

Role of Orc9 in DNA Damage Response in *Haloferax volcanii*

Çağla Tosun, BSc (Hons)

Student ID: 4337684

Thesis submitted to the University of Nottingham
for the degree of Master of Research

September 2019

23,888 words



The University of
Nottingham

UNITED KINGDOM • CHINA • MALAYSIA

Contents

Abstract	3
Acknowledgements	4
Chapter 1: Introduction	
1.1 Archaea	5
1.2 Replication Initiation	8
1.3. DNA Repair	18
1.4 Aims and Objectives	29
Chapter 2: Materials and Methods	
2.1 Materials	30
2.2 Methods	34
Chapter 3: DNA Damage Sensitivity Assays	
3.1 Background	49
3.2 Aims	49
3.3 Results	50
3.4 Discussion	63
Chapter 4: Interactions of Orc9	
4.1 Background	65
4.2 Aims	65
4.3 Results	66
4.4 Discussion	79
Chapter 5: Transcriptome profiling of <i>orc9</i>	
5.1 Background	81
5.2 Aims	82
5.3 Results	82
5.4 Discussion	89
Chapter 6: Conclusions and Future Perspectives	91
References	94
Appendix	100

Abstract

Archaeal species encode numerous copies of the DNA replication initiator protein Orc1. *Haloferax volcanii* contains 16 *orc1* genes, however only two of these are conserved throughout species: one is always linked to a replication origin (*Hvo-orc1*) while the other is never linked to a replication origin (*Hvo-orc9*). Unlike Orc1, Orc9 has been shown to play no role in replication, with preliminary data showing strains lacking Orc9 have an acute sensitivity to DNA damaging agents such as UV radiation. This suggests a role for Orc9 in the DNA damage response.

Here, we investigate the potential function of Orc9 as a DNA damage response factor in *Haloferax volcanii* through analysis of transcription profiles of wild type and $\Delta orc9$ strains in response to treatment with mitomycin-C (MMC). Cell killing assays using several DNA damaging agents such as MMS, phleomycin, H₂O₂ and gamma radiation were performed, and results showed that the $\Delta orc9$ was hypersensitive to all these mutagens except H₂O₂. Standard growth assays showed that in the presence of sub-lethal doses of MMC, $\Delta orc9$ strains have slower growth compared to wild type strains. These growth assays will inform the ideal concentration at which each DNA damaging agent induces a DNA damage response large enough to influence the cell's transcriptome profile and that also gives some cell viability. The correct concentrations of each DNA damaging agent were used to treat both the wild type and the $\Delta orc9$ cells prior to RNA extraction for downstream sequencing. The RNA-sequencing data revealed that Orc9 has an influence on the global genome expression of *H. volcanii*.

Overall, this study has provided insights into the role of Orc9 in DNA damage response. Orc9 is likely to be a key regulator that orchestrates a network of events that lead to repair of DNA damage. If this is the case, this means that DDR is conserved in all domains of life including Archaea.

Acknowledgements

First, I would like to thank my supervisor Prof. Thorsten Allers for letting me work on this interesting and challenging project. He has been a wonderful mentor to me this year. I am very appreciative of the things he taught me about science and academia in general. Thank you for being so patient with me every time I came to you to rant about yet another thing!

Thank you to members of the Allers lab (and beyond): Patricia, Ambika, Beccy, Nathan, Vicky, Laura, Chris, James and Jon. You have made my time in the lab so enjoyable. Thank you all for being such good friends and for letting me have a go at the crosswords first!

Special thanks to Vicky, who has been an incredible friend and a day-to-day supervisor to me all year. Thank you for being so patient with me when I worried about things a little too much. #TeamOrc9

Special thanks to Owen who has been the most supportive boyfriend anyone can have. I could not have done this without you. Also, to Jacqui and David who have given me a home away from home.

A very special thanks goes to my Mum! Dünyanın en fedakar ve çocuklarını en çok seven annesine sonsuz teşekkürler! Seni ve kardeşim Çağrı'yı çok seviyorum! Siz olmasanız bunu başaramazdım.

Finally, thank you Dad, for letting me chase my dream of coming to the UK to start my undergraduate degree 4 years ago. I know you would have been so proud of what I have achieved, so this is for you.

Chapter 1: Introduction

1.1 Archaea

Prior to the advent of nucleotide sequencing, organisms were classified using traditional identification methods that relied on visible characteristics such as cell morphology, physiology and pathogenicity. In 1977, Carl Woese and George Fox made a paradigm-shifting discovery and identified Archaea as a separate lineage from Bacteria (initially they used the terms Archaeobacteria and Eubacteria, respectively) (Woese and Fox, 1977). They used 16S rRNA (small subunit ribosomal RNA) as a molecular chronometer since it is a highly conserved sequence that is expressed ubiquitously in all free-living organisms. Phylogenetic relationships can be established between species by comparing 16S rRNA sequences owing to its slow-evolving, highly conserved nature. (Allers and Mevarech, 2005a). The three-domain tree of life identified Bacteria, Archaea and Eukarya as separate sister lineages (Figure 1.1). Numerous other studies supported the new classification and showed that Archaea were genetically more similar to Eukarya, especially in terms of their machinery for DNA repair and replication (Ishino *et al.*, 2013).

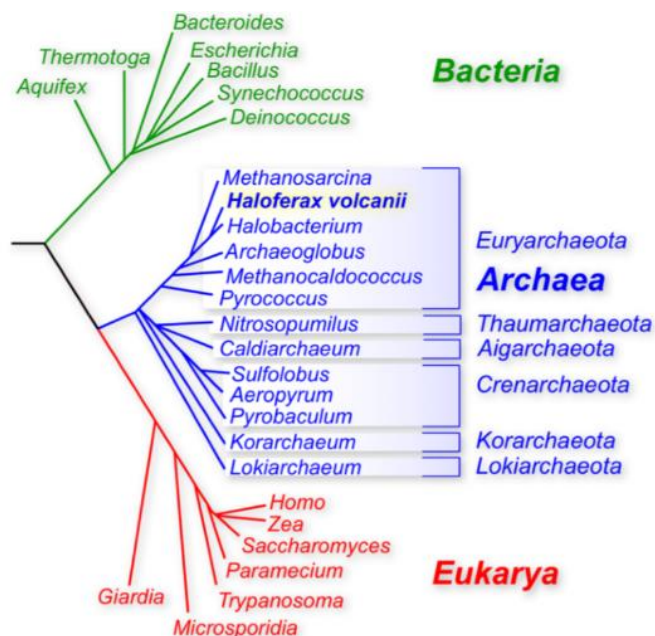


Figure 1.1: **Woese's three domain tree of life.** Phylogenetic tree based on 16S rRNA sequencing, defining Bacteria, Archaea and Eukarya as separate lineages. Adapted from (Allers and Mevarech, 2005b) .

Further 16S rRNA sequencing analyses identified two distinct groups in Archaea: the Crenarchaeota and the Euryarchaeota. The definition of the TACK (Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota) superphylum came when the advances in sequencing technologies allowed for collection of genomic data from several archaeal lineages (Guy and Ettema, 2011). Recently, a new group of Archaea named Lokiarchaeota was discovered in the Atlantic Ocean near a hydrothermal vent called Loki's Castle. Lokiarchaeota belongs to the Asgard superphylum along with Thorarchaeota, Odinararchaeota and Heimdallarchaeota. The genomes of these groups of archaea have high levels of similarities to eukaryotes and homologs of genes that were previously thought to be found only in eukaryotes. According to the recent findings, Heimdallarchaeota appears to be the closest archaeal relative of eukaryotes (Macleod *et al.*, 2019). Further analyses suggest a two domain tree of life where Eukaryotes are a subgroup of Archaea emerging from the Asgard superphylum (Zaremba-Niedzwiedzka *et al.*, 2017) (Figure 1.2).

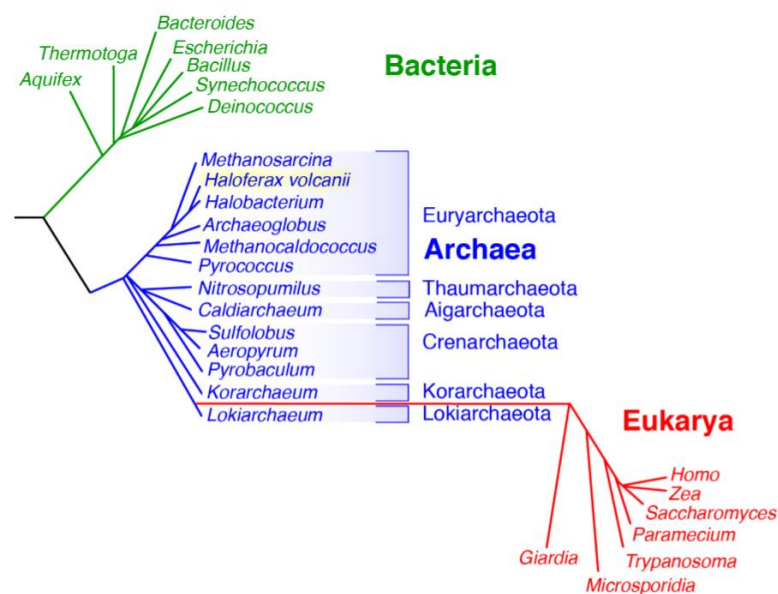


Figure 1.2: **Two domain tree of life.** Eukaryotes are shown branching from within Archaea. Adapted from (Allers and Mevarech, 2005b).

Halophiles

Archaea are renowned for inhabiting extreme environments such as hydrothermal vents, salt lakes and acid pools. However, they can be found in moderate environments as well, such as soil and marshes (DeLong and Pace, 2001). Halophiles are a group of Archaea that preferentially thrive in high salt environments. They maintain an osmotic

balance between intracellular and extracellular environments by using two alternative mechanisms. Cells can accumulate potassium ions inside the cell to maintain a high level of intracellular salt by using a ‘salt in’ mechanism (Oren, 1999). Alternatively, the ‘salt-out’ mechanism is used by many halophilic bacteria and eukaryotes whereby ions are constantly pumped out of the cell and organic solutes (such as glycerol) are kept inside the cell to maintain a high osmotic concentration (Oren, 2008). Halophiles using the ‘salt-in’ strategy possess proteins adapted to function in high intracellular salt conditions.

***Haloferax volcanii* as a model organism**

Haloferax volcanii is a halophilic euryarchaeon that was first discovered and isolated from the Dead Sea (Mullakhanbhai and Larsen, 1975). It was named after Ben Volcani who identified life in the Dead Sea for the first time (Wilkansky, 1936). It grows aerobically at 45°C and requires a salt concentration of 1.7 - 2.5 M NaCl in the media to survive. The generation time of *H. volcanii* is ~ 2.5 hours and it forms colonies on agar plates after 4-5 days of incubation. *H. volcanii* cell morphology is coccoid and the production of carotenoid pigments inside the cell gives cells a red colour. Its membrane is surrounded by a glycoprotein surface (S-) layer (Sleytr and Messner, 1983). *H. volcanii* may possess up to 20 copies of its genome due to its highly polyploid nature and this makes it difficult to study its cell cycle. Replication profiles of *H. volcanii* obtained by deep sequencing produced a maximum:minimum copy number ratio of >2:1, which suggests that there is no defined S-phase (Hawkins *et al.*, 2013).

The genome of the wild-type *H. volcanii* (DS2) is made up of 5 components: a main circular chromosome (2.9 Mb), 3 mega-plasmids: pHV1 (86 kb), pHV3 (442 kb), pHV4 (690 kb), and a small plasmid pHV2 (6.4 kb) (Hartman *et al.*, 2010). The laboratory strain H26 lacks pHV2 and has pHV4 integrated into the main chromosome (Hawkins *et al.*, 2013).

Haloferax volcanii is an ideal model organism to study DNA replication, repair and recombination. There are several genetic tools available for manipulation of the organism, such as auxotrophic selectable markers, reporter genes, promoters for inducible gene expression (Allers and Mevarech, 2005a).

1.2 Replication Initiation

All cells store genetic information in the form of a deoxyribonucleic acid (DNA) molecule, and it is essential to ensure that the genome is accurately replicated once every cell cycle in a timely fashion. DNA replication consists of three broad steps: initiation, elongation and termination. To ensure that all DNA in the genome replicates at an equal rate, cells have developed elaborate mechanisms to ensure the coordination of where and when replication initiation takes place. The first model of replication initiation was proposed by Jacob, Brenner and Cuzin, where they described a *trans*-acting initiator protein binding to a *cis*-acting replicator sequence (Jacob, Brenner and Cuzin, 1963). This is called the replicon model and is mostly accurate for how replication initiates in Bacteria (Figure 1.3). However, Archaea and especially Eukaryotes have more complex mechanisms for replication initiation.

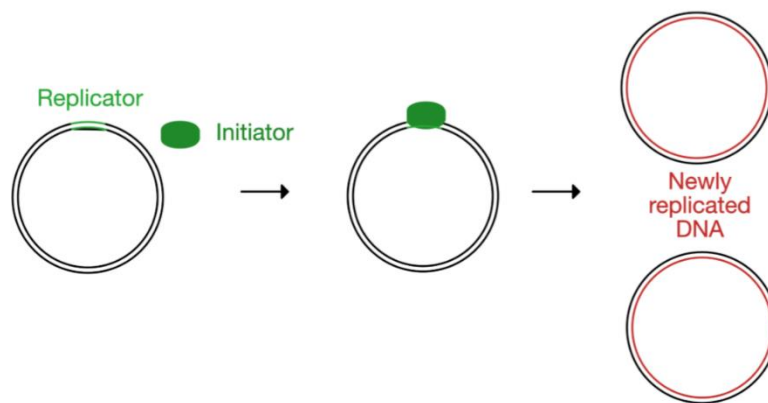


Figure 1.3: **The replisome model.** A replication initiator protein binding to a replicator sequence starts replication.

It is now widely accepted that DNA replication initiates at sites on the genome called origins of replication. Origins are sequences on the genome where DNA unwinding takes place to allow for the loading of the replication machinery and start DNA synthesis. Bacteria, Eukarya and Archaea all encode initiator proteins that seek and bind to these origins, they help to promote the DNA unwinding. Subsequent recruitment of several replisome components establish a bidirectional replication fork that moves away from the origin as it replicates DNA (Leonard and Méchali, 2013). The details of the elongation and termination phases of DNA replication are not of interest for this thesis, however replication initiation in Bacteria, Eukarya and Archaea will be discussed in detail (Figure 1.4).

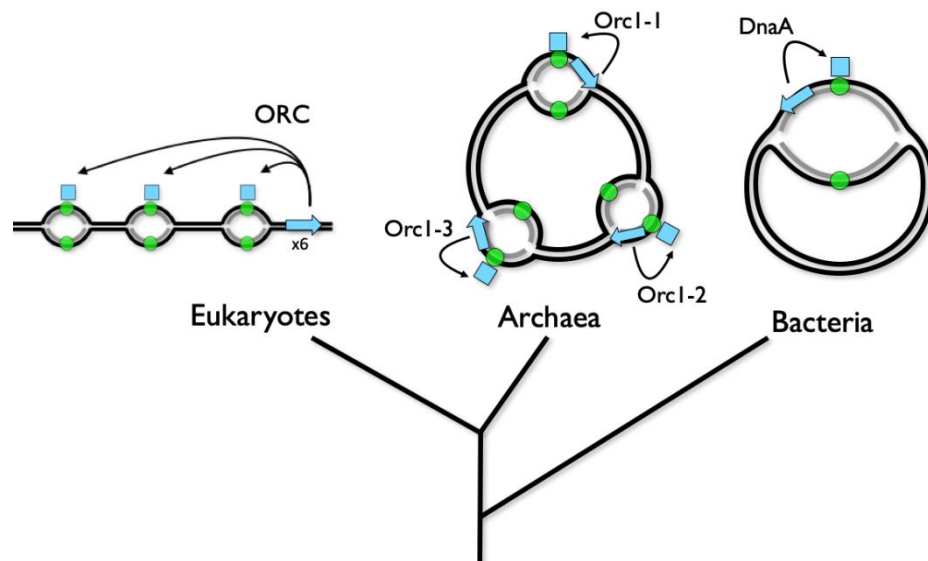


Figure 1.4: **Overview of replication initiation in Eukaryotes, Archaea and Bacteria.** Circular bacterial chromosomes initiate replication when the initiator protein DnaA binds to a single origin. Archaea initiates replication usually from multiple origins when the ORC (Origin Recognition Complex) proteins bind to their respective origins. Eukaryotes possess multiple linear chromosomes with multiple origins. Initiation is dependent on the ORC proteins binding to these origins.

Bacterial Replication Initiation

Typically, Bacteria possess a single circular chromosome. Bacterial DNA replication starts at a single origin (called *oriC* in *E. coli*). The *dnaA* gene that encodes the *E. coli* initiator protein DnaA (AAA+ ATPase) is located adjacent to the *oriC* origin (Hwang and Kornberg, 1992). *oriC* contains an AT-rich DNA unwinding element (DUE). DnaA binds origin recognition elements (or DnaA boxes) located around *oriC* in a sequence-specific fashion with its C-terminal helix-turn-helix motif. This binding bends the DNA molecule at the site of *oriC* by remodelling DUE and results in ‘strand melting’ to give a single stranded ‘bubble’ (Schaper and Messer, 1995). This allows the binding of the DnaC helicase loader, which then recruits the DnaB replicative helicase. The loading of DnaB replicative helicase unwinds DNA and allows the recruitment of other replisome components (primase, DNA polymerases, sliding clamp PCNA and single strand binding protein SSB) to start DNA synthesis (Figure 1.5) (Wegrzyn *et al.*, 2016).

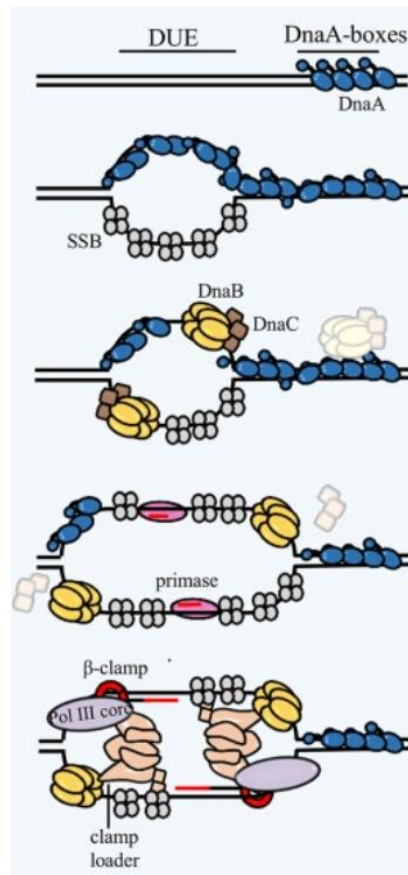


Figure 1.5: **Initiation of DNA replication in *Escherichia coli*.** DnaA binding to *oriC* leads to the sequential loading of the replisome components for DNA replication. Figure from (Wegrzyn *et al.*, 2016).

Eukaryotic Replication Initiation

Replication is much more complex in Eukaryotes than in Bacteria. Eukaryotes have multiple linear chromosomes that each has multiple origins of replication (O'Donnell, Langston and Stillman, 2013). Typically, not all origins fire at the same time in Eukaryotic chromosomes. Some origins fire during early S-phase while others fire later in S-phase. The choice of which origins fire first may be affected by factors such as chromatin structure, transcriptional status, developmental stage, and cell type (Barry and Bell, 2006). Eukaryotes utilise a six protein complex called ORC (OriRigin Recognition Complex) that specifically recognizes and binds to origins of replication with their C-terminal winged helix (WH) domain (Bleichert, Botchan and Berger, 2015). This binding occurs in G1 phase in the cell cycle and recruits the other initiation proteins Cdc6 and Cdt1 during S-phase. This aids the loading of the minichromosome maintenance (MCM) replicative helicase to the site to form the pre-replicative (pre-RC) complex. ORC, Cdc6 and Cdt1 dissociate and MCM is activated by cell cycle

regulators cyclin-dependant kinase (CDK) and Dbf4-dependent kinase (DDK). Activation of MCM recruits further replisome components called GINS and Cdc45. This forms a complex called CMG (Cdc45-MCM-GINS) after which several other replisome components are loaded such as DNA polymerases, primases and PCNA (clamp) proteins. Once the replisome complex is fully established, the replication fork starts to move away from the origin to synthesize DNA (Figure 1.6) (Méchali, 2010).

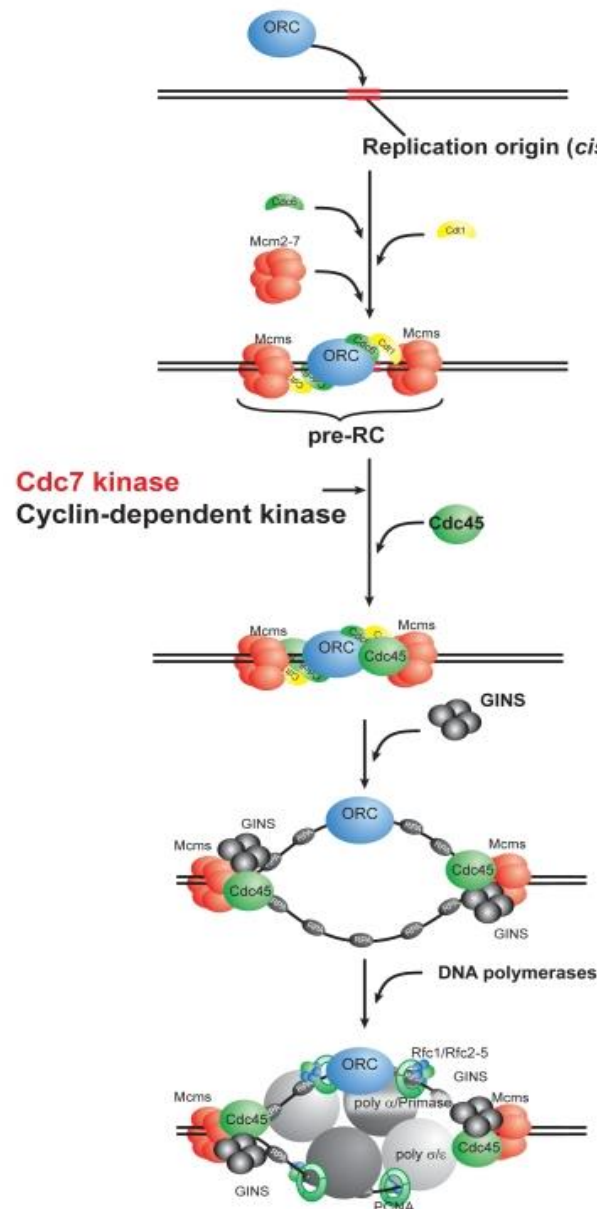


Figure 1.6: **Schematic of eukaryotic replication initiation.** ORC initiator protein binding to the replication origins results in the formation of the replisome complex. Figure from (Sawa and Masai, 2009).

Archaeal Replication Initiation

Archaea use either a single origin or multiple origins of replication depending on the species. Like Bacteria, archaeal origins contain an AT-rich DNA unwinding element (DUE) where binding of the initiator protein occurs in a sequence-specific fashion. However, several replication proteins share sequence homology with their eukaryotic counterparts (Ausani and Allers, 2017).

Orc1/Cdc6 (referred to as Orc or Cdc6 in the literature) is the archaeal initiator protein. The gene encoding Orc1/Cdc6 is usually located adjacent to the origin which it binds. It is most likely the ancestral version of the eukaryotic initiator ORC complex and Cdc6 (Barry and Bell, 2006). Orc1/Cdc6 may have a potential additional role in helicase loading, considering the role of the eukaryotic Cdc6 in loading the MCM helicase.

Archaea possess homologues of the MCM and GINS proteins. The Archaeal RecJ family is thought to be the ortholog of the eukaryotic Cdc45 (Makarova, Koonin and Kelman, 2012). Just like in eukaryotes, CMG complex forms and aids the unwinding of the DNA duplex to allow for the remaining replisome components to load to start DNA synthesis.

