Inactivation of a virulent wastewater *Escherichia coli* and non-virulent commensal *Escherichia coli* DSM1103 strains and their gene expression upon solar irradiation

Nada Al-Jassim [1], David Mantilla-Calderon [1], Tiannyu Wang [1], Pei-Ying Hong * [1]

1. King Abdullah University of Science and Technology (KAUST), Water Desalination and Reuse Center (WDRC), Biological and Environmental Sciences & Engineering Division (BESE), Thuwal, 23955-6900, Saudi Arabia

* Corresponding author:
Pei-Ying Hong
Email: peiying.hong@kaust.edu.sa
Phone: +966-12-8082218

Keywords: ultraviolet disinfection; decay kinetics; antibiotic-resistant bacteria; New Delhi metallo-beta-lactamase; virulence; transcriptomics

Running title: Decay kinetics and gene expression profiles of *E. coli* strains upon solar irradiation

Abstract

This study examined the decay kinetics and molecular responses of two *E. coli* strains upon solar irradiation. The first is a virulent and antibiotic resistant *E. coli* PI-7, carrying the emerging NDM-1 antibiotic resistance gene that was isolated from wastewater. The other strain, *E. coli* DSM1103, displayed lower virulence and antibiotic resistance compared to *E. coli* PI7. In buffer solution, *E. coli* PI-7 displayed a longer lag-phase prior to decay and a longer half-life compared with *E. coli* DSM1103 (6.64 ± 0.63 h and 2.85 ± 0.46 min vs. 1.33 ± 0.52 h and 2.04 ± 0.36 min). In wastewater, both *E. coli* decayed slower compared to when in buffer. Although solar irradiation remained effective in reducing *E. coli* cell numbers by more than 5-$\log_{10}$ in < 24 h, comparative genomics and transcriptomics revealed differences in the genomes and overall regulation of genes between the two *E. coli*. A wider arsenal of genes related to oxidative stress, cellular repair and protective mechanisms were upregulated in *E. coli* PI-7. Subpopulations of the *E. coli* PI-7 expressed genes related to dormancy and persister cells formation during the late decay phase, which may have accounted for its prolonged persistence. Upon prolonged solar irradiance, both *E. coli* strains displayed upregulation of genes related to horizontal gene transfer and antibiotic resistance. Virulence functions unique to *E. coli* PI-7 were also upregulated.

1. Introduction

Untreated municipal wastewater and wastewater treatment plants, with their abundance of nutrients, high bacterial numbers and sub-lethal levels of antibiotics, are considered favourable for both the survival of antibiotic-resistant bacteria (ARB) and the horizontal transfer of bacterial resistance. Collectively, the wastewater is implicated as a key player in the creation of superbugs and the dissemination of these emerging contaminants [1, 2]. The release of antibiotic-resistant organisms through wastewater effluents has previously been reported [3-5]. In an investigation of the performance of a conventional wastewater treatment plant using activated sludge process and chlorination as the tertiary treatment step, it was found that 2.8% of 72 isolates that remained viable in the chlorinated effluent were susceptible to all 8 tested antibiotics, while 27.8% were resistant to 5 or more antibiotics [4]. A separate study also reported that antimicrobial-resistant *E. coli* was not effectively eliminated by the wastewater treatment process, and that *E. coli* that are resistant to cefotaxime, ciprofloxacin and cefoxitin were present in the final treated effluent [3].

Among the ARB increasingly being identified as present in wastewater are bacteria coding for carbapenemases, specifically the New Delhi metallo-β-lactamase (NDM). The spread of mobile carbapenemases among bacterial pathogens is of great concern, not only because these enzymes confer resistance to carbapenems and other β-lactam antibiotics, but also because of the evident ease of mobility of this resistance. NDM-harboring bacteria typically are broadly resistant to multiple other antibiotic classes, leaving very few treatment options available [6-9]. Although mainly detected in clinical settings, there have been numerous reports of community-acquired infections from NDM-harboring opportunistic pathogens [6, 10] as well as reports of the presence of $bla_{\text{NDM-1}}$ and NDM1-harboring bacteria in urban wastewater [7, 11]. For example,

68 *bla*$_{\text{NDM-1}}$ genes was consistently detected throughout the different treatment stages, including
69 chlorinated effluent, in two wastewater treatment plants (WWTPs) in northern China [11]. These
70 findings implicating wastewater as a source of NDM-harboring bacteria and *bla*$_{\text{NDM-1}}$ genes
71 exacerbate the potential risks associated with the reuse of wastewater. This is of particular
72 relevance in countries like Saudi Arabia where the pressure on water sources is high due to water
73 scarcity, and there are plans to employ treated wastewater to alleviate considerable portions of
74 this water-scarcity pressure [12]. In addition to the earlier studies, an *E. coli*, named *E. coli* PI-7,
75 that displays resistance to a wide range of antibiotics was isolated from the influent stream of a
76 WWTP in Jeddah, Saudi Arabia [13]. Upon further investigation, it was found that this isolate
77 carries the *bla*$_{\text{NDM-1}}$ gene on a plasmid that, based on sequence comparisons, is thought to have
78 been horizontally acquired from a *Klebsiella oxytoca* isolate [14]. Moreover, genomic islands
79 associated with pathogenicity have also been found in *E. coli* PI-7, suggesting that this strain is a
80 potential pathogen that remains viable in the sanitary infrastructure.
81 To shed light on to what extent the virulent and antibiotic resistant bacteria presence in
82 treated wastewater would pose as a risk to public health during reuse events, it is key to
83 understand the fate and persistence of this emerging microbial contaminant. An example would
84 be to understand the effect of solar irradiation as a biocidal barrier to dissemination of ARB in
85 the environment. Studies on the efficacy of solar photoinactivation on pathogenic waterborne
86 bacteria and pathogen indicators have reported variable findings. Many studies found rapid
87 inactivation of fecal indicator organisms within a few hours' exposure to natural sunlight, and it
88 is reported that all of the classically defined waterborne pathogenic bacteria have been found to
89 be readily amenable to 6 h of solar disinfection under suitable field conditions [15-19]. However,
90 numerous studies also report that in certain instances, naturally occurring fecal coliforms have

shown much slower inactivation rates, and indicator bacteria may still remain detectable after a full day's sunlight exposure [20-25]. In particular, sub-populations of *E. coli* were found to be light-resistant and persisted for a longer period of time upon solar irradiation [16].

Although these studies demonstrate the decay kinetics of pathogenic waterborne bacteria and pathogen indicators, information is lacking on the effect of solar photoinactivation on ARBs, their fate and persistence, as well as an understanding of their response at the molecular level. This study examines the decay kinetics of the emerging contaminant, a virulent antibiotic resistant (carrying the concerning $bla_{\text{NDM-1}}$ gene) wastewater *E. coli* PI-7, upon solar irradiation in buffer and wastewater matrix so as to provide understanding on its fate and persistence upon dissemination into the natural environment. A non-virulent and less antibiotic resistant commensal *E. coli* isolate DSM1103 was also examined for its decay kinetics upon solar irradiation. The gene expression profiles of both isolates, which possess different genomic content, over an increasing duration of exposure to simulated solar irradiance were mapped by transcriptomics. This study serves to elucidate protective genetic mechanisms expressed by both isolates against solar biocidal effect, and expressions of concerning traits such as those related to virulence, horizontal gene transfer and antibiotic resistance mechanisms.

2. Materials and methods

2.1. Bacterial isolates and genome characteristics

E. coli DSM1103 (also known as ATCC25922) and *E. coli* strain PI-7, isolated from the influent stream of a WWTP in Jeddah, Saudi Arabia [13] were studied for their inactivation kinetics upon solar irradiation. A summarized comparison of the genomes of both isolates is provided in Table S1[26, 27]. Briefly, *E. coli* PI-7 was resistant to a wide spectrum of antibiotics (e.g. gentamicin, ampicillin, kanamycin, ceftazidime and sulfamethoxazole-trimethoprim,

chloramphenicol and erythromycin, tetracycline and carbapenem) and carries the *bla*$_{\text{NDM-1}}$ gene on an IncF plasmid of 110 kbp in size. *E. coli* PI-7 was also identified to contain genomic islands associated with pathogenicity (e.g. colonization fimbriae antigen I, two intimin-like proteins associated with EHEC and EOEC pathogens, type III secretion system), and demonstrated *in-vitro* cell invasiveness. *E. coli* DSM1103 was originally isolated from a human clinical sample collected in Seattle and WA (1946). In contrast to the *E. coli* PI-7, genomic characterization of *E. coli* DSM1103 did not reveal as wide repertoire of virulence-associated traits and antibiotic resistance genes as *E. coli* PI-7 [28]. Furthermore, *E. coli* DSM1103 is a recommended quality control strain for disc susceptibility testing and demonstrates susceptibility to a wide variety of antibiotics (http://www.atcc.org/products/all/25922.aspx).

2.2. Wastewater collection and quality testing

Wastewater samples were collected from the WWTP in KAUST, Saudi Arabia. The WWTP is equipped with the following process units: (i) a grid mesh screen to remove bulky items in the incoming wastewater stream, (ii) a primary clarifier, (iii) anoxic-oxic activated sludge tank, (iv) an aerobic membrane bioreactor (MBR), and (v) a holding tank for chlorination. Twenty L of effluent was collected after the aerobic MBR, and the same volume of chlorinated effluent was collected from the holding tank. Both wastewater samples were vacuum-filtered through 0.2 µm polycarbonate membranes filters (GE Healthcare Life Sciences, Little Chalfont, Buckinghamshire, UK) on the day of the collection, A number of water quality parameters including pH, chemical oxygen demand (COD), total dissolved organic carbon (DOC), total nitrogen (TN), specific ultraviolet absorbance (SUVA) and alkalinity (reported as mg $CaCO_3$/L) and chlorine residual were measured as detailed in Supplementary Information 1.

2.3. Decay kinetics upon solar inactivation

Solar inactivation trials were conducted as described in Supplementary Information 2. The number of colony forming units (CFU)/mL in the experimental microcosms was quantified over time. The results were converted to $\log_{10}$ and natural log (ln) curves of the CFU/mL obtained at each sampling time divided by the initial CFU_o/mL (i.e., $\log(CFU/CFU_0)$ or $\ln(CFU/CFU_0)$), and plotted against time to obtain the inactivation curves. Slopes (*k*) of dark control and irradiated sample inactivation curves were calculated from the ln curves, with the following to note: in cases where the length of lag prior to decay phase was longer than 2 h, the curve was considered to be bi-modal and the lag-phase data points were excluded from slope calculation. The attenuation of light penetration into microcosms due to sample turbidity was mathematically corrected for as described previously [29, 30]. Briefly, prior to each simulated solar inactivation trial, a sample from each microcosm was taken, and its absorbance at 280-700 nm was measured using UV-3600 UV-Vis spectrometer (Shimadzu, Kyoto, Japan). The readings were used to generate correction factors that were applied to slopes of decay curves prior to half-life calculations and statistical comparisons, which were conducted as described in Supplementary Information 3. Half-lives of the *E. coli*, or durations needed to reduce concentration by half, were calculated using first-order decay kinetics equation:

$$\ln(CFU_t/CFU_0) = -k^*t$$

Where k^* is the corrected slope of the inactivation curve, and t is time.

2.4. RNA-seq sampling and analysis

RNA-Seq was performed to provide comparative analysis of gene expression. Samples were taken at five points for *E. coli* PI-7 to capture the gene expression at different phases of the bacterial CFU inactivation curves, namely (i) prior to solar irradiation (0 h), and during the (ii) mid-lag phase (3 h), as well as (iii) early- (6.5 h), (iv) mid- (9 h) and (v) late-decay (14.5 h)

phases. A similar approach was performed for *E. coli* DSM1103 at four sampling points, capturing the gene expression of cultures (i) prior to solar irradiation (0 h), and during the (ii) early- (2h), (iii) mid- (3.5 h) and (iv) late-decay (6.5 h) phase. Although the exposure duration is different for the two strains, the main intention was to capture the expression profiles at different stages of each isolate's inactivation based on viable cell numbers in accordance to the decay kinetics observed in Figure 1. For each *E. coli* strain, two solar inactivation experiments were independently conducted for RNA sampling to obtain two biological replicates and two technical replicates at each sampling point. RNA preservation, extraction and sequencing were performed as described in Supplementary Information 4. The RNA-seq data was analyzed using CLC Genomics Workbench version 8.0.1 from CLC Bio (Cambridge, MA), as described in Supplementary Information 5 [31]. All transcriptomic sequences associated with *E. coli* PI-7 and *E. coli* DSM1103 are available on the ENA SRA public depository under accession numbers PRJEB13897 and PRJEB13888, respectively. Summarized RNA-seq mapping results for both strains are shown in Table S2 and S3, and Supplementary Information 6. Confirmation of RNA-seq results using reverse-transcription quantitative PCR (RT-qPCR) was performed as described in Supplementary Information 7 and 8, and Table S4.

2.5. Solar inactivation in presence of efflux pump inhibitor

Transcriptomic results showed upregulation of efflux pump genes in *E. coli* PI-7 and DSM1103. To verify the role of efflux pumps on the decay kinetics of the two *E. coli* upon solar irradiance, and as additional and non-molecular method of confirming transcriptomic results, inactivation trials were performed on *E. coli* PI-7 and DSM1103 in presence or absence of Phe-Arg-β-naphthylamide (PAβN), a known RND-type (Resistance-Nodulation-Division) efflux pump inhibitor. The bacteria was grown and prepared for inactivation trials in similar way as

detailed in Supplementary Information 2. PAβN was then added to the microcosms at 50 µg/mL or 0 µg/mL prior to simulated solar inactivation trials or as dark controls. The concentration of PaβN (i.e., 50 µg/mL) was decided based on preliminary trials, which determined no apparent change in the turbidity of overnight *E. coli* PI-7 cultures in the presence of PaβN alone, suggesting PaβN imposes no toxicity on *E. coli* PI-7 at this concentration. However, lower turbidity was observed when *E. coli* PI-7 was grown in presence of PaβN and 50 µg/mL tetracycline (a concentration lower than its minimum inhibitory concentration of 64 µg/mL[10]), suggesting a possible impact on the efflux pump mechanism.

3. Results

3.1. Inactivation kinetics of *E. coli* PI-7 and DSM1103 in different water matrices

The inactivation response of the bacterial isolates under simulated solar irradiation in buffer solution was examined (Figure 1A-C). *E. coli* PI-7 exhibited an initial lag phase of 6.64 ± 0.63 h, and achieved 5-$\log_{10}$ inactivation after > 12 h of sunlight irradiation (Figure 1A and C). In contrast, *E. coli* DSM1103 exhibited a significantly shorter lag phase of 1.33 ± 0.52 h (p < 0.0001), and achieved 5-$\log_{10}$ inactivation within 5 h of sunlight irradiation (Figure 1B and C). Half-life values derived from corrected decay constants showed that *E. coli* PI-7 had a half-life of 2.85 ± 0.46 min, which was significantly longer than the half-life of *E. coli* DSM1103 (2.04 ± 0.36 min, p < 0.0001) (Figure 1C). Additionally, for *E. coli* PI-7, a tailing effect starting at 11 h was observed in 35% of the trials. The half-life calculated from this portion of the decay curve for trials that displayed it was 7.76 ± 1.27 min. No such effect was detected in DSM1103 inactivation trials.

