NIH Public Access
Author Manuscript
Genomics. Author manuscript; available in PMC 2010 January 1.

Published in final edited form as:
Genomics. 2009 January ; 93(1): 42–51. doi:10.1016/j.ygeno.2008.09.001.

Systems based mapping demonstrates that recovery from alkylation damage requires DNA repair, RNA processing, and translation associated networks

John P. Rooney[a,b], **Ajish D. George**[a,b], **Ashish Patil**[a,b], **Ulrike Begley**[a,b], **Erin Bessette**[a,b], **Maria R. Zappala**[b,c], **Xin Huang**[b,c], **Douglas A. Conklin**[a,b], **Richard P. Cunningham**[b,c], and **Thomas J. Begley**[a,b,§]

a *Department of Biomedical Sciences, University at Albany, State University of New York, Rensselaer NY 12144-3456*

b *Gen*NY*Sis Center for Excellence in Cancer Genomics, University at Albany, State University of New York, Rensselaer NY 12144-3456*

c *Department of Biological Sciences, University at Albany, State University of New York, Albany NY 12222*

Abstract

The identification of cellular responses to damage can promote mechanistic insight into stress signalling. We have screened a library of 3,968 *E. coli* gene deletion mutants to identify 99 gene products that modulate the toxicity of the alkylating agent methyl methanesulfonate (MMS). We have developed an ontology mapping approach to identify functional categories over-represented with MMS-toxicity modulating proteins and demonstrate that, in addition to DNA re-synthesis (replication, recombination, and repair), proteins involved in mRNA processing and translation influence viability after MMS damage. We have also mapped our MMS-toxicity modulating proteins onto an *E. coli* protein interactome and identified a sub-network consisting of 32 proteins functioning in DNA repair, mRNA processing, and translation. Clustering coefficient analysis identified seven highly connected MMS-toxicity modulating proteins associated with translation and mRNA processing, with the high connectivity suggestive of a coordinated response. Corresponding results from reporter assays support the idea that the SOS response is influenced by activities associated with the mRNA-translation interface.

Keywords

Alkylation; *E. coli*; DNA damage; DNA repair; Gene deletion; Genomic phenotyping; Interactome mapping; mRNA processing; Systems biology; Translation

Introduction

Alkylating agents are electrophilic compounds that can modify cellular macromolecules and thereby initiate disease. Damage to both nucleic acids and proteins can occur after exposure to alkylating agents with alkylation of adenine, guanine, and cytosine bases in nucleic acids,

§Corresponding author: Thomas J. Begley, 518 591 7153 (phone), 518 591 7201 (fax), tbegley@albany.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and arginine, lysine, and cysteine residues in proteins as common sites of damage [1–8]. The alkylating agent methyl methanesulfonate (MMS) has been frequently employed to mimic the effects of both endogenous and environmental alkylating agents. MMS is an alkylating agent that damages both nucleic acids and proteins, thereby promoting mutagenesis and cell death [8,9]. MMS has proven to be a valuable tool for characterizing cellular damage-response machinery [10–13].

Cellular responses to alkylation damage play an important role in preventing mutations and cell death [14–17]. A number of enzyme activities from bacteria and mammals have been identified that modulate the toxicity and mutagenicity of alkylating agents [18–27], and some DNA repair proteins are examples of a conserved response to alkylation damage [28–33]. The alkylbase DNA glycosylases from *E. coli* (AlkA) and the mouse (Aag) are a case in point; inactivation of either renders cells sensitive to killing by alkylating agents [29]. A more recent example is the direct repair enzyme AlkB, which was initially identified in *E. coli* as an activity that modulates the toxicity of MMS [24]. AlkB repairs single and double stranded DNA and RNA lesions caused by MMS (1-methyladenine and 3-methylcytosine). Similar activities have been characterized in the mouse and in humans [30–32,34], and have been shown to complement the MMS sensitive phenotype of *alkB* deficient *E. coli*. The functional conservation of base excision and direct repair proteins between *E. coli* and mammals supports the concept that cells use common mechanisms to repair damage caused by alkylating agents.

Mechanistic studies in *E. coli* have previously demonstrated that components of the adaptive response and the SOS response are activated after MMS induced DNA damage [35–37]. Signalling proteins that initiate the adaptive and SOS responses are Ada and RecA respectively; both proteins recognize DNA damage and initiate downstream signalling to promote repair. Different types of DNA damage are detected by Ada and RecA, and activation of each protein will initiate transcriptional responses that facilitate cell survival after alkylation damage. Transcriptional reprogramming and increased repair in response to alkylation damage are well conserved themes across phylogeny [12,38–40].

We describe here a global study using a library of 3,968 unique *E. coli* gene deletion mutants to identify activities that prevent cell death after treatment with the alkylating agent MMS. We show that at least 99 different protein activities are important for preventing MMS-induced cell death. Functional and computational mapping of the MMS-toxicity modulating gene products identified protein networks specific to DNA repair, transcription, mRNA processing, and translation as being important after alkylation damage. Similarly, validation experiments that use newly constructed gene deletion mutants in cell killing assays and SOS-reporter assays demonstrate that cellular processes that promote the re-synthesis of DNA and proteins are essential for cell survival. Our results support the hypothesis that specific translational and mRNA processing activities, which are conserved from *E. coli* to humans, are utilized during the response to MMS damage.

Results and Discussion

Toxicity modulating genes identified by genomic phenotyping

We used a robotic plate-based screen of *E. coli* gene deletion mutants to identify genes and their associated proteins that modulate toxicity to MMS. *E. coli* gene deletion mutants were from the Keio library [41], which was generated using a targeted homologous recombination strategy and which consists of 8,640 mutants, with at least two independent isolates of each gene knockout represented in the library. The library we tested represented 3,968 *E. coli* genes and provided approximately 93% coverage of the genome. Mutants were grown to saturation in 96-well plates and 1 μl aliquots of a 1:10 dilution of the cell suspensions were robotically transferred onto agar plates containing two concentrations of MMS. Approximately 360 agar

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

plates, with 34,560 spotted cultures, were incubated overnight at 37°C and then digitally imaged for analysis. Images of each plate were compiled to create a data base (Supplemental Figure S1: Supplemental Table 1–2) and sensitive mutants were visually identified (Figure 1: Table 1). A virtual mutant representing at least two isolates of each gene deletion mutant in the library was given a MMS toxicity modulating score, which is based on the behaviour of all corresponding deletion mutants on two plates containing MMS (0.045 and 0.06% MMS). For example there were two Δ*alkA* mutants in the library, and a virtual mutant representing Δ*alkA* has a compiled MMS toxicity modulating score describing the behaviour of both Δ*alkA* mutants on two concentrations of MMS. The MMS-toxicity modulating score is a semi-quantitative measure of the sensitivity of a virtual mutant after MMS treatment, and consists of values from two concentrations of MMS for two independent mutants specific to each gene-product. Deletion mutants with reduced growth on MMS were given a score of 2, by our convention those with a color change from white to dark grey are associated with a growth defect and were scored 1 (See Figure 1B), while those showing unaffected growth were scored 0. In theory the virtual mutants most sensitive to MMS were scored as an eight (2 + 2 + 2 + 2), because two corresponding isolates were sensitive to both concentrations of MMS. We used a minimum cut-off of three to identify virtual mutants that were sensitive to MMS.

In all, we identified 99 virtual mutants that scored three or greater and that were classified as MMS sensitive. We also independently constructed 96 gene deletion mutants (Rooney et al., in press) which recapitulated some of the MMS-sensitive gene-deletion mutants. As a control, we also constructed gene deletion mutants that do not affect MMS sensitivity, as identified in our present study. We performed MMS sensitivity testing on these newly derived gene-deletion mutants (Supplemental Table S3) and determined that 90% displayed a phenotype similar to those virtual mutants derived from the Keio library, indicating that the Keio library is of high quality and that our results are highly reproducible.