Replication origins and initiator proteins in *Haloferax volcanii*

Replication of the multireplicon genome of *Haloferax volcanii* starts at several origins and uses replication proteins that are homologous to that found in eukaryotes (Robinson and Bell, 2005). A range of genome-wide approaches have been used to locate and characterize replication origins, such as chromatin immunoprecipitation, plasmid-based assays, sequencing of Okazaki fragments etc (Müller *et al.*, 2013). Using deep sequencing to map replication origins produces comparable data with high resolution. Flow cytometry is used to sort asynchronous cell cultures into replicating (S phase) and non-replicating (G2 phase) cell populations prior to purification of genomic DNA for sequencing. After deep sequencing, the resulting reads are mapped to reference genomes. Calculating the ratio of the uniquely mapped read counts in the replicating sample to the non-replicating sample generates a relative tag count, which is then normalized to a baseline of 1. To obtain a replication profile, these normalized relative copy numbers are plotted against a chromosomal coordinate. Peaks in the

relative copy number correspond to the locations of active origins. This method is called marker frequency analysis and identifies the locations of active origins in genomes (Müller and Nieduszynski, 2012) (Figure 1.7).

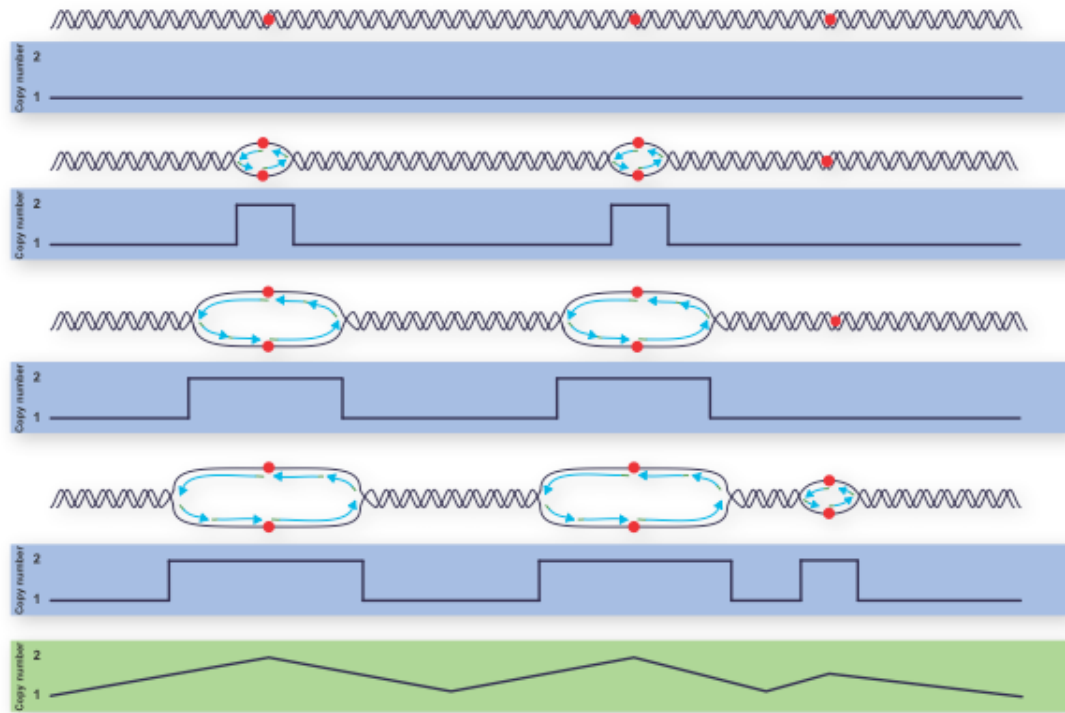


Figure 1.7 **Marker frequency analysis measures copy number differences.** Peaks indicate that corresponding sequences are over-represented in replicating samples compared to non-replicating samples. Deep sequencing detects a higher relative copy number around origins which in turn results in peaks. Adapted from (Müller *et al.*, 2013).

Marker frequency analysis determined that *Haloferax volcanii* has 3 replication origins on its main chromosome (*oriC1*, 2 and 3). Due to the integration of the pHV4 megaplasmid, there is an additional origin called *ori-pHV4* on the chromosome of the laboratory strain. Origins were believed to be essential for cell viability until it was shown that *H. volcanii* strains lacking all origins not only survive but grow faster than wild-type by 7.5% (Hawkins *et al.*, 2013) (Figure 1.8).

RadA (homologue of eukaryotic Rad51, bacterial RecA) is essential for homologous recombination. The inability to delete RadA from this originless strain led to the conclusion that cells lacking origins rely on homologous recombination to initiate replication. This is an alternative replication method called recombination-dependent replication (Hawkins *et al.*, 2013).

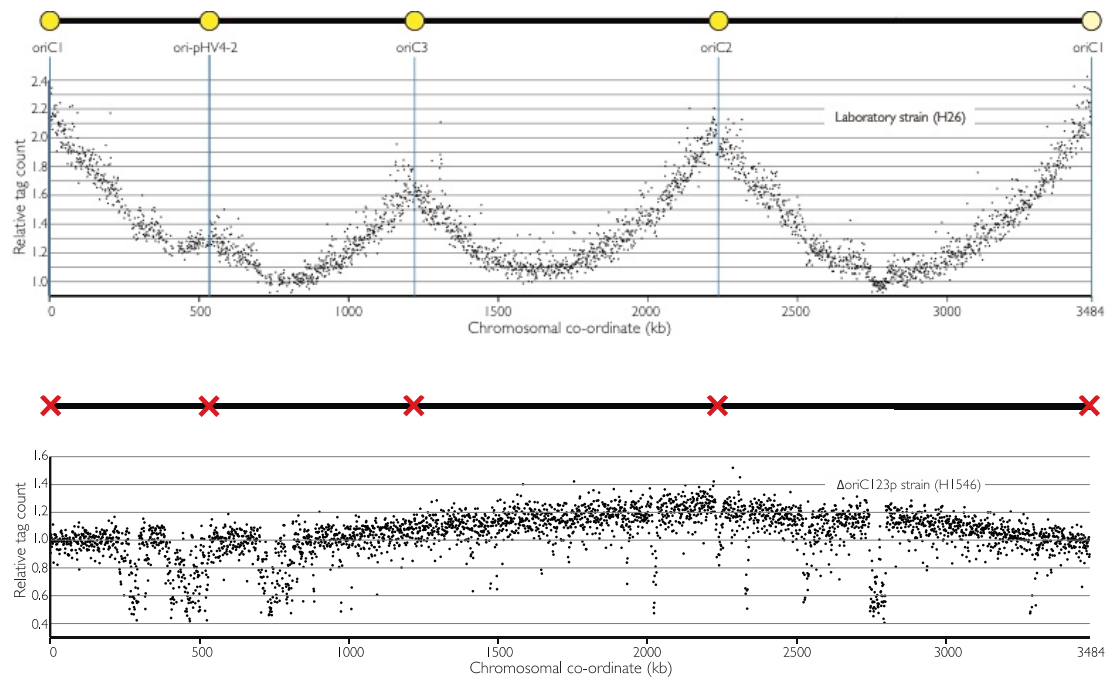


Figure 1.8 Replication profile of the laboratory strain (H26) and strain where all origins are deleted. Marker frequency analysis data obtained by deep sequencing shows the replication profiles of WT and originless strains. Peaks in the relative copy number indicate the location of a replication origin (annotated on figure). Deletion of all replication origins is possible in *H. volcanii* and results in the disappearance of peaks for each origin. From (Hawkins *et al.*, 2013).

H. volcanii encodes 16 Orc1/Cdc6 proteins (Table 1.1). There are 3 native *orc* genes (i.e. not derived from foreign genetic elements) that are not associated with origins. *orc15* and *orc16* are found to be associated with glycerol metabolism genes. *orc9* belongs to a family of Cdc6/Orc1-2 proteins that is highly conserved in all archaeal species, but unlike other *orc1/cdc6* genes, its members are never located adjacent to origins (Figure 1.9) (Raymann *et al.*, 2014). For this reason, it was proposed that Orc9 might have a role in negatively regulating the binding of replication initiator proteins to origins. However, the replication profile of an Δ *orc9* mutant of *H. volcanii* indicates that Orc9 plays no role in origin firing (Figure 1.10, unpublished data). This suggests that Orc9 might have an alternative role to replication initiation.

Table 1.1 **orc genes in *H. volcanii***. Origin-associated genes are in blue, native but not origin-associated genes are in pink, non-native (foreign) and not origin-associated genes are in orange.

Gene	Location	Features
<i>orc1</i>	Chromosome	<i>oriC1</i> , main chromosomal origin
<i>orc2</i>	Chromosome	<i>oriC3</i>
<i>orc3</i>	pHV4	<i>ori-pHV4</i> , in native region of pHV4.
<i>orc4</i>	Chromosome	Associated with non-native protein glycosylation genes
<i>orc5</i>	Chromosome	<i>oriC2</i>
<i>orc6</i>	pHV3	<i>ori-pHV3</i>
<i>orc7</i>	pHV4	Both associated with (non-native) transposases.
<i>orc8</i>	Phv1	
<i>orc9</i>	Chromosome	<u>Not</u> adjacent to origin. Potential role in DNA repair.
<i>orc10</i>	pHV1	<i>ori-pHV1</i> . Dominant in origin screen, difficult to delete.
<i>orc11</i>	Chromosome	Both associated with native glycerol metabolism genes.
<i>orc12</i>	Chromosome	
<i>orc13</i>	pHV4	Both associated with non-native DNA metabolism genes.
<i>orc14</i>	pHV4	
<i>orc15</i>	Chromosome	Both associated with native glycerol metabolism genes.
<i>orc16</i>	Chromosome	

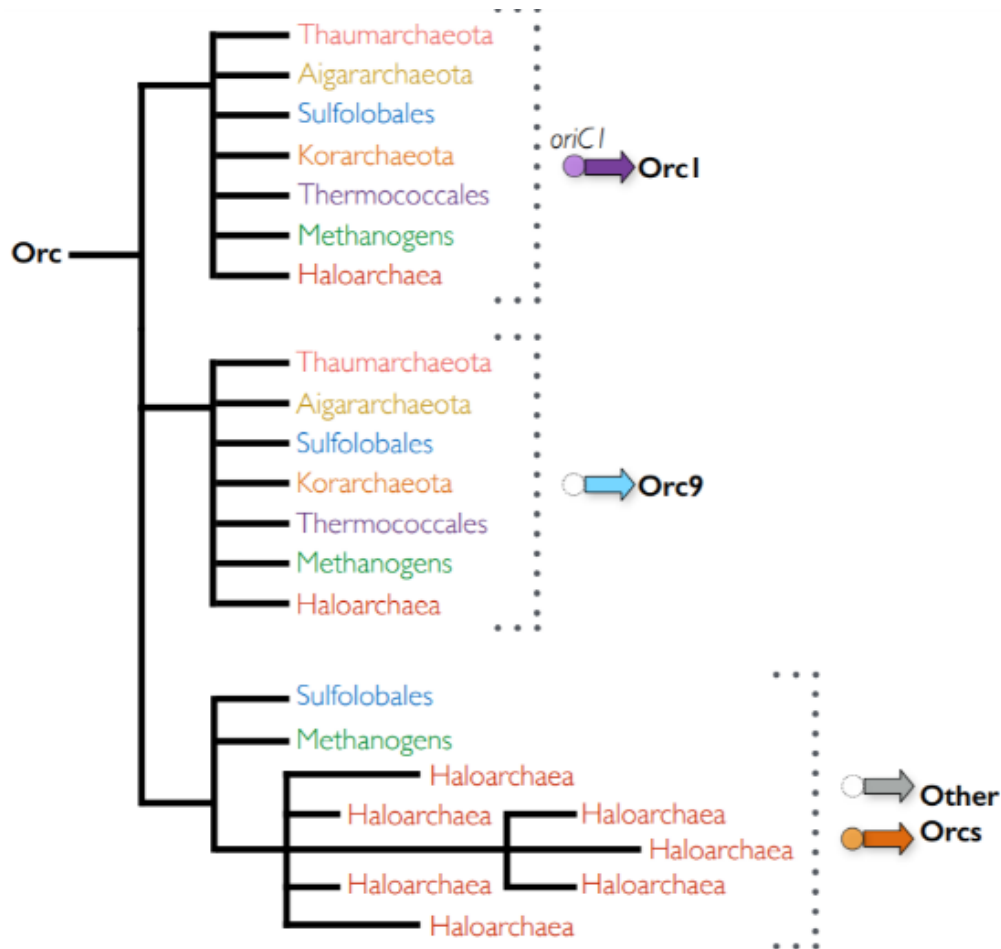


Figure 1.9 **Orc phylogeny in Archaea.** Orc1 (Cdc6-1) and Orc9 (Cdc6-2) are highly conserved in most Archaeal species. Orc1 is always found next to *oriC1* whereas Orc9 is never found next to a replication origin. Adapted from (Raymann *et al.*, 2014).

Role of Orc9 in *Haloferax volcanii*

orc9 is the only highly conserved *orc* gene that is not adjacent to a replication origin. Its conservation in all archaeal species suggests that Orc9 must play an important role. However, it is not essential since a *H. volcanii* strain deleted for *orc9* (Δ *orc9*) is viable and competition assays have shown that the wild-type strain grows only 4.8% faster than Δ *orc9* strain (unpublished data). If Orc9 played a role in replication, this would have reflected in the replication profile of Δ *orc9*. Deleting *orc9* does not have any impact on the replication profile of *H. volcanii*, which suggests that Orc9 does not have a role in replication (Figure 1.10, unpublished data).

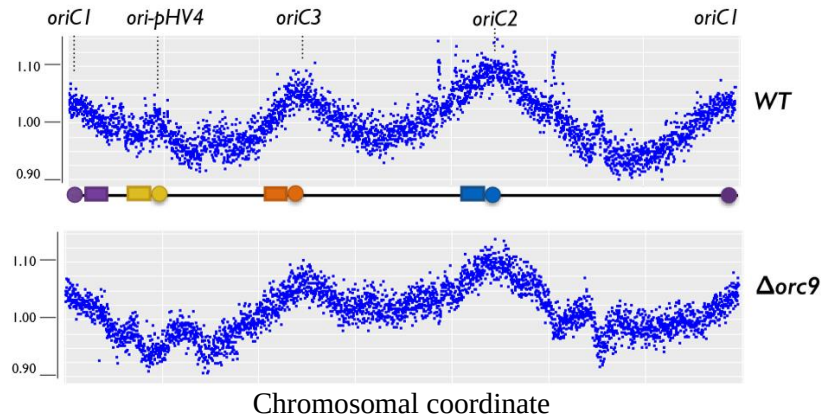


Figure 1.10: **Replication profiles of wild-type and $\Delta orc9$ strains.** Peaks in the relative copy number indicates the location of a replication origin (annotated on figure). Deletion of *orc9* has no impact on the replication profile of *H. volcanii*.

Preliminary data suggests that the $\Delta orc9$ strain exhibits sensitivity to DNA damaging agents: UV radiation, mitomycin C (MMC) and 4-nitroquinoline 1-oxide (4-NQO) (unpublished data). This implies a role for Orc9 in DNA repair. The $\Delta orc9$ strain was examined for genome arrangements after 1.65 $\mu\text{g/ml}$ MMC treatment (Figure 1.11). Pulsed field gel electrophoresis data suggested that Orc9 is involved in maintaining genomic stability.

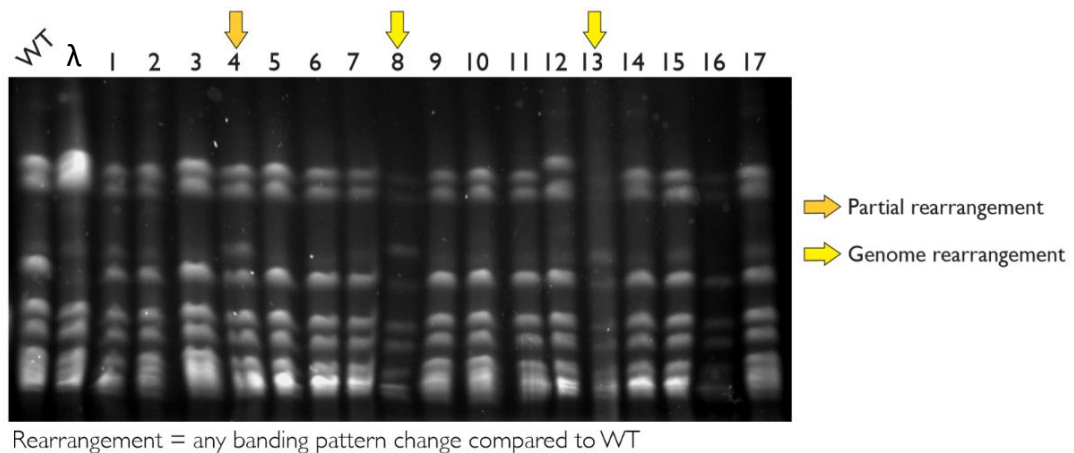


Figure 1.11: **Pulsed field gel electrophoresis of $\Delta orc9$ clones.** MMC-treated $\Delta orc9$ gave a genomic rearrangement rate of 14%. Lambda marker ladder (λ) can be seen on the second well from the left. Band sizes are not labelled since we are only interested in seeing if there is any change in banding pattern compared to the WT in samples 1-17. Data from Victoria Smith.

1.3 DNA Repair

DNA is an indispensable molecule for all cells and need to be accurately repaired when damaged to restore genetic information. All other cellular molecules can be degraded and made from scratch if required. For example, misfolded proteins are ubiquitinated and subsequently degraded by proteases. In contrast, DNA solely relies on repair mechanisms that are highly conserved in all domains of life.

DNA Damage

Sources of DNA damage can be endogenous or exogenous. Endogenous damage can arise from normal metabolic processes in cells such as the generation of reactive oxygen species during electron transport chain reactions or misincorporation of incorrect bases by DNA polymerases during replication. Exogenous damage arises from external sources such as UV radiation or chemical mutagens that are highly toxic to cells. Examples of chemical mutagens and the specific DNA lesions they produce are described in detail in *Chapter 3: DNA damage sensitivity tests*, where relevant. If unrepaired, both endogenous and exogenous DNA damage can lead to cell death. Figure 1.12 summarizes the causes of DNA damage, their consequences and corresponding repair pathways.

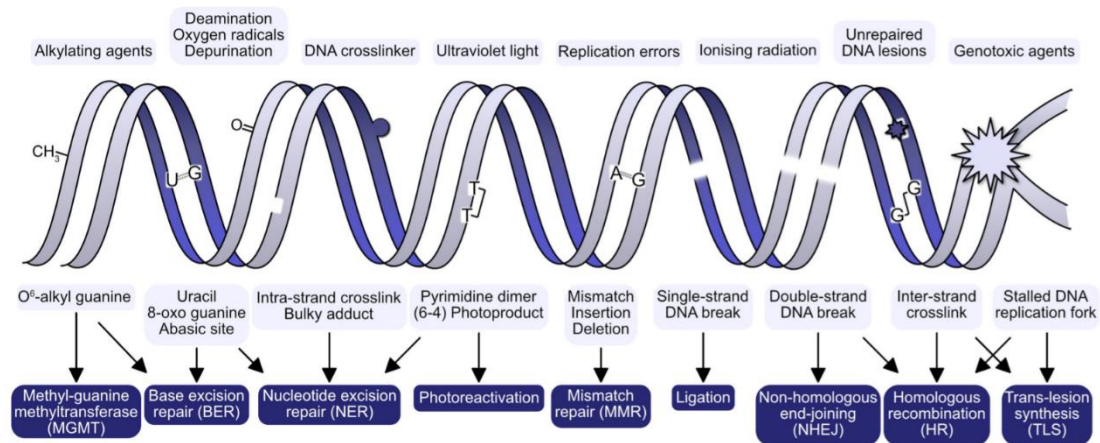


Figure 1.12: **Sources of DNA damage, resulting lesions and repair pathways.** Details of these pathways can be found later in this thesis. From (White and Allers, 2018).

DNA Repair Pathways

DNA damage is a threat to cell viability and multitude of DNA repair pathways have evolved to make sure DNA remains intact for the next round of replication.

Direct repair

There are only 3 well studied examples of direct repair:

- DNA ligation: Single-strand breaks (SSBs) can arise due to UV radiation or the incomplete ligation of Okazaki fragments in the lagging strand during replication. DNA ligases can re-ligate single-strand breaks and seal the gap. However, DNA ligases can only catalyse this reaction if the SSB has 'clean' (3' OH and a 5' phosphate) ends.
- Photoreactivation: Photolyase enzymes can directly reverse UV-induced cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-4PPs) by a process called photoreactivation. Photolyases break the covalent bond(s) between adjacent pyrimidine bases by transferring electron energy from chromophores. This process is light-dependent as photolyases require light to be activated (Essen and Klar, 2006).
- Reversal of methyl damage: Methyltransferases can directly remove methyl groups from DNA bases. Alkylating agents such as methyl methanesulfonate (MMS) adds methyl groups to bases in DNA and generates O⁶-methyl guanine as a minor adduct. Ada/MGMT directly removes the methyl group of O⁶-methylguanine (Sedgwick, 2004).

Excision Repair

DNA damage that is localized only on a single strand of DNA is generally repaired by one of the three excision repair pathways: Base excision repair (BER) removes a single damaged base and replaces it, nucleotide excision repair (NER) removes a short patch of nucleotides and replaces it, and mismatch repair (MMR) is interlinked with DNA replication and removes mis-incorporated bases in DNA after replication. Figure 1.13 compares BER and NER.

Base excision repair

Base excision repair is involved in the repair of the most diverse types of DNA lesions. It is a multistep process that involves several enzymes to remove damaged bases in DNA and incorporate correct nucleotides to restore the original sequence information. The core components of BER is highly conserved among eukaryotes, bacteria and archaea. The process of base excision repair is essentially the same in all domains of life and can be broken down to 3 broad stages: Base removal by DNA glycosylases,

generation of a single strand break by AP endonucleases or AP lyases, and gap filling by DNA polymerases and DNA ligases.

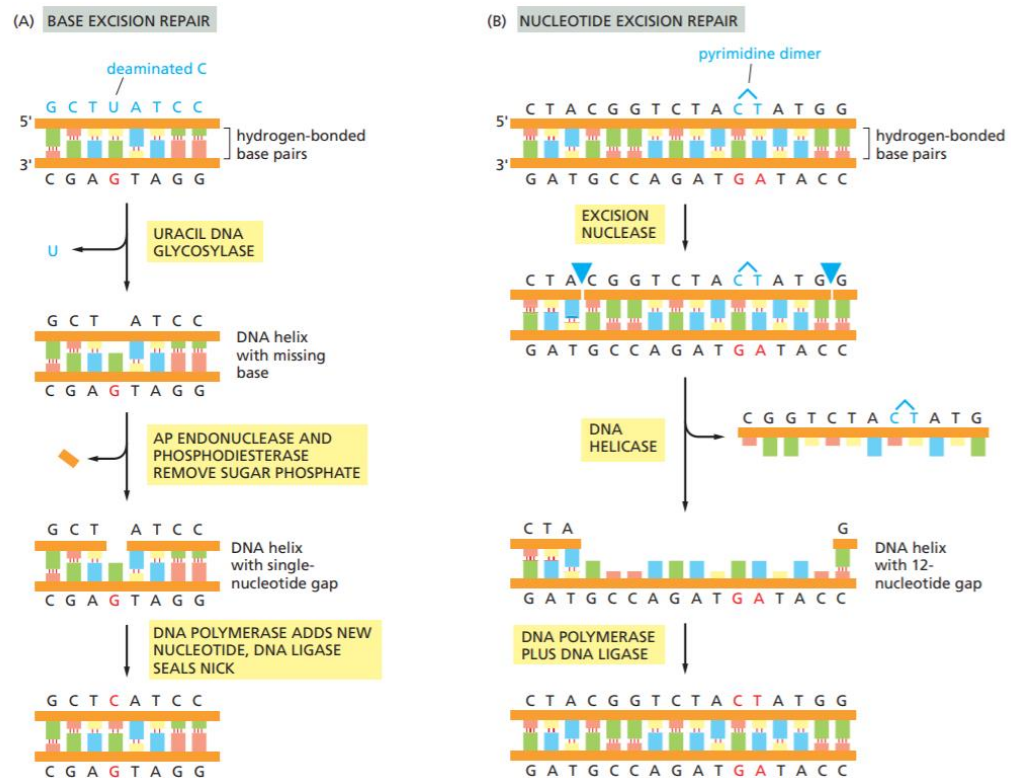


Figure 1.13: **A comparison of BER and NER.** From (Alberts *et al.*, 2014).

