Molecular BioSystems

Dynamic Article Links View Online

Cite this: DOI: 10.1039/c0mb00336k

www.rsc.org/molecularbiosystems

PAPER

# Genome-wide transcriptome and proteome analysis of *Escherichia coli* expressing IrrE, a global regulator of *Deinococcus radiodurans*†

**Zhengfu Zhou,‡[a] Wei Zhang,‡[ab] Ming Chen,[a] Jie Pan,[a] Wei Lu,[ab] Shuzhen Ping,[a] Yongliang Yan,[a] Xiaoguang Hou,[a] Menglong Yuan,[a] Yuhua Zhan[a] and Min Lin*[ab]**

***Received 17th December 2010, Accepted 15th February 2011***
**DOI: 10.1039/c0mb00336k**

Downloaded by University of Arizona on 09 March 2011
Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K

Gram-negative bacterium *Escherichia coli* and the Gram-positive *Deinococcus radiodurans* fundamentally differ in their cell structures and gene regulations. We have previously reported that IrrE, a *Deinococcus* genus-specific global regulator, confers significantly enhanced tolerance to various abiotic stresses. To better understand the global effects of IrrE on the regulatory networks, we carried out combined transcriptome and proteome analysis of *E. coli* expressing the IrrE protein. Our analysis showed that 216 (4.8%) of all *E. coli* genes were induced and 149 (3.3%) genes were repressed, including those for trehalose biosynthesis, nucleotides biosynthesis, carbon source utilization, amino acid utilization, acid resistance, a hydrogenase and an oxidase. Also regulated were the EvgSA two-component system, the GadE, GadX and PurR master regulators, and 10 transcription factors (AppY, GadW, YhiF, AsnC, BetI, CynR, MhpR, PrpR, TdcA and KdgR). These results demonstrated that IrrE acts as global regulator and consequently improves abiotic stress tolerances in the heterologous host *E. coli*. The implication of our findings is discussed in relation to the evolutionary role of horizontal gene transfer in bacterial regulatory networks and environmental adaptation.

## Introduction

Bacteria have existed on Earth for more than three billion years and are able to survive in very diverse and sometimes hostile environments. Two extreme examples are the Gram-negative *Escherichia coli* that tolerates varying osmolarities within the anaerobic gastrointestinal tracts of mammals,[1] and the Gram-positive *Deinococcus radiodurans* that survives extreme conditions of drought, radiation, and scarcity of nutrients.[2] The two species fundamentally differ with respect to their cell structures and gene regulations. *D. radiodurans* was originally identified as a contaminant of irradiated canned meat and has been isolated worldwide from locations both rich and poor in organic nutrients.[3] All species of the genus *Deinococcus* are extremely resistant to agents and conditions that damage DNA, including IR (ionizing radiation), UV (ultraviolet) and hydrogen peroxide.[4] Deinococci are the most radiation-resistant organisms, 200 times more resistant to IR and 20 times more resistant to UV than *E. coli*.[5] *E. coli* is a variable species with a highly dynamic genome.[6,7] It is known that metabolic networks in *E. coli* evolve, in response to changing environments, by the acquisition of heterologous genes and operons by horizontal gene transfer (HGT).[8] Most analyses have, however, ignored the influence of imported global regulators on the evolution of host regulatory networks.

Transcriptome analysis of *D. radiodurans* revealed a group of genes are induced or repressed by IR.[9] IrrE, a novel protein identified in *D. radiodurans*, is a central regulator of some of these genes.[10] The IrrE homologue was detected exclusively in *Deinococcus*, suggesting that it is a genus-specific regulator. Our previous study has shown that the heterologous expression of a single gene, IrrE, in *E. coli* induces significantly enhanced salt tolerance and protected the cells against other abiotic stresses, such as oxidative, osmotic and thermal shocks.[11] Studies of *E. coli* K-12 played a pioneering role in our understanding of DNA replication, transcription, translation, gene regulation, restriction enzymes and horizontal gene transfer.[12] Furthermore, *E. coli* K-12 is an extensively used model to study the molecular mechanisms underlying the adaptive evolution and the global regulation of bacterial metabolic networks. Here, we combine data from genome-wide mRNA and protein profiling and metabolite measurements to investigate the effects of the global regulator IrrE on the changes in gene expression in the heterologous *E. coli* host. This may shed light on the evolution of the regulatory network for environmental adaptation. We were surprised by the profound global changes IrrE made to *E. coli* gene expression.

[a] *Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Key Laboratory of Crop Biotechnology, Ministry of Agriculture, Beijing 100081, China*
[b] *National Centre for Plant Gene Research, Beijing 100101, China. E-mail: linmin57@vip.163.com; Fax: +86 10 82106142; Tel: +86 10 82106145*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c0mb00336k

‡ These authors contributed equally to this work.

View Online

Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K
Downloaded by University of Arizona on 09 March 2011

## Material and methods

### Strains, media and growth conditions

*Escherichia coli* JM109 harboring the IrrE-expressing plasmid pMG1-IrrE and the control plasmid pMG1[11] were used in this study. Cells were grown aerobically in Luria–Bertani (LB) broth or on Luria–Bertani agar at 37 °C. Ampicillin (50 μg $mL^{-1}$) was added to the culture media as appropriate. Overnight cultures were inoculated and sub-cultured to a starting OD600 (optical density at 600 nm) of 0.0001. Cultures were harvested at an OD600 = 0.4 after being maintained in an exponential phase for at least eight generations.

### Transcriptome analysis by DNA microarray hybridization

RNA extraction and processing. For microarray tests, overnight bacterial cultures grown in rich medium (LB) at 37 °C were diluted 100-fold into fresh medium and incubated until OD600 reached 0.4. Aliquots (15 mL) were separated for RNA extraction. Each sample had three biological replicates.

