MINI-REVIEW

The mechanisms underlying antigenic variation and maintenance of genomic integrity in *Mycoplasma pneumoniae* and *Mycoplasma genitalium*

Mohamad S. Hakim[1,2] · Luthvia Annisa[1] · Rizka O. A. Jariah[3] · Cornelis Vink[4]

Abstract

Mycoplasma pneumoniae and *Mycoplasma genitalium* are important causative agents of infections in humans. Like all other mycoplasmas, these species possess genomes that are significantly smaller than that of other prokaryotes. Moreover, both organisms possess an exceptionally compact set of DNA recombination and repair-associated genes. These genes, however, are sufficient to generate antigenic variation by means of homologous recombination between specific repetitive genomic elements. At the same time, these mycoplasmas have likely evolved strategies to maintain the stability and integrity of their 'minimal' genomes. Previous studies have indicated that there are considerable differences between mycoplasmas and other bacteria in the composition of their DNA recombination and repair machinery. However, the complete repertoire of activities executed by the putative recombination and repair enzymes encoded by *Mycoplasma* species is not yet fully understood. In this paper, we review the current knowledge on the proteins that likely form part of the DNA repair and recombination pathways of two of the most clinically relevant *Mycoplasma* species, *M. pneumoniae* and *M. genitalium.* The characterization of these proteins will help to define the minimal enzymatic requirements for creating bacterial genetic diversity (antigenic variation) on the one hand, while maintaining genomic integrity on the other.

Keywords Antigenic variation · DNA repair · Homologous DNA recombination · *M. genitalium* · *M. pneumoniae*

Introduction

Mycoplasma pneumoniae and *Mycoplasma genitalium* are pathogenic bacteria that cause significant health problems in the human population. These pathogens are genetically very similar (Himmelreich et al. 1997), and belong to the *Mollicutes* class of bacteria. While the genome of *M. pneumoniae* is significantly longer than that of *M. genitalium* (816 kb versus 580 kb) (Fraser et al. 1995; Himmelreich et al. 1996), the ~479 orthologous proteins encoded by these species on average display ~67% identity (Himmelreich et al. 1997).

M. pneumoniae causes both upper respiratory tract infections (RTIs), such as pharyngitis and tracheobronchitis, and lower RTIs, such as pneumonia. *M. pneumoniae* infection and transmission occurs during both endemic and epidemic settings in developed as well as developing countries. It has been reported that *M. pneumoniae* is responsible for up to 40% of the cases of community-acquired bacterial pneumonia (CABP) found during epidemic situations, many of which concern children (Waites et al. 2017). In addition, this bacterium was found to induce extrapulmonary infections

✉ Mohamad S. Hakim
m.s.hakim@ugm.ac.id

✉ Cornelis Vink
vink@euc.eur.nl

1   Department of Microbiology, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia

2   Postgraduate School of Molecular Medicine, Erasmus MC-University Medical Center Rotterdam, Rotterdam, The Netherlands

3   Department of Health Science, Faculty of Vocational Studies, Universitas Airlangga, Surabaya, Indonesia

4   Department of Life Sciences, Erasmus University College, Erasmus University, 3011 HP Rotterdam, The Netherlands

as well as post-infectious, immune-mediated diseases, such as Guillain–Barre syndrome (Meyer Sauteur et al. 2014a, 2016). Interestingly, *M. pneumoniae* can also be found in the respiratory tract of asymptomatic children (Spuesens et al. 2016, 2013), challenging both the diagnosis and treatment of *M. pneumoniae* infections in this patient group (Meyer Sauteur et al. 2014b, 2018; Spuesens et al. 2014).

M. genitalium is an etiological agent of various diseases of the human reproductive tract, such as non-gonococcal urethritis (NGU) in men, and cervicitis, endometritis, pelvic inflammatory disease (PID) and tubal-factor infertility in women (Deborde et al. 2019; McGowin and Totten 2017). The prevalence of *M. genitalium* infections is about 1.3–3.9% in the general population and significantly higher in specific groups such as female commercial sex workers (15.9%) (Baumann et al. 2018). *M. genitalium* infections are also relatively common in human immunodeficiency virus (HIV)-infected patients (Deborde et al. 2019). Therefore, it was considered to include *M. genitalium* screening and treatment interventions as part of HIV prevention strategies (Napierala Mavedzenge et al. 2015).

As human pathogens, *Mycoplasma* spp. continuously evolve due to external pressures exerted by the host immune system as well as the use of antibiotic drugs. Among the most common strategies employed by pathogenic microorganisms to evade immune surveillance and control by the host is antigenic variation. Antigenic variation is successfully employed by various bacterial pathogens, including *Neisseria* spp., *Mycoplasma* spp., and *Treponema pallidum* (Vink et al. 2012). It is also effectively used by viral pathogens, such as influenza virus (Shao et al. 2017). Antigenic variation leads to continuous alterations or modifications of the surface molecules that are mainly targeted by the host immune systems. Consequently, the humoral (antibody) response generated against the previous ("old") surface molecules cannot effectively recognize and neutralize the modified ("new") molecules, allowing the pathogen to persist in the infected host for a prolonged period of time (Dehon and McGowin 2017; Qin et al. 2019). Antigenic variation is also hypothesized to be involved in the repeated epidemics caused by *M. pneumoniae* (Dumke et al. 2008). In this species, as well as in *M. genitalium,* this antigenic variation is predicted to be generated through homologous recombination between specific, repetitive DNA elements that are dispersed throughout their genomes (Rocha and Blanchard 2002). These repetitive elements encode the antigenic surface molecules P1 and MgPa of *M. pneumoniae* and *M. genitalium*, respectively (Fig. 1).
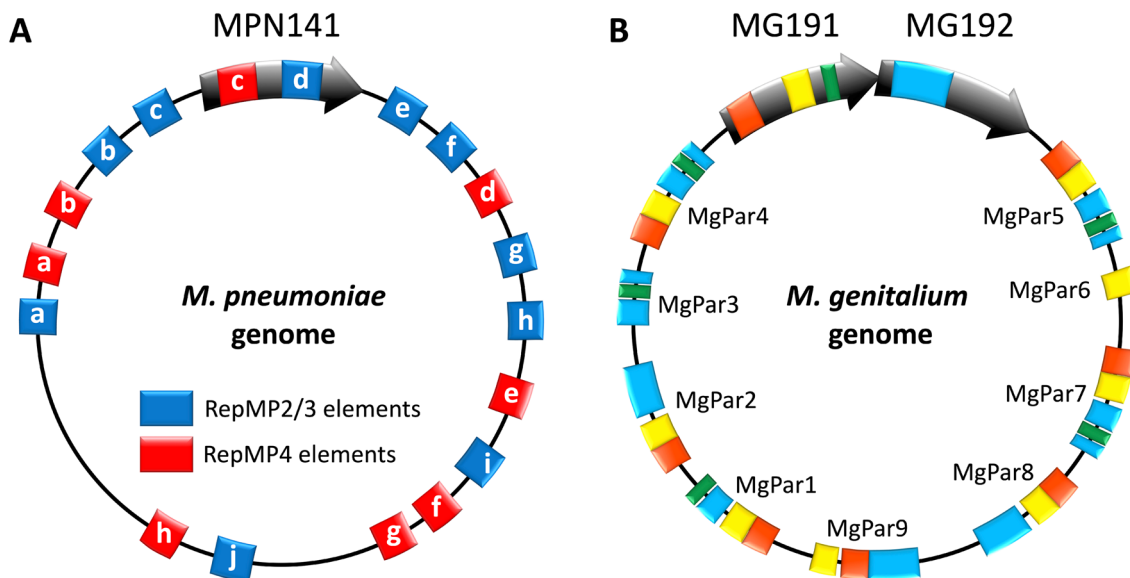
Fig. 1 Repetitive elements in the genome of *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. **a** The P1 gene (ORF MPN141) of *M. pneumoniae* contains two variable DNA elements, RepMP2/3 and RepMP4. Multiple variants of RepMP2/3 and RepMP4 elements are dispersed within the genome of *M. pneumoniae*. There are ten variants of RepMP2/3 and each variant is labeled 'a' to 'j' in blue. For RepMP4, there are eight variants and each variant is labeled 'a' to 'h' in red. The drawing is a modification of figures that were previously published by Spuesens et al. (Spuesens et al. 2009, 2011). **b** The MgPa operon of *M. genitalium* contains two variable genes, *mgpB* (ORF MG191) and *mgpC* (ORF MG192). There are three MgPar homologous regions within the MG191 gene, indicated as repeat regions B (orange), EF (yellow), and G (green), while there is only one, large MgPar homologous region within the MG192 gene, indicated as repeat region JKLM (light blue). Invariable, conserved regions within each gene are indicated in black. Nine homologous MgPar sequences, containing diverse copies of *mgpB*- and *mgpC*-associated homologous regions, are present in the genome of *M. genitalium*. Homologous sequences are indicated in the same color. The drawing is a modification of a figure that was previously published by McGowin and Totten (2017)

While antigenic variation through homologous recombination may be crucial for the propagation of *Mycoplasma* species in human populations, a major issue for these organisms is the maintenance of the integrity of their genomes, in particular because these genomes are much more compact than the genomes of most other bacterial taxa. Interestingly, the DNA repair systems that are involved in maintaining genome integrity in *M. pneumoniae* and *M. genitalium* may also be involved in the aforementioned recombination between repetitive DNA elements. Consequently, the DNA repair machinery of these organisms may be directly involved in antigenic variation.

The aim of this paper is to review our current knowledge on the genes and proteins that likely form part of the DNA repair and recombination pathways of *M. pneumoniae* and *M. genitalium.* Several previous studies have indicated that considerable differences exist between mycoplasmas and other bacteria in the composition of their sets of DNA recombination and repair-related genes. Thus, characterization of these genes and the encoded proteins will help to define the minimal enzymatic requirements for creating bacterial genetic diversity (antigenic variation) on the one hand, while maintaining genomic integrity on the other.

The role of homologous DNA recombination between repetitive elements in antigenic variation of *M. pneumoniae* and *M. genitalium*

Mechanism of antigenic variation and the role of repetitive element recombination

Mycoplasmas are the smallest known self-replicating organisms, both regarding cellular dimensions and genome size (Wilson and Collier 1976). As mentioned above, the genomes of *M. pneumoniae* (strain M129) and *M. genitalium* (strain G-37 [T]) are only 816 kb and 580 kb in length, respectively (Fraser et al. 1995; Himmelreich et al. 1996). By comparison, the genome of 'model' bacterium *Escherichia coli* strain K12 is 4,639 kb in length (Blattner et al. 1997). It is remarkable that, in spite of their limited sizes, the genomes of *M. pneumoniae* and *M. genitalium* consist of a significant portion of repeated DNA elements, which constitute approximately 8% and 4% of the genome, respectively (Himmelreich et al. 1996; Peterson et al. 1993). These multiple repeated elements, which are dispersed throughout the genome, display a high level of sequence similarity, but are not identical. In *M. pneumoniae,* these are referred to as Rep MP (RepMP2/3 and RepMP4) elements; while in *M. genitalium* they are termed MgPa repeats (MgPars) (Fraser et al. 1995; Peterson et al. 1995; Ruland et al. 1990; Su et al. 1988).

Interestingly, some of the copies of the repeated DNA elements were found to form part of genes encoding antigenic surface proteins. Among these proteins are the cytadhesins P1 of *M. pneumoniae* (Himmelreich et al. 1996) and MgPa of *M. genitalium* (Aparicio et al. 2018; Fraser et al. 1995). It has been demonstrated that the repeated DNA elements (including those in the P1 and MgPa operon) can undergo homologous DNA recombination, which may result in sequence variation (*i.e.* antigenic variation) of the P1 and MgPa proteins (Ma et al. 2007; Peterson et al. 1995; Ruland et al. 1990; Spuesens et al. 2009).

## Basic mechanisms of homologous DNA recombination

For general DNA editing and repair, bacteria can employ homologous DNA recombination, in which a stretch of DNA is eventually exchanged with an identical (or almost identical) sequence that originates either from another site of the genome or from extrachromosomal DNA. Much of our knowledge of homologous recombination is derived from studies performed in *E. coli* (Kowalczykowski et al. 1994). However, the basic mechanisms as well as genes or proteins involved in this process are conserved across eubacteria (Kuzminov 1999).

The initial step of homologous recombination requires single-stranded breaks or nicks at corresponding regions of two homologous DNA sequences by a specific endonuclease (Fig. 2). Subsequently, the two homologous single strands from each DNA molecule exchange positions in a reciprocal manner, resulting in basepair formation between the transferred (donor) and recipient DNA (heteroduplex formation). This reciprocal strand exchange produces a so-called Holliday junction (or Chi structure), a point at which two single DNA strands from two homologous double-stranded (ds) DNA molecules exchange. Migration of the Holliday junction (branch migration) allows further extension of the stretch of heteroduplex DNA. As the final step, resolution (cutting) of the Holliday junction by a resolvase enzyme, and subsequent ligation of the remaining single-stranded nicks by DNA ligase, will produce two intact, recombined DNA molecules (Fig. 2a) (Bianco et al. 1998; Kowalczykowski et al. 1994).

