Introduction of cytosine-5 DNA methylation sensitizes cells to oxidative damage

Krwawicz, J.[1*], Sheeba, C.J.[2*], Hains, K.[1], McMahon, T.[3], Zhang, Y.[1], Kriaucionis, S.[3], and Sarkies, P[1^].

[1]Department of Biochemistry, University of Oxford, Oxford, UK
[2]London Institute of Medical Sciences, Imperial College London, Du Cane Road, London, UK
[3]Ludwig Institute for Cancer Research, University of Oxford, Oxford, UK
*contributed equally
^Correspondence:
Peter.sarkies@bioch.ox.ac.uk

Abstract
DNA methylation at the 5 position of cytosine (5mC) is an ancient epigenetic mark in eukaryotes. The levels of total 5mC vary enormously between different species, and the DNA methyltransferases that introduce 5mC have been repeatedly lost in several independent lineages. DNA methyltransferases are a threat to genomic stability due to the increased mutagenicity of 5mC bases and the propensity of DNA methyltransferases themselves to introduce DNA alkylation damage as an off-target effect. However, whether alkylation damage explains why 5mC is frequently lost in evolution is unclear. Here we tested the fitness consequences of DNA methyltransferase-induced alkylation damage by introducing a eukaryotic-like 5mC system into *E. coli*. We showed that introducing 5mC genome-wide leads to increased sensitivity to alkylating agents, which is strongly enhanced by removal of the 3mC repair enzyme AlkB. Unexpectedly, we discovered that 5mC introduction led to increased sensitivity to oxidative stress. We showed that this is due to increased formation of reactive oxygen in the presence of 5mC. We determined that reactive oxygen species led to non-enzymatic oxidation of 5mC, producing modified cytosines such as 5fC that are recognised as DNA base damage in *E. coli*. Overall, our work identifies increased sensitivity to oxidative stress, as well as alkylating agents, as a negative consequence of genome-wide 5mC. Oxidative stress is frequently encountered by organisms in their environment, thus offering a plausible reason for total loss of 5mC in some species.

Introduction

Methylation of the 5th position of cytosine within DNA (5mC) is found in many species drawn from all domains of life(Jurkowski & Jeltsch, 2011). In prokaryotes (bacteria and archaea) methylation is carried out by DNA methyltransferases (DNMTs) that typically have sequence specificity targeting them to 4-8bp sequence motifs. 5mC introduced in this manner contributes to restriction modification systems, which protect genomes against invading viruses, plasmids or other mobile genetic elements. In contrast, in eukaryotes, DNMTs that introduce 5mC have more limited sequence specificity, and are capable of acting genome-wide. Across eukaryotes, 5mC is commonly introduced at CG dinucleotides(de Mendoza et al., 2020). The symmetrical nature of this sequence context enables 5mC in the CG context to act epigenetically, because maintenance methyltransferases (such as DNMT1 in mammals and Met1 in plants) have evolved to recognise hemi-methylated DNA after DNA replication and introduce methylation onto the

unmethylated strand(Law & Jacobsen, 2010). This means that 5mC in CG contexts can persist through cell division independent of the activity of the de novo methyltransferases (such as DNMT3A and B in mammals) that introduced them initially. DNA methylation in other sequence contexts is widespread in plants and is also found in some vertebrate genomes, particularly in non-dividing cells such as neurons(de Mendoza et al., 2020, 2021).

In eukaryotes, the functions of DNA methylation has not so far been linked to restriction modification systems analogous to bacteria. Instead, 5mC has roles in gene expression regulation(Sarkies, 2022b). These roles are complex, overlapping and not fully characterised, especially because they depend on genomic context. For example, in mammals DNA methylation at promoter regions is associated with transcriptional silencing during development(Weber et al., 2007), whereas methylation in gene bodies is linked to the presence of histone modifications associated with transcription, such as H3K36me3 and may contribute to noise suppression for example through suppression of spurious transcription (Baubec et al., 2015). In arthropods, 5mC is much less abundant genome-wide and is often confined to a small number of transcriptionally active housekeeping genes(Bewick et al., 2017; Lewis et al., 2020).

Eukaryotic 5mC methylation, and, in particular, epigenetic methylation in the CG context has an ancient origin. Nevertheless, DNA methylation is far from universal. Many species have lost DNA methylation altogether, including many commonly used as model organisms such as *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae and Schizosaccharomyces pombe*(Bewick et al., 2019; Feng et al., 2010; Keller et al., 2016; Zemach et al., 2010). As indicated above, the levels of 5mC vary enormously across different species, even if they possess methylation machinery(Bewick et al., 2017; Lewis et al., 2020). Often evolutionary transitions associated with either widespread increases in DNA methylation levels or complete loss have occurred very recently, suggesting ongoing processes selecting for or against DNA methylation in different species(de Mendoza, Hatleberg, et al., 2019; de Mendoza, Pflueger, et al., 2019). The reasons for such rapid diversification in methylation systems are still largely unknown.

One important possible explanation that has emerged is the propensity of 5mC to damage the genome. This is observed as an increased rate of mutations at methylated CpGs, when compared to unmodified contexts(Tomkova & Schuster-Böckler, 2018). In part this is due to the fact that cytosine deaminates to uracil whilst 5mC deaminates to thymine(Lindahl, 1996). Recognition of uracil as a foreign DNA base might be easier in cells than detecting a T-G mismatch, leading to an elevated rate of C to T mutations at 5mC(Schmutte et al., 1995; Tarantino et al., 2018). Potential evidence for this was first provided by the observation of depletion of CG dinucleotides in mammalian genomes relative to the expected proportion based on the total number of C and G nucleotides(Bird, 1980; Gardiner-Garden & Frommer, 1987). Subsequently, analysis of mutation rates in cancer quantified the excess C-T mutation rate(Alexandrov et al., 2013), leading to a more sophisticated understanding of its origin including the discovery that 5mC increases the rate of base pair misincorporation by replicative polymerases(Tomkova et al., 2018).

In addition to increased mutation of 5mC nucleotides, we recently discovered an additional source of genome instability due to DNMT activity. We discovered that across eukaryotes DNMTs co-evolve with alkylation repair, in particular the enzyme ALKBH2, which tends to be present whenever DNMTs are present and absent when DNMTs have been lost(Rošić et al., 2018). In arthropods, the presence of ALKBH2 correlates with higher levels of 5mC genome wide(Lewis et al., 2020). ALKBH2 repairs the damaged base 3mC, and

indeed we showed that DNMTs can introduce 3mC *in vitro* and *in vivo*(Dukatz et al., 2019; Rošić et al., 2018).

The mutagenic and genotoxic consequences of DNMT expression might be one reason why DNMTs, and 5mC could be selectively disadvantageous. This cost might be more important if the species is exposed to high levels of DNA damage in its environment, thus putting more strain on DNA repair mechanisms(Sarkies, 2022b). Here, we set out to explore this hypothesis to attempt to test directly the consequences of DNMT activity for sensitivity to genotoxic stress. In order to test this independent of functions of DNMT in gene regulation, we developed a system to introduce high levels of CG methylation into the model bacteria *E. coli*. We demonstrated that *E. coli* expressing DNMTs was hypersensitive to DNA alkylating agents that introduce 3mC into the DNA, and that this was exacerbated by loss of AlkB, the *E. coli* homologue of ALKBH2(Fedeles et al., 2015). Unexpectedly, we also discovered that DNMTs led to hypersensitivity to oxidative stress induction. We showed that DNMT expression leads to increased accumulation of reactive oxygen species, which is exacerbated by exogenous oxidising agents. We present biochemical and genetic evidence that the consequence of oxidative damage in genomes with 5mC is the emergence of 5hmC and 5fC - further oxidation products of 5mC, which results in toxicity. Taken together, our data points to an additional source of genotoxic stress beyond alkylation damage associated with DNMT activity, which could have particular relevance for explaining why some species have lost DNMTs and 5mC completely, and may also be important in the cellular consequences of DNMT activity in mammals.

Results

A system to introduce genome-wide CG methylation into E. coli

Studying the effect of DNMT activity on genotoxic stress sensitivity is complicated in eukaryotic cells because CG methylation is associated with gene expression so secondary effects on genotoxic stress due to expression changes are likely to be pervasive. We therefore sought to establish a system where we could introduce DNMT activity into cells and study its effects on genotoxic stress more straightforwardly. We took advantage of the highly active bacterial CG methyltransferase M.SssI(Darii et al., 2009). The catalytic residues of M.SssI are highly conserved with eukaryotic DNMTs (Figure 1A). We expressed M.SssI on a low copy number plasmid in a strain of *E. coli* lacking the McrABC restriction modification system (C2523). This resulted in robust induction of genome-wide 5mC in the CG content, as demonstrated by methylation sensitive restriction digest and mass spectrometry (Figure 1B,C). We also expressed the CG methyltransferase M.MpeI as an alternative model to test our conclusions(Wojciechowski et al., 2013). M.MpeI induced lower levels of 5mC genome-wide than M.SssI (Figure 1C). Together these two enzymes provided an opportunity to study the effects of genome-wide 5mC in the CG context on genotoxic stress.

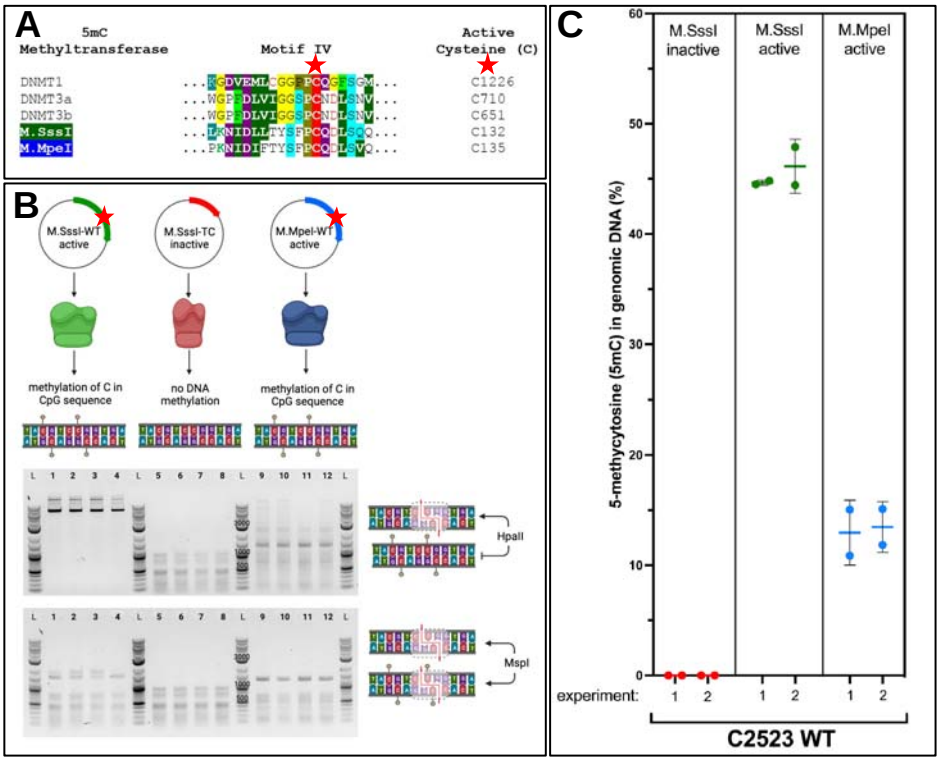


Figure 1- Introduction of genome-wide CG methylation into *E. coli*
A – Plasmids encoding active 5mC DNA methyltransferases (M.SssI and M.MpeI) and inactive M.SssI were introduced into *Esherichia coli* strain B C2523 without active 5mC nuclease McrA to protect methylated DNA before degradation. **B** Methyl-sensitive restriction enzyme digest. HpaII is blocked by 5mC in the recognized C5mCGG palindrome and MspI can digest methylated C5mCGG and unmethylated CCGG sequences. **C** Genomic DNA was isolated from strains expressing active methyltransferases (M.SssI and M.MpeI) and inactive M.SssI, and global 5mC level was calculated as a percentage of total cytosine.