We next catalogued the 99 virtual mutants by assigning the proteins corresponding to each catalogued deleted gene and assembled a list of 99 MMS-toxicity modulating proteins. We analyzed the type of protein activities important after MMS damage, using information supplied by the Ecogene database [42]. As expected, a number of previously identified DNA alkylation repair and recombination proteins were identified in our set of 99 MMS-toxicity modulating proteins, including Ada, AlkA, AlkB, and RecA. In all, we identified eleven different DNA repair/recombination activities that modulated the toxicity of MMS. These activities represented components of direct repair (DR), base excision repair (BER), mismatch repair (MMR) and recombinational repair (RR) (Ada, AlkA, AlkB, RecA, RecC, RuvA, RecB, RuvB, Dam, RecO, and RecN). In addition, four DNA replication proteins were identified (DnaT, PriA, DnaG, and ParC). The identification of DNA repair and DNA replication proteins confirms the model that the re-synthesis of damaged DNA is a priority after MMS exposure. Similar screens in budding yeast [10,11] also identified DR, BER, MMR, RR, and replication proteins as modulators of MMS-toxicity.

In addition to DNA repair, recombination, and replication proteins, we identified activities specific to transcriptional regulation, protein damage, and protein synthesis in our MMS screen. Transcriptional components important after MMS damage include members of the adaptive and SOS response (Ada and RecA). In addition, we found that previously identified and predicted transcriptional regulatory proteins (ArcA, Fis, CadC, UidR, and OxyR) modulated the toxicity of MMS. These proteins are associated with the regulation of a large (ArcA and Fis) or small number (CadC, UidR, and OxyR) of downstream targets, indicating that both global and specific transcriptional regulators were identified in our screen. ArcA can positively or negatively regulate transcription of aerobic enzymes and has been demonstrated to control the resistance of *E. coli* to specific dyes [43], presumably through regulation of envelope proteins. In theory, cells deficient in ArcA could contain higher levels of intracellular MMS

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and thus be susceptible to more DNA damage. The global regulator Fis was also identified as a protein that modulated MMS-toxicity, which is a novel observation, and Fis plays roles in the transcriptional activation of rRNA genes, site specific DNA inversion, and repression of DNA replication [44]. The precise role of Fis after DNA damage has yet to be determined, but Fis has great potential to influence DNA and protein metabolism after alkylation damage. CadC is a transcriptional activator for other cadaverin associated gene products (cadA and cadB) and is know to sense external stimuli associated with low pH and low oxygen [42], and we can speculate that this transcriptional activator can sense environmental conditions associated with MMS in the medium. Similarly the classification of the sensor for oxidative stress OxyR [42] as MMS toxicity modulating suggests that MMS damage can alter the levels of reactive oxygen species inside the cell.

A third category of proteins that modulates the toxicity of MMS included those specific to protein maintenance, protein stabilization, and translation. The Hsp70 chaperone protein DnaK was identified in our screen, suggesting that DnaK plays a role in stabilizing MMS modified proteins. DnaK has also been shown to play a regulatory role in DNA replication [45], which could account for the MMS sensitive phenotype of the corresponding deletion mutant. Eleven different activities specific to translation were identified as modulating MMS sensitivity. Basic ribosome machinery that includes major components of the 30S and 50S ribosome subunits (RplA, RplY, RpsO, RpsU and RpsT) were found to modulate cellular viability after MMS damage, with a *ΔrplA* mutant as MMS sensitive as the DNA alkylation repair deficient mutant *Δada*. In addition, translational components that affect ribosome activity, including ribosome binding (RbfA), peptide chain release (PrfB), trans-translation (SmbP), tRNA synthesis (PoxA), and tRNA modification (GidA/MnmG and YfcK/MnmC) enzymes, were classified as important after MMS treatment. The prominence of protein synthesis machinery in our list of MMS-toxicity modulating proteins suggests an important role for the cellular translational apparatus in the damage response.

We have also assayed all mutants corresponding to the 99 MMS modulating proteins for sensitivity to UV irradiation (254 nm) and the oxidizing agent tert-butyl hydroperoxide (t-BuOOH) (Supplemental Table S4). We have found that 40 of these mutants are sensitive to UV, 32 mutants are sensitive to t-BuOOH), with 21 sensitive to both UV and t-BuOOH (*ΔdnaT, ΔpriA, ΔrecA, ΔrecC, ΔruvC, ΔrbfA, Δfis, ΔruvA, ΔrecO, Δhfq, ΔrpsU, ΔrpsO, ΔrplA, Δrnt, ΔholC, ΔmtlA, ΔrecB, Δ rpsT, ΔyciF, ΔprfB and Δpnp*). MMS, UV, and t-BuOOH are all prototypical damaging agents that have been used to study DNA repair proficiency *in vivo* and sensitivity to all three agents suggests corresponding mutants are defective in the DNA damage response. Interestingly Fis, Hfq, MtlA, Pnp, PrfB, RbfA, Rnt, RplA, RpsO, RpsT, RpsU and YciF have no known association with DNA repair and the sensitivity of the corresponding mutants to three classic DNA damaging agents suggests that these transcription, RNA processing, and translation associated activities have an unknown yet important role in the DNA damage response.

Functional mapping identifies responses to alkylation damage

Trends observed in high throughput studies require statistical validation, and to do this we developed a functional mapping algorithm using GenProtEC protein annotation information [44]. The goal was to identify functional categories which have a significant over-representation of MMS-toxicity modulating proteins. GenProtEC protein annotation information represents a hierarchical classification of proteins which systematically describe a protein's role and biochemical function. We downloaded all of the *E. coli* annotations for our analysis. We used an intermediate level of the hierarchical protein information consisting of 130 functional categories, with some proteins classified in multiple functional categories. We cross referenced the GeneProtEC annotation entries with proteins corresponding to mutants

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

used in our screen, and identified 2,713 intersecting proteins, of which 73 modulated the toxicity of MMS (Supplemental Table S5). Those *E. coli* proteins that are not annotated in GeneProtEC were not used for functional mapping. Next we determined the number of MMS-toxicity modulating proteins found in each functional category. We also performed 10,000 random samplings of 73 proteins from our master list of 2,713 proteins, noted the number found in each functional category, and determined the average number and standard deviation for each functional category. The actual number of MMS toxicity modulating proteins, the random sampled average, and the standard deviation for each functional category were then used to determine whether there was significant enrichment for MMS-toxicity modulating proteins in each functional category (Table 2). We identified 12 different functional categories that were over-represented with MMS-toxicity modulating proteins (number in category ≥ 2, $p < 0.05$). All gene-deletion mutants whose corresponding proteins mapped to significantly over-represented categories were further validated using dilution based cytotoxicity assays (Figure 2; Supplemental Figure S2). All were verified as being MMS sensitive.

Some of the functional categories over-represented with MMS-toxicity modulating proteins included DNA recombination, DNA repair, DNA replication, DNA degradation, DNA structure, and RNA degradation, which cumulatively demonstrate a requirement for nucleic acid metabolism and genome maintenance after MMS damage. DNA damage is a known product of MMS exposure, thus these categories were expected and serve as a control to validate our algorithm. In addition to DNA damage, RNA has been shown to be a target for MMS. RNA processing activities that include Pnp and Rnt and the DNA/RNA demethylation enzyme AlkB were identified in our functional analysis [42]. Pnp is a component of the RNA degradosome and will hydrolyze mRNA to remove it from the transcript pool [46], while Rnt encodes a tRNA ribonuclease that plays a role in recycling uncharged adapter molecules [47]. Recently the DNA repair enzyme AlkB was shown to remove methyl groups from RNA, with methyl groups representing both damage and enzyme based modifications in tRNA. While *ΔalkB* cells are sensitive to MMS, the contribution of RNA damage or RNA modifications to cell death or viability after damage is unclear. Similarly, the precise role of Pnp and Rnt after MMS damage is unclear, as either could be used to remove damaged RNA or process RNA for signalling purposes after damage. Nonetheless, the significant theme of RNA degradation after MMS damage suggests that the removal and repair of RNA or RNA processing for signal transduction purposes play an important role after DNA damage.

Xenobiotic metabolism and stress signalling are common responses to exposure and damage and the removal of damaging agents and the increased transcription or post translational modification of downstream proteins are recognized damage responses observed across phylogeny. We identified both xenobiotic metabolism (AhpC, CpxA, and CynR) and stress based (ArcA and OxyR) regulatory proteins as being over-represented amongst the MMS-toxicity modulating proteins, which was expected and served as another validation for our algorithm.

