*Published online 19 April 2011* *Nucleic Acids Research, 2011, Vol. 39, No. 14* ***5991–6001***
*doi:10.1093/nar/gkr219*

# Evidence for an evolutionary antagonism between Mrr and Type III modification systems

**Mehari Tesfazgi Mebrhatu[1], Ewa Wywial[2], Anirban Ghosh[1], Chris W. Michiels[1], Ariel B. Lindner[3], François Taddei[3], Janusz M. Bujnicki[2,4], Laurence Van Melderen[5] and Abram Aertsen[1,3,5,*]**

[1]Laboratory of Food Microbiology and Leuven Food Science and Nutrition Research Centre (LFoRCe), Centre for Food and Microbial Technology, Department of Microbial and Molecular Systems ($M^2S$), Faculty of Bioscience Engineering, Katholieke Universiteit Leuven, Kasteelpark Arenberg 22, B-3001 Leuven, Belgium, [2]Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks Trojdena 4, 02-109 Warsaw, Poland, [3]Laboratoire Génétique Moléculaire Evolutive et Medicale, INSERM U571, Universite Rene Descartes-Paris V, rue du Faubourg St. Jacques 24, 75014 Paris, France, [4]Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 89, 61-614 Poznan, Poland and [5]Laboratoire de Génétique et de Bactérienne, Institut de Biologie et de Médecine Moléculaires, Université Libre de Bruxelles, rue des Professeurs Jeener et Brachet 12, 6041 Gosselies, Belgium

Received March 9, 2011; Revised March 23, 2011; Accepted March 24, 2011

## ABSTRACT

**The Mrr protein of *Escherichia coli* is a laterally acquired Type IV restriction endonuclease with specificity for methylated DNA. While Mrr nuclease activity can be elicited by high-pressure stress in *E. coli* MG1655, its (over)expression *per se* does not confer any obvious toxicity. In this study, however, we discovered that Mrr of *E. coli* MG1655 causes distinct genotoxicity when expressed in *Salmonella typhimurium* LT2. Genetic screening enabled us to contribute this toxicity entirely to the presence of the endogenous Type III restriction modification system (StyLTI) of *S. typhimurium* LT2. The StyLTI system consists of the Mod DNA methyltransferase and the Res restriction endonuclease, and we revealed that expression of the LT2 *mod* gene was sufficient to trigger Mrr activity in *E. coli* MG1655. Moreover, we could demonstrate that horizontal acquisition of the MG1655 *mrr* locus can drive the loss of endogenous Mod functionality present in *S. typhimurium* LT2 and *E. coli* ED1a, and observed a strong anti-correlation between close homologues of MG1655 *mrr* and LT2 *mod* in the genome database. This apparent evolutionary antagonism is further discussed in the light of a possible role for Mrr as defense mechanism against the establishment of epigenetic regulation by foreign DNA methyltransferases.**

## INTRODUCTION

In general, DNA restriction endonucleases (REases) team up with DNA methyltransferases (MTases) to form dedicated restriction modification (RM) systems, in which the MTase modifies DNA at specific sequences in order to protect it from cleavage by the cognate REase (1,2). Because of this interplay, bacterial RM systems behave as a primitive immune system in which the REase restricts the cellular entry of DNA lacking the proper methylation signature. As such, it can protect the host against deleterious phage infection or lateral acquisition of foreign DNA (3). However, as the genes encoding such RM systems are typically closely linked and often organized into an operon, they can also behave as selfish addiction modules that counteract their loss from the cell by post-segregational killing (4,5). In this phenomenon, siblings that lost the RM module will experience a decline in corresponding MTase activity, and eventually fail to protect all genomic target sites from cleavage by the residual cognate REase.

Depending on their operational characteristics, RM systems can be categorized into three main types (Type I–III) (6). As such, Type I systems are unique in requiring a dedicated specificity protein that directs activity of the corresponding MTase and REase to the target sequence, with cleavage itself occurring at variable distances from this site. In contrast, Type II MTase and REase pairs independently recognize the same target sequence and most often function as separate entities, with cleavage occurring at a defined position within or close to the recognition site.

*To whom correspondence should be addressed. Tel.: +32 16 32 17 52; Fax: +32 16 32 19 60; Email: abram.aertsen@biw.kuleuven.be

Finally, Type III systems display specificity for a short asymmetric recognition sequence, and cleave at a characteristic distance away from it. Importantly, binding to the target site is strictly mediated by the MTase, so that restriction is performed by a MTase–REase complex rather than by the REase alone. While all RM systems require AdoMet for donation of the methyl group, Type I and III systems additionally depend on ATP for their activity (6).

Interestingly, however, a small number of REases operate in the absence of a cognate MTase and have been grouped into a fourth category (Type IV). These REases invariantly bear a distinct specificity for modified (e.g. methylated, hydroxymethylated or glucosyl-hydroxymethylated) DNA, although their recognition site is often poorly defined (6). Well-known examples of such solitary and modification-dependent REases in *Escherichia coli* are McrA, McrBC and Mrr, of which the corresponding genes are typically located on laterally acquired genetic elements such as e14 (*mcrA*) and the 'Immigration Control Region' (*mcrBC* and *mrr*) (7–9).

While McrA and McrBC are relatively well-characterized, and shown to be involved in restricting phage infection (10–14), the physiological role of Mrr in the cell has been less obvious so far. Recently, however, our interest in Mrr was raised as it appears to be a driving part of the peculiar piezophysiology of *E. coli*. More specifically, it was shown that cellular filamentation and prophage induction caused by sub-lethal high-pressure shock in *E. coli* K12 were the result of a RecBCD-dependent SOS response that was instigated by Mrr endonuclease activity (15–18). Unfortunately, it so far remains unclear how exactly Mrr becomes activated by a physical stress such as high pressure. Further study, however, revealed that not all Mrr homologues behave the same way, and the Mrr protein encoded by *Salmonella typhimurium* LT2, in fact, could not be activated by high-pressure shock (19).

The current study starts with the serendipitous observation that the Mrr protein of *E. coli* MG1655 causes distinct genotoxicity when expressed in *S. typhimurium* LT2, and eventually results in postulating an evolutionary antagonism between Mrr and Type III MTases.

## MATERIALS AND METHODS

### Strains and construction of mutants

*Salmonella typhimurium* LT2 (20, laboratory collection), *Escherichia coli* K12 MG1655 (21, laboratory collection) and *E. coli* ED1A (22, kindly provided by Eric Denamur, Institut Pasteur, Paris, France) were used as parental strains in this study.

An LT2 derivative cured of its four active prophages (i.e. strain MA8508 constructed by Lemire and Bossi, manuscript submitted for publication) was kindly provided by Lionello Bossi (Centre National de la Recherche Scientifique, Gif-sur-Yvette, France), and is referred to as LT2Δφ in this study. Strain LT2K2 is a derivative of LT2 and harbors a *MudK* (23) based translational *lacZ* fusion to the SOS-responsive *STM0912* gene, and its use as an SOS indicator strain has been validated previously (19). LT2 *recA* was constructed by cotransducing the *recA1* and *srl*-202::*Tn*10 alleles from TT521 (kindly provided by John Roth, University of California at Davis, USA) with P22 *HT105/1 int-201* (24). In strain LT2 ΔStyLTI, the StyLTI *mod res* operon was deleted based on the protocol of Datsenko and Wanner (25), and using an amplicon prepared on pKD4 with the following primers: 5′-ATAAGAAATTGATTGCTGCACC TTAAAATTATCAACGGAGCGAAGATATCGTGTA GGCTGGAGCTGCTTC-3′ and 5′-ATTTTATATGGTC TATCAAAAAGATAACCGTTAAGGTTCATTCAAA CCCGCATATGAATATCCTCCTTAG-3′. This procedure replaced the *mod res* operon with an *frt* flanked kanamycin resistance cassette, which could subsequently be excised by transiently equipping this strain with plasmid pCP20 expressing the Flp site-specific recombinase (26). Strain LT2K2 Δ*mod res* was subsequently constructed by transducing *STM0912*::Mu*d*K to LT2 Δ*mod res* using P22 *HT105/1 int-201*.

