

Antibiotic resistance induced by *ptrR* mutation in *E. coli* BW25113 was detected. We propose that PtrR negatively controls the FnrS small RNA, which is involved in regulation of MarA mRNA. MarA is a global regulator of *E. coli* genes involved in resistance to antibiotics, oxidative stress, organic solvents, and heavy metals[30]. We tested the *ptrR* mutant and wild type *E. coli* strains for antibiotic resistance using the Biolog plate 11C[31]. Since FnrS is under positive control of the global anaerobic regulator Fnr, *E. coli* was grown under microaerobic conditions. The *ptrR* mutant showed increased resistance to high concentrations of demeclocycline, a tetracycline group antibiotic, which survived after 42 h, in contrast to the wild type. We also detected the increased resistance of

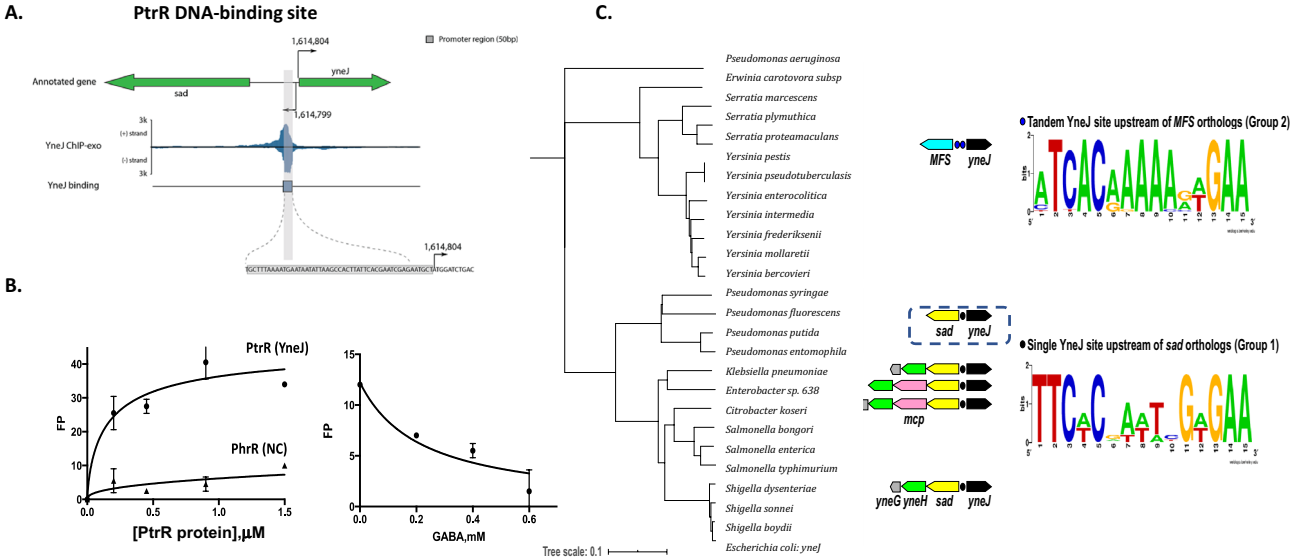

Figure 6. The systems approach for the function of the transcription factor PtrR. (**A**) The zoom-in of the PtrR-binding site at the promoter region of the *ptrR* (*yneJ*) and *sad* genes. Location of sigma-H and sigma-E promoters is from the EcoCyc database. (**B**) Fluorescent polarization assay of PtrR binding to the predicted operator site at *sad* gene. PhrR protein from *Halomonas* was used as a negative control. (**C**) The phylogenetic tree of PtrR orthologous proteins and predicted PtrR-binding motifs in *E. coli* and related genomes of Enterobacteria and *Pseudomonas* spp. The maximum likelihood phylogenetic tree was constructed using RAxML. The distinct genomic context of *ptrR* genes from two major tree branches (groups of PtrR orthologs) is shown by arrows with the following colors: black (*ptrR* regulator), yellow (*sad* for succinate semialdehyde dehydrogenase), green (*yneH* for glutaminase), pink (*mcp* for methyl-accepting chemotaxis protein), and blue (mfs for putative MFS-family transporter), while the predicted PtrR-binding sites are shown by black dots. Sequence logos of predicted DNA binding sites of PtrR from each of the two groups of analyzed species were constructed using WebLogo 2.0.

the *ptrR* mutant to chlorotetracycline (another tetracycline analog) (Fig. S11). However, with other antibiotics tested, no significant difference in growth of the mutant and wild type strains was observed.