DNMTs sensitize cells to alkylation damage

Our previous work demonstrated that DNMTs induce alkylation damage in the form of 3mC into genomic DNA(Rošić et al., 2018). To investigate the effect of this on genotoxic stress resistance we exposed cells expressing DNMT to the DNA alkylating agent Methyl-methan-sulfonate (MMS). MMS introduces a number of alkylation lesions into DNA including 3mC(Sedgwick, 2004). Cells expressing either M.SssI or M.MpeI showed increased sensitivity to MMS treatment compared to WT C2523 (Figure 2), supporting the conclusion that the expression of DNMTs increased the levels of alkylation damage. To test this conclusion further we inactivated AlkB, the *E. coli* homologue of the eukaryotic ALKB family(Ougland et al., 2015). *E. coli* AlkB, similarly to ALKBH2, is able to repair 3mC damage in DNA(Trewick et al., 2002), although it should be noted that its substrate specificity is much wider, with the ability to act on RNA and protein as well as DNA(Fedeles et al., 2015). AlkB mutants were hypersensitive to MMS treatment compared to WT C2523, consistent with previous studies of the *E. coli* AlkB enzyme(Dinglay et al., 2000; Sikora et al., 2010; Trewick et al., 2002). Expression of DNMTs strongly exacerbated this effect (DNMT p=6.4e-14 two-way Anova, Figure 2). Importantly, DNMT expression combined with deletion of

AlkB had a greater effect on MMS sensitivity than would be expected from the effect of each individually (DNMT*alkB interaction p=0.0044 three-way Anova). This non-additive interaction is consistent with the hypothesis that DNMTs introduce a form of DNA damage that is repaired by AlkB. Together, these data support the introduction of 3mC by DNMT activity and demonstrate that this leads to increased sensitivity to MMS, suggesting that the off-target induction of 3mC has a fitness cost for cells when exposed to genotoxic stress in their environment.

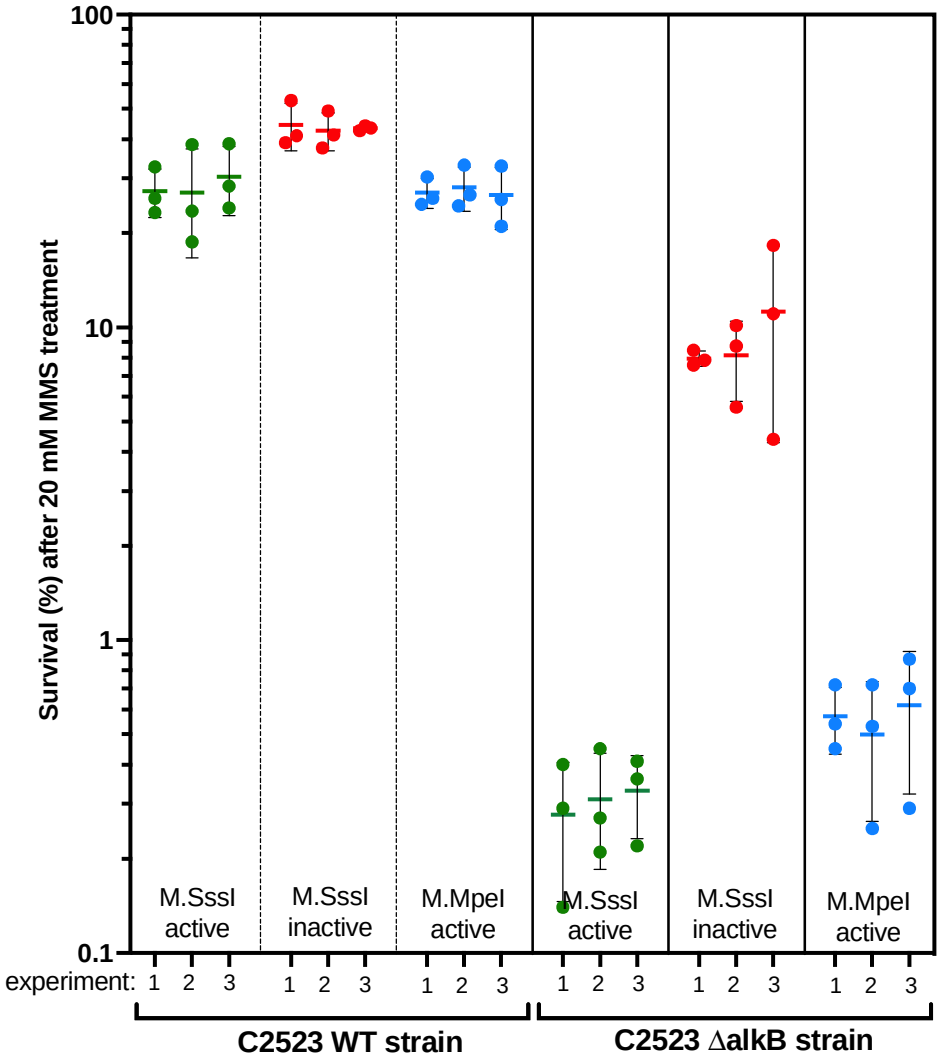


Figure 2 MMS-induced sensitivity in strains carrying active 5mC methyltransferases.
WT strains and *alkB* mutants with M.SssI and M.MpeI methyltransferases were treated with 20 mM MMS. The y-axis shows the percentage survival relative to the survival under the same conditions without MMS. Each set of 3 points shows three separate clones derived from a single transformation with the indicated plasmid. Three separate transformations were performed for each plasmid.

DNMT expression sensitizes cells to oxidative stress

We wondered whether the effect of DNMT expression was specific to DNA alkylation damage sensitivity. We therefore subjected DNMT expressing cells to other genotoxic stresses and tested their sensitivity. WT C2523 cells and C2523 expressing DNMT were equally sensitive to cisplatin (Supplemental Figure 1), which introduces intra and interstrand cross-links into the DNA(Dasari & Bernard Tchounwou, 2014). However, DNMTs demonstrated a marked sensitivity to induction of oxidative stress by exposure to H_2O_2 (Supplemental Figure 1; Figure 3). Genotoxic stress associated with H_2O_2 treatment is due to production of reactive oxygen species which generate a number of lesions in DNA. Oxidised purines in particular are excised by the base excision repair glycosylase Fpg (also known as MutM). Fpg deletion alone has little effect on sensitivity to oxidative stress, presumably due to redundancy across DNA repair pathways (Figure 3). However, induction of active DNMT expression through M.SssI led to a strongly exacerbated sensitivity to H_2O_2 treatment ($p<1e-16$ Two way anova; Figure 3). The combined effect of expression of DNMT and mutation of Fpg was greater than the sum of the individual effects (Fpg*DNMT interaction $p=0.032$; Figure 3). Together this suggested that DNMT expression increased sensitivity to oxidative stress-induced DNA damage.

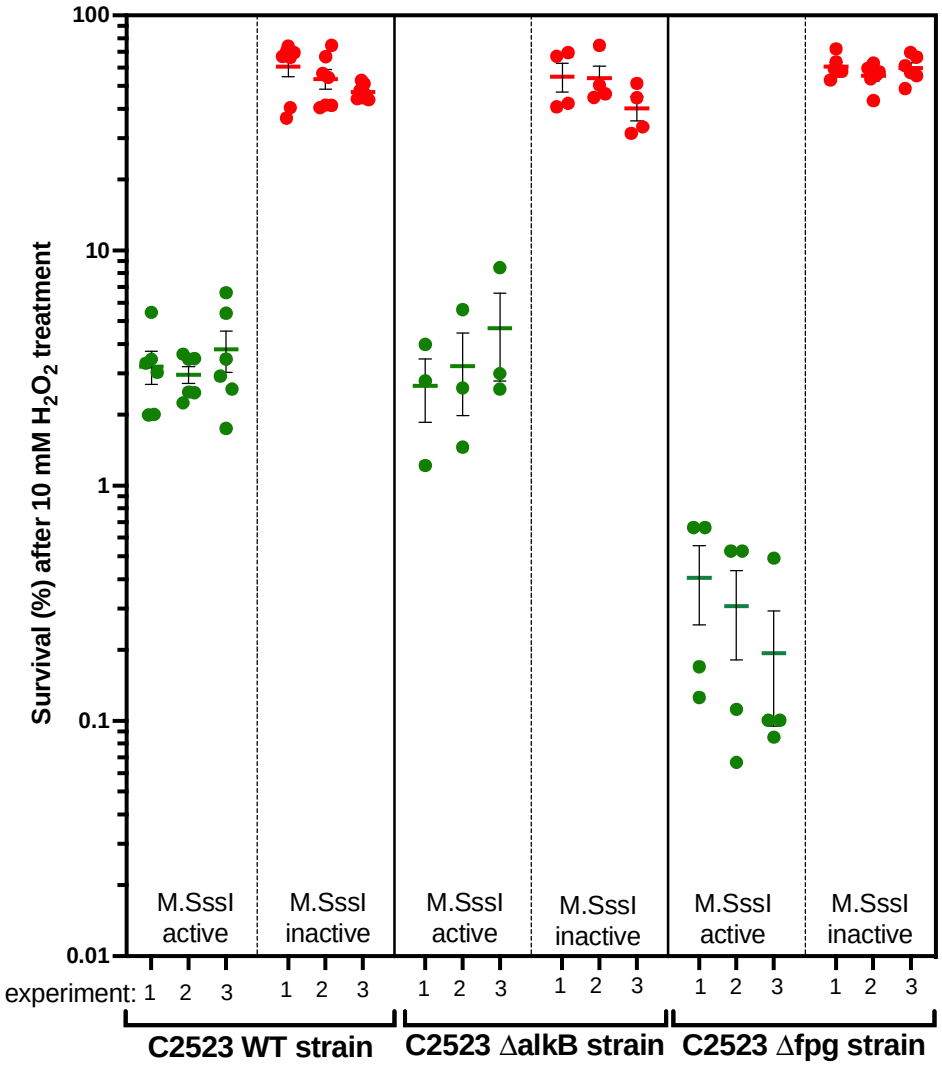


Figure 3 H_2O_2-induced sensitivity in strains carrying an active 5mC methyltransferase.
WT strains and *alkB* and *fpg* mutants with M.SssI methyltransferase were treated with 10 mM H_2O_2. The percentage of survivals was calculated for each strain and normalized to the number of colonies growing under the same conditions without H_2O_2 treatment. Each experiment shows independent clones from the same transformation. Three separate transformations were performed for each plasmid.