For RNA manipulation, the cells in the broth were centrifuged at 10 000 × *g* and quickly frozen in liquid nitrogen and stored at −70 °C. RNA was extracted from thawed cells using the *TRIzol* (Invitrogen) technique, and then purified with a RNeasy Mini Kit (QIAGEN) and DNase I. To perform this additional DNase digestion, RNA was precipitated and redissolved in 85 μL of nuclease-free water. We then added 10 μL of 10X DNase I buffer and 5 μL of (1 U $\mu L^{-1}$) DNase I (Ambion). The DNase reaction mixture was incubated at 37 °C for 30 min and then chilled on ice. A second RNeasy column purification was then performed. The RNA was finally resuspended in RNase-free water and quantitated with a Bioanalyzer 2100 (Agilent). The Affymetrix GeneChip *E. coli* Genome 2.0 Array was employed in this study. This array chip contains more than 10 000 probe sets that cover all genes in the genomes of four types of *E. coli* strains, K12 MG1655, O157:H7 EDL933, O157:H7 Sakai, and the uropathogen, CFT073. Samples containing 10 μg of RNA were reverse-transcribed and hybridized to the GeneChip. cDNA synthesis, fragmentation, labeling and washing and scanning of the *E. coli* GeneChip arrays were performed according to the instructions of the manufacturer (Affymetrix, Inc.).

Microarray data analysis. Data were scaled based on the total signal intensity, and data analysis was performed using Affymetrix software (MAS 5.0). Changes in gene expression were represented by mean expression ratios (MER). The *P* values were calculated by applying the Wilcoxon's signed rank test. Complete microarray data were registered at the National Center for Biotechnology Information (NCBI), Gene Expression Omnibus (GEO) and hybridization array data repository as GSE17465[NCBI GEO] (http://www.ncbi.nlm.nih.gov/geo).

Quantitative real-time PCR. To verify microarray results, gene expression was examined by quantitative real-time PCR (qPCR) using an ABI 7500 Real-Time PCR System (Applied Biosystems). Primers amplifying 100–200 bp of the target genes were designed using Primer Premier. RNA was reverse-transcribed into first-strand cDNA by a cDNA synthesis kit (New England Biolabs). The SYBR green fluorescence dye (Applied Biosystems) was used for the detection of the PCR amplification. The 16S RNA gene *rrsA* was used as reference for the normalization of the samples.[13]

### Proteome analysis by two-D PAGE

The *E. coli* control strain carrying only vector pMG1 and the transformant strain expressing IrrE, respectively, were grown to the mid-exponential phase (≈0.3 to 0.5, at OD600) with shaking at 37 °C in the LB medium. Cells were harvested by centrifugation (6000 × *g*, 10 min and 4 °C). The pellet was washed twice with buffer (10 mM Tris, 0.4 M sucrose, pH 7.0). The bacterial pellet was then suspended in a lysis buffer (8 M urea, 2 M thiourea, 4% w/v CHAPS, 1% w/v DTT, 0.8% v/v carrier ampholyte, 1 mM PMSF, 350 mM Tris base and 5 mM EDTA) and cell disruption was performed on ice with an ultrasonic homogenizer. The protein extract was aliquoted and frozen at −80 °C after cell debris was pelleted at 14 000 × g, 4 °C for 30 min. The protein concentration was then, measured using the Bradford method. For each experimental condition at least three samples were run in order to assess biological and analytical variations.

Protein separation. Two-dimensional protein gel electrophoresis was performed as described by Pan *et al.*[11] The first dimension (IEF) was carried out using immobilized pH gradient (IPG) strips (Amersham Pharmacia Biotech, Uppsala, Sweden) with a linear pH gradient between 4 and 7 (13 cm) in a Multiphor II electrophoresis unit (Amersham Pharmacia Biotech) at 12 °C. The strip was rehydrated for 12 h at room temperature (8 M urea, 2 M thiourea, 0.5% w/v CHAPS, 0.4% w/v DTT, 0.02% w/v Bromophenol blue and 0.52% v/v carrier ampholyte, pH 4 to 7). The same quantities of sample were applied by cup-loading (2 cm from the anode). The operating conditions were from 0 to 500 V in 2 h, 500 V for 5 h, from 500 to 3500 V in 5 h, 3500 V for 12 h under mineral oil. Thereafter, the IPG strips were incubated for 15 min in equilibration buffer I (6 M urea, 30% v/v glycerol, 2% w/v SDS, 0.02% w/v Bromophenol Blue and 1% w/v DTT in 50 mM Tris-HCl buffer pH 6.8) followed by 15 min in equilibration buffer II (6 M urea, 30% v/v glycerol, 2% w/v SDS,0.02% w/v Bromophenol Blue and 4% w/v iodoacetamide in 50 mM Tris-HCl buffer pH 6.8). The second dimension (SDS-PAGE) was accomplished with vertical 12% gels and sealed with 0.5% agarose. Electrophoresis was carried out in a Bio-Rad Protean II xi unit (Bio-Rad Laboratories, Hercules, CA). Protein spots were visualized by staining with CBB G-250. Gels were scanned with a PowerLook 1000 (UMAX Technologies Inc., Dallas, TX). 2D protein patterns were analyzed using the PDQuest V7.3.0 (Bio-Rad Laboratories) software. Only spots with intensity levels greater than 2.0 or less than 0.5 on stress treatments were considered.

Protein identification by MALDI-TOF MS. For MALDI-TOF MS analysis, protein spots were excised from the gels, digested with trypsin as described,[14] and the peptides generated were analyzed using a 4700 Proteomics Analyzer (Applied Biosystems). Proteins were identified by automated peptide mass fingerprinting using the Global Proteome Server Explorer software (Applied Biosystems).