Inactivation of the bacterial isolates under simulated solar irradiation in real wastewater effluent and chlorinated effluent was also examined (Figure 2A-D). Both types of treated wastewater had DOC, SUVA and alkalinity that ranged from 2.2- 3.1 mg/L, 1.9-4.4 L.mg-M and 69.3-82.9 mg $CaCO_3$/L, respectively (Table S5). No significant change in lag phase lengths was observed for ether isolate in either type of treated wastewater compared to the buffer ($p > 0.05$). *E. coli* PI-7 was observed to exhibit mean lag phase lengths of 6.64 ± 0.48 h and 6.30 ± 1.18 h in effluent and chlorinated effluent, respectively (Figure 2A and B). *E. coli* DSM1103 exhibited mean lag phase of 1.29 ± 0.49 h and 0.71 ± 0.76 h in effluent and chlorinated effluent, respectively (Figure 2C and D). After the lag phase, *E. coli* PI-7 exhibited a mean half-life of 4.36 ± 1.13 min and 5.23 ± 1.40 min in effluent and chlorinated effluent, respectively (Table S6). The half-life of *E. coli* PI-7 in chlorinated effluent was significantly longer than the half-life measured for *E. coli* PI-7 in buffer ($p = 0.0167$). For *E. coli* DSM1103, the half-life value in effluent was 4.03 ± 0.36 and 2.05 ± 0.16 min in chlorinated effluent (Table S6). The half-life in the effluent was significantly longer than in buffer and chlorinated effluent ($p = 0.0012$).

3.2. Extent of differential gene expression response throughout the different phases of solar irradiation

Prior to transcriptomic analysis, general genomic comparisons were performed to provide insight into the differences between the two isolates' genomes. The genome of *E. coli* DSM1103 (5.2 Mb) was found to be larger than the available genome of PI-7 (4.73 Mb) (Table S1). Based on the current available mapped genomes, *E. coli* DSM1103 exhibited a larger total number of annotated genes than PI-7 (4,808 vs. 4,471, respectively), as well as more unique gene sequences and annotated gene functions (910 vs. 489, and 185 vs. 155, respectively). Annotated functions

were further broken down per functional category (Table S7) to illustrate the core and unique genes for both bacterial strains.

RNA-seq was performed to examine the molecular responses of the two isolates to solar irradiation. Gene expression for each of the isolates was analysed independently by comparing irradiated sample gene response to its corresponding dark control at each time point. Results revealed increasing numbers of genes significantly ($p < 0.05$) upregulated or downregulated by ≥ 2-fold in both *E. coli* PI-7 and DSM1103 as the duration of solar irradiation increased (Figure S1). The results also showed an increase in the percentage of total genes that were upregulated during the mid-decay and late-decay phase.

A closer examination showed that *E. coli* PI-7 upregulated > 50 stress response and protection mechanism functions, > 130 cell repair, division and cell wall synthesis functions, and >50 virulence functions (virulence factors, motility, membrane transport systems) upon prolonged exposure to solar irradiation. *E. coli* DSM1103 also upregulated genes from the same categories. *E. coli* DSM1103 showed upregulation of ten functions related to stress response, 22 cell repair functions, and 25 virulence functions. Further elaborations on the differential gene expressions of each of these functional categories are detailed in sections 3.3 to 3.5.

3.3. Differential gene expression of stress response and protection mechanisms

Stress response and protection mechanisms (e.g. oxidative stress response, heat shock proteins and reactive oxidative species (ROS) scavengers) can be differentially expressed by either of the *E. coli* strains to counteract against the effects of solar irradiation.

In *E. coli* PI-7, the redox-sensitive transcriptional regulator gene *qor*R was downregulated at the early-decay phase (Table 1). Various genes related to oxidative stress

functions were upregulated at the decay phases, including a catalase, two peroxidases and three glutathione-related functions, all known to play a role in ROS scavenging, as well as 39 other stress/detox related functions (Table 2 and Table S8A:1-1.5). A superoxide dismutase [Cu-Zn] precursor gene in *E. coli* PI-7 was downregulated from early-decay to late-decay (Table S8A:1.5). In *E. coli* DSM1103, no response in regulatory oxidative stress genes was detected (Table 1). DSM1103 still activated genes in similar oxidative stress response categories as *E. coli* PI-7: a peroxidase, one glutathione-related functions and six other stress/detox functions. Despite DSM1103 possessing similar numbers of catalase, peroxidase and glutathione-related genes, and more than twice as many other stress/detox related functions compared with PI-7, the total number of upregulated genes for *E. coli* DSM1103 in these categories was lower than that observed for PI-7 (Table 2 and Table S8B:1-1.5).

Other mechanisms of protection and detoxification were also examined. A regulatory gene involved in biofilm formation was upregulated in both *E. coli* PI-7 and DSM1103 (Table 1). The *mar*A regulatory gene, a multiple antibiotic resistance protein and a central regulator for efflux pumps, were upregulated in *E. coli* PI-7 and were present but not upregulated in DSM1103 (Table 1). Twenty-two drug and heavy metal efflux pumps and transporters were correspondingly upregulated in *E. coli* PI-7; only one was upregulated in DSM1103 (Table 2). Considering this result and the potential of efflux pumps to contribute to antibiotic resistance, and as an additional and non-molecular method of confirming transcriptomic results, the effect of efflux pumps in facilitating *E. coli* PI-7's prolonged persistence was tested by performing simulated solar inactivation trials in presence of the RND-type efflux pump inhibitor PAβN. Results showed that with PAβN, lag phase length was significantly reduced to less than three hours (p-value = 0.0010) while half-life length was reduced to 1.77 ± 0.06 min (p = 0.0360)

(Figure S2). Five-$\log_{10}$ reduction in CFU numbers was achieved in under 6.5 h and 10-$\log_{10}$ of inactivation was achieved by 12 h. *E. coli* DSM1103 was also tested in the presence of PaβN. While lag phase length did not change significantly, half-life of *E. coli* DSM1103 significantly decreased to 1.38 ± 0.11 min (p = 0.0446) (Figure S3), suggesting a role in the efflux pump mechanisms for *E. coli* DSM1103 against solar irradiation, albeit its role may be less significant than that in *E. coli* PI-7.

3.4. Differential gene expression of cell repair and division mechanisms

Solar irradiance inactivates microorganisms by inflicting direct and indirect damage on various cellular components, compromising DNA and cellular membrane integrity. Therefore, gene expressions related to cellular repair and division categories were further examined to understand the mechanisms employed by both *E. coli* strains to counteract the solar inactivation effects.

In *E. coli* PI-7, the *rec*A essential DNA repair and maintenance protein gene was upregulated and the *lex*A SOS-response repressor gene was downregulated (Table 1). This response is required to activate the SOS-cellular and DNA repair network. Twenty-eight out of 66 DNA repair functions, such as three different DNA-damage inducible protein genes, various DNA repair genes, exonuclease and endonuclease genes, were upregulated in PI-7 (Table 2). Fifteen out of 21 genes with other DNA-related functions, such as a probable DNA recombinase and a DNA helicase *ruv*AB gene (Table S8A:2.2), were also upregulated in PI-7 (Table 2). *E. coli* DSM1103 downregulated its *lex*A gene but showed no differential expression of the *recA* gene. The observed response for DSM1103 out of 78 DNA repair and 32 other DNA-related

296 genes was limited to the upregulation of two endonuclease genes (Table S8B:2.3), and one
297 function related to DNA recombination (Table S8B:2.2).

298 Functions related to cell wall synthesis, cell division and DNA synthesis were examined.
299 In *E.coli* PI-7, over eighty cell wall synthesis functions were upregulated at different stages of
300 exposure to simulated solar irradiation (Table 2). Of those, 28 functions were related to capsular
301 and extracellular polysaccharides (Table S8A:5.2). Cell division gene *fts*I, which codes for a
302 peptidoglycan synthase and belongs to both the cell division and the cell wall synthesis
303 categories, was upregulated from mid-lag to mid-decay phases (Table S8A:3 and S8A:5.1). In
304 addition, replication regulatory protein *rep*A2 gene was downregulated in PI-7 (Table 1), and
305 nine DNA replication/synthesis functions (e.g. DNA gyrases, topoisomerases, helicases) were
306 also upregulated (Table 2). As for *E. coli* DSM1103, seven cell wall function genes, of which
307 five were related to capsular and extracellular polysaccharides, were upregulated from mid-decay
308 to late-decay phases (Table 2). In the cell division and replication category, one function related
309 to a septum formation protein gene and another putative replication protein, were upregulated
310 (Table 2).

311 Cells can also adopt dormancy or formation of persister cells in the event that their repair
312 mechanisms are not sufficient to counteract against oxidative stress. *E. coli* PI-7 upregulated 19
313 out of 57 gene with functions related to dormancy and persister cell formation (Table 2). One
314 possible way to achieve dormancy can be through programmed cell death and toxin-antitoxin
315 systems. The genes for the YoeB-YefM pair, the only complementary toxin-antitoxin pair
316 detected for *E. coli* PI-7 in this experiment, were upregulated at late-decay (Table S8A:6.1).
317 Furthermore, the *hipB* dormancy and persister-cell related gene in *E. coli* PI-7 was upregulated at
318 mid-decay and late-decay phase (Table S8A:4), as well as other transcription factors, heat/cold

shock, and SOS functions that have been implicated in persister formation. As for *E. coli* DSM1103, seven out of 44 dormancy and persister cell formation genes were upregulated (Table 2, and Table S8B:4). Five toxin-antitoxin functions (e.g. antitoxin YefM, three toxin Ldr functions) were upregulated at mid-decay and late-decay phases (Table S8B:6.1), but the available annotation did not reveal if there were any complementary toxin-antitoxin pairs.

3.5. Differential gene expression of virulence functions

Virulence-related functions like virulence-signaling, adhesion and invasion functions were examined for their gene expression response to simulated solar irradiation, particularly for the purpose of understanding the risk associated with *E. coli* PI-7's virulence and prolonged persistence. *E. coli* PI-7 showed upregulation of the *rpo*N regulatory gene, encoding a sigma factor known to play a role in bacterial virulence (Table 1). It also upregulated 25 virulence functions including four cell signaling functions related to virulence at mid-decay phase (Table S8A:6.3-7.3), and a virulence protein gene that reached 20-fold upregulation at high RPKM levels (Table S8A:7, and Table S9A:7). *E. coli* PI-7 also upregulated two flagellar regulatory functions, and 29 functions related to flagella and motility (Table 2). For *E. coli* DSM1103, there was a detectable virulence response in the upregulation of six virulence-related functions (Table 2, and Table S8B:6.3-7.3).

An earlier study to characterize the genomic and plasmidic content of *E. coli* PI-7 revealed a repertoire of pathogenic traits including colonization fimbriae antigen I (CFA/I), accessory colonization factor *acf*D precursor, type III secretion system and various putative virulence factors[13]. These virulence-associated traits are unique to *E. coli* PI-7 and not present in the *E. coli* DSM1103 genome. Both CFA/I fimbrial subunit genes were upregulated upon solar

342 irradiation (Table 1), with one upregulated at mid-decay and the other throughout the decay
343 phases, as well as 25 other fimbriae-related functions (Table 2). Accessory colonization factor
344 *acf*D precursor gene was upregulated at late-decay alongside other adhesion functions. For
345 example, the *fim*H adhesin gene was upregulated at late-decay, and four other adhesion functions
346 were upregulated at various decay phases (Table S8A:7.1). The four functions that belong to type
347 III secretion systems were also upregulated, two at mid-decay and another two at late-decay
348 phases (Table 2, and Table S8A:9.5). Moreover, upregulation of functions related to the
349 conjugative transfer of *E. coli* PI-7's IncF $bla_{\text{NDM-1}}$-carrying plasmid was also observed (Table
350 S8A:9.3), as well as over 30 functions belonging to secretion systems II, VI, VII and VIII (Table
351 S8A:9.4-9.8). Although *E.coli* DSM1103 has no functions belonging to the type III secretion
352 system, it was observed that genes related to the type VI and VII secretion systems were
353 upregulated at late-decay phases (Table S8B:9.6-9.7).
354
355 **3.6. RNA-seq confirmation using reverse transcription quantitative PCR (RT-qPCR)**
356 To validate RNA-seq results, RT-qPCR was performed on RNA from early-decay and
357 mid-decay phases for each of the two *E. coli*. For eight out of ten data points obtained for *E. coli*
358 PI-7, fold-change in gene expression from RNA-seq and RT-qPCR were in good agreement
359 (Figure S4A-E). This is with the slight exception in two instances. For example, RNA-seq results
360 showed significant upregulation for gene *hfl*C at both early and mid-decay phases, but only slight
361 upregulation for the same gene was observed in mid-decay when examined using RT-qPCR
362 (Figure S4B). For the second exception, RT-qPCR showed downregulation of the gene *mar*A at
363 mid-decay, while RNA-seq showed upregulation (Figure S4D). However, the upregulation in the
364 RNA-seq result was not statistically significant. For *E. coli* DSM1103, four out of eight data

points were in good agreement between RT-qPCR and RNA-seq results (Figure S5B-D). RT-qPCR consistently showed downregulation of the tested genes while RNA-seq showed upregulation of these genes. However, it is to be noted that most of these upregulation obtained by RNA-seq were not of significant fold-change values except for gene *fts*I at mid-decay (Figure 5A).

4. Discussion

Treated municipal wastewater has been increasingly recognized as a potential source for disseminating antibiotic-resistant bacteria (ARB) into the environment [32, 33]. Numerous studies have reported that, despite having undergone the entire treatment process, there was an increase in the relative proportion of ARB compared to the other antibiotic-susceptible bacteria in the wastewater, suggesting that ARB survives the treatment process better than the non-resistant bacteria [3, 4, 34, 35]. The presence of antibiotic-resistant bacteria in the treated effluent can be of a concern if the treated wastewater were to be reused for agricultural irrigation. In a 2013 report by the US Center of Disease Control and Prevention (CDC), it was estimated that more than two million people in the United States are sickened every year with antibiotic-resistant infections, with at least 23,000 dying as a result [36]. The CDC further accounts the major factor in the growth of antibiotic resistance to be due to dissemination of ARB from person to person, or from the non-human sources in the environment, including food. This includes the potential dissemination of ARB into environment through wastewater reuse.

To mitigate concerns arising from ARB in treated wastewater, sunlight irradiation can be utilized as one of the final barriers to prevent dissemination of contaminants into the environment during agricultural irrigation events. This is because the full spectrum of sunlight

388 includes radiation in the UV-A and UV-B spectrums, which are both known to impose biocidal
389 effect on microorganisms by causing inactivation through direct and indirect mechanisms [37-39]. In
390 direct mechanism, UV radiation penetrates cell walls of microorganisms and is absorbed by
391 nucleic acids to result in dimerization of adjacent cytosines and thymines [40]. In indirect
392 mechanisms, endogenous or exogenous photosensitizers absorb UV light [41-43] and generate
393 strong reactive oxidative species (ROS) that can damage the cellular membrane, nucleic acids,
394 proteins or other cellular materials [43, 44].

395 In this study, exposure to simulated solar irradiance achieved $\geq$ 5-$\log_{10}$ reduction in the
396 numbers of viable *E. coli* cells, but the response curves and irradiance doses required for the
397 reduction differed between both isolates. Specifically, *E. coli* PI-7 exhibited an extended lag-
398 phase before the onset of detectable inactivation while *E. coli* DSM1103 had a shorter lag-phase
399 and decay half-life than that of *E. coli* PI-7, in addition to a tailing zone at late-decay with
400 significantly slower inactivation in 35% of PI-7 inactivation trials. There are limited studies that
401 demonstrate the differences in the inactivation kinetics of different strains of *E. coli* upon solar
402 irradiation. However, our observations are in agreement with the few studies available that
403 compared the differences in the inactivation kinetics of strains within other bacterial species,
404 where it was found that different strains of *Acinetobacter johnsonii* and *Halomonas salina*
405 exhibited variations in their survival rates after UV-B irradiation [45, 46]. Variations in the decay
406 kinetics among different strains of the same bacterial species may be attributable to factors such
407 as genome size, nucleotide base compositions and genomic content [47]. In this study, the two *E.*
408 *coli* isolates tested have different genomic content and, therefore, may have differences in their
409 molecular responses to the simulated solar irradiance. To investigate the underlying mechanism

410 of each isolate's response to simulated solar irradiance, gene expression of irradiated samples
411 was compared with dark controls at the various points of the inactivation curve for each *E. coli*.

