

SCIENTIFIC REPORTS



OPEN

Global transcriptomic responses of *Escherichia coli* K-12 to volatile organic compounds

Received: 09 September 2015

Accepted: 21 December 2015

Published: 28 January 2016

Pui Yi Yung¹, Letizia Lo Grasso¹, Abeed Fatima Mohidin¹, Enzo Acerbi¹, Jamie Hinks¹, Thomas Seviour¹, Enrico Marsili^{1,2,3} & Federico M. Lauro^{1,4}

Volatile organic compounds (VOCs) are commonly used as solvents in various industrial settings. Many of them present a challenge to receiving environments, due to their toxicity and low bioavailability for degradation. Microorganisms are capable of sensing and responding to their surroundings and this makes them ideal detectors for toxic compounds. This study investigates the global transcriptomic responses of *Escherichia coli* K-12 to selected VOCs at sub-toxic levels. Cells grown in the presence of VOCs were harvested during exponential growth, followed by whole transcriptome shotgun sequencing (RNAseq). The analysis of the data revealed both shared and unique genetic responses compared to cells without exposure to VOCs. Results suggest that various functional gene categories, for example, those relating to Fe/S cluster biogenesis, oxidative stress responses and transport proteins, are responsive to selected VOCs in *E. coli*. The differential expression (DE) of genes was validated using GFP-promoter fusion assays. A variety of genes were differentially expressed even at non-inhibitory concentrations and when the cells are at their balanced-growth. Some of these genes belong to generic stress response and others could be specific to VOCs. Such candidate genes and their regulatory elements could be used as the basis for designing biosensors for selected VOCs.

Volatile organic compounds (VOCs) are low molecular weight molecules with a vapor pressure of ≥ 10 Pa at 20 °C¹, while compounds with a 6-months volatility between 5 and 95% at ambient temperature can be termed semi VOCs (sVOCs)². VOC such as toluene, methyl acetate, trichloroethylene, benzene, and phenol etc., are common indoor and urban contaminants³. Examples of common sVOCs include high molecular weight alkanes, polycyclic aromatic hydrocarbons (PAH), organochlorine pesticides, and substitute benzenes^{4,5}. While there are natural VOCs (e.g. cyclopentanone and dimethyl disulfide) and sVOCs (e.g., n-Heptadecane and 1-butyl-3-methyl-imidazolium hexafluorophosphate) produced biologically during degradation products of plant components or for biochemical signaling^{6,7}, many VOCs and sVOCs originate from fossil fuels, industrial chemicals and solvents. These compounds present a challenge to receiving environments and wastewater treatment processes, due to their toxicity and low bioavailability for degradation⁸.

The toxicity of VOCs and sVOCs has been evaluated in selected animal models. For example, cyclopentanone, N-methyl-2-pyrrolidone (NMP) and dimethylacetamide (DMA), were found to cause developmental toxicity in rat embryos⁹ and rabbits^{10,11}. The toxicity of sVOCs commonly found in indoor environment, such as plasticizers, solvents, and flame retardant is also well studied¹². In industrial settings, VOC and sVOCs have been shown to concentrate in both liquid and gas phases of wastewater treatment plant^{13,14}. Thus, industrial VOC and sVOC discharges present serious concerns for wastewater treatment.

Microorganisms are constantly sensing and responding to surrounding environmental conditions, including the presence of biologically toxic compounds. VOCs have been found to affect microbial diversity and biodegradation performance in activated sludge¹⁵ and in soil¹⁶. Microbial tolerance to various VOCs in bacteria falls into three broad mechanisms: 1) alteration of membrane protein composition¹⁷⁻¹⁹, 2) export of toxic compounds through membrane transporters^{20,21}; and 3) to a lesser extent, biotransformation of the compound to less toxic variant, which has been reported for soil microorganisms and a number of *Pseudomonas* species²². Expression of

¹Singapore Centre for Environmental Life Sciences Engineering (SCElse). 60 Nanyang Drive, SBS-01N-27, Singapore 637551. ²School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, Singapore 637459. ³School of Biotechnology, Dublin City University, Collins Avenue, Dublin 9, Ireland. ⁴Asian School of the Environment, Nanyang Technological University, 50 Nanyang Avenue, N2-01C-45, Singapore 639798. Correspondence and requests for materials should be addressed to F.M.L. (email: FLauro@ntu.edu.sg)

Chemical	Linear formula	MW	v/v%	ppm	mM	logP _{ow}	V (Pa)	Melting	Boiling
n-butanol (B)	CH ₃ (CH ₂) ₃ OH	74.12	0.1	1000	13.49	0.88	1200	−89.9	117.7
N-cyclohexyl-pyrrolidone (CHP)	C ₁₀ H ₁₇ NO	167.25	0.1	1000	5.98	1.41	6.7	12	284
Cyclopentanone (CP)	C ₅ H ₈ (=O)	84.12	0.1	1000	11.89	0.7	1520	−58	131
N,N-dimethylacetamide (DMA)	CH ₃ CON(CH ₃) ₂	87.12	0.5	5000	57.39	−0.7	300	−20	165
Dimethyl sulphide (DMS)	(CH ₃) ₂ S	62.13	0.1	1000	16.1	0.92	53700	−98	35–41
1-methyl-2-Pyrrolidone (NMP)	C ₅ H ₉ NO	99.13	0.1	1000	10.09	−0.46	42	23–24	201
N-methyl succinimide (NMS)	C ₅ H ₇ NO ₂	113.11	0.5	5000	44.2	NA	NA	65	235
Toluene (T)	C ₆ H ₅ CH ₃	92.14	0.02	200	2.17	2.69	2800	−95	110.6

Table 1. Basic properties of the VOCs used in the study. MW: Molecular weight; v/v% concentration in volume to volume ratio used for RNAseq based on a semi-MIC quantification assay; ppm: parts per million; mM: milliMolar; logP_{ow} value: the logarithm of partitioning coefficient in a defined octanol-water mixture values; V (Pa): Vapor pressure in Pa; Melting: melting temperature (°C); Boiling: boiling temperature (°C); NA: Not available.

detoxifying enzymes such as reductive dehalogenases²³ and oxygenases²⁴, have been exploited in the bioremediation of chlorinated aliphatic hydrocarbons and polycyclic aromatic contaminated soil and groundwater.

The *E. coli* K-12 MG1655 strain used in this study is the primary experimental reference model with a highly curated genome sequence with annotation²⁵. It is widely considered the *E. coli* strain of choice and its genome was the first published sequence of a wild-type laboratory strain of *E. coli* K-12 because it has relatively few genetic modifications compared to most other *E. coli* strains. *E. coli* has also been used extensively as biosensor due to its ease of genetic manipulation and availability of information²⁶. We chose *E. coli* K-12 also because the GFP: fusion library is readily available²⁷. Various genetic mechanisms have been identified to contribute to VOC tolerance in *E. coli*. For example, membrane transport proteins like the *acrAB-tolC* pump²⁸, mannose transporter²⁹ and phosphate transporters³⁰ in *E. coli* were found to confer tolerance to various VOCs. Regulatory elements such as the *FadR*, *MarR*³¹ and *purR* regulon³², were found to be involved in conferring tolerance to n-hexane, p-xylene and cyclohexane. Genes under the central metabolic processes, such as the *cyo* and *nuo* operons, responsible for energy conservation and production, and those under galactitol metabolic process (*gat* genes) were up-regulated in response to ethanol³⁰ and butanol³³, respectively. Overexpression of heat shock proteins, such as the *GrpE* and *GroESL* chaperone system also resulted in increased tolerance of various forms of butanol as well as ethanol^{34,35}. In addition, studies has been conducted to look at tolerance of *E. coli* to butanol using genomic library screening³⁶, microarray, and at proteomic, regulatory network and metabolite levels^{33,37,38}.

In most of these studies, genetic responses to sub-toxic VOC and sVOC concentrations have not been described. Studying gene activation/inactivation following exposures to sub-toxic levels will enable mechanisms of adaptation and enhanced tolerance to be decoupled from general stress responses, which would be expected at higher concentrations. In addition, information on the genetic responses of microorganisms to non-inhibitory levels would be relevant to understand and improve VOC and sVOC resistance in microorganisms that can be used for biocatalysis (e.g. for the removal of VOCs and sVOCs) applications. Such information would be preliminary to the development of rapid biosensing of VOC and sVOC in contaminated wastewater, offering protective measures for wastewater treatment plants and final users of reclaimed water³⁹.

In this study, we used transcriptomics to investigate the global gene expression of *E. coli* K-12 grown in the presence of industrially relevant VOCs and sVOCs. All of the selected compounds are commonly used as solvents or produced as by-products during manufacturing of polymers, cleaners and industrial chemicals, with an exception of N-methylsuccinimide (NMS), which is one of the metabolites commonly used as a biomarker for exposure of the solvent N-methyl-2-pyrrolidone (NMP)⁹. We aim to understand the specific and non-specific responses to the selected compounds. The focus in this study is to investigate genes that are responsive at non-growth inhibitory concentration, yet significant enough to induce a response at the transcriptome level.

Results and Discussion

Growth and overall transcriptome profile of *E. coli* grown with VOCs. We analyzed the transcriptome of *E. coli* K-12 grown in the presence and absence of selected VOCs (Supplementary Figure S1) using Illumina RNA-seq. Growth curve experiments were performed on *E. coli* with 0 (as control), 0.02, 0.1 and 0.5% (v/v) of the selected VOCs to determine the highest non-inhibitory concentration to be used in RNAseq experiments (Supplementary Methods and Figure S2). The concentrations were established to be: 0.02% for toluene (T), 0.1% for n-butanol (B), N-cyclohexyl-pyrrolidone (CHP), cyclopentanone (CP), dimethyl sulfide (DMS), N-methyl-2-pyrrolidone (NMP); 0.5% for N,N - Dimethylacetamide (DMA) and N-methyl succinimide (NMS) (Table 1). At these concentrations the cells reach optical density (600nm) of 0.4 in approximately 5–6 h from initial O.D. of 0.02 in MOPS media (Supplementary Figure S2). There was a slight growth inhibition on DMS and CHP treatment during mid-log growth at concentration of 0.1%. We have regarded this inhibition as non-significant and have chosen this concentration for subsequent RNA extraction. Previous work using *E. coli* to study the genes involved in tolerance (using microarray/genomic library screening) of selected VOCs uses a range of concentrations from 0.5%³⁶ to 1.7% butanol³⁶, and up to 10% for toluene²⁹. The concentration of n-butanol that caused a 50% growth decrease in M9 medium in *E. coli* DH1 was 0.8%(v/v)³³. Most of these studies used concentrations that are growth inhibitory to *E. coli*. We expect that the transcriptome of *E. coli* using non-inhibitory levels of compounds used in the current study would provide new insights compared to existing literature.