Damage recognition is the first step in BER and is carried out by DNA glycosylases. DNA glycosylases are lesion-specific, which means each enzyme recognises and removes a specific type of base damage (Sartori and Jiricny, 2003). For instance, uracil that arises in DNA due to spontaneous cytosine deamination is recognized and removed by uracil glycosylases (Savva *et al.*, 1995). Methylated adenine in DNA is recognized and removed by 3-methyladenine-DNA glycosylases (Karran *et al.*, 1980). Generally, DNA glycosylases flip the damaged base out of the DNA double helix and cleave the N-glycosidic bond between the base and the deoxyribose sugar in the DNA backbone. This generates what is called an abasic (apurinic/apyrimidinic (AP)) site. AP endonucleases are a set of enzymes that specifically recognize the AP sites in duplex DNA. They generate an incision in DNA by catalysing the cleavage of the phosphodiester backbone 5' to the AP site. This results in a single strand break with 3'OH and 5' deoxyribosephosphate (5'dRP) ends (Doetsch and Cunningham, 1990). Alternatively, some DNA glycosylases are bifunctional and possess an additional AP lyase activity. The AP lyase activity processes the AP site by cleaving the DNA

backbone 3' to the AP site. Subsequent β -elimination reaction results in a 3'- α , β -unsaturated aldehydic (3'-dRP) end which needs to be removed by an AP endonuclease. 3'-OH is an appropriate terminus for priming DNA synthesis, however the 5'-dRP end needs to be removed for the gap filling and subsequent ligation to take place. Different enzymes and strategies are used by bacteria and eukaryotes for the last stages of BER (Wallace, 2014).

The main polymerase used in BER by mammalian cells is DNA polymerase β (Pol β). Pol β removes 5'-dRPs since it possesses a dRPase activity. 3'-OH terminus of the AP site primes repair synthesis by Pol β . "Short-patch" BER takes place when Pol β inserts the single missing base and fills the gap. "Long-patch" BER takes place when Pol β inserts 2-8 nucleotides beyond the AP site by strand displacement synthesis. This results in a 5' nucleotide flap that is removed by the FEN1 (flap endonuclease 1) protein in eukaryotes and in archaea. FEN1 has a 5' to 3' exonuclease activity and can cleave DNA "flaps" at the junction with a duplex DNA. In *S. cerevisiae*, Rad27 is a homologue of FEN1 and cleaves 5'dRP ends (Liu *et al.*, 2005).

Whether the 5'-dRP intermediate is processed efficiently determines the pathway Pol β will carry out (short or long patch BER pathway completion). The type of lesion (and the choice of DNA glycosylase) also has an influence on the choice of which pathway is used. For example, repair of the 8-oxoG lesion only results in BER completion via the short-patch route. Both the short and the long patch BER pathways finalize when DNA ligases catalyse the joining of 3'-OH and the 5'-phosphate termini and seals the nick on the DNA backbone (Wallace, 2014).

Nucleotide Excision Repair

Nucleotide excision repair is the pathway that specializes in removing 'bulky' lesions (such as UV-induced pyrimidine dimers) from the DNA backbone. In contrast to BER, NER removes a short patch of nucleotides around the site of the DNA lesion, as opposed to a single base. It is made up of 3 broad steps: recognition DNA damage, incision and excision of the damaged DNA and gap filling. There are 3 well-studied human diseases associated with defects in the NER pathway: Xeroderma pigmentosum (XP), Cockayne Syndrome (CS) and photosensitive form of trichothiodystrophy

(TDD). All of these diseases have the symptom of extreme sunlight sensitivity and in the case of XP, a disposition to cancer (de Boer and Hoeijmakers, 2000).

The UvrA, UvrB, UvrC and UvrD proteins are responsible for NER in Bacteria. UvrA and UvrB forms a heterodimer UvrA₂B and recognizes damage as distortions to the DNA backbone. It binds to the site of damage and unwinds DNA. UvrA₂ dissociates but UvrB remains on the site of damage and recruits the UvrC nuclease. UvrC performs an incision on both sites of the damage. UvrD releases UvrC from the site and removes the incised oligonucleotide (12-13 nucleotides). DNA polymerases fill the gap and DNA ligases seal the DNA backbone. There is an alternative damage recognition mechanism named transcription-coupled repair (TCR). RNAP (RNA polymerase) stalls during transcription if there is a lesion blocking its path. This stalling acts as a damage recognition signal and a protein called Mfd aids the recruitment of UvrA to the site of damage (Marteiijn *et al.*, 2014).

Eukaryotic NER is similar to bacterial NER but more complex. There are two NER pathways: transcription-coupled repair (TCR) and global genome repair (GGR). They only differ in their method of recognising DNA damage (Figure 1.14). Just like in Bacteria, TCR in eukaryotes detects DNA damage through RNAP stalling. GGR utilizes XPC-Rad23B or DDB1-DDB2 to bind to DNA helix-distortions and detect DNA damage. After recognition of damage, both TCR and GGR recruit transcription factor IIH (TFIIH) to the site and DNA unwinding starts in opposite directions by the combined action of XPB and XPD helicases. Incision 5' to the damaged site is performed by a nuclease called XPF-ERCC1. Incision 3' to the site is performed by a nuclease called XPG. DNA polymerases δ and ϵ along with other helper proteins fill the gap and finalize the repair process (Schärer, 2013). Figure summarizes NER in eukaryotes.

Archaea have homologues of the eukaryotic XBF, XPB and XPD. Some archaea like *H. volcanii* possesses homologues of the bacterial genes *uvrABC* as well and they appear to be functional in NER (Lestini *et al.*, 2011).

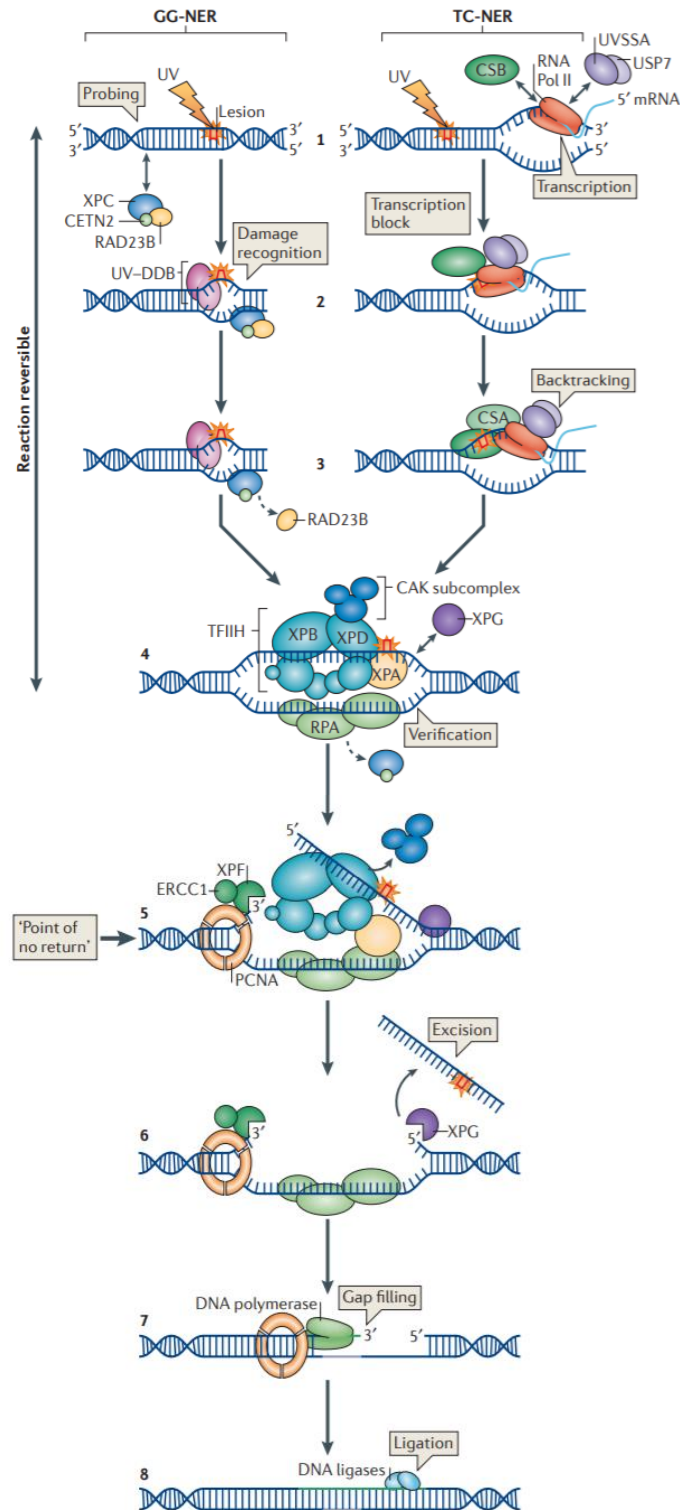


Figure 1.14: **Summary of eukaryotic NER.** The two sub-pathways (GG-NER and TC-NER) differ in their ways of detecting DNA damage. In GG-NER, XPC in complex with RAD23B constantly scans DNA for helix-distorting lesions. In TC-NER, damage is indirectly detected by the stalling of RNA Polymerase II (RNA Pol II) at a lesion. From (Martelijn *et al.*, 2014).

Mismatch Repair

Non-complementary bases can be mis-incorporated into newly synthesized DNA during replication by DNA polymerases. These mismatches can lead to mutations if they are not repaired before the next round of replication. Mismatches are recognised through distortion of the DNA backbone and repaired by a specialized excision repair pathway called Mismatch Repair Pathway (MMR). Components of MMR are highly conserved across all domains of life. Defects in MMR are associated with a familial type of colon cancer called hereditary nonpolyposis colorectal cancer (HNPCC).

Distinguishing the newly synthesized daughter strand from the template parent strand is an essential part of the MMR repair process. In *E. coli*, Dam methylase adds methyl groups to adenines in a sequence specific fashion (5'-GATC-3'). For a short period of time after DNA synthesis, the newly synthesized daughter strand remains unmethylated (Geier and Modrich, 1979). This differential post-replicative methylation status of the two strands in *E. coli* DNA aids the differentiation of the two strands to ensure the repair of only the daughter strand by MMR. In eukaryotes, the nicks in the newly synthesized DNA immediately after replication (which arise from nicks between Okazaki fragments in the lagging strand) are thought to aid this differentiation (Kunkel and Erie, 2015).

MutL-MutS are the core proteins involved in the MMR pathway and have homologs in all bacteria and eukaryotes (Figure 1.15). However, only some archaea encode homologs of MutS-MutL proteins. *Haloferax volcanii* has 4 MutS and 2 MutL homologs. Having multiple MutS-MutL homologs is a common pattern seen in eukaryotes where some MutS homologs are not involved in repair (Kelman and White, 2005). In *H. volcanii*, two of the MutS homologs belong to the MutS1 family which are involved in repair and the other two MutS homologs are poorly characterized and have no function in repair (Hartman *et al.*, 2010). Archaea lacking these canonical MutS-MutL homologs do exist and do not seem to show elevated spontaneous mutation rates, which suggests an alternative pathway (Grogan, 2004), (Busch and DiRuggiero, 2010). NucS (or EndoMS for endonuclease mismatch specific) is a recently discovered endonuclease that cleaves the DNA backbone at mismatched sites (Ishino *et al.*, 2016). Components of this novel NucS-dependent MMR pathway remains to be further identified.

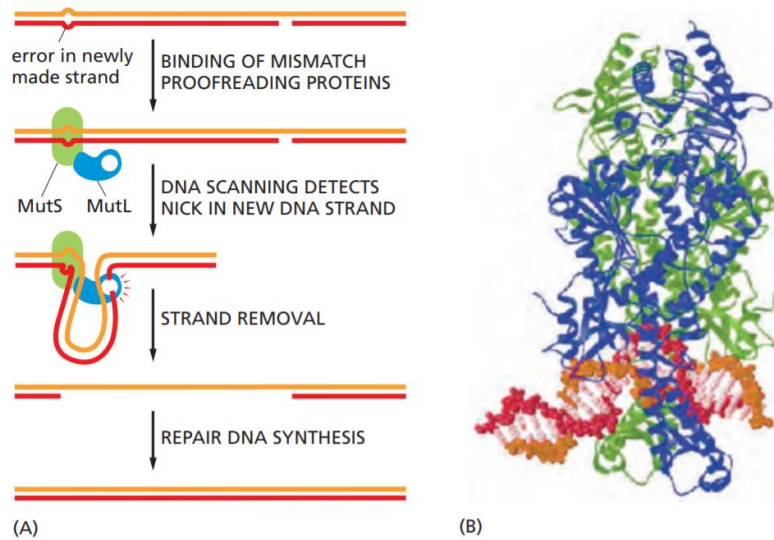


Figure 1.15: **Strand-directed mismatch repair.** (A) MutS scans the DNA for mismatches. MutL scans nearby DNA for nicks (indicative of newly synthesized DNA) and when found, triggers degradation of the nicked strand all the way to the mismatch. (B) Structure of a MutS protein bound to DNA. Notice how MutS bends DNA helix to produce a kink at the mismatched base. From (Alberts *et al.*, 2014).

Double Strand Break (DSB) Repair

Double strand breaks (DSBs) are the most cytotoxic lesions cells can encounter. Ionising radiation or DNA-cleaving mutagens such as phleomycin can induce DSBs exogenously. Another source of DSBs is collapsed replication forks. Since genetic information on both strands is missing, cells need to accurately repair both strands. Although the incidence of DSBs in nature is rare, it is extremely detrimental since it may result in genomic arrangements and cell death. Cells have developed two major mechanisms to ensure the repair of this type of damage: homologous recombination (HR) and non-homologous end joining (NHEJ). NHEJ involves processing and direct ligation of the two ends of the DSB so it is deleterious and potentially mutagenic. NHEJ will not be discussed in detail as it is mostly absent in bacteria and archaea (Brissett and Doherty, 2009) (Bowater and Doherty, 2006). HR is the most accurate method to repair a DSB because it uses an identical/homologous chromosome as a template to repair the DSB.

Homologous Recombination (HR)

Homologous recombination refers to the exchange of genetic information between homologous (identical) chromosomes. HR is the main pathway used to repair DSBs. HR also plays an important role in restarting collapsed or stalled replication forks. HR is utilized by eukaryotes to increase genetic diversity during meiosis.

Homologous recombination can be broken down to three broad stages: pre-synapsis, synapsis and post-synapsis (Figure 1.16). Pre-synapsis starts with a double-strand break (DSB) and involves the resection of the blunt dsDNA end by certain enzymes to generate a 3' overhang. In *E. coli*, the RecBCD complex (or Mre11/Rad50 in eukaryotes and archaea) generates this 3' overhang. Single stranded DNA that is produced by this end resection gets coated by single strand binding proteins (SSB/RPA). RecBCD or RecFOR in Bacteria load RecA onto the ssDNA as the SSB proteins are displaced. The RecA (eukaryotic Rad51/Dmc1, archaeal RadA) recombinase protein initiates the synapsis stage by facilitating strand exchange and subsequent D-loop formation which is the central reaction of homologous recombination. Therefore, deletion of recombinases gives rise to a complete inability to carry out homologous recombination and sensitivity to various DNA damaging agents. The loading of RecA leads to the 3' overhang base pairing with its homologous counterpart. This primes DNA synthesis by DNA polymerases and D-loop formation. Extension of the D-loop is carried out by branch migration which then forms another Holliday junction which is a 4-way DNA junction formed during synapsis). RuvABC in *E. coli* carries out both branch migration and the resolution of the Holliday junction in the post synapsis stage. RuvA binds to the junction and combined with the action of RuvB promotes branch migration. Holliday junctions are resolved at the post-synapsis stage in 2 different orientations. Depending on which orientation the cut is performed, either a crossover (CO) or a non-crossover product is produced. RuvC cleaves the Holliday junction in a symmetrical fashion. RecG is an alternative protein that carries out branch migration. In the absence of RuvC, RusA cleaves the Holliday junctions (Wright, Shah and Heyer, 2018).

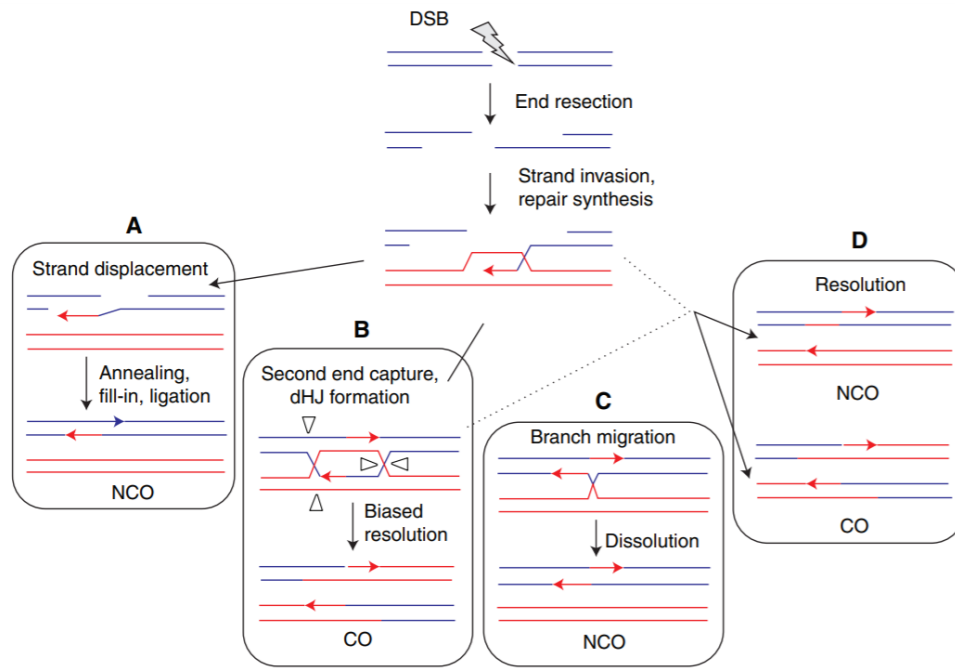


Figure 1.16 **Homologous recombination.** HR starts with a DSB which is followed by end resection leaving overhang ends which invade the corresponding sequence on the homologous chromosome. This is facilitated by essential recombinase proteins. Differential resolution of the double Holliday junction (dHJ) leads to different HR products. From (Jasin and Rothstein, 2013).

DNA Damage Response (DDR)

Repair of complex lesions in DNA require multiple repair proteins working together to restore genetic information. In Bacteria, Eukarya and Archaea, regulatory networks have evolved to produce coordinated DNA damage responses to various mutagens that cells get exposed to. The best example of a DNA damage response (DDR) in *E. coli* is the SOS response. A homologous recombination intermediate, ssDNA-RecA filaments, triggers this response by stimulating the cleavage of a transcriptional repressor called LexA. Subsequent induction of several genes (around 40) gives rise to a coordinated DNA damage response. The SOS response is widespread in many bacterial species (Butala, Žgur-Bertok and Busby, 2009).

Eukaryotes have more complex surveillance mechanisms for DDR. Three protein kinases are known to be involved in eukaryotic DDR: ataxia telangiectasia mutated (ATM), ATM and Rad3-related (ATR) and DNA-PKcs (DNA-dependent protein kinase). They sense DNA damage through direct protein-protein interactions and give

rise to a cascade of phosphorylation events that modify several repair enzymes to improve their efficiencies. They also induce changes in chromatin structure near the DNA lesion that increases the accessibility of DNA damage to repair enzymes (Sirbu and Cortez, 2013) (Blackford and Jackson, 2017).

Archaea do not possess any homologs of the bacterial or eukaryotic DDR response genes. This raises the question of whether Archaea have any DDR networks. UV radiation has been shown to induce a differential expression profile in *Sulfolobus solfataricus* (Salerno *et al.*, 2003) and *Sulfolobus acidocaldarius* (Götz *et al.*, 2007). Recently, the Orc9 homolog in *Sulpholobus islandicus* (*cdc6-2*) was studied for its potential role in DDR. The authors used RNA-sequencing to investigate the 4-nitroquinoline 1-oxide (4-NQO) induced DNA damage response and its regulation, and whether it is mediated by Cdc6-2. Their results showed that some repair genes were upregulated, while some cell cycle regulation and DNA replication genes were downregulated in a Cdc6-2-dependent fashion. They concluded that Orc9 (Cdc6-2) might have evolved to gain a function in regulating a global DNA damage response in *S. islandicus* (Sun *et al.*, 2018).

1.4 Aims and Objectives

Orc9 may have a role as a key regulator of DNA damage response in *Haloferax volcanii*. The aims of this study are:

- To test the sensitivity of the $\Delta orc9$ strain to DNA damaging agents such as methyl methane-sulphonate (MMS), hydrogen peroxide (H_2O_2), gamma radiation, phleomycin and UV radiation.
- To perform growth assays to test sensitivity of wild-type and $\Delta orc9$ strains to mitomycin C and H_2O_2 .
- To construct tryptophan-inducible alleles of the *orc9* gene and perform growth assays to determine whether the level of *orc9* transcript directly correlates with survival.
- To explore the synergistic genetic interaction of Orc9 and the recombinase RadA using strains where *radA* is under the control of a tryptophan inducible promoter.
- To tag Orc9 with epitope tags and express in *H. volcanii* under the control of its native promoter. Protein pulldown experiments will elucidate the protein interactors of Orc9.
- To perform RNA-sequencing on samples of wild-type and $\Delta orc9$ strains in the presence (and absence) of MMC to elucidate the differential gene expression profile and the possible downstream targets of Orc9.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Strains

Haloferax volcanii strains

Table 2.1 *Haloferax volcanii* strains used in this study

Strain	Parent	Genotype	Notes
H26	H18	Δ pyrE2	Constructed by TA
H53	H26	Δ pyrE2, Δ trpA	Constructed by TA
H2132	H2062	Δ pyrE2, Δ orc9	Constructed by DA
H2728	H2665	Δ pyrE2, Δ orc9, Δ orc3, Δ ori-pHV4-2, Δ orc2, Δ oriC3, Δ orc1, Δ oriC1, Δ orc5, Δ oriC2	Constructed by DA
H2655	H53	Δ pyrE2, Δ trpA,, Δ orc9	Constructed by DA
H2656	H2622	Δ pyrE2, Δ orc3, Δ ori-pHV4-2, Δ orc2, Δ oriC3, Δ orc1, Δ oriC1, Δ orc5, Δ oriC2	Constructed by DA
H4054	H26	Δ pyrE2, orc9+::[p.tnaM3-orc9+ pyrE2+]	Constructed by VS
H4055	H938	Δ fen1, Δ pyrE2, orc9+::[p.tnaM3-orc9+ pyrE2+]	Constructed by VS
H4056	H2656	Δ pyrE2, Δ orc3, Δ ori-pHV4-2, Δ orc2, Δ oriC3, Δ orc1, Δ oriC1, Δ orc5, Δ oriC2, orc9+::[p.tnaM3-orc9+ pyrE2+]	Constructed by VS
H4045	H53	Δ pyrE2, Δ trpA, Δ mrr	Constructed by VS
H4589	H4045	Δ pyrE2, Δ trpA, Δ mrr, orc9+::[7xHis 2xStrepII-orc9 pyrE2+]	Constructed by ÇT
H4414	H4054	Δ pyrE2, p.tnaM3-orc9+	Constructed by ÇT
H4419	H4056	Δ pyrE2, Δ orc3, Δ ori-pHV4-2, Δ orc2, Δ oriC3, Δ orc1, Δ oriC1, Δ orc5, Δ oriC2, p.tnaM3-orc9+	Constructed by ÇT
H1637	H1620	Δ pyrE2, Δ hdrB p.tnaA-radA+::hdrB+	Constructed by MH
H1642	H1622	Δ pyrE2, p.tnaA-radA+::hdrB+, Δ oriC2, Δ oriC1, Δ oriC3, Δ hdrB, Δ ori-pHV4-2,	Constructed by MH
H3414	H3359	Δ pyrE2, Δ orc9, p.tnaA-radA+	Constructed by DH

TA = Thorsten Allers, DA = Dasha Ausiannikava, VS = Victoria Smith, LM = Laura Mitchell, MH = Michelle Hawkins, DH = Daisy Haigh, ÇT = Çağla Tosun. [] indicated integrated plasmid.