412 The transcriptomic approach results showed that a large portion of the genes that were
413 differentially regulated by *E. coli* PI-7 was shared between genomes of the two *E. coli* strains.
414 Furthermore, despite *E. coli* DSM1103 possessing the larger genome size and higher number of
415 unique gene sequences and metabolic functions, *E. coli* PI-7 differentially regulated a larger
416 number of its available annotated genes compared with DSM1103, and also exhibited a higher
417 percentage of gene upregulation. A closer examination of specific gene categories revealed that
418 *E. coli* PI-7 upregulated a large number of DNA recombination and repair. This suggests that *E.*
419 *coli* PI-7's ability to activate a wide arsenal of DNA modification and DNA-damage repair
420 functions that are key factors contributing to the prolonged persistence of this isolate under solar
421 irradiance. DNA recombination and repair mechanisms are inducible as part of the SOS regulon
422 [48], which were observed in both *E. coli* strains to be downregulated for the associated SOS-
423 response repressor gene *lex*A. In addition, a large number of cell wall functions, and specifically
424 ones related to capsule and extracellular polysaccharides, was upregulated in *E. coli* PI-7. The
425 cell membrane is considered to be one of the key targets of damage induced by solar irradiance,
426 as it is the first point of contact with exogenous photosensitizers. Active synthesis of cell wall
427 components is indicative of continuous damage repair in *E. coli* PI-7. Production of extracellular
428 polysaccharides can also encapsulate the cells and provide a layer of protective shield against
429 solar irradiance [49, 50]. Similarly, both *E. coli* strains upregulated a number of biofilm formation
430 related genes, and the biofilm matrix can serve as physical shield against solar irradiance.

431 In addition to active cell repair, the longer persistence of *E. coli* PI-7 compared to *E. coli*
432 DSM1103 can be accounted for by the upregulation of oxidative stress response and protective

mechanisms. Interestingly, superoxide dismutase, which is known to convert superoxide O_2^- into less damaging oxygen and hydrogen peroxide [51], was downregulated in both *E. coli* strains, suggesting that the superoxide O_2^- ROS was not a key mechanism of damage in this study. Instead, genes implicated in the response to hydrogen peroxide radicals were upregulated in *E. coli* PI-7. Catalase is known to be a hydrogen peroxide scavenger that converts hydrogen peroxide to water and oxygen [52-54]. One catalase/peroxidase gene was upregulated during the decay phases in *E. coli* PI-7, but another was downregulated at mid-lag and early-decay. Although catalase scavenges hydrogen peroxide and protects bacterial cells against UV damage, excess catalase can act as a photosensitizer and cause increased sensitivity to UV damage when intracellular antioxidant moieties are limited [43, 55, 56]. Hence, a tight regulation of the catalase production may be undertaken by the *E. coli* PI-7 which, in turn, explains for the mixed response and upregulation of different catalase genes in varying decay phases. Cytochrome c peroxidase, another enzyme known to reduce hydrogen peroxide to water [57], was also upregulated in *E. coli* PI-7. Although no measurements were performed to determine the ROS abundance, these findings provide inferences to suggest that hydrogen peroxide may be the predominant ROS resulting in solar inactivation in this study. No evidence for specific hydrogen peroxide response was observed in *E. coli* DSM1103 and this may have attributed to the faster inactivation kinetics displayed by *E. coli* DSM1103. Our results are in good agreement with a microarray study on *Enterococcus faecalis*'s response to natural sunlight. Despite being a gram-positive bacteria, *E. faecalis* also demonstrated upregulation in DNA repair and replication, peptidoglycan and cell wall, cell redox homeostasis and oxidative stress functions (including thioredoxin and glutathione-related ones) upon solar irradiation. However, upregulation of superoxide dismutase was observed in *E. faecalis*, suggesting that superoxide O_2^- ROS may be implicated as a key

mechanism of cell damage against this bacterium [58]. This is in contrast to our findings obtained from both *E. coli* strains.

In addition to oxidative stress response and ROS scavenging, *E. coli* PI-7 activated a large number of multidrug efflux pumps and transporters, suggesting that cellular detoxification might also play a role in the response to solar irradiance. Conducting simulated solar inactivation trials on *E. coli* PI-7 in the presence of efflux pump inhibitor PAβN showed a great decrease in PI-7's persistence as it significantly decreased lag phase and half-life lengths, and achieved higher overall log-reduction in CFU numbers during a shorter duration of time. *E. coli* DSM1103 upregulated a single efflux pump functions, and therefore while the effect of PAβN was less extreme on *E. coli* DSM1103, it also experienced decreased half-life. Although efflux pumps' effects on inactivation have not been reported in solar irradiation studies, similar findings have been reported among microbial communities that were exposed to chlorination. Similar to the indirect inactivation mechanism of solar irradiation, chlorination results in oxidative stress. A study assessing the effect of chlorination found induced transcription of genes encoding major virulence factors in *S. aureus* [59] and chlorine-associated induction of antibiotic resistance in eight antibiotic-resistant *A. baumannii* pathogen isolates [60]. Tetracycline resistance conferred by means of efflux pump mechanism and encoded for by *tet*Z gene has also been documented to be enriched in bacterial DNA recovered from chlorinated effluent samples [4].

Besides active repair mechanisms and stress response and detoxification, entering into dormancy or formation of persister cells are documented responses to stress [61]. This is a particularly useful defense mechanism in the event that active cell repair remains insufficient to mitigate the UV-induced damage. Evidence suggesting formation of dormant or persister cells was observed during the late phase of decay where *E. coli* PI-7 exhibited an upregulation in

toxin-antitoxin systems for programmed cell death, and a number of dormancy and persister-cell related gene. These results may account for the tailing effect seen at the end of *E. coli* PI-7's inactivation curve, suggesting that a subpopulation of the *E. coli* cells entered into dormancy and persist indefinitely despite solar irradiation. The persistence of *E. coli* PI-7 subpopulations may be of a concern since *E. coli* PI-7 showed upregulation of virulence functions, a suite of genes involved in the replication and conjugative transfer of its NDM-1-carrying IncF plasmid and the CTX-M-15-carrying IncA plasmid.

The large number of virulence genes that were upregulated in *E. coli* PI-7 draws parallels with reported observations of pathogens activating virulence genes to evade the host's immune response during infection[62]. When pathogenic strains invade host, the immune response triggers production of ROS to induce cell apoptosis, which approximates the situation experienced by bacterial cells during solar irradiation[62]. It may be possible that the upregulation of virulent traits in *E. coli* PI-7 (a pathogenic strain) is a consequence of solar irradiation, and a possible interpretation would be that the virulence traits contribute to defense against solar irradiation or oxidative stress response. However, this hypothesized role of virulence genes in slowing decay kinetics remains to be further elucidated.

Inactivation of the two *E. coli* in wastewater was tested to examine the response in more realistic matrices. It was found that wastewater, compared with buffer solution, caused significant increase in the half-life of both *E. coli* strains. Specifically, the half-life of *E. coli* PI-7 was significantly longer in chlorinated wastewater treatment plant effluent, while the half-life for *E. coli* DSM1103 was significantly longer in non-chlorinated effluent. The slower decay experienced by both *E. coli* strains may be due to the high alkalinity present in the wastewater matrix (Table S5). Bicarbonates HCO_3^- present in water can react with hydroxyl radicals

producing $CO_3^{\bullet-}$, which has a slower reaction with organic molecules when compared to •O [19, 63]. On the other hand, the high SUVA content in both wastewater indicates the presence of dissolved organic matter, which may have shielded the microbial cells from sunlight irradiation or have cleaved into smaller, biologically labile, organic compounds that can in turn stimulate further bacterial growth [64, 65]. The longer half-life of both *E. coli* strains in wastewater matrix was observed even though the examined wastewater samples were filtered to remove turbidity. In actual instances of wastewater with high turbidity, the scattering effect of particulates would likely result in an even longer half-life than that observed in this study. Furthermore, in the case of *E. coli* PI-7, its longer half-life in the chlorinated effluent could be in part related to its upregulation of efflux pump and antibiotic resistance functions, which helped maintain homeostasis, not only against chlorine but also against ROS that may be present. Induction of these genes has also been reported in other bacterial species upon exposure to chlorine [59, 60], while an increase in the abundance of antibiotic resistance genes associated with efflux pump mechanisms was determined in the total microbial community recovered from chlorinated effluent compared to pre-chlorinated effluent [4]. These observations suggest that these traits may act as defense mechanisms against chlorine and other deleterious elements, hence resulting in a longer persistence of *E. coli* PI-7 in the chlorinated effluent.

Our findings collectively indicate that there is a need to place heightened levels of concern on dissemination risks arising from the ARB that may still remain viable in the treated wastewater.

5. Conclusion

This study revealed differences in the solar inactivation kinetics of a virulent antibiotic-resistant *E. coli* PI-7 strain, and a non-virulent and less antibiotic-resistant commensal *E. coli* DSM1103. The results showed that a minimum of half a day exposure to solar irradiation is required to overcome the lag phase of *E. coli* in the buffer solution and treated wastewater prior to achieving inactivation. Although solar irradiation remains useful and effective in reducing the number of *E. coli* cells by more than 5-$\log_{10}$, transcriptomics revealed differences in the overall upregulation of regulatory, protective and repair mechanisms between both *E. coli* strains, leading to different efficacies in damage sensing and response. Subpopulations of the *E. coli* PI-7 expressed genes related to dormancy and persister cells formation during the late-decay phase, which may have accounted for the prolonged persistence of *E. coli* PI-7. More importantly, this study showed that both *E. coli* strains displayed upregulation of virulence functions, with a wider arsenal of such genes being upregulated in the virulent *E. coli* PI-7, upon solar irradiation. The current guidance on the quality pertaining to treated wastewater for use in agricultural irrigation remains limited to assessing the abundance of fecal coliforms. Our findings suggest a need to expand efforts to monitor for the presence and abundance of ARB, particularly for those bacterial pathogens that are commonly associated with foodborne diarrheal diseases, in the treated wastewater that is to be reused.

Acknowledgements

This study is supported by KAUST baseline funding BAS/1/1033-01-01 awarded to P.-Y. Hong. The authors would to thank Dr Shengkun Dong and Professor Helen Nguyen for providing training to N.A-J on solar simulator in the earlier phase of this project; Dr Inhyuk Kwon and Professor Roderick Mackie for providing training to N.A-J on transcriptomics analysis; and Dr

Muhammad Raihan Jumat for providing assistance on reverse transcription. The authors would also like to express gratitude to Mr. George Princeton Dunsford for granting access to the wastewater samples, and to Dr Moataz El Ghany for ideas pertaining to the *E. coli* transcriptomics.

Supplementary Information Available

Methods: water quality testing; solar inactivation experiments; statistical comparisons of inactivation curves; RNA extraction and sequencing; RNA-seq analysis; RNA-seq mapping quality; cDNA synthesis; RT-qPCR confirmation experiments. Tables: genomic comparison overview; RNA-seq mapping quality, list of primers used for RT-qPCR; wastewater quality; inactivation kinetics in wastewater, genomic comparison in specific functional categories; fold-change of differentially regulated genes at select categories; RPKM values of differentially regulated genes at select categories. Figures: percentages and numbers of differentially regulated genes, PI-7 and DSM1103 inactivation with the efflux pump inhibitor, RNA-seq and RT-qPCR fold change for PI-7 and DSM1103.

Table 1. Regulatory genes with differential gene expression response in stress response, cell repair and virulence categories

a) Stress response	**PI-7**	**DSM 1103**
QorR, redox-sensing transcriptional regulator	-	NA
Biofilm formation regulator	+	+
MarA, multiple antibiotic resistance protein	+	o
b) Cellular repair		
RecA	+	o
LexA Repressor	-	-
Replication regulatory protein repA2	-	NA
c) Virulence		
Predicted regulator of CFA/I fimbriae	+	NA
FimB type 1 fimbriae regulatory proteins	+	NA
FimE type 1 fimbriae regulatory proteins	+	NA
FliA, RNA polymerase σ factor for flagellar operon	+	o
FlhC, flagellar transcriptional activator	+	o
FlgM, negative regulator of flagellin synthesis	o	-
RpoN, σ54 RNA polymerase	+	o

(+): upregulated; (-): downregulated; (o): present in genome but with no significant fold-change in expression; (NA) not found in the genome.
This list of regulatory genes is shortlisted from the full list of genes with significant fold-change in the selected categories (See Table S8).

569 **Table 2.** Numbers of upregulated genes out of the total number of present genes in stress
570 response, cell repair and virulence categories

571

a) Stress response	**PI-7**	**DSM 1103**
Biofilm	6/13	1/4
Catalases	1/2	0/2
Peroxidases	2/3	1/2
Glutathione	3/22	1/20
Other stress/detox	39/70	6/171
Multi drug and heavy metals efflux pumps/transporters	22/43	1/19
b) Cellular repair		
DNA repair	28/66	2/78
Cell division	17/38	1/36
Cell wall synthesis	87/191	7/192
DNA synthesis	9/14	1/15
Other DNA	15/21	4/32
Persistors	19/57	7/44
c) Virulence		
Virulence	25/41	6/38
Fimbriae/pili	30/53	13/42
Flagella and motility	29/61	6/60
Type III secretion system	4/4	0/NA

572 This list of regulatory genes is shortlisted from the full list of genes with significant fold-change
573 in the selected categories (See Table S8).
574 Genes were counted for inclusion in this table if they were upregulated at any of the examined
575 time points; total numbers are the total numbers of relevant genes that were found in the entire
576 genomes of the *E. coli* isolates.
577
578
579
580
581
582

583 **Figure legends**
584
585 **Figure 1.** Inactivation curves of *E. coli* PI-7 and DSM1103 under simulated solar irradiation in
586 buffer solution. *E. coli* PI-7 exhibits a longer lag phase than DSM1103 prior to decay. Irradiance
587 rate at 280-700 nm was 27.86 J/cm^2/h.
588
589 **Figure 2.** Inactivation curves of *E. coli* PI-7 and DSM1103 under simulated solar irradiation in
590 effluent and chlorinated effluent. Solar irradiation of *E. coli* PI-7 and DSM1103 in chlorinated
591 effluent and effluent, respectively, showed a longer half-life than that in buffer.
592
593

References

1. Baquero, F.; Martinez, J. L.; Canton, R., Antibiotics and antibiotic resistance in water environments. *Curr Opin Biotechnol* **2008,** *19*, (3), 260-5.

2. Schwartz, T.; Kohnen, W.; Jansen, B.; Obst, U., Detection of antibiotic-resistant bacteria and their resistance genes in wastewater, surface water, and drinking water biofilms. *FEMS Microbiol Ecol* **2003,** *43*, (3), 325-35.

3. Galvin, S.; Boyle, F.; Hickey, P.; Vellinga, A.; Morris, D.; Cormican, M., Enumeration and characterization of antimicrobial-resistant Escherichia coli bacteria in effluent from municipal, hospital, and secondary treatment facility sources. *Appl Environ Microbiol* **2010,** *76*, (14), 4772-9.

4. Al-Jassim, N.; Ansari, M. I.; Harb, M.; Hong, P. Y., Removal of bacterial contaminants and antibiotic resistance genes by conventional wastewater treatment processes in Saudi Arabia: Is the treated wastewater safe to reuse for agricultural irrigation? *Water Res* **2015,** *73*, 277-90.

5. Brechet, C.; Plantin, J.; Sauget, M.; Thouverez, M.; Talon, D.; Cholley, P.; Guyeux, C.; Hocquet, D.; Bertrand, X., Wastewater treatment plants release large amounts of extended-spectrum beta-lactamase-producing Escherichia coli into the environment. *Clin Infect Dis* **2014,** *58*, (12), 1658-65.

6. Kumarasamy, K. K.; Toleman, M. A.; Walsh, T. R.; Bagaria, J.; Butt, F.; Balakrishnan, R.; Chaudhary, U.; Doumith, M.; Giske, C. G.; Irfan, S., Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *The Lancet infectious diseases* **2010,** *10*, (9), 597-602.