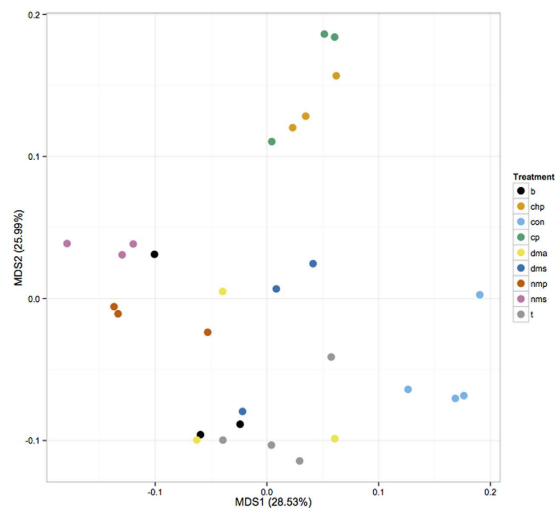


Figure 1. Multi-dimensional scale (MDS) plot of the global gene expression profiles of *E. coli* samples based on normalized feature count values. B: n-butanol; chp: N-cyclohexyl-pyrrolidone; con: no VOC control; cp: cyclopentanone; dma: N,N-dimethylacetamide; dms: dimethyl sulphide; nmp: 1-methyl-2-pyrrolidone; nms: N-methyl succinimide; t: toluene.

In the present transcriptomic analysis, read mapping against the *E. coli* K-12 MG1655 genome was performed which allowed us to identify differentially expressed genes. The analysis identified the expression of 4140 coding DNA sequence (CDS) tags. The non-metric multidimensional scaling (NMDS) plot of global mRNA expression profiles revealed separate clustering patterns on cells grown with VOC compared to the no VOC controls, with NMS and NMP-treated cells clustering furthest from the controls on the first dimension (Fig. 1). Biological replicates for most VOC treatments clustered tightly indicating consistency between the replicates, although the clustering for treatment DMS, DMA, B are not as tight compared to the rest of the treatments. The differentially expressed (DE) genes identified (with cut off at log fold change of greater than 1 or less than -1 , an average logCPM value of greater than or equal to 3, and a p -value less than 0.05) are distributed across a range of average logCPM values (Supplementary Figure S3). More DE genes were up rather than down regulated following treatment by B, DMA, DMS, and T. The converse was true for CHP, CP, NMP, and NMS-treated cells (Supplementary Figure S3, Table 2). The percentages of genes identified as significantly differentiated over the total gene expression profile in VOC treatments compared to the controls ranged from 9.28% (DMA) up to 25.94% (NMS) (Table 2). Similar trend was found for chemical-specific gene responses (identified based on Venn analysis of DE genes), with DMA having the lowest (1.30%) and NMS the highest percentage (24.21%) (Table 2, Fig. 2). In addition, a total of 625 DE genes were shared by four or more VOC treatments, suggesting a subset of common genetic responses. The expression pattern for these DE genes appears to be divided into two major clusters for the VOCs used in this study (Fig. 3). Cells grown with B, T, DMS and DMA elicited more similar transcription patterns than CHP, CP, NMP and NMS. (Fig. 3). These observations suggest that some VOCs might induce more cellular responses compared to others at non-growth inhibitory concentrations. Clustering of the overall transcriptome patterns of VOC treatments (Fig. 1) had some resemblance compared to the shared DE gene profiles (Fig. 3). For example, the profile of treatment CHP and CP, NMP and NMS are clustering closer to each other compared to other treatment in both the MNDS and heatmap plots. The relationship between the chemical properties of the compounds tested and the degree of cellular response in *E. coli* would be an interesting investigation in the future.

A number of genes relating to cold-shock responses were up regulated in our transcriptomic datasets (Supplementary Table S1). We have disregarded these genes as response to VOCs as the promoter clones for these genes failed to show an increase in GFP expression compared to the control in our bioassays at 37 °C (Supplementary Figure S4). These cold-shock related genes are likely to be an artifact of concentrating the biomass at 4 °C.

Functional gene categories induced by multiple chemical treatments. *Induction of iron-sulfur assembly system.* Fe/S proteins participate in diverse biological processes such as respiration, central metabolism, DNA repair and gene regulation⁴⁰. The iron-sulfur cluster (ISC) and sulfur mobilization (SUF) systems carry out biogenesis and maturation of all Fe/S clusters in prokaryotes. In the ISC system, IscU and IscS are required to build the Fe/S cluster, followed by release of Fe/S cluster by HscA and HscB. In the SUF system, SufSE forms the Fe/S cluster, and SufBCD complex is responsible for cluster transfer and release⁴⁰. The compounds used in the current study had a higher expression of genes under different Fe/S cluster biogenesis system compared to the no chemical treatment control. Both ISC and SUF systems were activated following B and CP treatment, while only the ISC system is activated following CHP, NMP, NMS and T treatment, and only SUF system was activated when cells were grown with DMA and DMS (Table 3). This suggests that different chemicals induced distinctive responses in Fe/S assembly systems. *IscR*, a gene encoding the regulator responsible for Fe/S homeostasis and regulates the expression of a number of Fe/S proteins⁴¹, was also up regulated in cells exposed to

Sample	Up-all	Down-all	Sum-all	Up-specific	Down-specific	Sum-specific	% Up-all	% Down-all	% Sum-all	% Up-specific	%Down specific	% Sum specific
B	430	249	679	25	12	37	10.39	6.01	16.40	3.68	1.77	5.45
CHP	324	388	712	33	37	70	7.83	9.37	17.20	4.63	5.20	9.83
CP	403	494	897	87	64	151	9.73	11.93	21.67	9.70	7.13	16.83
DMA	261	123	384	4	1	5	6.30	2.97	9.28	1.04	0.26	1.30
DMS	274	120	394	7	3	10	6.62	2.90	9.52	1.78	0.76	2.54
NMP	445	406	851	71	21	92	10.75	9.81	20.56	8.34	2.47	10.81
NMS	506	568	1074	113	147	260	12.22	13.72	25.94	10.52	13.69	24.21
T	418	223	641	39	17	56	10.10	5.39	15.48	6.08	2.65	8.74

Table 2. Summary of total the number of genes significantly up- or down-regulated (Up-/Down- all), and those responsive to specific VOCs in *E. coli*'s transcriptome (Up-/Down- specific). Sum: Sum of Up- and Down- regulated genes per treatment. “%-all” indicates the percentage of DE genes out of the 4140 locus tag analyzed, “%-specific” indicates the percentage of DE genes specific to the VOC out of the sum of DE genes of the particular VOC treatment [e.g. % Up-specific = “Up-specific”/“Sum-all” x 100].

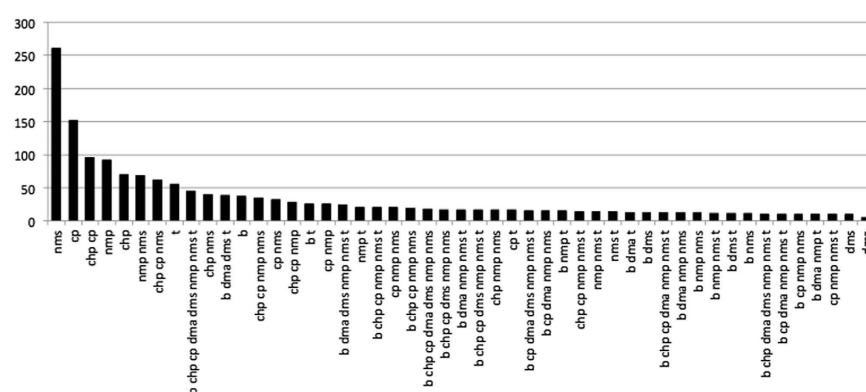


Figure 2. Number of differentially expressed genes based on Venn analysis of eight different treatments. Numbers of genes responsive only to single treatment are reported for all eight VOCs. For genes that were shared between two or more treatments, only those with number of shared genes exceeded ten were plotted.

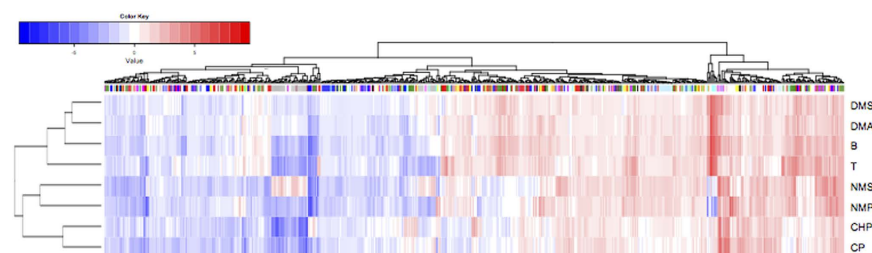


Figure 3. Heatmap of logFC values of 625 DE genes that occurred in at least four VOC treatments in *E. coli*'s transcriptome. Color bar on column dendrogram indicate COG functional category of the DE genes. Only COG categories containing more than 5 DE genes are colored. COG category with no known function, such as “R” and “S” are not colored. Color key: Green: Energy production and conversion [C]; Red: Amino acid transport and metabolism [E]; light green: Nucleotide transport and metabolism [F]; Blue: Carbohydrate transport and metabolism [G]; Orange: Coenzyme transport and metabolism [H]; Light purple: Lipid transport and metabolism [I]; Light blue: Translation, ribosomal structure and biogenesis [J]; Magenta: Transcription [K]; Pink: Replication, recombination and repair [L]; Maroon: Cell wall/membrane/envelope biogenesis [M]; Grey: Cell motility [N]; Black: Posttranslational modification, protein turnover, chaperones [O]; Yellow: Inorganic ion transport and metabolism [P]; Light orange: Secondary metabolites biosynthesis, transport and catabolism [Q]; Cyan: Signal transduction mechanisms [T]; Olive: Defense mechanisms [V].

the eight compounds tested in the current study. The up regulation of *iscR* is validated through promoter: GFP fused expression assays (Supplementary Figure S5). *IscR* represses its own expression when there is sufficient Fe/S cluster in the cell, and the *isc* operon is activated when cells are under Fe/S cluster-limiting and oxidative