Discussion

In this study, we applied a systems approach to characterize the transcriptional responses of seven putative LTFs: YbdO (CitR), YbeF, YbhD, YcaN, YiaU (LpsR), YgfI (DhfA), and YneJ (PtrR) (Table 6). The transcriptional response for the deletion of each LTF had been detected by RNAseq in M9 minimal medium supplemented by 7 mM L-threonine for all yTFs, except PtrR. The transcriptional analysis for the *ptrR* deletion mutant was detected in the M9 medium with Ptr and/or Glu as nitrogen source. For LTFs, conserved adjacent genes had been shown to be differentially expressed. For example, gene clusters encoding *ybdNM* and citrate utilization genes *citCDEF* (citrate lyase) were detected as transcriptionally regulated in the *citR* deletion mutant. The CitR DNA-binding site upstream of *citR* has been predicted and confirmed by ChIP-exo assay, suggesting autoregulation. CitR has been shown to be important for the growth in minimal medium supplemented by citrate at acidic conditions, suggesting citrate lyase regulation microaerobically. Additionally, flagella biosynthesis genes (FlhDC regulon) were differentially expressed in the mutant. The *citR* mutant in *E. coli* BW25113 fitness phenotype had been previously shown for motility in LB medium. Additionally, CitR is important for *E. coli* BW25113 fitness in minimal medium with glycolate as the carbon source (fit.genomics.lbl.gov), and D-glycine as the nitrogen source. The phenotype for formate and L-glutamate utilization *citR* deletion mutant was detected microaerobically (Fig. 7). YbeF is conserved in the gene cluster with citrate lyase encoding genes. The *ybeF* deletion leads to *lrhA* gene downregulation and upregulation of FlhDC regulated genes and, accordingly, the FlhDC iModulon.

The YgfI (DhfA) DNA-binding site upstream of the *dhaKLM* operon for dihydroxyacetone phosphotransferase was detected by ChIP-exo. The transcriptome analysis shows that DhfA was important for regulation of *dhaKLM, pflB, hycBCDEFG, narZYWV* and *adhB*. The *dhfA* mutant growth phenotype using glycerol as a carbon source had been detected (Fig. 3C). The common DNA-binding motif upstream of the *dhak, pflB, adhE, hycB,* and *narZ* genes was found, but future experiments to confirm DhfA binding to the predicted DNA-binding sites are essential (Table 6). The PflB and *hycBCDEFG* encoded hydrogenlyase are involved in pyruvate and Thr utilization as energy source (Fig. 3F). The *dhfA* mutant growth deficiency on glucose as the carbon source in minimal medium at microaerobic conditions had been shown, but supplementation by Thr reduced the growth phenotype (Fig. S7), suggesting DhfA dependent *pflB* and *hycBCDEFG* activation important for anaerobic metabolism. *dhfA* deletion mutant phenotype had been detected in microaerobic conditions for formate utilization in minimal medium after incubation at 37 °C for 24 h (Fig. 7).