Increased ROS production in E. coli expressing DNMTs.

We wondered why DNMT expression might induce sensitivity to oxidative stress. DNMT expression might somehow lead to production of ROS directly. We sought to minimise secondary effects of long-term growth with DNMTs so designed a system whereby we could induce DNMT expression. We expressed M.SssI under the control of the pBAD promoter (methods). When grown in glucose, M.SssI remained silent, but was induced by adding Arabinose (Supplementary Figure 2). We measured ROS production using a commercially available fluorescent sensor (see methods). We initiated expression of DNMT by adding arabinose and followed ROS accumulation over time under conditions in which baseline ROS

accumulation was minimised (see methods and Supplemental Figure 3). As expected, H_2O_2 treatment led to an accumulation of ROS. H_2O_2-induced ROS accumulation was markedly enhanced by addition of DNMT(Figure 4). Moreover, DNMT addition alone led to increased ROS accumulation, even without exposure to H_2O_2 (Figure 4). Interestingly, the effect of H2O2 and ROS was greater than the linear combination of the two effects separately (linear model interaction P-value<0.05). Together these data strongly suggested that DNMT increased ROS levels, and that there was a synergistic effect with H_2O_2.

The oxidative stress response relies on control of gene expression, in particular through the regulator oxyR which leads to transcriptional upregulation of a number of oxidative stress response genes including catalase which detoxifies H_2O_2 and reduces ROS production(Chiang & Schellhorn, 2012). Although 5mC is not recognised to have effects on transcription in *E. coli*, we considered the possibility that 5mC perturbed transcription of oxidative stress response genes, hence leading to failure to detoxify ROS. We therefore compared gene expression in the presence or absence of DNMT to test whether the transcriptional response to H2O2 treatment was impaired.

Induction of DNMT expression during logarithmic growth led to abundant transcriptional changes (Supplementary Figure 4A), suggesting that 5mC may have effects on gene expression. Despite our inducible system, some of these changes are likely to be secondary effects. For example, we observed upregulation of AlkB and its transcriptional regulator ada, although not aidB which is also regulated by ada, upon DNMT induction(Mielecki & Grzesiuk, 2014), which likely reflects production of alkylation damage in the form of 3mC by DNMT expression (Supplementary Figure 4A; Supplemental Table 4). However, the induction of oxyR regulated genes such as catalase was unaffected by 5mC (Supplementary Figure 4B). Thus, the increased sensitivity to H_2O_2 is not caused by failure of detoxification gene induction by DNMT expression.

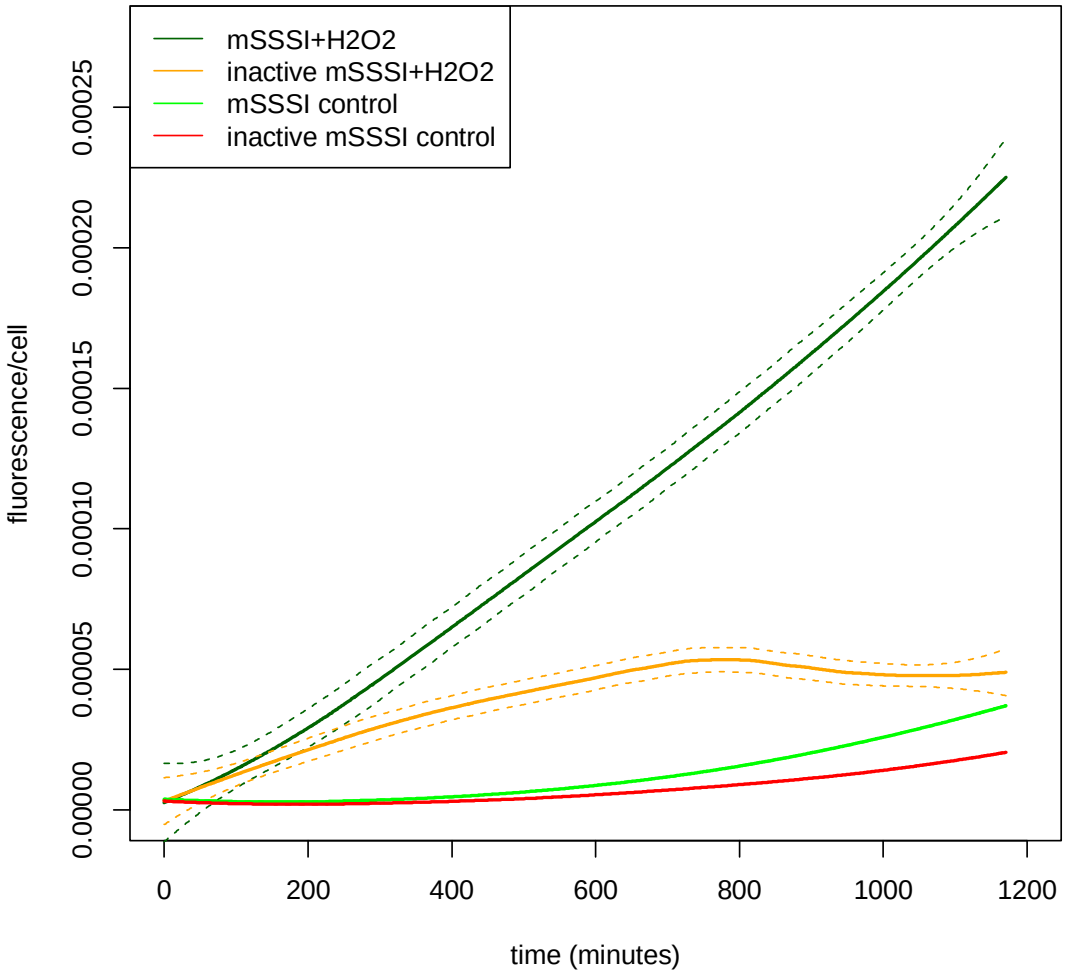


Figure 4 Reactive Oxygen Species production in the presence of M.SssI and H_2O_2. M.SssI was either induced with arabinose or kept repressed at the beginning of the experiment. Fluorescence per cell is plotted over 1200 minutes for each of the indicated conditions. The lines represent the loess fit from pooled data across two replicate experiments (two separate clones from a single transformation for each condition). The dashed lines show the standard error of the loess fit. Note that the error of the fit is too small to be visible as separate lines for conditions with very low reactive oxygen species.

DNMT sensitizes cells to ROS through 5hmC and 5fC production

The synergistic effect of DNMT expression with deletion of Fpg (Figure 3) suggested that DNA damage might be implicated in the increased sensitivity to ROS exhibited by DNMT-expressing cells. One possibility is that oxidation of 5mC might result in damaged DNA bases, which would not occur in the absence of DNMT expression and thus might account for increased sensitivity to oxidative stress. Supporting this hypothesis we observed that exposure of DNMT-expressing cells to H_2O_2 led to production of the 5mC oxidation products 5hmC and 5fC as measured by mass spectroscopy (Figure 5). In eukaryotes enzymes of the TET family catalyse sequential oxidation of 5mC to 5hmC and 5fC(He et al., 2011; Ito et al., 2011; Tahiliani et al., 2009). *E. coli* does not have TET enzymes. However, ROS have been shown to act non-enzymatically on 5mC to convert it directly to either 5hmC or 5fC(Madugundu et al., 2014). 5hmC is not recognised as DNA damage, but 5fC is a substrate for base excision repair(Bochtler et al., 2017). Moreover, its presence in the DNA might indicate the presence of other 5mC oxidation products which our mass spectrometry did not measure, which could also be recognised by cells as DNA damage.

We speculated that the production of 5fC and potentially other, downstream oxidation products of 5hmC or 5fC induced by H_2O_2 treatment might contribute to the excess sensitivity of DNMTs to H_2O_2 treatment. On the basis of this hypothesis, we predicted that artificially inducing 5hmC and 5fC production from 5mC would increase downstream oxidation products further and thus further increase DNMT sensitivity to ROS. To test this we expressed TET from *Naegleria gruberi*(Hashimoto et al., 2013) in *E. coli* under the control of an IPTG-inducible promoter, to induce sequential formation of 5hmC and 5fC from 5mC. NgTET expression alone a small but significant effect on sensitivity to H_2O_2 treatment, which may be due to the ability of NgTET to oxidise T to 5hmU(Pfaffeneder et al., 2014). However, when combined with DNMT expression there was a strong increase in sensitivity beyond what would be expected from either DNMT expression or NgTET expression alone (Interaction Tet*DNMT*IPTG p=1e-10 Anova; Figure 6A,B). Indeed, under the conditions of our assay there was essentially no growth at all when H_2O_2 was present- a synthetic lethal interaction (Figure 6A,B). We obtained similar results using a different TET enzyme, the phage-derived TET43 (Figure 6A) The production of 5hmC, 5fmC and further, downstream, oxidation products derived from 5mC may thus contribute to the excess sensitivity of cells expressing DNMTs to oxidative stress.