Strain MG1655 *mrr* harbors an *mrr*::*Tn*5 allele, which was obtained through an *in vitro* transposomics procedure using the EZ-Tn5 transposome kit (Epicentre, Landgraaf, The Netherlands) and an amplicon of the MG1655 *mrr* gene ($mrr^{MG1655}$). The latter was obtained by PCR using the following primers: 5′-CGATTCTAGACTGCAAGG ATGTACTATGACG-3′ and 5′-CGATAAGCTTGCGT TTGCGGGGTTGAGG-3′. The resulting *mrr*::*Tn*5 allele confers kanamycin resistance and was exchanged with the wild-type *mrr* allele according to the protocol of Datsenko and Wanner (25). Strain MG1655 λ*sfiA*::*lacZ* was constructed by lysogenizing strain DJ624 (MG1655 Δ*lac*X74 *mal*::$lacI^q$) (27) with λ*sfiA*::*lacZ*, originating from CSH50 λ*sfiA*::*lacZ* (28). Subsequently, Strain MG1655 λ*sfiA*::*lacZ mrr* was constructed by transducing the *mrr*::*Tn*5 allele from MG1655 *mrr* into MG1655 λ*sfiA*::*lacZ* with P1*vir* (29).

Strains were transformed with the necessary plasmids by electroporation, while curing of temperature-sensitive plasmids, such as pKD46 and pCP20, was performed by growing the corresponding strain at the non-permissive temperature in the absence of plasmid selection and subsequently isolating a clone that had lost the plasmid.

### Growth conditions

Stationary phase cultures were obtained by growth in lysogeny broth (LB) (30) for 21 h at 37°C under well-aerated conditions. Late exponential phase cultures in turn were prepared by diluting stationary phase cultures 1/100 or 1/1000 in fresh pre-warmed LB, and allowing further incubation at 37°C until $OD_{600} = 0.6$ as described earlier (31). When appropriate, the following chemicals (Applichem, Darmstadt, Germany) were added to the growth medium at the indicated final concentrations: ampicillin (100 μg/ml), chloramphenicol (30 μg/ml), kanamycin (50 μg/ml), glucose (0.2 or 0.02 %) and arabinose (0.2 or 0.02%).

Where indicated, quantitative measurement of growth was performed with either a Multiskan Ascent reader (Thermo, Aalst, Belgium) or a Bioscreen C Microbiology

Reader (Labsystems Oy, Helsinki, Finland), which automatically measure $OD_{600}$ of cultures growing in thermostatted microplate wells at regular time intervals.

### Construction of plasmids

Plasmid pFPV25-$P_{BAD}$-$mod^{LT2}$ is derived from the pFPV25 backbone (32), confers ampicillin resistance and conditionally expresses $Mod^{LT2}$ in the presence of arabinose. For its construction, an amplicon of the $mod^{LT2}$ allele was first obtained by PCR on genomic DNA of LT2, using the following primers: 5′- AAGTTCTAGAGGAG CGAAGATATCATGTTG -3′ and 5′- CGTAAAGCTTA TCATCCCCTCTTAATAACG -3′. Subsequently, this amplicon was digested with XbaI and HindIII, and cloned in pFPV25-$P_{BAD}$-*gfp* (i.e. pAA100; 19) digested with the same enzymes to replace the *gfp* gene and yield pFPV25-$P_{BAD}$-$mod^{LT2}$. Similarly, plasmid pFPV25-$P_{BAD}$-$mod^{ED1A}$ was constructed by preparing a PCR amplicon of the $mod^{ED1A}$ allele from genomic DNA of ED1A, using the following primers: 5′- ACGCTCTAGAGATACAGA TACGTGGAA-3′ and 5′- CGTACTGCAGCATCTCAT TCCCTTACAG-3′. Subsequently, this amplicon was digested with XbaI and PstI, and cloned in pFPV25-$P_{BAD}$-*gfp* digested with the same enzymes to replace the *gfp* gene and yield pFPV25-$P_{BAD}$-$mod^{ED1A}$. Plasmid pFPV25-$P_{BAD}$-$mrr^{MG1655}$ was constructed similarly, but using the *mrr*MG1655 amplicon obtained as described above. Plasmid pFPV25-$P_{BAD}$-$mrr^{LT2}$ was constructed earlier in a similar way (19). Where necessary, plasmid pFPV25-$P_{BAD}$ served as a control, and was obtained by removing the *gfp* gene from pAA100 with XbaI and HindIII, and reclosing the resulting fragment by self-ligation after blunting.

Plasmid pACYC184-$P_{BAD}$-$mrr^{MG1655}$ is derived from the pACYC184 backbone (33), confers chloramphenicol resistance and conditionally expresses $Mrr^{MG1655}$ in the presence of arabinose. For its construction, a fragment encoding $P_{BAD}$-$mrr^{MG1655}$ was cut out from the pFPV25-$P_{BAD}$-$mrr^{MG1655}$ vector with EcoRI and HindIII, and subsequently blunted. This fragment was then ligated into the pACYC184 vector, which was first opened with HindIII and BamHI and blunted, to yield pACYC184-$P_{BAD}$-$mrr^{MG1655}$. Plasmid pAA810, which is also pACYC184 based, was constructed earlier (18) and harbors the $mrr^{MG1655}$ locus under the control of its own promoter.

### Screening for suppression of $Mrr^{MG1655}$ lethality in LT2 *recA*

After we established that pACYC184-$P_{BAD}$-$mrr^{MG1655}$ conferred severe lethality in LT2 *recA* when plated on M9 minimal medium (30) containing 0.2 % arabinose, two different screens were initiated to find suppressors of this lethality. In a first screen, LT2 *recA* pACYC184-$P_{BAD}$-$mrr^{MG1655}$ was transformed with a library containing random fragments of MG1655 genomic DNA cloned in pBR322 (kindly obtained from Susan Gottesman, National Institutes of Health, Bethesda, USA). In a second screen, a random Mu*d*J knock-out library was constructed in LT2 *recA* pACYC184-$P_{BAD}$-$mrr^{MG1655}$ using the Mu*d*J transposon delivery system described by Hughes and Roth (23). Next, the corresponding plasmid and transposon libraries were plated on M9 medium complemented with 0.2% arabinose and appropriate antibiotics, and clones able to grow on such plates were isolated and purified. Subsequently, the corresponding pBR322-plasmids were transformed to a fresh LT2 *recA* pACYC184-$P_{BAD}$-$mrr^{MG1655}$ strain, while the corresponding transposon insertions were transduced to a fresh LT2 pACYC184-$P_{BAD}$-$mrr^{MG1655}$ strain after which the *recA1* allele was co-transduced into the resulting strain as described above. These resulting strains were examined to check whether the $Mrr^{MG1655}$ resistant phenotype could be confirmed.

One valid Mu*d*J clone arose from this selection procedure, and its insertion site was determined by subcloning the chromosome of the corresponding mutant in pUC18 (34) and selecting for the kanamycin resistance marker encoded on the Mu*d*J transposon.

### Measuring of SOS induction and DNA damage

SOS induction was measured via standard Miller assays (35), using LT2K2 or MG1655 λ*sfiA*::*lacZ*-based indicator strains. For the biochemical detection of DNA breaks and their distribution throughout a population, terminal deoxyribonucleotide transferase dUTP nick end labeling (i.e. TUNEL assay; in situ Cell Death Detection Kit, Fluorescein, Roche Mannheim, Germany) was used, based on the manufacturer's instructions and the amendments published earlier (36,37). Briefly, 1 ml of a culture was washed two times with phosphate-buffered saline (PBS; 0.14 M NaCl, 2.7 mM KCl, 1.5 mM $KH_2PO_4$, 9.6 mM $Na_2HPO_4$, pH 7.3), and fixed with ice-cold 4% paraformaldehyde for 30 min at room temperature. Cells were then pelleted and resuspended in ice-cold permeabilization solution (0.1% Triton X-100, 0.1% sodium citrate) and incubated for 2 min at 4°C. Again cells were washed one time with PBS and labeled using TUNEL reagents for 1 h in the dark at 37°C. Finally, cells were washed two times with PBS and analyzed using flow cytometry (Guava Easy Cyte Plus, Guava Tehnologies, Inc. Hayward, CA, USA). To generate a positive control, an aliquot of fixed and permeabilized control cells was first exposed to the blunt 4-cutter BsuRI for 1 h at 37°C, before being washed and labeled with TUNEL reagents.