Treatment	ID	Gene	B	CHP	CP	DMA	DMS	NMP	NMS	T	Gene description	ISC/SUF
b dms	b1679	sufE	1.49	0.71	0.80	0.79	1.02	-0.02	-0.91	0.91	Sulfur acceptor protein	SUF
b cp dma dms	b1680	sufS	1.83	0.68	1.21	1.21	1.35	0.28	0.36	0.91	Cysteine desulfurase, SufE induced	SUF
b cp dma dms	b1681	sufD	1.72	0.88	1.46	1.11	1.14	0.33	-0.16	0.71	Component of SufBCD Fe-S cluster assembly scaffold	SUF
b cp dms	b1682	sufC	1.77	0.99	1.97	1.11	1.30	0.42	-0.67	0.92	SufBCD Fe-S cluster assembly scaffold protein	SUF
b chp cp dma dms nmp t	b1683	sufB	2.71	1.54	3.12	1.87	2.02	1.44	0.11	1.77	Component of SufBCD Fe-S cluster assembly scaffold	SUF
b cp dma dms t	b1684	sufA	3.01	1.57	4.06	2.12	2.30	1.73	0.76	2.36	Fe-S cluster assembly protein	SUF
b chp cp nms	b2525	fdx	1.03	1.31	1.27	0.59	0.45	0.67	1.69	0.48	[2Fe-2S] ferredoxin	ISC
b chp cp nms	b2526	hscA	1.11	1.03	1.22	0.74	0.62	0.88	1.9	0.76	DnaK-like molecular chaperone specific for IscU	ISC
b chp cp nmp nms t	b2527	hscB	1.43	1.38	1.69	0.89	0.93	1.27	2.44	1.16	IscU-specific HscA co-chaperone Hsc56	ISC
chp cp nmp nms	b2528	iscA	0.97	1.05	1.33	0.41	0.51	1.39	1.52	0.72	FeS cluster assembly protein	ISC
b chp cp nmp nms t	b2529	iscU	1.31	1.50	1.78	0.59	0.62	1.79	2.07	1.14	Iron-sulfur cluster assembly scaffold protein	ISC
b chp cp dms nmp nms t	b2530	iscS	1.62	1.58	1.94	0.97	1.05	2.69	2.48	1.64	Cysteine desulfurase (tRNA sulfurtransferase)	ISC
b chp cp dma dms nmp nms t	b2531	iscR*	1.91	1.62	2.17	1.38	1.32	3.81	3.07	2.00	Isc operon repressor; suf operon activator	Regulator
chp cp nmp nms t	b3414	nfuA	0.66	1.31	1.71	0.5	0.23	1.75	1.41	1.32	Fe/S biogenesis protein; putative scaffold/chaperone	Fe/S carrier
b chp cp dms nmp nms	b4705	mntS	1.14	2.06	2.34	0.89	1.12	2.64	2.97	0.94	Mn(2)-response protein, MntR-repressed	Oxidative stress
b chp cp dma dms nmp nms t	b1778	msrB*	1.63	1.33	1.07	1.39	1.05	1.76	1.93	2.52	Methionine sulfoxide reductase B (EC:1.8.4.12)	Oxidative stress
b chp cp dma dms nmp nms	b0950	pqiA*	1.33	1.07	1.1	1.03	1.08	1.4	1.99	0.83	Paraquat-inducible, SoxRS-regulated inner membrane protein	Oxidative stress
b cp nmp nms	b0951	pqiB	1.14	0.8	1.05	0.75	0.71	1.19	1.5	0.77	Paraquat-inducible, SoxRS-regulated MCE domain protein	Oxidative stress
b chp cp dma dms nmp nms t	b2294	yfbU*	-1.84	-1.36	-1.18	-1.11	-1.39	-1.70	-1.99	-1.43	UPF0304 family protein; K09161 hypothetical protein	Oxidative stress
b dms nmp t	b3238	yhcN	1.92	1.11	1.39	1.39	1.6	2.25	1.26	1.69	Cadmium and peroxide resistance protein	Oxidative stress
b chp cp dma dms nmp nms t	b3495	uspA*	-1.79	-1.63	-1.49	-1.29	-1.63	-2	-3.16	-1.4	Universal stress global response regulator A	Usps
nms	b1895	uspC	0.78	0.34	0.75	-0.03	-0.24	0.61	-1.44	-0.55	Universal stress induced protein C	Usps
chp nms	b3923	uspD	-0.74	-1.23	-0.83	-0.73	-0.83	-0.58	-2.03	-0.42	Universal stress-induced protein D	Usps
b dma dms t	b1333	uspE*	7.16	1.96	1.39	5.79	6.71	-5.12	-1.81	6.96	Universal stress-induced protein E	Usps
b dma dms t	b1376	uspF	3.79	-0.09	-0.36	2.23	3.75	-6.06	-2.75	4.16	Universal stress-induced protein F, ATP-binding protein	Usps
b dma nmp nms t	b0607	uspG*	-1.48	-0.25	-0.03	-1.62	-0.88	-1.18	-2.4	-1.48	Universal stress protein UP12	Usps

Table 3. Genes responsible for Fe/S cluster biogenesis, oxidative and universal stress responses. *E. coli* contains the ISC and SUF Fe/S assembly system. Treatment: Chemical treatment associated with the DE genes; ID: Gene ID; the shaded cells and bolded numbers are not shown in the table. Genes marked with “*”: Gene promoter-fused GFP assays performed (Supplementary Figure S5).

stress conditions. Overexpression of *iscR* might indicate that the chemicals tested in the current study elicited an oxidative stress or iron-limiting condition on the cells. This could be caused by the action of the VOCs on outer membrane proteins³³. In addition, the SUF system is believed to provide better resistance to iron^{40,42} and oxidative stresses compared to the ISC system^{43–45}. Whether the induction of the SUF system when cells were exposed to DMA and DMS is directly linked to oxidative stress is unknown, as other regulators, like Fur, OxyR are also known to be involved in SUF-type Fe/S regulation⁴⁰.

Oxidative stress responses. A number of genes known to be induced by oxidative agents were up regulated in response to at least 4 VOCs used in the current study (Table 3 and Supplementary Figure S5). PqiAB is a SoxRS-regulated membrane protein known to be induced by paraquat and other superoxide generators, but it is not induced by hydrogen peroxide, ethanol and heat shock⁴⁶. YhcN was identified as a stress protein associated with hydrogen peroxide, cadmium and acid⁴⁷. MntS confers resistance to hydrogen peroxide by facilitating

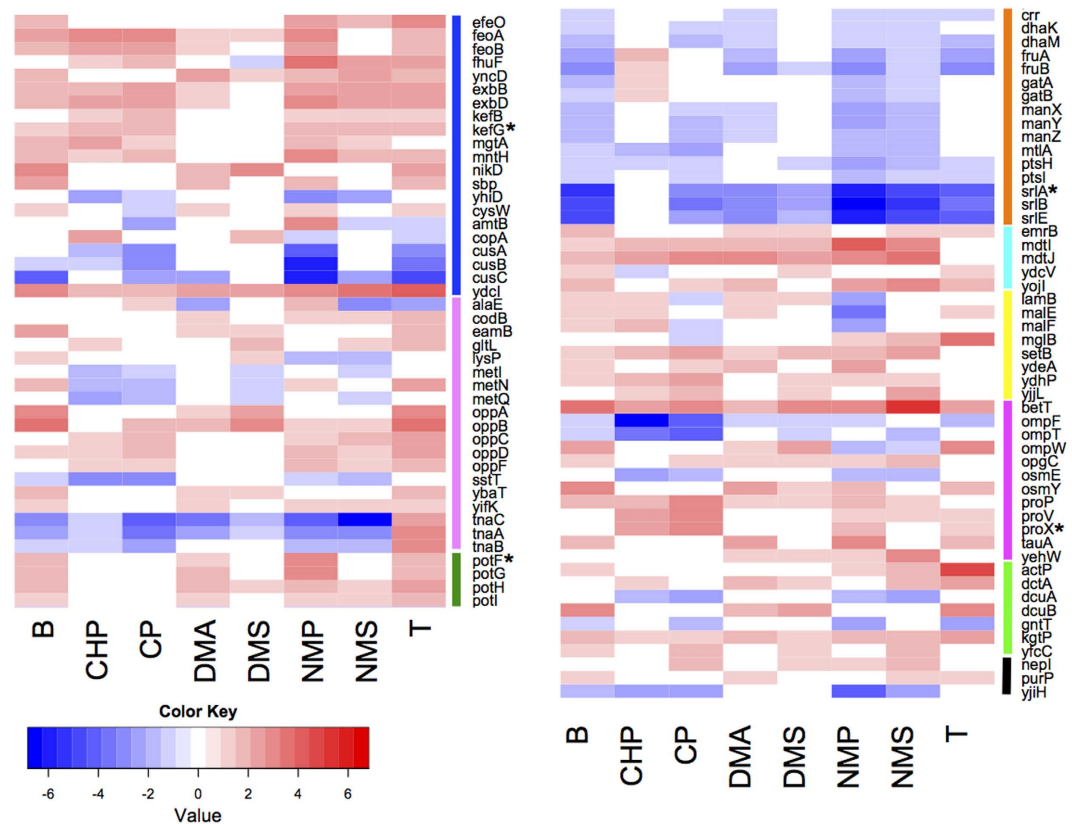


Figure 4. Heatmap of log fold-change values of transporter genes. Type of transporter proteins identified as differentially expressed in at least 4 VOC treatments. Color code for transporter type: blue: inorganic ion; pink: amino acid; green: putrescine; orange: PTS; cyan: multi-drug efflux; yellow: sugar; magenta: osmosis; lime: organic acids; black: others. Cells in white are values considered as insignificant (p -value < 0.05). Genes marked with “*”: Gene promoter-fused GFP assays performed (Supplementary Figure S5).

delivery of Mn^{2+} to Mn^{2+} -dependent enzymes⁴⁸. A gene encoding for methionine sulfoxide reductase, *msrB*, was up regulated as well. MsrB repairs methionine residues in proteins that have been oxidized by reactive oxygen species⁴⁹. Collectively, the results indicate that *E. coli* cells exposed to the compounds tested in the study induce oxidative stress responses even at non-inhibitory concentrations. In addition, there might be proteins oxidized by the presence of VOCs. *yfbU*, a gene known to be involved in cell death by oxidative DNA damage⁵⁰, was down regulated in all treatments, suggesting that the cells did not go through the toxin:antitoxin response when grown with chemical tested, but instead employ alternative oxidative stress responses as described.