7. Walsh, T. R.; Weeks, J.; Livermore, D. M.; Toleman, M. A., Dissemination of NDM-1 positive bacteria in the New Delhi environment and its implications for human health: an environmental point prevalence study. *The Lancet infectious diseases* **2011,** *11*, (5), 355-362.

8. Muir, A.; Weinbren, M., New Delhi metallo-β-lactamase: a cautionary tale. *Journal of Hospital Infection* **2010,** *75*, (3), 239-240.

9. Rasheed, J. K.; Kitchel, B.; Zhu, W.; Anderson, K. F.; Clark, N. C.; Ferraro, M. J.; Savard, P.; Humphries, R. M.; Kallen, A. J.; Limbago, B. M., New Delhi metallo-beta-lactamase-producing Enterobacteriaceae, United States. *Emerg Infect Dis* **2013,** *19*, (6), 870-8.

10. Poirel, L.; Hombrouck-Alet, C.; Freneaux, C.; Bernabeu, S.; Nordmann, P., Global spread of New Delhi metallo-β-lactamase 1. *The Lancet infectious diseases* **2010,** *10*, (12), 832.

11. Luo, Y.; Yang, F.; Mathieu, J.; Mao, D.; Wang, Q.; Alvarez, P., Proliferation of multidrug-resistant New Delhi metallo-β-lactamase genes in municipal wastewater treatment plants in northern China. *Environmental Science & Technology Letters* **2013,** *1*, (1), 26-30.

12. Al-Jassem, D., Treated sewage effluent initiative aims to ease water shortages. *Arab News* 6 Dec 2012, 2012.

13. Mantilla-Calderon, D.; Jumat, M. R.; Wang, T.; Ganesan, P.; Al-Jassim, N.; Hong, P.-Y., Isolation and characterization of NDM-positive Escherichia coli from municipal wastewater in Jeddah, Saudi Arabia. *Antimicrobial Agents and Chemotherapy* **2016,** *60*, (9), 5223-5231.

14. Huang, T. W.; Wang, J. T.; Lauderdale, T. L.; Liao, T. L.; Lai, J. F.; Tan, M. C.; Lin, A. C.; Chen, Y. T.; Tsai, S. F.; Chang, S. C., Complete sequences of two plasmids in a blaNDM-1-positive Klebsiella oxytoca isolate from Taiwan. *Antimicrob Agents Chemother* **2013,** *57*, (8), 4072-6.

15. Boyle, M.; Sichel, C.; Fernandez-Ibanez, P.; Arias-Quiroz, G. B.; Iriarte-Puna, M.; Mercado, A.; Ubomba-Jaswa, E.; McGuigan, K. G., Bactericidal effect of solar water disinfection under real sunlight conditions. *Appl Environ Microbiol* **2008,** *74*, (10), 2997-3001.

16. McGuigan, K. G.; Joyce, T. M.; Conroy, R. M.; Gillespie, J. B.; Elmore-Meegan, M., Solar disinfection of drinking water contained in transparent plastic bottles: characterizing the bacterial inactivation process. *J Appl Microbiol* **1998,** *84*, (6), 1138-48.

17. Ubomba-Jaswa, E.; Navntoft, C.; Polo-Lopez, M. I.; Fernandez-Ibanez, P.; McGuigan, K. G., Solar disinfection of drinking water (SODIS): an investigation of the effect of UV-A dose on inactivation efficiency. *Photochem Photobiol Sci* **2009,** *8*, (5), 587-95.

18. Wegelin, M.; Canonica, S.; Mechsner, K.; Fleischmann, T.; Pesaro, F.; Metzler, A., Solar water disinfection: scope of the process and analysis of radiation experiments. *Aqua* **1994,** *43*, (4), 154-169.

19. McGuigan, K. G.; Conroy, R. M.; Mosler, H. J.; du Preez, M.; Ubomba-Jaswa, E.; Fernandez-Ibanez, P., Solar water disinfection (SODIS): a review from bench-top to roof-top. *J Hazard Mater* **2012,** *235-236*, 29-46.

20. Oates, P. M.; Shanahan, P.; Polz, M. F., Solar disinfection (SODIS): simulation of solar radiation for global assessment and application for point-of-use water treatment in Haiti. *Water Res* **2003,** *37*, (1), 47-54.

21. Rijal, G. K.; Fujioka, R. S., Use of reflectors to enhance the synergistic effects of solar heating and solar wavelengths to disinfect drinking water sources. *Water Sci Technol* **2003,** *48*, (11-12), 481-8.

22. Fisher, M. B.; Keenan, C. R.; Nelson, K. L.; Voelker, B. M., Speeding up solar disinfection (SODIS): effects of hydrogen peroxide, temperature, pH, and copper plus ascorbate on the photoinactivation of E. coli. *J Water Health* **2008,** *6*, (1), 35-51.

23. Fisher, M. B.; Iriarte, M.; Nelson, K. L., Solar water disinfection (SODIS) of Escherichia coli, Enterococcus spp., and MS2 coliphage: effects of additives and alternative container materials. *Water Res* **2012,** *46*, (6), 1745-54.

24. Sinton, L. W.; Hall, C. H.; Lynch, P. A.; Davies-Colley, R. J., Sunlight inactivation of fecal indicator bacteria and bacteriophages from waste stabilization pond effluent in fresh and saline waters. *Appl Environ Microbiol* **2002,** *68*, (3), 1122-31.

667 25. Sommer, B.; Marino, A.; Solarte, Y.; Salas, M.; Dierolf, C.; Valiente, C.; Mora, D.;
668 Rechsteiner, R.; Setter, P.; Wirojanagud, W., SODIS- an emerging water treatment process.
669 *AQUA(OXFORD)* **1997,** *46*, (3), 127-137.

670 26. Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma,
671 K.; Gerdes, S.; Glass, E. M.; Kubal, M., The RAST Server: rapid annotations using subsystems
672 technology. *BMC genomics* **2008,** *9*, (1), 75.

673 27. Overbeek, R.; Olson, R.; Pusch, G. D.; Olsen, G. J.; Davis, J. J.; Disz, T.; Edwards, R.
674 A.; Gerdes, S.; Parrello, B.; Shukla, M., The SEED and the Rapid Annotation of microbial
675 genomes using Subsystems Technology (RAST). *Nucleic acids research* **2014,** *42*, (D1), D206-
676 D214.

677 28. Minogue, T. D.; Daligault, H. A.; Davenport, K. W.; Bishop-Lilly, K. A.; Broomall, S.
678 M.; Bruce, D. C.; Chain, P. S.; Chertkov, O.; Coyne, S. R.; Freitas, T.; Frey, K. G.; Gibbons, H.
679 S.; Jaissle, J.; Redden, C. L.; Rosenzweig, C. N.; Xu, Y.; Johnson, S. L., Complete Genome
680 Assembly of Escherichia coli ATCC 25922, a Serotype O6 Reference Strain. *Genome Announc*
681 **2014,** *2*, (5), e00969-14.

682 29. APHA; AWWA; WEF, Standard Method for the Examination of Water and Wastewater.
683 In American Water Works Association: 2012; pp 4-69.

684 30. Romero, O. C.; Straub, A. P.; Kohn, T.; Nguyen, T. H., Role of temperature and
685 Suwannee River natural organic matter on inactivation kinetics of rotavirus and bacteriophage
686 MS2 by solar irradiation. *Environmental science & technology* **2011,** *45*, (24), 10385-10393.

687 31. Baggerly, K. A.; Deng, L.; Morris, J. S.; Aldaz, C. M., Differential expression in SAGE:
688 accounting for normal between-library variation. *Bioinformatics* **2003,** *19*, (12), 1477-1483.

689 32. Rizzo, L.; Manaia, C.; Merlin, C.; Schwartz, T.; Dagot, C.; Ploy, M. C.; Michael, I.;
690 Fatta-Kassinos, D., Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria
691 and genes spread into the environment: a review. *Sci Total Environ* **2013,** *447*, 345-60.

692 33. Pruden, A., Balancing water sustainability and public health goals in the face of growing
693 concerns about antibiotic resistance. *Environmental science & technology* **2014,** *48*, (1), 5-14.

694 34. da Silva, M. F.; Vaz-Moreira, I.; Gonzalez-Pajuelo, M.; Nunes, O. C.; Manaia, C. M.,
695 Antimicrobial resistance patterns in Enterobacteriaceae isolated from an urban wastewater
696 treatment plant. *Fems Microbiol Ecol* **2007,** *60*, (1), 166-176.

697 35. Mokracka, J.; Koczura, R.; Kaznowski, A., Multiresistant Enterobacteriaceae with class 1
698 and class 2 integrons in a municipal wastewater treatment plant. *Water research* **2012,** *46*, (10),
699 3353-63.

700 36. CDC Antimicrobial resistance threat report
701 2013. http://www.cdc.gov/drugresistance/threat-report-2013/ (2 July 2014),

37. Fatta-Kassinos, D.; Vasquez, M. I.; Kummerer, K., Transformation products of pharmaceuticals in surface waters and wastewater formed during photolysis and advanced oxidation processes - degradation, elucidation of byproducts and assessment of their biological potency. *Chemosphere* **2011,** *85*, (5), 693-709.

38. Lam, M. W.; Mabury, S. A., Photodegradation of the pharmaceuticals atorvastatin, carbamazepine, levofloxacin, and sulfamethoxazole in natural waters. *Aquatic Sciences* **2005,** *67*, (2), 177-188.

39. Rizzo, L.; Fiorentino, A.; Anselmo, A., Advanced treatment of urban wastewater by UV radiation: Effect on antibiotics and antibiotic-resistant E. coli strains. *Chemosphere* **2013,** *92*, (2), 171-6.

40. Crittenden, J. C.; Harza, M. W.; Hand, D. W.; Howe, K. J., *MWH's Water Treatment: Principles and Design*. Wiley: 2012.

41. Bolton, N. F.; Cromar, N. J.; Hallsworth, P.; Fallowfield, H. J., A review of the factors affecting sunlight inactivation of micro-organisms in waste stabilisation ponds: preliminary results for enterococci. *Water Sci Technol* **2010,** *61*, (4), 885-90.

42. Muela, A.; Garcia-Bringas, J. M.; Seco, C.; Arana, I.; Barcina, I., Participation of oxygen and role of exogenous and endogenous sensitizers in the photoinactivation of Escherichia coli by photosynthetically active radiation, UV-A and UV-B. *Microb Ecol* **2002,** *44*, (4), 354-64.

43. Santos, A. L.; Gomes, N. C.; Henriques, I.; Almeida, A.; Correia, A.; Cunha, A., Contribution of reactive oxygen species to UV-B-induced damage in bacteria. *J Photochem Photobiol B* **2012,** *117*, 40-6.

44. Pattison, D. I.; Davies, M. J., Actions of ultraviolet light on cellular structures. *EXS* **2006**, (96), 131-57.

45. Fernandez Zenoff, V.; Sineriz, F.; Farias, M. E., Diverse responses to UV-B radiation and repair mechanisms of bacteria isolated from high-altitude aquatic environments. *Appl Environ Microbiol* **2006,** *72*, (12), 7857-63.

46. Wilson, C.; Caton, T. M.; Buchheim, J. A.; Buchheim, M. A.; Schneegurt, M. A.; Miller, R. V., DNA-repair potential of Halomonas spp. from the Salt Plains Microbial Observatory of Oklahoma. *Microb Ecol* **2004,** *48*, (4), 541-9.

47. McKinney, C. W.; Pruden, A., Ultraviolet disinfection of antibiotic resistant bacteria and their antibiotic resistance genes in water and wastewater. *Environ Sci Technol* **2012,** *46*, (24), 13393-400.

48. Janion, C., Inducible SOS response system of DNA repair and mutagenesis in Escherichia coli. *Int J Biol Sci* **2008,** *4*, (6), 338-44.

736 49. Ehling-Schulz, M.; Bilger, W.; Scherer, S., UV-B-induced synthesis of photoprotective
737 pigments and extracellular polysaccharides in the terrestrial cyanobacterium Nostoc commune. *J*
738 *Bacteriol* **1997,** *179*, (6), 1940-5.

739 50. Kehr, J. C.; Dittmann, E., Biosynthesis and function of extracellular glycans in
740 cyanobacteria. *Life (Basel)* **2015,** *5*, (1), 164-80.

741 51. McCord, J. M.; Fridovich, I., Superoxide dismutase. An enzymic function for
742 erythrocuprein (hemocuprein). *J Biol Chem* **1969,** *244*, (22), 6049-55.

743 52. Triggs-Raine, B. L.; Loewen, P. C., Physical characterization of katG, encoding catalase
744 HPI of Escherichia coli. *Gene* **1987,** *52*, (2-3), 121-8.

745 53. Loewen, P. C.; Triggs, B. L., Genetic mapping of katF, a locus that with katE affects the
746 synthesis of a second catalase species in Escherichia coli. *J Bacteriol* **1984,** *160*, (2), 668-75.

747 54. Loewen, P. C., Isolation of catalase-deficient Escherichia coli mutants and genetic
748 mapping of katE, a locus that affects catalase activity. *J Bacteriol* **1984,** *157*, (2), 622-6.

749 55. Heck, D. E.; Shakarjian, M.; Kim, H. D.; Laskin, J. D.; Vetrano, A. M., Mechanisms of
750 oxidant generation by catalase. *Ann N Y Acad Sci* **2010,** *1203*, 120-5.

751 56. Heck, D. E.; Vetrano, A. M.; Mariano, T. M.; Laskin, J. D., UVB light stimulates
752 production of reactive oxygen species: unexpected role for catalase. *J Biol Chem* **2003,** *278*, (25),
753 22432-6.

754 57. Partridge, J. D.; Poole, R. K.; Green, J., The Escherichia coli yhjA gene, encoding a
755 predicted cytochrome c peroxidase, is regulated by FNR and OxyR. *Microbiology* **2007,** *153*, (Pt
756 5), 1499-507.

757 58. Sassoubre, L. M.; Ramsey, M. M.; Gilmore, M. S.; Boehm, A. B., Transcriptional
758 response of Enterococcus faecalis to sunlight. *Journal of Photochemistry and Photobiology B:*
759 *Biology* **2014,** *130*, 349-356.

760 59. Chang, M. W.; Toghrol, F.; Bentley, W. E., Toxicogenomic response to chlorination
761 includes induction of major virulence genes in Staphylococcus aureus. *Environmental science &*
762 *technology* **2007,** *41*, (21), 7570-7575.

763 60. Karumathil, D. P.; Yin, H.-B.; Kollanoor-Johny, A.; Venkitanarayanan, K., Effect of
764 chlorine exposure on the survival and antibiotic gene expression of multidrug resistant
765 Acinetobacter baumannii in water. *International journal of environmental research and public*
766 *health* **2014,** *11*, (2), 1844-1854.

767 61. Lewis, K., Persister cells, dormancy and infectious disease. *Nat Rev Microbiol* **2007,** *5*,
768 (1), 48-56.

769 62. Fang, F. C., Antimicrobial actions of reactive oxygen species. *MBio* **2011,** *2*, (5), e00141-
770 11.