Locus tag	Gene name	Gene function	*P* value	Base mean	log^2 fold change
b0123	***cueO***	Blue copper oxidase CueO	1.53E−04	178	1.8
b0484	***copA***	Copper-exporting P-type ATPase	3.00E−05	521	1.8
b0570	***cusS***	Sensor histidine kinase CusS	4.39E−04	112	2.2
b0571	***cusR***	Transcriptional regulatory protein	2.97E−05	181	2.5
b0572	***cusC***	Cation efflux system protein CusC	1.42E−05	949	7.8
b0574	***cusB***	Cation efflux system protein CusB	3.81E−05	208	6.0
b0575	***cusA***	Cation efflux system protein CusA	6.52E−05	281	4.8
b0778	***bioD1***	ATP-dependent dethiobiotin synthetase BioD 1	1.11E−04	126	− 1.4
b1297	***puuA***	Gamma-glutamylputrescine synthetase	4.64E−07	706	1.9
b1298	***puuD***	Gamma-glutamyl-gamma-aminobutyrate hydrolase PuuD	6.21E−05	337	1.9
b1299	***puuR***	HTH-type transcriptional regulator PuuR	9.24E−05	113	1.3
b1302	***puuE***	4-aminobutyrate aminotransferase PuuE	3.16E−04	161	1.6
b1495	***nuoK***	NADH-quinone oxidoreductase subunit K	4.59E−07	73	− 1.9
b1496	***yddA***	Inner membrane ABC transporter ATP-binding protein YddA	1.25E−06	63	− 2.3
b1526	***yneJ***	Transcriptional regulator YneJ	9.41E−06	26	− 6.2
b1596	***ynfM***	Inner membrane transport proteinYnfM	3.06E−04	221	− 1.5
b1656	***sodB***	Superoxide dismutase [Fe]	5.46E−05	323	4.8
b1717	***rpmI***	50S ribosomal protein L35	3.57E−05	55	2.4
b1886	***tar***	Methyl-accepting chemotaxis protein II	4.20E−05	298	1.4
b1889	***motB***	Motility protein B	3.08E−04	61	1.6
b2094	***gatA***	PTS system galactitol-specific EIIA component	6.09E−05	1719	− 1.4
b2106	***rsnA***	Nickel/cobalt efflux system RcnA	5.47E−07	323	− 1.5
b3858	***yihD***	Protein YihD	1.29E−06	42	2.6
b3938	***metJ***	Met repressor	3.84E−04	191	1.4
b4142	***groS***	10 kDa chaperonin	1.40E−04	197	1.9
b4207	***fklB***	FKBP-type peptidyl-prolyl cis–trans isomerase	2.03E−05	539	1.5
b4314	***fimA***	Type-1 fimbrial protein, A chain	3.46E−04	1118	1.4

Table 4. Differentially expressed genes revealed by RNA-Seq of a *ptrR* knockout and wild type *E. coli* strains during growth in M9 medium with L-glutamate and putrescine as nitrogen sources and glycerol as the primary carbon source.

Locus tag	Base mean	log^2 fold change	*p* value	Gene name
b1525	222.2	1.32	0.022	*sad*
b1526	26.1	− 6.69	0.0003	*yneJ*
b1527	14.1	2.78	0.032	*yneK*

Table 5. Differentially expressed *ptrR* adjacent genes revealed by RNA-Seq *ptrR*(*yneJ)* deletion mutant strain and wild type *E. coli* strains during growth in M9 medium with glucose as the primary carbon source and 20 mM L-glutamate as nitrogen source.

yTF gene name	Number of DEGs	DNA binding sites (ChIP-exo/DNA-binding predicted)	Predicted roles	Related references
YbdO	86	2/1 *ybdO*	Citrate utilization related, Flagella biosynthesis	2
YgfI	1402	2/9 *dhaK, pflB, adhE, hycBCDEF, narZ*	Dihydroxyacetone, glycerol or Thr utilization	2
YiaU	674	9/5 *waaPSBOJYZU* (operon), *yjiT, adeP, yiaT, gltD*	Membrane modification/ LPS biosynthesis	2
YcaN	702	6/2, *ycaC, ycaD*	unknown	2
YbhD	238	1/2, *ybhH, ybhI*	L-malate utilization related	2
YbeF	67	3/-	Flagella biosynthesis, putative citrate utilization related	2
YneJ	121	9/2, *sad, fnrS*	Putrescine utilization related	12–15, 17–19, 24–27

Table 6. Summary of the yTFs newly characterized in this study. Abbreviation: LPS-lipopolysaccharide.