Induction of various transporter proteins. Transporter proteins for inorganic ions, amino acids, and the PTS systems were among the top three categories to be differentially expressed in at least 4 chemical treatments compared to the control (Fig. 4 and Supplementary Figure S5). Genes involved in the uptake of both inorganic iron (e.g. *feoA*, *feoB* and *efeO*), and siderophores (*exbBD*, *yncD* and *fhuF*) were up regulated. Genes involved in iron uptake have been shown to increase *E. coli*'s tolerance to environmental stresses. For example, over expression of *feoA* increases the tolerance of *E. coli* to butanol³⁶, and *efeO* confers resistance to mitomycin C and other stresses such as UV irradiation compared to wild type cells⁵¹. ExbB and ExbD proteins are required to provide energy for the import of iron-siderophore complexes and vitamin B12 across the outer membrane via TonB^{52–54}. YncD, a putative TonB-dependent outer membrane transporter for iron⁵⁵, could be one of the protein targets of TonB-ExbB-ExbD. The FhuF protein is required for cells to use hydroxamate-type siderophores as iron source⁵⁶. Collectively, up regulation of iron uptake genes implies that the cells are actively utilizing iron, possibly for the formation of Fe/S cluster containing proteins as described above.

Transporters for other inorganic ions were also up regulated (Fig. 4), e.g., genes for magnesium (*mgfA*) and manganese (*mntH*) uptake. MntH was shown to support the growth of *E. coli* cells encountering iron-deficiency and oxidative stress⁵⁷. During H_2O_2 stress, mutants lacking ability to import manganese and iron suffer high rates of protein oxidation, implying the role of MntH in preventing protein damage. Potassium efflux genes (*kefB* and *kefG*) were up regulated too. Efflux of potassium is known to play a role in protecting the cell from electrophilic toxicity through acidification of the cytoplasm⁵⁸, suggesting cells grown with VOC might be undergoing electrophilic stress.

The second largest transporter type relates to amino acids (Fig. 4). In particular, the dipeptide ABC transporter, encoded by the *oppABCDF* operon, was up regulated in most VOC treatments. The OppABCDF system function in oligopeptide uptake as well as recycling of cell wall peptides⁵⁹. Expression of *opp* genes was up

regulated in cells treated with 1% isobutanol as an early stage response³⁸, and *oppD* increased antibiotic resistance in *E. coli* during biofilm formation⁶⁰. Increased expression of the *opp* genes support previous findings that these transporters are involved in VOC resistance. The *tnaCAB* gene cluster, responsible for the uptake of tryptophan, was down regulated in response to most VOC used. Mutants lacking *tnaCAB* had increased isobutanol tolerance⁶¹, supporting our finding that *tnaCAB* plays a negative role in VOC tolerance. The cytoplasmic putrescine transporter protein, encoded by *PpotFGHI*, was significantly up regulated following n-butanol, DMA, NMP and T treatment. The up regulation of *potG* stimulates cell growth in the presence of phenylpropanoids, which indicates that PotFGHI might also be involved in the import of this compound class⁶². Cells grown with VOCs could either have an elevated concentration of putrescine inside the cell, or could also play a role in transport of VOCs.

The third most abundant transporter class containing DE genes identified belong to the phosphotransferase (PTS) system, which is an active transport system responsible for uptake of nutrients in bacteria (Fig. 4). The PTS system is activated when ambient nutrient level is low⁶³. In this study, most of the DE genes under the PTS systems were down regulated in most VOC treatments, including those responsible for glucose, dihydroxyacetone, fructose, galactitol, mannose and glucitol. Down regulation of these systems could be explained by the high nutrient media utilized in growing the cells (1.5% glucose), hence the cells does not require active transport for nutrient uptake.

Other transporter types with differential gene responses include multidrug efflux proteins and those related to osmotic response (Fig. 4). Three genes related to multidrug efflux proteins, *mdtI*, *mdtJ* and *emrB*, were up regulated in most chemical treatment used in the current study. MdtJ and I are two components of a spermidine exporter⁶⁴ and *emrB* is known to increase tolerance to hydrophobic compounds, such as organomercurials and nalidixic acid⁶⁵ and thiolactomycin⁶⁶. Multidrug exporters are capable of exporting compounds consisting different structural components, hence they could potentially export the compounds tested in the current study. Genes known to be associated with maintaining appropriate osmotic conditions in cells, for example, *osmY*, and ABC transporters for transport of osmoprotectants like proline, glycine betaine, and taurine (*proP*, *proV*, *proX* and *tauA*) were up regulated. The VOC used in the current study might have an effect in the osmotic condition in *E. coli* cells, hence inducing the expression of this gene class. In addition, the expression of a DNA-binding transcriptional repressor known to confer organic and inorganic acid stress, *ycdI*, was up regulated in all VOC treatment. YdcI protein is conserved across gram-negative bacteria and a *S. typhimurium* mutant lacking this gene had decreased resistance to acid stress⁶⁷. Up regulation of *ycdI* genes in our study imply that this gene may also be a response to VOC.

Universal stress proteins. *E. coli* harbors six *usp* genes – *uspA*, *C*, *D*, *E*, *F* and *G*.^{68,69} The functions of Usps overlap to some extent, e.g. both UspA and UspE are involved in oxidative stress defense⁶⁸, while UspG and UspF are associated with fimbriae-associated adhesion^{68,70}. From the transcriptomic results of the current study, we observed a down regulation of *uspA* and *uspG* in most VOC treatments, while *uspE* and *uspF* were up-regulated in B, DMA, DMS and T (Table 3, Supplementary Figure S5). As UspA have functions that overlaps with UspE, down regulation of *uspA* can be compensated for by the expression in *uspE*. Similarly, down regulation of *uspG* expression can be compensated by up regulation of *uspF*.

Flagella and cellular motility. Many genes relating to flagella biosynthesis and motility (the *flg*, *flh* and *fli* genes) were significantly down regulated in all VOC treatments, with the exception of treatment NMS (Table S2). Previous studies have found that flagellar biosynthesis was down regulated in *E. coli* exposed to ethanol³⁰ as well as heat stress⁷¹. Since NMS is not a VOC, it is not surprising that these genes were not repressed. However, a decrease in expression of flagella genes did not result in a reduction in motility in soft agar motility assays (Supplementary Methods and Figure S6). It is possible that the *E. coli* cells have already synthesized the flagellum before flagellar gene repression occurring in the assay. Other possible reasons include the differences in growth condition of *E. coli* due to the nature of the motility assay, e.g. surface-associated soft agar versus liquid, and the time of incubation.

Functional gene categories induced by specific chemical treatments. *Shared DE genes responsive to CHP and CP.* A total of 96 genes responded significantly with specificity to CP and CHP, which shared the highest number of genes compared to other chemical treatment combinations (Figs 2 and 5). Top COG categories of the shared DE genes belong to Post-translational modification, protein turnover and chaperones (O), Amino acid transport and metabolism (E), Cell wall/membrane biogenesis (M) (Fig. 5). A few DE genes identified gave indications that CP and CHP might interfere with protein structure and outer membrane integrity. For example, the *mld* and *mldF* genes, which prevent accumulation of phospholipids (PLs) in the outer leaflet of the outer membrane in *E. coli* cells, were up regulated. Cells accumulate PLs in the outer leaflet of the OM when exposed to harsh chemical treatments. This process would disrupt the LPS organization and increasing sensitivity to small toxic molecules⁷². Up regulation of *mld* genes imply that the cells' OM lipid asymmetry could be disrupted in the presence of the chemicals tested. In addition, a number of genes encoding for molecular chaperones were significantly up regulated in response to CHP and CP. These include the protein (re)-folding chaperones (*htpG*, *fkpA*, *dnaK-DnaJ-GrpE* and the *GroES*), protein resolubilization chaperones (*clpB*) and a protease involving in clearing the defective peptides (*hslU*). Up regulation of these genes imply that CHP and CP cause cellular protein misfolding in *E. coli*.

Transporter-related genes specifically up regulated in response to CHP and CP include genes encoding for peptide transport proteins (*dtpD*), and a putative drug efflux system protein (*mdtG*). Overexpression of *mdtG* has been found to increase resistance to deoxycholate (bile acid) and the broad spectrum antibiotic fosfomycin⁷³. Up regulation of such multidrug efflux genes could imply that cells perceive CP and CHP compounds as drugs and attempt to export them out of the cells.

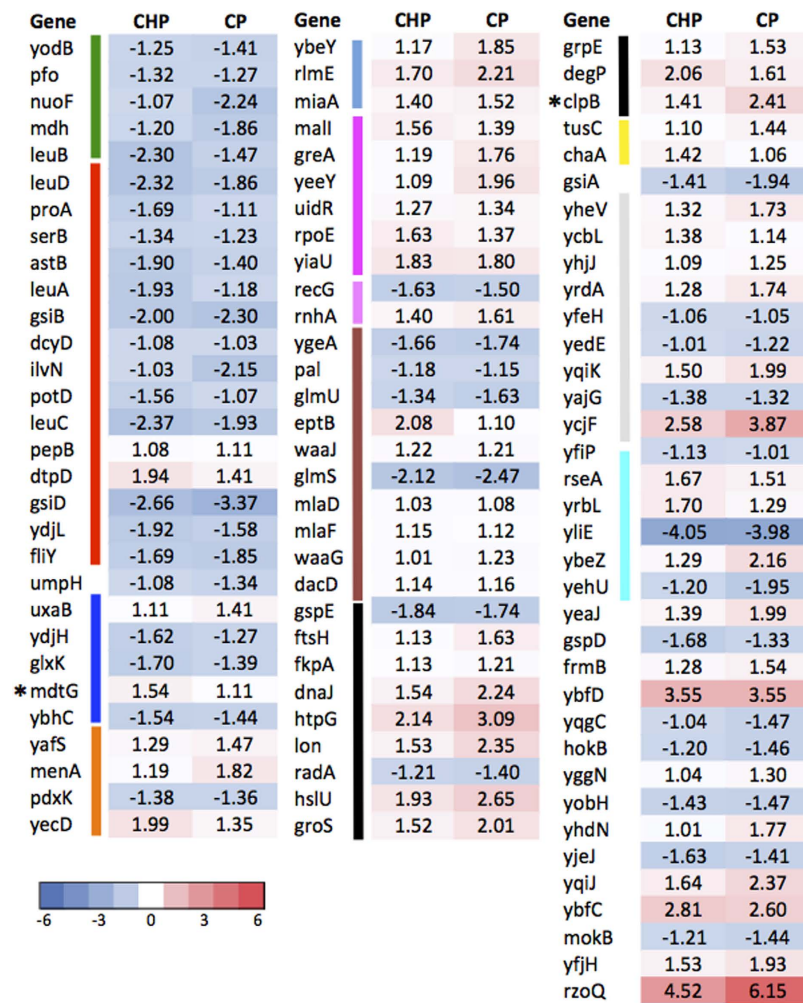


Figure 5. Log fold change values of the 96 shared DE genes between cells exposed to CHP and CP treatment. Color bars beside gene names indicate COG categories (labeling same as Fig. 4, except that COG category “R” and “S” are labeled as light grey). Gene name with “*”: expression tested on promoter:GFP fused *E. coli* clones.