In *E. coli,* homologous recombination requires a coordinated action of more than 20 different proteins. Of these, RecA is involved in the initial step of pairing of the two homologous DNA sequences and single-strand invasion onto the other DNA template. While RuvA and RuvB are proteins that are necessary for branch migration, the resolution of the Holliday junction is executed by RuvC (Eggleston and West 1996; Kowalczykowski et al. 1994).

Although the aforementioned process of homologous recombination involves the reciprocal, bidirectional

A Double-reciprocal homologous recombination
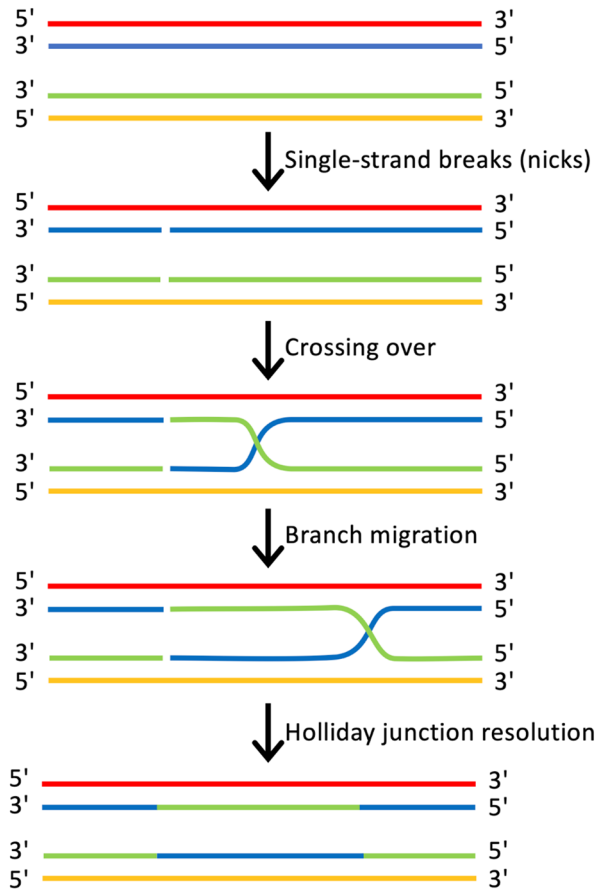
B Gene conversion
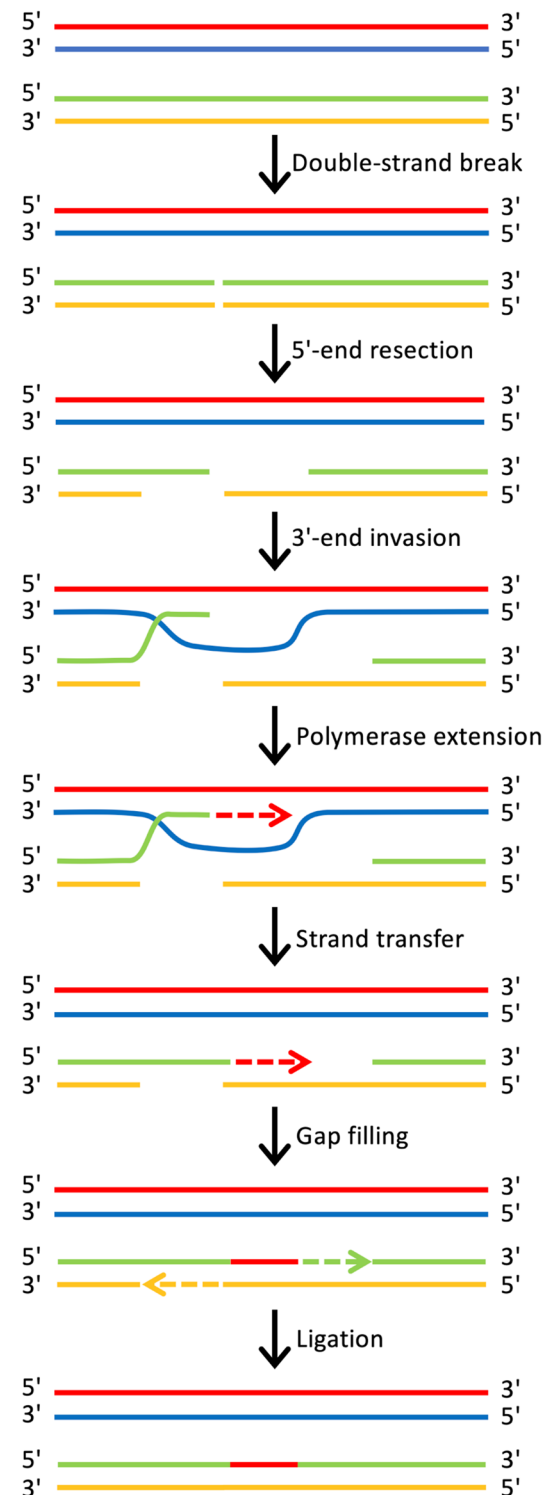
exchange of DNA strands, homologous recombination can also be non-reciprocal and unidirectional, through a process

termed gene conversion. In gene conversion, the original recipient sequences are replaced by the donor sequences

◄**Fig. 2** The basic mechanism of homologous DNA recombination. **a** Reciprocal homologous recombination. This process is initiated by generation of single-strand breaks or nicks at corresponding regions of two homologous sequences by a specific endonuclease. Subsequently, the two homologous single strands from each DNA molecule exchange positions in a reciprocal manner (crossing over), resulting in heteroduplex formation. This reciprocal strand exchange produces a Holliday junction. Branch migration of the Holliday junction allows further extension of the stretch of heteroduplex DNA. Resolution of the Holliday junction is executed by a resolvase enzyme. As the final step, ligation of the remaining single-strand nicks is performed by DNA ligase, resulting in two intact, recombined DNA molecules. **b** Non-reciprocal homologous recombination (gene conversion). This process is initiated by the generation of a double-strand break. The resulting 5′ ends are then resected, generating 3′ ssDNA tails. In the example shown, one of the 3′ tails transfers to -and basepairs with- another, homologous DNA molecule, forming a displacement (D)-loop. The newly formed heteroduplex DNA is extended by DNA polymerase. The invading strand, including the newly synthesized stretch of DNA, is then displaced form the template strand and reanneals to the DNA strand it was originally attached to (strand transfer). Remaining single-stranded gaps are subsequently filled by the combined action of DNA polymerase I and DNA ligase

and are eventually lost from the bacterial genome (Fig. 2b) (Bianco et al. 1998; Chen et al. 2007).

Evidence of homologous recombination-induced antigenic variation in *M. pneumoniae* and *M. genitalium*

For *M. pneumoniae*, it has been hypothesized that homologous recombination between RepMP elements may occur through gene conversion-like processes. This hypothesis was supported by analysis of the RepMP elements from a collection of 23 M. pneumoniae isolates. Sequence analysis of all 10 variants of RepMP2/3 and all 8 variants of RepMP4 from these isolates indicated that one or more 'donor' RepMP variants appeared to have been copied to other ('recipient') RepMP variants, including those located within the P1 operon (Spuesens et al. 2009). The same phenomenon was also observed among RepMP5 variants, resulting in amino acid changes in another surface-exposed cytadhesin protein, named P40 (Spuesens et al. 2011). In addition, analysis of an additional set of 23 *M. pneumoniae* clinical isolates also showed a significant rate of reorganization through recombination events between these repetitive elements (Lluch-Senar et al. 2015).

Sequence analysis of the *M. genitalium mgpB* and *mgpC* genes (the second and third gene within the MgPa operon) provided strong support for the hypothesis that recombination occurs between MgPar sequences, and results in antigenically distinct MgPa variants both in vitro and in clinical isolates (Fookes et al. 2017; Iverson-Cabral et al. 2006, 2007; Ma et al. 2007, 2014). In an experimental chimpanzee model of *M. genitalium* infection, sequence variation within the *mgpC* gene initially occurred within five weeks post infection and accumulated progressively (Ma et al. 2015). Interestingly, in the majority of the cases, MgPa recombination events in *M. genitalium* were found to be caused by reciprocal DNA recombination events, in contrast to the homologous recombination events found in *M. pneumoniae*, which all appeared to have resulted from gene conversion events. Another difference between the recombination processes in these mycoplasmas is that recombination events occur at a significantly higher frequency in *M. genitalium* than in *M. pneumoniae* (Vink et al. 2012).

Because both the *M. pneumoniae* P1 protein and the *M. genitalium* MgPa protein are highly immunogenic (Razin and Jacobs 1992), the recombination-induced antigenic variation of these proteins may play a crucial role in the pathogenicity of both pathogens and their evasion from the host's immune system (Vink et al. 2012). Consequently, antigenic variation may be a critical factor in allowing *M. pneumoniae* and *M. genitalium* to persist in infected humans for prolonged periods of time (Atkinson et al. 2008; Hardy et al. 2002; Iverson-Cabral et al. 2006; Vink et al. 2012). In support of this notion, persistent infections with these bacteria were observed in animal models of infection (Hardy et al. 2002; McGowin et al. 2010; Wood et al. 2017). In addition, the antigenic variation of P1 and MgPa may facilitate adaptation of the bacteria to different host microenvironments, because both P1 and MgPa proteins are surface-exposed proteins that function in the attachment of the bacteria to host cells (Ma et al. 2014).

The *Mycoplasma* genes and proteins involved in DNA recombination

Due to the compact nature of all metabolic pathways in mycoplasmas, it is likely that the enzymatic machinery that governs recombination between repeated DNA elements in *M. pneumoniae* and *M. genitalium* largely overlaps with the machinery involved in general DNA recombination and repair in these bacteria (Estevao et al. 2011). As shown in Table 1, the predicted set of genes and proteins involved in DNA recombination and repair in these mycoplasmas is limited in number. Notably, *M. pneumoniae* and *M. genitalium* lack a significant number of enzymes known to be involved in DNA repair in other bacterial classes. These enzymes include LexA, PhrI, PhrII, RecBCD, AddAB, RecFOR, RecQ and RecJ, as well as proteins involved in mismatch repair, such as MutS, MutL, and MutH (Carvalho et al. 2005).

The functions of several of the proteins listed in Table 1 have been investigated. These proteins include SSB (Sluijter et al. 2008), RecA (Sluijter et al. 2009), RuvA (Ingleston et al. 2002; Sluijter et al. 2012), RuvB (Estevao et al. 2011), and RecU (Sluijter et al. 2011, 2010). Remarkably, as detailed below, some of these proteins showed significantly

Table 1 *M. pneumoniae* and *M. genitalium* genes encoding proteins putatively involved in DNA recombination and repair

M. pneumoniae gene	*M. genitalium* ortholog	Name and predicted function of the encoded proteins	Name and established function of the encoded proteins	*E. coli* orthologue	References
MPN229	MG091	SSB_{Mge}: a putative single-stranded DNA binding protein	SSB_{Mpn}: single-stranded DNA binding protein	SSB_{Eco}	(Sluijter et al. 2008)
MPN490	MG339	–	$RecA_{Mpn}$ and $RecA_{Mge}$: homologous DNA strand transferase	$RecA_{Eco}$	(Sluijter et al. 2009)
MPN535	MG358	–	$RuvA_{Mpn}$ and $RuvA_{Mge}$: Holliday junction (HJ)-binding proteins	$RuvA_{Eco}$	(Ingleston et al. 2002; Sluijter et al. 2012)
MPN536	MG359	–	$RuvB_{FH}$, $RuvB_{M129}$, and $RuvB_{Mge}$: DNA helicases	$RuvB_{Eco}$	(Estevao et al. 2011)
MPN528a	MG352	–	$RecU_{Mge}$: a HJ resolvase; $RecU_{Mpn}$ is a non-functional protein	–	(Sluijter et al. 2011, 2010)
MPN340	MG244	–	$PcrA1_{M129}$ and $PcrA_{FH}$: DNA helicases with UvrD-like activity	$UvrD_{Eco}$	(Estevao et al. 2013)
MPN341	MG244	–	$PcrA2_{Mpn}$ and $PcrA_{Mge}$: DNA helicases with UvrD-like activity	$UvrD_{Eco}$	(Estevao et al. 2013)
MPN328	MG235	–	Nfo_{Mpn} and Nfo_{Mge}: endonuclease IV proteins	Nfo_{Eco}	(Estevao et al. 2014)
MPN619	MG421	UvrA: a putative ATP-dependent DNA-binding protein	–	$UvrA_{Eco}$	This study
MPN211	MG073	UvrB: a putative DNA-binding protein	–	$UvrB_{Eco}$	This study
MPN125	MG206	UvrC: a putative excinuclease	–	$UvrC_{Eco}$	This study

The genes and their names were taken from Fraser et al. and Himmelreich et al. (Fraser et al. 1995; Himmelreich et al. 1996)

different activities than their homologs from other bacterial classes.