 This journal is © The Royal Society of Chemistry 2011

View Online

Downloaded by University of Arizona on 09 March 2011
Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K

**Fig. 1 Up- and down-regulation of the metabolic pathway and the transport system genes *E. coli* expressing IrrE.** The colour indicates significant changes ($P > 0.05$) in gene expression or genes and pathways. Red, overexpression; blue, reduced expression; black, change < 1.8-fold. Dashed arrows indicate multiple enzymatic steps. Ovals in the inner membrane, ABC-type transport systems; large squares, phosphotransferase permease systems; small squares, multi-subunit transporters; large rectangles in the outer menbrane, iron transport systems; small rectangles, transport proteins. Abbreviations: G-3-P, glyceraldehyde-3P; UDP-G, UDP glucose; CBDII, cytochrome bd-II terminal oxidase; DK-I, 5-keto-4-deoxy urinate; DK-II, 2, 5-diketo-3-deoxygluconate; KDG, 2-keto-3-deoxygluconate; KDGP, 2-keto-3-deoxygluconate-6P. The pathway map was adapted from the BioCyc database.

### BIOLOG GN2 assays

Colonies from the test strains were inoculated from pregrowth LB agar plates into Biolog inoculating fluid, and the optical density of the inoculating suspension was adjusted to 60% transmittance by use of a turbidimeter. Each suspension of bacterial cells was grown in Biolog defined minimal medium in BIOLOG GN2 plates (150 ml per well), which were then incubated at 37 °C for 24 h. The OD595 was measured with a microtiter plate reader (model Vmax; Molecular Devices Corp., Menlo Park, Calif.) after 6 and 24 h of incubation.

### Measurement of intercellular trehalose content

The cell cultures were sampled (100 ml each) and centrifuged at 4 °C and 4000 × g for 10 min immediately after collection. The bacterial pellet was washed twice with PBS buffer and then resuspended in 10 ml of PBS buffer. In each instance, a beaker with a layer of sample 5 mm thick was irradiated in a microwave oven (750 w, 2450 MHz) for 30–300 s.[15] The samples were centrifuged for 10 min in a bench centrifuge to remove cell debris, and the supernatants were stored at −70 °C. The intracellular trehalose was determined by high-performance liquid chromatography HPLC analysis of the supernatant. Quantitative HPLC analysis of the trehalose was performed as described.[16] At the same time, the cultures were diluted, plated in triplicate on LB agar plates, and incubated for 1 day at 37 °C before scoring for colonies. Data represent the means and standard errors of three replicate experiments.

## Results and discussion

### Expression profiling analysis

IrrE of *D. radiodurans* acts as a global regulator that controls a large set of genes encoding different metabolic and signaling pathways. IrrE dramatically enhanced the salt tolerance in heterologous hosts.[11] Aiming for a comprehensive understanding of the transcriptional response of *E. coli* expressing IrrE, we monitored the mRNA expression profile using whole-genome microarrays. This identified 216 genes whose expression was at least 1.8-fold induced in the IrrE-expressing

View Online

strain (MER $\geqslant$ 1.8 $P$ < 0.05), and 149 genes were repressed by IrrE (MER $\leqslant$ −1.8 $P$ < 0.05) (ESI: Table S1 and S2†). These regulated genes were classified using the COG database. Up-regulated genes were involved in carbohydrate metabolism (20 of 216 induced genes, 9%), regulatory functions and signal transduction (8%), and amino acid biosynthesis (7%), nucleotide transport and metabolism (7%), the cell envelope (4%) and energy production and conversion (4%). Down-regulated genes were involved in amino acid biosynthesis (23 of 149 repressed genes, 15%), energy production and conversion (14%), carbohydrate metabolism (10%), regulatory functions and signal transduction (8%), and nucleotide transport and metabolism (4%). Note that IrrE significantly repressed 90 contiguous *E. coli* genes (b0260 to b0349). About 20% of the up-regulated and 8% of the down-regulated genes encoded proteins of hypothetical or unknown functions. No functional COG categories existed for 46 up-regulated genes and 29 down-regulated genes. IrrE thus controls multiple metabolic functions including trehalose biosynthesis, nucleotide biosynthesis, carbon source utilization, amino acid utilization, acid resistance, oxidoreductases, and transporters in *E. coli* (Fig. 1).

### The respiratory hydrogenase and oxidase genes

The *hyaABCDEF* operon encodes hydrogenase-1 and is induced by IrrE (MER > 2.5). HyaA and HyaB are the small and large subunits, respectively.[17] HyaCDE are also required for the activity hydrogenase-1, and HyaF is required to achieve the maximal wild-type level of hydrogenase-1 activity.[18] The *app* genes are also induced by IrrE. They specify Cytochrome bd-II oxidase (encoded by the *appCB* operon, MER = 3.2; 3.4), and phytase/acid phosphatase (encoded by *appA*, MER = 1.8).

Hydrogenase-1 is a Fe–S cluster-containing anaerobic respiratory enzyme that mediates hydrogen uptake in the presence of high-potential acceptors such as ferricyanide and phenazine methosulfate.[19] The *sufABCDSE* operon, which encodes the minor Fe–S cluster assembly system, was up-regulated by IrrE. In *E. coli*, SufA accepts 2Fe–2S clusters from SufBCD and transfers the 2Fe–2S cluster to apoferredoxin, apoaconitase and other proteins.[20] SufS alone exhibits cysteine desulfurase activity. SufE, a homodimeric protein without a cofactor, enhances the cysteine desulfurase activity of SufS. SufBCD is an ATPase that further increases the cysteine desulfurase activity of the SufS SufE complex.[21] The *suf* operon protects cells against oxidative damage[22] and was found to be induced by simultaneous high osmolarity and heat stress.[23]

### Osmoprotection including trehalose biosynthesis

Trehalose is a nonreducing, particularly stable disaccharide synthesized from two glucose moieties. Like other compatible osmolites such as glycine betaine, trehalose efficiently stabilizes proteins, dehydrated enzymes,[24] and lipid membranes, as well as protecting biological structures from damage during abiotic stress in diverse organisms ranging from bacteria and fungi to invertebrates and plants.[24–27] The IrrE-induced *otsAB* operon, encodes the two enzymes OtsA (trehalose-6-phosphate synthase) and OtsB (trehalose-6-phosphate phosphatase) which convert UDP-glucose to trehalose and are unique to trehalose biosynthesis. OtsA and OtsB are also induced by osmotic shock, extreme heat, extreme cold, desiccation, and entry into the stationary phase.[24,27] OtsA produces trehalose-6-phosphate which induces the *treBC* operon (Trehalose transport and degradation), but OtsB very effectively hydrolyzes trehalose-6-phosphate. Therefore induction of *otsAB* results in repression of the *treBC* operon.[28] Induction of the trehalose biosynthesis operon was thus expected to result in the accumulation of trehalose. This was confirmed by HPLC analysis. The IrrE-expressing strain accumulated 2.1-fold more trehalose than the control (0.24 ± 0.01 mg per $10^{11}$ cells instead of