### Constructing the phylogenetic tree of *E. coli* and *S. enterica* strains

To construct the phylogenetic tree of *E. coli* and *S. enterica* strains, all full-length *Escherichia* and *Salmonella* 16 S rRNA gene sequences were downloaded from the Greengenes database core set (38) in the FASTA aligned format. Next, *E. coli* ED1A 16S rRNA gene sequence was identified using NCBI nblast (with *E. coli* CFT073 16S rRNA as a seed) and added manually to the dataset. Gene sequences that could not be successfully aligned originated from partially sequenced genomes or were duplicative, and were removed from the dataset. Next, the dataset was re-aligned using Greengenes (38), which aligns 16S rRNA gene sequences to 7682 characters

full-length gene templates. All thresholds were kept at default values. Finally, based on the alignment, the phylogenetic tree was calculated using MEGA 4.0 (39) employing the Minimum Evolution method and assuming a Jukes–Cantor model of nucleotide substitution. Bootstrap values based on 1000 replications are listed as percentages at the branching points.

### Calculating the anti-correlation of $Mrr^{MG1655}$ and $Mod^{LT2}$ homologues

Sequences of *E. coli* Mrr (gi: 127320) and *S. typhimurium* Mod (gi: 300193) were used as queries for PSI-BLAST (40) searches ($E$-value $\leq 1e^{-3}$) against the nr database using default parameters and run until convergence. We retrieved 2001 Mrr and 6878 Mod sequences. After removing sequences from partially sequenced genomes and redundant sequences, we performed sequence clustering based on pair-wise BLAST similarity scores, using Cluster Analysis of Sequences (CLANS) (41). The clustering was completed at $P = 0.012$ for Mrr and $P = 0.004$ for Mod, respectively. The reported $P$-values give a well-resolved separation of multiple distinct clusters. The $P$-value was chosen empirically, given the $P$-value plot for each data set, which shows a histogram of the number of sequences for each $E$-value below a given value. For example, a cut-off of 0.004 will exclude connections worse than 0.004. Clusters are thought of as robust if small changes in cut-off values do not result in major changes in their content. Next, we extracted members of clusters containing the Mrr query sequence (gi: 127320) and the Mod query sequence (gi: 300193) for correlation analysis. At the end, we obtained 75 $Mrr^{MG1655}$ and 211 $Mod^{LT2}$ homologues derived from 272 fully sequenced genomes and eliminated duplicate species to obtain 45 $Mrr^{MG1655}$ and 156 $Mod^{LT2}$ homologues derived from 192 fully sequenced genomes.

The Pearson $r$ correlation coefficient measures the degree to which values of two variables are linearly related to each other. It is defined as the covariance of two variables divided by the product of their standard deviations and was calculated for the $Mrr^{MG1655}$ and $Mod^{LT2}$ families using STATISTICA 8 (StatSoft, Inc., Tulsa, OK, USA). We also determined the probability that the observed correlation is real and not a chance occurrence. The obtained correlation of $-0.52$ for 45 $Mrr^{MG1655}$ and 156 $Mod^{LT2}$ homologues, derived from 192 fully sequenced genomes, is less than the critical value for $df = 190$, $\alpha = 0.05$, two-tailed test with $P < 0.0001$.

### Determination of the horizontal transfer index for homologs of $mrr^{MG1655}$ and $mod^{LT2}$

Nakamura *et al.* (42) have published an analysis of 116 completely sequenced prokaryotic genomes, in which they calculated an index of potential distant horizontal transfer for all genes, by comparing the frequency of 'words' of pentanucleotide length within each gene with the average word frequency of the entire genome. This index can be used to detect only recent transfers of genes from species with distinct gene sequence compositions, because genes transferred between similar genomes usually display similar sequence compositions and genes transferred a long time ago tend to accommodate their sequence composition to that of the new host. We examined the horizontal transfer index of genes encoding $Mrr^{MG1655}$ and $Mod^{LT2}$ homologues in the updated set of 165 genomes (Yoji Nakamura and Takashi Gojobori, personal communication).

## RESULTS

### Expression of $Mrr^{MG1655}$ is harmless in *E. coli* MG1655 but genotoxic in *S. typhimurium* LT2

Previously, we have shown that, in contrast to MG1655, LT2 is unable to mount the SOS response after high-pressure shock (19). However, during our initial attempts to possibly restore this feature by complementing LT2 with a plasmid encoding the $mrr^{MG1655}$ gene under the control of its own promoter (i.e. pAA810), we obtained colonies with a sick translucent appearance. Moreover, when transformants were grown in liquid broth, lysis of the culture (Figure 1A) and concomitant accumulation of cellular debris could be clearly observed.

As the genome of LT2 contains four functional prophages (20), we postulated that cell lysis was caused by prophage activation due to $Mrr^{MG1655}$-mediated DNA damage. Upon further investigation, an LT2 mutant deprived of its active prophages (LT2Δφ) indeed failed to show lysis (Figure 1B). Moreover, we could clearly detect SOS induction (Figure 1C) and even confirm physical DNA damage (Figure 1D) upon expression of $Mrr^{MG1655}$ in LT2, both revealing the severe genotoxicity of this protein in LT2.

Despite its constitutive toxicity in LT2, it is important to note that $Mrr^{MG1655}$ was harmless upon similar ectopic expression in its natural MG1655 host (data not shown), unless such strains were exposed to high-pressure shock (18). We, therefore, assumed this surprising dissimilarity to stem from differences in the physiology of both strains, resulting either in dedicated mechanisms of $Mrr^{MG1655}$-restriction alleviation in MG1655 or dedicated mechanisms of $Mrr^{MG1655}$ activation in LT2.

### The endogenous StyLTI Type III restriction modification system is responsible for $Mrr^{MG1655}$ toxicity in *S. typhimurium* LT2

In order to find out why $Mrr^{MG1655}$ behaves so differently in closely related species, a selection scheme was devised in which we tried to suppress the lethality of $Mrr^{MG1655}$ in LT2. First of all, delivery of $Mrr^{MG1655}$ was made dependent on the presence of arabinose (using pACYC184-$P_{BAD}$-$mrr^{MG1655}$), in order to clearly define the onset of genotoxicity and avoid trivial suppressor mutations in the $mrr^{MG1655}$ locus. As such, conditional expression of $Mrr^{MG1655}$ reduced viability of LT2 wild-type with almost 99.99% (i.e. ca. 4-log reduction), although this lethality could be mitigated in the LT2Δφ strain (Figure 2). Subsequently, in order to avoid picking up suppressor mutants that would trivially be affected in prophage induction, we used an LT2 *recA* mutant as the