SSB_{Mpn}

The MPN229 ORF of *M. pneumoniae* encodes a single-stranded DNA-binding protein (SSB_{Mpn}) consisting of 166 amino acids. Functional characterization of this ~18-kDa protein revealed that it possesses similar activities as its counterpart from *E. coli.* SSB_{Mpn} was reported to form tetramers that efficiently and selectively bind to single-stranded DNA (ssDNA) substrates. This activity was independent of the presence of divalent cations, such as Mg^{2+} (Sluijter et al. 2008). Importantly, SSB_{Mpn} efficiently supported *E. coli* Recombinase A ($RecA_{Eco}$)-promoted homologous DNA recombination.

The *M. genitalium* counterpart of SSB_{Mpn}, SSB_{Mge}, is encoded by the MG091 gene. Although SSB_{Mpn} and SSB_{Mge} are 61% identical on the amino acid sequence level (Sluijter

et al. 2008), the latter protein has not yet been subjected to functional analyses.

$RecA_{Mpn}$ and $RecA_{Mge}$

The MPN490 and MG339 ORFs of *M. pneumoniae* and *M. genitalium*, respectively, were reported to encode homologs of Recombinase A (RecA) proteins. These ~37-kDa proteins, which were designated $RecA_{Mpn}$ and $RecA_{Mge}$, respectively, were found to have a high similarity on the amino acid sequence level (79% identity) (Sluijter et al. 2009). Functional in vitro studies demonstrated that both $RecA_{Mpn}$ and $RecA_{Mge}$ possess similar activities as their counterpart from *E. coli*, by mediating recombination events between homologous DNA substrates in an ATP-, pH- and divalent cation (Mg^{2+})-dependent manner. Also, the recombinase activity of $RecA_{Mpn}$ and $RecA_{Mge}$ was found to be dependent on the presence of SSB_{Mpn} (Sluijter et al. 2009). In vivo, it was shown that $RecA_{Mge}$ efficiently mediates MgpB and MgpC phase and antigenic variation in *M. genitalium*. However,

this protein was not found to play a role in DNA repair (Burgos et al. 2012).

RuvA$_{Mpn}$ and RuvA$_{Mge}$

Both *M. pneumoniae* and *M. genitalium* were also found to encode homologs of RuvA and RuvB. The RuvA homologs of these species are encoded by ORFs MPN535 and MG358, and were named RuvA$_{Mpn}$ and RuvA$_{Mge}$, respectively. Although the sequence similarity between these ~23.7-kDa proteins is high (68.8% identity), both proteins demonstrated notably different in vitro activities. RuvA$_{Mpn}$ exhibits a relatively high binding affinity for both Holliday junction (HJ) and ssDNA substrates, in contrast to both RuvA$_{Mge}$ and RuvA$_{Eco}$, which preferentially bind to HJs. Additionally, while RuvA$_{Mpn}$ was not found to stimulate RuvB activity, RuvA$_{Mge}$ was shown to stimulate the helicase and ATPase activities of RuvB$_{Mge}$ (Ingleston et al. 2002; Sluijter et al. 2012).

RuvB$_{FH}$, RuvB$_{M129}$, and RuvB$_{Mge}$

The ORFs MPN536 and MG359 have been shown to encode RuvB homologs of *M. pneumoniae* and *M. genitalium*, respectively, with a molecular mass of ~35.0 kDa. Interestingly, the two main subtypes of *M. pneumoniae* were found to encode variants of RuvB that only differ at a single amino acid position; subtype 1 strains were shown to encode a RuvB homolog (termed RuvB$_{M129}$) carrying a tryptophan residue at position 140, whereas subtype 2 strains have a leucine residue at this position in the protein (RuvB$_{FH}$) (Estevao et al. 2011). Interestingly, this sequence difference was reported to have a major impact on the in vitro activities of these proteins. While RuvB$_{FH}$ was found to have significant divalent cation- and ATP-dependent DNA helicase activity, RuvB$_{M129}$ only displayed marginal levels of DNA-unwinding activity. The *M. genitalium*-encoded RuvB homolog (RuvB$_{Mge}$), which shares 84.4% of its amino acid sequence with its *M. pneumoniae* counterparts, also showed a relatively low DNA-unwinding activity (Estevao et al. 2011).

RecU$_{Mpn}$ and RecU$_{Mge}$

The MG352-encoded RecU protein of *M. genitalium* (RecU$_{Mge}$) was found to be a potent HJ-resolving enzyme. However, this protein has unique properties compared to its homologs from other bacterial classes. Specifically, the HJ-resolving activity of RecU$_{Mge}$ was only detected in the presence of Mn^{2+} and not Mg^{2+}. Remarkably, the RecU protein encoded by the MPN528a ORF of *M. pneumoniae* subtype 2 strains (RecU$_{Mpn}$) did not possess any in vitro HJ-binding or cleavage activities. It was demonstrated that this inactivity was caused by the presence of a glutamic

acid residue at position 67 of the protein, which is not conserved in RecU$_{Mge}$. Even more striking was the finding that the MPN528a gene of *M. pneumoniae* subtype 1 strains is unable to encode a full-length RecU homolog, because it contains a premature TAA translation termination codon at position 181–183. Consequently, subtype 1 strains of *M. pneumoniae* cannot produce a functional RecU protein (Sluijter et al. 2010). Taken together, *M. pneumoniae* strains may not encode functional homologs of RecU, which may explain the relatively low level of recombination that is observed within the *M. pneumoniae* genome (Sluijter et al. 2010).

In conclusion, functional studies of the various proteins that are putatively involved in DNA recombination in *Mycoplasma* spp. revealed significantly different in vitro activities between these proteins and their homologs from other bacterial classes. These findings emphasize the important notion that functionalities or activities of certain genes or proteins can not be based solely on sequence homology. Moreover, these studies also showed that the DNA recombination machineries of *M. pneumoniae* and *M. genitalium* differ markedly from those of other bacteria (Sluijter et al. 2012).

Another important aspect of DNA recombination is its regulation. In 2014, it was shown that MgPar recombination in *M. genitalium* was positively regulated by a novel putative transcription factor encoded by ORF MG428 (Burgos and Totten 2014). Notably, mutants lacking MG428 were defective in generating *mgpBC* gene variants (Burgos and Totten 2014). Interestingly, overexpression of MG428 increased expression of *ruvA, ruvB, recA,* as well as other genes, and was associated with increased *mgpBC* gene variation (Burgos and Totten 2014; Torres-Puig et al. 2015). These findings highlight the complexity of the regulation of recombination events in *M. genitalium,* and possibly in *M. pneumoniae,* that should be extensively explored.

Additionally, it is important to note that homologous recombination can be employed for experimental genome alteration of mycoplasmas. Experimental gene modification through homologous recombination has previously been performed to analyze the function of specific genes or proteins of *M. pneumoniae* and *M. genitalium* (Dhandayuthapani et al. 1999; Krishnakumar et al. 2010).

The role of DNA repair machinery to maintain genomic integrity of *Mycoplasma* spp.

Because maintenance of the integrity of the genome is essential for all living organisms, they all have DNA repair systems in place that both detect and repair DNA lesions. These DNA lesions can arise due to physical triggers (e.g. UV light, extreme temperatures and desiccation) and

various chemical agents (e.g. H_2O_2 and cisplatin) causing chemical modifications of the DNA structure. The most important physical trigger of DNA damage is UV light, which is able to induce the covalent linkage of two adjacent pyrimidines (pyrimidine dimers), including cyclobutane pyrimidine dimers (CPD) and pyrimidine–pyrimidone (6–4) photoproducts [(6–4)PP] (Goosen and Moolenaar 2008). Chemical damage to DNA can be caused by compounds such as cisplatin (cis-platinum (II) diaminodichloride), methylmethanesulfonate, ethylmethanesulfonate, N-ethyl-N-nitrosourea, leading to lesions like cisplatin cross-links, thymine glycol products, psoralen monoadducts, O^6-methyl guanine, and abasic sites. For pathogenic bacteria, including *Mycoplasma* spp*.,* such lesions can also be induced through immune responses of the host (e.g., through reactive oxygen species [ROS] generated by macrophages) and the action of antibiotics.

In *Mycoplasma* spp*.,* DNA repair is probably also involved in the recombination-induced antigenic variation. As the recombining repeated elements are not identical (Spuesens et al. 2009), heteroduplex DNA formed at the sites of homologous DNA recombination between donor and recipient repeats will also include mismatched base pairs which have to be corrected. Therefore, a DNA repair system is required to correct mismatched bases after each recombination event. This repair system may not only be involved in the recombination of repeated DNA elements, but also in the maintenance of the integrity of the *Mycoplasma* genomes (Carvalho et al. 2005). This maintenance is especially important for the mycoplasmas, as their genomes (which are often designated as 'minimal genomes') have a limited length and gene content, and do not exhibit significant redundancy of gene or protein functions (Fraser et al. 1995; Glass et al. 2006; Himmelreich et al. 1996).

Various DNA repair systems have evolved in prokaryotes. Of these, nucleotide excision repair (NER) and base excision repair (BER) represent the major DNA repair pathways. Another mechanism includes direct repair by a photolyase (photoreactivation) and UV-damage endonuclease (UVDE) (Goosen and Moolenaar 2008; Hu et al. 2017).

A comparative analysis of the gene content of nine different *Mycoplasma* species has indicated that both *M. pneumoniae* and *M. genitalium* contain genes that are putatively involved in NER. This pathway may be the only 'complete' DNA repair pathway in these species. While these species also have genes potentially involved in BER and recombination repair, they do not encode the full set of BER proteins, as found in other groups of prokaryotes. These isolated genes could complement the NER activity in the mycoplasmas (Carvalho et al. 2005).

The NER pathway in *E. coli* and *Mycoplasma* spp.

NER is a universal process that is involved in the repair of a wide variety of DNA lesions produced by different DNA damaging agents. These lesions include single-base modifications (caused by ionizing radiation, psoralen, etc.), intra- and inter-strand cross-links (caused by UV irradiation, mitomycin C, cisplatin, etc.), mismatched bases, and backbone modifications (Van Houten et al. 2005).

Basically, NER is a multi-step process, involving (1) damage recognition, (2) dual incision of a damaged DNA-containing oligomer, (3) removal of the incised oligomer, and (4) repair DNA synthesis, followed by (5) ligation (Hu et al. 2017). In *Homo sapiens*, NER requires the participation of at least 16 proteins (Mu et al. 1995). Prokaryotes, by contrast, require only three proteins, UvrA, UvrB, and UvrC, which are collectively termed the UvrABC system. The UvrABC system of *E. coli* has been studied extensively and serves as a model for NER (Van Houten 1990; Van Houten et al. 2005).

In *E. coli*, the DNA repair pathway is initiated through the formation of a DNA damage-recognition protein complex containing the UvrA and UvrB proteins (Cordone et al. 2011; Verhoeven et al. 2002) (Fig. 3). UvrA plays a key role in damage recognition, since it preferentially binds to damaged DNA in the absence of other NER components (Stracy et al. 2016). UvrA seems to recognize the unwinding and bending of the 'damaged' DNA, as well as aberrations of the global conformation of the double helix, but does not seem to probe the stability of the base interactions. This mechanism allows UvrA to detect various DNA lesions and achieves broad specificity (Jaciuk et al. 2011). UvrA is also required for the loading of UvrB to form the preincision complex at the site of a DNA lesion (Truglio et al. 2006).

After damage identification, the UvrAB complex unwinds the DNA around the lesion, allowing direct access of UvrB to the lesion (Zou and Van Houten 1999). UvrA then loads UvrB onto the damaged DNA site (Stracy et al. 2016). Following dissociation of UvrA from the complex, UvrB forms a stable UvrB-DNA preincision complex (Theis et al. 2000). The pre-incision complex is subsequently bound by UvrC, which makes incisions around the damaged DNA. It first makes a cut at the fourth or fifth nucleotide from the 3′ side of the damage, followed by incision at the eighth nucleotide from the 5′ side of the damage (Verhoeven et al. 2000). The resulting 12- to 13-nt fragment is ultimately removed by UvrD (helicase II) and the resulting single-nucleotide gap is filled by DNA polymerase I. In the final step, DNA ligase joins the two ends of the nick **(Fig. 3)**.