Shared DE genes responsive to NMP and NMS. The next chemical pair sharing the highest number of DE genes is NMP and NMS, sharing 68 genes based on Venn analysis (Figs 2 and 6). NMP is an organic compound consisting of a 5-membered lactam and NMS is a metabolite of NMP biodegradation⁹. Although NMS is not considered as a VOC, it is cyclic. Most DE genes under energy production and conversion responding specifically to NMP and NMS were down regulated (e.g. *hyaDC*, *cbdAB* and *frdAD* genes), except for *rsxC*, which is part of the *rsx* operon (Fig. 6). The *rsxABCDEF* gene cluster is involved in switching off the SoxR-mediated induction of SoxS transcription factor when cells are deficient of oxidizing agents⁷⁴. Up regulation of these genes could imply that the cells cultured with NMP and NMS were less prone to oxidative stress and require SoxR reduction to repress downstream activation of SoxS. Interestingly, *rsxA* was shown to be important for survival of cells exposed to ionizing radiation⁷⁵.

Genes encoding for Tola-TolQ-TolR complex, were up regulated in cells treated with NMP and NMS (Fig. 6). The Tol-Pal cell envelope complex is known to be involved in maintaining cell envelope integrity, and mutants have greatly increase sensitivity to drugs and detergents and are prone to periplasmic leakage^{66,77}. Cells treated with NMP and NMS might respond differently to membrane disruption compared to that of CP and CHP. NMP and NMS activate the TolaAQR complex whereas cells exposed to CP and CHP activate the Mla pathway. The molecular mechanisms behind activation of different gene clusters in response to maintenance of cell envelope integrity would be an interesting area for future investigations.

Genes under “Defense mechanisms” that were up regulated include *arnE* and *nudE* which belong to the drug/metabolite transporter superfamily and the Nudix hydrolases family respectively (Fig. 6). Genes relating to iron-enterobactin transporter (*fepB* and *fepD*) and thiosulfate: cyanide (*glpE*) sulfurtransferase were up regulated specifically following NMP and NMS treatment. The *fepBCDG* complex together with the TonB-dependent outer-membrane transporter, and *fepA*, is responsible for the import of ferric enterobactin across the cell

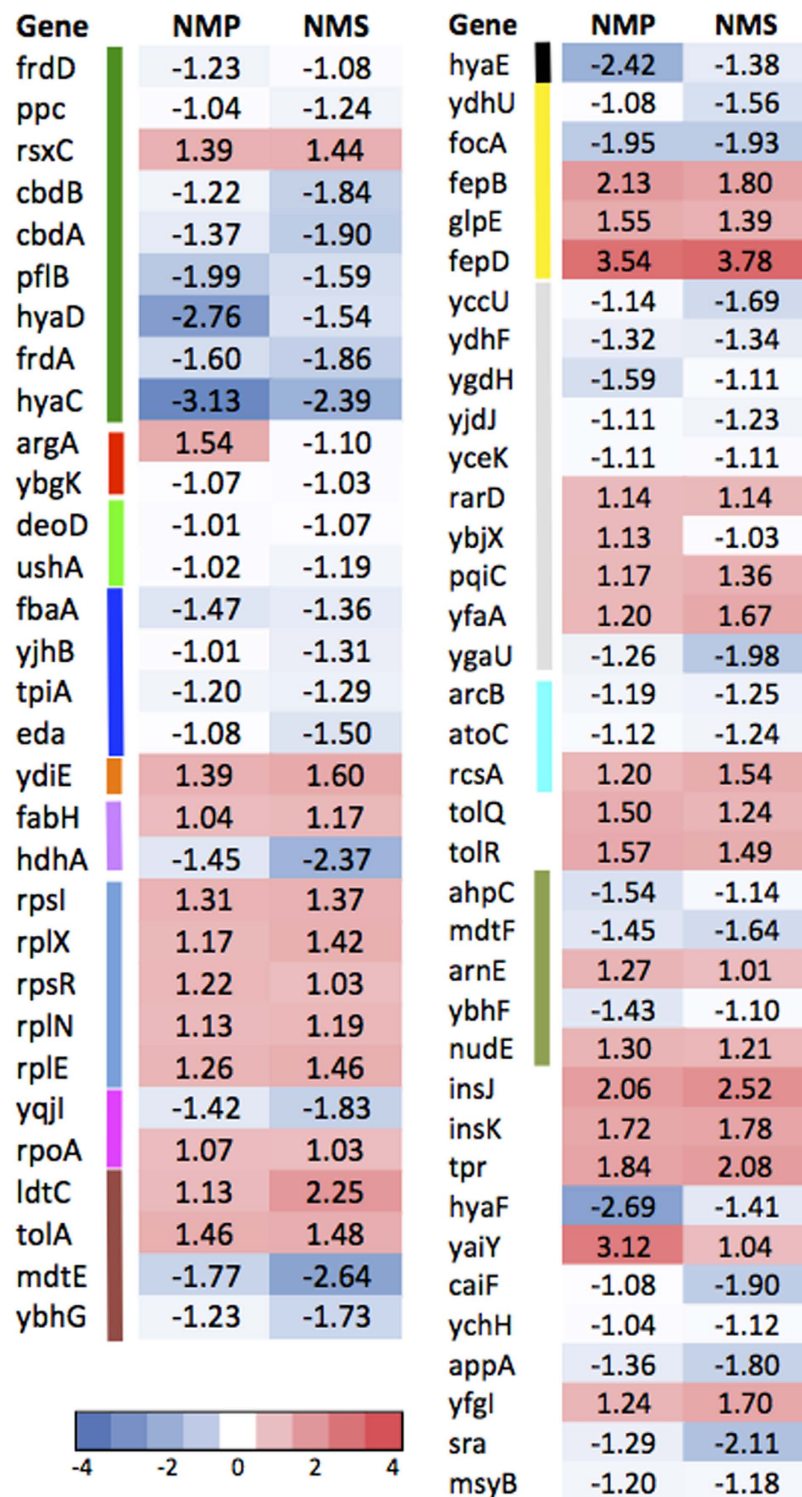


Figure 6. Log fold change values of the 68 shared DE genes between cells exposed to NMP and NMS treatment. Color bars beside gene names indicate COG categories (labeling same as Fig. 4, except that COG category “R” and “S” are labeled as light grey).

envelope. In addition to the iron-uptake system discussed in the previous sections, cells treated with NMP and NMS appear to have an additional iron-enterobactin transporter up regulated in the conditions tested in this study.

Stress and membrane repair-related DE genes responsive to one chemical treatment. Genes that responded positively to one particular VOCs were identified, with a number of them related to stress (*oxyR*, *dinF*, *ydiY*),

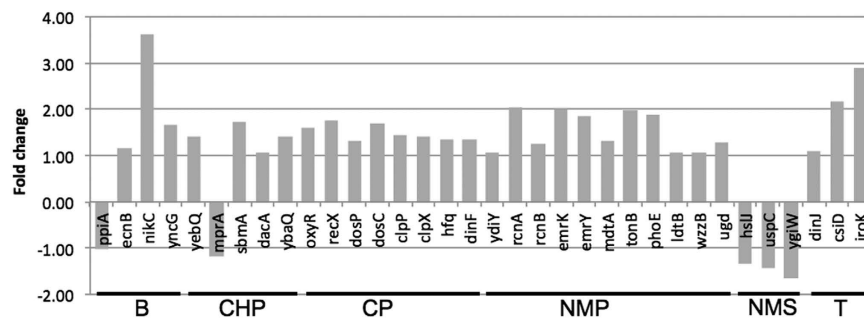


Figure 7. Fold change of the DE genes related to stress response and membrane repair induced by specific VOCs.

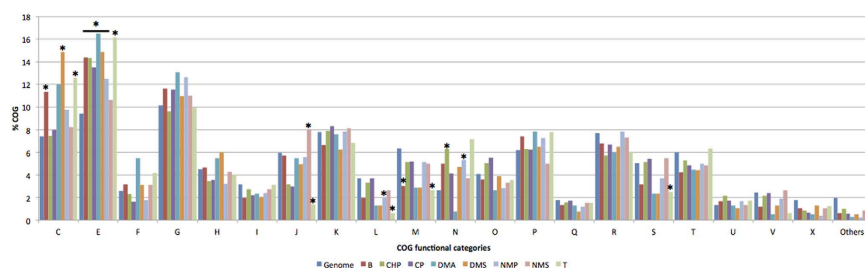


Figure 8. Percentage distribution of COG categories of the DE genes under selected VOC treatments.

Asterisks marked the COG categories in treated samples that had a significantly different COG distribution from the background composition of the *E. coli* genome (with confidence level of 0.99 and bootstrap replicates of 10000). Key for the various COG functional categories are as described in Fig. 4, with additional categories as follows: U: Intracellular trafficking, secretion, and vesicular transport; X: mobilome, phages, and transposons.

transport pumps for metals (*nikC*, *rcnA* and *rcnB*) and transport pumps for drugs (*emrKY*, *mdtA*, *sbmA*, *yebQ*) (Fig. 7). Expression of *emrK* (part of the EmrKY-TolC multidrug efflux transport system) was found to increase in the presence of sub-inhibitory concentration of a number of antibiotics⁷⁸. As the concentration of the chemical used in this study is considered non-inhibitory, results supported the conclusion that low concentrations of compounds are sufficient to induce a transcriptional response in various functional categories from the cell (Supplementary Figures S7). A number of genes relating to cell wall biogenesis were specifically up regulated when cells were exposed to NMP (*tonB*, *phoE*, *ldtB*, *wzzB*, *ugd*). Induction of these genes could imply that the cells have membrane component biosynthetic pathways activated specifically when exposed to NMP, implying that the involvement of NMP damages cell wall component, thus requiring repair.