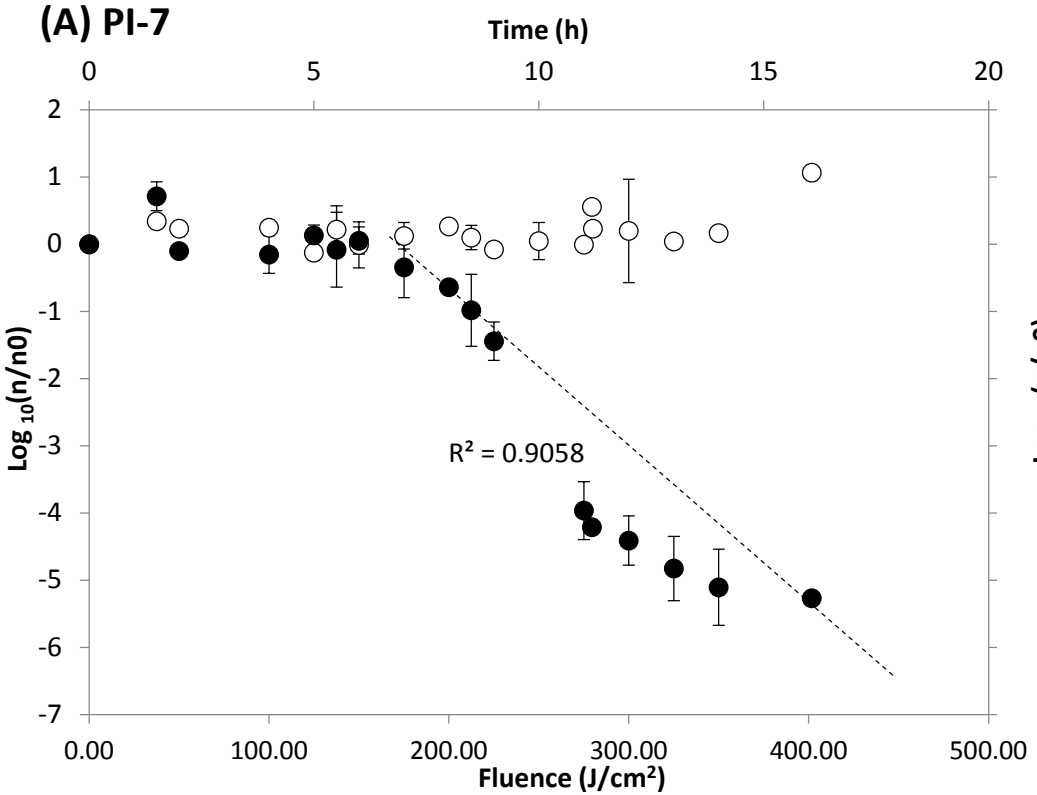
771 63. Canonica, S.; Kohn, T.; Mac, M.; Real, F. J.; Wirz, J.; von Gunten, U., Photosensitizer
772 method to determine rate constants for the reaction of carbonate radical with organic compounds.
773 *Environ Sci Technol* **2005,** *39*, (23), 9182-8.

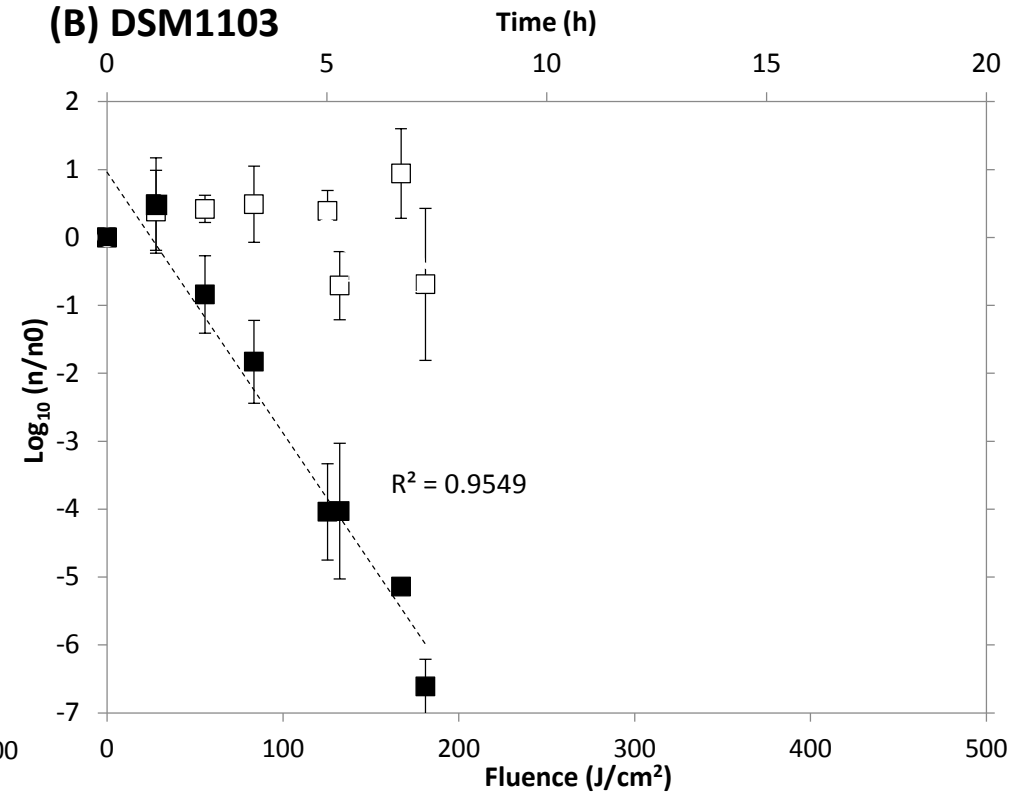
774 64. Hunting, E. R.; White, C. M.; van Gemert, M.; Mes, D.; Stam, E.; van, H. G.; der, G.;
775 Kraak, M. H.; Admiraal, W., UV radiation and organic matter composition shape bacterial
776 functional diversity in sediments. *Front Microbiol* **2013,** *4*, 317.

777 65. Wetzel, R. G.; Hatcher, P. G.; Bianchi, T. S., Natural photolysis by ultraviolet irradiance
778 of recalcitrant dissolved organic matter to simple substrates for rapidbacterial metabolism.
779 *Limnology and Oceanography* **1995,** *40*, (8), 1369-1380.

780 66. Neusser, T.; Polen, T.; Geissen, R.; Wagner, R., Depletion of the non-coding regulatory
781 6S RNA in E. coli causes a surprising reduction in the expression of the translation machinery.
782 *BMC genomics* **2010,** *11*, (1), 1.

783 67. Gottesman, S.; Storz, G., Bacterial small RNA regulators: versatile roles and rapidly
784 evolving variations. *Cold Spring Harbor perspectives in biology* **2011,** *3*, (12), a003798.
785





(C) Decay kinetics

	Lag phase length (h)	Lag phase fluence (J/cm²)	Decay constant	Half-life length (min)	Half-life fluence (J/cm²)
PI-7 (n=7)	6.64 ± 0.63	184.96 ± 17.55	-14.96 ± 2.61	2.85 ± 0.46	1.32 ± 0.21
DSM1103 (n=6)	1.33 ± 0.52	37.05 ± 14.49	-21.01 ± 4.54	2.04 ± 0.36	0.95 ± 0.17

Figure 1.

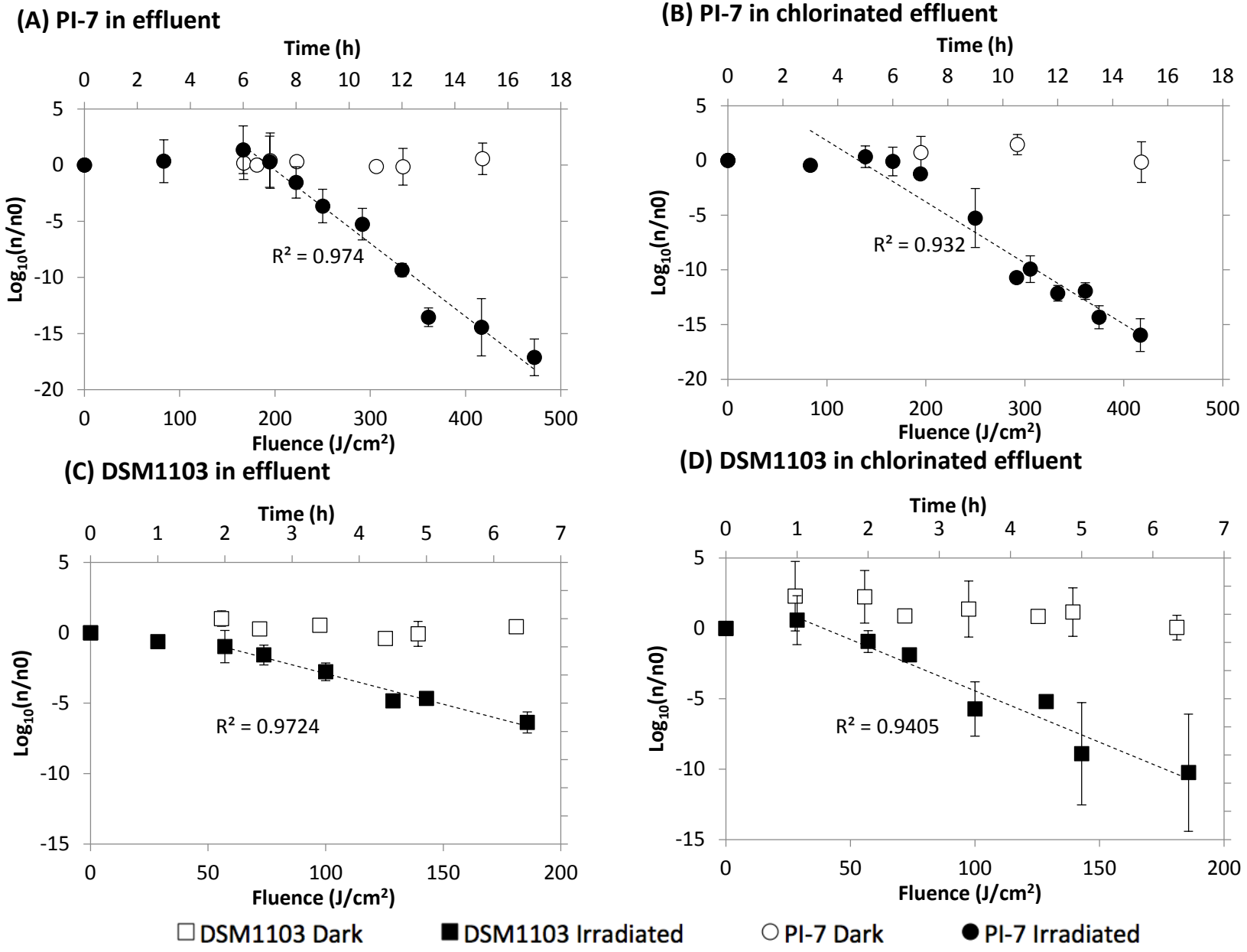

(A) PI-7 in effluent
Time (h)
Log10(n/n0)
R² = 0.974
Fluence (J/cm²)
(B) PI-7 in chlorinated effluent
R² = 0.932
(C) DSM1103 in effluent
R² = 0.9724
(D) DSM1103 in chlorinated effluent
R² = 0.9405
DSM1103 Dark
DSM1103 Irradiated
PI-7 Dark
PI-7 Irradiated

Figure 2.

Supplementary Information

Inactivation of a virulent wastewater *Escherichia coli* and non-virulent commensal *Escherichia coli* DSM1103 strains and their gene expression upon solar irradiation

Nada Al-Jassim [1], David Mantilla-Calderon [1], Tiannyu Wang [1], Pei-Ying Hong * [1]

1. King Abdullah University of Science and Technology (KAUST), Water Desalination and Reuse Center (WDRC), Biological and Environmental Sciences & Engineering Division (BESE), Thuwal, 23955-6900, Saudi Arabia

Summary of contents

Supplementary Information – 25 pages
Includes: Supplementary methodology – 7 pages
Supplementary tables – 7 tables
Supplementary figures – 5 figures
Additional supplementary files – 1 file
Microsoft Excel spreadsheet – 2 tables

Table of Contents

Separately from this document:

1. Water quality testing

COD was measured using LCK 314 (15-150 mg/L) cuvette test vials based on protocols specified by the manufacturer (Hach-Lange, Manchester, UL). Concentrations of DOC and TN were measured by the high-temperature catalytic oxidation (HTCO) method using a commercially available automatic TOC-V_{CPH} analyzer with a TNM-1 add-on (Shimadzu, Japan), following the methodology recommended by Shimadzu. The SUVA (in L/mg-M) was calculated by dividing the UV absorbance at 254 nm of the sample (in cm^{-1}) by the DOC (in mg/L) of the sample and then multiplying by 100 cm/M. Alkalinity was measured based on standard method 2320. Total chlorine concentration in chlorinated effluent samples was determined using the DPD/KI (N, N-diethyl-p-phenylenediamine/potassium iodide) colorimetric standard method 4500-CI G 29.

2. Solar inactivation experiments

Bacterial isolates were streaked onto nutrient agar (Sigma–Aldrich, Buchs, Switzerland) plates for *E. coli* DSM 1103, and nutrient agar plates supplemented with 8 µg/mL meropenem for *E. coli* PI-7 in order to maintain selective pressure to prevent loss of the *bla*$_{NDM-1}$-containing plasmid. Pure colonies of *E. coli* PI-7 and *E. coli* DSM 1103 were picked and inoculated into the corresponding nutrient with and without meropenem, respectively. Liquid cultures were incubated at 37 °C to obtain a bacterial culture at the late exponential growth phase and with an optical density at 600 nm of close to 1.0. Cells were washed three times with 1 mM $NaHCO_3$, and were then resuspended in 1X phosphate buffered saline, PBS (Thermo Fisher Scientific, CA, USA). In the case of solar inactivation trials in wastewater, effluent or chlorinated effluent that was filtered through 0.2 µm polycarbonate membranes was used in place of 1X PBS. Homogenous suspensions of either

E. coli DSM 1103 or PI-7 were then split into 80 mL portions over a number of microcosms. Microcosms used for the experiments were made from 100 mL glass beakers wrapped in black scotch tape to prevent unwanted light penetration. Each microcosm contained a magnetic stir bar and the top was either covered with aluminum foil in the case of dark controls, or with a glass filter (Newport Corporation, Irvine, CA) that allows light wavelengths ≥ 280 nm to pass through. Solar exposure was conducted in an Atlas Suntest ® XLS+ photosimulator (Chicago, IL) equipped with a xenon arc lamp. The solar irradiation intensity was measured by Spectroradiometer ILT950 (International Light Technologies, Peabody, MA, USA) and shown in figure below. The UV irradiance provided by the photosimulator approximates to the irradiance measurements of direct and indirect (shaded) noon sunlight that were measured in the KAUST campus during the month of December 2015.

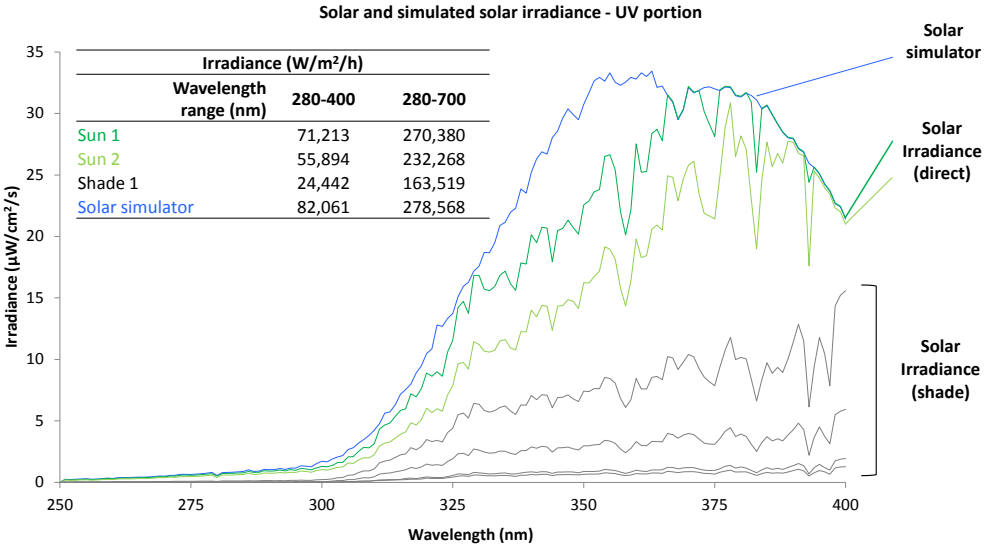


Figure. Solar simulator validation. Solar irradiance was measured at noon during December 2015 in King Abdullah University of Science and Technology, Saudi Arabia. Irradiance was measured in two spots under direct sunlight, and four spots in the shade.
Table shows total irradiance (in Watts/m²/h) for environmental and simulated solar irradiation.