One of the most prominent categories identified by functional mapping was specific to protein synthesis, and included ribosomal proteins and translational machinery. Six activities specific to protein synthesis were represented in our functional mapping results (PrfB, RplA, RpsO, RpsT, RpsU, RplY) and the MMS sensitivity of the corresponding mutants supports our hypothesis that translational machinery plays an important role in recovery from damage. Clearly, protein synthesis machinery responds to transcriptional cues and is involved in the synthesis of important toxicity modulating proteins. It is known that reactive electrophiles will damage proteins and the replacement of damaged proteins is certainly an important cellular activity after MMS damage [5,7]. Additionally, we can speculate that ribosomal proteins sense cellular stress or that translational regulation occurs after damage, but these roles for ribosomal and protein synthesis machinery in the damage response have yet to be proven.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Interactome analysis identifies vital MMS-toxicity modulating networks

Species-specific protein interaction information has been demonstrated to be an effective tool for analyzing global data sets [48–54], and can assist in the identification of protein networks activated by damage. Protein-protein interaction information can be compiled to generate an interactome *in silico*, and the resulting structure is a static blueprint of potential signalling pathways and protein complexes inside the cell. We have compiled all reported protein-protein interactions for *E. coli* available in the Database of Interacting Proteins (DIP) [55] and supplemented them with a large protein-protein interaction study [56]. In all we compiled 18,161 interactions between 3,467 *E. coli* proteins (Supplemental Table S6). The compiled interactome can be considered a non saturated structure with regard to molecular interactions, but it is an extensive framework that can be used to identify protein networks activated by MMS damage.

The compiled interactome was mapped with MMS-toxicity modulating proteins, and filtered to show only MMS-toxicity modulating proteins and their corresponding protein-protein interactions. Next we colored each node according to cellular function, to demonstrate that activities involved in DNA metabolism (repair, recombination, and replication), transcription, translation, and mRNA processing, among others, are found in the filtered interactome (Figure 3, top). Further analysis of the filtered interactome identified a large connected component of 32 proteins along with two two-protein modules. The 32 protein sub-network identified by interactome filtering is connected by 37 protein-protein interactions. It should be noted that interactome filtering does not use statistical validation to assign p-values to sub-networks. Instead, the filtering step identifies all connected MMS-toxicity modulating proteins in the interactome to provide a global view of how different functional activities are potentially coordinated.

We used clustering coefficient analysis of the mapped and filtered structure to identify highly connected and statistically significant protein architectures that respond to MMS treatment. In general, interactome mapping has been shown to identify biologically important architectures [49,50]. Clustering coefficient analysis of a mapped interactome has been demonstrated to identify signatures of protein pathways and complexes responding to damage and can identify local areas of high connectivity in a mapped network [10,11,57]. We analyzed our 32 protein sub-network using clustering coefficient analysis, to identify proteins whose interacting neighbors share protein-protein interactions. We identified a significantly clustered group of seven proteins (Hfq, Pnp, RplA, RpsO, DeaD, ParC, and SmpB) centred on RplA and connected by 13 protein-protein interactions (Figure 3, bottom). The significance ($p < 10^{-6}$) of the highly clustered RplA-centred subnetwork was validated by network randomizations and random samplings. The 50S subunit protein RplA is the focal point of a highly clustered subnetwork that contains components of the 30S and 50S ribosome (RpsO and RplA) and RNA processing activities (Hfq, Pnp, and SmpB) [42]. The 30S ribosome protein RpsT was also closely associated with the RplA-centred sub-network, via interactions with Pnp and DeaD, and was added to the subnetwork based on functional overlap with other members. The RplA-centred sub-network of eight MMS-toxicity modulating proteins is suggestive of a coordinated pathway specific to mRNA processing and protein synthesis machinery. Importantly, these activities have been identified again as being important after MMS exposure, albeit using a different mapping approach based on protein-protein interactions.

Defective SOS Responses in the RplA-Focused Sub-Network

The SOS response to DNA damage caused by MMS is one of the major response pathways promoting cell viability after DNA alkylation damage. Defects in the SOS response could be responsible for the MMS-sensitive phenotype of gene deletion mutants specific to members of the RplA-centred sub-network. Thus we analyzed the induction of the SOS response in each

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

gene deletion mutant (Figure 4) specific to the RplA-centred sub-network, along with wild-type and *ΔrecA* control strains. We recorded a ~4-fold induction in the SOS-response in wild-type cells treated with MMS, relative to untreated wild-type, and no SOS induction in *ΔrecA* cells after MMS treatment, indicating our assay was working properly. Next we looked at the basal levels of the SOS reporter in each of our cell types and determined that *ΔrpsT* cells had a ~4-fold induction, relative to untreated wild-type, with this induction similar to what was observed for wild-type cells treated with MMS. Clearly, *ΔrpsT* cells have a hyper active SOS response under normal conditions suggesting a faulty DNA damage response in these cells. Next we analyzed the MMS induced levels of the SOS response for each of the gene-deletion mutants specific to the eight proteins found in the RplA-centred subnetwork. We determined that six of the gene-deletion mutants have a modest (*ΔdeaD, Δhfq, Δpnp*) or slight (*ΔrplA, ΔrpsO, ΔsmpB*) decrease in their MMS induced SOS response, relative to wild-type. Based on fold-change there appears to be two groups of SOS corrupted gene deletion mutants, with the level of the SOS-response in *ΔdeaD, Δhfq, Δpnp* mutants about 63% of wild-type, while the level of the SOS-response in *ΔrplA, ΔrpsO, ΔsmpB* mutants about 80% of wild-type. The decreased level of the MMS-induced SOS response for *ΔdeaD, Δhfq, Δpnp, ΔrplA, ΔrpsO, ΔsmpB* could explain the MMS sensitive phenotype for each mutant, due to decreased DNA repair capacity. The precise roles of DeaD, Hfq, Pnp, RplA, RpsO, and SmpB in the SOS response are unknown. DeaD, Hfq, Pnp, and SmpB are activities associated with translation and the metabolism of RNA, while RplA and RpsO are part of the ribosome. In theory there could be specific roles for each in damage-induced transcription or translation, with deficiencies in each activity directly or indirectly affecting the SOS response. The finding that they are all highly connected suggests a coordinated role of at least six activities. It is interesting to note that Pnp is directly connected to RecA by one interaction in our filtered network, and this network could be passing RNA or protein damage signals to RecA or influencing the levels of SOS machinery; however these conclusions are highly speculative.

Human Counterparts of DeaD, Hfq, Pnp, and SmpB

The RplA-centred network identified in this study contains six MMS-toxicity modulating proteins required for an optimal SOS response. One of the reasons we performed an MMS screen was our desire to identify human proteins with the potential to modulate the toxicity of alkylating agents. We used the Basic Local Alignment Search Tool (BLAST) to determine if similar proteins were found in humans (Table 3) and identified highly similar amino acid sequences ($E < 10^{-19}$) related to *E. coli* DeaD, Hfq, Pnp, and SmpB. Corresponding human activities are the translation initiation factor 4A, isoform 3 (EIF4A3), delta 2-isopentenyl adenosine tRNA-like protein (AAM13690), polyribonucleotide nucleotidyltransferase 1 (PNPT1), and a protein of unknown function (Z22851). Both the translation initiation factor 4A, isoform 3 (EIF4A3) and delta 2-isopentenyl adenosine tRNA-like protein (AAM13690) should be associated with translation, but to date these are predicted activities based on homology. Human polyribonucleotide nucleotidyltransferase can process mRNA to affect stability [58,59], which places this activity as a regulator of protein levels. Three of the identified human proteins are linked to protein synthesis, suggesting that similar to DNA repair proteins, EIF4A3, PNPT1, and AAM13690 could be conserved activities associated with modulating the toxicity of MMS. While this hypothesis and the identification of protein-protein interactions between these human proteins are the focus of future experiments, our study highlights how computational and systems based studies in *E. coli* can be used to identify proteins of interest for study in mammalian systems.