Escherichia coli strains

Table 2.2 *Escherichia coli* strains used in this study

Strain	Genotype	Notes
XL-1 Blue MRF'	<i>endA1</i> , <i>gyrA96</i> (NalR), <i>lac</i> [F' <i>proAB lacIqZΔM15 tn10</i> (TetR)], <i>Δ(mcrA)183</i> , <i>Δ(mcrCBhsdSMR-mrr)173</i> , <i>recA1</i> , <i>relA1</i> , <i>supE44</i> , <i>thi-1</i>	Standard cloning strain for blue/white screening using pBluescript based plasmids. Tetracycline resistant. Restriction endonuclease and recombination deficient. <i>dam</i> ⁺ . From Stratagene.
N2338 (GM121)	<i>F</i> -, <i>ara-14</i> , <i>dam-3</i> , <i>dcm-6</i> , <i>fhuA31</i> , <i>galK2</i> , <i>galT22</i> , <i>hsdR3</i> , <i>lacY1</i> , <i>leu-6</i> , <i>thi-1</i> , <i>thr-1</i> , <i>tsx-78</i>	<i>dam</i> - <i>dcm</i> - mutant for preparing DNA for <i>Haloferax volcanii</i> transformations (Allers et al. 2004). From RG Lloyd

2.1.2 Plasmids

Table 2.3 Plasmids used in this study

Name	Use	Notes
pTA1432	Genomic clone of <i>orc9</i>	Constructed by TA
pTA1705	Contains upstream genomic sequence of <i>orc9</i> , used to construct pTA2292	Constructed by LM
pTA1771	For the expression of N-terminally tandem 7xHis 2xStrepII tagged genes under the control of their native chromosomal promoter	Constructed by TA (Wardell et al., 2017)
pTA2261	First step in constructing a tagged <i>orc9</i> gene replacement construct	Constructed by ÇT
pTA2292	N-terminally tandem 7xHis 2xStrepII tagged <i>orc9</i> gene replacement construct	Constructed by ÇT

TA = Thorsten Allers, LM = Laura Mitchell, ÇT = Çağla Tosun.

2.1.3 Oligonucleotides

Table 2.4 Oligonucleotides used in this study

Name	Sequence (5' to 3')	Use	Notes
orc9BspF	CTTTACT ^{tc} ATGAGGGAGGAC CCCGAAG	Used to construct pTA2292	Made by TA
orc9EcoR	GAGTAA ^g AAT ^{Tc} TCCTGAAAC GGGGCGG	Used to construct pTA2292	Made by TA
orc9extF	GGCTCAACTCCTGATGGGCTG TC	For colony PCR to check inducible <i>orc9</i>	Made by TA
orc9R	GGACGTACTTCGACTTGTCGT AGGC	For colony PCR to check inducible <i>orc9</i>	Made by TA

Restriction sites engineered into the oligonucleotide is marked in red. Mismatches against the template are marked as lower-case letters. TA = Thorsten Allers.

2.1.4 Chemicals and Enzymes

All enzymes were purchased from New England BioLabs (NEB) and all chemicals from Sigma, unless otherwise stated. Appropriate methods give details of specific chemicals, buffers and solutions.

2.1.5 Media

Haloferax Media

Media are sterilised by autoclaving for 1 minute at 121°C. Liquid media are stored at room temperature in the dark and solid media in the form of plates are stored at 4°C in sealed bags as to prevent desiccation. All plates were dried for at least 30 mins before use.

30% saltwater (SW): 4 M NaCl, 148 mM MgCl₂·6H₂O, 122 mM MgSO₄·7H₂O, 94 mM KCl, 20 mM Tris.HCl pH7.5.

18% saltwater (SW): Made with 30% SW, 3 mM CaCl₂. CaCl₂ added after autoclaving.

Trace elements: 1.82 mM MnCl₂·4H₂O, 1.53 mM ZnSO₄·7H₂O, 8.3 mM FeSO₄·7H₂O, 200 µM CuSO₄·5H₂O. Filter sterilised and stored at 4°C.

Hv-Min Salts: 0.4 M NH₄Cl, 0.25 M CaCl₂, 8% v/v of trace element solution. Stored at 4°C.

Hv-Min carbon source: 10% DL-lactic acid Na₂ salt, 8% succinic acid Na₂ salt·6H₂O, 2% glycerol, pH to 7.0 with NaOH. Filter sterilised.

10 x YPC: 5% yeast extract (Difco), 1% peptone (Oxoid), 1% casamino acids, 17.6 mM KOH. Not autoclaved, used immediately.

10 x Ca: 5% casamino acids, 17.6 mM KOH. Not autoclaved, used immediately.

Hv-Ca salts: 362 mM CaCl₂, 8.3% v/v of trace elements, 615 µg/ml thiamine 77 µg/ml biotin.

KPO₄ Buffer: 308 mM K₂HPO₄, 192 mM KH₂PO₄ pH7.0.

Hv-YPC agar: 1.6% agar (Bacto), 18% SW, 1 x YPC, 3mM CaCl₂. Microwaved without 10 x YPC to dissolve agar. 10 x YPC added, then autoclaved. CaCl₂ added prior to pouring.

Hv-Ca agar: 1.6% agar (Bacto), 18% SW, 1 x Ca, 0.84% v/v of Hv-Ca salts, 0.002% v/v of KPO₄ buffer (pH 7.0). Microwaved without 10 x Ca to dissolve agar. 10 x Ca added, then autoclaved. Hv-Ca salts and KPO₄ buffer added prior to pouring.

Hv-YPC broth: 18% SW, 1 x YPC, 3 mM CaCl₂. CaCl₂ added after autoclaving, when cool.

Hv-Ca+ broth: 18% SW, 30 mM Tris.HCl pH 7.0, 1 x Ca, 2.5% v/v of Hv-Min carbon source, 1.2% v/v of Hv-Min Salts, 0.002% v/v of KPO₄ buffer (pH 7.0), 444 nM biotin, 2.5 µM thiamine. 30% SW, dH₂O and Tris-HCl pH 7.0 autoclaved. All other components added when cool.

Tryptophan (trp), uracil (ura), 5-Fluoroorotic acid (5-FOA) are all added to *Haloferax volcanii* media to a final concentration of 50 µg/ml if needed.

Escherichia coli Media

Sterilised by autoclaving at 121°C for 15 mins. Stored at room temperature until supplements added, then stored at 4°C. Ampicillin (Amp) is added to a final concentration of 50 µg/ml in *E.coli* media.

LB (Lysogeny Broth): 1% tryptone (Bacto), 0.5% yeast extract (Difco), 170 mM NaCl, 2mM NaOH, pH 7.0.

LB agar: 300 ml of LB broth, 1.5% agar (Bacto).

2.1.6 Other Solutions

TE: 10 mM Tris.HCl pH 8.0, 1 mM EDTA.

Sodium Acetate: 3 M NaAc pH 5.2, filter sterilised.

2.2 Methods

2.2.1 General *Escherichia coli* Microbiology

Growth and Storage

Cultures on solid media were grown in a static incubator (LEEC) at 37°C overnight. Small liquid cultures (1-10 ml) were grown overnight at 37°C in the same incubator with 8 rpm rotation. Large scale liquid cultures (300 ml) were grown overnight at 37°C with 110 rpm shaking in an Innova 4330 floor-standing incubator (New Brunswick Scientific). For short term storage, all cultures were stored at 4°C. For long term storage, glycerol was added (20% v/v) to liquid cultures, snap frozen on dry ice and stored at -80°C.

Preparation of Electrocompetent Cells

The two strains of *E. coli*, XL-1 Blue (*dam*⁺, tetracycline resistant) and N2338 (*dam*⁻), were used to prepare electrocompetent *E. coli* cells.

A 5 ml culture was grown overnight with appropriate antibiotic selection at 37°C with 8 rpm rotation. Cells were diluted 1/100 in LB broth supplemented with appropriate antibiotics. These were grown at 37°C to $A_{650} = 0.5-0.8$. Cells were pelleted at 6000 x g for 12 minutes at 4°C. The supernatant was removed and the pellet resuspended in an equal volume of ice-cold sterile 1 mM HEPES (pH 7.5). This process was repeated using two thirds volume 1 mM HEPES (pH 7.5), one third volume 1 mM HEPES (pH 7.5), 0.1 volume 1 mM HEPES (pH 7.5) + 10% glycerol and finally 0.001 volume 1 mM HEPES (pH 7.5) + 10% glycerol. Cells were aliquoted into 100 µl aliquots, snap frozen on dry ice and stored at -80°C.

*Transformation of *Escherichia coli* by Electroporation*

1-2 µg of DNA in 5 µl of sterile dH₂O was added to 40 µl of electrocompetent cells on ice. The DNA and cells were gently mixed and transferred to a prechilled sterile electroporation cuvette (1 mm electrode gap, GENEFLOW). The cuvette was placed in an *E. coli* gene pulser (BioRad) and pulsed at 1.8 kV. 1 ml of SOC was immediately added and samples were incubated at 37°C with 8 rpm rotation for 1 hour, to allow for recovery of the cells. Cells were plated onto LB+2xAm^r plates and incubated at 37°C overnight.

Buffers and solutions:

SOC Broth: 2% tryptone (Bacto), 0.5% yeast extract (Difco), 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose.

2.2.2 General *Haloferax volcanii* Microbiology

Growth and Storage

Cultures on solid media were grown in a static incubator (LEEC) at 45°C for 514 days in a plastic bag to prevent drying of plates. Small liquid cultures (1-10 ml) were grown overnight at 45°C in the same incubator with 8 rpm rotation. Large scale liquid cultures (50-600 ml) were grown overnight at 45°C with 110 rpm shaking in an Innova 4330 floor-standing incubator (New Brunswick Scientific). For short term storage, all cultures were stored at room temperature. For long term storage, glycerol (80% glycerol in 6% SW) was added (20% v/v) to liquid cultures, snap frozen on dry ice and stored at -80°C.

Transformation of Haloferax volcanii Using PEG600

Haloferax volcanii can be easily and efficiently transformed by using PEG600 (Cline et al. 1989). DNA must first be passaged through a *dam*- *E. coli* host prior to transformation as *H. volcanii* encodes a restriction endonuclease (Mrr) that targets methylated DNA (Holmes et al. 1991). Δmrr strains of *H. volcanii* (deficient in Mrr) can be transformed directly with *dam*⁺ plasmid DNA.

A 5-10 ml culture in Hv-YPC broth (+Thy) of *H. volcanii* was grown overnight at 45°C with 8 rpm rotation until $A_{650} = 0.6-0.8$. Cells were pelleted by centrifugation at 3300 x g for 8 minutes in a 15 ml round-bottomed tube. The supernatant was removed and the cells gently resuspended in 2 ml buffered spheroplasting solution. Cells were transferred to a fresh 2 ml round-bottomed tube, pelleted again and supernatant removed. Cells were gently resuspended in 400-800 μ l buffered spheroplasting solution. A 200 μ l aliquot per transformation was transferred to a fresh 2 ml round-bottomed tube and a 20 μ l drop of EDTA pH 8.0 was pipetted onto the side of the tube, gently inverted and incubated for 10 minutes at room temperature. Transforming DNA was added in the same manner as EDTA and incubated at room temperature for 5 minutes. 250 μ l of 60% PEG600 was added to the side of the tube and mixed by gentle rocking followed by incubation at room temperature for 30 minutes. 1.5 ml of spheroplast dilution solution was added, mixed by gentle inversion and incubated at

room temperature for 2 minutes. It was then centrifuged at 3300 x g for 8 minutes at 25°C. The cell pellet was then transferred whole to a sterile 4 ml tube containing 1 ml regeneration solution. The cells were left to recover undisturbed for 90 minutes at 45°C after which the pellet was resuspended by gently tapping the side of the tube and incubated for a further 3-4 hours at 45°C with 8 rpm rotation. Cells were transferred to a fresh 2ml round-bottomed tube and centrifuged at 3300 x g for 8 minutes at 25°C. The supernatant was removed and the cell pellet gently resuspended in 1 ml of transformation dilution solution. Appropriate dilutions were made and 100 µl of each dilution plated on appropriate media. Plates were incubated for at least 5 days at 45°C.

Buffers and Solutions:

All buffers and solutions sterilised by filtration through a 0.2 µm Minisart filter (Sartorius).

Buffered Spheroplasting Solution: 1 M NaCl, 27 mM KCl, 50 mM Tris.HCl pH 8.5, 15% sucrose.

Unbuffered Spheroplasting Solution: 1 M NaCl, 27 mM KCl, 15% sucrose, pH 7.5.

Transforming DNA: 5 µl 0.5 M EDTA pH 8.0, 15 µl unbuffered spheroplasting solution, 10 µl DNA (~1-2 µg).

60% Polyethylene Glycol 600 (PEG600): 150 µl PEG600, 100 µl unbuffered spheroplasting solution.

Spheroplast Dilution Solution: 23% SW, 15% sucrose, 37.5 mM CaCl₂.

Regeneration Solution: 18% SW, 1 x YPC, 15% sucrose, 30 mM CaCl₂.

Transformation Dilution Solution: 18% SW, 15% sucrose, 30 mM CaCl₂.

2.2.3 DNA Extraction from Cells

Plasmid Extraction from Escherichia coli

Plasmid DNA extraction from *E. coli* was performed using Macherey-Nagel NucleoSpin Plasmid (Mini) and NucleoBond Xtra (Midi) kits. Protocol was followed as described in the manufacturer's guidelines. For minipreps 1-2 ml *E. coli* cell culture (LB broth+Amp) was used and eluted in 30 µl elution buffer. For midipreps 300 ml *E. coli* cell culture (LB broth+Amp) was used. Midipreps were eluted using isopropanol. The DNA was ethanol precipitated, resuspended in 200 µl of TE, and stored at -20°C.

Genomic DNA Extraction from Haloferax volcanii by Spooling

A 5 ml culture of *H. volcanii* was grown overnight in Hv-YPC (+Thy) at 45°C until $A_{650} = 0.6-0.8$. 2 ml of culture was transferred to a round-bottomed 2 ml tube and centrifuged at 3300 x g for 8 minutes at 25°C. The supernatant was removed and the pellet resuspended in 200 µl of ST buffer. 200 µl of lysis solution was added, the tube mixed by inversion, and the cell lysate overlaid with 1 ml of 100% EtOH. DNA was spooled at the interface with a capillary tip until the liquid was homogenous and clear. The spool of DNA was washed twice in 100% EtOH, and excess EtOH allowed to drain from the DNA. The DNA was air-dried, resuspended in 500 µl of TE, and precipitated with 400 µl isopropanol and 50 µl sodium acetate. The samples were centrifuged at 11,000 x g for 5 minutes, washed in 1 ml 70% EtOH, and dried thoroughly. The pellet was resuspended in 100-500 µl of TE and stored at 4°C.

Buffers and Solutions:

ST Buffer: 1 M NaCl, 20 mM Tris.HCl pH 7.5.

Lysis Buffer: 100 mM EDTA pH 8.0, 0.2% SDS.

2.2.4 Nucleic Acid Manipulation

PCR Amplification

Amplification of DNA was carried out using Q5 Hotstart or OneTaq. These enzymes are suitable for amplifying templates with high GC content. Q5 Hotstart were used for amplifications that required high fidelity. OneTaq were used for diagnostic amplifications. Reaction conditions are shown below. All reactions were carried out using a Techne TC-512 thermocycler.

Table 2.4 **PCR components for Q5 Hot Start and OneTaq**

Q5 Hot Start	OneTaq
200 µM of each dNTP	200 µM of each dNTP
0.5 µM of each primer	0.5 µM of each primer
1 ng – 1 µg genomic DNA or 1 pg – 1 ng plasmid DNA template	10 ng of template DNA
1 x Q5 Reaction Buffer	1 x OneTaq GC Buffer
1 x Q5 High GC Enhancer	-
0.02 U/µl Q5 Hotstart	0.025 U/µl OneTaq

Table 2.5 **PCR conditions for Q5 Hot Start and OneTaq**

Step	Q5 Hot Start	OneTaq
Initial Denaturation	98°C, 30 seconds	94°C, 30 seconds
Denaturation	98°C, 5-10 seconds	94°C, 30 seconds
Annealing	T _m °C, 10-30 seconds	T _m °C, 30 seconds
Extension	72°C, 20-30 seconds/kb	68°C, 60 seconds/kb
Final Extension	72°C, 5 minutes	68°C, 5 minutes

30 cycles

Annealing temperatures for primers (T_m°C) were calculated using the equation below (Equation 2.1) (Howley *et al.*, 1979).

$$81.5 + (16.6 \times \log_{10}[\text{Na}^+]) + (0.41 \times \%GC) - (100 - \%homology) - \left(\frac{600}{\text{length}}\right)$$

Equation 2.1 **Calculating annealing temperature of primers.** %GC =

percentage guanine and cytosine in the primer, Homology = percentage homology shared between primer and template, Length = length of primer in bases.

Touchdown PCR

When primers did not have 100% homology to the template DNA (e.g. when introducing restriction sites or mutations), two annealing temperatures were calculated. T_mS was based on the original percentage of homology between primer and template, T_mE was based on 100% homology. The reaction temperature started at T_mS and increased linearly to T_mE across 10 cycles. The remaining 20 cycles used T_mE as the annealing temperature.

Colony PCR

Colony PCRs were used for screening large numbers of colonies for a desired plasmid or chromosomal gene. *H. volcanii* or *E. coli* colonies were grown on solid media and touched lightly with a sterile yellow tip, ensuring only a small number of cells were collected and the colony was not disturbed. The yellow tip was pipetted up and down in 100 µl of sterile dH₂O to wash cells off the yellow tip. This was then boiled at 100°C to lyse cells for 10 minutes and immediately cooled on ice. 1 µl of the preparation was then used in a PCR with DyNAzyme EXT polymerase or OneTaq polymerase.

Restriction Digests

Restriction digests were carried out following manufacturer's instructions (NEB). All digests were supplemented with 200 ng/μl BSA (NEB). For double digests NEB buffers were selected so that each enzyme had at least 75% activity. Plasmid DNA was digested for at least one hour. Genomic DNA and PCR products were digested for 16 hours.

Dephosphorylation of Vector DNA

To prevent self-ligation of vector DNA, Shrimp alkaline phosphatase was used to remove 5' phosphate groups. Samples were incubated with 5 units of Shrimp alkaline phosphatase / μg of DNA and 1 x Antarctic phosphatase buffer for 30 minutes at 37°C. Phosphatase was heat inactivated at 65°C for 10 minutes.

Ligation of DNA

Ligations were performed using T4 DNA ligase. For each μg of DNA, 5 units of ligase were used in a reaction with 1 x T4 ligase buffer. For vector:insert ligations, reactions contained a molar ratio of >3:1 insert to vector DNA. Ligations were carried out at 15°C overnight or 4°C for 36 hours. Ligations were ethanol precipitated, resuspended in 5 μl sterile dH₂O and transformed into *E. coli*.

Ethanol Precipitation of DNA

To ethanol precipitate DNA, 2 volumes of 100% EtOH and 1/10 volume of 3 M sodium acetate (pH 5.2) were added to DNA and incubated at -20°C for at least 1 hour. Samples were centrifuged at 20,000 x g for 30 minutes at 4°C and the supernatant removed. Pellets were washed in 400 μl 70% EtOH followed by centrifugation at 20,000 x g for 10 minutes at 4°C. The supernatant was removed and the pellets air-dried thoroughly before resuspension in sterile dH₂O.

Nucleic Acid Purification

PCR products, ligations, restriction digests and dephosphorylated DNA products were purified using Macherey-Nagel DNA purification kits. Protocol was followed according to manufacturer's instructions. In these kits DNA is bound pH dependently to a silica membrane and is separated from contaminants by washing with ethanol. DNA was eluted in 30 μl of the provided elution buffer.

DNA Sequencing

All DNA sequencing reactions and analysis were performed by the Biopolymer Synthesis and Analysis Unit, University of Nottingham. Sequencing was carried out using the dideoxy chain termination method (Sanger et al. 1977).

Oligonucleotide Synthesis

All oligonucleotides were synthesised by Eurofins MWG, Germany.

Nucleic Acid Quantification

To determine the concentration and purity of plasmid preparations the absorbance at 260 nm and the 260:280 nm absorbance ratio, respectively, were measured by spectrometer (Beckman Coulter DU 530).

Agarose Gel Electrophoresis

TBE buffer was used as standard practice for casting and running agarose gels. TAE buffer was used when high quality resolution and/or Southern blotting was required. Agarose gels were cast using agarose powder (SeaKem Lonza) and either TBE or TAE buffer. Gel loading dye was added to the DNA samples to give a final concentration of 1 x. All samples and molecular markers, either 1 kb (NEB) or 100 bp (NEB), were loaded onto the gel. TBE gels (10 cm) were run at 110 V for ~1 hour. TAE gels (25 cm) were run overnight (16 hours) at 50 V with buffer circulation. For visualisation of bands, gels were stained with ethidium bromide at a final concentration of 0.5 µg/ml or SYBR Safe (Invitrogen) at a 0.5 x final concentration. Gels used for DNA extraction were stained with ethidium bromide after the removal of sample lanes.

Buffers and Solutions:

TBE (Tris/Borate/EDTA): 89 mM Tris, 89 mM boric acid, 2 mM EDTA pH 8.0.

TAE (Tris/Acetic acid/EDTA): 40 mM Tris, 20 mM acetic acid, 1 mM EDTA pH 8.0.

Gel loading dye (5x): 50 mM Tris.HCl, 100 mM EDTA pH 8.0, 15% Ficoll (w/v), 0.25% Bromophenol Blue (w/v), 0.25% Xylene Cyanol FF (w/v).

Agarose Gel Extraction and Purification of DNA

To purify DNA from agarose gels without UV exposure, sample lanes were protected using foil while the appropriate band was excised. DNA was visualised by UV light using a UV transilluminator (UVP inc.). DNA was purified using the Macherey-Nagel DNA purification kit following manufacturer's guidelines.

2.2.5 Genetic Manipulation of *Haloferax volcanii*

2.2.5.1 Plasmid Construction

Generating a Gene Replacement Construct

Gene replacement constructs were made by inserting the gene of interest (usually with the addition of an inducible promoter or tags) along with its flanking regions into pTA131 or derivative. The exact protocol for doing this was different for each gene replacement construct and is detailed in relevant chapters.