The importance of the proteins involved in NER was previously shown in several studies using mutant bacteria that were unable to express these proteins (LeCuyer et al. 2010). Missing one component of the system causes reduced survival rates in the bacteria that were exposed to
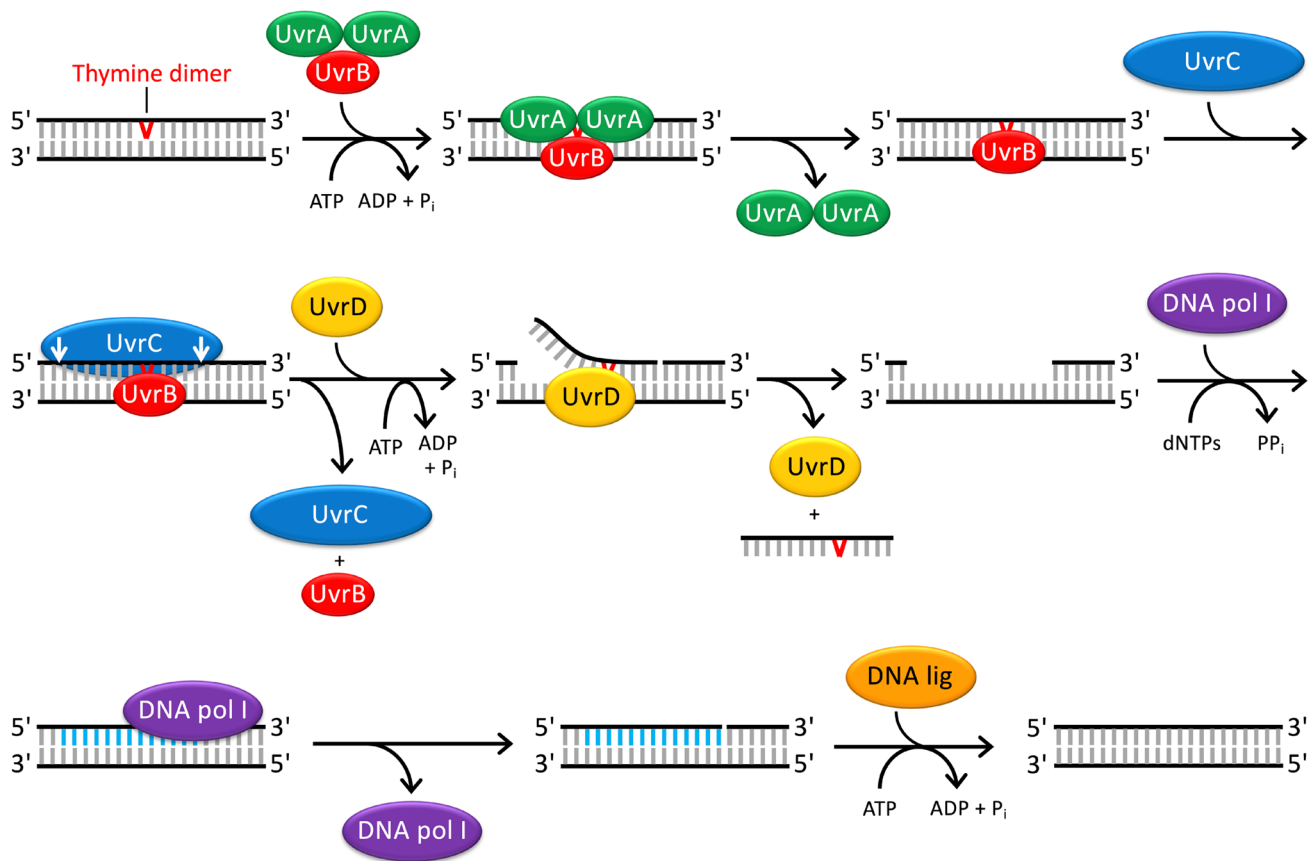
Fig. 3 Recognition and repair of damaged DNA by the UvrABC system of *E. coli*. The DNA repair pathway is initiated through the formation of a DNA damage-recognition protein complex containing the UvrA and UvrB proteins. After damage identification (in this example, a thymine dimer), UvrA dissociates from the complex and UvrB forms a stable UvrB-DNA preincision complex. The pre-incision complex is subsequently bound by UvrC, which makes incisions around the damaged DNA. It first makes an incision at the fourth or fifth nucleotide (nt) from the 3′ side of the damage, followed by incision at the eighth nt from the 5′ side of the damage. The resulting 12- to 13-nt fragment is then removed by UvrD (helicase II). The remaining single-stranded gap is then filled by the action of DNA polymerase I and DNA ligase

DNA-damaging agents. For example, *Mycobacterium tuberculosis* strains that lack either *uvrA* or *uvrB* are more sensitive than the wild-type strain to UV light and other DNA-damaging agents, such as mitomycin C, reactive oxygen, and nitrogen intermediates (Rossi et al. 2011). Similar results were reported for a *Mycobacterium smegmatis uvrA* mutant (Cordone et al. 2011). Furthermore, *E. coli uvrA⁻, uvrB⁻* or *uvrC⁻* mutants were unable to excise UV-induced pyrimidine dimers from their genomes, and were highly mutable (Howard-Flanders and Boyce 1966; Ishii and Kondo 1975; Kato 1972).

UvrA homologs from *M. pneumoniae* and *M. genitalium*

The UvrA proteins belong to the ATP-binding cassette (ABC) superfamily of ATPases (ABC-type ATPases).

Proteins that belong to this superfamily use energy derived from ATP hydrolysis to catalyze a variety of biochemical reactions (Thomas et al. 1986; Wilkens 2015). ABC ATPases share several conserved functional regions in their structures, such as the Walker A/P loop, Q loop, ABC signature, Walker B, D loop, and H loop (Hopfner and Tainer 2003; Thomas et al. 1986; Wilkens 2015). As shown in **Fig. 4**, those regions are highly conserved among the predicted UvrA proteins from several bacterial species, including the representatives from *M. genitalium* and *M. pneumoniae*, UvrA$_{Mge}$ and UvrA$_{Mpn}$, respectively. The latter two proteins are encoded by ORFs MG421 and MPN619, respectively (Fraser et al. 1995; Himmelreich et al. 1996), and are 85% identical on the amino acid level. To date, the functions of these predicted mycoplasma proteins have not yet been determined.
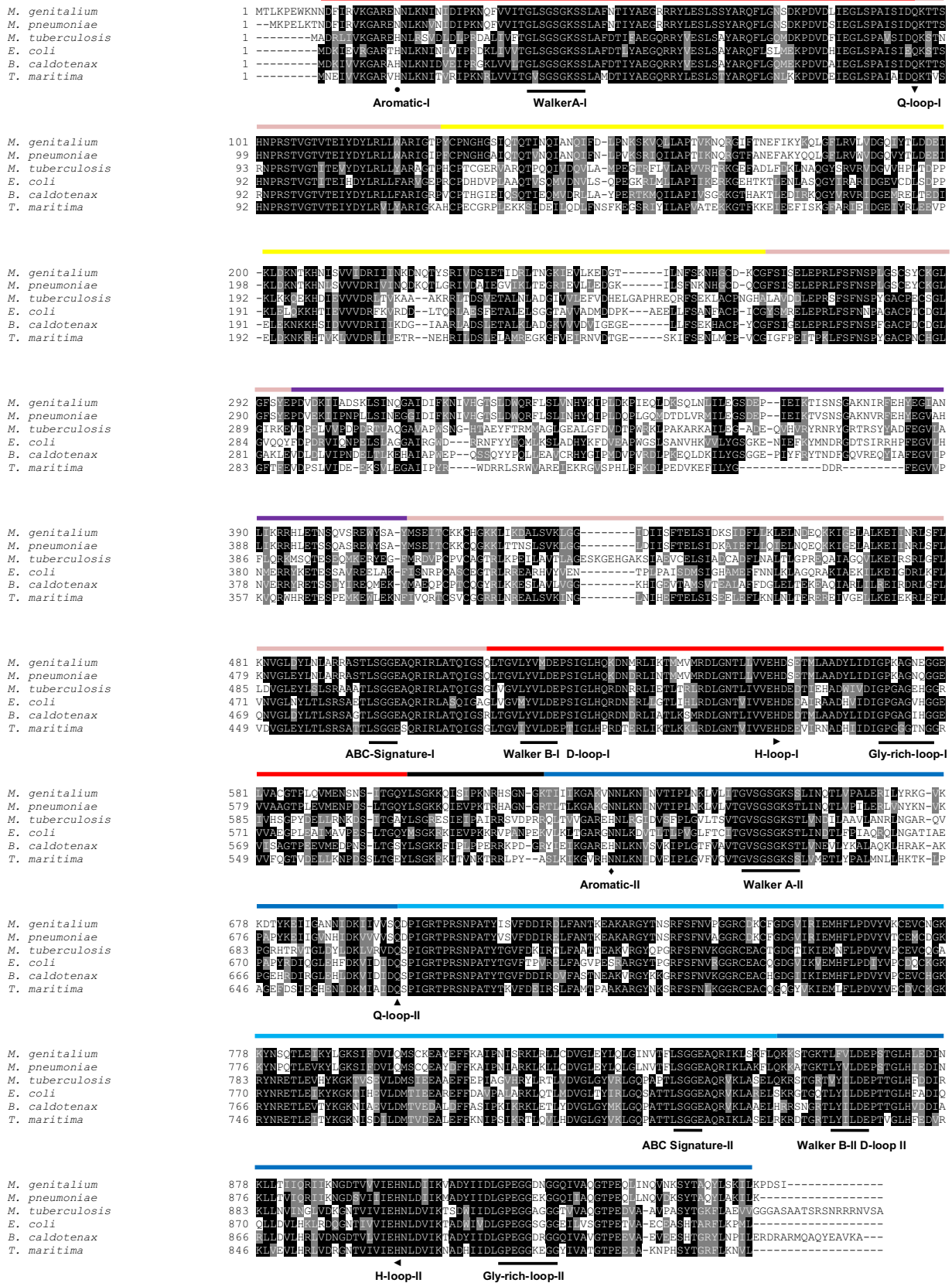
◄**Fig. 4** Multiple alignment of UvrA(-like) amino acid sequences. The multiple alignment was generated with the amino acid sequences predicted to be encoded by the following ORFs [The code in parentheses represents the GenBank accession numbers (https://www.ncbi.nlm.nih.gov/)]: *Mycoplasma genitalium* G37 (NP_073092), *Mycoplasma pneumoniae* M129 (NP_110308.1), *Mycobacterium tuberculosis* (P63380.1), *Escherichia coli* str. K-12 substr. MG1655 (NP_418482.1), *Bacillus caldotenax* (AAK29748.1), and *Thermotaga maritima* (Q9WYV0.1). The program Clustal W (https://www.ebi.ac.uk/Tools/msa/clustalw2) was used to generate a multiple alignment of the amino acid sequences. The program BOXSHADE, version 3.21 (https://www.ch.embnet.org/software/BOX_form.html), was used to generate white letters on black boxes (for residues that are identical in at least four out of eight sequences) and white letters on gray boxes (for similar residues). Secondary structure elements are based on the crystal structure of UvrA from *T. maritima* (Jaciuk et al., 2011 [41]) and indicated as colored-lines above the sequences (ATP-binding I, red; signature I, pink; UvrB binding, yellow; insertion (DNA binding), purple; linker, black; ATP-binding II, blue; and signature II, cyan)

UvrB homologs from *M. pneumoniae* and *M. genitalium*

As explained above, *E. coli* UvrB is a specific DNA damage-binding protein with helicase and strand-separating activities that plays a central role in the multistep process of DNA damage recognition and incision. It interacts first with UvrA, then with UvrC, and finally with UvrD and DNA polymerase I to complete the excision repair (Theis et al. 1999).

UvrB proteins contain five domains, termed 1a, 1b, 2, 3 and 4, with ATP-binding sites located between domains 1a and 3. These proteins have six helicase motifs, three of which [helicase motif I (Walker A), II (Walker B), and III] are located within domain 1a, whereas the other three (helicase motif IV, V, and VI) are located within domain 3 (Linton 2007; Theis et al. 1999, 2000), indicating that UvrB is a member of the helicase superfamily. As depicted in **Fig. 5**, the six helicase motifs are more conserved among bacterial species than are the other regions, indicating the important functional role of these motifs. UvrB uses its helicase-like activity to locally unwind DNA at the site of DNA damage (Linton 2007; Machius et al. 1999). The multiple sequence alignment shown in Fig. 5 also indicates the relatively conserved β-hairpin region of UvrB proteins. It has been shown for *E. coli* UvrB that this region is involved in DNA damage recognition and UvrC-mediated incision (Gordienko and Rupp 1997).

The MG073 and MPN211 ORFs of *M. genitalium* of *M. pneumoniae*, respectively, have previously been predicted as genes that encode UvrB homologs (Fraser et al. 1995; Himmelreich et al. 1996). These homologs, termed UvrB$_{Mpn}$ and UvrB$_{Mge}$, respectively, display 76% identity, while the similarity between UvrB$_{Mge}$ and UvrB$_{Eco}$ is 45%. These *Mycoplasma* genes have not yet been subjected to functional analyses.

UvrC homologs from *M. pneumoniae* and *M. genitalium*

As described above, *E. coli* UvrC can produce single-strand cuts on either side of a DNA lesion. These cuts are executed by two endonuclease domains located at the N-terminal and C-terminal parts of the protein. These domains are separated by a highly variable linker region (Goosen and Moolenaar 2008). While UvrC can bind the DNA-containing lesions alone, the binding efficiency is increased significantly in complex with UvrB (Uphoff and Sherratt 2017).