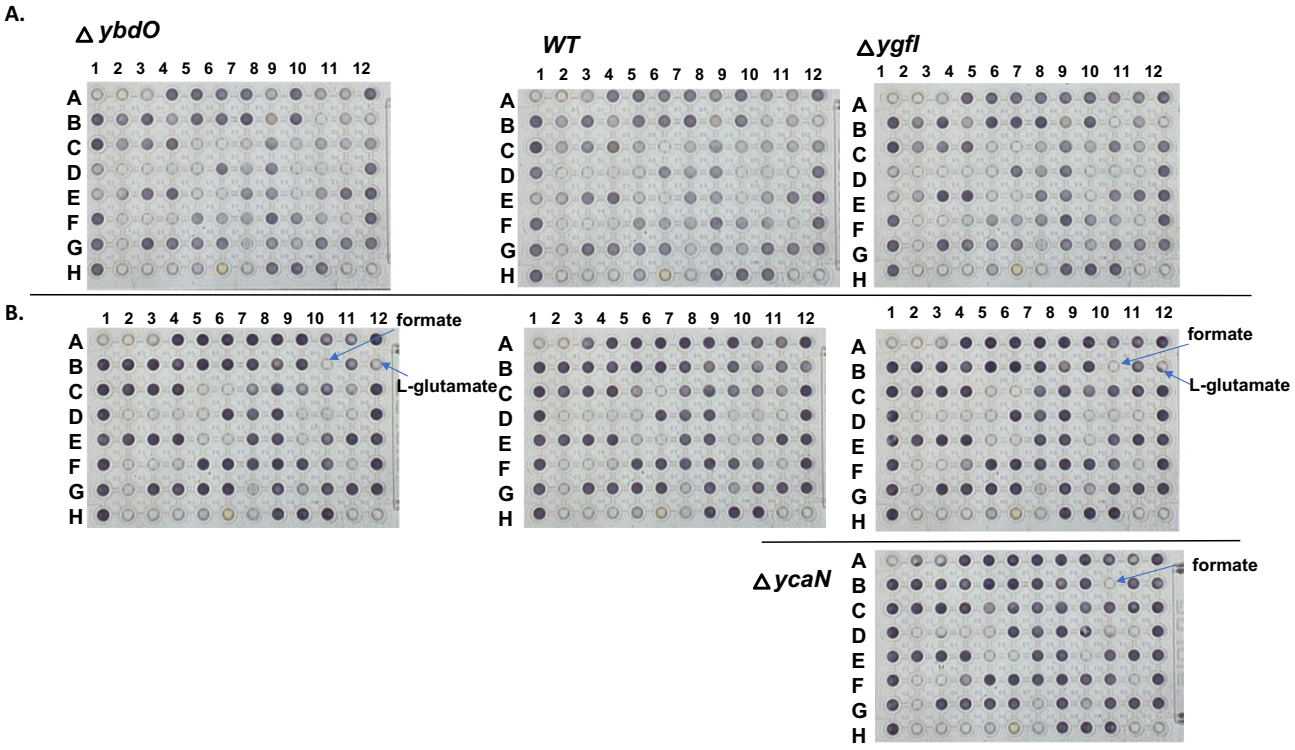

Figure 7. The systems approach for the phenotype detection of the transcription factors *ybdO*, *ycaN* and *ygfI* deletion mutants. The phenotype microarray (Biolog PM1 plate) (**A.** 12 h and **B.** 24 h growth) in M9 medium (95 carbon sources screening for *ybdO* and *ygfI, ycaN* deletion mutant strains compared to *E. coli* BW25113 (wild type, *WT*) strain. The carbon sources in PM1 plate: L-arabinose, *N*-acetyl-glucosamine, succinate, D-galactose, L-aspartate, L-proline, D-alanine, D-trehalose, D-mannose, D-serine, acetate, D-fructose, L-rhamnose, L-lactate, D-xylose, D-sorbitol, L-fucose, D-gluconate, D-glucuronate, D-glycerol-P, D-glucose, D-mellibiose, Lactose, maltose, uridine, L-glutamine, adenosine, L-glutamate, adenosine, L-serine, L-threonine, etc.