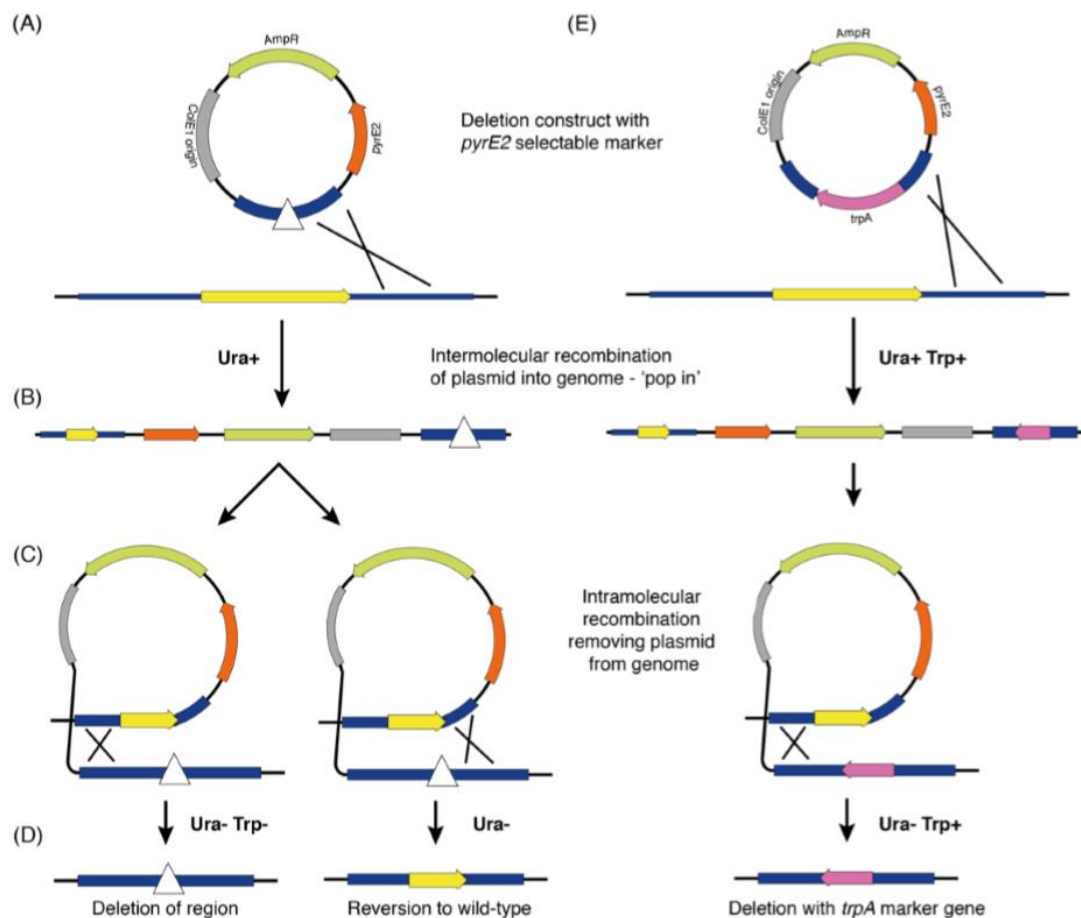


Figure 2.6 **Gene deletion by pop-in/pop-out method.** (A) A $\Delta pyrE2$ strain is transformed with *pyrE2*⁺ deletion construct. (B) Pop-in colonies selected using *ura*⁺ phenotype. Cells undergo pop-out when uracil selection is relieved. This is selected for by plating on 5-FOA. (C) Recombination can be in the downstream region (left) or the upstream region (right). (D) The gene is deleted (left) or it reverts to wild type (right). (E) A *trpA* marker can be used in the deletion construct to allow for direct selection of the deletion.

2.2.6 Genotype Screening

There are variety of methods that can be used to screen for the presence or absence of a specific gene in *Haloferax volcanii*. If the strain has a selectable phenotype e.g. $\Delta trpA$ colonies can be plated onto selectable media. However, due to the polyploid nature of *H. volcanii* strains can be mero-diploid meaning that they have a mixture of mutant and wild-type alleles present on different chromosome copies. To screen for colonies in which all chromosome copies carry the same allele, colony hybridisation and Southern blotting are used. For both of these methods DNA is denatured and transferred to a positively charged membrane by either colony lift or vacuum transfer, respectively.

Southern Blot Vacuum Transfer

Purified *H. volcanii* DNA was digested with enzymes that cut either side of the region of interest. The resulting DNA fragments were separated on a 200 ml 0.75% TAE gel for 16 hours at 50 V, with buffer circulation. The gel was poststained with 0.5 $\mu\text{g/ml}$ ethidium bromide for 30 minutes and visualised. The gelembedded DNA was acid nicked for 20 minutes in 0.25 M HCl, followed by washing for 10 minutes in dH₂O. The DNA was then denatured in denaturing solution for 45 minutes. Membrane (BioRad Zeta-Probe GT or Amersham Hybond-XL) was soaked in dH₂O for 5 minutes before equilibrating in denaturing solution for a further 2 minutes. Vacuum transfer was carried out using a Vacugene XL gel blotter and Vacugene Pump (Pharmacia Biotech) for 1 hour 15 minutes at 50 mBar. Following transfer the membrane was washed briefly in 2 x SSPE for <30 seconds and air-dried before DNA was cross-linked with 120 mJ/cm² UV.

Buffers and Solutions:

20 x SSPE: 3 M NaCl, 230 mM NaH₂PO₄, 32 mM EDTA, pH 7.4.

Denaturing Solution: 1.5 M NaCl, 0.5 M NaOH.

Hybridisation

Membranes from either colony lifts or Southern blots were pre-hybridised for >3 hours at 65°C in 40 ml pre-hybridisation solution. Radiolabelled DNA probes were made with 50 ng of DNA and 0.74 MBq of [α -³²P] dCTP (Perkin Elmer). DNA was denatured at 100°C for 5 minutes then incubated with the radioisotope and HiPrime random priming mix (Roche) for 15-20 minutes at 37°C. The radiolabelled probe was

then purified on a BioRad P-30 column and mixed with 10 mg/ml salmon sperm DNA, followed by denaturing at 100°C for 5 minutes and quenching on ice. For Southern blots 3 µl of 1 µg/ml 1 kb ladder or 3 µl of 1 µg/ml lambda DNA for Pulsed-Field Gel Southern blots was also included in the radiolabelling reaction. The pre-hybridisation solution was replaced with 30 ml of hybridisation solution, the probe DNA was added and then the membranes were incubated overnight at 65°C. The membranes were washed twice with 50 ml low stringency wash solution, once for 10 minutes and once for 30 minutes. This was followed by another two washes with high stringency wash solution, both for 30 minutes. Membranes were air-dried before being wrapped in Saran wrap and exposed to a phosphorimager screen (Fujifilm BAS Cassette 2325) for >24 hours. The screen was scanned using a Molecular Dynamics STORM 840 scanner. Alternatively, membranes were visualized using Amersham Hyperfilm MP (GE Healthcare).

Buffers and Solutions:

100 x Denhardt's Solution: 2% Ficoll 400, 2% PVP (Polyvinyl Pyrrolidone) 360, 2% BSA (Bovine Serum Albumin, Fraction V).

20 x SSPE: 3 M NaCl, 230 mM NaH₂PO₄, 32 mM EDTA, pH 7.4.

Pre-hybridisation Solution: 6 x SSPE, 1% SDS. 5 x Denhardt's solution, 200 µg/ml salmon sperm DNA (Roche, boiled for 5 minutes at 100°C prior to addition).

Hybridisation Solution: 6 x SSPE, 1% SDS, 5% dextran sulfate.

Low Stringency Wash Solution: 2 x SSPE, 0.5% SDS.

High Stringency Wash Solution: 0.2 x SSPE, 0.5% SDS.

2.2.7 Phenotyping of *Haloferax volcanii*

Tryptophan Gradient Plates

Tryptophan gradient plates were poured with 10 ml Hv-Ca + Trp (in appropriate concentration) on a slant forming a tapered wedge of agar (Figure 2.7). Once set, this was covered with Hv-Ca agar (without tryptophan) with the plate flat, therefore forming a gradient across the plate, with the desired tryptophan concentration at one side of the plate reducing to no tryptophan on the other side (Adapted from Bryson & Szybalski 1952). Plates were dried for 15 minutes.

Cultures of strains were set up in 5 ml Hv-YPC and incubated overnight at 45 °C with 8 rpm rotation until $A_{650} = 0.6-0.8$. These were then diluted into fresh Hv-YPC and incubated at 45°C until $A_{650} = 1.0$. Cultures were then diluted appropriately in 18% SW. Autoclaved paintbrushes were first dipped in 18% SW, before being dipped into the diluted culture and, then painted in one direction (from Trp+ to Trp-) across the gradient plate. The same paintbrush was then dipped into the diluted culture again and painted over the top of the first streak in the other direction. The plates were then incubated at 45°C for 4-5 days.

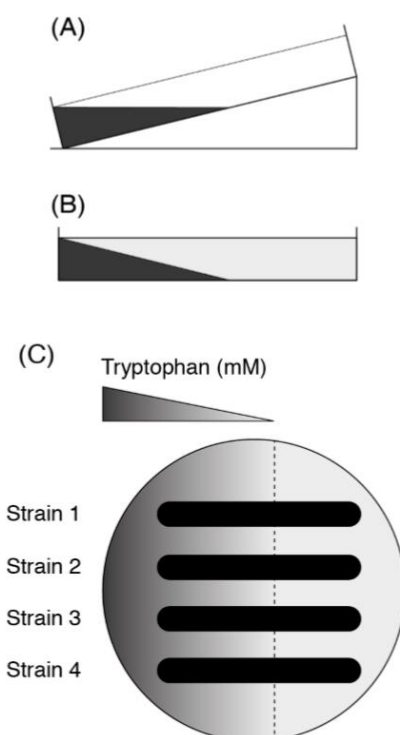


Figure 2.7 **Tryptophan gradient plates.** (A) A wedge of Hv-Ca media with the desired concentration of tryptophan is poured while the plate is slanted. (B) The plate is then laid flat and Hv-Ca media (without tryptophan) is poured over the rest of the plate to form a flat surface. (C) Strains are painted across the tryptophan gradient.

Growth Curves

For testing growth rates in liquid media, cultures were set up in 5 ml Hv-YPC broth and grown at 45°C to an $A_{650} = 0.2-0.8$ (exponential phase). Cells were diluted and grown again to $A_{650} = 0.2-0.8$. Cells were diluted again and grown overnight at 45°C until they reached $A_{650} = 0.4$. 250 μ l of culture and appropriate blanks were added to the wells of a 98-well microtiter plate (Corning). The plate was sealed with 1.25 cm microporous tape (Boots) and incubated at 45°C with double orbital shaking at 425

rpm for between 48-72 hours in an Epoch 2 Microplate Spectrophotometer (BioTek). Readings were taken every 15 minutes at A₆₀₀. The generation time was calculated by plotting the growth on a log₂ scale and using the equation below (Equation 2.2).

$$G = \frac{t}{n}$$

$$n = \frac{\log b - \log B}{\log 2}$$

G = generation time
t = time
n = number of generations
B = start OD
b = end OD

Equation 2.2 **Calculating the generation time of a liquid culture using the OD A₆₀₀**

It had previously been noted that generation times of strains varied between experiments when measuring growth in this method (Gamble-Milner, 2016). Therefore, in this study, comparisons are made only between duplicates of strains within the same experiment i.e. strains that have been grown on the same microtiter plate and the A₆₀₀ measured simultaneously during a single run on the Epoch2 Microplate Spectrophotometer (BioTek). As generation times are seen to vary between experiments, generation times stated are not absolute. However, the relationship between different strains is observed to be consistent, meaning that reliable comparisons are able to be made between different strains in the same experiment. This means that the generation times and growth curves generated by this method are used to illustrate the differences between different strains rather than an exact generation time. It is also worth noting that this method is not sensitive enough to detect very small differences in growth that can be observed with growth competition methods (Delmas et al. 2009).

DNA Damage Sensitivity Assays

UV radiation sensitivity assays

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh culture was diluted again (1/500) and grown at 45°C overnight. The second overnight was diluted to OD₆₅₀ ≈ 0.1 and then put back at 45°C till OD₆₅₀ ≈ 0.4. Serial dilutions were made from 10⁻¹ to 10⁻⁸ and duplicate 20 µl drops of culture were pipetted onto Hv-YPC agar. Spots were allowed to dry at room temperature and plates were either exposed to ultraviolet light at a rate of 1 J/m²/second (254 nm peak)

or not exposed to UV light, as a control. Plates were shielded from visible light by use of a black plastic bag to prevent photoreactivation of DNA, and incubated at 45°C until colonies were visible.

Gamma radiation survival assays

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh culture was diluted again (1/500) and grown at 45°C overnight. The second overnight was diluted to $OD_{650} \approx 0.1$ and then put back at 45°C till $OD_{650} \approx 0.8$. Serial dilutions were made from 10^{-1} to 10^{-8} and duplicate 20 μ l drops of culture were pipetted onto small Hv-YPC agar. Plots were allowed to dry at room temperature. Plates were either exposed to gamma radiation (^{137}Cs as the source) at a rate of 4.75 Gy/min or not exposed to gamma radiation, as a control. Plates were incubated at 45°C until colonies were visible.

Phleomycin sensitivity assays

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh culture was diluted again (1/500) and grown at 45°C overnight. The second overnight was diluted to $OD_{650} \approx 0.1$ and then put back at 45°C till $OD_{650} \approx 0.5$. The culture was divided into several aliquots of 0.5 ml and transferred to 4 ml culture tubes. 10 μ l serial dilution of phleomycin in 18% SW were added to each tube to the final concentration of 0.5, 1 and 2 mg/ml or 20 μ l 18% SW were added to the control tube. Then the tubes were incubated on Eppendorf thermomixer at 45°C at 450 rpm for 1 hour. After 1 hour, 1 ml incubated cultures were transferred to 2 ml round bottomed tubes and spun for 8 minutes at 6000 rpm to stop the treatment. Cells were resuspended in 1 ml Hv-YPC broth and serial dilutions were made from 10^{-1} to 10^{-8} in 18% saltwater. Duplicate 20 μ l drops of culture were pipetted onto Hv-YPC agar. Plots were allowed to dry at room temperature. Plates were incubated at 45°C until colonies were visible.

H₂O₂ sensitivity assays

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh culture was diluted again (1/500) and grown at 45°C overnight. The second overnight was diluted to $OD_{650} \approx 0.1$ and then put back at 45°C till $OD_{650} \approx 0.4$. The culture was divided into several aliquots of 1 ml and transferred to 4 ml culture tubes. 20 μ l serial dilution of H₂O₂ in 18% SW were added to each tube to the final concentration of 2, 4 and 6 mM H₂O₂ or 20 μ l 18% SW were added to the control tube. The tubes were incubated on Eppendorf thermomixer at 45°C for 1 hour. . After

1 hour, 1 ml incubated cultures were transferred to 2 ml round bottomed tubes and spun for 8 minutes at 6000 rpm to stop the treatment. Cells were resuspended in 1 ml Hv-YPC broth and serial dilutions were made from 10^{-1} to 10^{-8} in 18% saltwater. Duplicate 20 μ l drops of culture were pipetted onto Hv-YPC agar. Plots were allowed to dry at room temperature. Plates were incubated at 45° until colonies were visible.

MMS sensitivity assays

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh culture was diluted again (1/500) and grown at 45°C overnight. The second overnight was diluted to $OD_{650} \approx 0.1$ and then put back at 45°C till $OD_{650} \approx 0.6$. The culture was divided into several aliquots of 1 ml and transferred to 4 ml culture tubes. 20 μ l serial dilution of MMS in 18% SW were added to each tube to the final concentration of 0.02, 0.03, 0.04, 0.05% or 20 μ l 18% SW were added to the control tube. Then the tubes were incubated on Eppendorf thermomixer at 45°C 450 rpm for 1 hour. After 1 hour, 1 ml incubated cultures were transferred to 2 ml round bottomed tubes and spun for 8 minutes at 6000 rpm to stop the treatment. Cells were resuspended in 1 ml Hv-YPC broth and serial dilutions were made from 10^{-1} to 10^{-8} in 18% saltwater. Duplicate 20 μ l drops of culture were pipetted onto Hv-YPC agar. Spots were allowed to dry at room temperature. Plates were incubated at 45° until colonies were visible.

Total RNA Extraction from H. volcanii for RNA Sequencing

Colonies were used to inoculate 5 ml of Hv-YPC broth and incubated at 45°C overnight. The fresh cultures were diluted to $OD_{650} \approx 0.1$ and then put back at 45°C till $OD_{650} \approx 0.4$. Cultures were increased by diluting the $OD_{650} \approx 0.4$ cultures in 50 ml Hv-YPC broth in sterile flasks aiming for an $OD_{650} \approx 0.2$ for the next day. RNA extraction (or treatment, if required) starts when the cultures are at $OD_{650} \approx 0.2$.

To start the extraction, cells were pelleted at 6,000 rpm for 8 minutes at 4 °C. The supernatant was removed and cells were resuspended in 1 ml of Trizol (Invitrogen) and incubated for 5 minutes at room temperature. Then 200 μ l chloroform was added. The tube was vortexed for 15 seconds and further incubated at room temperature for 3-5 minutes. Cell lysate was centrifuged at 4°C at 12,000 x g for 15 minutes. The top aqueous layer was transferred very carefully to a clean tube and overlaid with 0.5 ml isopropanol and incubated for 10-15 minutes to precipitate RNA. The pellet was

precipitated again by centrifugation at 12,000 x g at 4 °C for 10 minutes. Supernatant was removed and pellet was washed twice with 1.2 ml 75% ethanol (ice-cold) by inverting the tube a few times and centrifugation at 7,500 x g for 5 minutes at 4 °C between washes. Supernatant was removed completely and the pellet was air dried for 10 minutes by inverting the tube on clean tissue. The pellet was resuspended in 50 µl RNAase-free water. 5 µl of 10X Ambion TURBO DNase buffer and 1 µl of TURBO DNase (Ambion) were added. The sample was incubated at 37°C for 30 minutes and then 1 µl of TURBO DNase were added before a further 30 minute-incubation. 5 µl of DNase inactivation reagent was added and the sample was incubated for 3 minutes at room temperature. The samples were centrifuged for 1.5 minutes at 14,000 x g at RT and supernatant was transferred to a new tube. Samples were separated into small aliquots and stored at 80°C until required to minimise degradation. The concentration and purity ($A_{260/280} \approx 2$) of each RNA sample was measured using Epoch 2 Microplate Spectrophotometer (BioTek).

The RNA-Sequencing was carried out by the Deep Seq facility in the School of Life Sciences at the University of Nottingham. Sequencing was performed using Illumina NextSeq500 instrument with a 150-cycle mid output sequencing kit to generate approx. 10million paired end reads per sample (2 x 75bp). This is the standard level of sequencing for expression profiling of prokaryotes.

Chapter 3: DNA Damage Sensitivity Assays

3.1 Background

Historically, DNA damaging agents have been used extensively to reveal and characterize the function of proteins that are involved in key DNA repair pathways. If a mutant strain deleted for a gene is hypersensitive to a specific form of DNA damage, it suggests a role for that gene in a repair pathway induced by the mutagen used. The biological significance of a lesion depends on the extent at which it distorts the DNA double helix structure and whether it causes a blockage for transcription or DNA replication. A multitude of repair pathways have evolved to correct the plethora of DNA lesions arising either endogenously or exogenously.

Deletion of the *orc9* gene does not alter the replication profile of *Haloferax volcanii*, which suggests that Orc9 does not have a role in DNA replication. Preliminary data suggests that Orc9 might have a role in DNA damage repair. Δ *orc9* strains exhibit sensitivity to UV radiation, 4-NQO and MMC. Using pulsed field gel electrophoresis, Δ *orc9* mutants were shown to have increased genomic instability when treated with MMC (Darya Ausiannikava, Victoria Smith, unpublished data). Testing the sensitivity of Δ *orc9* mutant strains to mutagens that induce a wider range of lesions might help elucidate the role of Orc9 in DNA damage repair.

3.2 Aims

Several mutagens that induce different types of DNA damage were used to test the activity of Orc9 in multiple repair pathways. The overall aim of this chapter was to determine which DNA damage repair pathway(s) *H. volcanii* Orc9 might be involved in. The aims of this chapter are:

- Perform gamma radiation, phleomycin, MMS and H₂O₂ sensitivity assays on WT and Δ *orc9* strains, to test their sensitivity to these mutagens.
- Repeat the UV sensitivity assay using WT and Δ *orc9* strains to confirm the hypersensitivity phenotype of Δ *orc9* to UV-induced DNA damage.
- Perform growth assays to test sensitivity of WT and Δ *orc9* strains to mitomycin C and H₂O₂.

3.3 Results

3.3.1 MMS and H₂O₂ Sensitivity Assays

Background

Methyl methane-sulphonate (or **methanesulfonic acid methyl ester**, MMS) is a monofunctional alkylating agent with a single electrophilic group that reacts covalently with one of the nucleophilic sites in DNA. This reaction adds a methyl group to one of the 4 bases, adenine and guanine being the most frequently methylated (Table 3.1) (Wyatt and Pittman, 2006). Not all alkylation adducts have the same level of toxicity in cells. N3-MeA blocks replication and induces low levels of A:T to T:A transitions. N7-MeG does not block or miscode, however depurination might lead to a toxic abasic site (Boiteux and Laval, 1983) (Larson *et al.*, 1985). O6-MeG is known to produce mispairing of G with T instead of a C (Loechler, Green and Essigmann, 1984). The biological significance of adducts that have lower occurrence is not well studied (Wyatt and Pittman, 2006).

Table 3.1 **Relative proportions of methylated bases in DNA after reaction with MMS.**
Data from (Beranek, 1990).

Type of Adduct	Percentage of total methylation (%)
N7-methylguanine (N7-MeG)	83
N3-methyladenine (N3-MeA)	11
N7-methyladenine (N7-MeA)	1.9
N1-methyladenine (N1-MeA)	1.2
N3-methylguanine (N3-MeG)	0.7
O ⁶ -Methylguanine (O6-MeG)	0.3

Alkylation damage in DNA is removed via direct reversal by alkyltransferases or the BER pathway. Unlike direct reversal of alkylation damage, BER is not a single-step process and generates intermediates that are potentially toxic to cells (such as AP sites). If replication forks encounter BER intermediates or methyl damage, strand breaks may be produced. In *Saccharomyces cerevisiae*, *rad52* mutants (which are hypersensitive to radiation) were shown to be hypersensitive to MMS as well (Krogh and Symington, 2004) (Xiao, Chow and Rathgeber, 1996). For this reason, MMS was

misleadingly defined as a ‘radiomimetic’ chemical, however MMS can only produce strand breaks indirectly.

Hydrogen peroxide (H_2O_2) is an oxidizing agent that can be produced endogenously in cells as a by-product of the electron transport chain (Figure 3.1). Hydrogen peroxide breaks down in an iron-mediated Fenton reaction to generate a highly toxic hydroxyl radical that causes oxidative base damage and strand breaks in DNA (Goldstein, Meyerstein and Czapski, 1993). There are more than 20 base modifications caused by oxidative damage. The most common damage to purines and pyrimidines are 7,8-dihydro-8-oxoguanine, more commonly known as 8-oxoguanine (8-oxoG), and thymine glycol respectively (Henle and Linn, 1997).