Although *M. genitalium* MG206 was annotated as an ORF that has the potential to encode a homolog of UvrC proteins (Fraser et al. 1995), the MG206-derived amino acid sequence demonstrated only 25% similarity with the UvrC protein from *E. coli* (UvrC$_{Eco}$). A significantly higher similarity was observed between UvrC$_{Eco}$ and the UvrC homolog from *M. pneumoniae* M129 (59%) **(Fig. 6)**. Interestingly, a *M. genitalium* mutant with a deletion of MG206 was reported to have a growth deficiency as well as an increased sensitivity to UV-induced DNA damage (Burgos et al. 2012), indicating that the MG206-encoded protein (UvrC$_{Mge}$) has a potential role in DNA repair in *M. genitalium*.

Functional UvrD homologs from *M. pneumoniae* and *M. genitalium*

Interestingly, both *M. pneumoniae* and *M. genitalium* do not possess obvious *uvrD* gene homologs in their genomes (Table 1) (Carvalho et al. 2005). Instead, these bacteria possess ORFs (MPN340 and MPN341 of *M. pneumoniae,* and MG244 of *M. genitalium*; Table 1) that encode PcrA helicases that belong to the same family (superfamily 1, SF1) as UvrD (Estevao et al. 2013). PcrA homologs are found in all gram-positive bacteria, including *Bacillus subtilis* and *Staphylococcus aureus,* as well as in bacteria belonging to the *Firmicutes* and *Mollicutes* classes (Petit and Ehrlich 2002; Singleton et al. 2007).

M. genitalium MG244 encodes a single PcrA helicase (PcrA$_{Mge}$) that represents the ortholog of the *M. pneumoniae* MPN341-encoded protein termed PcrA$_{Mpn}$. The second ORF encoding a PcrA homolog in *M. pneumoniae*, MPN340, is not found in other *Mycoplasma* spp. Interestingly, the length of this ORF (1,590 bp) is considerably shorter than that of MPN341 (2,148 bp). Sequence analysis of the MPN340-encoded proteins (PcrA2$_{M129}$ and PcrA2$_{FH}$ from subtype 1 and subtype 2 strains, respectively) showed that they lacked a so-called 2B subdomain that is found in most SF1 DNA helicases. Surprisingly, all four proteins were found to have divalent cation- and ATP-dependent DNA helicase activity (Estevao et al. 2013). It is therefore possible that these proteins may be involved in UvrD-like activities in NER in both *M. pneumoniae* and *M. genitalium*.
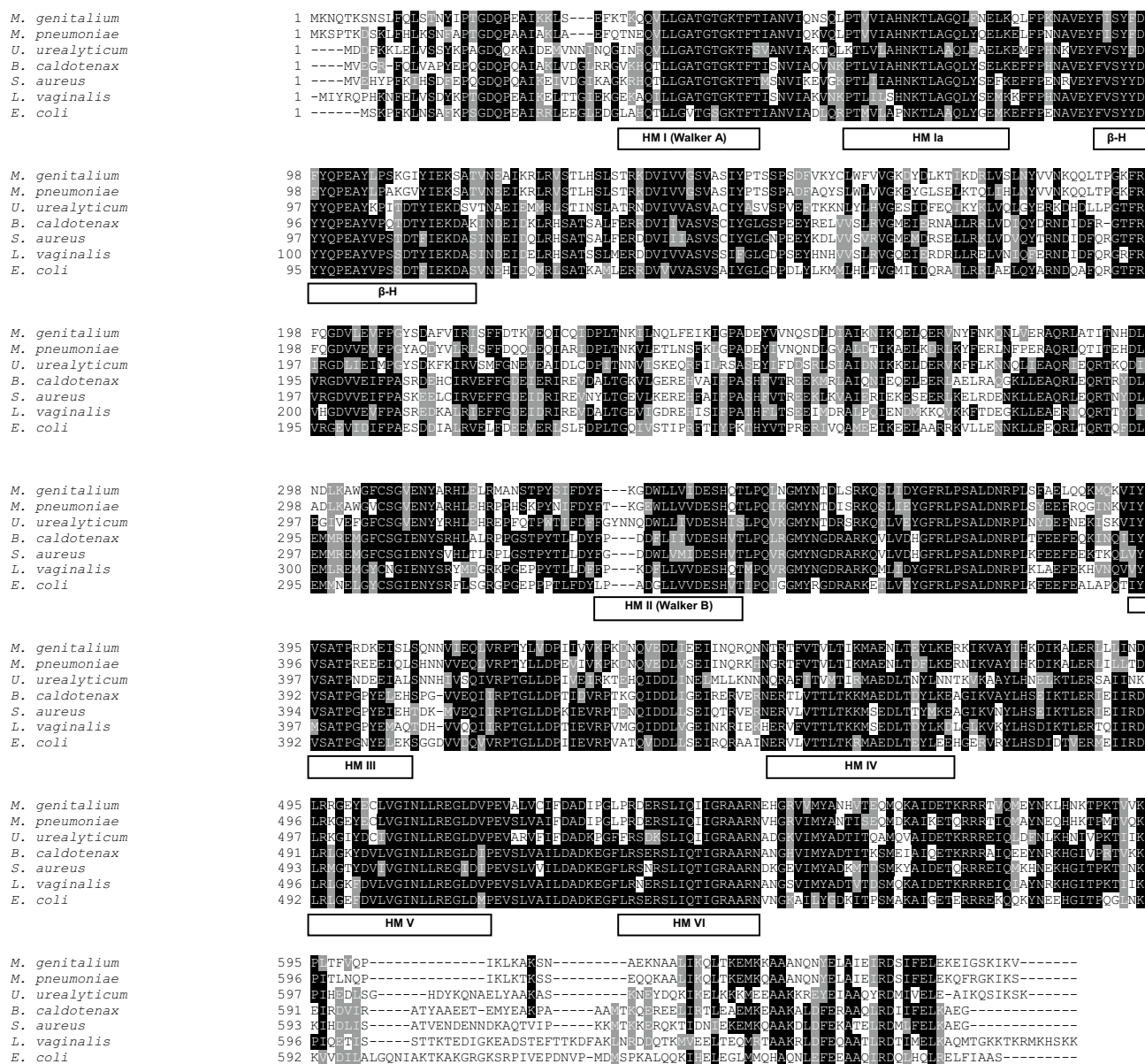
Fig. 5 Multiple alignment of UvrB(-like) amino acid sequences. The multiple alignment was generated with the amino acid sequences predicted to be encoded by the following ORFs [The code in parentheses represents the GenBank accession numbers (https://www.ncbi.nlm.nih.gov/)]: *Mycoplasma genitalium* G37 (AAC71291.1), *Mycoplasma pneumoniae* M129 (NP_109899.1), *Ureaplasma urealyticum* serovar 13 str. (ZP_02931962.1), *Bacillus caldotenax* (2FDC_A), *Staphylococcus aureus* subsp. aureus VRS4 (EIK14428.1), *Lactobacillus vaginalis* ATCC49540 (ZP_03959610.1), and *Escherichia coli* str. K-12 (POA8F8.2). The program Clustal W (https://www.ebi.ac.uk/Tools/msa/clustalw2) was used to generate the multiple alignment of amino acid sequences. The program BOXSHADE, version 3.21 (https://www.ch.embnet.org/software/BOX_form.html), was used to generate white letters on black boxes (for residues that are identical in at least four out of seven sequences) and white letters on gray boxes (for similar residues). The annotation of the helicase motifs I-VI (HM I-VI) and β-hairpin (β-H) is based on the crystal structure of UvrB from *B. caldotenax* (Theis et al. 1999)

Although the *E. coli* UvrD protein only has functional counterparts (but not orthologs) in both *Mycoplasma* spp., the conservation of the other actors of the UvrABC system suggests that the NER pathway in the mycoplasmas may function in a similar way as in *E. coli* (Carvalho et al. 2005).

The BER pathway in *E. coli* and *Mycoplasma* spp.

BER is a major pathway to repair DNA damage due to nucleobase hydrolysis, alkylations, deaminations, or oxidation processes (Wallace 2014). Due to its essential role in maintaining the genomic integrity, BER pathway is highly

M. genitalium	1	---MTTNLKQKLKTAEKKPGCYLKDSNGKVLYVGKASNLFNRVLQYF-KNNPYKTQLLSSQLSDVDEFLLKDENALALNLEAKLINQYQPRLNLVLKQN
M. pneumoniae	1	MLIVNNTLAFKLKNAEHKPGCYLKLDACOVLYVGKAKDLFKRVLHYFL-PNRSFKTRALVERHADLEYVILKNENLALNLEAKLLKQMKPRLNLVLKSN
M. gallisepticum	1	---MFNLRQKLDLAEKKPGCYLKNHLNELIYLGKAKNLYKRVLQYDV-GPKDLKTSKLVNELFDLEBELVNNELEALLEANLIKKHKPRYNILLKDN
M. putrefaciens	1	-----MDLKEKLNLIPNKPGCYLLNSLNQVIYVGKAKNLKKRVSSYFKL-RAHNIKTTLVREISDLNYPVVNNEKESLLDENLIKKYHLPLYNILNDD
B. cereus	1	-VHEHLKERL-LAIIPDQPGCYLMKDKQCTVIYVGKAKVLKNRVSSYFTGSHDG-KTLLVGELIVDLFEYLVTSSNLEALLLDENLIKKYHDPLYNIQLKDD
B. mycoides	1	-VHEHLKERL--ISILPDQPGCYLMKDKQCTVIYVGKAKVLKNRVSYFTGSHDG-KTLLVGELIVDLFEYLVTSSNLEALLLDENLIKKYHDPLYNIQLKDD
E. coli	1	-LSDQFDAKAFLKTLTSQPGVYRMYDAGGTVIYVGKAKDLKKRLSSYFRSNLASRKTEALVAQLQQLLDVTVTHTETALLLBHNYIKTLVQPRYNVLLDD

M. genitalium	97	NLYLYFYLLKAKKPTLELARKLQIKTTKCFGPBASSKFKLRELHDLLLKLFPLRKCAP---HQLNLPCFYLQLGLGLGQQKQ-TDIKEKYQQVLSNLLQF
M. pneumoniae	100	NLYLMLFTTASVKPTLELGRRYEFSKNKYFGPBASSKFKLRDLYDLLLKLFPLRKCAP---HELGLPCFYYQLKMCLGQQAG-BDLPPRYQTTVKGTLQF
M. gallisepticum	97	NLYPYILITLKLKYPRLIYTRNELDPKKGKHLGPBASSEMKAYDLYNLLLKLFPLLNCF----NKLGRKGELMDLNLGLKACTH-EVLBADYEVLKKKLLYF
M. putrefaciens	95	KLYPYIVITNCKLDESYKYLRKYDLKALRNLGPLFQ-KSNARALLLTLQLFPLKFQN---NKDHLPCLYYHLNQCSGAGFK-LVDFSYYEQQLKQLLKF
B. cereus	97	KTYPLIKITALKQPRLLLTRNVKLDKGKYFGPLPN-AQSAHLTKKLLDRMLPLRKCSN----MPDLVCLYYHLGQCLAFQVK-EVTLEQNQEIVDEIIKF
B. mycoides	97	KTYPLIKITALKQPRLLLTRNVKLDKGKYFGPLPN-AQSAHLTKKLLDRMLPLRKCTN----IPDLVCLYYHLGQCLAFQVK-EVTLEQNQEIVDGTVKF
E. coli	100	KSYPLTFLSGLTHPRIALHRGAKLAKGEYFGPFPN-GYAVRETLALLLQKLFBLRLQCENSVYRNLSRPCLQYQLGRCLGPVEGLVLKLBEYAQQVEYVRLF

M. genitalium	193	FNDD-PSVVLNYLKAAEKKASLNQBFEKAQQFLTLQKLAVLELLTKTHHTTLIKQKSSHLBIGYVFQNNVIALLTLLCYEKGELLTDKLQAVLETLEQTDIVELES
M. pneumoniae	196	FNHGLPEOVLNHLQQQEIKASEQQNFEAARHLFLDLQKLAVLELVNMCQTAFIKAKQSELBIGYVFLKNVLALTVBAYVDNQLLLGKNQQVLFLLPQDDEKELES
M. gallisepticum	192	FHNGLADQVLKLLKELELIASEKFLFECAKKYLLQLQKLANLLFLKQILNLYSAKGRIDVLNLQLKENVLCLVLBSYVSSQLLKN----THLCFYYGGEQF
M. putrefaciens	189	FKGDISQVKLTSLINQMNKASBMLQFEGAQRIKELOQSLBFLTTKQNLEFQSDKN-IDLINLQLSENKLVFVLLEYRAGOLTFKLEY---ACDYDQQDIVE
B. cereus	191	LNGGHKEVLRSBLETLMYBASBKLEFERAKELRDQLAHLDAIMLKQKVLIVSDLVGL-RDVFGYAVLKGWLCLVQVFFVRKGKGLLELRLVSLMLPLYD----BFEE
B. mycoides	191	LNGGHKEVLRSBEAAMYBASBKLEFERAKELRDQISHLDAIMLKQKVMIVSDLVBL-RDVFGYVVLKGWLCLVQVFFVRKGKGLLELRLVSLMFPLYD----BFEE
E. coli	199	LSGKDDQVLTQLISLMETASQNLBFBEAAARIRDQLQALVRRYTLKQ-FVSNTGDLL-LDVIGVAFLAGLAQLHVLBLQQGKVLIG-SRSYFPKVPGG-TBLPSE