We suggest re-name YiaU to LpsR, lipopolysaccharide biosynthesis regulator. ChIP-exo detected multiple LpsR DNA-binding sites (Fig. 4A). The RNAseq transcriptomic analysis and ChIP-Exo (LpsR DNA-binding) additionally detected direct regulation of the *waaPSBOJYZU* operon *gltB* and *adeP* genes, suggesting LpsR relationship to glutamate metabolism. *lpsR* deletion mutant transcriptomic analysis suggested the regulation of *yiaT*, the conserved adjacent gene to *lpsR*, which is divergently transcribed and *yeaK*, encoding deacylase for mischarged aminoacyl-tRNA (L-serine, L-threonine) (Table 3, Fig. 4D). Thr supplement in minimal medium could lead to mischarged tRNA and LpsR is important for the *yeaK* regulation in the presence of Thr. The *lpsR* deletion leads to increased sensitivity to oxacillin in RPMI_10LB medium microaerobically and LpsR was found to have the effect on the biofilm formation previously[3].

The *ybhD* adjacent gene *ybhH* conserved in enterobacterial genomes was upregulated in the *ybhD* deletion mutant. The *ybhH, ybhI,* and *ybhD* genes are conserved adjacent genes and potentially regulated by Nac (EcoCyc). YbhI is the putative tricarboxylate transporter, homologous to 2-oxoglutarate/malate translocator, (id. 35%) (Fig. S8)[32]. We detected the *ybhD* deletion strain growth phenotype in the M9 minimal medium with glycerol (carbon source), supplemented by L-malate (Fig. S9), but not in the absence of L-malate, suggesting that the de-repressed *ybhI* and *ybhH* genes are probably involved in L-malate utilization (Table 6).

We detected that PtrR (YneJ) is the transcriptional regulator for Sad and the small RNA, FnrS. PtrR was predicted to be a repressor of the *fnrS* gene, encoding a small regulatory RNA (Table 6). We demonstrated PtrR binding to the predicted DNA-binding site. According to RNA-Seq data, PtrR is a repressor for *sad* under the nutrient limitation-stress conditions. The *ptrR* gene deletion effect was shown by growth phenotype (aerobically) and phenotype microarray data (micro-aerobically).

PtrR-mediated regulation appears to be important for Ptr utilization as an energy source. A pleiotropic effect of the PtrR-dependent regulation of the *sad* gene under nitrogen/carbon starvation has been investigated and discussed. The known stress/starvation sigma σ^S-controlled *csiD-ygaF-gabDTP* region is related to GABA utilization, while Sad is important for Ptr utilization[25,33–35].

Extracellular Ptr alters the OmpF porin charge and pore size, resulting in partial pore closure and a consequent decrease in outer membrane permeability[15,36]. Our results demonstrated that PtrR is important for the growth of the *E. coli* BW25113 strain with Glu as the sole nitrogen source and glycerol as the carbon source and resistance to the tetracycline group of antibiotics (i.e., demeclocycline and chlortetracycline), but not to chloramphenicol, erythromycin, and other antibiotics. PtrR is potentially important for the regulation of the highly conserved, anaerobically induced small RNA- *fnrS*, which is likely important for regulation under anaerobic

growth conditions[37]. Interestingly, a *ptrR* mutant was shown previously to be resistant to bacteriophage lambda infection[38] and we found that PtrR is potentially related to tetracycline resistance[39]. ChIP-exo and RNAseq results have been analyzed, providing a hypothesis for the physiological functions of YneJ (tentatively re-named PtrR, putrescine related regulator), YgfI (tentatively re-named DhfA, dihydroxyacetone phosphotransferase and formate utilization activator), and YbdO (tentatively re-named CtrR, citrate utilization related regulator).

The identification of the additional DNA-binding sites for YgfI, YcaN, YiaU, YbeF, YneJ by gSELEX (genomic SELEX) method in the presence of Thr in the minimal medium possibly will provide additional information about novel LTFs transcriptional regulatory network[6,40]. LTFs are not always expressed under laboratory growth conditions (for instance, see Ishihama et al. J. Bacteriol. 196, 2718–2727, 2014).

Taken together, the systems analysis of the *E. coli* BW25113 and MG1655 strains transcriptomic data, ChIP-exo DNA-binding data for LTF, and the regulated biochemical pathway reconstruction and fitness/phenotype of the LTF deletion mutants strains produce fruitful hypotheses for the yTF function prediction that is important for TRN reconstruction in *E. coli*.