3. Statistical comparisons of inactivation curves

Statistical analyses were performed using Minitab version 1.4.0. One-way ANOVA was performed to compare the lengths of lag-phases, the corrected decay constants and the half-life values of the different experimental conditions and isolates. Single-regression analysis was performed to compare each slope to 0 at $\alpha = 0.05$ so as to determine if inactivation or persistence is observed. Multiple-regression analysis was also performed on decay curves from each experimental batch to determine if control and irradiated samples were significantly different within and between groups ($\alpha = 0.05$). Lag-phase length and half-life were not analyzed for dark control samples due to the lack of decay in these samples.

4. RNA extraction and sequencing

Sample volumes varying between 1 to 4 mL were taken from the experimental microcosms and then centrifuged at 10,000 *g* for 10 minutes at room temperature. The pellets were resuspended in 1 mL of RNAProtect® preservation reagent (Qiagen, Hilden, Germany) and then preserved at 4 °C until RNA extraction. The total RNA was extracted using the RNeasy Mini kit (Qiagen, Hilden, Germany) with the on-column DNase treatment. The RNA was eluted with RNase-free water and RNA quality was determined using 2200 Tapestation Bioanalyzer (Agilent Technologies, Santa Clara, CA). All samples that passed the quality control were submitted to KAUST Genomics Core lab for RNA-Seq on Illumina HiSeq platform. Prior to sequencing, the bacterial ribosomal RNA was removed from 1 µg of total RNA using the Ribo Zero rRNA removal kit (Illumina, San Diego, CA, USA). The enriched mRNA fraction was converted to RNA-seq libraries compatible for runs on Illumina HiSeq platforms based on protocols described by manufacturer. Fastq files were generated and demultiplexed by KAUST Bioinformatics core lab.

5. RNA-seq analysis

The RNA-seq data was analyzed using CLC Genomics Workbench version 8.0.1 from CLC Bio (Cambridge, MA). For reference genome sequences, the published annotated sequences were used for *E. coli* DSM1103 [28] and genomic DNA sequences annotated using RAST were used for *E. coli* PI-7 [10]. RNA-seq reads were mapped onto the reference sequences using CLC. Reads were only assembled if the fraction of read that aligned to the reference genome was greater than 0.9 and if the read matched other regions of the reference genome at fewer than 10 nucleotide positions. The RNA-seq output files were analyzed for statistical significance, separately for each of the two experimental batches for each *E. coli*, based on the proportion-based test described by Baggerly *et al.* [31]. Gene expression levels in irradiated samples were compared with their respective dark controls at each time point for each experimental batch. Change in expression level was taken into consideration if it was statistically significant (p-value <0.05), and showed ≥ 2-fold change consistently in both biological replicates. Afterwards, CLC was used again to generate normalized mean expression values based on both experimental batches for each *E. coli*, and fold-changes in expression levels were calculated based on those means. Thus, some of the results showed high yet statistically insignificant fold-changes in expression levels. This might have been caused by too large of a difference in normalized expression values within irradiated samples in a replicate or within dark samples in a replicate, or caused by only of the experimental batches and not the other having large and significant fold-change.

6. RNA-seq read mapping quality

Summarized RNA-seq mapping results for both strains are shown in Table S2 and S3. The percentage of reads mapped to the genome for each isolate was always 85% or higher (Table

S2 and S3), indicating good read quality and no significant contamination of the samples. Average percentages of reads mapped to CDS (coding sequence) at each phase were significantly lower than reads mapped to the genome in both *E. coli*, which may indicate expression of unidentified/non-annotated genes, especially with PI-7's genome not being closed. The percentage of RNA reads that mapped to coding sequence decreased over increasing duration of exposure to irradiance. This might have been a result of accumulation of irradiance-incurred damage to the genomes and transcripts. Bacterial expression of non-coding sequences and regulatory RNAs (which have been reported to be induced under some stress conditions and to play a role in activation/repression of membrane, membrane transport, cell division and sporulation functions [66, 67]) may partially explain the low read coverage of CDS regions, although it is unlikely that it would account for a majority of the percentage of non-CDS reads.

7. cDNA synthesis

For RT-qPCR confirmation experiments, RNA samples which were preserved and extracted as described in Supplementary Information 4 were used for cDNA generation according to Invitrogen's SuperScript First-Strand Synthesis System for RT-PCR kit reagents and instructions. Briefly, each mRNA sample was mixed with 100 pmol of random hexamers, incubated at 65 °C for 5 min then chilled on ice for 1 min. Each reaction was then mixed with 2 µL 10X RT buffer, 4 µL of 25 mM $MgCl_2$, 2 µL of 0.1 M DTT and 1 µL of 40U/µL RNaseOUT, incubated at room temperature for 2 min, then 1 µL of SuperScript II RT was added. Reactions were incubated at room temperature for 10 min, at 42 °C for 50 min, then the reactions were terminated at 70 °C for 15 min and chilled on ice. Next, 1 µL of RNase H

was added to each reaction and they were incubated at 37 °C for 20 min. The generated cDNA was stored at -20 °C until reverse transcription quantitative PCR (RT-qPCR).

8. RT-qPCR confirmation experiments

After RNA-seq and cDNA synthesis, there was sufficient cDNA to analyze only the early-decay and mid-decay phase samples for the selected genes present in both isolates using RT-qPCR. Selected genes for RT-qPCR include *fts*I, *hfl*C, *lex*A , *mar*A and *uid*A for both *E. coli* PI-7 and DSM1103, and gene CFA was additionally analyzed for *E. coli* PI-7. All primers are listed in Table S4. qPCR standards for the respective genes were prepared by cloning target gene amplicons into pGEM®-T easy vectors (Promega, Madison, WI, USA). Plasmids carrying the gene inserts were harvested from transformed cells using PureYield™ Plasmid Miniprep System (Promega, Madison, WI, USA). The extracted plasmids were sequenced to verify the insertions. Plasmid copy numbers were calculated based on the concentration of extracted plasmid DNA and the known sizes of vector and insert, and were diluted in series from 10^2 to 10^9 copies/µL to produce a 6-points qPCR standard curves for each gene. For gene *uid*A qPCR, the TaqMan assay was used. Triplicate 20 µL reactions for each sample included 10 µL Applied Biosystems' TaqMan Fast Advanced master mix (Thermo Fisher Scientific, Carlsbad, CA, USA), 1 µL of each primer (10 µM), 0.8 µl of probe (10 µM), 1 ng standard or cDNA template (or 2 µL H_2O in non-template controls) and topped up to 20 µL using molecular-biology grade water. For all remaining genes, triplicate 20 µL qPCR reactions per sample were set up using 10 µL of Applied Biosystems' PowerSYBR Green master mix (Thermo Fisher Scientific, Carlsbad, CA, USA), 0.4 µL of each primer (10 µM), 1 ng standard or cDNA template (or 2 µL H_2O in non-template controls) and topped up to 20 µL using molecular-biology grade water. qPCR, following the relative standard curve

method, was performed using Applied Biosystems®7900HT Fast Real-Time PCR system with 384-well block module (Thermo Fisher Scientific, Carlsbad, CA, USA) using the cycling conditions listed in Table S4. For those genes using a Sybr Green reporter, melting curve analysis was performed with a dissociation cycle that included an increment of temperature from 60 °C to 95 °C, at an interval of 0.5 °C for 5 s. After optimization, copy numbers and relative concentrations of each gene to house-keeping gene *uid*A were calculated using Microsoft Excel. These proportions were compared in irradiated and dark samples to generate fold-change values that were compared with fold-change values generated from RNA-seq analysis.

Table S1. Summary of sequence- and function-based genome comparisons between *E. coli* PI-7 and DSM1103

Isolate	**PI-7** [a]	**DSM1103** [b]
Genome	Not closed (6 contigs)	Closed
Genome size	4.73 Mb	5.20 Mb
Plasmids	>1 plasmid - 110,768 bp, (not closed), carries NDM-1 gene - Probable second plasmid, best match pkoxR1	2 plasmids - 48,488 bp - 24,185 bp
Total number of annotated genes [*]	4,471	4,808
Number of core gene sequences [*]	3971 PI-7 gene sequences have a match in DSM 1103. 4139 DSM 1103 gene sequences have a match in PI-7.	
Number of unique gene sequences [*]	489	910
Number of core annotated functions [*]	3860 functions are shared between both isolates.	
Number of unique annotated functions [*]	155	185

[a]: *E. coli* PI-7's genome reference [10].
[b]: *E. coli* DSM1103's full genome reference [28].
[*]: Numbers of annotated genes, gene sequences and functions were obtained through comparisons performed using RAST [26, 27].

Table S2. RNA-seq mapping results for *E. coli* PI-7

Time (h)	Microcosm sample		Reads after trimming [a]	Reads mapped to CDS regions [b]			Reads mapped to genome		
				Uniquely mapped reads (%)	Nonspecifically mapped reads (%)	Unmapped reads (%)	Uniquely mapped reads (%)	Nonspecifically mapped reads (%)	Unmapped reads (%)
0	*Exp1*	D-a	15,444,160	4,845,780 (31.38)	732,991 (4.75)	9,865,389 (63.88)	13,131,686 (85.03)	849,585 (5.50)	1,462,889 (9.47)
		D-b	13,588,000	3,926,546 (28.90)	722,996 (5.32)	8,938,458 (65.78)	11,587,430 (85.28)	787,544 (5.80)	1,213,026 (8.93)
	Exp2	D-c	16,378,002	4,527,904 (27.65)	642,683 (3.92)	11,207,415 (68.43)	15,073,086 (92.03)	390,225 (2.38)	914,691 (5.58)
		D-d	13,660,428	3,875,758 (28.37)	565,103 (4.14)	9,219,567 (67.49)	12,611,660 (92.32)	315,938 (2.31)	732,830 (5.36)
3	*Exp1*	D-a	14,088,460	4,428,958 (31.44)	821,530 (5.83)	8,837,972 (62.73)	12,476,438 (88.56)	588,176 (4.17)	1,023,846 (7.27)
		D-b	25,042,288	7,571,428 (30.23)	1,727,803 (6.90)	15,743,057 (62.87)	22,218,450 (88.72)	1,087,498 (4.34)	1,736,340 (6.93)
	Exp2	D-c	40,238,272	1,0840,606 (26.94)	2,150,318 (5.34)	27,247,348 (67.72)	37,396,502 (92.94)	871,408 (2.17)	1,970,362 (4.90)
		D-d	13,432,292	4,155,798 (30.94)	813,407 (6.06)	8,463,087 (63.01)	12,339,104 (91.86)	331,396 (2.47)	761,792 (5.67)
	Exp1	I-a	26,,429,352	4,241,696 (16.05)	822,211 (3.11)	21,365,445 (80.84)	23,556,312 (89.13)	882,411 (3.34)	1,990,629 (7.53)
		I-b	23,,977,652	3,891,282 (16.23)	770,154 (3.21)	19,316,216 (80.56)	21,390,502 (89.21)	722,734 (3.01)	1,864,416 (7.78)
	Exp2	I-c	12,291,282	1,570,730 (12.78)	332,702 (2.71)	10,387,850 (84.51)	11,465,608 (93.28)	157,596 (1.28)	668,078 (5.44)
		I-d	36,637,170	4,571,852 (12.48)	929,162 (2.54)	31,136,156 (84.99)	34,169,310 (93.26)	505,830 (1.38)	1,962,030 (5.36)

(Continued) Table S2. RNA-seq mapping results for *E. coli* PI-7

6.5	*Exp1*	D-a	13,949,762	4,289,990 (30.75)	1,056,767 (7.58)	8,603,005 (61.67)	12,585,402 (90.22)	486,699 (3.49)	877,661 (6.29)
		D-b	27,042,942	8,000,056 (29.58)	1,945,657 (7.19)	17,097,229 (63.22)	24,324,334 (89.95)	924,013 (3.42)	1,794,595 (6.64)
	Exp2	D-c	22,959,496	6,401,804 (27.88)	1,422,866 (6.20)	15,134,826 (65.92)	21,588,026 (94.03)	331,214 (1.44)	1,040,256 (4.53)
		D-d	13,131,734	3,601,840 (27.43)	789,544 (6.01)	8,740,350 (66.56)	12,211,974 (93.00)	345,898 (2.63)	573,862 (4.37)
	Exp1	I-a	18,844,808	1,621,614 (8.61)	258,924 (1.37)	16,964,270 (90.02)	17,091,268 (90.69)	594,743 (3.16)	1,158,797 (6.15)
		I-b	20,577,680	1,758,406 (8.55)	291,780 (1.42)	18,527,494 (90.04)	18,479,942 (89.81)	860,386 (4.18)	1,237,352 (6.01)
	Exp2	I-c	21,039,096	1,434,716 (6.82)	240,914 (1.15)	19,363,466 (92.04)	19,910,708 (94.64)	249,888 (1.19)	878,500 (4.18)
		I-d	8,605,052	327,498 (3.81)	50,071 (0.58)	8,227,483 (95.61)	8,099,418 (94.12)	201,274 (2.34)	304,360 (3.54)
9	*Exp1*	D-a	23,269,980	6,507,794 (27.97)	1,561,815 (6.71)	15,200,371 (65.32)	20,998,674 (90.24)	888,322 (3.82)	1,382,984 (5.94)
		D-b	13,150,170	3,513,402 (26.72)	1,006,087 (7.65)	8,630,681 (65.63)	11,709,264 (89.04)	627,492 (4.77)	813,414 (6.19)
	Exp2	D-c	15,922,346	4,229,432 (26.56)	821,459 (5.16)	10,871,455 (68.28)	14,944,442 (93.86)	285,998 (1.80)	691,906 (4.35)
		D-d	14,023,418	3,835,200 (27.35)	880,342 (6.28)	9,307,876 (66.37)	13,136,832 (93.68)	272,688 (1.94)	613,898 (4.38)
	Exp1	I-a	16,463,292	565,574 (3.44)	122,619 (0.74)	15,775,099 (95.82)	14,336,964 (87.08)	907,337 (5.51)	1,218,991 (7.40)
		I-b	22,237,810	1,195,962 (5.38)	184,255 (0.83)	20,857,593 (93.79)	20,340,490 (91.47)	701,532 (3.15)	1,195,788 (5.38)

(Continued) Table S2. RNA-seq mapping results for *E. coli* PI-7

	Exp2	I-c	45,847,192	1,602,732 (3.50)	272,692 (0.59)	43,971,768 (95.91)	43,744,596 (95.41)	449,582 (0.98)	1,653,014 (3.61)
		I-d	43,281,446	878,172 (2.03)	154,410 (0.36)	42,248,864 (97.61)	41,690,864 (96.33)	372,320 (0.86)	1,218,262 (2.81)
14.5	*Exp1*	D-a	15,532,630	3,899,944 (25.11)	1,006,179 (6.48)	10,626,570 (68.41)	14,207,438 (91.47)	477,267 (3.07)	847,925 (5.46)
		D-b	14,308,202	3,726,754 (26.05)	885,293 (6.19)	9,696,155 (67.77)	12,601,118 (88.07)	746,680 (5.22)	960,404 (6.71)
	Exp2	D-c	20,836,926	4,778,376 (22.93)	1,050,474 (5.04)	15,008,076 (72.03)	19,683,606 (94.47)	257,530 (1.24)	895,790 (4.30)
		D-d	15,030,408	3,480,272 (23.15)	783,125 (5.21)	10,767,011 (71.63)	13,809,288 (91.88)	501,684 (3.34)	719,436 (4.79)
	Exp1	I-a	22,524,494	1,175,550 (5.22)	194,864 (0.87)	21,154,080 (93.92)	20,223,970 (89.79)	726,414 (3.22)	1,574,110 (6.99)
		I-b	15,188,868	1,598,506 (10.52)	223,504 (1.47)	13,366,858 (88.00)	13,711,276 (90.27)	417,563 (2.75)	1,060,029 (6.98)
	Exp2	I-c	15,252,230	377,914 (2.48)	59,849 (0.39)	14,814,469 (97.13)	14,551,670 (95.41)	129,243 (0.85)	571,317 (3.75)
		I-d	16,332,674	897,944 (5.50)	93,127 (0.57)	15,341,603 (93.93)	15,115,662 (92.55)	467,128 (2.86)	749,884 (4.59)

Time (h) denotes exposure to irradiance durations corresponding to the decay phases as follows: 0 h = prior to solar irradiation; 3 h = mid-lag phase; 6 h = early-decay phase; 9 h = mid-decay phase; 14.5 = late-decay phase.
D/I: Microcosm samples are labeled as D- or I- to denote dark and irradiated microcosms, respectively.
CDS: coding sequence.
[a] All reads are 101 nucleotides long.
[b] Reads were mapped to PI-7 reference genome using CLC Genomics Workbench version 8.0.1 with a minimum length fraction of 0.9, a minimum similarity fraction of 0.8, and a maximum number of hits for a read of 10.