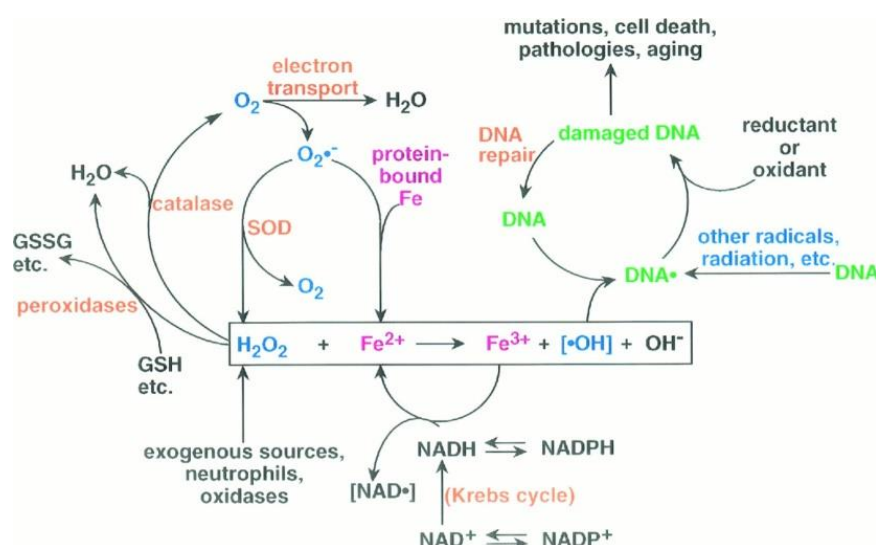


Figure 3.1 **H_2O_2 causing oxidative damage to DNA via the Fenton reaction.** Superoxide dismutase (SOD) generates H_2O_2 and superoxide ($\text{O}_2^{\bullet-}$) releases protein bound Fe to form Fe^{+2} . H_2O_2 and Fe^{+2} can react to generate $\bullet\text{OH}$ hydroxyl type radicals which causes oxidative DNA damage. Catalases and peroxidases break down and deplete H_2O_2 . Figure from (Henle and Linn, 1997).

Catalase, superoxide dismutase and peroxidase enzymes are the primary defence against reactive oxygen species and protect cells from their lethal effects. Several repair pathways are involved in oxidative DNA damage repair: Base excision repair (BER), transcription-coupled repair (TCR), global genome repair (GGR), mismatch repair (MMR), translesion synthesis (TLS), homologous recombination (HR) and non-homologous end-joining (NHEJ) all contribute to repair of oxidative DNA damage in cells. However, BER is the major pathway for the repair of oxidative DNA damage (Slupphaug, Kavli and Krokan, 2003).

Results

Strains H53 (WT) and H2655 ($\Delta orc9$) were treated with different concentrations of MMS for 1 hour to test their sensitivity to MMS-induced DNA damage (See *Chapter 2: Materials and Methods*). The $\Delta orc9$ strain was consistently more sensitive to MMS than the WT strain at all treatment doses (Figure 3.2). The difference between the survival fractions of H53 and H2655 at each MMS concentration is significant as the error bars of each data point do not overlap. At the highest MMS dose (0.05%), the average survival fraction of the WT strain was about 65 times higher than the $\Delta orc9$ strain.

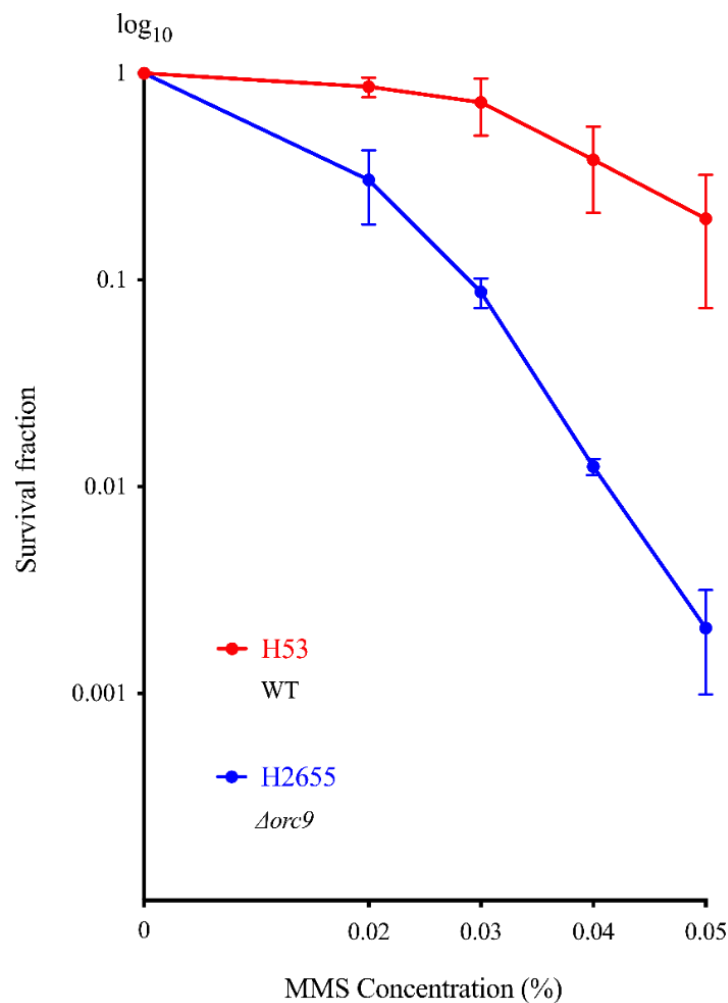


Figure 3.2: **Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains upon 1-hour acute treatment with MMS.** Data points show the mean of three independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a log₁₀ scale.

Strains H53 (WT) and H2655 ($\Delta orc9$) were treated with different concentrations of H_2O_2 for 1 hour to test their sensitivity to H_2O_2 -induced DNA damage (See *Chapter 2: Materials and Methods*). The $\Delta orc9$ strain was not more sensitive to H_2O_2 than the WT strain at any of the treatment concentrations (Figure 3.3). At all concentrations, the error bars of each data point overlap, which means the difference between the survival fractions of the two strains at each H_2O_2 concentration is not significant. Strains H53 (WT) and H2655 ($\Delta orc9$) were used to set up a growth assay under increasing concentrations of H_2O_2 (0, 0.9, 1.1, 1.3 mM) in the media (Figure 3.4). Growth of H2655 ($\Delta orc9$) does not slow down more than growth of H53 (WT) in the presence of increasing concentrations of H_2O_2 . Growth curves are parallel to each other which suggests similar generation times.

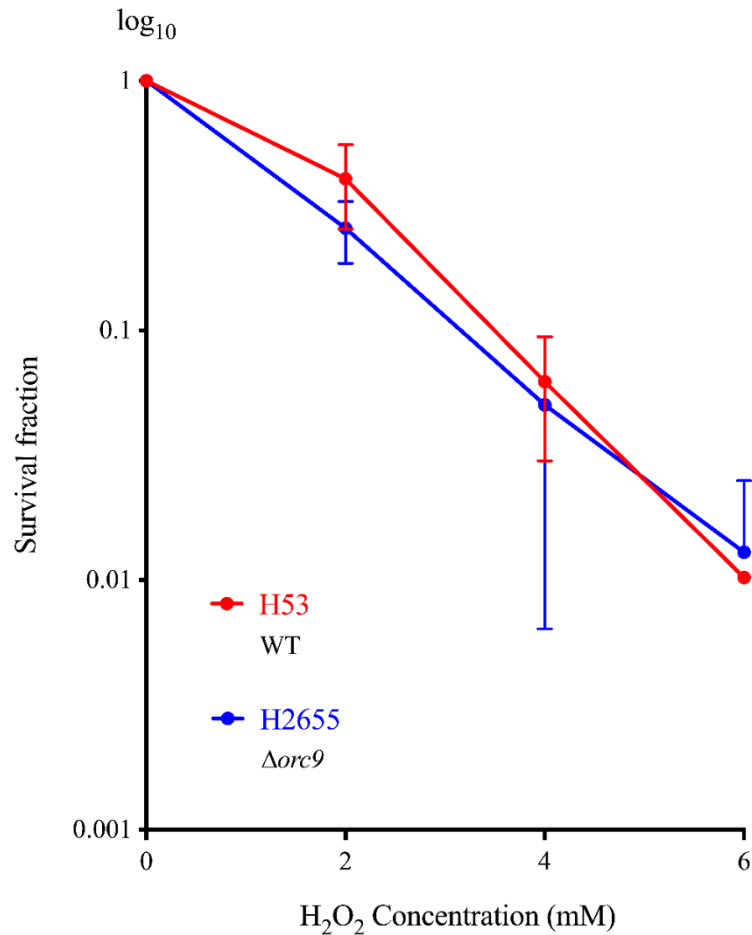


Figure 3.3: **Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains upon 1-hour acute treatment with H_2O_2 .** Data points show the mean of three independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a log₁₀ scale.

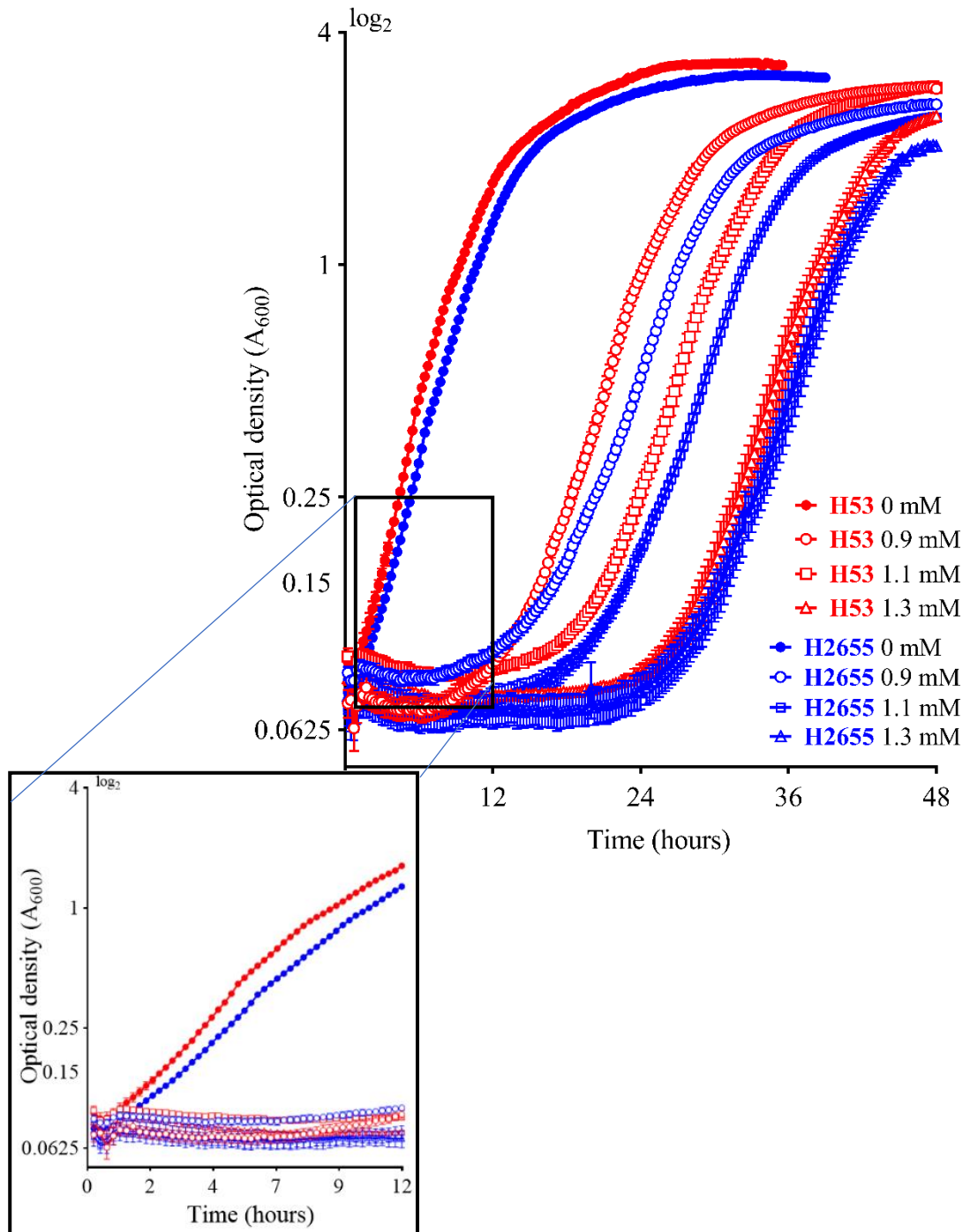


Figure 3.4: **Growth curves of H53 (WT) and H2655 ($\Delta orc9$) under increasing concentrations of H_2O_2 .** Optical density is given on a $\log_2$ scale. Error bars represent the standard error of the mean of three repeats. Small graph shows the growth curves in the first 12 hours of the assay.

3.3.2 UV Radiation and 4-NQO Sensitivity Assays

Background

UV radiation induces many types of lesions in DNA. Cyclobutane pyrimidine dimers (CPDs) and (6-4) pyrimidine photoproducts (6-4PPs) are the most common types caused by UV damage (Figure 3.5).

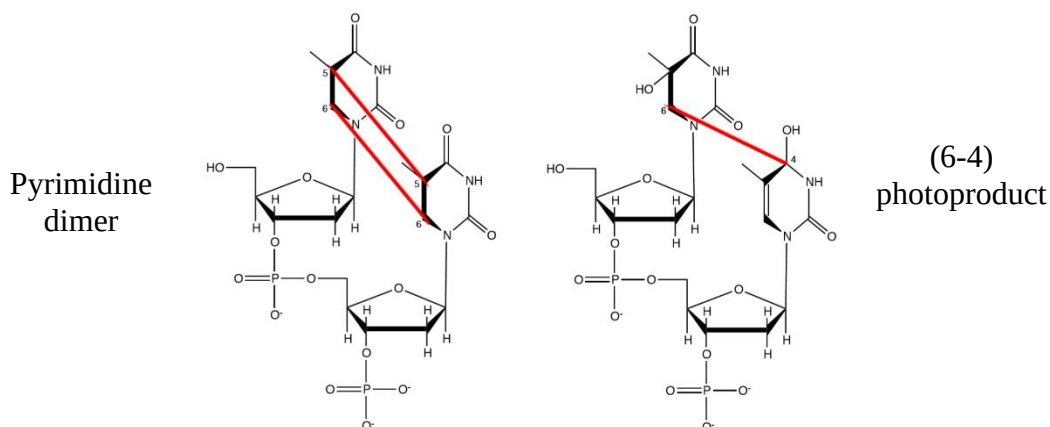


Figure 3.5: **UV radiation induces bulky lesions in DNA.**

DNA molecules absorb UV light, and this results in the breakage of the C=C double bond in pyrimidine bases. If this results in a covalent linkage between adjacent pyrimidine bases, a pyrimidine dimer is formed. If a single bond is formed between C⁶ and C⁴ carbons of adjacent pyrimidine bases, 6-4 photoproducts form. Both types of lesion are called ‘bulky’ lesions and they cause perturbations of the normal DNA structure. If these lesions remain unrepaired, stalled replication forks, mispairing of bases during DNA replication and strand breaks can occur (Sinha and Häder, 2002). UV radiation can indirectly form double-strand breaks when replication forks collapse at the site of a single strand break induced by UV radiation, or if there are two single strand breaks close in proximity (Lankinen, Vilpo and Vilpo, 1996).

Photolyases carry out photoreactivation and directly repair CPDs and 6-4 PPs. Nucleotide excision repair (NER) is the major pathway that is involved in carrying out UV-induced DNA damage repair. DSBs are repaired by homologous recombination. 4-nitroquinoline 1-oxide (4-NQO) is a UV-mimetic that induces an adduct named hydroxydeoxyguanosine in DNA, which is predominantly repaired by NER. It is also known to induce single and double strand breaks (Arima *et al.*, 2006) (Andoh and Ide, 1972).

Results

Strains H53 (WT) and H2655 ($\Delta orc9$) were exposed to different doses of UV radiation to test their sensitivity to UV-induced DNA damage (See *Chapter 2: Materials and Methods*). The results confirmed preliminary findings. The $\Delta orc9$ strain was more sensitive to UV radiation than the WT strain (Figure 3.6).

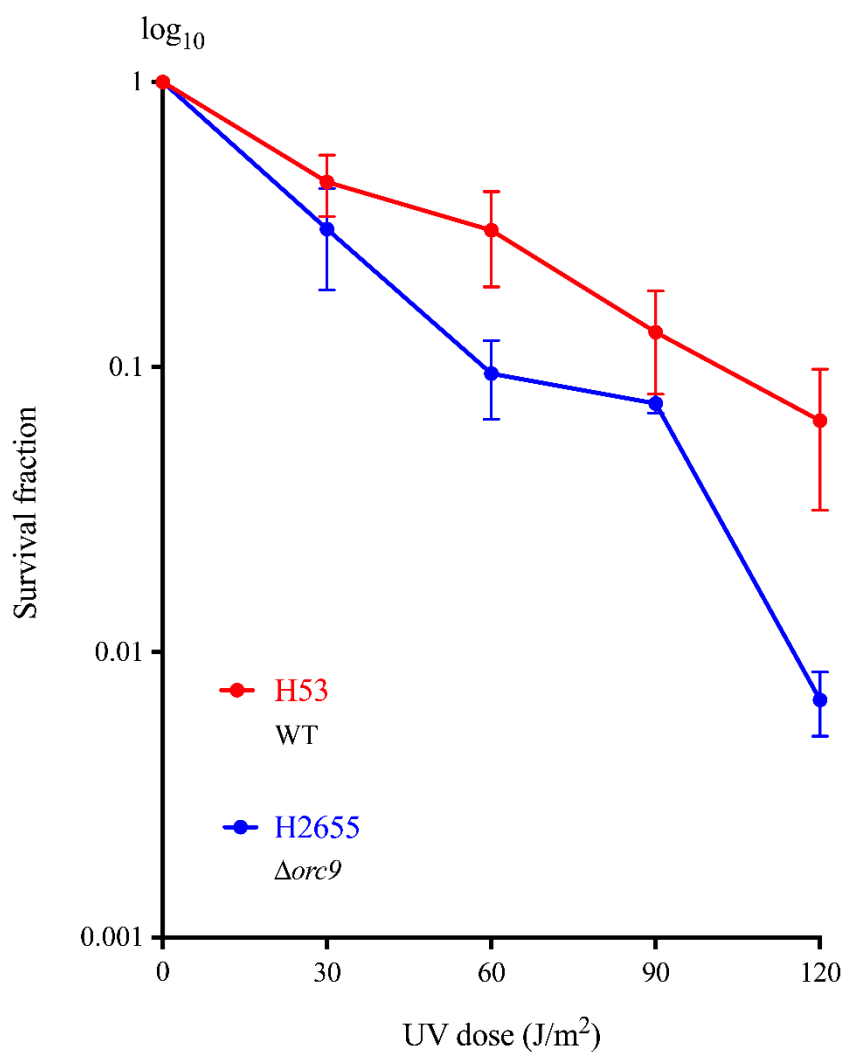


Figure 3.6: **Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains after exposure to different doses of UV radiation.** Data points show the mean of three independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a $\log_{10}$ scale.

The 4-NQO sensitivity assay was performed by Victoria Smith prior to this study (unpublished data). Strains H53 (WT) and H2655 ($\Delta orc9$) were treated with different concentrations of 4-NQO for 1 hour to test their sensitivity to 4-NQO-induced DNA damage. Consistent with the results from the UV radiation sensitivity assay, the $\Delta orc9$ strain was more sensitive to 4-NQO than the WT strain (Figure 3.7).

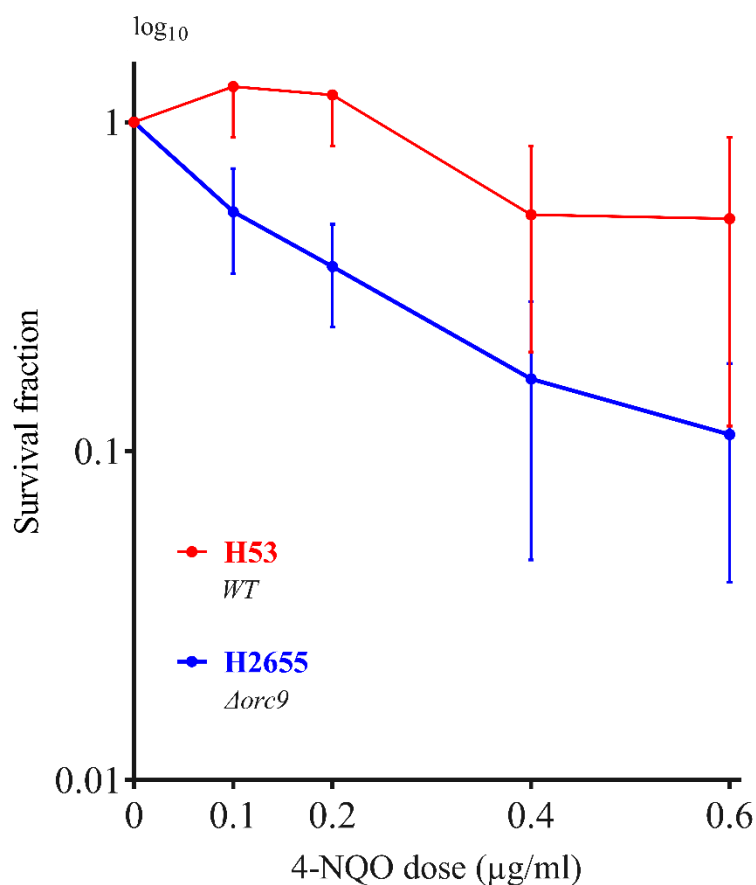


Figure 3.7: **Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains upon 1-hour acute treatment with 4-NQO.** Data points show the mean of two independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a log₁₀ scale. Data from Victoria Smith (unpublished).

3.3.3 Gamma Radiation and Phleomycin Sensitivity Assays

Background

Gamma radiation is a type of ionising radiation whose damage to DNA has been extensively studied. Gamma radiation induces lesions in DNA via direct and indirect effects. The major source of indirect effects to DNA by ionising radiation is reactive oxygen species generated by the radiolysis of water. The reactive oxygen species attack DNA bases and form several adducts. However, this type of damage is not the most serious lesion induced by gamma radiation, which is double strand breaks (DSBs).

Phleomycin is a mutagen that is structurally related to bleomycin group of antibiotics (Figure 3.8). Bleomycins have DNA-cleaving activity and generate lesions including apurinic/apyrimidinic sites and single and double-strand breaks. Due to their antitumour activity, bleomycins are used to treat different types of cancers. The type of phleomycin used in this assay is phleomycin D1 (zeocin) which is naturally produced by the bacterium *Streptomyces verticillus*. Phleomycin D1 is soluble in water and its solution has a blue colour due to the presence of chelated Cu^{+2} . Like bleomycin, phleomycin binds DNA via its bithiazole moiety (Povirk, 1996). Phleomycin is a radiomimetic chemical that is used to induce single and double-strand breaks in DNA (Sleigh, 1976).

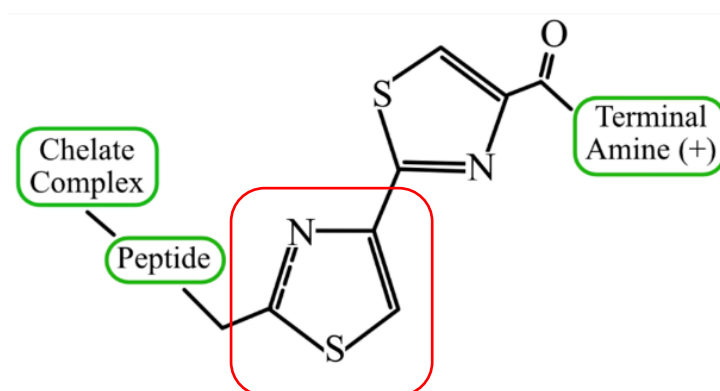


Figure 3.8: **Structure of phleomycin.** The left side of the molecule is involved in metal chelation and the bithiazole moiety (encircled in red) is important in DNA binding. Phleomycin differs from bleomycin by the absence of a double bond (shown in dotted line). Adapted from (Povirk, 1996).

Results

Strains H53 (WT) and H2655 ($\Delta orc9$) were treated with different doses of gamma radiation to test their sensitivity to ionising radiation-induced DNA damage (See *Chapter 2: Materials and Methods*). At the lowest gamma radiation dose (375 Gy), there was no significant difference between the survival fractions of the two strains. At doses 750 Gy and 1500 Gy, the $\Delta orc9$ strain was significantly more sensitive than the WT strain, as the error bars of each data point do not overlap (Figure 3.9).