**Table 1** Up-regulation of osmotically active genes and acid resistance genes in the IrrE-expressing *E. coli*

| Locus | Gene | Induction ratio[a] | Function[b] |
|---|---|---|---|
| **Osmotically inducible protein genes** | | | |
| b1283 | *osmB* | 5.0 | Osmotically inducible lipoprotein |
| b1482 | *osmC* | 3.3 | Predicted redox protein, regulator of disulfide bond formation |
| b1739 | *osmE* | 2.9 | DNA-binding transcriptional activator OsmE |
| b2131 | *osmF* | 2.3 | Periplasmic glycine betaine/choline-binding protein of an ABC-type transport system |
| b4376 | *osmY* | 2.6 | Predicted periplasmic or secreted lipoprotein |
| b1896 | *otsA* | 3.4 | Trehalose-6-phosphate synthase |
| b1897 | *otsB* | 4.7 | Trehalose-6-phosphatase |
| b1732 | *katE* | 2.5 | Catalase |
| **Acid resistance genes** | | | |
| b3506 | *slp* | 6.6 | Starvation-inducible outer membrane lipoprotein |
| b3507 | *yhiF* | 2.4 | DNA-binding HTH domain-containing proteins |
| b3508 | *yhiD* | 2.2 | Uncharacterized membrane protein |
| b3512 | *gadE* | 5.7 | DNA-binding HTH domain-containing proteins |
| b3515 | *gadW* | 2.5 | Putative ARAC-type regulatory protein |
| b3516 | *gadX* | 4.9 | DNA-binding transcriptional regulator GadX |
| b3517 | *gadA* | 2.9 | Glutamate decarboxylase and related PLP-dependent proteins |
| b1493 | *gadB* | 3.8 | Glutamate decarboxylase and related PLP-dependent proteins |
| b1492 | *gadc* | 3.9 | Amino acid transporters |
| b4452 | *gadY* | 2.9 | Small RNA |

[a] The transcriptional ratios obtained from DNA microarray experiment (see Material and methods) are indicated. [b] Information regarding gene function is from the EcoCyc database.

Downloaded by University of Arizona on 09 March 2011
Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K

 This journal is © The Royal Society of Chemistry 2011

View Online

Downloaded by University of Arizona on 09 March 2011
Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K

0.11 ± 0.01 mg per $10^{11}$ cells). Together, these data suggested that the accumulation of trehalose could be attributed to the IrrE expression in growing cells.

IrrE induces another subset of *E. coli osm* (osmoprotection) genes that are dispersed through the genome. The genes induced were *osmB*,[29] *osmC*,[30] *osmE*,[31] *osmF*[32] and *osmY*.[33] The data from the microarray analysis are listed in Table 1. Surprisingly, *betTIBA* (osmoprotectant glycine betaine biosynthesis) were down-regulated. This needs to be further investigated. The global regulator IrrE also induces multiple stress genes, including *katE* (catalase), *pspA* (phage shock protein A), in *E. coli* growing under non stressed conditions. IrrE thus enables the cells to overcome sudden stresses much more quickly than would be the case without IrrE, as was described.[11]

### Acid resistance

Many genes involved in acid resistance, including *slp*, *yhiF*, *yhiD*, *gadE*, *gadW*, *gadX*, *gadA*, *gadBC* and *gadY* (small RNA) were induced by IrrE (Table 1). The *E. coli* genes *slp* to *gadA* form the genomic acid fitness island (AFI) containing several transcription units.[34] Four AFI genes (*gadA*, *gadX*, *gadW*, and *gadE*) and *gadB* are involved in an acid resistance system that consumes intracellular protons for the decarboxylation of glutamic acid.[34] Slp, an outer membrane lipoprotein, and YhiF, a negative regulator for the C4-dicarboxylate transport, have more specific roles in reducing metabolite stress.[34] YhiD is a putative membrane protein, and *gadY* specifies a small RNA that induces the expression of *gadX*.[35]

### Carbohydrate utilization

The carbon source utilization profile of *E. coli* expressing IrrE and the control strain was determined using Biolog assays. The control strain utilized 45 different carbon sources. Among those, seven carbon sources were utilized more efficiently (biggest increase was for the compatible solute sorbitol), eight less efficiently and 14 remained almost unused in the *E. coli* expressing IrrE (Table S5†). The change in sorbitol utilization was particularly interesting because the IrrE-expressing strain had an increased resistance to osmotic shock caused by sorbitol.[11] The microarray data show that the *srlAEBD* operon, which encodes the SrlABE sorbitol permease, was upregulated by IrrE. SrlABE, belongs to the functional superfamily of the phosphoenolpyruvate (PEP)-dependent, sugar transporting phosphotransferases (PTS). SrlABE, which is located in the inner membrane, phosphorylates sorbitol from the periplasm and releases the phosphate ester into the cytoplasm. The glucitol-6-P dehydrogenase SrlD then oxidizes sorbitol-6-phosphate to the glycolytic intermediate fructose-6-P. SrlD can also catalyze the reverse reaction and synthesize sorbitol.[36] It has been reported that expression of the *E. coli srlD* gene increased the sorbitol content and improved salt tolerance in *Pinus taeda*.[37]