COG category enrichment of DE genes. We performed COG enrichment analysis of total DE genes induced by individual chemical treatments against the *E. coli*'s genome copy of COG distribution (Fig. 8). More than half of the chemicals tested had amino acid related genes overrepresented compared to the *E. coli* genomic background. Amino acid metabolism is central to cellular survival and it is related to many parts of cellular metabolism. Genes under this category have been found to be differentially expressed in *E. coli* cells exposed to butanol^{33,36}. Cells exposed to n-butanol, DMS and toluene have a significant higher number of DE genes belonging to COG category of energy conversion, implying that genes under energy conversion are responsive to these VOCs. NMS is the only treatment that had genes relating to translation overrepresented. A total of 31 genes under translation and ribosomal biogenesis category were specifically responsive to NMS, many of them encode for ribosomal subunit proteins, implying that the cells were actively synthesizing proteins. Being a metabolic by-product of NMP, NMS is not a VOC, and appears that this metabolite does not impair cellular metabolism/growth at all. Both CHP and NMP had motility gene class overrepresented compared to *E. coli*'s background as well. COG categories were under represented compared to *E. coli*'s genomic background including those related to replication and repair for treatment NMP and T, and cell wall biogenesis were underrepresented in treatment n-butanol and toluene. Collectively these results could imply that different VOCs induced genes under specific COG categories.

Catabolic pathways of VOCs and aromatic compounds. Little is known about the biodegradation of VOCs focused on in this study. The genome of *E. coli* K-12 contains neither the genes responsible for the degradation of DMS (e.g. DMS monooxygenase, DMS dehydrogenase and DMS methyltransferase)⁷⁹, nor for toluene (i.e. toluene-2,3-dioxygenase)⁸⁰. Transcriptomic profiles of genes encoding for ring-hydroxylating oxygenases and transformation of aromatic compounds revealed very few differentially expressed genes in cells treated with the VOCs in the current study, with the exception of *tnaA* and *entA*, which were up-regulated following toluene treatment, and *ubiX*, *ubiB*, which were up-regulated with CHP and NMS treatment (Supplementary Table S3). However, these genes are also involved in generic cellular metabolism and their direct involvement in the

transformation of VOCs in this study is yet to be determined. A search for xenobiotics degradation pathways in KEGG (according to which some pyrrolidones have been classified), revealed that most of the described xenobiotics in KEGG are structurally very different from the VOC used here. Hence to the best of our knowledge, this study is the first to describe transcriptomic responses of *E. coli* K-12 exposed to VOCs with pyrrolidone backbone.

In conclusion, RNA-seq data in this study suggested that a variety of genes relating to Fe/S cluster biogenesis, oxidative and universal stress responses, as well as transport and membrane bound proteins are responsive to selected VOCs in *E. coli*. These genes were differentially expressed when the cells were in balanced-growth and at the highest non-inhibitory concentrations, which is well above the basal detectable environmental levels (PUB, personal communications). By identifying the transcriptional responses occurring between the basal levels and high concentration spikes, we have set the framework for the analysis of the dose dependent response, a key element in biosensor development. The numerous changes in gene expression upon exposure to the different VOCs suggests that *E. coli* might exhibit analogous response when exposed to chemical compounds of similar nature. It is interesting to speculate that the clustering of DE genes in response to different VOC tested could be related to the overall physical properties (polarity, volatility) and to the structure of the VOCs (i.e. linear chain vs cyclic compounds) used in the current study. Further studies are necessary to uncover the specific molecular mechanisms of *E. coli*'s cellular responses to chemical compounds of different structures. In addition, a number of DE genes described in this study, for example, those related to Fe/S cluster biogenesis, and various transporter genes, are conserved in other environmentally relevant bacteria, such as *Pseudomonas* species²⁰. Results from the current study hence could also be applicable to future biosensor development in bacteria other than *E. coli*. However, one should note that some *Pseudomonas* species are known to be able to metabolize a number of VOCs and cyclic hydrocarbons via enzymatic conversions^{22,81}, hence their global genetic response to VOCs might be different from *E. coli*.

Experimental Procedures

Chemicals. Chemicals used were as follows: n-butanol (B), N-cyclohexyl-pyrrolidone (CHP), Cyclopentanone (CP), N,N-Dimethylacetamide (DMA), Dimethyl sulphide (DMS), N-methyl-2-pyrrolidone (NMP), and N-methyl-succinimide (NMS) and Toluene (T). All were purchased from Sigma-Aldrich (Taufkirchen, Germany) and were of analytical purity.

***E. coli* cultivation and RNA extraction.** *E. coli* K-12 strain MG1655 was cultured in 10 mL LB5 broth within a shaking incubator at 150 rpm and at 37 °C for 16 h. The overnight culture was diluted (1:100) in 10 mL MOPS medium (Neidhardt *et al.* 1974) supplemented with 1.5% glucose. Based on the MIC assays (Supplementary Methods), different VOC concentrations were added at the beginning of cultivation (Table 1) and three biological replicates were used for each chemical treatment. Cells were grown in Balch-type tubes (18 × 150 mm) with 20 mm butyl rubber stopper and aluminum seal to minimize leakage of VOCs during the cultivation time. Cells were incubated in a shaking incubator at 37 °C and were harvested for RNA extraction when OD₆₀₀ reached 0.4. The RNA extraction was as follows: 5 mL aliquots of the cultures were added to two volumes of RNeasy Protect Bacteria Reagent (Qiagen). The mixture was incubated at room temperature for 5 min followed by centrifugation at 4,000 × g for 10 min at 4 °C. The supernatant was removed and the cell pellets were stored at –80 °C until RNA extraction. RNA was extracted using the RNeasy[®] Mini Kit (Qiagen), following the manufacturer's recommendations. Contaminating DNA was removed using DNase (Qiagen) until DNA concentration was less than 5% of the RNA. DNA and RNA concentrations were measured using Picogreen and Ribogreen assays (Invitrogen), respectively.

RNA sequencing. The quality of the RNA samples was determined by running the samples on a Bioanalyzer RNA 6000 Pico Chip (Agilent). Next-generation sequencing library preparation was performed following Illumina's TruSeq Stranded mRNA Sample Preparation protocol with the following modifications: RNA samples were added to the elute-fragment-prime step. The PCR amplification step, which selectively enriches for library fragments that have adapters ligated on both ends, was performed according to the manufacturer's recommendation. Each library was uniquely tagged with one of Illumina's TruSeq LT RNA barcodes to allow library pooling for sequencing. Library quantitation was performed using Invitrogen's Picogreen assay and the average library size was determined by running the libraries on a Bioanalyzer DNA 1000 chip (Agilent). Library concentrations were normalized to 2 nM and validated by qPCR on a ViiA-7 real-time thermocycler (Applied Biosystems), using qPCR primers recommended in Illumina's qPCR protocol, and Illumina's PhiX control library as standard. Libraries were then pooled and sequenced in one lane of an Illumina HiSeq2500 rapid sequencing run at a read-length of 101bp paired-end. Sequencing data have been submitted to GenBank SRA archive with the BioProject ID: PRJNA286974 and SRP accession SRP059483.

RNAseq data analysis. Quality trimming and adaptor removal were done using Cutadapt v1.9.0⁸² with the following parameters: –q 20, –m 30, –overlap 10, –quality-base 33. Sequences were mapped to the *E. coli* str. K-12 MG1655 genome (NCBI accession: NC_00913.3) by bowtie2⁸³ with end-to-end and very-sensitive modes. The alignments were converted to .bam and .bam-indexed files using Samtools⁸⁴. Sorted alignment files were imported into R to calculate overlapping reads as counts per gene using a combination of the following R packages: Rsamtools, GenomicFeatures and GenomicAlignments⁸⁵. Only the concordant pairs in the sorted *.bam files were imported using the function “readGAlignmentPairsFromBam”. Differential genes were identified from the tabular output of gene count abundance using edgeR package⁸⁶. The edgeR package implements a quantile-adjusted conditional maximum likelihood (qCML) estimator for the dispersion parameter of the negative binomial distribution^{86,87}. Testing for DE genes from biological replicates is based on the exact test derived based on these models. To calculate differentially expressed genes, all VOC treatments were compared to the

control in which the cells did not have any exposure to VOCs. DE genes that have at least 2-fold change, *p*-value less than 0.05 and logCPM value greater than 3 were considered significantly different from the no VOC control. Genes were mapped to COG and KEGG IDs using the December 2014 release of COG database⁸⁸ and June 2013 release of the KEGG database (Kanehisa Laboratories). Principal component analysis (PCoA), Venn and heatmap analysis were performed using R packages (vegan, venn, heatmap.2, respectively), and pathway maps were plotted using iPATH⁸⁹.

GFP kinetics using fluorescent transcriptional reporter *E. coli* clones. Selected *E. coli* clones with transcriptional fusions of GFP to relevant promoters of the identified DE genes were used to validate the RNAseq results²⁷. Reporter strains were inoculated from frozen stocks into 2× LB broth and incubated for 16 h at 37 °C. The cells were diluted (1:100) into fresh 1× MOPS medium supplemented with 25 μg/mL kanamycin and 1.5% glucose and grown as described previously. The VOC were added at the same concentration used in RNA experiments. When the OD_{600nm} reaches 0.35, an aliquot of culture was transferred to 96-well microplate. Optical densities were measured at OD_{595nm} and GFP intensity were measured at 485/535nm at 15 min interval for 4 h. Triplicates were performed and cells grown without VOC were used for comparison. *E. coli* clone with the same vector backbone without any promoter was used as background noise subtraction. Fold-change analysis was performed and maximum fold-change was recorded.