Table S3. RNA-seq mapping results for *E. coli* DSM1103

Time (h)	Microcosm sample		Reads after trimming [a]	Reads mapped to CDS regions [b]			Reads mapped to genome		
				Uniquely mapped reads (%)	Nonspecifically mapped reads (%)	Unmapped reads (%)	Uniquely mapped reads (%)	Nonspecifically mapped reads (%)	Unmapped reads (%)
0	*Exp1*	D-a	15,577,476	8,251,626 (52.97)	789,570 (5.07)	6,536,280 (41.96)	14,395,262 (92.41)	976,153 (6.27)	206,061 (1.32)
		D-b	14,522,776	7,893,406 (54.35)	763,538 (5.26)	5,865,832 (40.39)	13,368,748 (92.05)	928,354 (6.39)	225,674 (1.55)
	Exp2	D-d	17,869,770	7,891,508 (44.16)	1,025,039 (5.74)	8,953,223 (50.10)	15,538,486 (86.95)	1,935,050 (10.83)	396,234 (2.22)
		D-e	10,496,694	4,253,152 (40.52)	608,307 (5.80)	5,635,235 (53.69)	8,898,126 (84.77)	1,187,113 (11.31)	411,455 (3.92)
2	*Exp1*	D-a	14,284,694	7,498,322 (52.49)	665,151 (4.66)	6,121,221 (42.85)	13,241,972 (92.70)	907,406 (6.35)	135,316 (0.95)
		D-b	10,306,088	5,409,756 (52.49)	472,237 (4.58)	4,424,095 (42.93)	9,508,874 (92.26)	701,133 (6.80)	96,081 (0.93)
	Exp2	D-d	6,170,952	2,569,350 (41.64)	310,346 (5.03)	3,291,256 (53.33)	5,411,222 (87.69)	666,573 (10.80)	93,157 (1.51)
		D-e	14,543,920	6,046,188 (41.57)	613,308 (4.22)	7,884,424 (54.21)	12,897,028 (88.68)	1,428,372 (9.82)	218,520 (1.50)
	Exp1	I-a	12,747,082	6,372,030 (49.99)	592,439 (4.65)	5,782,613 (45.36)	11,745,066 (92.14)	829,194 (6.50)	172,822 (1.36)
		I-b	22,373,358	11,230,918 (50.20)	939,664 (4.20)	10,202,776 (45.60)	20,898,294 (93.41)	1,261,462 (5.64)	213,602 (0.95)
		I-c	10,069,090	5,000,604 (49.66)	485,103 (4.82)	4,583,383 (45.52)	9,209,868 (91.47)	722,957 (7.18)	136,265 (1.35)

(Continued) Table S3. RNA-seq mapping results for *E. coli* DSM1103

	Exp2	I-d	11,430,604	4,254,842 (37.22)	431,007 (3.77)	6,744,755 (59.01)	10,089,312 (88.27)	1,116,973 (9.77)	224,319 (1.96)
		I-e	19,154,408	5,993,354 (31.29)	584,983 (3.05)	12,576,071 (65.66)	17,051,366 (89.02)	1,784,080 (9.31)	318,962 (1.67)
3.5	*Exp1*	D-a	4,281,046	2,212,340 (51.68)	228,614 (5.34)	1,840,092 (42.98)	3,872,306 (90.45)	335,730 (7.84)	73,010 (1.71)
		D-b	53,984,356	28,117,410 (52.08)	2,135,663 (3.96)	23,731,283 (43.96)	*50,412,262* (93.38)	*3,046,976* (5.64)	*525,118* (0.97)
	Exp2	D-d	16,398,470	6,487,384 (39.56)	797,994 (4.87)	9,113,092 (55.57)	15,016,286 (91.57)	1,072,106 (6.54)	310,078 (1.89)
		D-e	22,297,174	8,783,598 (39.39)	960,498 (4.31)	12,553,078 (56.30)	19,765,792 (88.65)	2,209,917 (9.91)	321,465 (1.44)
	Exp1	I-a	13,198,954	3,523,012 (26.69)	257,317 (1.95)	9,418,625 (71.36)	12,374,002 (93.75)	703,293 (5.33)	121,659 (0.92)
		I-b	20,685,618	7,046,708 (34.07)	616,399 (2.98)	13,022,511 (62.95)	19,503,692 (94.29)	963,144 (4.66)	218,782 (1.06)
		I-c	37,858,090	13,448,812 (35.52)	1,186,996 (3.14)	23,222,282 (61.34)	35,318,916 (93.29)	2,046,288 (5.41)	492,886 (1.30)
	Exp2	I-d	16,810,704	3,540,368 (21.06)	432,636 (2.57)	12,837,700 (76.37)	15,712,042 (93.46)	827,640 (4.92)	271,022 (1.61)
		I-e	21,468,104	4,605,986 (21.46)	520,553 (2.42)	16,341,565 (76.12)	18,916,324 (88.11)	2,090,585 (9.74)	461,195 (2.15)

(Continued) Table S3. RNA-seq mapping results for *E. coli* DSM1103

6.5	*Exp1*	D-a	24,198,370	12,097,824 (49.99)	954,153 (3.94)	11,146,393 (46.06)	22,261,312 (92.00)	1,580,405 (6.53)	356,653 (1.47)
		D-b	13,760,650	6,653,986 (48.36)	750,266 (5.45)	6,356,398 (46.19)	11,907,806 (86.54)	1,388,576 (10.09)	464,268 (3.37)
	Exp2	D-d	16,468,340	5,253,550 (31.90)	556,058 (3.38)	10,658,732 (64.72)	14,715,592 (89.36)	1,536,683 (9.33)	216,065 (1.31)
		D-e	19,055,320	5,904,822 (30.99)	639,872 (3.36)	12,510,626 (65.65)	17,852,460 (93.69)	944,166 (4.95)	258,694 (1.36)
	Exp1	I-a	32,314,798	8,717,558 (26.98)	684,457 (2.12)	22,912,783 (70.90)	28,639,597 (88.63)	3,195,208 (9.89)	479,994 (1.49)
		I-b	42,969,260	9,202,048 (21.42)	814,705 (1.90)	32,952,507 (76.69)	38,220,086 (88.95)	3,987,918 (9.28)	761,256 (1.77)
		I-c	8,107,936	1,951,550 (24.07)	178,994 (2.21)	5,977,392 (73.72)	7,015,762 (86.53)	858,839 (10.59)	233,335 (2.88)
	Exp2	I-d	17,717,024	2,961,582 (16.72)	456,468 (2.58)	14,298,974 (80.71)	16,558,142 (93.46)	846,220 (4.78)	312,662 (1.76)
		I-e	7,723,678	983,154 (12.73)	168,367 (2.18)	6,572,157 (85.09)	7,404,144 (95.86)	194,526 (2.52)	125,008 (1.62)

Time (h) denotes exposure to irradiance durations corresponding to the decay phases as follows: 0 h = prior to solar irradiation; 2 h = early decay phase; 3.5 h = mid-decay phase; 6.5 h = late-decay phase.
D/I: Microcosm samples are labeled as D- or I- to denote dark and irradiated microcosms, respectively.
CDS: coding sequence.
[a] All reads are 101 nucleotides long.
[b] Reads were mapped to DSM 1103 reference genome using CLC Genomics Workbench version 8.0.1 with a minimum length fraction of 0.9, a minimum similarity fraction of 0.8, and a maximum number of hits for a read of 10

Table S4. List of primers used for RT-qPCR in this study

Primer	Gene target	Gene function	Amplicon size (bp)	Sequence	Cycling conditions*	Amplification efficiency*
FtsI146-F FtsI146-R	*fts*I	Cell-division protein/peptidoglycan synthesis	146	5`-GCGTTTTGCGTTGTTATGCG-3` 5`-GGGAGGTGGAAACTTGCTCA-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 55 °C x 20 s	99%
HflC154-F HflC154-R	*hfl*C	Modulator of ftsH protease/control of lysogenization	154	5`-TCAACCCGAACAGTATGGCG-3` 5`-TCCTGACCTTGTGAACGGTG-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 55 °C x 20 s	92%
LexA145-F LexA145-R	*lex*A	Transcriptional repressor of SOS response genes	145	5`-GGGATTCGTCTGTTGCAGGA-3` 5`-GCAGGAAATCAGCATTCGGC-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 55 °C x 20 s	93%
MarA127-F MarA127-R	*mar*A	Transcriptional activator /multiple antibiotic resistance protein	127	5`-GACAACCTGGAATCGCCACT-3` 5`-TACGGCTGCGGATGTATTGG-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 55 °C x 20 s	98%
CFA157-F CFA157-R	CFA	Colonization Factor Antigen	157	5`-TCATCACGGTATCGCCAGTT-3` 5`-GTGGCTATCGAGGGTGACTC-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 55 °C x 20 s	93%
UidA159-F UidA159-R	*uid*A	Housekeeping gene beta-D-glucuronidase	159	5`-CGAATCCTTTGCCACGCAAG-3` 5`-TCACAGCCAAAAGCCAGACA-3`	50 °C x 2 min; 95 °C x 20 s; 40 cycles of 95 °C x 1 s and 60 °C x 20 s	93%

Probe	Gene target	Sequence	Cycling conditions*	Amplification efficiency*
uidA-23	*uid*A	5`-/56FAM/TCGCCCTCC/ZEN/ACTGCCACTGACCG/3IABkFQ/-3`	As described for uidA159 primer pair	

*: Cycling conditions and amplification efficiency for qPCR

Table S5. Wastewater quality parameters

Quality Parameter	Influent	Effluent	Chlorinated effluent
pH	-	7.54 ± 0.20	7.57 ± 0.29
COD (mg/L)	137.50 ± 28.28	13.03 ± 2.33	7.57 ± 3.66
TOC (mg/L)	6.86 ± 3.57	2.16 ± 0.26	3.08 ± 1.31
TN (mg/L)	12.87 ± 4.40	17.86 ± 0.81	14.37 ± 1.63
SUVA (L/mg-M)	1.18	4.36 ± 1.78	1.86 ± 1.63
Alkalinity as mg CaCO3/L	28.89	69.25 ± 7.66	81.87 ± 4.17
Free Cl residual (mg/L)	-	-	0.24 ± 0.06
Total Cl residual (mg/L)	-	-	0.36 ± 0.09

COD: chemical oxygen demand; TOC: total organic carbon; TN: total nitrogen; SUVA: specific ultra-violet absorbance; Cl: chlorine; C. Effluent: chlorinated effluent.

Table S6. Inactivation kinetics of *E. coli* PI-7 and DSM1103 in treated wastewater

(A) PI-7 inactivation		Effluent	C. Effluent
Dark[α]	Sample size	6	5
	Decay constant *k**	0.02 ± 0.10	-0.05 ± 0.10
Irradiated	Sample size	6	7
	Lag phase length (h)	6.64 ± 0.48	6.30 ± 1.18
	Lag phase fluence (J/cm^2)	184.97 ± 13.37	175.50 ± 32.87
	Decay constant *k**	-10.42 ± 4.00	-8.38 ± 2.00
	Half-life length (min)	4.36 ± 1.13	5.23 ± 1.40
	Half-life fluence (J/cm^2)	2.02 ± 0.52	2.43 ± 0.65
(B) DSM1103 inactivation		Effluent	C. Effluent
Dark[α]	Sample size	4	5
	Decay constant *k**	0.00 ± 0.06	-0.06 ± 0.14
Irradiated	Sample size	7	7
	Lag phase length (h)	1.29 ± 0.49	0.71 ± 0.76
	Lag phase fluence (J/cm^2)	35.94 ± 13.65	19.78 ± 21.17
	Decay constant *k**	-10.83 ± 2.32	-18.83 ± 6.23
	Half-life length (min)	4.02 ± 0.96	2.40 ± 0.78
	Half-life fluence (J/cm^2)	1.87 ± 0.45	1.11 ± 0.36

[α]Lag phase length and fluence, and half-life length and fluence were not calculated for dark controls.
*Decay constants (negative slopes of natural log inactivation over time curves) were mathematically corrected for light attenuation. Each set of inactivation data is based on at least three biological replicates.

Table S7. Numbers of unique metabolic functions in the genomes of *E. coli* PI-7 and DSM1103 in the various functional categories

Functional Category	Core	PI-7	DSM1103
Amino Acids and Derivatives	307	1	6
Carbohydrates	571	15	43
Cell Division and Cell Cycle	37	0	0
Cell Wall and Capsule	224	6	9
Clustering-Based Subsystems	620	15	21
Cofactors, Vitamins, Prosthetic Groups, Pigments	252	2	2
DNA Metabolism	109	4	1
Dormancy and Sporulation	4	0	0
Fatty Acids, Lipids, and Isoprenoids	96	0	19
Iron acquisition and metabolism	23	1	26
Membrane Transport	141	35	21
Metabolism of Aromatic Compounds	3	17	0
Miscellaneous	51	6	0
Motility and Chemotaxis	72	2	0
Nitrogen Metabolism	52	4	0
Nucleosides and Nucleotides	133	0	1
Phages, Prophages, Transposable elements, Plasmids	7	5	19
Phosphorus Metabolism	42	0	0
Potassium metabolism	24	0	0
Protein Metabolism	263	1	1
Regulation and Cell signaling			
Programmed Cell Death and Toxin-antitoxin Systems	13	6	7
Quorum Sensing and Biofilm Formation	4	8	0
Regulation of Cell Virulence	8	0	0
No Subcategory	94	2	3
Regulons	11	0	0
Respiration	152	6	2
RNA Metabolism	237	0	1
Secondary Metabolism	5	8	0
Stress Response	163	2	1
Sulfur Metabolism	54	0	1
Virulence, Disease and Defense	88	7	1

Numbers were obtained through comparisons performed using RAST [26, 27].
"Core" column refers to functions that were found to be shared by the two *E. coli* isolates.
0 = no unique functions found for that bacterium in that particular functional category.