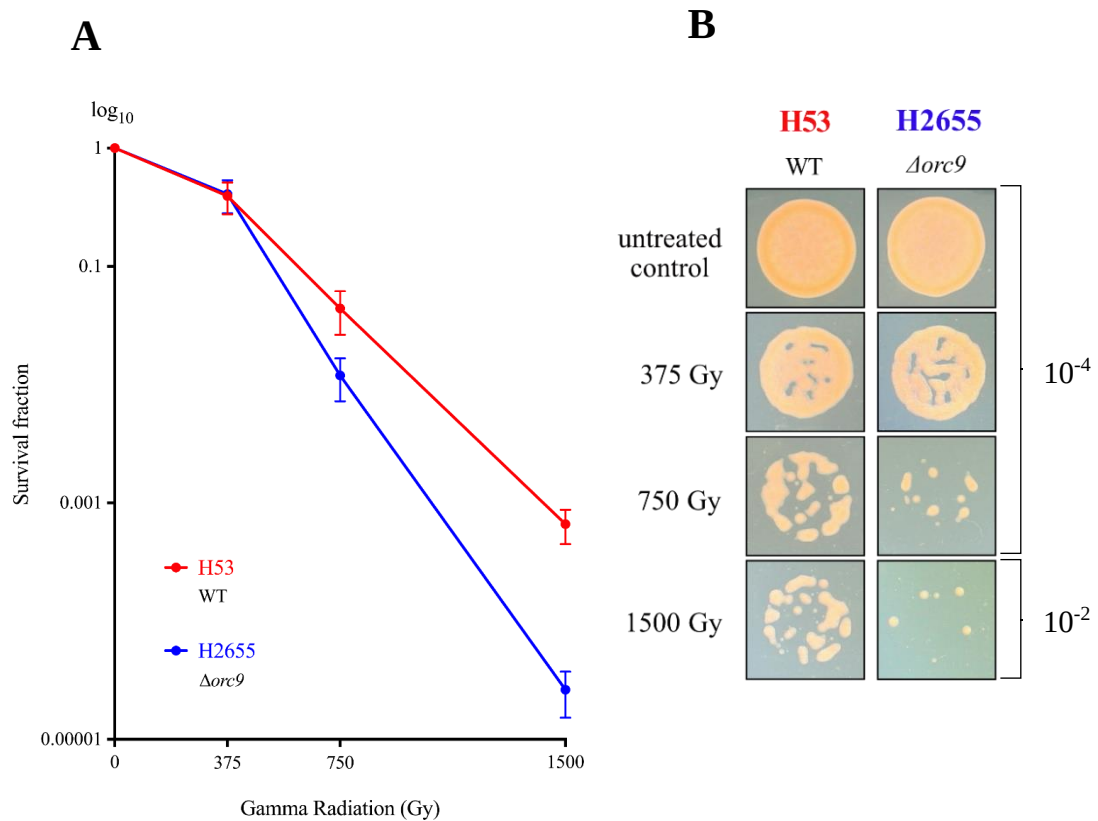


Figure 3.9: Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains after exposure to different doses of gamma radiation. (A) Data points show the mean of three independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a log₁₀ scale. (B) Cell viability is measured by counting colonies after 3-5 days of incubation at 45 °C. Figure shows dilution of 10⁻⁴ for doses 0, 375 and 750 Gy. At the highest dose of 1500 Gy, 10⁻⁴ dilution did not produce any colonies so the figure shows the colonies counted from the 10⁻² dilution.

Strains H53 (WT) and H2655 ($\Delta orc9$) were treated with different concentrations of phleomycin for 1 hour to test their sensitivity to phleomycin-induced DNA damage (See *Chapter 2: Materials and Methods*). Consistent with the results from the gamma radiation sensitivity assay, the $\Delta orc9$ strain was more sensitive to phleomycin than the wt strain at all treatment doses (Figure 3.10). The difference between the survival fractions of H53 and H2655 at each phleomycin concentration is significant, as the error bars of each data point do not overlap.

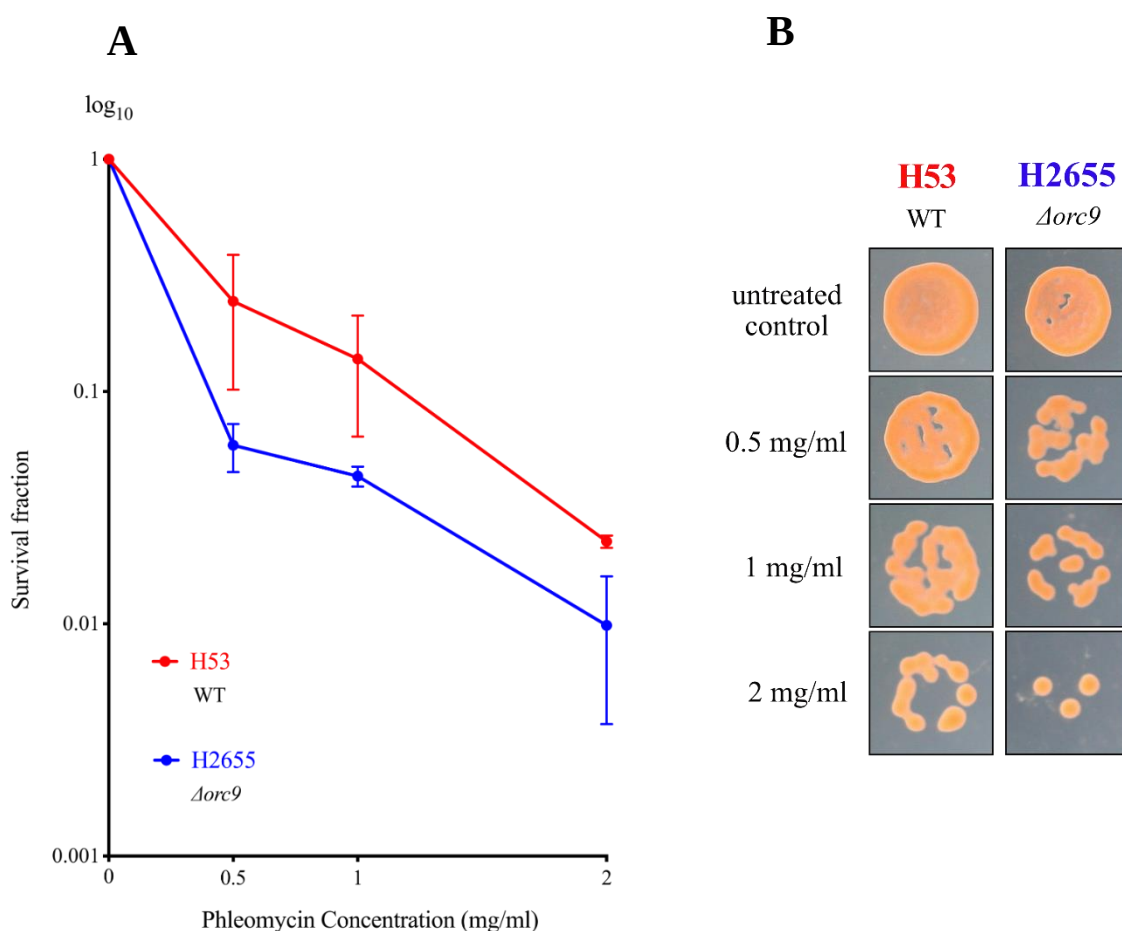


Figure 3.10: Comparison of the survival fractions of H53 (WT) and H2655 ($\Delta orc9$) strains after 1 hour treatment with phleomycin. (A) Data points show the mean of two independent experiments. Error bars represent standard error of the mean. Survival fractions are given on a log₁₀ scale. (B) Cell viability is measured by counting colonies after 3-5 days of incubation at 45 °C. Figure shows the dilution of 10⁻⁴ for all doses of phleomycin.

3.3.4 Growth assays Mitomycin C Sensitivity Test

Background

Mitomycin C (MMC) is a bifunctional alkylating agent that has two electrophilic sites that can react with two nucleophilic sites in DNA (Tomasz, 1995). MMC produces interstrand cross-links when these two sites are on the opposite DNA strands (Figure 3.11). Interstrand cross-links are an important class of DNA lesions because they stop the double helix from unwinding, completely blocking transcription and DNA replication (Tomasz *et al.*, 1987). Both strands need to be repaired for this type of lesion to be fixed. Interstrand crosslinks in DNA are repaired by a combination of NER and homologous recombination.

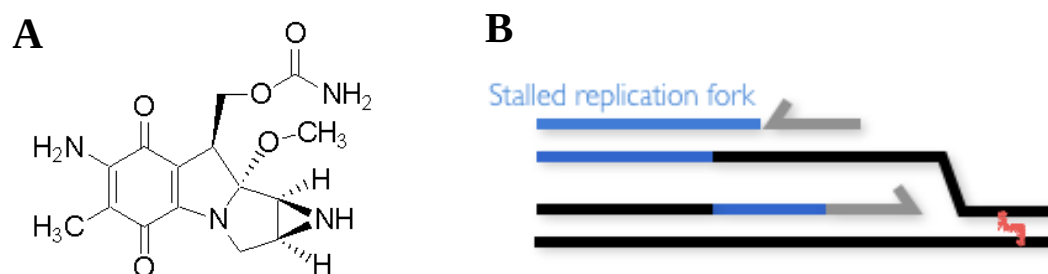


Figure 3.11: **Mitomycin C.** (A) Chemical structure of the mitomycin C molecule. (B) Diagram showing a stalled replication fork for upstream of an interstrand crosslink lesion.

Results

Strains H53 (WT) and H2655 ($\Delta orc9$) were grown for 72 hours under different concentrations of MMC. At a concentration of 0.05 $\mu\text{g/ml}$ MMC, there is a substantial difference between the growth curves of the two strains as depicted on Figure 3.12.

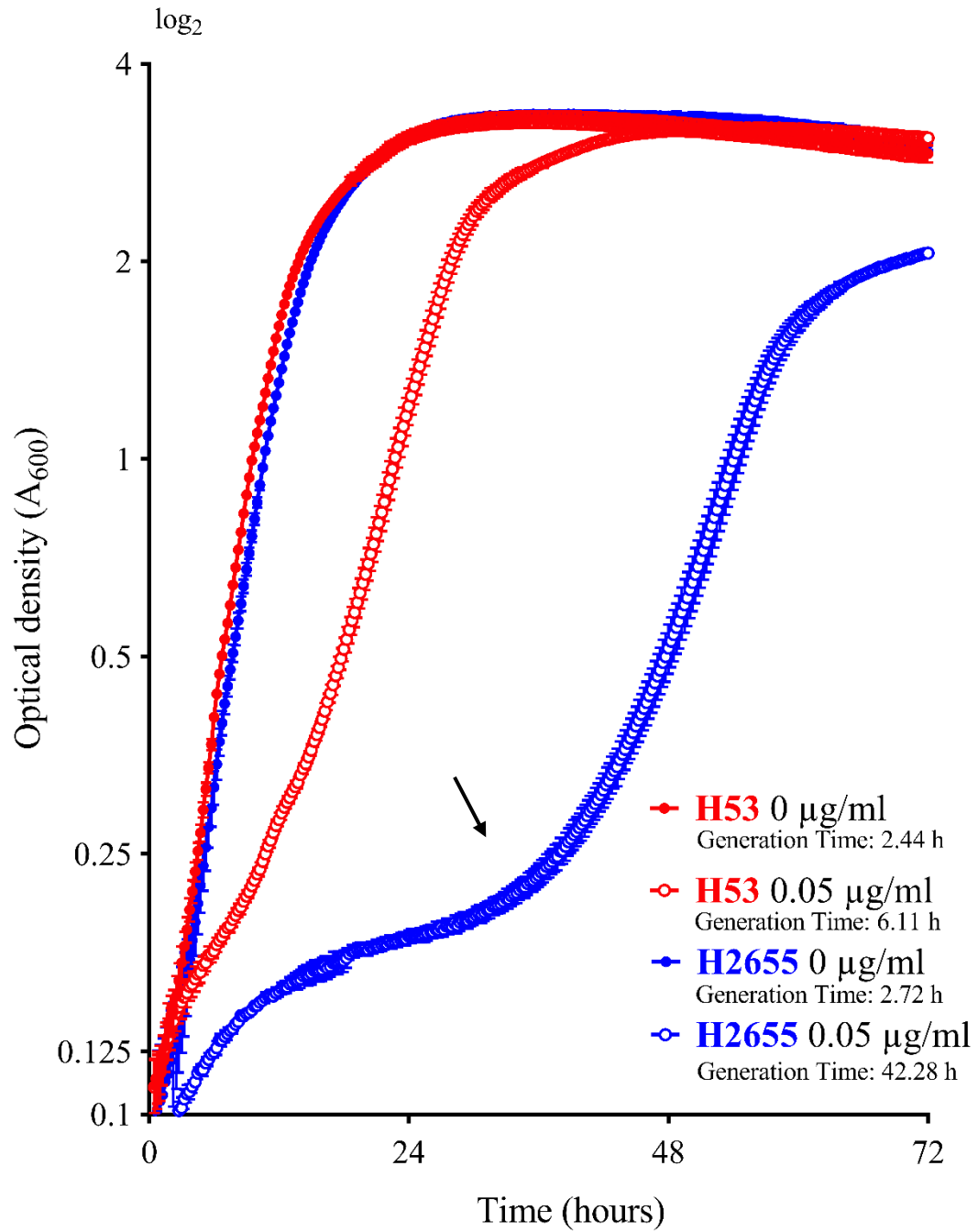


Figure 3.12: **Growth curves of H53 (WT) and H2655 (*Aorc9*) under increasing concentrations of MMC.** Optical density is given on a $\log_2$ scale. Error bars represent the standard error of the mean of five repeats. Generation times of each curve are given on graph. The arrow indicates the point where MMC presumably degrades.

3.4 Discussion

MMS and H₂O₂ Sensitivity Assays

It is clear from the MMS sensitivity assay that MMS-induced DNA damage repair in *Haloferax volcanii* is Orc9-dependent. This suggests that Orc9 may have a role in BER. However, no increased sensitivity to H₂O₂ was detected in the Δ *orc9* strain compared to the wild type. This is supported by the growth assay result where the growth curves for H53 (WT) and H2655 (Δ *orc9*) have similar generation times under increasing doses of H₂O₂. This suggests that oxidative damage in *Haloferax volcanii* is repaired or prevented efficiently via an Orc9-independent pathway. This might be because the enzymatic inactivation of H₂O₂ in *Haloferax volcanii* is efficient. It might also be due to an alternative oxidative damage repair pathway that Orc9 does not have a role in.

UV and 4-NQO Sensitivity assays

The Δ *orc9* strain exhibited hypersensitivity to both UV and 4-NQO induced DNA damage. This suggests that the repair of these type of DNA damage is Orc9-dependent. NER is the predominant pathway used to repair UV and 4-NQO induced damage; however, indirect effects of UV and 4-NQO result in DSBs which require the involvement of homologous recombination for repair. Orc9 might have a role in any or both of these pathways of DNA repair. To distinguish between these pathways, a mutagen was used that predominantly induces DSBs.

Gamma Radiation and Phleomycin Sensitivity Assays

Gamma radiation and phleomycin sensitivity assays collectively showed that the Δ *orc9* strain is hypersensitive to both of these mutagens. This suggests that Orc9 must have a significant role in the repair of DSBs. Phleomycin induces predominantly DSBs and so it can be concluded that in *Haloferax volcanii*, Orc9 plays a role in homologous recombination.

MMC Growth Assay

MMC retards the growth of the H2655 (Δ *orc9*) more than it does H53 (WT), when grown with 0.05 µg/ml MMC. This suggests that Orc9 has a role in interstrand crosslink repair. Interstrand crosslinks require the involvement of both NER and HR

for repair. Considering Orc9 is involved in NER and HR, it is no surprise that the *Δorc9* strain was severely sensitive to MMC. Interstrand crosslinks stall replication forks if unrepaired before the next round of replication. Orc9 might be involved in a pathway involved in restarting replication forks.

Chapter 4: Interactions of Orc9

4.1 Background

The genome of *Haloferax volcanii* encodes sixteen Orc proteins, two of which are conserved in all Archaeal species: Orc1 (Cdc6/Orc1-1) and Orc9 (Cdc6/Orc1-2). Orc1 is always linked to the main replication origin (*oriC1*) and conversely, Orc9 is never located next to an origin (Raymann *et al.*, 2014). Unlike Orc1, Orc9 does not have a role in replication initiation as it does not alter the replication profile of *H. volcanii* (See Chapter 1: Introduction). Preliminary data implicated a role in DNA repair as deletion of *orc9* results in acute sensitivity to MMC and UV-induced DNA damage (unpublished data). As discussed in Chapter 3, deletion of *orc9* increases sensitivity of *H. volcanii* to MMS, gamma radiation and phleomycin, as well as MMC and UV radiation. This suggests that Orc9 has a role in multiple DNA damage repair pathways.

4.2 Aims

This chapter will further explore the role of Orc9 in DNA repair and aim to:

- Construct inducible alleles of the *orc9* gene and use this to replace the native *orc9* gene in strains with and without DNA replication origins. Growth assays will determine whether the level of *orc9* expression affects the growth rate of *H. volcanii* cells, under DNA damage with mitomycin-C, and whether this differs between wild-type and originless strain backgrounds.
- Explore the synergistic genetic interaction of Orc9 and RadA using strains where *radA* is under the control of a tryptophan inducible promoter, *p.tnaA*. Tryptophan gradient plates will examine the growth defects of these strains by carrying out gene depletion.
- Tag Orc9 with epitope tags and express in *H. volcanii* under the control of its native promoter. The purified Orc9 will be examined for interactions with other proteins by co-purification.

4.3 Results

4.3.1 Tryptophan Inducible Orc9

Construction of *p.tnaAM3-orc9* strains

The plasmid pTA2074 (constructed by Victoria Smith, Figure 4.1) was transformed into strains H26 ($\Delta pyrE2$), H938 ($\Delta pyrE2$, $\Delta fen1$) and H2656 ($\Delta pyrE2$, $\Delta orc3$, $\Delta ori-pHV4-2$, $\Delta orc2$, $\Delta oriC3$, $\Delta orc1$, $\Delta oriC1$, $\Delta orc5$, $\Delta oriC2$) to generate pop-in strains: H4054, H4055 and H4056 respectively (Victoria Smith, unpublished data).

(A)

<i>tnaA</i>	CCTGCCGATTACTTCACATTTCGCGGACCTATTGCGCA	37
<i>tnaAM3</i>	CCTGCCGATGACTTCACATTTCGCGGACCTATTGCGCA	37

(B)

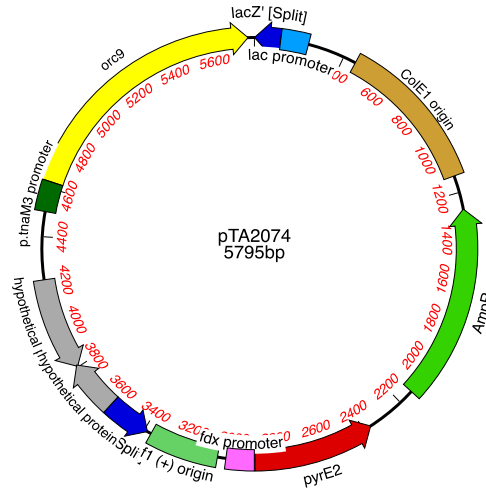


Figure 4.1: **p.tnaAM3 promoter.** (A) TATA box sequence of *tnaA* and *tnaAM3* promoters showing point mutation from T > G, respectively. (B) pTA2074, gene replacement construct with *orc9* under the control of the *p.tnaAM3* promoter (Constructed by Victoria Smith, 2017).

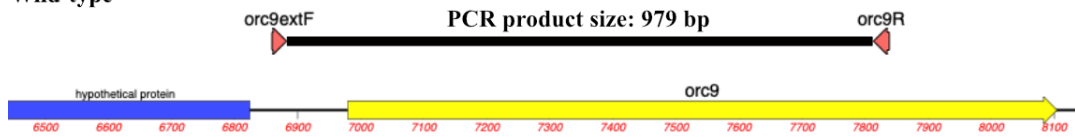
p.tnaA is a native *H. volcanii* promoter and it regulates the expression of the tryptophanase gene, *tnaA*. The tryptophanase protein degrades excess tryptophan in cells. As tryptophan is an energetically costly amino acid for cells to make, the *tnaA* gene is tightly regulated by its native *p.tnaA* promoter. This makes *p.tnaA* a useful genetic tool to study the effects of different transcript levels of genes of interest (Large

et al., 2007). Alternatively, the *p.tnaAM3* promoter is used when lower expression levels of a gene are required. *p.tnaAM3* differs from *p.tnaA* by a single point mutation (Figure 4.1), which reduces its activity by ~50%.

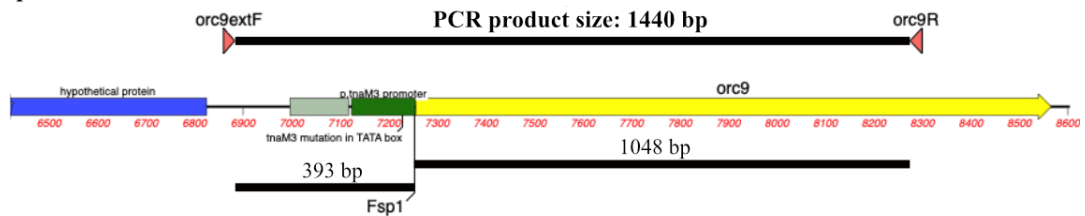
The pop-in strains H4054 ($\Delta pyrE2$, $orc9+::[p.tnaM3-orc9+ pyrE2+]$) and H4056 ($\Delta pyrE2$, $\Delta orc3$, $\Delta ori-pHV4-2$, $\Delta orc2$, $\Delta oriC3$, $\Delta orc1$, $\Delta oriC1$, $\Delta orc5$, $\Delta oriC2$, $orc9+::[p.tnaM3-orc9+ pyrE2+]$) were cultured for 4 overnights in YPC broth to allow for recombination events (See Chapter 2: *Materials and Methods*) to replace the native *orc9* promoter with *p.tnaAM3*. Pop-out cultures were plated on Hv-Ca + 5-FOA plates and colonies formed were patched on Hv-YPC plates to relieve selection pressure. To screen for the strains with the *p.tnaAM3-orc9* allele, colony PCR using primers *orc9extF* and *orc9R* was performed (Figure 4.2). A follow-up digest with *Fsp1* was performed with the colony PCR products of the pop-out candidate patches. Finally, a Southern blot confirmed the strains H4414 and H4419 (Figure 4.3).