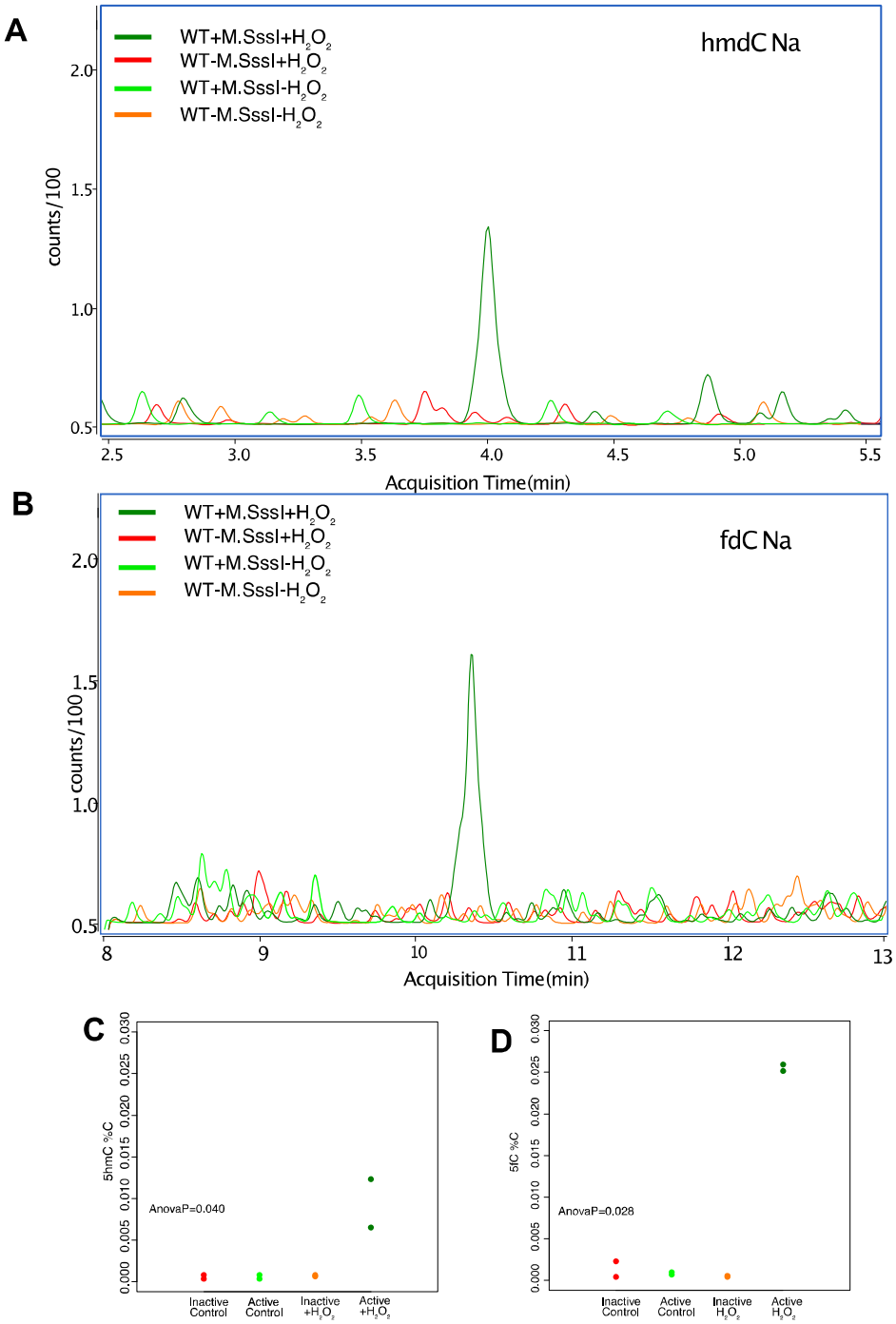


Figure 5- Production of 5hmC and 5fC from upon H2O2 treatment in the presence of M.SssI. A and B show representative mass spectrometry traces for peaks identified as 5hmdC and 5fdC respectively. Quantitation is in figure C and D, where data points are the mean of two technical replicates and each data point is a separate clone from the transformation. Y-axis shows level of 5hmC or 5fC measured by mass spectrometry as a percentage of unmodified cytosine. P-value from a one way anova for effect of the experimental condition on the measurement is indicated on each plot.

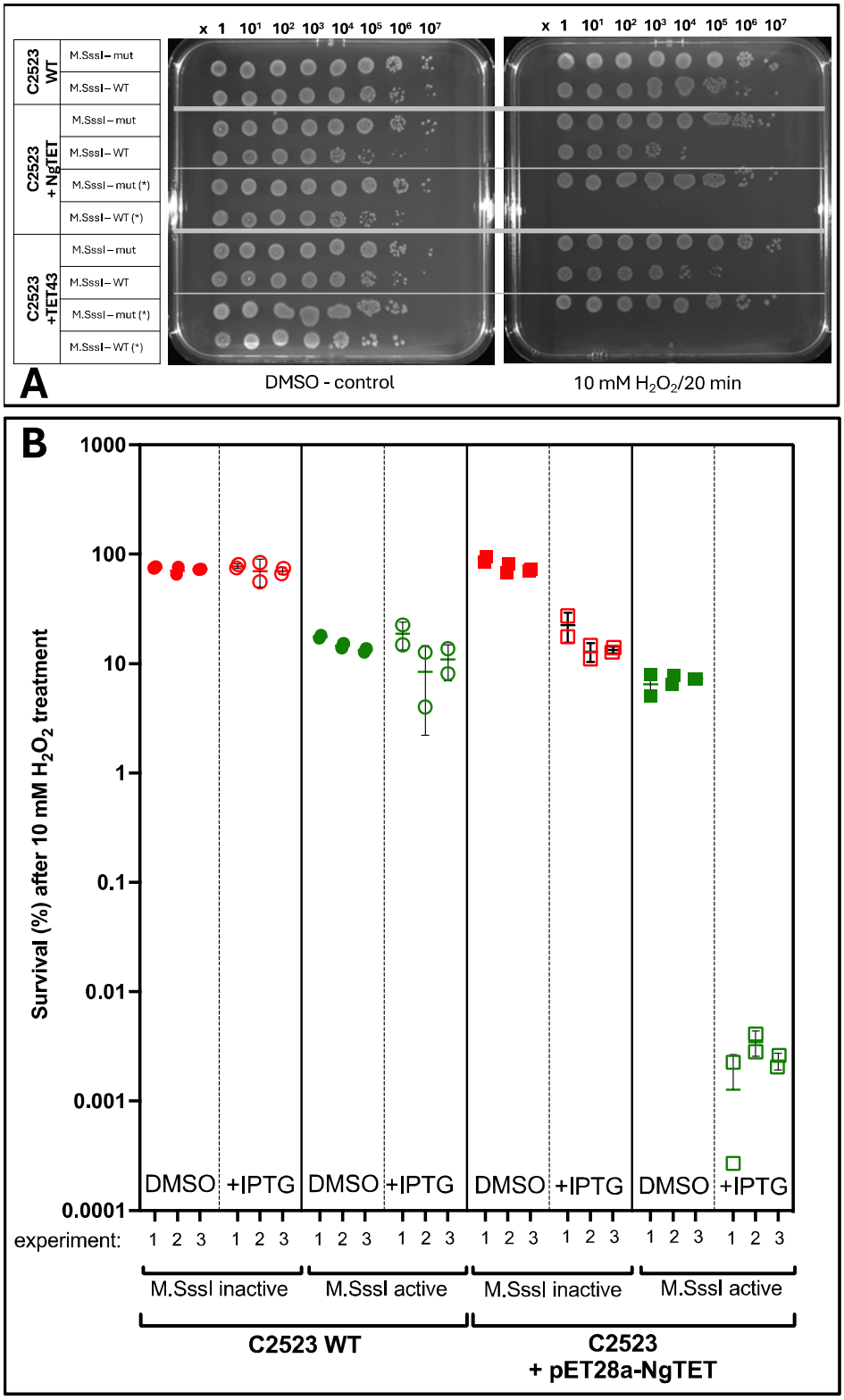

A
x 1 10¹ 10² 10³ 10⁴ 10⁵ 10⁶ 10⁷
C2523 WT
M.SssI – mut
M.SssI – WT
C2523 + NgTET
M.SssI – mut
M.SssI – WT
M.SssI – mut (*)
M.SssI – WT (*)
C2523 +TET43
M.SssI – mut
M.SssI – WT
M.SssI – mut (*)
M.SssI – WT (*)
DMSO - control
10 mM H2O2/20 min
B
Survival (%) after 10 mM H2O2 treatment
1000
100
10
1
0.1
0.01
0.001
0.0001
DMSO
+IPTG
experiment:
1 2 3
M.SssI inactive
M.SssI active
C2523 WT
C2523
+ pET28a-NgTET

Figure 6 H_2O_2-induced sensitivity in strains carrying an active 5mC methyltransferase and expressing TET responsible for oxidative derivatives of 5mC.
The WT strain and a strain carrying IPTG-inducible TETs (Negleria TET – NgTET or phage TET – TET43, more details in Materials section), which modifies 5mC to its oxidative derivatives that, in addition, also overexpressed inactive or active M.SssI methyltransferase, were treated with 10 mM H_2O_2. **A** – Drop plate test for *E. coli* cells without and after 10 mM H_2O_2 treatment (more technical details on Figure 3). **B** – The percentage of survivals from WT and NgTET cells was calculated for each strain and each condition and normalized to the number of colonies growing under the same conditions without H_2O_2 treatment. Each set of 3 points represents 3 independent clones from a single transformation and three separate transformations were performed.

Discussion

In addition to its roles in regulating transcription epigenetically, cytosine-5 DNA methylation is increasingly recognised as a source of DNA damage. The heterologous DNMT expression system that we established in this manuscript enabled us to probe how the propensity of cytosine DNA methyltransferases to damage DNA affects sensitivity to exogenous genotoxic stress. Below we discuss how our findings indicate a novel connection between DNMTs and DNA damage through their impact on oxidative stress sensitivity. We also discuss the implications of our results for the evolution of DNMT activity and expression.

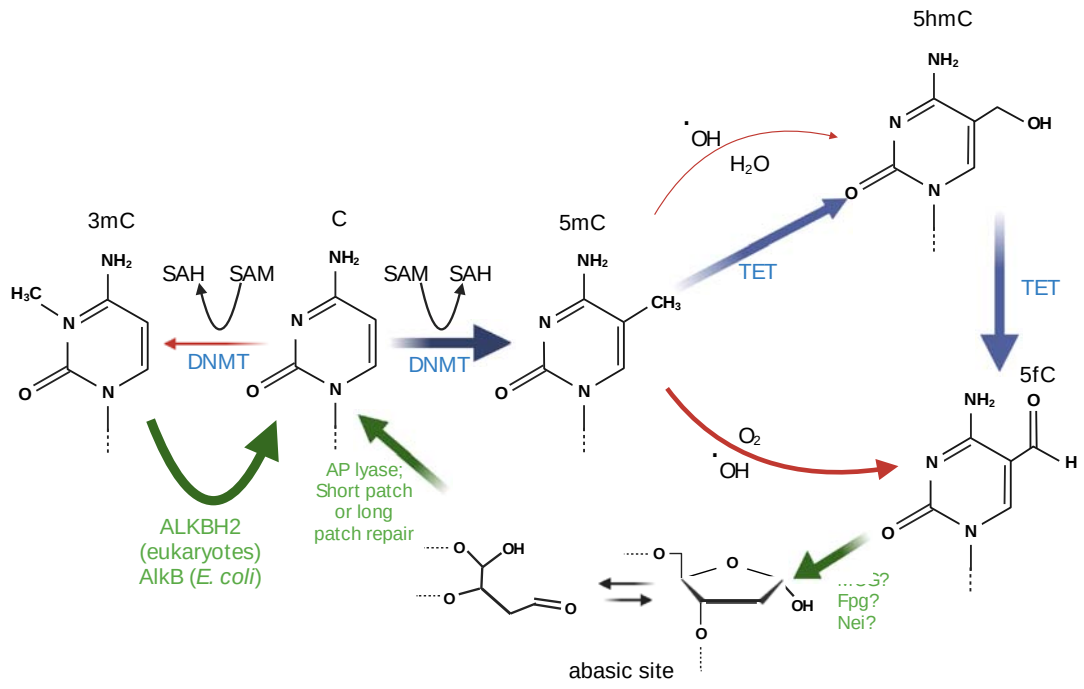
Oxidative stress as a new source of DNA damage induction by DNMT expression

A major new finding of our paper is that DNMT expression sensitizes cells to oxidative stress. Importantly, we demonstrated that this does not proceed through the induction of DNA alkylation damage (3mC) by DNMTs because AlkB did not exacerbate the sensitivity. What, then, might be the mechanism? A straightforward explanation is offered by our observation that H_2O_2 treatment increases the levels of 5hmC and 5fmC in the DNA. Previously, free radical chemistry has been demonstrated to lead to formation of these oxidation products(Madugundu et al., 2014), thus we propose that the same process occurs in our system when H_2O_2 is added via H_2O_2-induced ROS formation. The relatively higher levels of 5fC rather than 5hmC that we observed is consistent with this suggestion as $OH^{\cdot}$ radicals were shown to produce 5fC directly from 5mC(Madugundu et al., 2014), in contrast to the enzymatic oxidation which proceeds sequentially(Ito et al., 2011). In mammals 5fC is a lesion in the DNA which is excised by base excision repair, most prominently by the activity of TDG aided by NEIL1 and NEIL2(Bochtler et al., 2017) Whilst no direct counterpart of TDG exists in *E. coli*, mismatched-uracil glycosylase (MUG) is structurally similar. Moreover, NEIL1 and NEIL2 are related to Fpg in *E. coli*(Jacobs & Schär, 2012). We provided further evidence for this model by accelerating 5fC formation through co-expression of TET with DNMT, which resulted in synthetic lethality upon exposure to H_2O_2. Importantly, although our mass spectrometry investigations were limited to 5hmC and 5fC their presence likely indicates that other oxidation products of 5mC would be created which would also be substrates for glycosylases(Slyvka et al., 2017).