M. genitalium	292	AIITFLYHHYKT----TPLBSKTTVSLDETN--------LKTLLSDSLKLGVFKPKNGNEKLLTQTVIDNAKHALATKYLKFTSNYDLTQLHK-DLLQLLN
M. pneumoniae	296	ALVTFLYHLYST----NKIPKTLTVSLSEBN--------LSLLANSLKLNVTQPKNGEQKSILQTVIDNALYALNTKWTGFINNLNRAEVHQ-LLAQLLQ
M. gallisepticum	288	VLTSYISQVKTLSLINQMNKASBMLQFEQAQRLK--------ATLLKDSLGLELLIPTSGKMNELMSLALQNVTNELSQKLDSLVKKEQRINLALBLDLKKQLLK
M. putrefaciens	285	LLNLFLQQLYAK----NITELDVLLVPNDLBL--------LELDENILKFSSHALNL-QENLLVNLAKNNAKELLLKAAISKNVNVGDEQEILNLLHQLLSN
B. cereus	286	GFLTLTFTGCILYENSSHFK--PKELLVPGSIL---------SELLERFLELEATLPKKGKKKDLVBLANKNAKIALEEKLYLIERDEERTIKAVELLGKQLLG
B. mycoides	286	GFLTLTFTGCLYEQNNHFK--PKELLVPGSIL---------PELLERFLELDYTLPKKGKKKDLVBLANKNAKIALGEKLYLIERDEERTIKAVELLGKQLLD
E. coli	295	VLBTFLGQLYLQGSQMRTLLPGEILLLDFNLSDKTLLADSLSBTAGRKLNLQTKP--BGDRARYLRLARTNAATALTSLKLSQQSTVHQR----TALASLVLK

M. genitalium	379	TLYLH-SLBLIDYSFYDQNHVVLGGMLRBEDGKKLKLLSRLYNIN-SLKKGLTNHLALLVYRRILSALQTKANLPFSDLLIIDGGKLQLKSVKLQVFSLFSN
M. pneumoniae	383	VPSLQ-SLBLIDYSFYDKDHVVLGAMLRVLPNGKRLKALSRLYNIN-LDHKGLTNYAMLVVYRRIISSLQTHKQLPLSDLLIILDGGKLHQLNTVTKVFASFPN
M. gallisepticum	376	VLKLN-HLBVLDNSNLFNTDKLLSAMLVFLNNQFNKLKYRKYKLKDQQALGDYHLMLYVIYRRLYQALKNN-FLDLPDLLIIDGGKHQLLAALKALVDLQI
M. putrefaciens	372	LKRYPGYLBLFDLSNLYNQFITLACLVLWKNAKPTLSNBLPRKYNIDL-ASYTSLDLRLKFLINLRTTKKLQAD--EKLPDLIIVLDGAKLQLTAAEIISELN-
B. cereus	375	ILTPYLRLBAFDNSNLQGSNEVLSAMLAFIDKGKPKALEYRKYKLKTVLQGPDDYESLRLVRRRYTALKEG--LPLPDLIIIDGGKHLAAASDVLENELG
B. mycoides	375	ILTPYLRLBAFDNSNLQGTNEVLSAMLVFIDGEKSAKLEYRKYKLKTLQGPDDYESLRLVRRRYTALKEN--LPLPDLILIDGGKHLAAASDVLENELG
E. coli	389	LPBLVK-RLBCFDLSHTMGEQTVLSCVVFLANGPLRAEYRYNLTGLTPGDDYAALNQYLRRRYGKATLDDS--K-LPDVILIDGGKLQLAQAANVFA-ELD

M. genitalium	477	VKPPILIGL----LVKL---NKNHQTDLIVLSDFQVKKLAINSPLFHYLATIGTEVDGFAKRSAFNKLSNHQLQNPLLQTFGVGKLITAQLLFDNEQPLNNLK
M. pneumoniae	481	VTQEILIGL----LAKL---NTLHQTDLIVLTDNTTLNLLDNLPLBFLLTTIQLEVDSPAKHNAFLKRVSBARFQNPLLQTBGVGLKTVQLLLDNBQLN----
M. gallisepticum	474	DQRINLIGL----LAKL---NNLHQTDLIVTSNLLBLSLDKLSSVLYFLLANLQLEVHKFAILSFFRLTKAKSLYDSLLQIKLGLKRLQQLLEFLKKLLDEKL
M. putrefaciens	468	-LEIDLIGL----LAKL---DQLHKQDLKLIDVYGQILBVHKQFKKLLVNLSNIQLRVDQLAKSL-GFKKLYHNQLDQDLLTIKLGLKTSQELYKHFKKLLDNLK
B. cereus	472	LYIBL-----MAGLLVKL---DDKLKTSLHLBLGDPPBPVMILBNSQEBYLLQRIQLBVHFFAILTFHRQLLHCKLSVLQSALLDLIPCLGLDKRLKLLKHFGSLKKLK
B. mycoides	472	LYIBL-----MAGLLVKL---DDKLRTSLHLVLGDPPBPVMLLBNSQEBYLLQRIQLBVHFFAILTFHRQLLHCKLSVLQSALLGLIPGVGEKRLKVLLKHFGSLKKLK
E. coli	484	VSWDKNHPLLLGVAKGADRKAGLETLFFEPEGLGFSLPPDSPALHVLLQHIRLLESBDHALGGHRKRAKVKNTLSSLLTIBGVGPKRLQLLLKYMGGLQGLAR

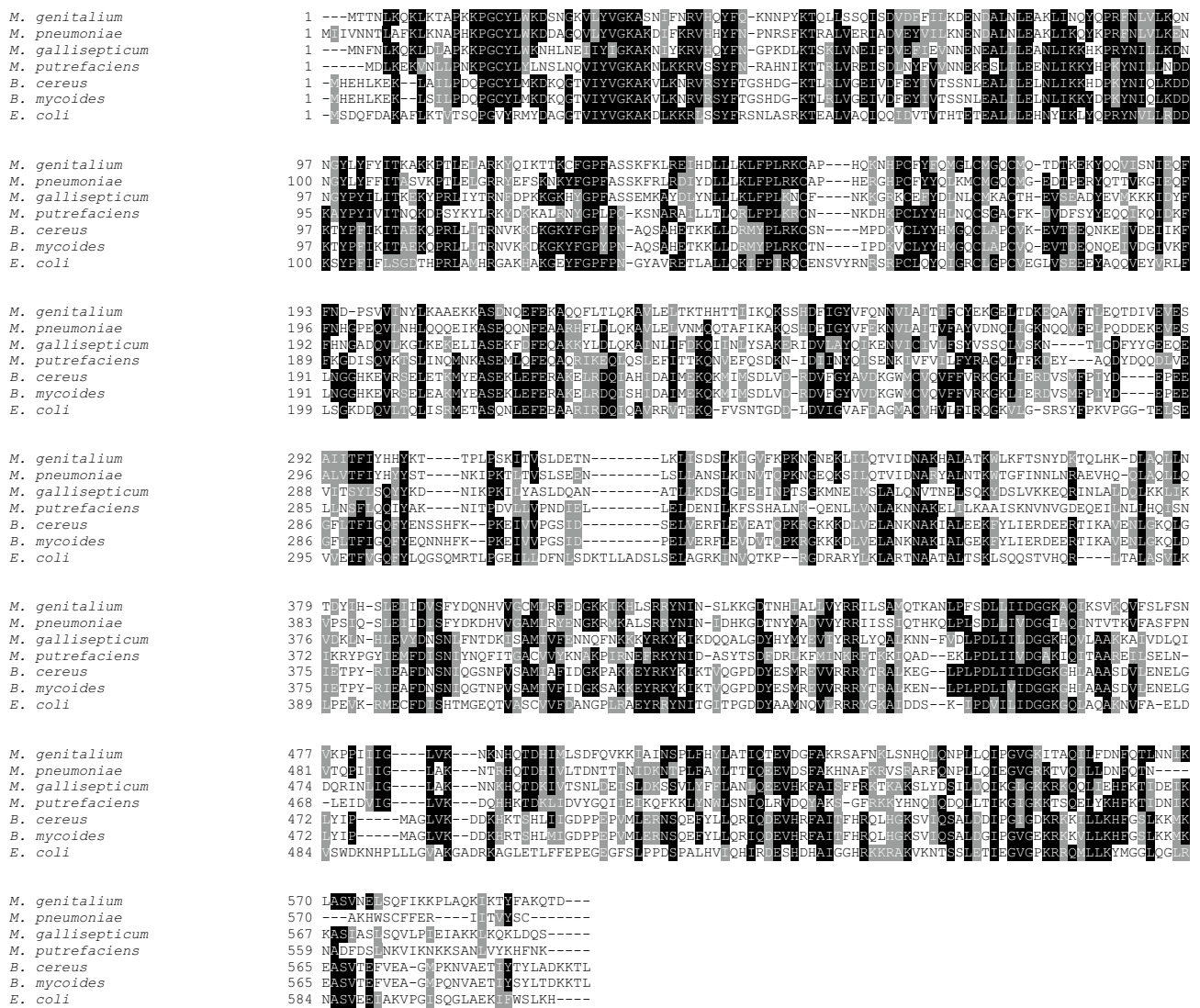|		
M. genitalium	570	LASVNELSQFIKKPLAQKLKTYFAKQTD---
M. pneumoniae	570	---AKHWSCFFER----ILTVYSC-------
M. gallisepticum	567	KASLASLSQVLPLEIAKKLKQKLDQS-----
M. putrefaciens	559	NADFDSLNKVIKNKKSANLVYKHFNK-----
B. cereus	565	EASVBFVEA-GVPKNVAETLYTYLADKKTL
B. mycoides	565	EASVTFVEA-GVPQNVAETLYSYLTDKKTL
E. coli	584	NASVBBLAKVPGLSQGLAEKLEWSLKH----

Fig. 6 Multiple alignment of UvrC(-like) amino acid sequences. Multiple alignment was generated with the amino acid sequences predicted to be encoded by the following ORFs [The code in parentheses represents the GenBank accession numbers (https://www.ncbi.nlm.nih.gov/)]: *Mycoplasma genitalium* G37 (AAC71424.1), *Mycoplasma pneumoniae* M129 (NP_109813.1), *Mycoplasma gallisepticum* str. F (YP_005880751.1), *Mycoplasma putrefaciens* KS1 (YP_004790386.1), *Bacillus cereus* ATCC 4342 (ZP_04286196.1), *Bacillus mycoides* Rock 3–17 (ZP_04158875.1) and *Escherichia coli* str. K-12 (POA860.1). The program Clustal W (https://www.ebi.ac.uk/Tools/msa/clustalw2) was used to generate multiple alignment of amino acid sequences. The program BOXSHADE, version 3.21 (https://www.ch.embnet.org/software/BOX_form.html), was used to generate white letters on black boxes (for residues that are identical in at least four out of seven sequences) and white letters on gray boxes (for similar residues)

conserved during evolution. BER fixes small base lesions that do not induce large distortions in the DNA helix structure (Krokan and Bjoras 2013). In short, BER involves four steps of repair: (1) recognition and incision at the abasic site, (2) gap generation, (3) repair synthesis, and (4) DNA ligation.