Maltose comes from the hydrolysis of starch. The six genes for maltose utilization, *lamB* and *malEFGKM*, were down-regulated by IrrE. The Biolog data showed that the utilization of maltose was decreased in the IrrE-expressing strain (Table S5†). Maltose is normally the preferred carbon source for cells growing in LB.[38] The maltose transport system consists of LamB and MalEFGK. LamB is an outer membrane porin for maltodextrins and other carbohydrates. Following transport into the periplasm, maltose binds to the periplasmic maltose-binding protein (MBP) MalE. MalE bound to maltose undergoes a conformational change and interacts with the two inner membrane proteins, MalF and MalG. These proteins bind to MalK which hydrolyses ATP to ADP in the cytoplasm.[39] The function of *malM*, the last gene in the operon, is unknown.[40] Maltose can also diffuse through the trehalose transporter TreB.[41]

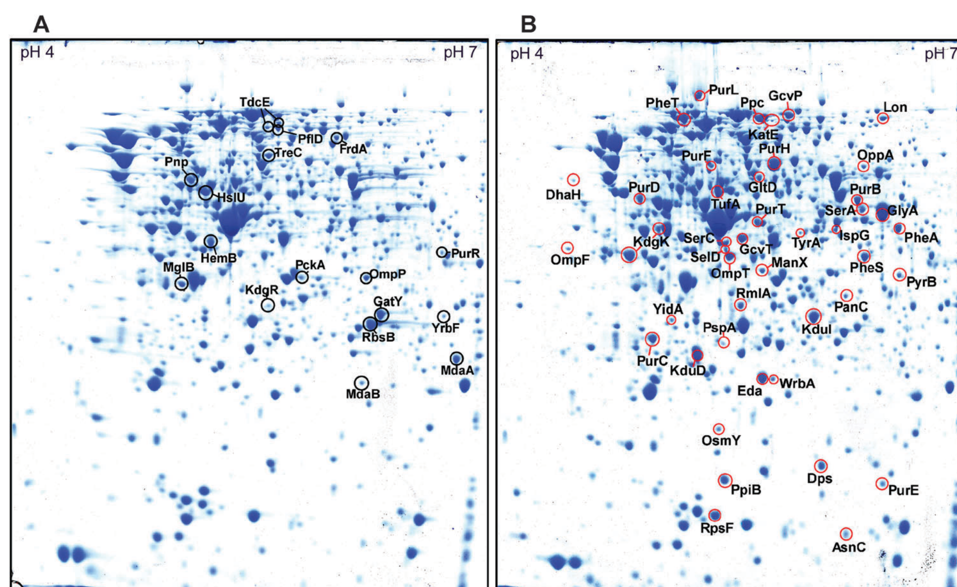


**Fig. 2** Changes in protein expression caused by the *D. radiodurans* IrrE global regulator in *E. coli*. Coomassie-blue stained two-dimensional protein gels (size fractionation and pH gradient) of total protein. A, control strain (vector only); B, strain expressing IrrE. Red circles, induced by IrrE; black circles, repressed by IrrE. The circled spot changes were observed in three repeat experiments. The proteins were identified by MALDITOF MS analysis.

View Online

Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K
Downloaded by University of Arizona on 09 March 2011

The repression of the *treBC* operon and induction of the *otsAB* operon in the IrrE-expressing strain resulted in a decrease in the concentration of trehalose-6-phosphate, the inducer of maltose uptake, by converting it to trehalose. IrrE thus caused the down-regulation of the maltose uptake system (Fig. 1).

### Amino acid utilization

12 amino acids, 7 amino acid analogs and 3 dipeptides were included as substrates in GN2 MicroPlates. The utilization of the amino acids L-alanine, L-asparagine, L-aspartic acid, D-serine, L-serine, L-threonine; and the dipeptides L-alanyl-glycine, glycyl-L-aspartic acid was severely decreased in the IrrE-expressing strain (Table S5†). IrrE downregulated the *tdcABCDEFG* operon, which is involved in the transport and degradation of L-serine and L-threonine (Table S2†). TdcC is a member of the STP family of amino acid transporters, likely to function as a proton-driven serine/threonine symporter.[42] TdcB, catabolic threonine dehydratase, catalyzes the conversion of L-threonine to 2-oxobutanoate and ammonia and also catalyzes the conversion of L-serine to pyruvate and ammonia.[43] TdcG, L-serine deaminase III, is one of four enzymes converting serine to pyruvate and ammonia. TdcE has both pyruvate formate-lyase and 2-ketobutyrate formate-lyase activity, whereas TdcD is a propionate/acetate kinase.[44] The function of TdcF is unknown. The enzymes, encoded by the *tdcBCDE* participate in the degradation of threonine to propionate. Threonine degradation generates ATP, and threonine can be used as the sole source of carbon and energy. Note that *prpBCDE*, involved in the oxidation of propionate to pyruvate, was also down-regulated by IrrE (MER $< -13$). The microarray data also showed that *dctA*, which encodes a L-aspartate transporter,[45] was repressed by IrrE. Generally, all the above genes involved in amino acid metabolism were repressed by IrrE.

### Proteome profiles

To further investigate the role of IrrE in *E. coli*, protein extracts were fractionated by 2D gel electrophoresis and the protein spots were identified by MALDI-TOF MS. Nearly 700 protein spots were resolved on all gels (Fig. 2). Compared to the control, the intensity of 60 of the protein spots was increased, and 44 of them were successfully identified by strong hits to the *E. coli* genome database (Fig. 2B, Table S3†). The intensity of 33 protein spots was reduced in the IrrE strain, and 7 of them disappeared completely (Fig. 2A). The 17 identified proteins are shown in Fig. 2 and listed in Table S4.† There was a good correlation between the mRNA transcriptome and the proteome data. In particular, the expression of almost all the genes of the inosine-5′-phosphate biosynthesis pathway was enhanced by IrrE at both the transcriptional and translational levels (Fig. 2, Table S3†). Inosine-5′-phosphate is a key intermediate in the biosynthesis of purine nucleotides. In *E. coli*, inosine-5′-phosphate is synthesized from 5-phospho-α-D-ribose 1-diphosphate (PRPP) in 11 steps. In bacteria, the majority of *de novo* purine biosynthetic genes are unlinked, but they are coordinately regulated by the PurR repressor. Expression of *purR* was repressed by IrrE. Proteomic analysis also showed a 5.6-fold induction of OsmY (response to hyperosmotic stress, Table S3†), and TdcE (serine degradation) was 7.1-fold decreased (Table S4†), consistent with the above microarray data. For a few genes, the transcriptome and proteome data were contradictory, probably because the translation of these genes was regulated. For example, OppA, the periplasmic oligopeptides binding component,