This explanation would suggest that the excess sensitivity to oxidative stress caused by DNMT expression is due to 5mC being oxidised to form modified nucleotides that may be recognised and excised as DNA damage (Figure 7). However, we also observed that DNMT expression itself increased the levels of reactive oxygen species. This occurred even in cells

that had not been subjected to exogenous oxidative stress. A key question therefore is how this is linked to the excess sensitivity of DNMT-expressing cells.

Figure 7- Model: sources of DNA damage due to DNMT expression. Repair pathways are shown in green, enzymatic pathways in blue and DNA damage mechanisms in red.



One possibility is that DNMTs through their catalytic mechanism somehow induce ROS production. This would result in higher endogenous ROS levels when DNMT is expressed, so that exogenous oxidative stress overwhelms cellular detoxification systems resulting in increased sensitivity. In this model the ROS produced by DNMTs are an additional source of toxicity beyond 5mC oxidation. The increased ROS could also participate in 5mC oxidation conversion, in addition to other forms of DNA, RNA and protein damage. ROS production as part of the DNMT catalytic mechanism has not been documented, so further work on the enzymology would be required to test this hypothesis.

An alternative hypothesis is that the increased ROS levels due to DNMTs are directly due to the presence of 5mC and its ability to react with radical species to form 5hmC and 5fC(Madugundu et al., 2014). These reactions proceed via a radical cytosine derivative, with alternative pathways potentially generating other free radicals via reactions with oxygen or water in addition to the formation of 5mC and 5hmC. As a result, it is possible that the presence of 5mC itself could be sufficient to result in increased levels of ROS, by participating in chain reactions that interfere with the ability of cellular detoxification mechanisms such as catalase to remove the free radicals. The advantage of this mechanism is that it directly links the increased ROS levels we observed in DNMT-expressing cells to the increased genotoxic stress via 5fC formation, thus removes the need to invoke a speculative

mechanism for direct ROS production through DNMT activity. A more detailed chemical and kinetic study of the ROS formation in DNMT-expressing cells would be needed to resolve these questions.

Implications of the effect of DNMTs on sensitivity to genotoxic stress for the evolution of DNMTs

The key enzymes that introduce and maintain cytosine DNA methylation in animals are found across the eukaryotic tree indicating an ancient origin at the dawn of eukaryotic cells(Ponger & Li, 2005). However, the evolution of DNA methyltransferase pathways across eukaryotes is strikingly rapid, with many different lineages having lost DNMTs completely(de Mendoza et al., 2020). To what extent might the propensity of DNMTs to damage DNA be responsible for the frequent losses of DNMTs? Due to the coevolution of DNMTs with the alkylation repair enzyme ALKBH2, we speculated that DNMT's off-target activity may pose a threat to genome stability which imposes a fitness cost, even in the presence of the ALKBH2 repair pathway(Sarkies, 2022a). A key prediction of this model is that DNMT expression would be disadvantageous to cells in the absence of ALKBH2 activity. Our *E. coli* model of DNMT expression enabled us to test this and indeed we found that deletion of AlkB leads to a large increase in sensitivity to the alkylating agent MMS when combined with DNMT expression. Thus the propensity of DNMTs to alkylate DNA could lead to a fitness cost at least in the presence of environmental sources of alkylation damage. It is notable that there is also a small decrease in fitness due to combined expression of DNMT and deletion of AlkB, even without MMS treatment. Nevertheless, for this to make loss of DNMT expression favourable the benefit of DNMT expression in terms of other functions such as gene expression would have to be quite small.

High concentrations of alkylating agents such as MMS are not a common component of ecologically relevant niches. However, oxidative stress is much more commonly experienced by organisms(Samet & Wages, 2018). Our observation that DNMT expression renders *E. coli* hypersensitive to oxidative stress may therefore provide a plausible reason why organisms with DNMTs might be at a selective disadvantage in certain environments. Switching into an environment with higher levels of oxidative stress might therefore result in the cost of DNMT expression outweighing its benefit leading, over evolutionary time, to its loss. It would be very interesting to establish the environments of species with and without DNMTs to test whether exposure to oxidative stress might be associated with high rates of DNMT loss.

Materials and methods

Bacterial strains

Escherichia coli B wild-type (WT) strain C2523 which does not cleave methylated DNA (*mcrA*$^{-}$, *mcrBC*$^{-}$, *mrr*$^{-}$, *dcm*$^{-}$, *dam*$^{+}$) with the genotype *fhuA2 [lon] ompT gal sulA11 R(mcr-73::miniTn10--Tet)2 [dcm] R(zgb-210::Tn10--Tet) endA1 Δ(mcrC-mrr)114::IS10* (New Engalad Bioloabs, NEB) was used in this study. Single deletions of *alkB* gene (*ΔalkB::kan*^{R}) and *fpg* gene (*Δfpg::kan*^{R}) were carried out by P1-mediated transduction using the corresponding knockout strains from the Keio collection as donors(Baba et al., 2006). In this system the targeted gene is replaced with the kanamycin gene flanked by the FRT sequence for

subsequent removal of kanamycin cassette. To remove the kanamycin cassette, pCP20, an ampicillin and chloramphenicol resistant plasmid with temperature-sensitive replication and flippase recombinase synthesis(Cherepanov & Wackernagel, 1995) was transformed into the *ΔalkB::kan*^{R} mutant and selected at 30°C, after which pCP20 was purified at 43°C. Removal of knockout genes and the antibiotic resistance gene were confirmed by PCR using and by culturing the strains in antibiotic containing LB media.
To control DNA methyltransferase expression from arabinose inducible promotor *E. coli* TOP10 K12 (NEB) strain with the genotype *F- mcrA (mrr-hsdRMS-mcrBC) 80lacZ M15 lacX74 recA1 ara 139 (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG* was used. This strain also does not cleave 5mC methylated DNA.

Plasmids

The bacterial CpG DNA 5-methylcytosine (5mC) methyltransferase from *Spiroplasma sp.* (M.SssI) (UNIOROT accession number P15840) cloned into plasmid pAIT2, a low copy number plasmid with kanamycin (Kan) and chloramphenicol (Chl) resistance was used in this study. M.SssI is expressed under lacZ promoter. The sequence of M.SssI ORF is 1158 bps long and has 386 amino acids(Renbaum et al., 1990). pAIT2 with truncated M.SssI (M.SssI-inactive) was created by removing the first 570 bps of M.SssI, which includes the SAM binding domain of the protein. The full length of M.SssI coding sequence was re-cloned into pBAD plasmid with ampicillin (Amp) resistance to control DNA methyltransferase expression from arabinose (P_{BAD}) promoter by addition of arabinose (ara) to the media(Guzman et al., 1995). Plasmid with bacterial CpG DNA 5-methylcytosine (5mC) methyltransferase from *Mycoplasma penetrans* (M.MpeI) was kindly provided by Matthias Bochtler (IIMCB, Warsaw, Poland). Briefly, the 1185 bp sequence encoding 395 aa long M.MpeI protein was cloned into pET-28a (Novagen, Km^R) plasmid. This M.MpeI variant differed from the UNIPROT protein (accession Q8EVR5) by the Q68R, K71R, S295P point mutations without affecting an enzymatic properties(Wojciechowski et al., 2013).
The degree of CpG methylation from pAIT2-M.SssI-active, pAIT2-M.SssI-inactive, pBAD-M.SssI-active and pET28a-M.MpeI-active plasmids were assessed by restriction endonuclease digestion with the methylation sensitive enzyme, HpaII (CCGG target sequence), or the methylation insensitive isoschizomeric endonuclease, MspI and subsequent agarose gel electrophoresis monitoring visualized by GelDoc system (BioRad).
The heterolobosean amoeboflagellate *Naegleria gruberi* ten-eleven translocation (TET) methylcytosine dioxygenase (NgTET) cloned into IPTG-inducible pJS119K plasmid (Km^R) and the phage-derived TET43(Burke et al., 2021) were kind gift from Lana Saleh (New England Biolabs, Massachusetts, USA).

Bacterial media and culture conditions

All strains were cultured in Luria-Bertani (LB) medium (1% NaCl, 0.5% yeast extract, 1% tryptone) with or without appropriate antibiotics (Kanamycin (Kan) 50 µg/ml, Ampicillin (Amp) 100 µg/ml, Chloramphenicol (Chl) 30 µg/ml) at 37°C with 220 rpm. Similarly, the bacteria were plated in LB plates (1.6% agar) with or without the aforementioned concentrations of antibiotics and incubated at 37°C overnight.
In order to tightly control expression of M.SssI methyltransferase under arabinose promoter P_{BAD} the strain TOP10 with pBAD-M.SssI-WT plasmid was cultured in LB supplemented with 0.5% glucose (to hold the expression) or 0.5% L-arabinose (to induce protein production) additionally supplemented with Amp. Briefly, a strain with inducible DNA methyltransferase

was cultured overnight in LB medium supplemented with 0.5% glucose to maintain promoter suppression. The bacteria were then diluted 1:100 and a new culture was started under the same conditions to an OD600 of about 0.3. The bacteria were then centrifuged and washed, and the medium was changed to LB with 0.5% arabinose to induce protein expression from the P_{BAD} promoter.

Treatment with DNA damaging agents

MMS acute treatment

Sensitivity to alkylation DNA damage was tested for C2523 WT and *ΔalkB* strains expressing either low (pAIT2) copy number plasmids containing M.SssI or its inactive version (M.SssI-tc) or high copy number plasmid (pET28a) containing active M.MpeI. All experiments were repeated at least 3 times, each time with 3 repeats of each condition, always including WT and mutant parent strains. Bacteria were cultured in LB medium with appropriate antibiotics to OD600 0.6 to 0.8. One half of the culture served as untreated control and the other was treated with 20 mM methylmethane sulphonate (MMS, Sigma) for 30 min, during which both sets of the bacterial culture continued to grow at 37°C with 220 rpm. After treatment, cultures were directly and quickly diluted in PBS to allow visualization of individual colonies and placed on LB agar plates supplemented with the appropriate antibiotic. After overnight incubation at 37°C, the colonies were counted and the colony forming unit (CFU/ml) was calculated. To normalize the effect of MMS, the percentage of survival was calculated by taking the number of bacteria growing under control conditions as 100% and calculating the percentage surviving after mutagen treatment.