References

- Heinrich-Ramm, R. *et al.* Biological monitoring for exposure to volatile organic compounds (VOCs). *Pure Appl Chem* **72**, 385–436, doi: 10.1351/Pac200072030385 (2000).
- Vo, U. U. & Morris, M. P. Nonvolatile, semivolatile, or volatile: redefining volatile for volatile organic compounds. *J Air Waste Manag Assoc* **64**, 661–669 (2014).
- Ye, W., Won, D. Y. & Zhang, X. A preliminary ventilation rate determination methods study for residential buildings and offices based on VOC emission database. *Build Environ* **79**, 168–180, doi: 10.1016/j.buildenv.2014.05.009 (2014).
- Wu, Y. F., Jia, Y. Z. & Lu, X. W. Assessment of semi-volatile organic compounds in drinking water sources in Jiangsu, China. *Ecotox Environ Safe* **94**, 138–146, doi: 10.1016/j.ecoenv.2013.04.016 (2013).
- He, X. R. *et al.* Distribution, sources, and ecological risk assessment of SVOCs in surface sediments from Guan River Estuary, China. *Environ Monit Assess* **186**, 4001–4012, doi: 10.1007/s10661-014-3675-z (2014).
- Wheatley, R. E. The consequences of volatile organic compound mediated bacterial and fungal interactions. *Anton Leeuw Int J G* **81**, 357–364, doi: Doi 10.1023/A:1020592802234 (2002).
- Dicke, M., van Loon, J. J. A. & Soler, R. Chemical complexity of volatiles from plants induced by multiple attack. *Nat Chem Biol* **5**, 317–324, doi: Doi 10.1038/Nchembio.169 (2009).
- Quigley, C. J. & Corsi, R. L. Emissions of Vocs from a Municipal Sewer. *J Air Waste Manag* **45**, 395–403 (1995).
- Flick, B., Talsness, C. E., Jackh, R., Buesen, R. & Klug, S. Embryotoxic potential of N-methyl-pyrrolidone (NMP) and three of its metabolites using the rat whole embryo culture system. *Toxicology and applied pharmacology* **237**, 154–167, doi: 10.1016/j.taap.2009.02.024 (2009).
- Klimisch, H. J. & Hellwig, J. Developmental toxicity of dimethylacetamide in rabbits following inhalation exposure. *Hum Exp Toxicol* **19**, 676–683, doi: Doi 10.1191/096032700668432103 (2000).
- Panel, R. E. *et al.* A toxicologic and dermatologic assessment of cyclopentanones and cyclopentenones when used as fragrance ingredients. *Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association* **50** Suppl 3, S517–556, doi: 10.1016/j.fct.2012.04.019 (2012).
- Fournier, K., Glorennec, P. & Bonvallot, N. An exposure-based framework for grouping pollutants for a cumulative risk assessment approach: Case study of indoor semi-volatile organic compounds. *Environ Res* **130**, 20–28, doi: 10.1016/j.envres.2014.01.007 (2014).
- Barco-Bonilla, N. *et al.* Study of the distribution of 204 organic contaminants between the aqueous phase and the suspended particulate matter in treated wastewater for proper environmental control. *Desalin Water Treat* **51**, 2497–2515, doi: 10.1080/19443994.2013.765095 (2013).
- Barco-Bonilla, N., Romero-Gonzalez, R., Plaza-Bolanos, P., Martinez Vidal, J. L. & Garrido Frenich, A. Systematic study of the contamination of wastewater treatment plant effluents by organic priority compounds in Almeria province (SE Spain). *The Science of the total environment* **447**, 381–389, doi: 10.1016/j.scitotenv.2013.01.027 (2013).
- Estrada, J. M., Rodriguez, E., Quijano, G. & Munoz, R. Influence of gaseous VOC concentration on the diversity and biodegradation performance of microbial communities. *Bioproc Biosyst Eng* **35**, 1477–1488, doi: Doi 10.1007/S00449-012-0737-X (2012).
- Sun, D. Q. *et al.* Effect of volatile organic compounds absorbed to fresh biochar on survival of *Bacillus mucilaginosus* and structure of soil microbial communities (vol 15, pg 271, 2015). *J Soil Sediment* **15**, 490–490, doi: 10.1007/S11368-014-1046-6 (2015).
- Kabelitz, N., Santos, P. M. & Heipieper, H. J. Effect of aliphatic alcohols on growth and degree of saturation of membrane lipids in *Acinetobacter calcoaceticus*. *FEMS microbiology letters* **220**, 223–227 (2003).
- Pini, C. V., Bernal, P., Godoy, P., Ramos, J. L. & Segura, A. Cyclopropane fatty acids are involved in organic solvent tolerance but not in acid stress resistance in *Pseudomonas putida* DOT-T1E. *Microbial biotechnology* **2**, 253–261, doi: 10.1111/j.1751-7915.2009.00084.x (2009).
- Heipieper, H. J., Meinhardt, F. & Segura, A. The cis-trans isomerase of unsaturated fatty acids in *Pseudomonas* and *Vibrio*: biochemistry, molecular biology and physiological function of a unique stress adaptive mechanism. *FEMS microbiology letters* **229**, 1–7 (2003).
- Ramos, J. L. *et al.* Mechanisms of solvent tolerance in gram-negative bacteria. *Annual review of microbiology* **56**, 743–768, doi: 10.1146/annurev.micro.56.012302.161038 (2002).
- Nikaido, H. & Takatsuka, Y. Mechanisms of RND multidrug efflux pumps. *Biochimica et biophysica acta* **1794**, 769–781, doi: 10.1016/j.bbapap.2008.10.004 (2009).
- Segura, A., Duque, E., Mosqueda, G., Ramos, J. L. & Junker, F. Multiple responses of gram-negative bacteria to organic solvents. *Environmental microbiology* **1**, 191–198 (1999).
- Mayer-Blackwell, K. *et al.* Nanoliter qPCR platform for highly parallel, quantitative assessment of reductive dehalogenase genes and populations of dehalogenating microorganisms in complex environments. *Environmental science & technology* **48**, 9659–9667, doi: 10.1021/es500918w (2014).
- Di Gennaro, P. *et al.* Dynamic changes in bacterial community structure and in naphthalene dioxygenase expression in vermicompost-amended PAH-contaminated soils. *Journal of hazardous materials* **172**, 1464–1469, doi: 10.1016/j.jhazmat.2009.08.013 (2009).
- Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
- Bereza-Malcolm, L. T., Mann, G. & Franks, A. E. Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach. *ACS Synth Biol* **4**, 535–546, doi: 10.1021/sb500286r (2015).

27. Zaslaver, A. *et al.* A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nature methods* **3**, 623–628, doi: 10.1038/nmeth895 (2006).
28. Takatsuka, Y., Chen, C. & Nikaido, H. Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6559–6565, doi: 10.1073/pnas.1001460107 (2010).
29. Okochi, M., Kurimoto, M., Shimizu, K. & Honda, H. Increase of organic solvent tolerance by overexpression of manXYZ in *Escherichia coli*. *Applied microbiology and biotechnology* **73**, 1394–1399, doi: 10.1007/s00253-006-0624-y (2007).
30. Horinouchi, T. *et al.* Transcriptome analysis of parallel-evolved *Escherichia coli* strains under ethanol stress. *BMC genomics* **11**, 579, doi: 10.1186/1471-2164-11-579 (2010).
31. Oh, H. Y., Lee, J. O. & Kim, O. B. Increase of organic solvent tolerance of *Escherichia coli* by the deletion of two regulator genes, fadR and marR. *Applied microbiology and biotechnology* **96**, 1619–1627, doi: 10.1007/s00253-012-4463-8 (2012).
32. Shimizu, K., Hayashi, S., Doukyu, N., Kobayashi, T. & Honda, H. Time-course data analysis of gene expression profiles reveals purR regulon concerns in organic solvent tolerance in *Escherichia coli*. *Journal of bioscience and bioengineering* **99**, 72–74, doi: 10.1263/jbb.99.072 (2005).
33. Rutherford, B. J. *et al.* Functional genomic study of exogenous n-butanol stress in *Escherichia coli*. *Applied and environmental microbiology* **76**, 1935–1945, doi: 10.1128/AEM.02323-09 (2010).
34. Zingaro, K. A. & Terry Papoutsakis, E. GroESL overexpression imparts *Escherichia coli* tolerance to i-, n-, and 2-butanol, 1,2,4-butanetriol and ethanol with complex and unpredictable patterns. *Metabolic engineering* **15**, 196–205, doi: 10.1016/j.ymben.2012.07.009 (2013).
35. Zingaro, K. A. & Papoutsakis, E. T. Toward a semisynthetic stress response system to engineer microbial solvent tolerance. *mBio* **3**, doi: 10.1128/mBio.00308-12 (2012).
36. Reyes, L. H., Almario, M. P. & Kao, K. C. Genomic library screens for genes involved in n-butanol tolerance in *Escherichia coli*. *PLoS one* **6**, e17678, doi: 10.1371/journal.pone.0017678 (2011).
37. Wang, J. *et al.* Global metabolomic and network analysis of *Escherichia coli* responses to exogenous biofuels. *Journal of proteome research* **12**, 5302–5312, doi: 10.1021/pr400640u (2013).
38. Brynildsen, M. P. & Liao, J. C. An integrated network approach identifies the isobutanol response network of *Escherichia coli*. *Molecular systems biology* **5**, 277, doi: 10.1038/msb.2009.34 (2009).
39. Tuduri, L., Millet, M., Briand, O. & Montury, M. Passive air sampling of semi-volatile organic compounds. *Trac-Trend Anal Chem* **31**, 38–49, doi: 10.1016/J.Trac.2011.08.007 (2012).
40. Roche, B. *et al.* Reprint of: Iron/sulfur proteins biogenesis in prokaryotes: formation, regulation and diversity. *Biochimica et biophysica acta* **1827**, 923–937, doi: 10.1016/j.bbabi.2013.05.001 (2013).
41. Giel, J. L., Rodionov, D., Liu, M., Blattner, F. R. & Kiley, P. J. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O₂-regulated genes in *Escherichia coli*. *Molecular microbiology* **60**, 1058–1075, doi: 10.1111/j.1365-2958.2006.05160.x (2006).
42. Outten, F. W., Djaman, O. & Storz, G. A suf operon requirement for Fe-S cluster assembly during iron starvation in *Escherichia coli*. *Molecular microbiology* **52**, 861–872, doi: 10.1111/j.1365-2958.2004.04025.x (2004).
43. Lee, K. C., Yeo, W. S. & Roe, J. H. Oxidant-responsive induction of the suf operon, encoding a Fe-S assembly system, through Fur and IscR in *Escherichia coli*. *Journal of bacteriology* **190**, 8244–8247, doi: 10.1128/JB.01161-08 (2008).
44. Yeo, W. S., Lee, J. H., Lee, K. C. & Roe, J. H. IscR acts as an activator in response to oxidative stress for the suf operon encoding Fe-S assembly proteins. *Molecular microbiology* **61**, 206–218, doi: 10.1111/j.1365-2958.2006.05220.x (2006).
45. Dai, Y. & Outten, F. W. The *E. coli* SufS-SufE sulfur transfer system is more resistant to oxidative stress than IscS-IscU. *FEBS letters* **586**, 4016–4022, doi: 10.1016/j.febslet.2012.10.001 (2012).
46. Koh, Y. S. & Roe, J. H. Isolation of a novel paraquat-inducible (pqi) gene regulated by the soxRS locus in *Escherichia coli*. *Journal of bacteriology* **177**, 2673–2678 (1995).
47. Lee, J., Hiibel, S. R., Reardon, K. F. & Wood, T. K. Identification of stress-related proteins in *Escherichia coli* using the pollutant cis-dichloroethylene. *Journal of applied microbiology* **108**, 2088–2102, doi: 10.1111/j.1365-2672.2009.04611.x (2010).
48. Martin, J. E., Waters, L. S., Storz, G. & Imlay, J. A. The *Escherichia coli* small protein MntS and exporter MntP optimize the intracellular concentration of manganese. *PLoS genetics* **11**, e1004977, doi: 10.1371/journal.pgen.1004977 (2015).
49. Ezraty, B., Grimaud, R., El Hassouni, M., Moinier, D. & Barras, F. Methionine sulfoxide reductases protect Ffh from oxidative damages in *Escherichia coli*. *The EMBO journal* **23**, 1868–1877, doi: 10.1038/sj.emboj.7600172 (2004).
50. Amitai, S., Kolodkin-Gal, I., Hananya-Meltabashi, M., Sacher, A. & Engelberg-Kulka, H. *Escherichia coli* MazF leads to the simultaneous selective synthesis of both “death proteins” and “survival proteins”. *PLoS genetics* **5**, e1000390, doi: 10.1371/journal.pgen.1000390 (2009).
51. Han, X. *et al.* *Escherichia coli* genes that reduce the lethal effects of stress. *BMC microbiology* **10**, 35, doi: 10.1186/1471-2180-10-35 (2010).
52. Held, K. G. & Postle, K. ExbB and ExbD do not function independently in TonB-dependent energy transduction. *Journal of bacteriology* **184**, 5170–5173 (2002).
53. Larsen, R. A., Letain, T. E. & Postle, K. *In vivo* evidence of TonB shuttling between the cytoplasmic and outer membrane in *Escherichia coli*. *Molecular microbiology* **49**, 211–218 (2003).
54. Letain, T. E. & Postle, K. TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Molecular microbiology* **24**, 271–283 (1997).
55. McHugh, J. P. *et al.* Global iron-dependent gene regulation in *Escherichia coli*. A new mechanism for iron homeostasis. *The Journal of biological chemistry* **278**, 29478–29486, doi: 10.1074/jbc.M303381200 (2003).
56. Matzanke, B. F., Anemuller, S., Schunemann, V., Trautwein, A. X. & Hantke, K. PhuF, part of a siderophore-reductase system. *Biochemistry* **43**, 1386–1392, doi: 10.1021/bi0357661 (2004).
57. Anjem, A., Varghese, S. & Imlay, J. A. Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Molecular microbiology* **72**, 844–858, doi: 10.1111/j.1365-2958.2009.06699.x (2009).
58. Ferguson, G. P., Nikolaev, Y., McLaggan, D., Maclean, M. & Booth, I. R. Survival during exposure to the electrophilic reagent N-ethylmaleimide in *Escherichia coli*: role of KefB and KefC potassium channels. *Journal of bacteriology* **179**, 1007–1012 (1997).
59. Goodell, E. W. & Higgins, C. F. Uptake of cell wall peptides by *Salmonella typhimurium* and *Escherichia coli*. *Journal of bacteriology* **169**, 3861–3865 (1987).
60. Ito, A., Taniuchi, A., May, T., Kawata, K. & Okabe, S. Increased antibiotic resistance of *Escherichia coli* in mature biofilms. *Applied and environmental microbiology* **75**, 4093–4100, doi: 10.1128/AEM.02949-08 (2009).
61. Atsumi, S. *et al.* Evolution, genomic analysis, and reconstruction of isobutanol tolerance in *Escherichia coli*. *Molecular systems biology* **6**, 449, doi: 10.1038/msb.2010.98 (2010).
62. Zhou, J. *et al.* Identification of membrane proteins associated with phenylpropanoid tolerance and transport in *Escherichia coli* BL21. *Journal of proteomics* **113**, 15–28, doi: 10.1016/j.jprot.2014.09.012 (2015).
63. Shimizu, K. Regulation systems of bacteria such as *Escherichia coli* in response to nutrient limitation and environmental stresses. *Metabolites* **4**, 1–35 (2014).
64. Higashi, K. *et al.* Identification of a spermidine excretion protein complex (MdtJI) in *Escherichia coli*. *Journal of bacteriology* **190**, 872–878, doi: 10.1128/JB.01505-07 (2008).