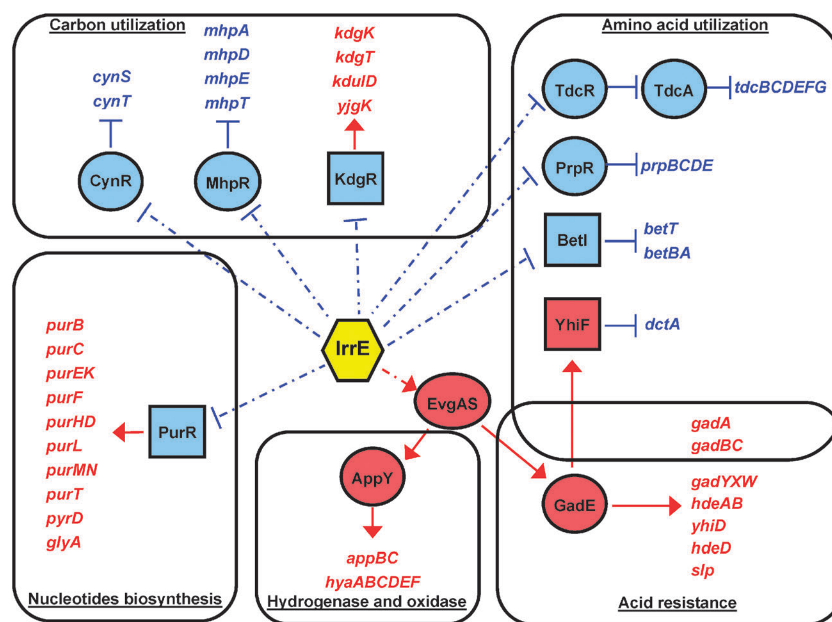


**Fig. 3** IrrE regulation in *E. coli*. The transcriptional regulatory network was constructed from mRNA and protein expression data. Red arrows, induction; blue lines with a cross bar, repression. Solid lines, proven regulatory interactions; broken lines, predicted interactions. Oval shapes, *E. coli* inducers; squares, repressors. Italic, regulated genes. Gene names or backgrounds: red, induced; blue, repressed. Yellow hexagon, dual function induction and repression.


This journal is © The Royal Society of Chemistry 2011

View Online

Downloaded by University of Arizona on 09 March 2011
Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K

was increased in the proteome, but decreased 5.9-fold in the transcriptome.

### IrrE is a global regulator that initiates a complex transcriptional cascade

In their natural environment, bacteria are often challenged by changing nutrient availability, and by exposure to various physical stresses, including osmotic, oxidative and temperature stress. Successful adaptation to the environment requires the sensing and correct response to multiple stimuli through the simultaneous expression of many genes. Transcriptional regulation is the primary adaptation mechanism which must be studied in its entirety as a complex network of molecular interactions. It is now understood that localized regulatory networks in prokaryotic genomes evolve rapidly by mutation or horizontal gene transfer, and they can differ significantly between closely related strains.[46] The EcoCyc database describes 21 two-component systems, 35 master regulators and 183 transcription factors which are often integrated in complex regulatory networks. We have used transcriptome, proteome and metabolome data to construct a regulatory map for *E. coli* expressing IrrE from *D. radiodurans*.

IrrE altered the expression of many transcriptional regulators including the EvgSA two-component system, the master regulators GadE, GadX and PurR and gene/operon-specific transcription factors AppY, GadW, YhiF, AsnC, BetI, CynR, MhpR, PrpR, TdcA, KdgR. Among the listed regulators, 7 (*appY*, *evgA*, *gadE*, *gadW*, *gadX*, *yhiF* and *asnC*) were induced, and another 7 (*betI*, *cynR*, *mhpR*, *prpR*, *tdcA*, *kdgR* and *urr*) were repressed (Table S1 and Table S2†). We suggest that IrrE may control most genes indirectly through these regulators (Fig. 3). *E.g.* IrrE induces the regulator AppY which then induces *appCB* and *hyaABCDEF*.[47] Another example, *kdgR* encodes a DNA-binding transcription factor that represses transcription of the operons involved in transport and catabolism of 2-keto-3-deoxy gluconate (KDG).[48] Among the IrrE induced genes identified here, 5 (*kdgK*, *kdgT*, *kduID* and *yjgK*) have been previously shown to be controlled by KdgR.[49]

In another example, EvgS (MER = 1.8) and EvgA (MER = 2.54) are the sensor kinase and response regulator for the EvgSA two-component system (TCS). The EvgSA TCS is believed to initiate a complex transcriptional cascade that provides acid and drug resistance in *E. coli*.[50] IrrE probably regulates the EvgSA TCS which then induces multiple genes including *gadE*, *gadX* and *appY* (Fig. 3).[51] In addition, expression of genes involved in acid resistance was increased during the exponential phase by IrrE. Among these acid resistance genes, increased expression of *yhiF* represses the transcription of *dctA* (transportor of L-aspartate transport). Consistently, utilization of L-aspartate was severely decreased in the IrrE-expressing strain. However, only a part of the EvgSA TCS regulatory circuit was induced by IrrE, *e.g.* the expression of the AraC-like transcriptional regulator YdeO,[51,52] which is controlled by the inducer EvgA, did not change significantly (MER = 1.2). Maybe other IrrE induced repressors also act on this and other genes.