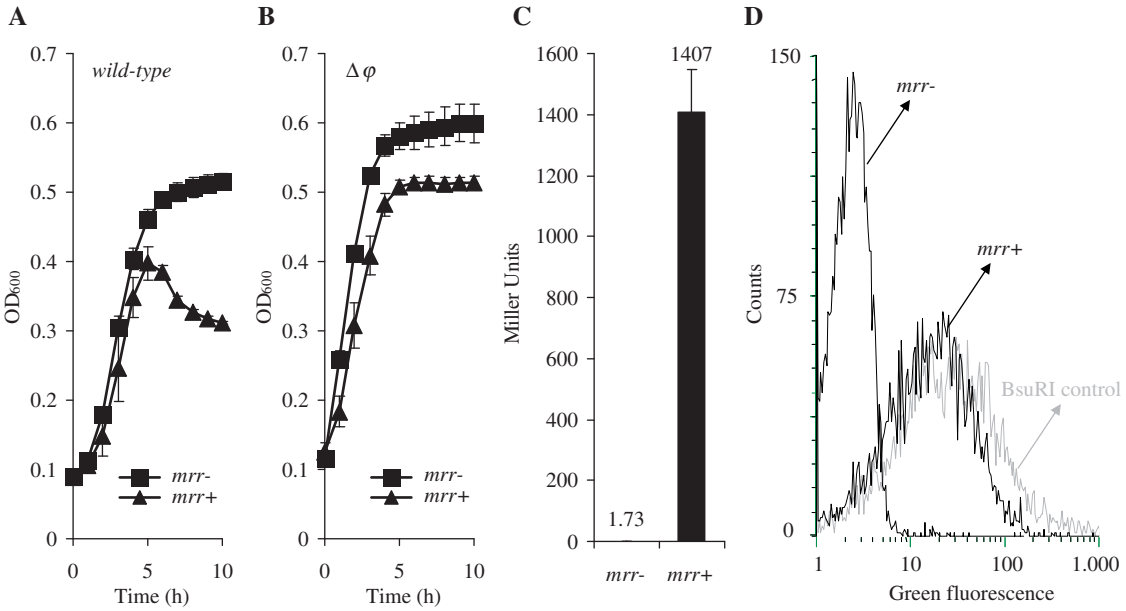


**Figure 1.** Mrr$^{MG1655}$-mediated genotoxicity in *S. typhimurium* LT2. Growth of LT2 wild-type (**A**) or LT2Δφ (**B**) equipped with either pACYC184 (*mrr*−) or pAA810 (*mrr*+). (**C**) Activation of the SOS response in late exponential phase cultures of LT2K2 equipped with either pACYC184 (*mrr*−) or pAA810 (*mrr*+). β-Galactosidase activity in the LT2K2 background serves as a measure of SOS induction, and is expressed as Miller Units. (**D**) Flow cytometry of TUNEL-labeled late exponential phase cultures of LT2 pACYC184-$P_{BAD}$-*mrr*$^{MG1655}$ in LB in the presence of 0.02% glucose (*mrr*−) or 0.02% arabinose (*mrr*+). Fluorescence is a measure of DNA damage, and permeabilized cells of LT2 treated with the blunt four-cutter *Bsu*RI to deliberately introduce DNA breaks *in vitro* were used as a positive control.

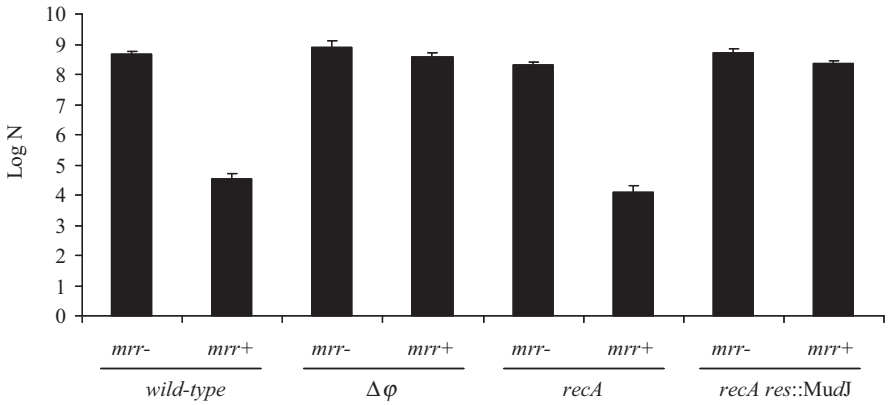


**Figure 2.** Mrr$^{MG1655}$-mediated lethality in *S. typhimurium* LT2. Survival of indicated LT2 mutants, carrying pACYC184-$P_{BAD}$-*mrr*$^{MG1655}$ and grown to stationary phase in LB, when plated on M9 minimal medium agar plates supplemented with 0.2% glucose (*mrr*−) or 0.2% arabinose (*mrr*+). Cell counts were expressed as Log *N*, in which *N* represents cfu/ml.

latter does not support prophage activation but proved to be similarly hypersensitive to Mrr$^{MG1655}$ because of its inability to repair DNA damage (Figure 2). Finally, suppression of Mrr$^{MG1655}$ toxicity in LT2 *recA* pACYC184-$P_{BAD}$-*mrr*$^{MG1655}$ was attempted by introducing (i) random transposon knock-outs or (ii) a multi-copy plasmid library of the MG1655 genome. As such, we hoped to identify cellular functions that either trigger or mitigate Mrr$^{MG1655}$ activity in LT2, respectively.

Of both strategies, only the knock-out approach resulted in a *bona fide* Mrr$^{MG1655}$ resistant mutant, and subcloning enabled us to map its transposon insertion site to the LT2 *res* gene (Figure 2). This gene is located immediately downstream of the *mod* gene, and the *mod res* operon encodes the Type III StyLTI restriction–modification system of LT2. In order to unequivocally corroborate the interference of the StyLTI system with Mrr$^{MG1655}$, clean Δ*mod*, Δ*res* and Δ*mod res* mutants were constructed *de novo* in wild-type LT2 and examined. Confirmingly, Mrr$^{MG1655}$ lethality (Figure 3A) and genotoxicity (Figure 3B) proved to be completely abolished in LT2 Δ*mod* and Δ*mod res* mutants, and to be greatly attenuated in an LT2 Δ *res* mutant. These results clearly delineate the Type III StyLTI system as the sole feature responsible for constitutive Mrr$^{MG1655}$ toxicity in LT2.

## Functional incompatibility between Mod$^{LT2}$ and Mrr$^{MG1655}$

Based on the above results, it was postulated that Mod$^{LT2}$ MTase activity would render the chromosome susceptible

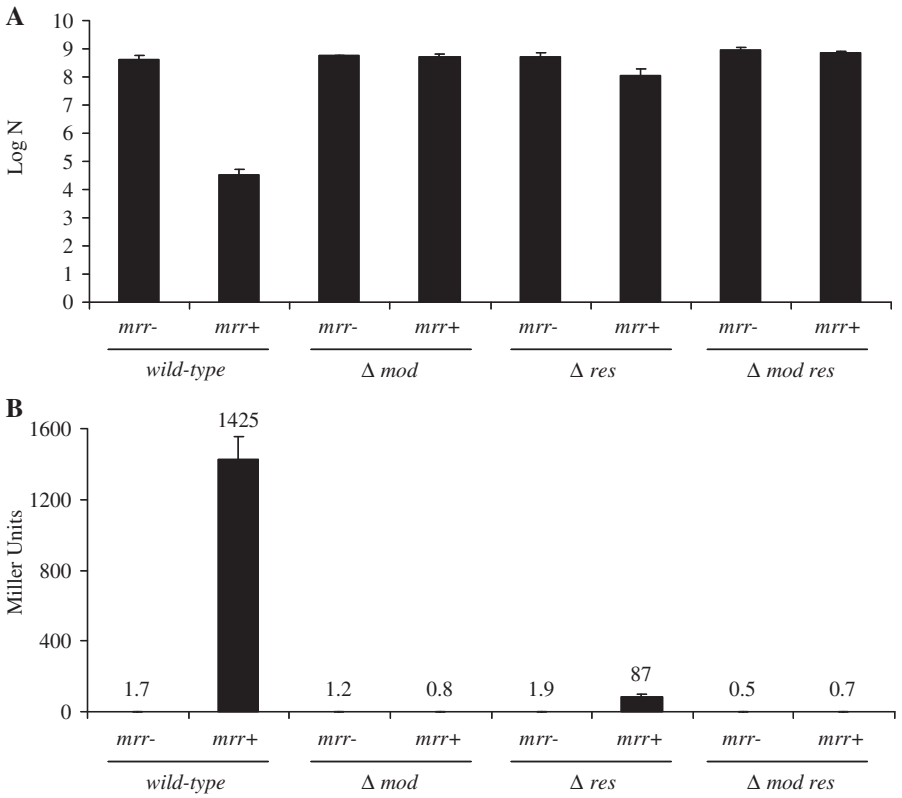


**Figure 3.** The *S. typhimurium* StyLTI RM system is responsible for $Mrr^{MG1655}$-mediated toxicity. (**A**) Survival of indicated LT2 mutants, carrying pACYC184-$P_{BAD}$-$mrr^{MG1655}$ and grown to stationary phase in LB, when plated on M9 minimal medium agar plates supplemented with 0.2% glucose (*mrr*−) or 0.2% arabinose (*mrr*+). Cell counts were expressed as Log *N*, in which *N* represents cfu/ml. (**B**) Activation of the SOS response in late exponential phase cultures of indicated mutants of LT2K2 equipped with either pACYC184 (*mrr*−) or pAA810 (*mrr*+). β-Galactosidase activity in the LT2K2 background serves as a measure of SOS induction, and is expressed as Miller Units.

to degradation by $Mrr^{MG1655}$, and the corresponding genes were conditionally co-expressed in MG1655 *mrr*. While expression of $mod^{LT2}$ or $mrr^{MG1655}$ alone was completely harmless in this background, their co-expression resulted in poor growth (data not shown) and massive SOS induction (Figure 4). These data confirm the apparent incompatibility between both proteins, and identified $Mod^{LT2}$ activity as instigator of $Mrr^{MG1655}$ restriction.