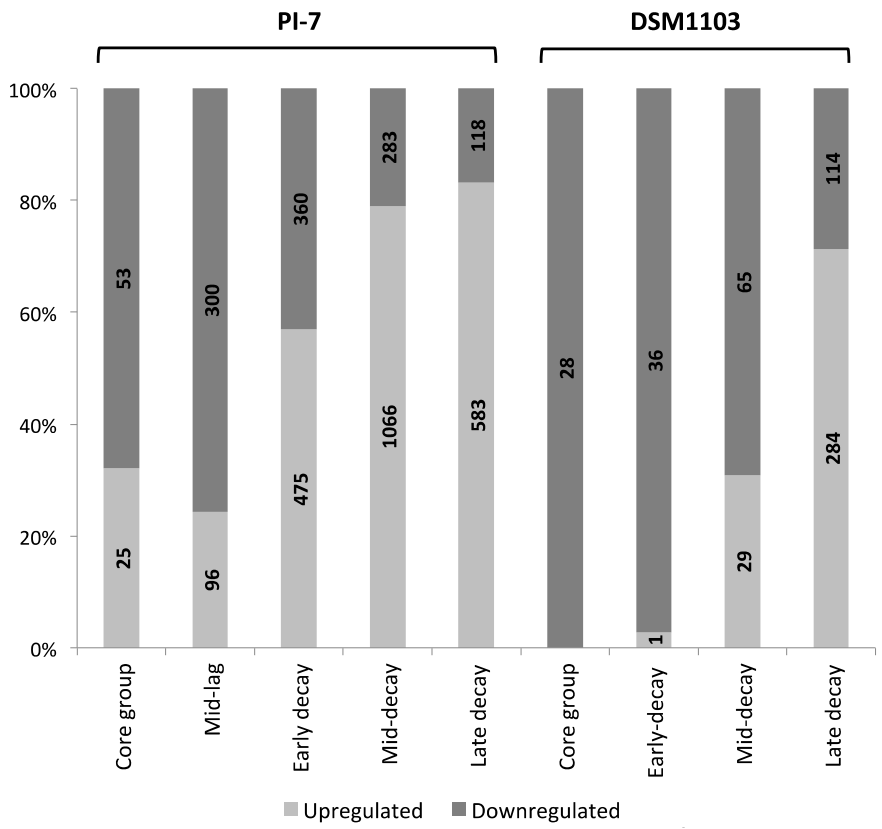
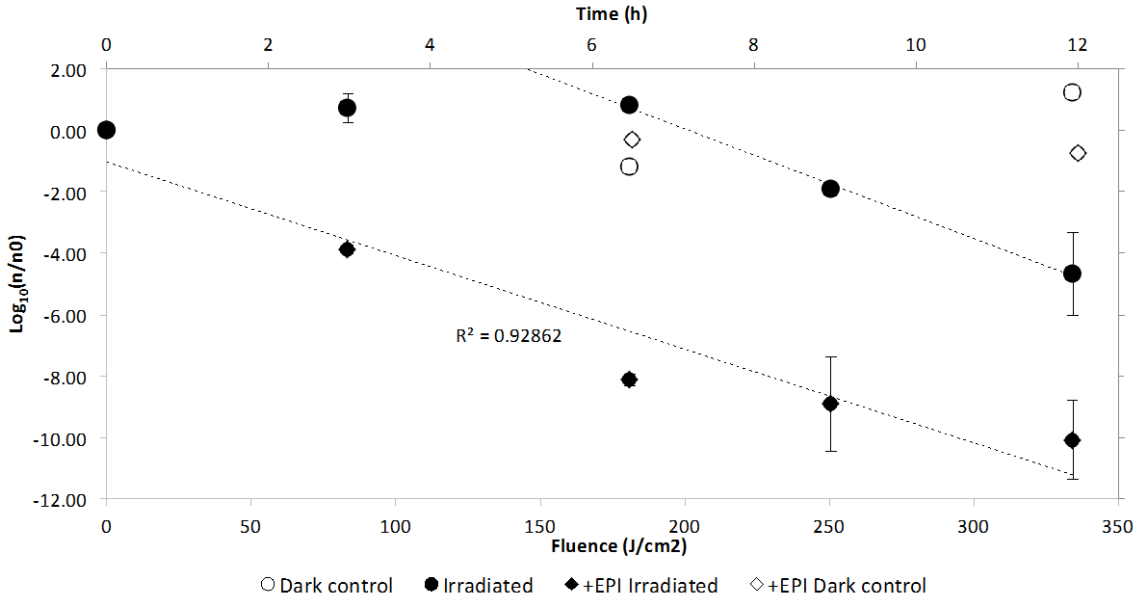


Figure S1. Percentage of genes significantly up- and down-regulated for each of *E. coli* DSM1103 and PI-7 at different phases of exposure to simulated solar irradiance in buffer solution

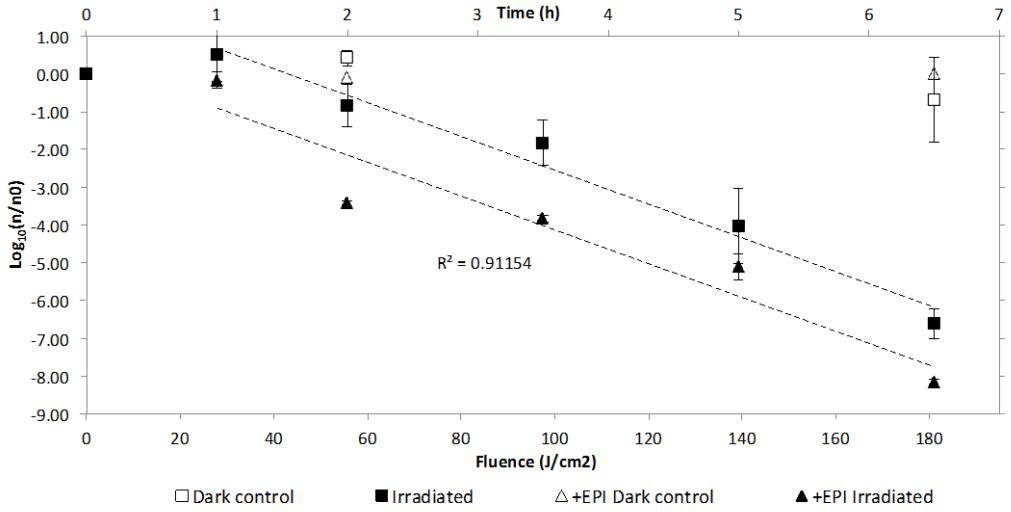
The normalized bar lengths indicate the proportion of genes that were either upregulated (light) or downregulated (dark), for the indicated *E. coli*, out of the total number of differentially regulated genes for that particular *E. coli* for the indicated phase only.
Numbers shown inside the bars refer to the numbers of genes differentially regulated in the corresponding direction.
"Core group" category refers to genes that were significantly differentially regulated (by ≥2-fold) in the same direction consistently at all phases.
Decay phases for PI-7 are as follows: 0 h = prior to solar irradiation; 3 h = mid-lag phase; 6 h = early-decay phase; 9 h = mid-decay phase; 14.5 = late-decay phase.
Decay phases for DSM1103 are as follows: 0 h = prior to solar irradiation; 2 h = early decay phase; 3.5 h = mid-decay phase; 6.5 h = late-decay phase.



	Lag phase length (h)	Lag phase fluence (J/cm²)	Decay constant	Half-life length (min)	Half-life fluence (J/cm²)
PI-7 (n=7)	6.64 ± 0.63	184.96 ± 17.55	-14.96 ± 2.61	2.85 ± 0.46	1.32 ± 0.21
PI-7 with EPI (n=4)	< 3	< 85	-23.55 ± 0.80	1.77 ± 0.06	0.82 ± 0.03

Figure S2. Inactivation curves of *E. coli* PI-7 in PBS in absence or presence of efflux pump inhibitor PAβN

EPI: 50 μg/mL efflux pump inhibitor.
E. coli PI-7 with EPI showed reduced lag phase and half-life compared with PI-7 without EPI.
One-way ANOVA analysis on log-reduction in dark samples with and without the EPI showed that 50 μg/mL of PAβN had no toxic effect on the bacteria (p = 0.9200).



	Lag phase length (h)	Lag phase fluence (J/cm²)	Decay constant	Half-life length (min)	Half-life fluence (J/cm²)
DSM1103 (n=6)	1.33 ± 0.52	37.05 ± 14.49	-21.01 ± 4.54	2.04 ± 0.36	0.95 ± 0.17
DSM1103 with EPI (n=3)	1.00 ± 0	27.86 ± 0	-30.23 ± 0.21	1.38 ± 0.11	0.64 ± 0.05

Figure S3. Inactivation curves of *E. coli* DSM1103 in PBS in absence or presence of efflux pump inhibitor PAβN

EPI: 50 µg/mL efflux pump inhibitor.
E. coli DSM1103 with EPI showed no significant change in lag phase length but a reduced half life compared with DSM1103 without EPI.
One-way ANOVA analysis on log-reduction in dark samples with and without the EPI showed that 50 µg/mL of PAβN had no toxic effect on the bacteria (p = 0.2124).

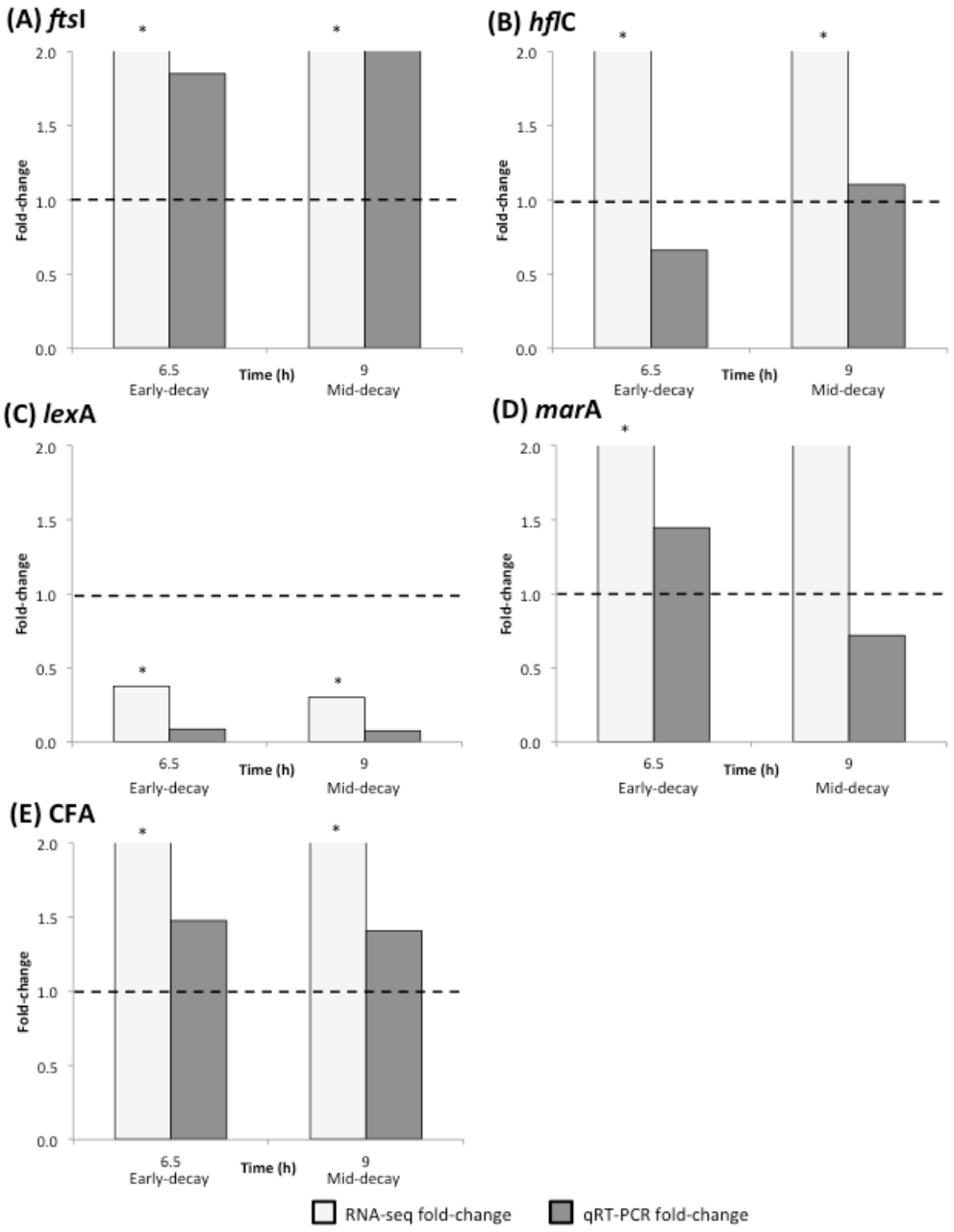


Figure S4. RNA-seq and RT-qPCR fold-change in *E. coli* PI-7 at the early-decay and mid-decay phases

RNA-seq fold change was calculated as explained in Supplementary Information 5. RT-qPCR fold-change values were generated as explained in Supplementary Information 8. Dotted line at fold-change =1 indicates where expression values (RNA-seq) or proportion to housekeeping gene uidA (RT-qPCR) are equal between irradiated and dark samples. Values above the line indicate upregulation, while those below indicate downregulation. Asterisks (*) denote significant differential expression by 2-fold for RNA-seq data.

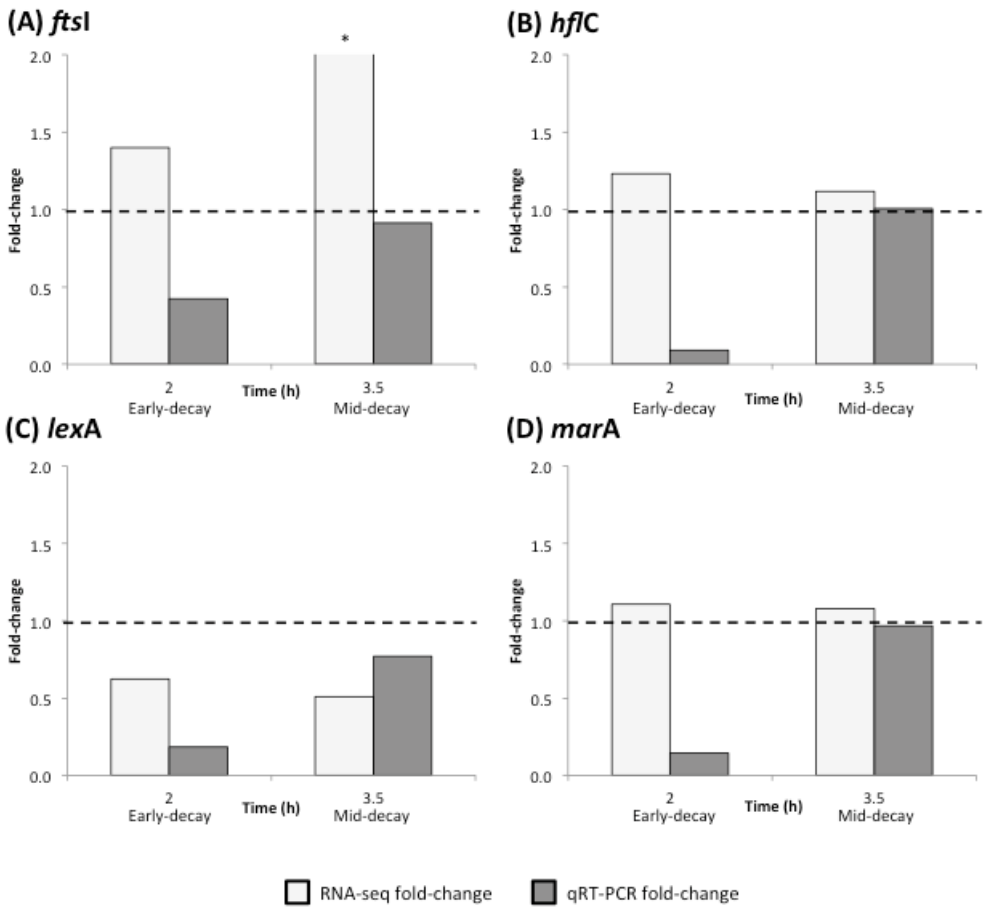


Figure S5. RNA-seq and RT-qPCR fold-change in *E. coli* DSM1103 at the early-decay and mid-decay phases

RNA-seq fold change was calculated as explained in Supplementary Information 5. RT-qPCR fold-change values were generated as explained in Supplementary Information 8. Dotted line at fold-change =1 indicates where expression values (RNA-seq) or proportion to housekeeping gene uidA (RT-qPCR) are equal between irradiated and dark samples. Values above the line indicate upregulation, while those below indicate downregulation. Asterisks (*) denote significant differential expression by 2-fold for RNA-seq data.