The pathway is initiated by the search of DNA lesions by specific DNA glycosylases (for example, uracil DNA glycosylase of *E. coli).* After recognition, damage-specific DNA glycosylases remove the damaged bases by cleaving the *N*-glycosyl bond between the base and the sugar, which results in an abasic or apurinic/apyrimidinic (AP) site in the DNA. Generation of the gap at this specific lesion is performed by class II AP endonucleases (endonuclease IV and exonuclease III), which specifically cleave at abasic sites, and RecJ protein, which excises a 5′-terminal deoxyribose-phosphate residue. Class I AP lysases can also participate in this process by making incisions at the 3′ side of AP sites. Finally, repair synthesis and ligation are performed by DNA polymerase I and DNA ligase, respectively (Dianov and Lindahl 1994; Kow 1994) **(Fig. 7).**
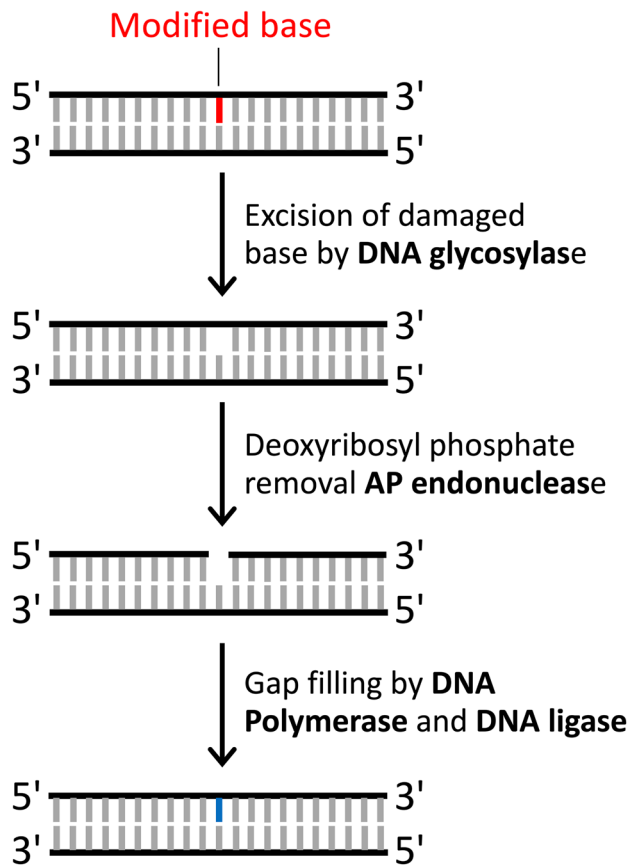
Fig. 7 Recognition and repair of damaged DNA by the Base Excision Repair (BER) system. The pathway is initiated by the recognition of a damaged base by DNA glycosylases. Subsequently, the damaged base is removed, resulting in an abasic or apurinic/apyrimidinic (AP) site in the DNA. Then, the deoxyribosyl phosphate backbone is cleaved by AP endonucleases, followed by repair synthesis and ligation (by DNA polymerase I and DNA ligase, respectively)

Failure to remove AP sites in genomic DNA will result in blockade of DNA replication or mutation of the genome.

Nfo_{Mpn} and Nfo_{Mge}

In contrast to other bacterial classes, which are known to harbor multiple enzymes coordinately involved in BER, *M. pneumoniae* and *M. genitalium* were hypothesized to possess only a single BER-associated enzyme, i.e. a homolog of Nfo (or EndoIV) proteins (Fraser et al. 1995; Himmelreich et al. 1996).

The best characterized Nfo protein is the one derived from *E. coli* (Nfo_{Eco}). Nfo_{Eco} is a multifunctional protein that is known to participate in BER by recognizing and removing AP sites at 5′ of damaged residues. It also has an intrinsic 3′ → 5′ exonuclease activity (Kerins et al. 2003). In addition to BER, the Nfo protein has been demonstrated to be involved in an alternative, overlapping pathway of DNA repair,

termed nucleotide incision repair (NIR). In this system, Nfo functions in the initial incision step of various types of oxidative stress-induced DNA damage to provide target sites for DNA polymerase to finalize the repair process (Golan et al. 2010; Ischenko and Saparbaev 2002). Characterization of Nfo homologs derived from other distantly related species, such as *Thermus thermophilus* (Nfo_{Tth}), *Thermotoga maritima* (Nfo_{Tma}), and *Chlamydophila pneumoniae* (Nfo_{Cpn}), demonstrated that they all have similar activities as Nfo_{Eco} (Back et al. 2006; Kerins et al. 2003; Liu et al. 2007).

The Nfo homologs from *M. pneumoniae* (Nfo_{Mpn}) and *M. genitalium* (Nfo_{Mge}) are encoded by ORFs MPN328 and MG235, respectively. Both proteins possess a high degree of similarity (65% identity) and are capable of removing AP sites at the phosphodiester bond immediately 5′ to the damaged DNA. Nfo_{Mpn} and Nfo_{Mge} were also found to possess 3′ → 5′ exonuclease activity in the presence of Mg^{2+}. In addition, both proteins were shown to recognize and remove larger DNA lesions, such as cholesteryl-modified bases in the DNA (Estevao et al. 2014).

Conclusion and future perspectives

Antigenic variation in *M. pneumoniae* and *M. genitalium* is likely generated through homologous recombination between specific, repetitive DNA elements that are dispersed throughout the bacterial genomes. Characterization of the complete set of proteins involved in homologous DNA recombination in *M. pneumoniae* and *M. genitalium* has indicated that the functional activities of at least some of these proteins are different compared to those of other bacteria.

Additionally, mycoplasmas have evolved strategies to maintain the integrity of their 'minimal' genomes through an efficient DNA repair system. Comparative genomic analysis indicated that the NER pathway may be the only 'complete' DNA repair pathway in these species. While both human mycoplasmas also harbor genes potentially involved in BER, they do not encode a full set of BER-associated proteins, as found in other bacterial taxa.

Further characterization of the protein repertoire involved in homologous DNA recombination and repair in *M. pneumoniae* and *M. genitalium* is important, as it will identify the minimal enzymatic requirements for both generating bacterial genetic diversity (antigenic variation) and maintaining genomic integrity.

Compliance with ethical standards

References

Aparicio D, Torres-Puig S, Ratera M, Querol E, Pinol J, Pich OQ, Fita I (2018) *Mycoplasma genitalium* adhesin P110 binds sialic-acid human receptors. Nat Commun 9(1):4471

Atkinson TP, Balish MF, Waites KB (2008) Epidemiology, clinical manifestations, pathogenesis and laboratory detection of *Mycoplasma pneumoniae* infections. FEMS Microbiol Rev 32(6):956–973

Back JH, Chung JH, Park JH, Han YS (2006) A versatile endonuclease IV from *Thermus thermophilus* has uracil-excising and 3'–5' exonuclease activity. Biochem Biophys Res Commun 346(3):889–895

Baumann L, Cina M, Egli-Gany D, Goutaki M, Halbeisen FS, Lohrer GR, Ali H, Scott P, Low N (2018) Prevalence of *Mycoplasma genitalium* in different population groups: systematic review and meta-analysis. Sex Transm Infect 94(4):255–262

Bianco PR, Tracy RB, Kowalczykowski SC (1998) DNA strand exchange proteins: a biochemical and physical comparison. Front Biosci 3:D570–603

Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277(5331):1453–1462

Burgos R, Totten PA (2014) MG428 is a novel positive regulator of recombination that triggers mgpB and mgpC gene variation in *Mycoplasma genitalium*. Mol Microbiol 94(2):290–306

Burgos R, Wood GE, Young L, Glass JI, Totten PA (2012) RecA mediates MgpB and MgpC phase and antigenic variation in *Mycoplasma genitalium,* but plays a minor role in DNA repair. Mol Microbiol 85(4):669–683

Carvalho FM, Fonseca MM, Batistuzzo De Medeiros S, Scortecci KC, Blaha CA, Agnez-Lima LF (2005) DNA repair in reduced genome: the *Mycoplasma* model. Gene 360(2):111–119

Chen JM, Cooper DN, Chuzhanova N, Ferec C, Patrinos GP (2007) Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet 8(10):762–775

Cordone A, Audrain B, Calabrese I, Euphrasie D, Reyrat JM (2011) Characterization of a *Mycobacterium smegmatis* uvrA mutant impaired in dormancy induced by hypoxia and low carbon concentration. BMC Microbiol 11:231

Deborde M, Pereyre S, Puges M, Bebear C, Desclaux A, Hessamfar M, Le Roy C, Le Marec F, Dabis F, Cazanave C (2019) High prevalence of *Mycoplasma genitalium* infection and macrolide resistance in patients enrolled in HIV pre-exposure prophylaxis program. Med Mal Infect 49(5):347–349

Dehon PM, McGowin CL (2017) The immunopathogenesis of *Mycoplasma genitalium* infections in women: a narrative review. Sex Transm Dis 44(7):428–432

Dhandayuthapani S, Rasmussen WG, Baseman JB (1999) Disruption of gene mg218 of *Mycoplasma genitalium* through homologous recombination leads to an adherence-deficient phenotype. Proc Natl Acad Sci USA 96(9):5227–5232

Dianov G, Lindahl T (1994) Reconstitution of the DNA base excision-repair pathway. Curr Biol 4(12):1069–1076

Dumke R, Schurwanz N, Jacobs E (2008) Characterisation of subtype- and variant-specific antigen regions of the P1 adhesin of *Mycoplasma pneumoniae*. Int J Med Microbiol 298(5–6):483–491

Eggleston AK, West SC (1996) Exchanging partners: recombination in *E. coli.* Trends Genet 12(1):20–6

Estevao S, Sluijter M, Hartwig NG, van Rossum AM, Vink C (2011) Functional characterization of the RuvB homologs from *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. J Bacteriol 193(23):6425–6435

Estevao S, van der Heul HU, Sluijter M, Hoogenboezem T, Hartwig NG, van Rossum AM, Vink C (2013) Functional analysis of the superfamily 1 DNA helicases encoded by *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. PLoS ONE 8(7):e70870

Estevao S, van der Spek PE, van Rossum AM, Vink C (2014) Uncoupling of the apyrimidinic/apurinic endonucleolytic and 3'–>5' exonucleolytic activities of the Nfo protein of Mycoplasma pneumoniae through mutation of specific amino acid residues. Microbiology 160(Pt 6):1087–1100

Fookes MC, Hadfield J, Harris S, Parmar S, Unemo M, Jensen JS, Thomson NR (2017) *Mycoplasma genitalium:* whole genome sequence analysis, recombination and population structure. BMC Genom 18(1):993

Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, Fritchman RD, Weidman JF, Small KV, Sandusky M, Fuhrmann J, Nguyen D, Utterback TR, Saudek DM, Phillips CA, Merrick JM, Tomb JF, Dougherty BA, Bott KF, Hu PC, Lucier TS, Peterson SN, Smith HO, Hutchison CA, Venter JC (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270(5235):397–403

Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, Smith HO, Venter JC (2006) Essential genes of a minimal bacterium. Proc Natl Acad Sci USA 103(2):425–430

Golan G, Ishchenko AA, Khassenov B, Shoham G, Saparbaev MK (2010) Coupling of the nucleotide incision and 3'–>5' exonuclease activities in *Escherichia coli* endonuclease IV: structural and genetic evidences. Mutat Res 685(1–2):70–79

Goosen N, Moolenaar GF (2008) Repair of UV damage in bacteria. DNA Repair (Amst) 7(3):353–379

Gordienko I, Rupp WD (1997) The limited strand-separating activity of the UvrAB protein complex and its role in the recognition of DNA damage. EMBO J 16(4):889–895

Hardy RD, Jafri HS, Olsen K, Hatfield J, Iglehart J, Rogers BB, Patel P, Cassell G, McCracken GH, Ramilo O (2002) *Mycoplasma pneumoniae* induces chronic respiratory infection, airway hyperreactivity, and pulmonary inflammation: a murine model of infection-associated chronic reactive airway disease. Infect Immun 70(2):649–654

Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC, Herrmann R (1996) Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res 24(22):4420–4449

Himmelreich R, Plagens H, Hilbert H, Reiner B, Herrmann R (1997) Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. Nucleic Acids Res 25(4):701–712

Hopfner KP, Tainer JA (2003) Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. Curr Opin Struct Biol 13(2):249–255

Howard-Flanders P, Boyce RP (1966) DNA repair and genetic recombination: studies on mutants of *Escherichia coli* defective in these processes. Radiat Res:Suppl 6:156+

Hu J, Selby CP, Adar S, Adebali O, Sancar A (2017) Molecular mechanisms and genomic maps of DNA excision repair in *Escherichia coli* and humans. J Biol Chem 292(38):15588–15597

Ingleston SM, Dickman MJ, Grasby JA, Hornby DP, Sharples GJ, Lloyd RG (2002) Holliday junction binding and processing by

the RuvA protein of *Mycoplasma pneumoniae*. Eur J Biochem 269(5):1525–1533

Ischenko AA, Saparbaev MK (2002) Alternative nucleotide incision repair pathway for oxidative DNA damage. Nature 415(6868):183–187

Ishii Y, Kondo S (1975) Comparative analysis of deletion and base-change mutabilities of *Escherichia coli* B strains differing in DNA repair capacity (wild-type, uvrA-, polA-, recA-) by various mutagens. Mutat Res 27(1):27–44

Iverson-Cabral SL, Astete SG, Cohen CR, Rocha EP, Totten PA (2006) Intrastrain heterogeneity of the mgpB gene in *Mycoplasma genitalium* is extensive *in vitro* and *in vivo* and suggests that variation is generated via recombination with repetitive chromosomal sequences. Infect Immun 74(7):3715–3726

Iverson-Cabral SL, Astete SG, Cohen CR, Totten PA (2007) mgpB and mgpC sequence diversity in *Mycoplasma genitalium* is generated by segmental reciprocal recombination with repetitive chromosomal sequences. Mol Microbiol 66(1):55–73

Jaciuk M, Nowak E, Skowronek K, Tanska A, Nowotny M (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. Nat Struct Mol Biol 18(2):191–197

Kato T (1972) Excision repair characteristics of recB - res - and uvrC - strains of *Escherichia coli*. J Bacteriol 112(3):1237–1246