Conclusions

Gene deletion libraries are valuable tools that can be used to asses the functional importance of specific proteins after experimental perturbations, and in conjunction with Systems Biology

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

based approaches can be used to identify protein pathways and protein complexes responding to damage. We have used high-throughput screening of an *E. coli* gene deletion library to identify 99 proteins that modulate the toxicity of the alkylating agent MMS. In addition, we have used both functional and interactome mapping of identified MMS toxicity modulating proteins to demonstrate that mRNA processing and translation specific proteins participate in the response to macromolecular alkylation damage. Further, we have used clustering coefficient analysis to identify a highly connected group of activities associated with mRNA processing and translation, and demonstrated that the corresponding proteins influence the efficiency of the SOS response. Activities associated with protein synthesis have the potential to play an important role in signal transduction after damage, both general and specific, and our work supports the idea that in addition to DNA repair, mRNA processing, and translational components are vital after alkylation damage. In addition, we show that systems based approaches coupled to homology searches can be used to identify putative alkylation resistance proteins in humans.

Materials and Methods

E. coli mutants, high throughput screening, and validation

Luria Bertani broth (LB) (BP1426-3, Fisher Scientific, Waltham, MA) was used in both liquid and plate form to culture *E. coli*. The library of *E. coli* gene deletion mutants was acquired from the Genome Analysis Project in Japan [41] and supplied in 96-well plate format. High throughput genomic phenotyping was performed similarly to that described for *Saccharomyces cerevisiae* gene deletion mutants [10,11]. Briefly, 96-well plates containing the gene deletion mutants were replicated into liquid medium (LB-kanamycin) and grown for 16-hours at 37ºC. The saturated cultures were then diluted 10-fold into the same medium and 1 μl cell suspensions were robotically (Matrix Hydra) spotted on LB-kanamycin agar plates containing increasing concentrations of MMS (0, 0.045, and 0.060 % MMS) or tBuOH (150 and 165 μM). UV exposures were supplied using a Stratalinker (Stragene, Cedar Creek, Texas) with 254 nm bulbs, with mutants initially spotted on LB-kanamycin agar plates and then exposed to UV (6 and 8 J/m^2).

Inoculated plates were incubated for 16-hours at 37 ºC and the resulting plates were imaged using an AlphaImager (Alpha Innotech Corporation, San Leandro, CA). Images of plates were compiled into a visual database and analyzed to identify mutants with reduced growth after MMS treatment. Reduced growth for a specific gene-deletion mutant was identified relative to other mutants found on the 96-well plate and wild-type BW25113 cells, and was also relative to growth of the mutant on an untreated plate. To identify MMS-toxicity modulating proteins, we linked sensitive mutants to their corresponding deleted gene and assumed the protein encoded by the deleted gene was responsible for the observed phenotype. Scores represent semi-quantitative measures. For example, if a deletion mutant had reduced growth on a given concentration of MMS it was scored a two for that MMS concentration, while mutants showing only a change in color were scored one, and all others were scored a zero. It is important to note that our screen was designed to only identify mutants with increased sensitivity to MMS. Future screens that utilize higher MMS concentrations could be undertaken to identify mutants that grow better than wild-type in the presence of alkylating agents.

Newly constructed *E. coli* mutants were made using homology based recombination promoted by the λ-Red systems as described [60] and verified by PCR and DNA sequence analysis (Rooney et al., in press). Mutants were assayed in 96-well format as described above. In addition, 59 mutants from the original Keio library were further analyzed for MMS sensitivity. Each mutant was serially diluted (five 10-fold dilutions) and 5 μl of each dilution was manually applied to LB plates containing increasing concentrations of MMS. Plates were incubated at 37ºC for 16 hours, imaged, and analyzed. Sensitivity was determined by identifying the last

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

dilution at which a deletion mutant grew on MMS containing plates, relative to the diluted wild-type cells on MMS plates. Similarly, growth of both wild-type and deletion mutants on untreated plates was used to control for the viability and growth rates of gene deletion mutants under normal conditions.

Functional mapping of MMS-toxicity modulating data

All GenProtEC functional classifications specific to *E. coli* proteins were downloaded and cross referenced to proteins represented in our study. We choose to use an intermediate level of functional classification because it represented a broad yet specific spectrum of cellular processes. First we determined the actual number of MMS-toxicity modulating proteins found in each functional category. Next, statistics were compiled using random sampling and a normal curve approximation, as described previously [11]. Briefly, the base set of 2,713 proteins represented in our data set were randomly sampled to pick N = 73 proteins. The number of proteins specific to each functional category was determined and values were compiled for M = 10,000 iterations of N = 73 random proteins. Average values and standard deviations for each functional category were then generated and Z-scores were compiled for each functional category using the following formula:

$$\mathbf{Z}=\frac{\mathbf{Actual}-\mathbf{Average}}{\mathbf{Standard\ Deviation}}$$

Z-scores measure whether a functional category is over- or under-represented with MMS-toxicity modulating proteins. Corresponding p-values were determined for all over-represented categories using a one-tailed test and normal approximation.

Protein interactome mapping

E. coli protein interaction information was downloaded from the Database of Interacting Proteins (DIP) [61] and combined with the large protein-protein interaction data set published by Mori and co-workers [56]. Protein-protein interaction information was imported into Cytoscape for network visualization and sub-network filtering. Filtering was performed by highlighting MMS-toxicity modulating proteins and their associated protein-protein interactions. Clustering coefficient analysis (C) was then performed as described [57] on all filtered nodes and those nodes with C > 0 were visualized, along with corresponding interactions, using Cytoscape [55]. Network randomizations and the significance of highly connected protein groups was determined as described, with some minor additions [57]. The RplA-centred network was identified by clustering coefficient analysis of the 32 node structure identified by interactome mapping and filtering. The average clustering coefficient of all nodes found in the RplA-focused sub-network is $C_{Rpla\text{-}focused} = 0.229$, and 1000 random samplings of 32 nodes was performed to generate an average clustering coefficient for a random 32 node structure ($C_{avg} = 0.051$) and standard deviation ($S_{avg} = 0.016$). These values were then used to obtain p-values using a one-tailed test and normal approximation. A second set of randomizations was used to further validate the significance of our findings. These randomizations involved taking the base 32 node structure, identified by interactome mapping and filtering, and used 1000 randomized sets of interactions to this base unit. After each randomization, clustering coefficient analysis was performed ($C_{avg} = 0.019$), and upon completion of 1000 iterations a standard deviation ($S_{avg} = 0.0046$) was determined and used to generate a significance value, as described above.

SOS-Reporter Assays

The *sulA*-GFP reporter system used for the SOS studies was purchased from Open Biosystems (pMS201_sulA_GFP). We modified the plasmid by adding a chloramphenicol (CAM) resistant

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

cassette to generate pMS201_sulA_GFP_CAM. Plasmids were transformed into mutants of interest and selected on LB-CAM plates. Transformants were grown to mid log phase, split in two, and then mock or MMS treated (0.015%) for 30 minutes. Cells were then harvested by centrifugation, washed, and suspended in phosphate buffered saline (PBS). GFP levels in 30,000 cells were analyzed by fluorescent activated cell sorting (FACS) analysis using a Becton Dickinson LSRII Benchtop Flow Cytometer.

BLAST Analysis

Each *E. coli* protein sequence was analyzed by BLAST using the tBLASTn program, which is available from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) [62]. The nucleotide collection specific to humans was used; with E-values of less than 10^{-1} serving as our cut-off to identify similar proteins between *E. coli* and humans.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Work in the Begley lab is supported by the NIH (1K22ES01225101 and 1R01ES015037) and a James D. Watson Award through NYSTAR. Work in the Cunningham lab is supported by the NIH (CRR1C06RR0154464 and GM46312).