Methods

RNA sequencing. The wild type BW25113 strain was grown as a control for the isogenic *ptrR* mutant strain. Pre-cultures were obtained by scraping frozen stocks and growing the cells in LB medium. Cells were washed twice with M9 medium and inoculated to an OD_{600} of 0.05. The cells were collected at an OD_{600} of 0.9 (only *ycaN, yiaU, ybhD* mutants and the WT control were collected at OD_{600} of 2.0; late-exponential phase) and were harvested using the Qiagen RNA-protect bacteria reagent according to the manufacturer's specifications. Pelleted cells were stored at − 80 °C, and after cell resuspension and partial lysis, they were ruptured with a beat beater; the total RNA was extracted using a Qiagen RNA purification kit. After total RNA extraction and subsequent ribosomal RNA removal, the quality was assessed using an Aglient Bioanalyser using an RNA 6000 kit. The data processing is described in Supplemental materials.

Protein purification. The PrtR-producing strain was grown overnight, re-inoculated into 50 mL of the fresh medium, and induced with 0.6 mM IPTG after an OD_{600} of 0.6 was reached. The cells were harvested after 4 h and lysed, the cell pellet was resuspended in the lysis buffer. Rapid purification of recombinant proteins on Ni-nitrilotriacetic acid-agarose minicolumns was performed. The protein was refolded on a mini-column, and the buffer was changed to a buffer containing 0.1 M Tris–HCl, 0.1 M NaCl, 10 mM urea.

Fluorescent polarization assay. The purified PtrR protein and 10 nM fluorescently labeled DNA fragment (5′-gggTTCTCGATTCGTGAAggg-3′) were incubated in the assay mixture. The PtrR binding assay mixture (0.1 ml) contained Tris buffer, pH 7.5, 0.1 M NaCl, 10 mM $MgSO_4$, 5 mg/ml sperm DNA and 1uM of the fluorescently labeled predicted PtrR binding DNA fragment as well as 0–0.6 mM GABA. Then the PtrR protein (0–1.5 uM) was added to the assay mixture, and it was incubated for 1 h at 30 °C in the presence or the absence of GABA.

Targeted high-performance liquid chromatography. For organic acid and carbohydrate detection, samples were collected after 4 h for every 30–45 min. The filtered samples were loaded onto a 1260 Infinity series (Agilent Technologies) high-performance liquid chromatography (HPLC) system with an Aminex HPX-87H column (Bio-Rad Laboratories) and a refractive index detector and HPLC was operated using ChemStation software. The HPLC was run with a single mobile phase composed of HPLC grade water buffered with 5 mM sulfuric acid (H_2SO_4). The flow rate was held at 0.5 mL/minute, the sample injection volume was 10 uL, and the column temperature was maintained at 45 °C. The identities of compounds were determined by retention time comparison to standard curves of acetate, ethanol, glucose, lactate, pyruvate, formate and succinate. The peak area integration and resulting chromatograms were generated within ChemStation and compared to that of the standard curves to determine the concentration of each compound in the samples.

Phenotype microarray. The *E. coli* BW25113 wild type and *ybdO, ygfI, ptrR* mutant strains were grown overnight in M9 glucose medium, washed with M9 medium (PM1, PM2) or RPMI with 10% LB (RPMI_LB) for PM12B and inoculated as recommended to the Omnolog plates PM1, PM2, PM11C or PM12B, for the antibiotic resistance measurements at 37 °C. The experiments were repeated two times.

Received: 16 November 2021; Accepted: 14 April 2022
Published online: 04 May 2022

References

1. Gao, Y. *et al.* Unraveling the functions of uncharacterized transcription factors in *Escherichia coli* using ChIP-exo. *Nucleic Acids Res.* **49**, 9696–9710 (2021).
2. Rodionova, I. A. *et al.* Identification of a transcription factor, PunR, that regulates the purine and purine nucleoside transporter punC in *E. coli*. *Commun. Biol.* **4**, 991 (2021).
3. Schellhorn, H. Faculty opinions recommendation of novel regulators of the csgD gene encoding the master regulator of biofilm formation in *Escherichia coli* K-12. In *Faculty Opinions—Post-Publication Peer Review of the Biomedical Literature* (2020). https://doi.org/10.3410/f.738307210.793578431.