Cisplatin and H_2O_2 treatment

Cisplatin and H_2O_2 were carried out following the protocol for MMS acute treatment. Cisplatin (Merck) was dissolved in DMF (Sigma) and the cells were treated with 100 µM cisplatin for 1h. 30% solution of H_2O_2 (Sigma) was used to treated bacteria with final concentration of 10 mM H_2O_2 for 20 min.

Statistical analysis was performed in R. A two way anova was used to test the effect of M.SssI expression on sensitivity to DNA damaging agents and a three way anova used to incorporate the effect of specific mutations on the effect of M.SssI. Presence of a synthetic interaction was determined by significance of the relevant interaction term such that the combined effect of the mutant and the M.SssI on sensitivity was greater than additive i.e. greater than the sum of the mutant and M.SssI alone. The code used took the form: Model<-aov(growth~drug*M.SssI*mutant) where %growth is the number of colonies normalised to the WT control from the same experiment.

Drop plate test for quick enumeration

The drop plate method after mutagen treatment was perform for quick enumeration of *E. coli* cells.

WT strains and *alkB* and *fpg* mutants with M.SssI inactive and active methyltransferases were cultured to reach OD600 of 0.4 and then treated with: DMSO (as control), 20 mM MMS, 10 mM H_2O_2 or 100 µM cisPt, next the series of dilutions were made, and the drops (5 µl) were plated on LB (Chl/Kan) square plates. The dose of spots from left to right corresponded to 1, 2, 3, 4, 5, 6, 7 log10 dilutions of *E. coli.*

Reactive Oxygen Species (ROS) production monitoring

WT Top10 cells were made competent and transformed with the pBAD M.SssI construct or dH_2O. The transformed bacteria were plated on LB High Salt Agar + 0.8% Glucose (+100µg/mL Carbenicillin for pBAD M.SssI transformants) and incubated overnight at 37°C. Individual colonies from Top10 and Top10 + pBAD M.SssI plates were then inoculated overnight in LB High Salt media +0.5% Glucose (+100µg/mL Carbenicillin for pBAD M.SssI transformants). The next day, cultures were diluted 25X and grown to OD600 = 0.4. 3ml of the original overnight culture was taken for plasmid DNA extraction and methylation analysis. Cultures were then split into 3x6ml, centrifuged at 4700 rpm for 3 mins and resuspend in LB High Salt media + 0.5% Arabinose, 0.5% Arabinose + 0.15% Glucose or 0.15% Glucose respectively. Cultures were grown to OD600 = 0.6 and split into 2x3ml. Cultures were treated with 10mM or 0mM H2O2 then incubated at 37°C , 200rpm for 30 mins. 3ml was then taken for plasmid DNA extraction and methylation analysis. In a 96 well plate, 2 x 100µL was plated for each condition. 2µL 50X ROS Dye from the Total Reactive Oxygen Species (ROS) Assay Kit (520nm, Cat. No. 88-5930) was added to each well. The plate reader was heated to 37°C and fluorescence measured by excitation at 475nm and emission at 500-550nm. Absorbance was then measured by OD600. The plate was then shaken for 10 mins at 200 cycles/min (orbital, 2mm shaking diameter). This loop was repeated 100 times.

Genomic DNA isolation and Mass Spectrometry analysis of DNA epigenetic modifications

C2523 WT cells expressing M.SssI methyltransferase or its inactive version (M.SssI-tc) as well as M.MpeI metyltransferase, both H_2O_2 treated and untreated, were cultured overnight in 10ml of LB at 37C with shaking. Cells were lysed and proteins simultaneously denatured in the appropriate lysis buffer according to the commercial kit DNeasy Blood & Tissue Kit (Qiagen, cat. no. 69504). Proteinase K then was added and after incubation in 55 °C for 30 min, lysates were loaded onto the mini columns and centrifuged. After several washing steps according to the protocol, pure, high-molecular-weight DNA was eluted and precipitated with isopropanol. Then, precipitates were diluted in water and 1 µg of pure DNA was checked for its quality and quantity on Genomic DNA ScreenTape (Agilent, cat. no. 5067-5365) using TapeStation 4150 (Agilent).

DNA nucleoside analysis by mass spectrometry (HPLC–QQQ)

1500ng of DNA was treated with 5U DNA Degradase Plus™ (Zymo Research, E2020) in 120µl 1X DNA Degradase Plus™ Reaction Buffer. The mixture was incubated at 37°C for four hours and then lyophilised by SpeedVac to a volume of 3µl. 14.5µl of bufferA (10 mM ammonium acetate, pH 6) was added and the mixture was then filtered through a 96-well filter microplate with a 0.2µm polyethersulfone membrane via centrifugation at 3200g for 35 minutes at 4°C. Finally, a further 4µl of bufferA was added to the approximately 10µl of filtrate following centrifugation.

For the analysis by HPLC–QQQ mass spectrometry, a 1290 Infinity UHPLC was fitted with a Zorbax Eclipse plus C18 column, (1.8µm, 2.1mm 15mm; Agilent) and coupled to a 6495a Triple Quadrupole mass spectrometer (Agilent Technologies) equipped with a Jetstream ESI-AJS source. The data were acquired in dMRM mode using positive electrospray ionisation (ESI1). The AJS ESI settings were as follows: drying gas temperature 230°C, the drying gas flow 14 $lmin^{-1}$, nebulizer 40 psi, sheath gas temperature 400°C, sheath gas flow 11 $lmin^{-1}$, Vcap 2,000 V and nozzle voltage 0 V. The iFunnel parameters were as follows: high pressure RF 110 V, low pressure RF 80 V. The fragmentor of the QQQ mass spectrometer was set to

380 V and the delta EMV set to +200. The gradient used to elute the nucleosides started by a 5-min isocratic gradient composed with 100% bufferA (10 mM ammonium acetate, pH 6) and 0% buffer B (100% methanol – 106035, Merck) with a flow rate of 0.400 ml min^{-1} and was followed by the subsequent steps: 8–9 min, 94.4 % A; 16 min 86.3 % A; 17– 21 min 0 % A; 24.3–25 min 100 % A. The gradient was followed by a 3-min post time to re-equilibrate the column. The raw mass spectrometry data was analysed using the MassHunter Quantitative Analysis Software package (Agilent Technologies, version 10.0). The transitions and retention times used for the characterization of nucleosides and their adducts are summarized in [table ref]. For the identification of compounds, raw mass spectrometry data was processed using the dMRM extraction function in the MassHunter software. For each nucleoside, precursor ions corresponding to the $M\text{-}H^+$ and $M\text{-}Na^+$ species were extracted, and the signal observed was quantified against a standard curve. The chromatograms were generated using the MassHunter Qualitative Analysis Software package (Agilent Technologies, version 10.0).

Induction of DNMT activity and preparation of cells for transcriptomics

Bacterial TOP10 cells with pBAD-M.SssI-WT plasmid or empty pBAD plasmid were cultured to stationary phase in 5 ml LB supplemented with 0.1% glucose to maintain promoter suppression and further supplemented with Amp. The cells were then diluted 1:100 and re-cultured to logarithmic phase (OD600 approximately 0.3) in 20 ml LB supplemented with 1% glucose and Amp, centrifuged, washed with glucose and divided in half. One half (10 ml) was cultured in LB with 1% glucose and Amp to maintain suppression of the P_{BAD} promoter, and the other 10 ml was cultured in LB with 0.1% arabinose and Amp to induce expression of DNA methyltransferase 5mC from the P_{BAD} promoter. During the 2 hours of growth, cells growing on glucose were diluted several times after reaching an OD600 of about 1 to 0.5, while cells growing on arabinose and producing M.SssI were unable to double their amount due to the production of a protein that was toxic to them and remained at the OD600 about 0.3). Nevertheless, the cultures were again divided in half, one half (5 ml) was left as a control and the other half (5 ml) was treated with 10 mM H_2O_2 for 20 minutes. After this time, all cultures were centrifuged, glucose or arabinose and H_2O_2 were washed off, and cells were resuspended in 250 µ of PBS, and next used for RNA isolation to RNA sequencing.

RNA isolation

RNA was isolated using TRIzol purification technique. Briefly, the cell pellet was resuspended in 250 µl of PBS, then 1 ml of TRIzol reagent was added per sample. After 5 min of incubation at room temperature (RT) 0.2 ml of chloroform was added followed by 15 s variously shaking and incubation at RT for 3 min. Next, sample was centrifuged in microfuge on maximal speed for 15 min at 4 °C and the aqueous phase (top) of the sample was place into a new tube followed by 0.5 ml addition of 100 % isopropanol for RNA precipitation. After 30 min incubation in RT sample was pelleted by centrifugation at maximal speed for 15 min at 4 °C. The pellet was washed twice with 1 mL of 75% ethanol. Air dried the RNA pellet was resuspended in in RNase-free water (100 µL). Next, RNA was treated with 1 unit of DNaseI (DNAse I recombinant, RNase-free, ROCHE) per sample, incubated at 30 °C for 20 min and transferred to ice. DNaseI was removed from sample by mixing the sample with 350 µl of RLT buffer (QIAGEN RNeasy Mini kit) followed by addition of 350 µl ethanol (96 %) without centrifugation. 700 µl of sample was transferred into RNeasy Mini spin column (QIAGEN RNeasy Mini kit, cat. no. 74104), centrifuged for 15 s at

8000 x g. RNA on mini-column was twice washed with 500 µl of RPE buffer (QIAGEN RNeasy Mini kit). RNA was removed from the column by 30 µl of RNase free water. Quality and quantity of RNA was checked on RNA ScreenTape (Agilent, cat. no. 5067-5576) using TapeStation 4150 (Agilent). 1 µg of RNA was sent for RNA sequencing (Novogene).

Transcriptomic data analysis

RNA sequencing reads were aligned to the *E. coli* genome using bowtie2 using the preset –very_fast. SAM files were processed to BAM files using samtools (Li) and bam files processed to bed files using bedtools. Counts for each annotated *E. coli* gene were calculated by intersecting the *E. coli* gene coordinate gff file with the bed file using bedtools. Subsequently, data analysis was performed in R, using DESeq2 to normalize the count data across samples and calculate log2 fold change for the different comparisons. Scripts for the R analysis are provided on the Sarkieslab github page:
https://github.com/SarkiesLab/EcoliM.SssI_geneexpression

Acknowledgements

We thank Dr Lana Saleh and Professor Matthias Bochtler for sharing plasmids used in this experiment. We thank Dr Ben Berks for help with the pBAD expression system and Dr Stefan Uphoff for sharing clones from the Keio library for knockout generation.

Funding
This work was funded by a grant from the WorldWide Cancer Fund (20-0006) to PS and a grant from the Royal Society (RGS\R2\222189) to PS, Ludwig Institute for Cancer Research to SK. Katie Hains is funded by a PhD scholarship from Lincoln College Oxford.