References

1. Brennand J, Margison GP. Reduction of the toxicity and mutagenicity of alkylating agents in mammalian cells harboring the Escherichia coli alkyltransferase gene. Proceedings of the National Academy of Sciences USA 1986;83(17):6292–6296.
2. Boffa LC, Bolognesi C. Methylating agents: their target amino acids in nuclear proteins. Carcinogenesis 1985;6(9):1399–401. [PubMed: 4028338]
3. Saffhill R, Margison GP, O'Connor PJ. Mechanisms of carcinogenesis induced by alkylating agents. Bioch Biophys Acta 1985;823:111–145.
4. Liebler DC. Proteomic approaches to characterize protein modifications: new tools to study the effects of environmental exposures. Environ Health Perspect 2002;110(Suppl 1):3–9. [PubMed: 11834459]
5. Shin NY, et al. Protein targets of reactive electrophiles in human liver microsomes. Chem Res Toxicol 2007;20(6):859–67. [PubMed: 17480101]
6. Szapacs ME, et al. Covalent adduction of human serum albumin by 4-hydroxy-2-nonenal: kinetic analysis of competing alkylation reactions. Biochemistry 2006;45(35):10521–8. [PubMed: 16939204]
7. Dennehy MK, et al. Cytosolic and nuclear protein targets of thiol-reactive electrophiles. Chem Res Toxicol 2006;19(1):20–9. [PubMed: 16411652]
8. Burgis NE, Samson LD. The Protein Degradation Response of Saccharomyces cerevisiae to Classical DNA-Damaging Agents. Chem Res Toxicol. 2007
9. Smith GJ, Grisham JW. Cytotoxicity of monofunctional alkylating agents. Methyl methanesulfonate and methyl-N′-nitro-N-nitrosoguanidine have different mechanisms of toxicity for 10T1/2 cells. Mutation Research 1983;111(3):405–17. [PubMed: 6646150]
10. Begley TJ, et al. Recovery Pathways in S. cerevisiae Revealed by Genomic Phenotyping and Interactome Mapping. Mol Canc Res 2002;1
11. Begley TJ, et al. Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping. Mol Cell 2004;16(1):117–25. [PubMed: 15469827]
12. Jelinsky SA, Samson LD. Global response of Saccharomyces cerevisiae to an alkylating agent. Proc Natl Acad Sci U S A 1999;96(4):1486–91. [PubMed: 9990050]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

13. Chang M, et al. A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. Proc Natl Acad Sci U S A 2002;99 (26):16934–9. [PubMed: 12482937]
14. Engelward BP, et al. A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome. J Biol Chem 1998;273(9):5412–8. [PubMed: 9479003]
15. Chen J, Samson L. Induction of S. cerevisiae MAG 3-methyladenine DNA glycosylase transcript levels in response to DNA damage. Nucleic Acids Res 1991;19(23):6427–32. [PubMed: 1754379]
16. Jeggo P, et al. An adaptive response of E. coli to low levels of alkylating agent: comparison with previously characterised DNA repair pathways. Mol Gen Genet 1977;157(1):1–9. [PubMed: 414071]
17. Begley TJ, Samson LD. AlkB mystery solved: oxidative demethylation of N1-methyladenine and N3-methylcytosine adducts by a direct reversal mechanism. Trends Biochem Sci 2003;28(1):2–5. [PubMed: 12517444]
18. Sekiguchi M, Sanada M. Alkylation carcinogenesis in mice with altered levels of DNA repair methyltransferase. Prog Exp Tumor Res 1999;35:25–36. [PubMed: 10377749]
19. Nakatsuru Y, et al. Characterization of O^6-methylguanine-DNA methyltransferase in transgenic mice introduced with the E. coli ada gene. Mutation Research 1991;254(3):225–30. [PubMed: 2052012]
20. Rebeck GW, et al. A second DNA methyltransferase repair enzyme in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 1988;85(9):3039–43. [PubMed: 3283737]
21. Potter PM, et al. Characterisation and nucleotide sequence of ogt, the O6-alkylguanine-DNA-alkyltransferase gene of E. coli. Nucleic Acids Research 1987;15(22):9177–93. [PubMed: 2825131]
22. McCarthy TV, Lindahl T. Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of E. coli. Nucleic Acids Res 1985;13(8):2683–98. [PubMed: 2987862]
23. Nakabeppu Y, Kondo H, Sekiguchi M. Cloning and characterization of the alkA gene of Escherichia coli that encodes 3-methyladenine DNA glycosylase II. Journal of Biological Chemistry 1984;259 (22):13723–9. [PubMed: 6389535]
24. Kataoka H, Yamamoto Y, Sekiguchi M. A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate. Journal of Bacteriology 1983;153(3):1301–7. [PubMed: 6337994]
25. Evensen G, Seeberg E. Adaptation to alkylation resistance involves the induction of a DNA glycosylase. Nature 1982;296(5859):773–5. [PubMed: 7040984]
26. Sobol RW, et al. Base excision repair intermediates induce p53-independent cytotoxic and genotoxic responses. J Biol Chem 2003;278(41):39951–9. [PubMed: 12882965]
27. Sobol RW, Wilson SH. Mammalian DNA beta-polymerase in base excision repair of alkylation damage. Prog Nucleic Acid Res Mol Biol 2001;68:57–74. [PubMed: 11554313]
28. Riazuddin S, Lindahl T. Properties of 3-methyladenine-DNA glycosylase from Escherichia coli. Biochemistry 1978;17:2110–2118. [PubMed: 352392]
29. Engelward BP, et al. Repair-deficient 3-methyladenine DNA glycosylase homozygous mutant mouse cells have increased sensitivity to alkylation-induced chromosome damage and cell killing. EMBO Journal 1996;15(4):945–52. [PubMed: 8631315]
30. Duncan T, et al. Reversal of DNA alkylation damage by two human dioxygenases. Proc Natl Acad Sci U S A 2002;99(26):16660–5. [PubMed: 12486230]
31. Trewick SC, et al. Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 2002;419(6903):174–8. [PubMed: 12226667]
32. Falnes PO, Johansen RF, Seeberg E. AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature 2002;419(6903):178–82. [PubMed: 12226668]
33. Aravind L, Koonin EV. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol 2001;2(3)
34. Drablos F, et al. Alkylation damage in DNA and RNA--repair mechanisms and medical significance. DNA Repair (Amst) 2004;3(11):1389–407. [PubMed: 15380096]
35. Jeggo P, et al. The adaptive response of E. coli to low levels of alkylating agent: the role of polA in killing adaptation. Molecular & General Genetics 1978;162(3):299–305. [PubMed: 355835]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