4. Yamanaka, Y., Shimada, T., Yamamoto, K. & Ishihama, A. Transcription factor CecR (YbiH) regulates a set of genes affecting the sensitivity of *Escherichia coli* against cefoperazone and chloramphenicol. *Microbiology* **162**, 1253–1264 (2016).
5. Maddocks, S. E. & Oyston, P. C. F. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**, 3609–3623 (2008).
6. Ishihama, A., Shimada, T. & Yamazaki, Y. Transcription profile of *Escherichia coli*: Genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res.* **44**, 2058–2074 (2016).
7. Knapp, G. S. & Hu, J. C. Specificity of the *E. coli* LysR-type transcriptional regulators. *PLoS ONE* **5**, e15189 (2010).
8. Lachnit, T., Bosch, T. C. G. & Deines, P. Exposure of the host-associated microbiome to nutrient-rich conditions may lead to dysbiosis and disease development—an evolutionary perspective. *mBio* **10**, 0035519 (2019).
9. Rychel, K. *et al.* iModulonDB: A knowledgebase of microbial transcriptional regulation derived from machine learning. *Nucleic Acids Res.* **49**, D112–D120 (2021).
10. Website. Price, Morgan (2020): The Fitness Browser: Genome-wide mutant fitness data for diverse bacteria (November 2020 release). figshare. Dataset. https://doi.org/10.6084/m9.figshare.13172087.v1.
11. Rodionova, I. A. *et al.* The uridylyltransferase GlnD and tRNA modification GTPase MnmE allosterically control folylpoly-γ-glutamate synthase FolC. *J. Biol. Chem.* **293**, 15725–15732 (2018).
12. Li, J. *et al.* Polyamines disrupt the KaiABC oscillator by inducing protein denaturation. *Molecules* **24** (2019).
13. Latour, Y. L., Gobert, A. P. & Wilson, K. T. The role of polyamines in the regulation of macrophage polarization and function. *Amino Acids* **52**, 151–160 (2020).
14. Publio, B. C., Moura, T. A., Lima, C. H. M. & Rocha, M. S. Biophysical characterization of the DNA interaction with the biogenic polyamine putrescine: A single molecule study. *Int. J. Biol. Macromol.* **112**, 175–178 (2018).
15. Tkachenko, A. G., Pozhidaeva, O. N. & Shumkov, M. S. Role of polyamines in formation of multiple antibiotic resistance of *Escherichia coli* under stress conditions. *Biochemistry* **71**, 1042–1049 (2006).
16. Nemoto, N. *et al.* Mechanism for regulation of the putrescine utilization pathway by the transcription factor PuuR in *Escherichia coli* K-12. *J. Bacteriol.* **194**, 3437–3447 (2012).
17. Kurihara, S. *et al.* gamma-Glutamylputrescine synthetase in the putrescine utilization pathway of *Escherichia coli* K-12. *J. Biol. Chem.* **283**, 19981–19990 (2008).
18. Kurihara, S., Kato, K., Asada, K., Kumagai, H. & Suzuki, H. A putrescine-inducible pathway comprising PuuE-YneI in which gamma-aminobutyrate is degraded into succinate in *Escherichia coli* K-12. *J. Bacteriol.* **192**, 4582–4591 (2010).
19. Gao, Y. *et al.* Systematic discovery of uncharacterized transcription factors in *Escherichia coli* K-12 MG1655. *Nucleic Acids Res.* **46**, 10682–10696 (2018).
20. Romiyo, V. & Wilson, J. W. Phenotypes, transcriptome, and novel biofilm formation associated with the ydcI gene. *Antonie Van Leeuwenhoek* **113**, 1109–1122 (2020).
21. Lehnen, D. *et al.* LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. *Mol. Microbiol.* **45**, 521–532 (2002).
22. Gonzalez, R., Murarka, A., Dharmadi, Y. & Yazdani, S. S. A new model for the anaerobic fermentation of glycerol in enteric bacteria: Trunk and auxiliary pathways in *Escherichia coli*. *Metab. Eng.* **10**, 234–245 (2008).
23. Lee, J.-H., Lee, K.-L., Yeo, W.-S., Park, S.-J. & Roe, J.-H. SoxRS-mediated lipopolysaccharide modification enhances resistance against multiple drugs in *Escherichia coli*. *J. Bacteriol.* **191**, 4441–4450 (2009).