65. Lomovskaya, O. & Lewis, K. Emr, an Escherichia coli locus for multidrug resistance. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 8938–8942 (1992).
66. Furukawa, H., Tsay, J. T., Jackowski, S., Takamura, Y. & Rock, C. O. Thiolactomycin resistance in Escherichia coli is associated with the multidrug resistance efflux pump encoded by emrAB. *Journal of bacteriology* **175**, 3723–3729 (1993).
67. Jennings, M. E. *et al.* Characterization of the Salmonella enterica serovar Typhimurium ydcI gene, which encodes a conserved DNA binding protein required for full acid stress resistance. *Journal of bacteriology* **193**, 2208–2217, doi: 10.1128/JB.01335-10 (2011).
68. Nachin, L., Nannmark, U. & Nyström, T. Differential roles of the universal stress proteins of Escherichia coli in oxidative stress resistance, adhesion, and motility. *Journal of bacteriology* **187**, 6265–6272, doi: 10.1128/JB.187.18.6265-6272.2005 (2005).
69. Gustavsson, N., Diez, A. & Nyström, T. The universal stress protein paralogs of Escherichia coli are co-ordinately regulated and co-operate in the defence against DNA damage. *Molecular microbiology* **43**, 107–117 (2002).
70. Hodges, A. P. *et al.* Bayesian network expansion identifies new ROS and biofilm regulators. *PLoS one* **5**, e9513, doi: 10.1371/journal.pone.0009513 (2010).
71. Gunasekera, T. S., Csonka, L. N. & Paliy, O. Genome-wide transcriptional responses of Escherichia coli K-12 to continuous osmotic and heat stresses. *Journal of bacteriology* **190**, 3712–3720, doi: 10.1128/JB.01990-07 (2008).
72. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8009–8014, doi: 10.1073/pnas.0903229106 (2009).
73. Nishino, K. & Yamaguchi, A. Analysis of a complete library of putative drug transporter genes in Escherichia coli. *Journal of bacteriology* **183**, 5803–5812, doi: 10.1128/JB.183.20.5803-5812.2001 (2001).
74. Koo, M. S. *et al.* A reducing system of the superoxide sensor SoxR in Escherichia coli. *The EMBO journal* **22**, 2614–2622, doi: 10.1093/emboj/cdg252 (2003).
75. Byrne, R. T., Chen, S. H., Wood, E. A., Cabot, E. L. & Cox, M. M. Escherichia coli genes and pathways involved in surviving extreme exposure to ionizing radiation. *Journal of bacteriology* **196**, 3534–3545, doi: 10.1128/JB.01589-14 (2014).
76. Davies, J. K. & Reeves, P. Genetics of resistance to colicins in Escherichia coli K-12: cross-resistance among colicins of group B. *Journal of bacteriology* **123**, 96–101 (1975).
77. Cascales, E., Gavioli, M., Sturgis, J. N. & Llobes, R. Proton motive force drives the interaction of the inner membrane TolA and outer membrane pal proteins in Escherichia coli. *Molecular microbiology* **38**, 904–915 (2000).
78. Tanabe, H. *et al.* Growth phase-dependent transcription of emrKY, a homolog of multidrug efflux emrAB genes of Escherichia coli, is induced by tetracycline. *The Journal of general and applied microbiology* **43**, 257–263 (1997).
79. Schafer, H., Myronova, N. & Boden, R. Microbial degradation of dimethylsulphide and related C1-sulphur compounds: organisms and pathways controlling fluxes of sulphur in the biosphere. *J Exp Bot* **61**, 315–334, doi: 10.1093/jxb/erp355 (2010).
80. Diaz, E., Ferrandez, A., Prieto, M. A. & Garcia, J. L. Biodegradation of aromatic compounds by Escherichia coli. *Microbiol Mol Biol Rev* **65**, 523–569, table of contents, doi: 10.1128/MMBR.65.4.523-569.2001 (2001).
81. Ramos, J. L. *et al.* Mechanisms of solvent resistance mediated by interplay of cellular factors in Pseudomonas putida. *FEMS microbiology reviews* **39**, 555–566, doi: 10.1093/femsre/fuv006 (2015).
82. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
83. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357–359, doi: 10.1038/nmeth.1923 (2012).
84. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079, doi: 10.1093/bioinformatics/btp352 (2009).
85. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLoS computational biology* **9**, e1003118, doi: 10.1371/journal.pcbi.1003118 (2013).
86. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140, doi: 10.1093/bioinformatics/btp616 (2010).
87. Robinson, M. D. & Smyth, G. K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**, 321–332, doi: 10.1093/biostatistics/kxm030 (2008).
88. Galperin, M. Y., Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic acids research* **43**, D261–269, doi: 10.1093/nar/gku1223 (2015).
89. Yamada, T., Letunic, I., Okuda, S., Kanehisa, M. & Bork, P. iPath2.0: interactive pathway explorer. *Nucleic acids research* **39**, W412–415, doi: 10.1093/nar/gkr313 (2011).

Acknowledgements

The authors would like to acknowledge financial support from PUB (Award No: M4340001.C70) and Singapore Centre on Environmental Life Sciences Engineering (SCELSE), whose research is supported by National Research Foundation Singapore, Ministry of Education, Nanyang Technological University and National University of Singapore, under its Research Centre of Excellence Programme. We also acknowledge Martin Tay for providing the revised mapping table of the COG's 2014 database.

Author Contributions

F.M.L., P.Y., L.G., J.H., T.S. and E.M. planned the experiments. L.G., P.Y. and A.M. performed laboratory experiments. P.Y. and E.A. performed the data analysis. P.Y., E.M. and F.M.L. wrote the paper. F.M.L., P.Y., L.G., J.H., T.S. and E.M. discussed the results and revised the manuscript.

Additional Information

Accession codes: Sequencing data have been submitted to GenBank SRA archive with the BioProject ID: PRJNA286974 and SRP accession SRP059483.

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Yung, P. Y. *et al.* Global transcriptomic responses of *Escherichia coli* K-12 to volatile organic compounds. *Sci. Rep.* **6**, 19899; doi: 10.1038/srep19899 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>