36. Samson L, Cairns J. A new pathway for DNA repair in Escherichia coli. Nature 1977;267(5608): 281–3. [PubMed: 325420]
37. Courcelle J, et al. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli. Genetics 2001;158(1):41–64. [PubMed: 11333217]
38. Gasch AP, et al. Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. Mol Biol Cell 2001;12(10):2987–3003. [PubMed: 11598186]
39. Jelinsky SA, et al. Regulatory networks revealed by transcriptional profiling of damaged saccharomyces cerevisiae cells: rpn4 links base excision repair with proteasomes [In Process Citation]. Mol Cell Biol 2000;20(21):8157–67. [PubMed: 11027285]
40. Workman CT, et al. A systems approach to mapping DNA damage response pathways. Science 2006;312(5776):1054–9. [PubMed: 16709784]
41. Baba T, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2006;2:2006 0008.
42. Rudd KE. EcoGene: a genome sequence database for Escherichia coli K-12. Nucleic Acids Res 2000;28(1):60–4. [PubMed: 10592181]
43. Drury LS, Buxton RS. DNA sequence analysis of the dye gene of Escherichia coli reveals amino acid homology between the dye and OmpR proteins. J Biol Chem 1985;260(7):4236–42. [PubMed: 2984198]
44. Serres MH, Goswami S, Riley M. GenProtEC: an updated and improved analysis of functions of Escherichia coli K-12 proteins. Nucleic Acids Res 2004;32(Database issue):D300–2. [PubMed: 14681418]
45. Ezaki B, et al. Involvement of DnaK protein in mini-F plasmid replication: temperature-sensitive seg mutations are located in the dnaK gene. Mol Gen Genet 1989;218(2):183–9. [PubMed: 2674651]
46. Carpousis AJ, Vanzo NF, Raynal LC. mRNA degradation. A tale of poly(A) and multiprotein machines. Trends Genet 1999;15(1):24–8. [PubMed: 10087930]
47. Zuo Y, Deutscher MP. The physiological role of RNase T can be explained by its unusual substrate specificity. J Biol Chem 2002;277(33):29654–61. [PubMed: 12050169]
48. Boulton SJ, et al. Combined functional genomic maps of the C. elegans DNA damage response. Science 2002;295(5552):127–31. [PubMed: 11778048]
49. Rual JF, et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature 2005;437(7062):1173–8. [PubMed: 16189514]
50. Gunsalus KC, et al. Predictive models of molecular machines involved in Caenorhabditis elegans early embryogenesis. Nature 2005;436(7052):861–5. [PubMed: 16094371]
51. Ideker T, et al. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 2001;292(5518):929–34. [PubMed: 11340206]
52. Lim J, et al. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. Cell 2006;125(4):801–14. [PubMed: 16713569]
53. Goh KI, et al. The human disease network. Proc Natl Acad Sci U S A 2007;104(21):8685–90. [PubMed: 17502601]
54. Yildirim MA, et al. Drug-target network. Nat Biotechnol 2007;25(10):1119–26. [PubMed: 17921997]
55. Shannon P, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498–504. [PubMed: 14597658]
56. Arifuzzaman M, et al. Large-scale identification of protein-protein interaction of Escherichia coli K-12. Genome Res 2006;16(5):686–91. [PubMed: 16606699]
57. Said MR, et al. Global network analysis of phenotypic effects: protein networks and toxicity modulation in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2004;101(52):18006–11. [PubMed: 15608068]
58. Leszczyniecka M, et al. Identification and cloning of human polynucleotide phosphorylase, hPNPase old-35, in the context of terminal differentiation and cellular senescence. Proc Natl Acad Sci U S A 2002;99(26):16636–41. [PubMed: 12473748]
59. Sarkar D, et al. Down-regulation of Myc as a potential target for growth arrest induced by human polynucleotide phosphorylase (hPNPaseold-35) in human melanoma cells. J Biol Chem 2003;278 (27):24542–51. [PubMed: 12721301]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

60. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 2000;97(12):6640–5. [PubMed: 10829079]

61. Xenarios I, et al. DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res 2002;30(1):303–5. [PubMed: 11752321]

62. Altschul SF, et al. Basic local alignment search tool. J Mol Biol 1990;215(3):403–10. [PubMed: 2231712]

63. Boiteux S, Huisman O, Laval J. 3-Methyladenine residues in DNA induce the SOS function sfiA in Escherichia coli. EMBO 1984;3:2569–2573.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

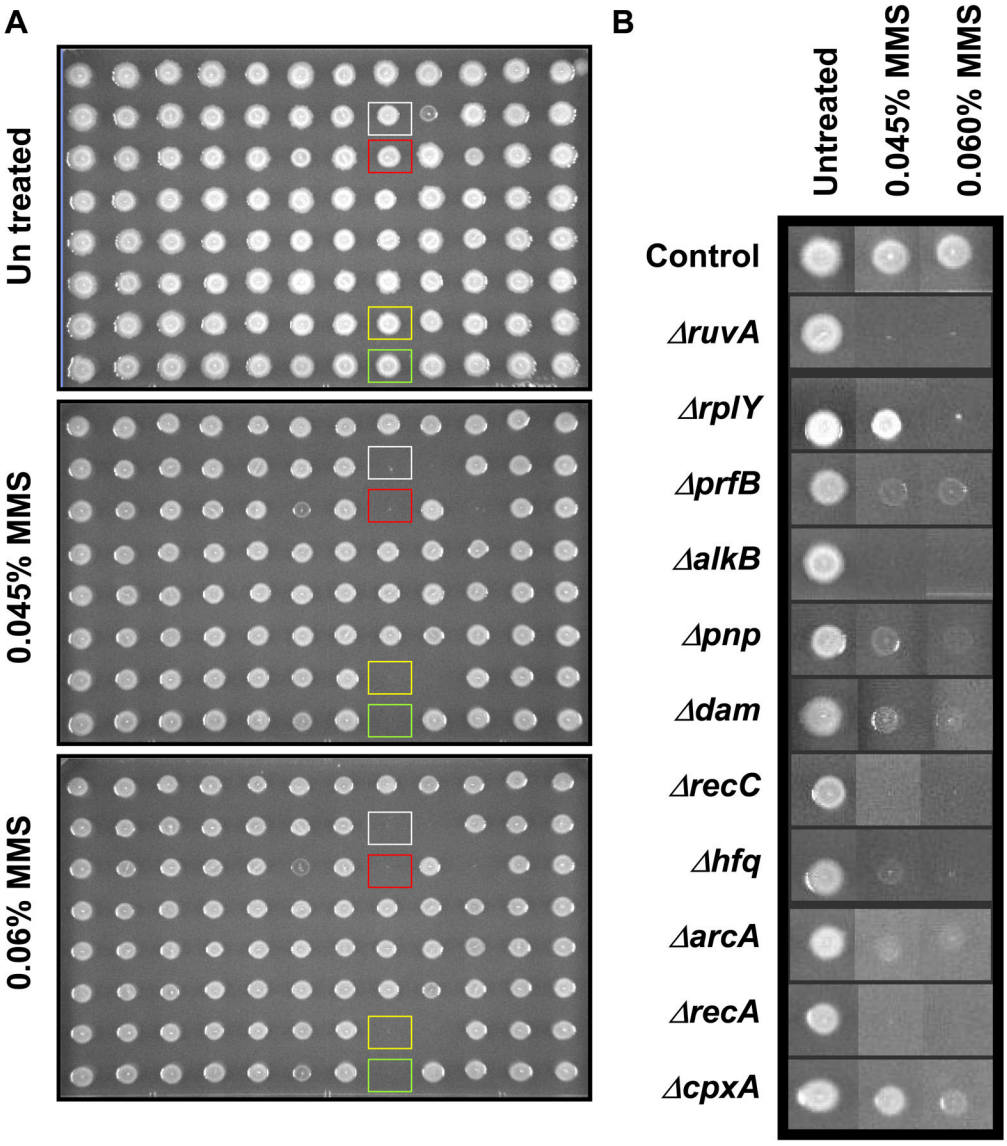

Figure 1. Genomic phenotyping of *E. coli* with MMS

(A) 96 gene deletion mutants were spotted onto agar plates containing increasing concentrations of MMS, incubated at 37ºC for 16 hours, and imaged. White, red, yellow and green squares identify the MMS sensitive gene deletion mutants *ΔruvA*, *ΔruvC*, *ΔalkA*, and *ΔalkB* mutants, respectively. Note that there are a total of 9 MMS sensitive deletion mutants on the plate. (B) Images of specific mutants were recompiled and demonstrate that varying degrees of MMS sensitivity were observed in the tested gene deletion mutants. The *Δpnp* and *Δdam* mutants are examples where a color change from white to dark grey was indicative of a growth defect after MMS exposure.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