There is a strong selection for regulated and coordinated gene expression in ever-changing, highly competitive environments.[53] For example, intracellular pathogens and endosymbionts survive in a less variable environment and have fewer regulators per gene, compared with many free living bacteria, which have more regulators in their genomes.[54] The most detailed information is available about the evolution of regulatory networks of *E. coli*.[55] In general, *E. coli* populations are ecologically and metabolically versatile and can adapt to growth under a wide range of conditions. Many *E. coli* genes, including most pathway-specific regulators, were acquired by horizontal gene transfer from other bacteria instead of being inherited by continuous vertical descent from an ancient ancestor.[56] In contrast, HGT of global regulators was rare. Indeed, global regulators are responsible for about two-thirds of the known regulation of *E. coli* and probably evolved vertically within the γ-Proteobacteria.[56] *E. coli* is not only the extremely well studied model organism, but is also an extensively used microbe for industrial applications, especially for the production of small molecules of interest. *E. coli* and *D. radiodurans* belong to two phyla, Proteobacteria and Deinococcus-Thermus, which have remarkably different phenotypes. In a previous study, we found that the IrrE of *D. radiodurans* induces a significantly enhanced salt tolerance and protected the *E. coli* cells against other abiotic stresses. As a case study of horizontal gene transfer of a single gene under laboratory condition, our work opens the door for the use of *E. coli* K-12 as a model organism to further elucidate the effect of IrrE on the global transcriptional response in response to various abiotic stresses, and will provide interesting insights into the evolutionary mechanisms of bacterial regulatory networks and environmental adaptation.

## Conclusion

*D. radiodurans* is a Gram-positive bacterium belonging to the phylum Deinococcus-Thermus, and is phylogenetically very distant from *E. coli*. IrrE is a well-conserved Deinococcus-specific global regulator. To our knowledge, this is the first report of a Genus-specific global regulator that causes regulatory changes in unstressed *E. coli*. We observed that in one two-component system, three master regulators and 10 transcription factors were up- or down-regulated, suggesting that IrrE functioned as a global regulator initiating a complex transcriptional cascade in the heterologous *E. coli* host. Furthermore, there are a large number of host genes with a $\geq$ 1.8-fold change in expression, including those responsible for trehalose biosynthesis, nucleotides biosynthesis, carbon source utilization, amino acid utilization, and acid resistance. This current analysis is only a beginning, and further research is needed to investigate the genome-wide transcriptional response of the heterologous host *E. coli* expressing IrrE at various stresses. These findings have the potential to improve many industrial applications of *E. coli* and other bacteria.

## Acknowledgements

We would like to thank Dr Tobias Kieser for many helpful discussions and Dr Darrell Leong for critical reading of the manuscript. This work was supported by grants from the

View Online

National Natural Science Foundation of China (No. 30925002, 30970069 and 30970093), the National Basic Research (973) Program of China (No. 2007CB707805 and 2010CB126504) and the National Research and Development Project of Transgenic Crops of China (2008ZX08003-003B and 2009ZX08012-008B).