References

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V, Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*(7463), 415–421. https://doi.org/10.1038/nature12477

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, *2*. https://doi.org/10.1038/MSB4100050/SUPPL_FILE/MSB4100050-SUP-0010.XLS

Baubec, T., Colombo, D. F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A. R., Akalin, A., & Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature*, *520*(7546), 243–247. https://doi.org/10.1038/nature14176

Bewick, A. J., Hofmeister, B. T., Powers, R. A., Mondo, S. J., Grigoriev, I. V, James, T. Y., Stajich, J. E., & Schmitz, R. J. (2019). Diversity of cytosine methylation across the fungal tree of life. *Nature Ecology & Evolution*, *3*(3), 479–490. https://doi.org/10.1038/s41559-019-0810-9

Bewick, A. J., Vogel, K. J., Moore, A. J., & Schmitz, R. J. (2017). Evolution of DNA methylation across insects. *Molecular Biology and Evolution*, *34*(3), 654–665. https://doi.org/10.1093/molbev/msw264

Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research*, *8*(7), 1499–1504. https://doi.org/10.1093/nar/8.7.1499

Bochtler, M., Kolano, A., & Xu, G. L. (2017). DNA demethylation pathways: Additional players and regulators. *BioEssays*, *39*(1), 1–13. https://doi.org/10.1002/BIES.201600178

Burke, E. J., Rodda, S. S., Lund, S. R., Sun, Z., Zeroka, M. R., O'Toole, K. H., Parker, M. J., Doshi, D. S., Guan, C., Lee, Y. J., Dai, N., Hough, D. M., Shnider, D. A., Corrêa, I. R., Weigele, P. R., & Saleh, L. (2021). Phage-encoded ten-eleven translocation dioxygenase (TET) is active in C5-cytosine hypermodification in DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(26). https://doi.org/10.1073/PNAS.2026742118/-/DCSUPPLEMENTAL

Cherepanov, P. P., & Wackernagel, W. (1995). Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, *158*(1), 9–14. https://doi.org/10.1016/0378-1119(95)00193-A

Chiang, S. M., & Schellhorn, H. E. (2012). Regulators of oxidative stress response genes in Escherichia coli and their functional conservation in bacteria. *Archives of Biochemistry and Biophysics*, *525*(2), 161–169. https://doi.org/10.1016/J.ABB.2012.02.007

Darii, M. V, Cherepanova, N. A., Subach, O. M., Kirsanova, O. V, Raskó, T., Ślaska-Kiss, K., Kiss, A., Deville-Bonne, D., Reboud-Ravaux, M., & Gromova, E. S. (2009). Mutational analysis of the CG recognizing DNA methyltransferase SssI: Insight into enzyme–DNA interactions. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1794*(11), 1654–1662. https://doi.org/https://doi.org/10.1016/j.bbapap.2009.07.016

Dasari, S., & Bernard Tchounwou, P. (2014). Cisplatin in cancer therapy: molecular mechanisms of action. *European Journal of Pharmacology*, *740*, 364. https://doi.org/10.1016/J.EJPHAR.2014.07.025

de Mendoza, A., Hatleberg, W. L., Pang, K., Leininger, S., Bogdanovic, O., Pflueger, J., Buckberry, S., Technau, U., Hejnol, A., Adamska, M., Degnan, B. M., Degnan, S. M., & Lister, R. (2019). Convergent evolution of a vertebrate-like methylome in a marine sponge. *Nature Ecology & Evolution*, *3*(10), 1464–1473. https://doi.org/10.1038/s41559-019-0983-2

de Mendoza, A., Lister, R., & Bogdanovic, O. (2020). Evolution of DNA Methylome Diversity in Eukaryotes. In *Journal of Molecular Biology* (Vol. 432, Issue 6, pp. 1687–1705). https://doi.org/10.1016/j.jmb.2019.11.003

de Mendoza, A., Pflueger, J., & Lister, R. (2019). Capture of a functionally active methyl-CpG binding domain by an arthropod retrotransposon family. *Genome Research*, *29*(8), 1277–1286. https://doi.org/10.1101/gr.243774.118

de Mendoza, A., Poppe, D., Buckberry, S., Pflueger, J., Albertin, C. B., Daish, T., Bertrand, S., de la Calle-Mustienes, E., Gómez-Skarmeta, J. L., Nery, J. R., Ecker, J. R., Baer, B., Ragsdale, C. W., Grützner, F., Escriva, H., Venkatesh, B., Bogdanovic, O., & Lister, R. (2021). The emergence of the brain non-CpG methylation system in vertebrates. *Nature Ecology and Evolution*, *5*(3), 369–378. https://doi.org/10.1038/s41559-020-01371-2

Dinglay, S., Trewick, S. C., Lindahl, T., & Sedgwick, B. (2000). Defective processing of methylated single-stranded DNA by E. coli alkB mutants. *Genes and Development*, *14*(16), 2097–2105. https://doi.org/10.1101/gad.14.16.2097

Dukatz, M., Requena, C. E., Emperle, M., Hajkova, P., Sarkies, P., & Jeltsch, A. (2019). Mechanistic Insights into Cytosine-N3 Methylation by DNA Methyltransferase DNMT3A. *Journal of Molecular Biology*, *431*(17), 3139–3145. https://doi.org/10.1016/j.jmb.2019.06.015

Fedeles, B. I., Singh, V., Delaney, J. C., Li, D., & Essigmann, J. M. (2015). The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: Repairing nucleic acid alkylation damage and beyond. *Journal of Biological Chemistry*, *290*(34), 20734–20742. https://doi.org/10.1074/jbc.R115.656462

Feng, S., Cokus, S. J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M. G., Hetzel, J., Jain, J., Strauss, S. H., Halpern, M. E., Ukomadu, C., Sadler, K. C., Pradhan, S., Pellegrini, M., & Jacobsen, S. E. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences*, *107*(19), 8689–8694. https://doi.org/10.1073/pnas.1002720107

Gardiner-Garden, M., & Frommer, M. (1987). CpG Islands in vertebrate genomes. *Journal of Molecular Biology*, *196*(2), 261–282. https://doi.org/10.1016/0022-2836(87)90689-9

Guzman, L. M., Belin, D., Carson, M. J., & Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology*, *177*(14), 4121–4130. https://doi.org/10.1128/JB.177.14.4121-4130.1995

Hashimoto, H., Pais, J. E., Zhang, X., Saleh, L., Fu, Z. Q., Dai, N., Corrêa, I. R., Zheng, Y., & Cheng, X. (2013). Structure of a Naegleria Tet-like dioxygenase in complex with 5-methylcytosine DNA. *Nature 2013 506:7488*, *506*(7488), 391–395. https://doi.org/10.1038/nature12905

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, N., Sun, Y., Li, X., Dai, Q., Song, C. X., Zhang, K., He, C., & Xu, G. L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, *333*(6047), 1303–1307. https://doi.org/10.1126/SCIENCE.1210944/SUPPL_FILE/HE.SOM.PDF

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C., & Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, *333*(6047), 1300–1303. https://doi.org/10.1126/SCIENCE.1210597

Jacobs, A. L., & Schär, P. (2012). DNA glycosylases: in DNA repair and beyond. *Chromosoma*, *121*(1), 1. https://doi.org/10.1007/S00412-011-0347-4

Jurkowski, T. P., & Jeltsch, A. (2011). On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0028104

Keller, T. E., Han, P., & Yi, S. V. (2016). Evolutionary transition of promoter and gene body DNA methylation across invertebrate-vertebrate boundary. *Molecular Biology and Evolution*, *33*(4), 1019–1028. https://doi.org/10.1093/molbev/msv345

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. In *Nature Reviews Genetics* (Vol. 11, Issue 3, pp. 204–220). https://doi.org/10.1038/nrg2719

Lewis, S. H., Ross, L., Bain, S. A., Pahita, E., Smith, S. A., Cordaux, R., Miska, E. A., Lenhard, B., Jiggins, F. M., & Sarkies, P. (2020). Widespread conservation and lineage-specific diversification of genome-wide DNA methylation patterns across arthropods. *PLoS Genetics*, *16*(6). https://doi.org/10.1371/journal.pgen.1008864

Lindahl, T. (1996). The Croonian Lecture, 1996: endogenous damage to DNA. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *351*(1347), 1529–1538. https://doi.org/10.1098/rstb.1996.0139

Madugundu, G. S., Cadet, J., & Wagner, J. R. (2014). Hydroxyl-radical-induced oxidation of 5-methylcytosine in isolated and cellular DNA. *Nucleic Acids Research*, *42*(11), 7450. https://doi.org/10.1093/NAR/GKU334

Mielecki, D., & Grzesiuk, E. (2014). Ada response – a strategy for repair of alkylated DNA in bacteria. *FEMS Microbiology Letters*, *355*(1), 1. https://doi.org/10.1111/1574-6968.12462

Ougland, R., Rognes, T., Klungland, A., & Larsen, E. (2015). Non-homologous functions of the AlkB homologs. In *Journal of Molecular Cell Biology* (Vol. 7, Issue 6, pp. 494–504). https://doi.org/10.1093/jmcb/mjv029

Pfaffeneder, T., Spada, F., Wagner, M., Brandmayr, C., Laube, S. K., Eisen, D., Truss, M., Steinbacher, J., Hackner, B., Kotljarova, O., Schuermann, D., Michalakis, S., Kosmatchev, O., Schiesser, S., Steigenberger, B., Raddaoui, N., Kashiwazaki, G., Müller, U., Spruijt, C. G., … Carell, T. (2014). Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nature Chemical Biology 2014 10:7*, *10*(7), 574–581. https://doi.org/10.1038/nchembio.1532

Ponger, L., & Li, W. H. (2005). Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. *Molecular Biology and Evolution*, *22*(4), 1119–1128. https://doi.org/10.1093/molbev/msi098

Renbaum, P., Abrahamove, D., Fainsod, A., Wilson, G. G., Rottem, S., & Razin, A. (1990). Cloning, characterization, and expression in Escherichia coli of the gene coding for the CpG DNA methylase from Spiroplasma sp. strain MQ1(M Sssl). *Nucleic Acids Research*, *18*(5), 1145–1152. https://doi.org/10.1093/NAR/18.5.1145

Rošić, S., Amouroux, R., Requena, C. E., Gomes, A., Emperle, M., Beltran, T., Rane, J. K., Linnett, S., Selkirk, M. E., Schiffer, P. H., Bancroft, A. J., Grencis, R. K., Jeltsch, A., Hajkova, P., & Sarkies, P. (2018). Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. *Nature Genetics*, *50*(3), 452–459. https://doi.org/10.1038/s41588-018-0061-8

Samet, J. M., & Wages, P. A. (2018). Oxidative Stress from Environmental Exposures. *Current Opinion in Toxicology*, *7*, 60. https://doi.org/10.1016/j.cotox.2017.10.008

Sarkies, P. (2022a). DNA Methyltransferases and DNA Damage. *Advances in Experimental Medicine and Biology*, *1389*, 349–361. https://doi.org/10.1007/978-3-031-11454-0_14

Sarkies, P. (2022b). Encyclopaedia of eukaryotic DNA methylation: from patterns to mechanisms and functions. *Biochemical Society Transactions*, *50*(3), 1179–1190. https://doi.org/10.1042/BST20210725

Schmutte, C., Yang, A. S., Beart, R. W., & Jones, P. A. (1995). Base Excision Repair of U:G Mismatches at a Mutational Hotspot in the p53 Gene Is More Efficient Than Base Excision Repair of T:G Mismatches in Extracts of Human Colon Tumors1. *CANCER RESEARCH*, *55*, 3742–3746. http://aacrjournals.org/cancerres/article-pdf/55/17/3742/2457615/cr0550173742.pdf