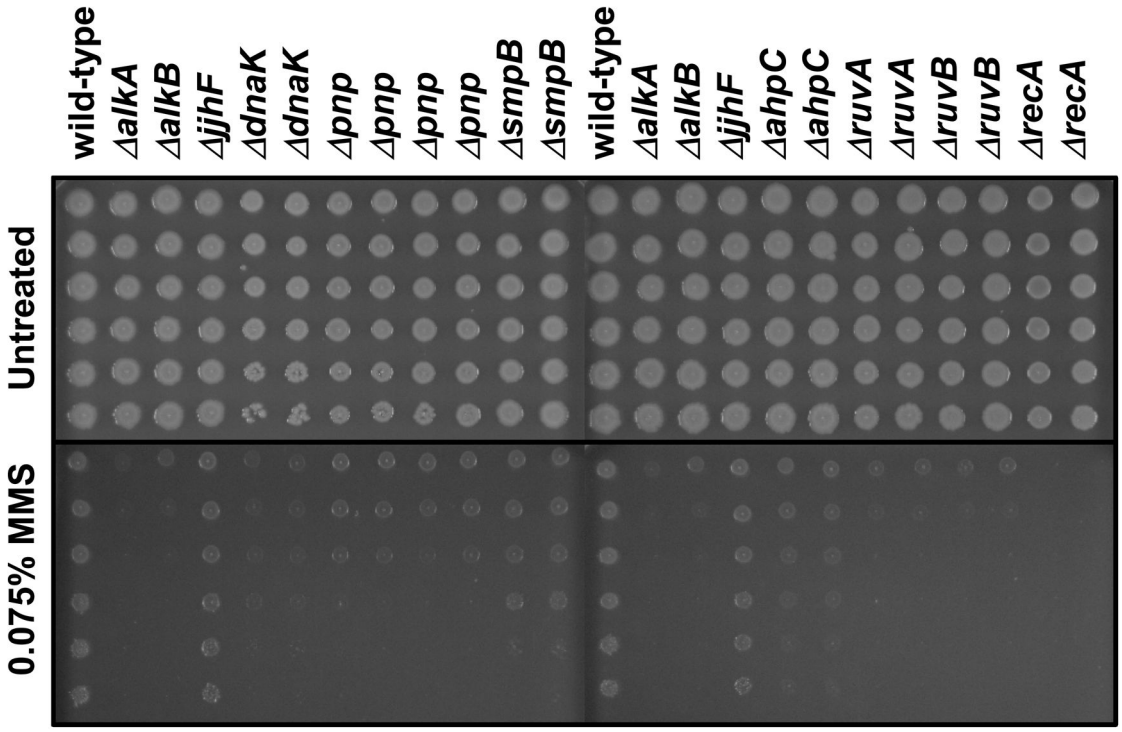

Figure 2. Dilution screen validates high-throughput results

Deletion mutants specific to the GeneProtEC functional categories over-represented with MMS-toxicity modulating proteins were grown overnight and serially diluted onto LB plates without and with 0.075% MMS. Wild-type and *ΔjjhF* serve as controls and are both resistant to MMS. All others were classified as MMS sensitive in both the initial and secondary screening assays. Data specific to all mutants represented in Table 2 can be found in Supplemental Figure S2.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

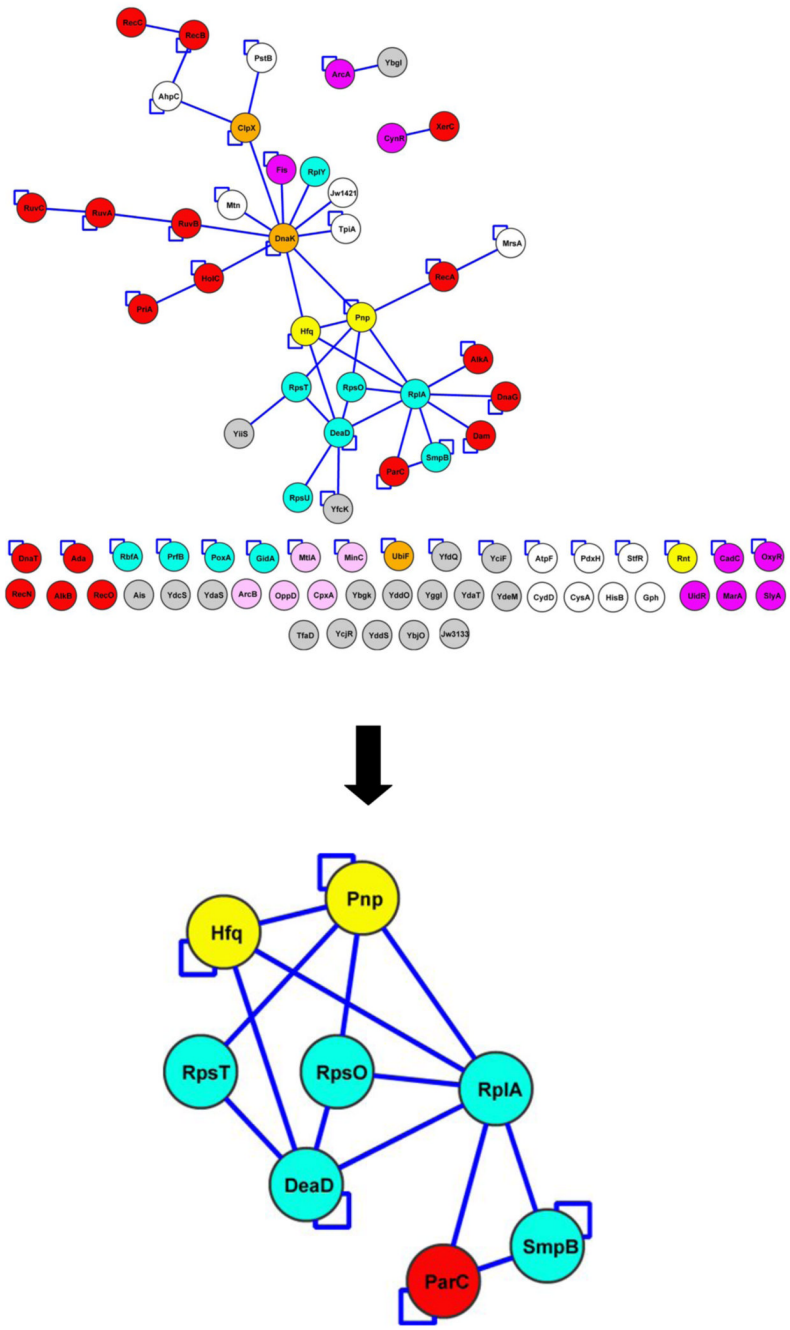

Figure 3. Functional themes and highly clustered proteins in the MMS-toxicity modulating sub-network

(Top) Each of the MMS toxicity modulating proteins found in the filtered interactome was colored according to its predominant functional theme, as defined by EcoGene or SwissProt. Red circle = DNA repair, replication, and recombination; blue circle = protein synthesis; purple circle = transcription; yellow circle = RNA processing; orange circle = protein stabilization; grey circle = unknown; white circle = other. (Bottom) MMS-toxicity modulating proteins with clustering coefficients greater then zero were determined using MATLAB algorithms and then visualized using Cytoscape. This was done to identify groups of proteins that have the potential to be part of a complex or pathway. RpsT was included due to its connectivity to the sub-

network via Pnp and DeaD. Ultimately, the highly clustered sub-network centred on RplA was identified and contained activities involved in protein synthesis and RNA metabolism.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

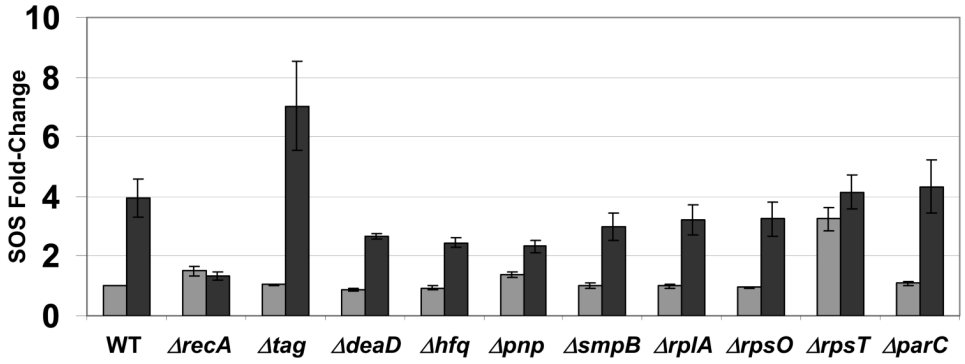

Figure 4. SOS Reporter analysis of mutants specific to the RplA-centred sub-network

A plasmid based SOS reporter, *sulA*-GFP, was transformed into each cell type, transformants were grown to mid log phase and mock (grey bars) or 0.015% MMS treated (black bars) for 30 minutes. Wild-type, *Δtag* (hyper-SOS after MMS[63]), and *ΔrecA* (hypo-SOS after MMS) serve as controls. FACS analysis of 30,000 cells was then used to quantitate GFP levels, and fold change relative to untreated wild-type was plotted for each cell type. Error bars represent standard deviations between three biological replicates of 30,000 cells each.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Proteins corresponding to *E. coli mutants* sensitive to MMS