24. Yeo, S.-J., Jeong, J.-H., Yu, S.-N. & Kim, Y.-G. Crystallization and preliminary X-ray crystallographic analysis of YgjG from *Escherichia coli*. *Acta Crystallogr Sect. F Struct. Biol. Cryst. Commun.* **68**, 1070–1072 (2012).
25. Schneider, B. L. & Reitzer, L. Pathway and enzyme redundancy in putrescine catabolism in *Escherichia coli*. *J. Bacteriol.* **194**, 4080–4088 (2012).
26. Park, S. A., Park, Y. S. & Lee, K. S. Kinetic characterization and molecular modeling of NAD(P)(+)-dependent succinic semialdehyde dehydrogenase from Bacillus subtilis as an ortholog YneI. *J. Microbiol. Biotechnol.* **24**, 954–958 (2014).
27. Fuhrer, T., Chen, L., Sauer, U. & Vitkup, D. Computational prediction and experimental verification of the gene encoding the NAD^+/$NADP^+$-dependent succinate semialdehyde dehydrogenase in *Escherichia coli*. *J. Bacteriol.* **189**, 8073–8078 (2007).
28. Anand, A. *et al.* Pseudogene repair driven by selection pressure applied in experimental evolution. *Nat. Microbiol.* **4**, 386–389 (2019).
29. Romine, M. F. *et al.* Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism. *Proc. Natl. Acad. Sci.* **114**, E1205–E1214 (2017).
30. Barbosa, T. M. & Levy, S. B. Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. *J. Bacteriol.* **182**, 3467–3474 (2000).
31. Berger, P. *et al.* Carriage of Shiga toxin phage profoundly affects *Escherichia coli* gene expression and carbon source utilization. *BMC Genomics* **20**, 504 (2019).
32. Weber, A. *et al.* The 2-oxoglutarate/malate translocator of chloroplast envelope membranes: Molecular cloning of a transporter containing a 12-helix motif and expression of the functional protein in yeast cells. *Biochemistry* **34**, 2621–2627 (1995).
33. Schneider, B. L. *et al.* The *Escherichia coli* gabDTPC operon: Specific gamma-aminobutyrate catabolism and nonspecific induction. *J. Bacteriol.* **184**, 6976–6986 (2002).
34. Iyer, S., Le, D., Park, B. R. & Kim, M. Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in *Escherichia coli*. *Nat. Microbiol.* **3**, 741–748 (2018).
35. Shimizu, K. Metabolic regulation and coordination of the metabolism in bacteria in response to a variety of growth conditions. In *Bioreactor Engineering Research and Industrial Applications I* 1–54 (2015). https://doi.org/10.1007/10_2015_320.
36. Iyer, R. & Delcour, A. H. Complex inhibition of OmpF and OmpC bacterial porins by polyamines. *J. Biol. Chem.* **272**, 18595–18601 (1997).
37. Durand, S. & Storz, G. Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol. Microbiol.* **75**, 1215–1231 (2010).
38. Banzhaf, W. *et al. Evolution in Action: Past, Present and Future: A Festschrift in Honor of Erik D. Goodman* (Springer, New York, 2020).
39. Rodionov, D. A., Gelfand, M. S., Mironov, A. A. & Rakhmaninova, A. B. Comparative approach to analysis of regulation in complete genomes: Multidrug resistance systems in gamma-proteobacteria. *J. Mol. Microbiol. Biotechnol.* **3**, 319–324 (2001).
40. Ishihama, A. Prokaryotic genome regulation: Multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiol. Rev.* **34**, 628–645 (2010).

Author contributions

I.R., Y.G., M.S., H.L. and B.P. wrote the manuscript text and I.R., B.P. and D.R. prepared the figures, N.W., R.S., Z.Z. and Y.H. provide experimental analysis and experiments , H. L., D.R., J. M. provide bioinformatics analysis, Hyun Gyu Lim (H.L).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-11134-7.

Correspondence and requests for materials should be addressed to I.A.R. or B.O.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.