In order to examine whether this functional incompatibility could also extend into mutual exclusiveness between $mrr^{MG1655}$ and $mod^{LT2}$ loci, LT2 was equipped with pAA810 and 10 independent clones were followed for a number of generations. After only two passages, cultures grew without lysis (Figure 5A) or the accumulation of cellular debris, and gave rise to healthy colonies upon plating. Importantly, four out of these 10 independent suppressor mutants carried mutations in $mod^{LT2}$ (Figure 5B), while others were affected in the $mrr^{MG1655}$ locus encoded on pAA810. Interestingly, two of the compromised $mod^{LT2}$ alleles suffered frame-shift mutations within short homopolymeric tracts present in the open reading frame (Figure 5B). These data indicate that $Mrr^{MG1655}$ and $Mod^{LT2}$ activities are mutually exclusive, and that acquisition of one could lead to functional inactivation of the other.

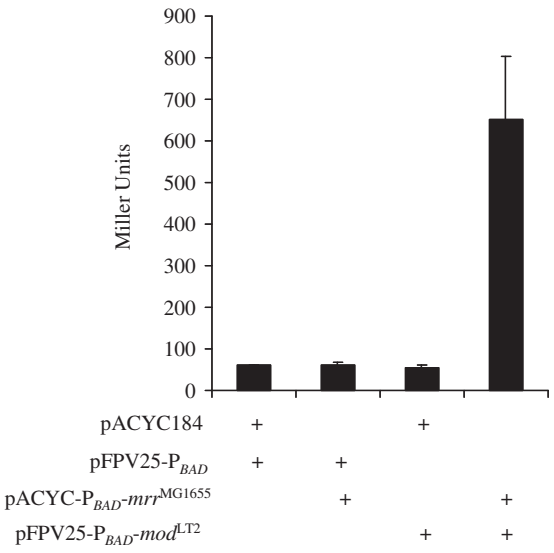


**Figure 4.** Expression of the $Mod^{LT2}$ MTase triggers $Mrr^{MG1655}$ activity in *E. coli* MG1655. Activation of the SOS response in late exponential phase cultures of MG1655 *mrr* λ*sfiA*::*lacZ* equipped with the indicated combination of plasmids and grown in the presence of 0.02% arabinose. β-galactosidase activity in the MG1655 λ*sfiA*::*lacZ* background serves as a measure of SOS induction, and is expressed as Miller Units.

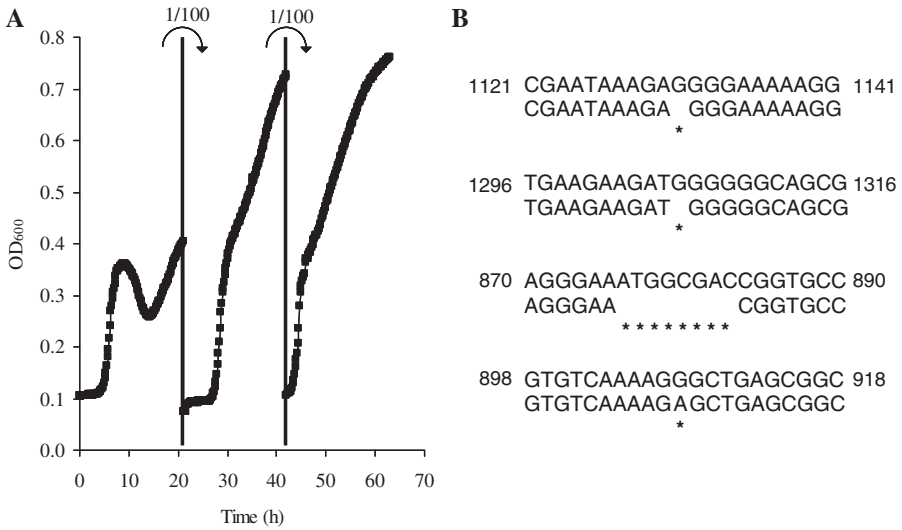


**Figure 5.** Spontaneous alleviation of $Mrr^{MG1655}$-mediated toxicity. (**A**) Typical restoration of normal growth of LT2 pAA810 during consecutive passages (1/100 dilution at indicated time points) in LB medium. Representative data are shown. (**B**) Mutations identified in the $mod^{LT2}$ alleles of four independent LT2 pAA810 mutants in which $Mrr^{MG1655}$ toxicity was suppressed. Numbers indicate the position of the base at the beginning and end of the displayed wild-type (i.e. upper) sequence with respect to the start of the open reading frame. Asterisks highlight the actual deletions or substitution in each of the four alleles (i.e. lower sequence).

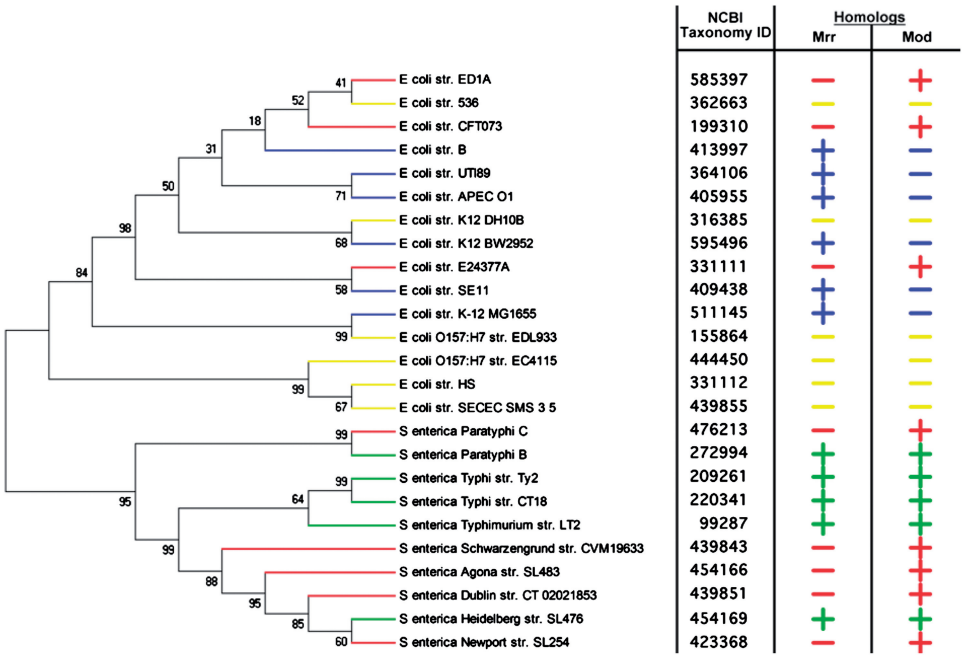


**Figure 6.** Minimal evolution phylogenetic tree of 16S rRNA derived from fully sequenced *Escherichia coli* and *Salmonella enterica* genomes. The number at each node corresponds to the percent bootstrap value (1.000 bootstraps). Additionally, NCBI Taxonomy ID as well as Mrr and Mod protein content is indicated in the lineage of each strain. The branches are colored blue for genomes with Mrr homologues, red with Mod homologues, green with both Mrr/Mod homologues, and yellow with no Mrr/Mod homologues.

In the same context, it was noteworthy to find that close homologues of $mrr^{MG1655}$ and $mod^{LT2}$ often seem to co-exist within *S. enterica* (Figure 6), which could implicate the functional loss of one of these alleles. Indeed, most *S. enterica mrr* alleles very closely resemble that of strain LT2, and we were unable to observe any $Mod^{LT2}$-dependent (or in fact any) genotoxicity upon deliberate over-expression of $Mrr^{LT2}$ in LT2 (data not shown), further supporting the hypothesis that the latter protein could have been degenerated as the result of functional counter-selection by an active Mod in *S. enterica*. Despite recent advancements in the structure–function relationship of Mrr proteins (43), we were unable to pinpoint the exact reason for the inactivity of $Mrr^{LT2}$.