Sedgwick, B. (2004). Repairing DNA-methylation damage. In *Nature Reviews Molecular Cell Biology* (Vol. 5, Issue 2, pp. 148–157). https://doi.org/10.1038/nrm1312

Sikora, A., Mielecki, D., Chojnacka, A., Nieminuszczy, J., Wrzesiński, M., & Grzesiuk, E. (2010). Lethal and mutagenic properties of MMS-generated DNA lesions in Escherichia coli cells deficient in BER and AlkB-directed DNA repair. *Mutagenesis*, *25*(2), 139–147. https://doi.org/10.1093/mutage/gep052

Slyvka, A., Mierzejewska, K., & Bochtler, M. (2017). Nei-like 1 (NEIL1) excises 5-carboxylcytosine directly and stimulates TDG-mediated 5-formyl and 5-carboxylcytosine excision. *Scientific Reports 2017 7:1*, *7*(1), 1–13. https://doi.org/10.1038/s41598-017-07458-4

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., & Rao, A. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science (New York, N.Y.)*, *324*(5929), 930. https://doi.org/10.1126/SCIENCE.1170116

Tarantino, M. E., Dow, B. J., Drohat, A. C., & Delaney, S. (2018). Nucleosomes and the three glycosylases: High, medium, and low levels of excision by the uracil DNA glycosylase superfamily. *DNA Repair*, *72*, 56–63. https://doi.org/10.1016/j.dnarep.2018.09.008

Tomkova, M., McClellan, M., Kriaucionis, S., & Schuster-Böckler, B. (2018). DNA Replication and associated repair pathways are involved in the mutagenesis of methylated cytosine. *DNA Repair*, *62*, 1–7. https://doi.org/10.1016/j.dnarep.2017.11.005

Tomkova, M., & Schuster-Böckler, B. (2018). DNA Modifications: Naturally More Error Prone? In *Trends in Genetics* (Vol. 34, Issue 8, pp. 627–638). https://doi.org/10.1016/j.tig.2018.04.005

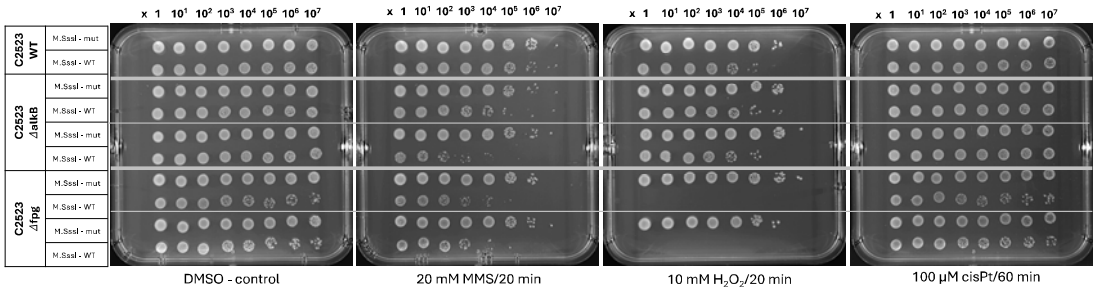
Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., & Sedgwick, B. (2002). Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature*, *419*(6903), 174–178. https://doi.org/10.1038/nature00908

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Pääbo, S., Rebhan, M., & Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*, *39*(4), 457–466. https://doi.org/10.1038/ng1990

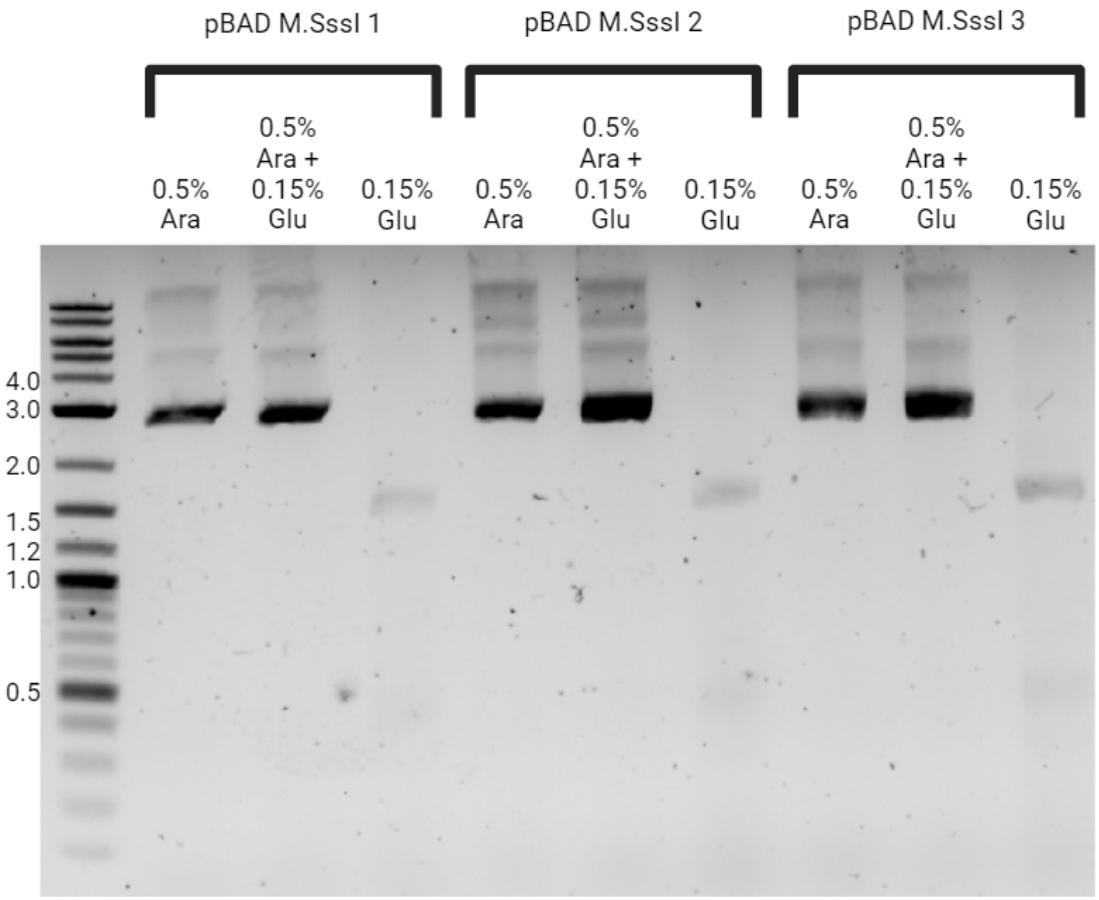
Wojciechowski, M., Czapinska, H., & Bochtler, M. (2013). CpG underrepresentation and the bacterial CpG-specific DNA methyltransferase M.Mpel. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(1), 105–110. https://doi.org/10.1073/PNAS.1207986110/SUPPL_FILE/SAPP.PDF

Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, *328*(5980), 916–919. https://doi.org/10.1126/science.1186366

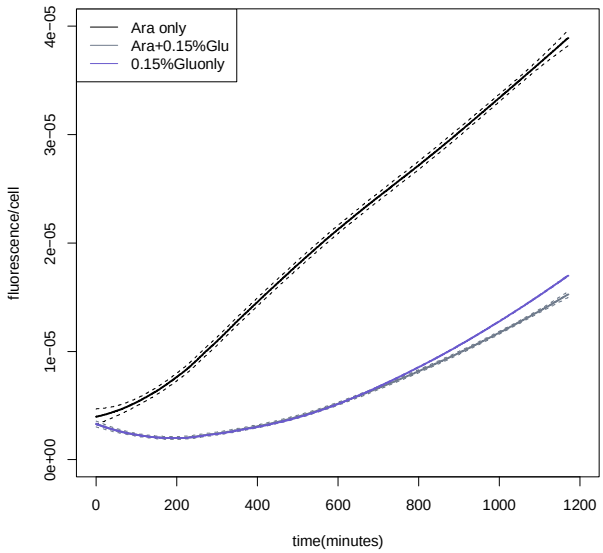
Supplemental Figures



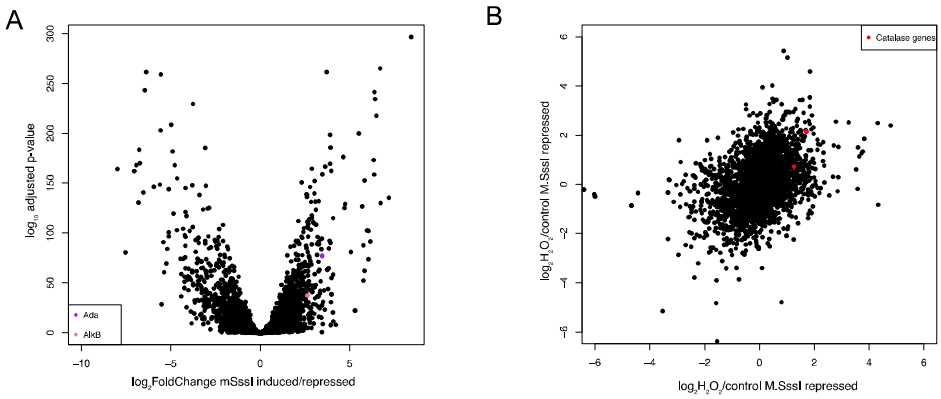
Supplemental Figure 1- Drop test screening sensitivity of M.SssI expressing cells to various DNA damaging agents. WT, alkB mutants and Fpg mutants are shown.



Supplemental Figure 2- Induction of M.SssI expression under control of pBAD promoter by arabinose. The plasmid was digested with methyl-sensitive restriction enzyme HpaII. The presence of the upper band indicates undigested plasmid. Arabinose maintains methyltransferase expression even in the presence of 0.15% glucose, whereas removal of arabinose leads to repression of the methyltransferase. Three independent clones from a single transformation are shown.



Supplemental Figure 3- Removal of glucose affects ROS production in *E. coli*. Fluorescence per cell was measured over 1200 minutes for TOP10 cells in the absence of the M.SssI plasmid. 0.15% glucose is sufficient to repress ROS levels, demonstrating that it is removal of glucose rather than addition of arabinose that causes ROS accumulation.



Supplemental Figure 4- Transcriptomics analysis. **A** shows a volcano plot of genome wide transcript level changes measured by RNAseq comparing M.SssI expressing cells (arabinose) to M.SssI repressed cells (glucose). The x axis shows $\log_2$ fold change and the y axis shows the Benjamani Hochberg adjusted p-value calculated through DESeq2. *alkB* and *ada* are highlighted. **B** shows the log2 fold change in transcript levels upon H2O2 treatment of cells in which M.SssI is repressed (x axis) compared to cells where M.SssI is expressed (y axis). Catalase genes *katG* and *katE* are highlighted.

Supplemental Material

SupTables1.xls Spreadsheets containing all the raw data underlying figures 1,2,3,5 and 6.
SupTables2.xls Spreadsheets containing raw and processed data for figure 4 (ROS measurements)
SupTable3.xls Mass spectrometry data
SupFile4.Rdata Transcriptomics data raw counts