## References


1 F. R. Blattner, G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao, *Science*, 1997, **277**, 1453–1462.
2 O. White, J. A. Eisen, J. F. Heidelberg, E. K. Hickey, J. D. Peterson, R. J. Dodson, D. H. Haft, M. L. Gwinn, W. C. Nelson, D. L. Richardson, K. S. Moffat, H. Qin, L. Jiang, W. Pamphile, M. Crosby, M. Shen, J. J. Vamathevan, P. Lam, L. McDonald, T. Utterback, C. Zalewski, K. S. Makarova, L. Aravind, M. J. Daly, K. W. Minton, R. D. Fleischmann, K. A. Ketchum, K. E. Nelson, S. Salzberg, H. O. Smith, J. C. Venter and C. M. Fraser, *Science*, 1999, **286**, 1571–1577.
3 M. Blasius, S. Sommer and U. Hubscher, *Crit. Rev. Biochem. Mol. Biol.*, 2008, **43**, 221–238.
4 M. M. Cox and J. R. Battista, *Nat. Rev. Microbiol.*, 2005, **3**, 882–892.
5 M. J. Daly, *Nat. Rev. Microbiol.*, 2009, **7**, 237–245.
6 J. G. Lawrence and H. Ochman, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 9413–9417.
7 U. Dobrindt, F. Agerer, K. Michaelis, A. Janka, C. Buchrieser, M. Samuelson, C. Svanborg, G. Gottschalk, H. Karch and J. Hacker, *J. Bacteriol.*, 2003, **185**, 1831–1840.
8 C. Pal, B. Papp and M. J. Lercher, *Nat. Genet.*, 2005, **37**, 1372–1375.
9 M. Tanaka, A. M. Earl, H. A. Howell, M. J. Park, J. A. Eisen, S. N. Peterson and J. R. Battista, *Genetics*, 2004, **168**, 21–33.
10 A. M. Earl, M. M. Mohundro, I. S. Mian and J. R. Battista, *J. Bacteriol.*, 2002, **184**, 6216–6224.
11 J. Pan, J. Wang, Z. Zhou, Y. Yan, W. Zhang, W. Lu, S. Ping, Q. Dai, M. Yuan, B. Feng, X. Hou, Y. Zhang, R. Ma, T. Liu, L. Feng, L. Wang, M. Chen and M. Lin, *PLoS One*, 2009, **4**, e4422.
12 J. L. Hobman, C. W. Penn and M. J. Pallen, *Mol. Microbiol.*, 2007, **64**, 881–885.
13 Y. Yan, J. Yang, Y. Dou, M. Chen, S. Ping, J. Peng, W. Lu, W. Zhang, Z. Yao, H. Li, W. Liu, S. He, L. Geng, X. Zhang, F. Yang, H. Yu, Y. Zhan, D. Li, Z. Lin, Y. Wang, C. Elmerich, M. Lin and Q. Jin, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 7564–7569.
14 L. Feng, W. Wang, J. Cheng, Y. Ren, G. Zhao, C. Gao, Y. Tang, X. Liu, W. Han, X. Peng, R. Liu and L. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 5602–5607.
15 L. Chuanbin, X. Jian, B. Fengwu and S. Zhiguo, *Biotechnol. Tech.*, 1998, **12**, 941–943.
16 H. Yu, Y. S. Ding, S. F. Mou, P. Jandik and J. Cheng, *J. Chromatogr., A*, 2002, **966**, 89–97.
17 N. K. Menon, J. Robbins, H. D. Peck, Jr., C. Y. Chatelus, E. S. Choi and A. E. Przybyla, *J. Bacteriol.*, 1990, **172**, 1969–1977.
18 N. K. Menon, J. Robbins, J. C. Wendt, K. T. Shanmugam and A. E. Przybyla, *J. Bacteriol.*, 1991, **173**, 4851–4861.
19 T. V. Laurinavichene and A. A. Tsygankov, *FEMS Microbiol. Lett.*, 2001, **202**, 121–124.
20 H. K. Chahal, Y. Dai, A. Saini, C. Ayala-Castro and F. W. Outten, *Biochemistry*, 2009, **48**, 10644–10653.
21 F. W. Outten, M. J. Wood, F. M. Munoz and G. Storz, *J. Biol. Chem.*, 2003, **278**, 45713–45719.
22 J. H. Lee, W. S. Yeo and J. H. Roe, *Mol. Microbiol.*, 2004, **51**, 1745–1755.
23 T. S. Gunasekera, L. N. Csonka and O. Paliy, *J. Bacteriol.*, 2008, **190**, 3712–3720.
24 O. Kandror, A. DeLeon and A. L. Goldberg, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9727–9732.
25 M. Iordachescu and R. Imai, *J. Integr. Plant Biol.*, 2008, **50**, 1223–1229.
26 C. De Virgilio, T. Hottiger, J. Dominguez, T. Boller and A. Wiemken, *Eur. J. Biochem.*, 1994, **219**, 179–186.
27 E. N. Miller and L. O. Ingram, *Biotechnol. Lett.*, 2008, **30**, 503–508.
28 B. Kempf and E. Bremer, *Arch. Microbiol.*, 1998, **170**, 319–330.
29 J. U. Jung, C. Gutierrez, F. Martin, M. Ardourel and M. Villarejo, *J. Biol. Chem.*, 1990, **265**, 10574–10581.
30 C. Gutierrez and J. C. Devedjian, *J. Mol. Biol.*, 1991, **220**, 959–973.
31 C. Gutierrez, S. Gordia and S. Bonnassie, *Mol. Microbiol.*, 1995, **16**, 553–563.
32 C. Checroun and C. Gutierrez, *FEMS Microbiol. Lett.*, 2004, **236**, 221–226.
33 J. T. Oh, T. K. Van Dyk, Y. Cajal, P. S. Dhurjati, M. Sasser and M. K. Jain, *Biochem. Biophys. Res. Commun.*, 1998, **246**, 619–623.
34 A. K. Mates, A. K. Sayed and J. W. Foster, *J. Bacteriol.*, 2007, **189**, 2759–2768.
35 J. A. Opdyke, J. G. Kang and G. Storz, *J. Bacteriol.*, 2004, **186**, 6698–6705.
36 P. W. Postma, J. W. Lengeler and G. R. Jacobson, *Microbiol. Rev.*, 1993, **57**, 543–594.
37 W. Tang, X. Peng and R. J. Newton, *Plant Physiol. Biochem.*, 2005, **43**, 139–146.
38 M. V. Baev, D. Baev, A. J. Radek and J. W. Campbell, *Appl. Microbiol. Biotechnol.*, 2006, **71**, 310–316.
39 W. Boos and H. Shuman, *Microbiol. Mol. Biol. Rev.*, 1998, **62**, 204–229.
40 E. Schneider, E. Francoz and E. Dassa, *Biochim. Biophys. Acta.*, 1992, **1129**, 223–227.
41 K. Decker, F. Gerhardt and W. Boos, *Mol. Microbiol.*, 1999, **32**, 777–788.
42 V. N. Sumantran, H. P. Schweizer and P. Datta, *J. Bacteriol.*, 1990, **172**, 4288–4294.
43 L. S. Park and P. Datta, *J. Bacteriol.*, 1979, **138**, 1026–1028.
44 C. Hesslinger, S. A. Fairhurst and G. Sawers, *Mol. Microbiol.*, 1998, **27**, 477–492.
45 S. J. Gutowski and H. Rosenberg, *Biochem. J*, 1975, **152**, 647–654.
46 M. S. Gelfand, *Curr. Opin. Struct. Biol.*, 2006, **16**, 420–429.
47 T. Atlung and L. Brondsted, *J. Bacteriol.*, 1994, **176**, 5414–5422.
48 J. Pouyssegur and F. Stoeber, *J. Bacteriol.*, 1974, **117**, 641–651.
49 D. A. Rodionov, M. S. Gelfand and N. Hugouvieux-Cotte-Pattat, *Microbiology*, 2004, **150**, 3571–3590.
50 K. Nishino and A. Yamaguchi, *J. Bacteriol.*, 2002, **184**, 2319–2323.
51 K. Nishino, Y. Inazumi and A. Yamaguchi, *J. Bacteriol.*, 2003, **185**, 2667–2672.
52 M. T. Gallegos, R. Schleif, A. Bairoch, K. Hofmann and J. L. Ramos, *Microbiol. Mol. Biol. Rev.*, 1997, **61**, 393–410.
53 V. Torsvik, L. Ovreas and T. F. Thingstad, *Science*, 2002, **296**, 1064–1066.
54 I. Cases, V. de Lorenzo and C. A. Ouzounis, *Trends Microbiol.*, 2003, **11**, 248–253.
55 M. Madan Babu and S. A. Teichmann, *Nucleic Acids Res.*, 2003, **31**, 1234–1244.
56 M. N. Price, P. S. Dehal and A. P. Arkin, *Genome biology*, 2008, **9**, R4.


Published on 07 March 2011 on http://pubs.rsc.org | doi:10.1039/C0MB00336K
Downloaded by University of Arizona on 09 March 2011

 This journal is © The Royal Society of Chemistry 2011