### Functional incompatibility between Mod$^{ED1A}$ and Mrr$^{MG1655}$

When looking at the phylogenetic distribution in *E. coli* (Figure 6), on the other hand, it seems that homologues of *mrr*$^{MG1655}$ and *mod*$^{LT2}$ are frequently lost or acquired throughout evolution. Moreover, in agreement with their proposed incompatibility, there is a noticeable anti-correlation between both loci, and some *E. coli* strains (e.g. ED1A), in fact, lack an *mrr* allele while harboring a *mod* allele instead. Importantly, when the *mod*$^{ED1A}$ gene was cloned and co-expressed it with *mrr*$^{MG1655}$ in MG1655 *mrr*, we observed clear evidence for genotoxicity (Figure 7). This indicates that, despite bearing only 34% amino acid sequence identity with Mod$^{LT2}$, Mod$^{ED1A}$ is also capable of rendering the chromosome susceptible for cleavage by Mrr$^{MG1655}$.

Furthermore, similar to LT2, ED1A also became sick upon acquisition of pAA810 (carrying the *mrr*$^{MG1655}$ locus), and readily yielded suppressor mutants (data not shown). Interestingly, while half (i.e. five out of 10) of these mutants were affected in the *mrr*$^{MG1655}$ locus on pAA810, the other half failed to yield an amplicon of the *mod*$^{ED1A}$ gene by PCR. In this context, it is noteworthy that the *mod*$^{ED1A}$ locus is located near a prophage integrase gene, perhaps indicating that it forms part of a mobile genetic element that can easily be excised from the chromosome.

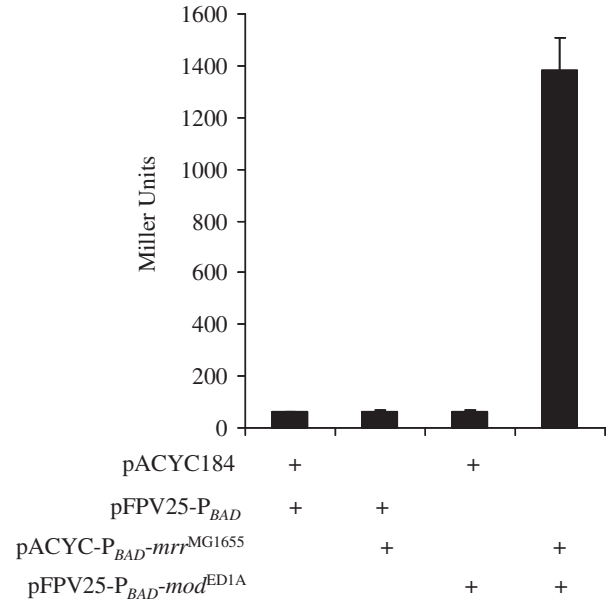


**Figure 7.** Expression of the Mod$^{ED1A}$ MTase triggers Mrr$^{MG1655}$ activity in *E. coli* MG1655. Activation of the SOS response in late exponential phase cultures of MG1655 *mrr* λ*sfiA*::*lacZ* equipped with the indicated combination of plasmids and grown in the presence of 0.02% arabinose. β-galactosidase activity in the MG1655 λ*sfiA*::*lacZ* background serves as a measure of SOS induction, and is expressed as Miller Units.

### Evolutionary antagonism between close homologues of Mrr$^{MG1655}$ and Mod$^{LT2}$

Spurred by the clear asymmetric phylogenetic distribution of *mrr*$^{MG1655}$ and *mod*$^{LT2}$ homologs within *E. coli* and *S. enterica* (Figure 6), we subsequently examined their horizontal transfer index throughout 165 sequenced prokaryotic genomes according to the method of Nakamura *et al.* (42). Although this index only permits detection of recent gene transfers from species with distinct gene sequence compositions, we nevertheless found that 11% (2/18) of Mrr$^{MG1655}$ homologues and 28% (9/32) Mod$^{LT2}$ homologues were most likely the subject of such recent horizontal transfer. Accordingly, we inferred that close homologues of *mrr*$^{MG1655}$ and *mod*$^{LT2}$ could more generally act as mobile elements, and subsequently wondered whether the mutual exclusiveness of both genes would extend beyond strains of *E. coli* and *S. enterica*. To approach this question, we searched 272 completely sequenced bacterial and archaeal genomes, and listed those in which Mrr$^{MG1655}$ and/or Mod$^{LT2}$ homologues were present (Figure 8). Using this list, the Pearson *r* correlation coefficient was calculated for 192 of these genomes (with one sampling per genome) and shown to be −0.52, which indicated a strong and statistically significant ($P < 0.0001$) anti-correlation between both loci. As such, the presence of an Mrr homolog tends to be commonly paired with the absence of a Mod homolog in the same genome, and *vice versa*.

## DISCUSSION

During our attempts to rescue the high-pressure induced SOS response in *S. typhimurium* LT2 by introducing Mrr$^{MG1655}$ of *E. coli* MG1655, we observed that the presence of this Type IV REase conferred constitutive (i.e. without the need for activation by high pressure) genotoxicity in LT2, while a similar toxicity was totally absent in MG1655. Since our knowledge about Mrr functionality is still limited and fragmentary, we decided to dissect the molecular basis of this apparent discrepancy between its impact on MG1655 and LT2. As such, we were able to show that the observed genotoxicity was elicited by the activity of a Type III MTase present in LT2 (i.e. Mod$^{LT2}$ of the StyLTI system), which most likely makes the chromosome susceptible to degradation by Mrr$^{MG1655}$. This feature, however, was not limited to Mod$^{LT2}$ as another Type III MTase encoded by *E. coli* ED1A (i.e. Mod$^{ED1A}$) could similarly elicit Mrr$^{MG1655}$ activity. Furthermore, because of this functional antagonism, loci encoding Mrr$^{MG1655}$ and Mod$^{LT2/ED1A}$ were shown to readily counter-select each other. In the same context, we revealed a strong and more general anti-correlation between close homologues of *mrr*$^{MG1655}$ and *mod*$^{LT2}$ in the genome database, suggesting that their incompatibility might well-extend beyond the genus of *Escherichia* and *Salmonella*. As such, this study is the first to provide evidence for Type III MTases as natural triggers of Mrr activity.

The observed interference between Mrr$^{MG1655}$ and Mod$^{LT2/ED1A}$ has a number of interesting repercussions. As such, Type III MTases only methylate one specific strand of its 5–6 bp asymmetric target sequence, indicating that Mrr$^{MG1655}$ is, in fact, able to restrict hemi-methylated DNA. Moreover, during the writing of this article, *in vitro*