Kerins SM, Collins R, McCarthy TV (2003) Characterization of an endonuclease IV 3'-5' exonuclease activity. J Biol Chem 278(5):3048–3054

Kow YW (1994) Base excision repair in *E. coli*-an overview. Ann N Y Acad Sci 726:178–180

Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM (1994) Biochemistry of homologous recombination in *Escherichia coli*. Microbiol Rev 58(3):401–465

Krishnakumar R, Assad-Garcia N, Benders GA, Phan Q, Montague MG, Glass JI (2010) Targeted chromosomal knockouts in *Mycoplasma pneumoniae*. Appl Environ Microbiol 76(15):5297–5299

Krokan HE, Bjoras M (2013) Base excision repair. Cold Spring Harb Perspect Biol 5(4):a012583

Kuzminov A (1999) Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. Microbiol Mol Biol Rev 63(4):751–813

LeCuyer BE, Criss AK, Seifert HS (2010) Genetic characterization of the nucleotide excision repair system of *Neisseria gonorrhoeae*. J Bacteriol 192(3):665–673

Linton KJ (2007) Structure and function of ABC transporters. Physiology (Bethesda) 22:122–130

Liu X, Zhang Y, Liang R, Hou J, Liu J (2007) Characterization of the 3' exonuclease of *Chlamydophila pneumoniae* endonuclease IV on double-stranded DNA and the RNA strand of RNA/DNA hybrid. Biochem Biophys Res Commun 361(4):987–993

Lluch-Senar M, Cozzuto L, Cano J, Delgado J, Llorens-Rico V, Pereyre S, Bebear C, Serrano L (2015) Comparative "-omics" in *Mycoplasma pneumoniae* clinical isolates reveals key virulence factors. PLoS ONE 10(9):e0137354

Ma L, Jensen JS, Mancuso M, Myers L, Martin DH (2015) Kinetics of genetic variation of the *Mycoplasma genitalium* MG192 gene in experimentally infected chimpanzees. Infect Immun 84(3):747–753

Ma L, Jensen JS, Myers L, Burnett J, Welch M, Jia Q, Martin DH (2007) *Mycoplasma genitalium:* an efficient strategy to generate genetic variation from a minimal genome. Mol Microbiol 66(1):220–236

Ma L, Mancuso M, Williams JA, Van Der Pol B, Fortenberry JD, Jia Q, Myers L, Martin DH (2014) Extensive variation and rapid shift of the MG192 sequence in *Mycoplasma genitalium* strains from patients with chronic infection. Infect Immun 82(3):1326–1334

Machius M, Henry L, Palnitkar M, Deisenhofer J (1999) Crystal structure of the DNA nucleotide excision repair enzyme UvrB from *Thermus thermophilus*. Proc Natl Acad Sci USA 96(21):11717–11722

McGowin CL, Spagnuolo RA, Pyles RB (2010) *Mycoplasma genitalium* rapidly disseminates to the upper reproductive tracts and knees of female mice following vaginal inoculation. Infect Immun 78(2):726–736

McGowin CL, Totten PA (2017) The unique microbiology and molecular pathogenesis of *Mycoplasma genitalium*. J Infect Dis 216(Suppl 2):S382–S388

Meyer Sauteur PM, Huizinga R, Tio-Gillen AP, Roodbol J, Hoogenboezem T, Jacobs E, van Rijn M, van der Eijk AA, Vink C, de Wit MC, van Rossum AM, Jacobs BC (2016) *Mycoplasma pneumoniae* triggering the Guillain-Barre syndrome: a case-control study. Ann Neurol 80(4):566–580

Meyer Sauteur PM, Jacobs BC, Spuesens EB, Jacobs E, Nadal D, Vink C, van Rossum AM (2014a) Antibody responses to *Mycoplasma pneumoniae:* role in pathogenesis and diagnosis of encephalitis? PLoS Pathog 10(6):e1003983

Meyer Sauteur PM, Unger WWJ, van Rossum AMC, Berger C (2018) The art and science of diagnosing *Mycoplasma pneumoniae* infection. Pediatr Infect Dis J 37(11):1192–1195

Meyer Sauteur PM, van Rossum AM, Vink C (2014b) *Mycoplasma pneumoniae* in children: carriage, pathogenesis, and antibiotic resistance. Curr Opin Infect Dis 27(3):220–227

Mu D, Park CH, Matsunaga T, Hsu DS, Reardon JT, Sancar A (1995) Reconstitution of human DNA repair excision nuclease in a highly defined system. J Biol Chem 270(6):2415–2418

Napierala Mavedzenge S, Muller EE, Lewis DA, Chipato T, Morrison CS, Weiss HA (2015) *Mycoplasma genitalium* is associated with increased genital HIV type 1 RNA in Zimbabwean women. J Infect Dis 211(9):1388–1398

Peterson SN, Bailey CC, Jensen JS, Borre MB, King ES, Bott KF, Hutchison CA (1995) Characterization of repetitive DNA in the *Mycoplasma genitalium* genome: possible role in the generation of antigenic variation. Proc Natl Acad Sci USA 92(25):11829–11833

Peterson SN, Hu PC, Bott KF, Hutchison CA (1993) A survey of the *Mycoplasma genitalium* genome by using random sequencing. J Bacteriol 175(24):7918–7930

Petit MA, Ehrlich D (2002) Essential bacterial helicases that counteract the toxicity of recombination proteins. EMBO J 21(12):3137–3147

Qin L, Chen Y, You X (2019) Subversion of the immune response by human pathogenic *Mycoplasmas*. Front Microbiol 10:1934

Razin S, Jacobs E (1992) *Mycoplasma* adhesion. J Gen Microbiol 138(3):407–422

Rocha EP, Blanchard A (2002) Genomic repeats, genome plasticity and the dynamics of *Mycoplasma* evolution. Nucleic Acids Res 30(9):2031–2042

Rossi F, Khanduja JS, Bortoluzzi A, Houghton J, Sander P, Guthlein C, Davis EO, Springer B, Bottger EC, Relini A, Penco A, Muniyappa K, Rizzi M (2011) The biological and structural characterization of *Mycobacterium tuberculosis* UvrA provides novel insights into its mechanism of action. Nucleic Acids Res 39(16):7316–7328

Ruland K, Wenzel R, Herrmann R (1990) Analysis of three different repeated DNA elements present in the P1 operon of *Mycoplasma pneumoniae:* size, number and distribution on the genome. Nucleic Acids Res 18(21):6311–6317

Shao W, Li X, Goraya MU, Wang S, Chen JL (2017) Evolution of influenza A virus by mutation and re-assortment. Int J Mol Sci. https://doi.org/10.3390/ijms18081650

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem 76:23–50

Sluijter M, Aslam M, Hartwig NG, van Rossum AM, Vink C (2011) Identification of amino acid residues critical for catalysis of Holliday junction resolution by *Mycoplasma genitalium* RecU. J Bacteriol 193(15):3941–3948

Sluijter M, Estevao S, Hoogenboezem T, Hartwig NG, van Rossum AM, Vink C (2012) The RuvA homologues from *Mycoplasma genitalium* and *Mycoplasma pneumoniae* exhibit unique functional characteristics. PLoS ONE 7(5):e38301

Sluijter M, Hoogenboezem T, Hartwig NG, Vink C (2008) The *Mycoplasma pneumoniae* MPN229 gene encodes a protein that selectively binds single-stranded DNA and stimulates recombinase A-mediated DNA strand exchange. BMC Microbiol 8:167

Sluijter M, Kaptein E, Spuesens EB, Hoogenboezem T, Hartwig NG, Van Rossum AM, Vink C (2010) The *Mycoplasma genitalium* MG352-encoded protein is a Holliday junction resolvase that has a non-functional orthologue in *Mycoplasma pneumoniae*. Mol Microbiol 77(5):1261–1277

Sluijter M, Spuesens EB, Hartwig NG, van Rossum AM, Vink C (2009) The *Mycoplasma pneumoniae* MPN490 and *Mycoplasma genitalium* MG339 genes encode reca homologs that promote homologous DNA strand exchange. Infect Immun 77(11):4905–4911

Spuesens EB, Brouwer RW, Mol KH, Hoogenboezem T, Kockx CE, Jansen R, Van IWF, Van Rossum AM, Vink C (2016) Comparison of *Mycoplasma pneumoniae* genome sequences from strains isolated from symptomatic and asymptomatic patients. Front Microbiol 7:1701

Spuesens EB, Fraaij PL, Visser EG, Hoogenboezem T, Hop WC, van Adrichem LN, Weber F, Moll HA, Broekman B, Berger MY, van Rijsoort-Vos T, van Belkum A, Schutten M, Pas SD, Osterhaus AD, Hartwig NG, Vink C, van Rossum AM (2013) Carriage of *Mycoplasma pneumoniae* in the upper respiratory tract of symptomatic and asymptomatic children: an observational study. PLoS Med 10(5):e1001444

Spuesens EB, Meyer Sauteur PM, Vink C, van Rossum AM (2014) *Mycoplasma pneumoniae* infections-does treatment help? J Infect 69(Suppl 1):S42–S46

Spuesens EB, Oduber M, Hoogenboezem T, Sluijter M, Hartwig NG, van Rossum AM, Vink C (2009) Sequence variations in RepMP2/3 and RepMP4 elements reveal intragenomic homologous DNA recombination events in *Mycoplasma pneumoniae*. Microbiology 155(Pt 7):2182–2196

Spuesens EB, van de Kreeke N, Estevao S, Hoogenboezem T, Sluijter M, Hartwig NG, van Rossum AM, Vink C (2011) Variation in a surface-exposed region of the *Mycoplasma pneumoniae* P40 protein as a consequence of homologous DNA recombination between RepMP5 elements. Microbiology 157(Pt 2):473–483

Stracy M, Jaciuk M, Uphoff S, Kapanidis AN, Nowotny M, Sherratt DJ, Zawadzki P (2016) Single-molecule imaging of UvrA and UvrB recruitment to DNA lesions in living *Escherichia coli*. Nat Commun 7:12568

Su CJ, Chavoya A, Baseman JB (1988) Regions of *Mycoplasma pneumoniae* cytadhesin P1 structural gene exist as multiple copies. Infect Immun 56(12):3157–3161

Theis K, Chen PJ, Skorvaga M, Van Houten B, Kisker C (1999) Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. EMBO J 18(24):6899–6907

Theis K, Skorvaga M, Machius M, Nakagawa N, Van Houten B, Kisker C (2000) The nucleotide excision repair protein UvrB, a helicase-like enzyme with a catch. Mutat Res 460(3–4):277–300

Thomas DC, Kunkel TA, Casna NJ, Ford JP, Sancar A (1986) Activities and incision patterns of ABC excinuclease on modified DNA containing single-base mismatches and extrahelical bases. J Biol Chem 261(31):14496–14505

Torres-Puig S, Broto A, Querol E, Pinol J, Pich OQ (2015) A novel sigma factor reveals a unique regulon controlling cell-specific recombination in *Mycoplasma genitalium*. Nucleic Acids Res 43(10):4923–4936

Truglio JJ, Croteau DL, Van Houten B, Kisker C (2006) Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev 106(2):233–252

Uphoff S, Sherratt DJ (2017) Single-molecule analysis of bacterial DNA repair and mutagenesis. Annu Rev Biophys 46:411–432

Van Houten B (1990) Nucleotide excision repair in *Escherichia coli*. Microbiol Rev 54(1):18–51

Van Houten B, Croteau DL, DellaVecchia MJ, Wang H, Kisker C (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. Mutat Res 577(1–2):92–117

Verhoeven EE, van Kesteren M, Moolenaar GF, Visse R, Goosen N (2000) Catalytic sites for 3' and 5' incision of *Escherichia coli* nucleotide excision repair are both located in UvrC. J Biol Chem 275(7):5120–5123

Verhoeven EE, Wyman C, Moolenaar GF, Goosen N (2002) The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. EMBO J 21(15):4196–4205

Vink C, Rudenko G, Seifert HS (2012) Microbial antigenic variation mediated by homologous DNA recombination. FEMS Microbiol Rev 36(5):917–948

Waites KB, Xiao L, Liu Y, Balish MF, Atkinson TP (2017) *Mycoplasma pneumoniae* from the respiratory tract and beyond. Clin Microbiol Rev 30(3):747–809

Wallace SS (2014) Base excision repair: a critical player in many games. DNA Repair (Amst) 19:14–26

Wilkens S (2015) Structure and mechanism of ABC transporters. F1000Prime Rep 7:14

Wilson MH, Collier AM (1976) Ultrastructural study of *Mycoplasma pneumoniae* in organ culture. J Bacteriol 125(1):332–339

Wood GE, Patton DL, Cummings PK, Iverson-Cabral SL, Totten PA (2017) Experimental infection of pig-tailed macaques *(Macaca nemestrina)* with *Mycoplasma genitalium*. Infect Immun. https://doi.org/10.1128/IAI.00738-16

Zou Y, Van Houten B (1999) Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. EMBO J 18(17):4889–4901