Gene Name	Description	MMS Sensitive
alkA	3-methyl-adenine DNA glycosylase II	++++
alkB	oxidative demethylase of N1-methyladenine or N3-methylcytosine DNA lesions	++++
arcA	DNA-binding response regulator in two-component regulatory system with ArcB or CpxA	++++
cydD	fused cysteine transporter subunits of ABC superfamily: membrane component/ATP-binding component	++++
cysA	sulfate/thiosulfate transporter subunit/ATP-binding component of ABC superfamily	++++
dnaT	DNA biosynthesis protein (primosomal protein I)	++++
priA	primosome factor n′ (replication factor Y)	++++
recA	DNA strand exchange and recombination protein with protease and nuclease activity	++++
recC	exonuclease V (RecBCD complex), gamma chain	++++
rplA	50S ribosomal subunit protein L1	++++
ruvA	component of RuvABC resolvasome, regulatory subunit	++++
ruvC	component of RuvABC resolvasome, endonuclease	++++
aroK	shikimate kinase I	+++
dcuC	anaerobic C4-dicarboxylate transport	+++
dnaG	DNA primase	+++
fis	global DNA-binding transcriptional dual regulator	+++
hfq	HF-I, host factor for RNA phage Q beta replication	+++
holC	DNA polymerase III, chi subunit	+++
JW5183	unknown function	+++
mtlA	fused mannitol-specific PTS enzymes: IIA components/IIB components/IIC components	+++
mtn	5′-methylthioadenosine/S-adenosylhomocysteine nucleosidase	+++
parC	DNA topoisomerase IV, subunit A	+++
pdxH	pyridoxine 5′-phosphate oxidase	+++
pnp	polynucleotide phosphorylase/polyadenylase	+++
prfB	peptide chain release factor RF-2	+++
rbfA	30s ribosome binding factor	+++
recB	exonuclease V (RecBCD complex), beta subunit	+++
rplY	50S ribosomal subunit protein L25	+++
rpsO	30S ribosomal subunit protein S15	+++
rpsU	30S ribosomal subunit protein S21	+++
ruvB	ATP-dependent DNA helicase, component of RuvABC resolvasome	+++
ybhH	unknown function	+++
ybjO	predicted inner membrane protein	+++
yihX	predicted hydrolase	+++
yjjY	unknown function	+++
ytfA	predicted transcriptional regulator	+++
ada	fused DNA-binding transcriptional dual regulator/O6-methylguanine-DNA methyltransferase	++
arcB	tripartite sensor/histidine protein kina	++

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Gene Name	Description	MMS Sensitive
atpF	F0 sector of membrane-bound ATP synthase, subunit b	++
clpX	ATPase and specificity subunit of ClpX-ClpP ATP-dependent serine protease	++
dam	DNA adenine methylase	++
dnaK	chaperone Hsp70, co-chaperone with DnaJ	++
hisB	fused histidinol-phosphatase/imidazoleglycerol-phosphate dehydratase	++
mrsA	phosphoglucosamine mutase	++
oppD	oligopeptide transporter subunit/ATP-binding component of ABC superfamily	++
recO	gap repair protein	++
rnt	ribonuclease T (RNase T)	++
tpiA	triosephosphate isomerase	++
ubiF	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol oxygenase	++
uidR	DNA-binding transcriptional repressor	++
yaiS	unknown function	++
ybgK	predicted enzyme subunit	++
yciF	unknown function	++
ydcS	predicted spermidine/putrescine transporter subunit/periplasmic-binding component of ABC superfamily	++
yjiW	unknown function	++
ahpC	alkyl hydroperoxide reductase, C22 subunit	+
ais	conserved protein	+
bdm	osmoresponsive gene with reduced expression in biofilms, function unknown	+
blr	beta-lactam resistance protein	+
cadC	DNA-binding transcriptional activator	+
cpxA	sensory histidine kinase in two-component regulatory system with CpxR	+
cynR	transcriptional activator of cyn operon	+
deaD	ATP-dependent RNA helicase	+
fliO	flagellar biosynthesis protein	+
gidA	5-methylaminomethyl-2-thiouridine modification at tRNA U34 (MnmG)	+
gph	phosphoglycolate phosphatase	+
JW2207	unknown function	+
lpcA	D-sedoheptulose 7-phosphate isomerase	+
marA	Transcription activator of multiple antibiotic resistance	+
minC	cell division inhibitor	+
oxyR	DNA-binding transcriptional dual regulator	+
potH	putrescine transporter subunit: membrane component of ABC superfamily	+
poxA	predicted lysyl-tRNA synthetase	+
pstB	phosphate transporter subunit/ATP-binding component of ABC superfamily	+
recN	recombination and repair	+
rpsT	30S ribosomal subunit protein S20	+
rzpD	DLP12 prophage; predicted murein endopeptidase	+
slyA	global transcriptional regulator	+

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Gene Name	Description	MMS Sensitive
smpB	trans-translation protein	+
stfR	Rac prophage; predicted tail fiber protein	+
tfaD	pseudogene, tail fiber assembly gene	+
xerC	site-specific tyrosine recombinase	+
yaiU	predicted protein	+
ybgI	conserved metal-binding protein	+
ybhR	predicted transporter subunit: membrane component of ABC superfamily	+
ycfC	predicted lysogenization regulator	+
ycjR	unknown function	+
ydaS	Rac prophage	+
ydaT	Rac prophage; predicted protein	+
yddO	D-ala-D-ala transporter subunit/ATP-binding component of ABC superfamily	+
yddS	D-ala-D-a la transporter subunit/periplasmic-binding component of ABC superfamily	+
ydeM	unknown function	+
yecN	predicted inner membrane protein	+
yfcK	fused 5-methylaminomethyl-2-thiouridine-forming enzyme methyltransferase(MnmC)	+
yfdQ	CPS-53 (KpLE1) prophage; predicted protein	+
yggA	arginine transporter	+
yggL	unknown function	+
yiiS	unknown function	+
yliD	predicted peptide transporter subunit: membrane component of ABC superfamily	+

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

GenProtEC functional categories over-represented with MMS-toxicity modulating proteins

Classifcation	Total Proteins in Category	MMS Toxicity Modulating	Proteins
DNA recombination	19	9	Fis, RecA, RecB, RecC, RecO, RuvA, RuvB, RuvC, XerC
Ribosomal proteins	15	5	RplA, RplY, RpsO, RpsT, RpsU
Translation	28	6	PrfB, RplA, RplY, RpsO, RpsT, RpsU
DNA repair	48	8	AlkA, AlkB, Gph, RecA, RecO, RuvA, RuvB, RuvC
DNA replication	36	5	DnaG, DnaT, HolC, ParC, PriA
RNA degradation	10	2	Pnp, Rnt
DNA structure level	13	2	Dam, ParC
DNA degradation	27	3	RecB, RecC, Rnt
Nucleoproteins, basic proteins	16	2	Fis, Hfq
Other stresses (mechanical, nutritional, oxidative)	18	2	ArcA, OxyR
Posttranscriptional	63	4	CpxA, Hfq, RecA, RplA
Detoxification (xenobiotic metabolism)	42	3	AhpC, CpxA, CynR

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Human proteins similar to the highly clustered MMS-toxicity modulating proteins from *E. coli*

***E. coli* Protein**	**Description**	**Human Protein**	**E-Value**
DeaD	Translation factor W2, putative RNA helix-destabilizer;facilitates translation of mRNAs with 5′ secondary structures	Eukaryotic translation initiation factor 4A, isoform 3 (EIF4A3)	4.00E-70
Hfq	Global regulator of sRNA function; host factor for RNA phage Q beta replication; HF-I; DNA- and RNA-binding protein; RNA chaperone; multiple regulatory roles	Delta 2-isopentenyl adenosine tRNA-like protein (AAM13690)	1.00E-19
Pnp	Polynucleotide phosphorylase; exoribonuclease; PNPase component of RNA degradosome	polyribonucleotide nucleotidyltransferase1 (PNPT1)	6.00E-126
SmpB	tmRNA RNA-binding protein; required for peptide-tagging trans-translation and association of tmRNA(SsrA) with ribosomes	Unknown (Z22851)	3.00E-20