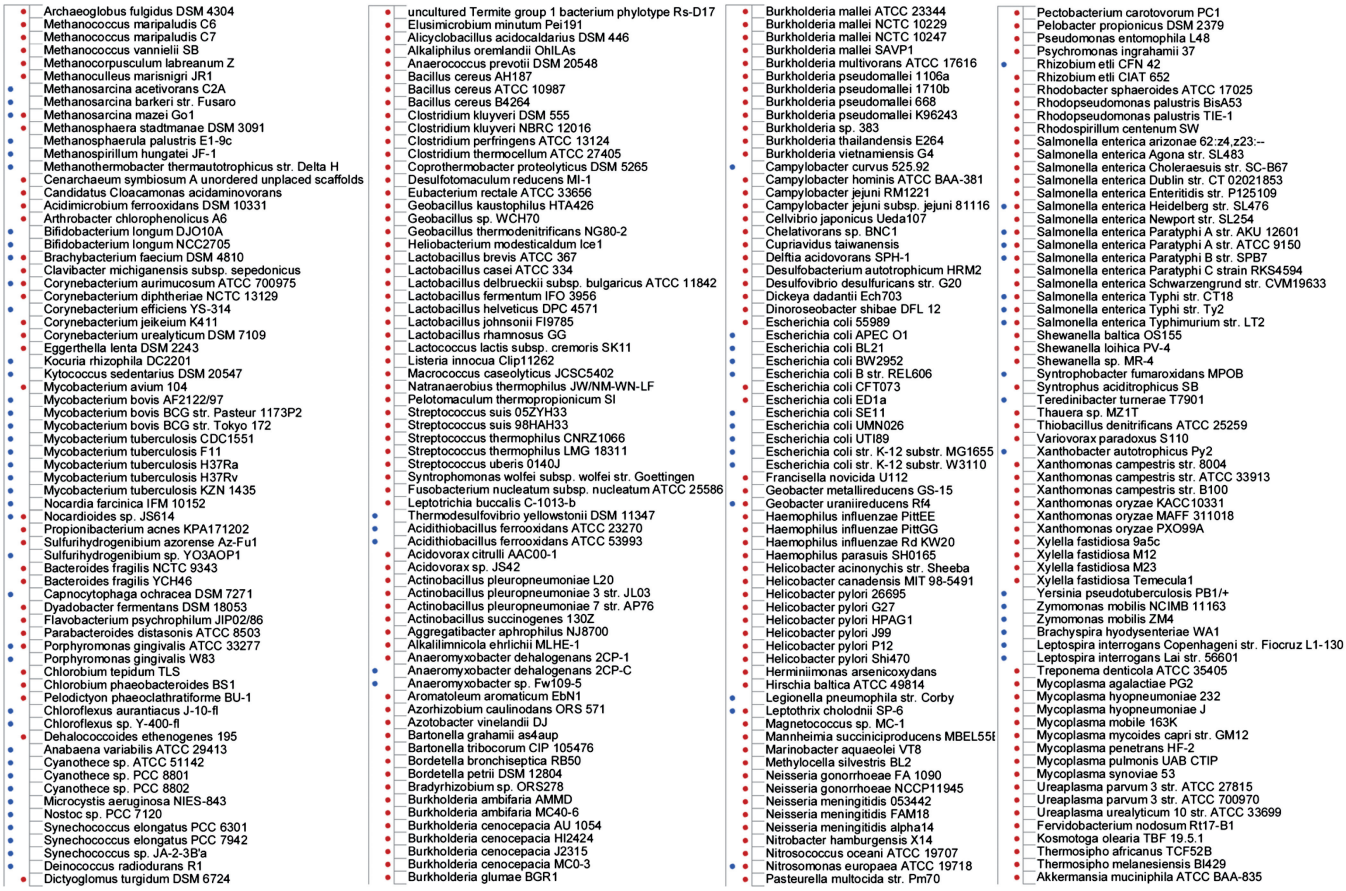


**Figure 8.** Occurrence and anti-correlation of close homologues of $Mrr^{MG1655}$ (blue dots) and $Mod^{LT2}$ (red dots) in the genome database.

evidence was published that revealed a similar activity for an Mrr-like REase (i.e. MspJI) from *Mycobacterium sp JLS* (44). In addition to hemi-methylation, all currently characterized Type III MTases seem to modify adenine (45), indicating that their activity remains undetected by other characterized Type IV REases (i.e. McrBC and McrA) which specifically target modified cytosines (46,47). Finally, our results also indicate that inactivation or deletion of the $res^{LT2}$ gene greatly attenuates $Mod^{LT2}$-dependent $Mrr^{MG1655}$ toxicity in LT2. In turn, this might suggest that the expression and/or activity of $Mod^{LT2}$ could be affected by $Res^{LT2}$, which challenges the current assumption that Type III MTases function independently of their corresponding REases (48).

Interestingly, their specificity for methylated or modified DNA has allowed the solitary Type IV REases to fulfill a number of different activities or roles in the cell. As such, it was demonstrated that McrA, McrBC and GmrSD are able to restrict invading phages that chemically disguise their chromosome with increasingly complex base modifications (8,49). Indeed, while McrA and McrBC are able to restrict DNA containing hydroxymethyscytosine (HMC), GmrSD has developed specificity for glucosylated HMC, and this differentiation is thought to reflect the evolutionary arms race between phage modification systems and the host REases. Most recently, however, Fukuda *et al.* (50) demonstrated that the deliberate introduction of Type II MTases sharing the $R^mC$ sequence specificity of McrBC could trigger cell death in *E. coli* by eliciting McrBC-dependent chromosome degradation, thereby counter-selecting the establishment or maintenance of the MTase. As a result, these authors postulated a novel function for Type IV REases in warding off epigenetic regulation by incoming MTases (50,51). The functional and evolutionary antagonism between Type IV Mrr and Type III Mod systems observed in this report seems to corroborate this hypothesis. Moreover, aside from being subject to horizontal gene transfer, some Type III MTases have actually been identified as dedicated epigenetic regulators, and have recently been shown to direct the expression of a number of genes (including virulence genes) in species of *Neisseria* and *Haemophilus* (48,52,53).

Interestingly, and fuelling their importance as epigenetic regulators, the coding sequence of many Type III MTases harbors homopolymeric tracts or tandem repeats, allowing frequent frame-shifts to occur and making their expression prone to phase variation at the translational level (48,52). In turn, this results in a phase variable ON/OFF switching of Mod-dependent regulons (i.e. phasevarions), which is believed to be a strategy that increases the adaptive potential of populations facing fluctuating environmental conditions. Although these phenomena have not been thoroughly investigated in

*S. enterica* or *E. coli*, it should be noted that the corresponding genes contain a number of 6–7 bp homopolymeric tracts. Moreover, we have shown such tracts to support phasevariation in the $mod^{LT2}$ gene upon functional counter-selection of its gene product by $Mrr^{MG1655}$. Furthermore, it is clear from the genome database that a number of *S. enterica* strains naturally contain out-of-frame *mod* alleles. The foil pathogen *S.* Gallinarum str. 287/91, for example, harbors an inactive *mod* allele because of a +1 frame shift in a poly-G tract. Whether Mod also contributes to epigenetic regulation in strains of *E. coli* and *S. enterica* still remains to be established.

Aside from the control of phasevarions, however, frequent phase variation of Mod activity would have a number of interesting consequences with respect to its antagonism with Mrr. As such, lateral acquisition of Mrr could readily select for a population phenotypically remaining in the OFF state by counteracting those cells in which the *mod* allele switches ON. On the other hand, transient OFF states might actually have evolved in order to allow safe horizontal passage or dissemination of *mod* alleles through strains exhibiting Mrr activity, thus mitigating restriction or cell death. Finally, one could even imagine that switching Mod activity in the presence of a cognate and functional Mrr would temporarily allow a small subpopulation to experience bursts of DNA damage and increased mutation rates, and to behave as transient mutators that can generate genotypic variability. The transient nature of increased mutation rates is essential, as constitutive mutators eventually succumb to the accumulation of deleterious random mutations (54). Obviously, the validity and relevance of the possible phenomena stated above require further investigation.

In summary, we present data supporting the functional interference and corresponding genomic incompatibility between Type IV Mrr and Type III Mod activity within *E. coli* and *S. enterica*, and extrapolate that these observations might extend to other species. Furthermore, this apparent antagonism fits within the recently emerging view of Type III MTases as phase variable epigenetic regulators and the hypothesis that Type IV REases might have evolved to counteract the establishment of specific genome methylation systems.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr Régis Hallez for helpful discussions, and Kelly Hughes, Lionello Bossi, John Roth and Eric Denamur for sharing strains.

## FUNDING

The Polish Ministry of Science and Higher Education (Grant POIG.02.03.00-00-003/09 to E.W.); The EU 7FP (Grant number 229676 to J.M.B.); Erasmus Mundus External Cooperation Window (EMECW13-Eurindia to A.G.); The Research Foundation - Flanders (FWO-Vlaanderen, 1.5.258.08N to A.A.); The K. U. Leuven Research Fund (CREA/09/017 to A.A.). Funding for open access charge: Partially waived by Oxford University Press.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Bickle,T.A. and Krüger,D.H. (1993) Biology of DNA restriction. *Microbiol. Rev.*, **57**, 434–450.
2. Kobayashi,I. (2001) Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res.*, **29**, 3742–3756.
3. Hoskisson,P.A. and Smith,M.C. (2007) Hypervariation and phase variation in the bacteriophage 'resistome'. *Curr. Opin. Microbiol.*, **10**, 396–400.
4. Naito,T., Kusano,K. and Kobayashi,I. (1995) Selfish behavior of restriction-modification systems. *Science*, **267**, 897–899.
5. Asakura,Y. and Kobayashi,I. (2009) From damaged genome to cell surface: transcriptome changes during bacterial cell death triggered by loss of a restriction-modification gene complex. *Nucleic Acids Res.*, **7**, 3021–3031.
6. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S.K., Dryden,D.T.F., Dybvig,K. *et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, **31**, 1805–1812.
7. Heitman,J. and Model,P. (1987) Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. *J. Bacteriol.*, **169**, 3243–3250.
8. Raleigh,E.A., Trimarchi,R. and Revel,H. (1989) Genetic and physical mapping of the *mcrA* (*rglA*) and *mcrB* (*rglB*) loci of *Escherichia coli* K-12. *Genetics*, **122**, 279–296.
9. Waite-Rees,P.A., Keating,C.J., Moran,L.S., Slatko,B.E., Hornstra,L.J. and Benner,J.S. (1991) Characterization and expression of the *Escherichia coli* Mrr restriction system. *J. Bacteriol.*, **173**, 5207–5219.
10. Revel,H.R. (1967) Restriction of nonglucosylated T-even bacteriophage: properties of permissive mutants of *Escherichia coli* B and K12. *Virology*, **31**, 688–701.
11. Raleigh,E.A. and Wilson,G. (1986) *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc. Natl Acad. Sci. USA*, **83**, 9070–9074.
12. Stewart,F.J. and Raleigh,E.A. (1998) Dependence of McrBC cleavage on distance between recognition elements. *Biol. Chem.*, **379**, 611–616.
13. Panne,D., Raleigh,E.A. and Bickle,T.A. (1999) The McrBC endonuclease translocates DNA in a reaction dependent on GTP hydrolysis. *J. Mol. Biol.*, **290**, 49–60.
14. Anton,B.P. and Raleigh,E.A. (2004) Transposon-mediated linker insertion scanning mutagenesis of the *Escherichia coli* McrA endonuclease. *J. Bacteriol.*, **186**, 5699–5707.
15. Aertsen,A., Van Houdt,R., Vanoirbeek,K. and Michiels,C.W. (2004b) An SOS response induced by high pressure in *Escherichia coli*. *J. Bacteriol.*, **186**, 6133–6141.
16. Aertsen,A., Faster,D. and Michiels,C.W. (2005) Induction of Shiga toxin-converting prophage in *Escherichia coli* by high hydrostatic pressure. *Appl. Environ. Microbiol.*, **71**, 1155–1162.
17. Aertsen,A. and Michiels,C.W. (2005) SulA-dependent hypersensitivity to high pressure and hyperfilamentation after high-pressure treatment of *Escherichia coli* lon mutants. *Res. Microbiol.*, **156**, 233–237.
18. Aertsen,A. and Michiels,C.W. (2005) Mrr instigates the SOS response after high-pressure stress in *Escherichia coli*. *Mol. Microbiol.*, **58**, 1381–1391.
19. Aertsen,A., Tesfazgi Mebrhatu,M. and Michiels,C.W. (2008) Activation of the *Salmonella* Typhimurium Mrr protein. *Biochem. Biophys. Res. Commun.*, **367**, 435–439.
20. McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W., Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F. *et al.* (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*, **413**, 852–856.

21. Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1462.
22. Touchon,M., Hoede,C., Tenaillon,O., Barbe,V., Baeriswyl,S., Bidet,P., Bingen,E., Bonacorsi,S., Bouchier,C., Bouvet,O. *et al.* (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.*, **5**, e1000344.
23. Hughes,K.T. and Roth,J.R. (1988) Transitory cis complementation: a method for providing transposition functions to defective transposons. *Genetics*, **119**, 9–12.
24. Davis,R.W., Botstein,D. and Roth,J.R. (1980) *Advanced Bacterial Genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
25. Datsenko,K.A. and Wanner,B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA.*, **97**, 6640–6645.
26. Cherepanov,P.P. and 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**, 9–14.
27. Vanderpool,C.K. and Gottesman,S. (2007) The novel transcription factor SgrR coordinates the response to glucose-phosphate stress. *J. Bacteriol.*, **189**, 2238–2248.
28. Van Melderen,L., Bernard,P. and Couturier,M. (1994) Lon dependent proteolysis of CcdA is the key control for activation of CcdB in plasmid-free segregant bacteria. *Mol. Microbiol.*, **11**, 1151–1157.
29. Silhavy,T.J., Berman,M.L. and Enquist,L.W. (1984) *Experiments with Gene Fusions*. Cold Spring Harbor Laboratory Press, Plainview, NY.
30. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York, NY.
31. Aertsen,A., Vanoirbeek,K., De Spiegeleer,P., Sermon,J., Hauben,K., Farewell,A., Nyström,T. and Michiels,C. (2004a) Heat shock protein-mediated resistance to high hydrostatic pressure in *Escherichia coli*. *Appl. Environ. Microbiol.*, **70**, 2660–2666.
32. Valdivia,R.H. and Falkow,S. (1996) Bacterial genetics by flow cytometry: rapid isolation of *Salmonella* Typhimurium acid-inducible promoters by differential fluorescence induction. *Mol. Microbiol.*, **22**, 367–378.
33. Rose,R.E. (1988) The nucleotide sequence of pACYC184. *Nucleic Acids Res.*, **16**, 355.
34. Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
35. Miller,J.H. (1992) *A Short Course in Bacterial Genetics*. Cold Spring Harbor Laboratory Press, New York, NY.
36. Rhower,F. and Azam,F. (2000) Detection of DNA damage in prokaryotes by terminal deoxyribonucleotide transferase-mediated dUTP nick end labeling. *Appl. Environ. Microbiol.*, **66**, 1001–1006.
37. Gunderson,C.W. and Segall,A.M. (2006) DNA repair, a novel antibacterial target: Holliday junction-trapping peptides induce DNA damage and chromosome segregation defects. *Mol. Microbiol.*, **59**, 1129–1148.
38. DeSantis,T.Z., Hugenholtz,P., Larsen,N., Rojas,M., Brodie,E.L., Keller,K., Huber,T., Dalevi,D., Hu,P. and Andersen,G.L. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.*, **72**, 5069–5072.
39. Tamura,K., Dudley,J., Nei,M. and Kumar,S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Bioland. Evol.*, **24**, 1596–1599.
40. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
41. Frickey,T. and Lupas,A. (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics*, **20**, 3702–3704.
42. Nakamura,Y., Itoh,T., Matsuda,H. and Gojobori,T. (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.*, **36**, 760–766.
43. Orlowski,J., Mebrhatu,M.T., Michiels,C.W., Bujnicki,J.M. and Aertsen,A. (2008) Mutational analysis and a structural model of methyl-directed restriction enzyme Mrr. *Biochem. Biophys. Res. Commun.*, **377**, 862–866.
44. Zheng,Y., Cohen-Karni,D., Xu,D., Chin,H.G., Wilson,G., Pradhan,S. and Roberts,R.J. (2010) A unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Res.*, **38**, 5527–5534.
45. Madhusoodanan,U.K. and Rao,D.N. (2010) Diversity of DNA methyltransferases that recognize asymmetric target sequences. *Crit. Rev. Biochem. Mol. Biol.*, **45**, 125–145.
46. Raleigh,E.A., Murray,N.E., Revel,H., Blumenthal,R.M., Westaway,D., Reith,A.D., Rigby,P.W., Elhai,J. and Hanahan,D. (1988) McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucleic Acids Res.*, **25**, 1563–1575.
47. Raleigh,E.A. (1992) Organization and function of the *mcrBC* genes of *Escherichia coli* K-12. *Mol. Microbiol.*, **6**, 1079–1086.
48. Srikhanta,Y.N., Fox,K.L. and Jennings,M.P. (2010) The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nat. Rev. Microbiol.*, **8**, 196–206.
49. Bair,C.L. and Black,L.W. (2007) A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J. Mol. Biol.*, **23**, 768–778.
50. Fukuda,E., Kaminska,K.H., Bujnicki,J.M. and Kobayashi,I. (2008) Cell death upon epigenetic genome methylation: a novel function of methyl-specific deoxyribonucleases. *Genome Biol.*, **9**, R163.
51. Ishikawa,K., Fukuda,E. and Kobayashi,I. (2010) Conflicts targeting epigenetic systems and their resolution by cell death: novel concepts for methyl-specific and other restriction systems. *DNA Res.*, **17**, 325–42.
52. Srikhanta,Y.N., Maguire,T.L., Stacey,K.J., Grimmond,S.M. and Jennings,M.P. (2005) The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc. Natl Acad. Sci. USA*, **12**, 5547–5551.
53. Bayliss,C.D., Callaghan,M.J. and Moxon,E.R. (2006) High allelic diversity in the methyltransferase gene of a phase variable type III restriction-modification system has implications for the fitness of *Haemophilus influenzae*. *Nucleic Acids Res.*, **34**, 4046–4059.
54. Giraud,A., Radman,M., Matic,I. and Taddei,F. (2001) The rise and fall of mutator bacteria. *Curr. Opin. Microbiol.*, **4**, 582–585.