(A)

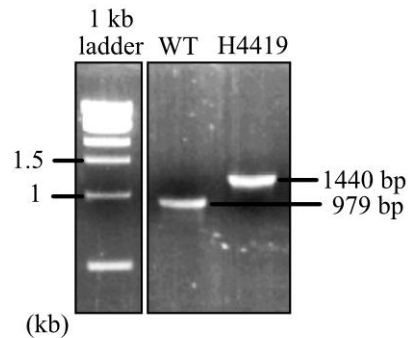
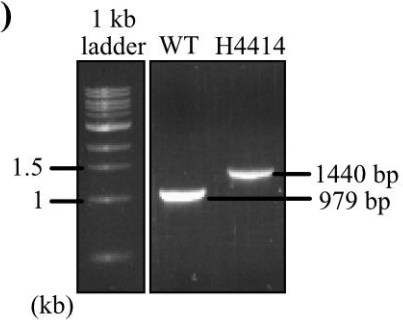
Wild-type



p.tnaAM3-orc9



(B)



(C)

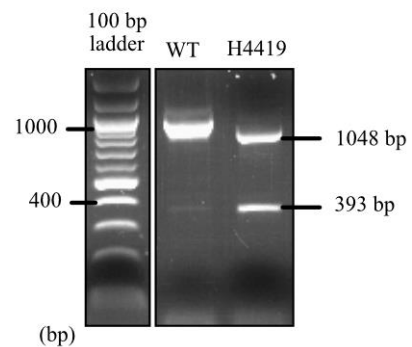
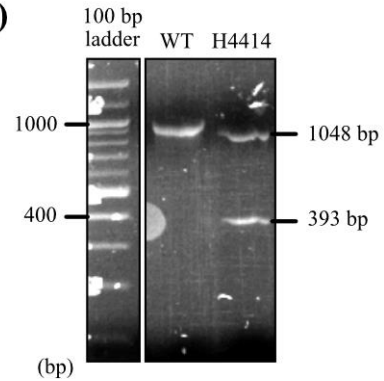
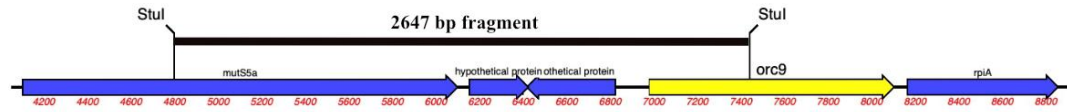


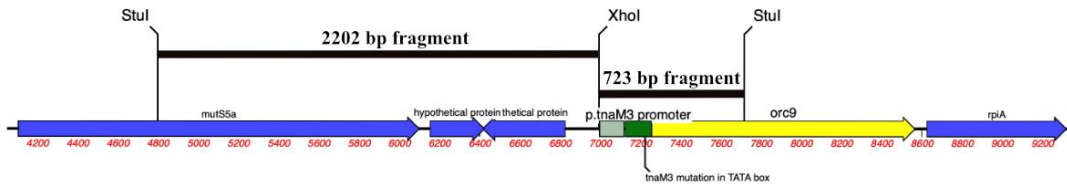
Figure 4.2: *p.tnaAM3-orc9* strains checked by colony PCR and *FspI* digest. (A) Predicted product sizes of wild-type (979 bp) and *p.tnaAM3-orc9* (1440 bp) strains after colony PCR using primers *orc9extF* and *orc9R*. (B) Agarose gel images showing expected band sizes of the colony PCR products of *p.tnaAM3-orc9* strains H4414 and H4419 compared to WT. (C) Agarose gel images showing the *FspI* digest of the colony PCR products. Wild type *orc9* does not have an *FspI* site (undigested product).

(A)

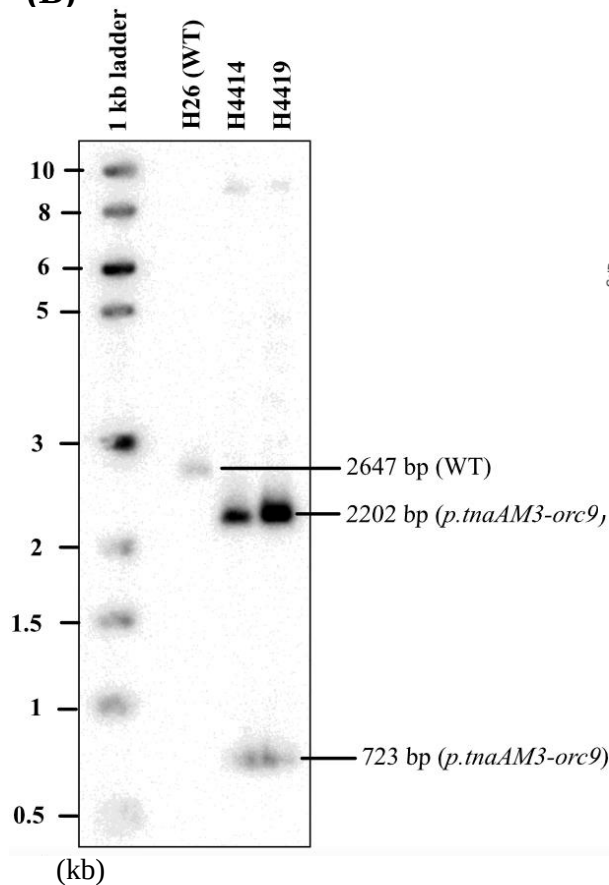
Wild-type



p.tnaAM3-orc9



(B)



(C)

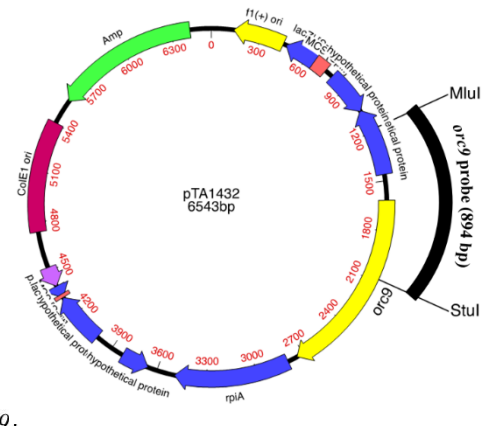


Figure 4.3: *p.tnaAM3-orc9* strains checked by Southern blot. (A) Predicted fragment sizes of wild-type (2647 bp) and *p.tnaAM3-orc9* strains (2202 bp and 723 bp) with *StuI* and *XhoI* digests of genomic DNA. (B) Southern blot showing *p.tnaAM3-orc9* strains H4414 and H4419 compared to wild-type, giving the expected band sizes (indicated on figure). (C) Probe used in Southern blot detection was *MluI* and *StuI* digested fragment of pTA1432 (genomic clone of *orc9*) (894 bp).

Growth Assays

Prior to setting up the growth assay using inducible *orc9* strains, a mitomycin C potency test was necessary. The mitomycin C (MMC) used (from Sigma-Aldrich) comes in powder form and different batches of MMC proved to have different potency levels. A test growth assay was performed to determine the dose at which wild-type and $\Delta orc9$ strains show a clear difference between growth rates without resulting in too much cell death (Figure 4.4). The dose 0.05 $\mu\text{g/ml}$ was picked for the growth assay with inducible *orc9* strains.

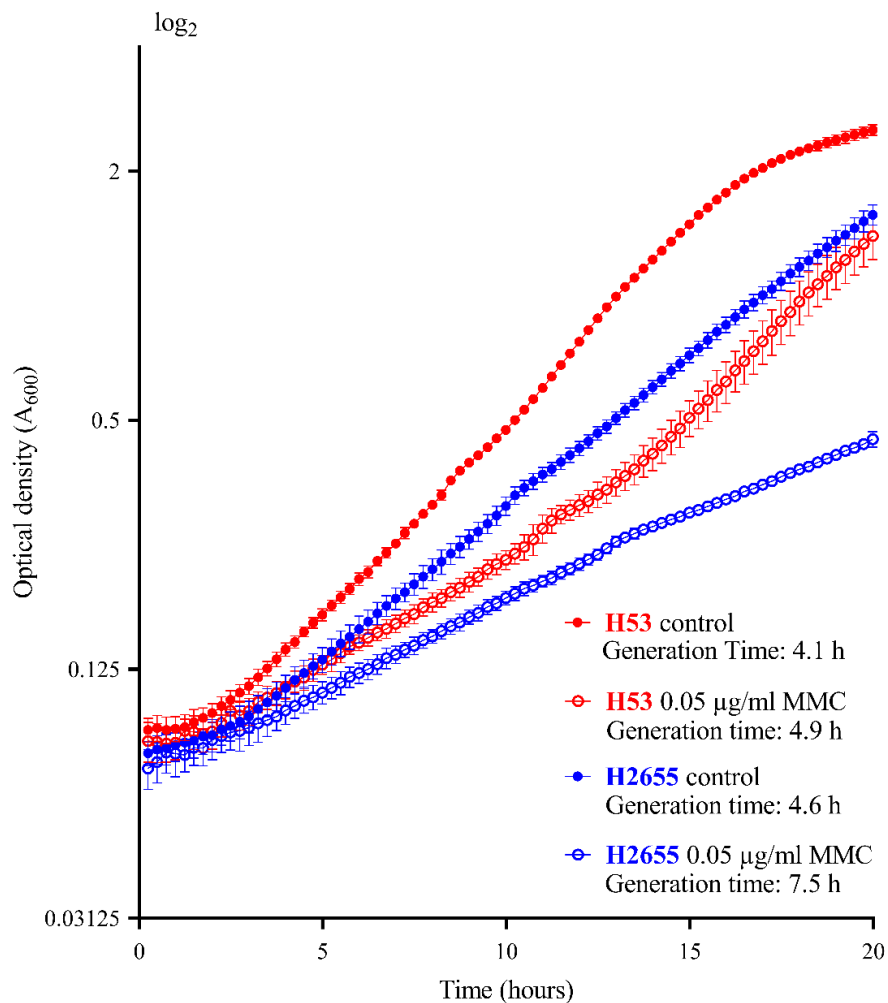


Figure 4.4: **Growth assay for dose optimisation of the new MMC batch.** Growth of strains H53 (*orc9+*) and H2655 ($\Delta orc9$) was measured with different levels of MMC in the media. A dose of 0.05 $\mu\text{g/ml}$ was chosen for further use. Both strains ($n=5$) were incubated on the same 96 well plate and measured simultaneously using an Epoch2 Microplate Spectrophotometer (BioTek). Calculated generation times are given next to the corresponding strain on graph. Error bars represent standard error of the mean (SEM). Optical density (A_{600}) is given on a $\log_2$ scale.

Strains H26 ($\Delta pyrE2$), H2132 ($\Delta pyrE2$, $\Delta orc9$), H2656 ($\Delta pyrE2$, $\Delta orc3$, $\Delta ori-pHV4-2$, $\Delta orc2$, $\Delta oriC3$, $\Delta orc1$, $\Delta oriC1$, $\Delta orc5$, $\Delta oriC2$), H2728 ($\Delta pyrE2$, $\Delta orc3$, $\Delta ori-pHV4-2$, $\Delta orc2$, $\Delta oriC3$, $\Delta orc1$, $\Delta oriC1$, $\Delta orc5$, $\Delta oriC2$, $\Delta orc9$) were used as controls in these growth assays along with the inducible strains H4414 ($\Delta pyrE2$, $p.tnaM3-orc9+$) and H4419 ($\Delta pyrE2$, $\Delta orc3$, $\Delta ori-pHV4-2$, $\Delta orc2$, $\Delta oriC3$, $\Delta orc1$, $\Delta oriC1$, $\Delta orc5$, $\Delta oriC2$, $p.tnaM3-orc9+$). These strains were grown in Hv-Ca + ura broth for 2 consecutive overnights. Second overnights were diluted into fresh Hv-Ca + ura +/- trp (0.25, 0.5, 1 mM) and grown until they reached an OD₆₅₀ of about 0.05. Using an Epoch2 Microplate Spectrophotometer (BioTek), growth rate of these strains was measured under control (no MMC) and mitomycin C (0.05 μ g/ml) treatment conditions.

Results show that growth rates of strains H26 and H2132 under MMC and control conditions does not change with changing tryptophan concentration in the media, as expected (data not shown). As depicted in Figure 4.5, these results show that under +/- MMC conditions, H4414 strain grown with 0 mM tryptophan produced similar generation times as that of H2132. This suggests that when there is no tryptophan in the media, $p.tnaAM3$ promoter is completely turned off and $orc9$ is not expressed, hence the growth resembles the deletion phenotype. When there is 0.25 mM tryptophan in the media, growth rate of H4414 under MMC treatment is faster than when there is no tryptophan in the media (compare 6.8 h generation time to 9 h). However, it is not as fast as the growth rate of the H26 strain treated with MMC, where $orc9$ is expressed under its native promoter. This suggests that 0.25 mM of tryptophan is not enough to induce wildtype levels of $orc9$. This experiment was repeated with 0.5 mM and 1 mM tryptophan in the media. No significant increase in the growth rates were seen under MMC treatment (compared to 0.25 mM levels) when the tryptophan concentrations were increased (data not shown).

Growth rates of strains H2656 and H2728 under MMC and control conditions does not change with changing tryptophan concentration in the media, as expected (data not shown). As shown in Figure 4.6, H4419 ($p.tnaAM3-orc9$) strain grown with 0 mM tryptophan produced similar generation times as that of H2728 ($\Delta orc9$). However, there seems to be no increase in the growth rate of MMC treated H4419 strain in the presence of tryptophan (0.25 mM). Increasing tryptophan concentration did not increase the growth rate of H4419 under MMC conditions (data not shown).

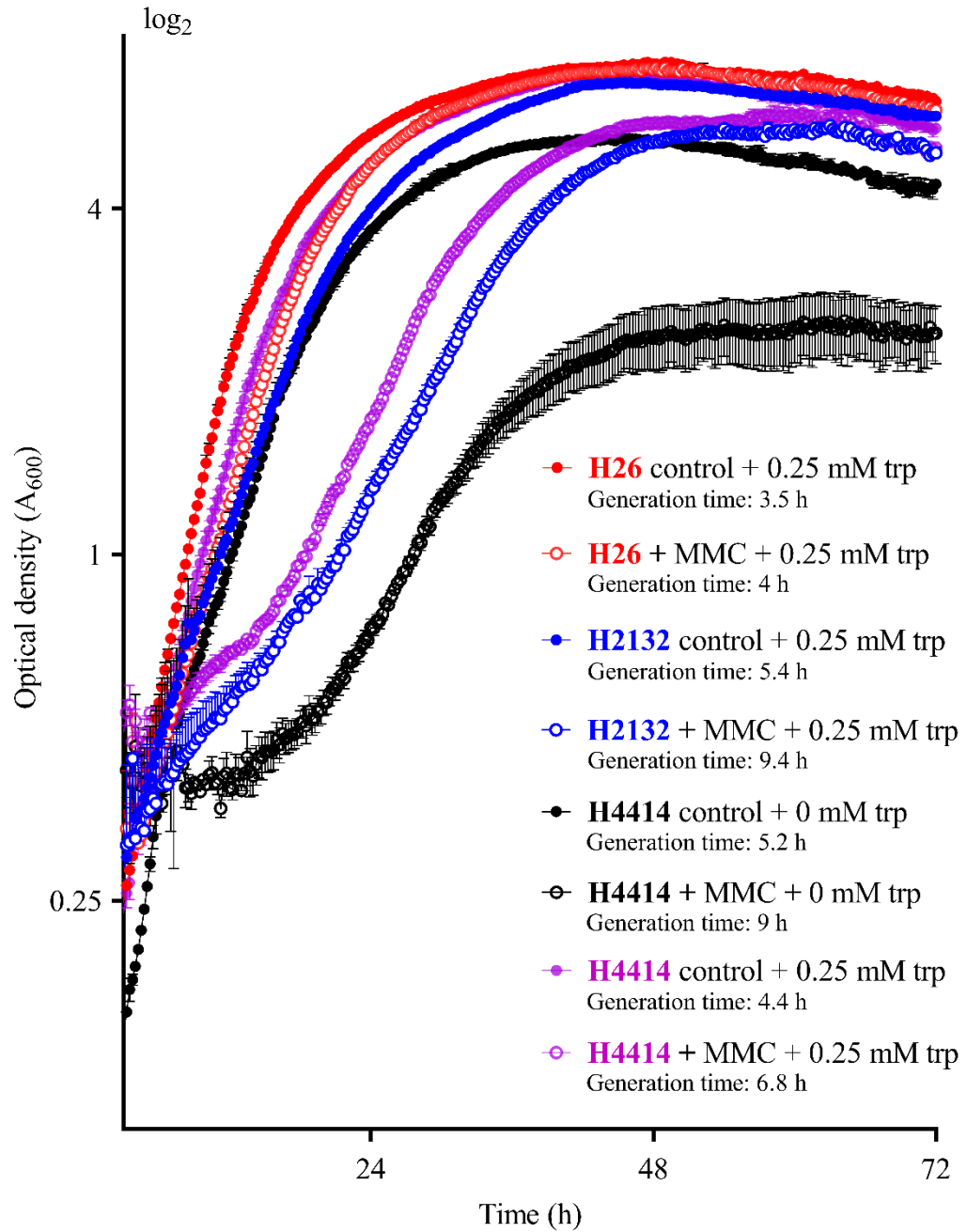


Figure 4.5: Exponential growth rates of the *p.tnaAM3-orc9* strain (H4414) in the presence of 0.25 mM tryptophan under +/- MMC in media. Growth of strains H26 (wild-type), H2132 ($\Delta orc9$) and H4414 (*p.tnaAM3-orc9*) was measured with 0.05 $\mu\text{g/ml}$ MMC (or no MMC for controls) in the media. Strains ($n=2$) were incubated on the same 96 well plate and measured simultaneously using an Epoch2 Microplate Spectrophotometer (BioTek). Calculated generation times are given next to the corresponding strain on graph. Error bars represent standard error of the mean of two repeats. Optical density (A_{600}) is given on a $\log_2$ scale.

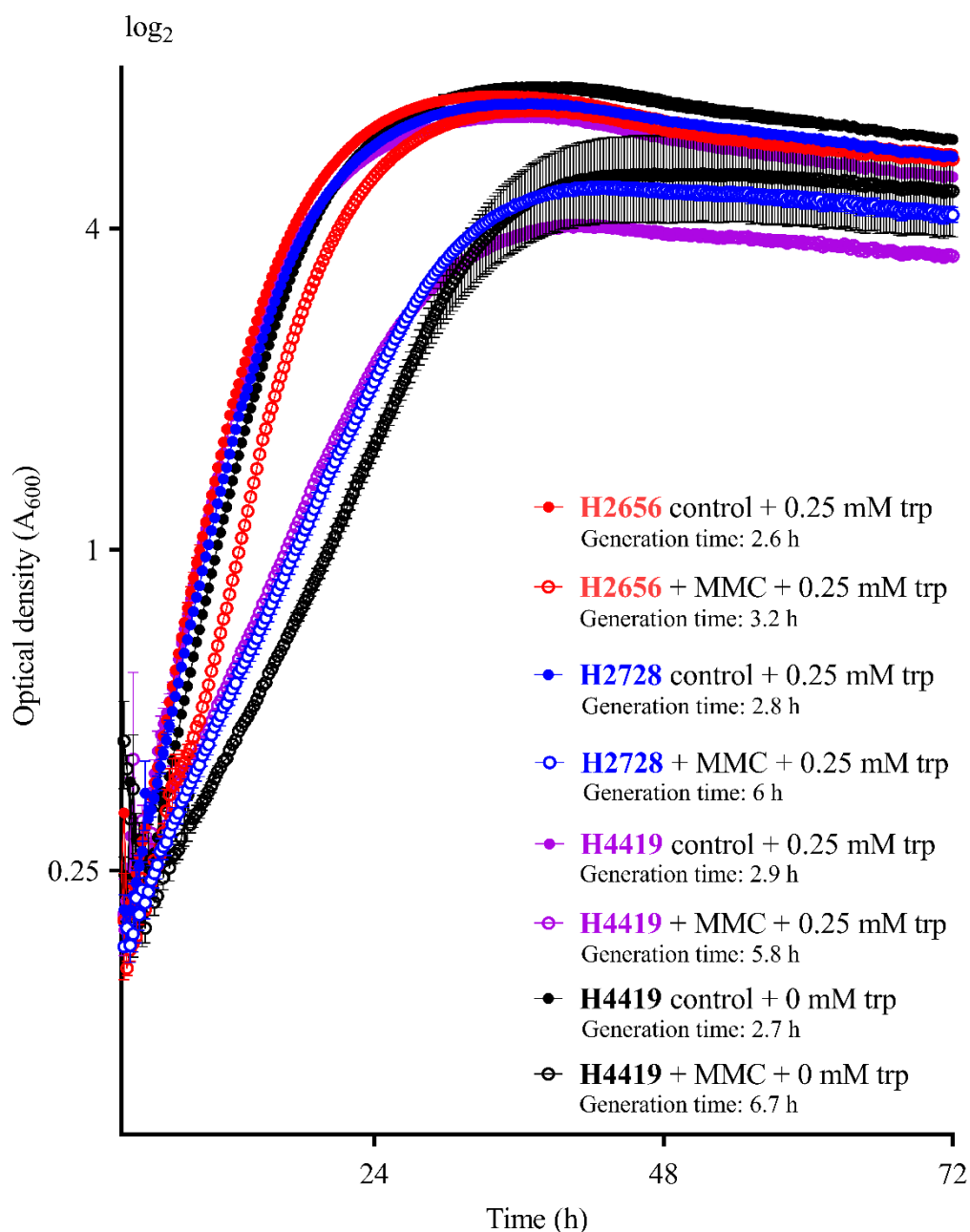


Figure 4.6: Exponential growth rates of the *p.tnaAM3-orc9* strain (H4419) in the presence of 0.25 mM tryptophan under +/- MMC in media. Growth of strains H2656, H2728 and H4419 was measured with 0.05 μ g/ml MMC (or no MMC for controls) in the media. Strains (n=2) were incubated on the same 96 well plate and measured simultaneously using an Epoch2 Microplate Spectrophotometer (BioTek). Calculated generation times are given next to the corresponding strain on graph. Error bars represent standard error of the mean of two repeats. Optical density (A₆₀₀) is given on a log₂ scale.

4.3.2 *p.tnaA-radA* strains on tryptophan gradient plates

Strains were painted onto tryptophan gradient plates (See Methods) to examine the requirement for *radA* in different *H. volcanii* strain backgrounds. Strains H26 (Δ *pyrE2*), H53 (Δ *pyrE2*, Δ *trpA*) and H2132 (Δ *pyrE2*, Δ *orc9*) were used as controls (top 3 strains on Figure 4.7). The *radA* gene is under the control of the tryptophan inducible promoter, *p.tnaA* in strains H1637 (Δ *pyrE2*, *oriC*⁺), H1642 (Δ *pyrE2*, Δ *oriC1*, Δ *oriC2*, Δ *oriC3*, Δ *ori-pHV4*) and H3414 (Δ *pyrE2*, Δ *orc9*) (all constructed by Michelle Hawkins, (Hawkins et al., 2013)).

Control strains grew as expected on the tryptophan gradient plates. H26 (Δ *pyrE2*) and H2132 (Δ *pyrE2*, Δ *orc9*) are viable across the plate regardless of tryptophan concentration, as they are both *trp*⁺. H53 (Δ *pyrE2*, Δ *trpA*) grew well in the presence of tryptophan and its growth stopped as the tryptophan levels came to a zero, as expected. This is because H53 has the *trpA* gene deleted, which encodes tryptophan synthase for cells to produce the amino acid tryptophan. H1637 (Δ *pyrE2*, Δ *hdrB* *p.tnaA-radA*⁺::*hdrB*⁺) grew well in the presence of tryptophan. As the tryptophan levels depleted, there were less colonies and the size of the colonies were smaller. This phenotype resembles the Δ *radA* deletion phenotype where cells grow poorly. H1642 (Δ *pyrE2*, *p.tnaA-radA*⁺::*hdrB*⁺, Δ *oriC2*, Δ *oriC1*, Δ *oriC3*, Δ *hdrB*, Δ *ori-pHV4-2*), however, showed no growth in the absence of tryptophan. This result repeated what was shown before and confirmed that *radA* is essential in strains where all replication origins are deleted, because cells lacking origins rely on RadA recombinase to initiate replication (Hawkins et al., 2013). H3414 (Δ *pyrE2*, Δ *orc9*, *p.tnaA-radA*⁺) grew well in the presence of tryptophan, but cells grew poorly as the tryptophan levels depleted. There was some growth in the absence of tryptophan, but the colonies only appeared after 7-10 days of incubation at 45°C and were very small. On the 0 mM *trp* side of the plates, the colony sizes of H1637 was a lot bigger than the H3414. This suggests that there is an increased requirement for *radA* in the Δ *orc9* background. However, the phenotype is not lethal since there are some colonies forming where there is no *orc9* or *radA*. This synthetic growth defect phenotype suggests that *Orc9* and *RadA* have roles in alternative pathways of DNA repair and cells need at least one of them to grow as normal.

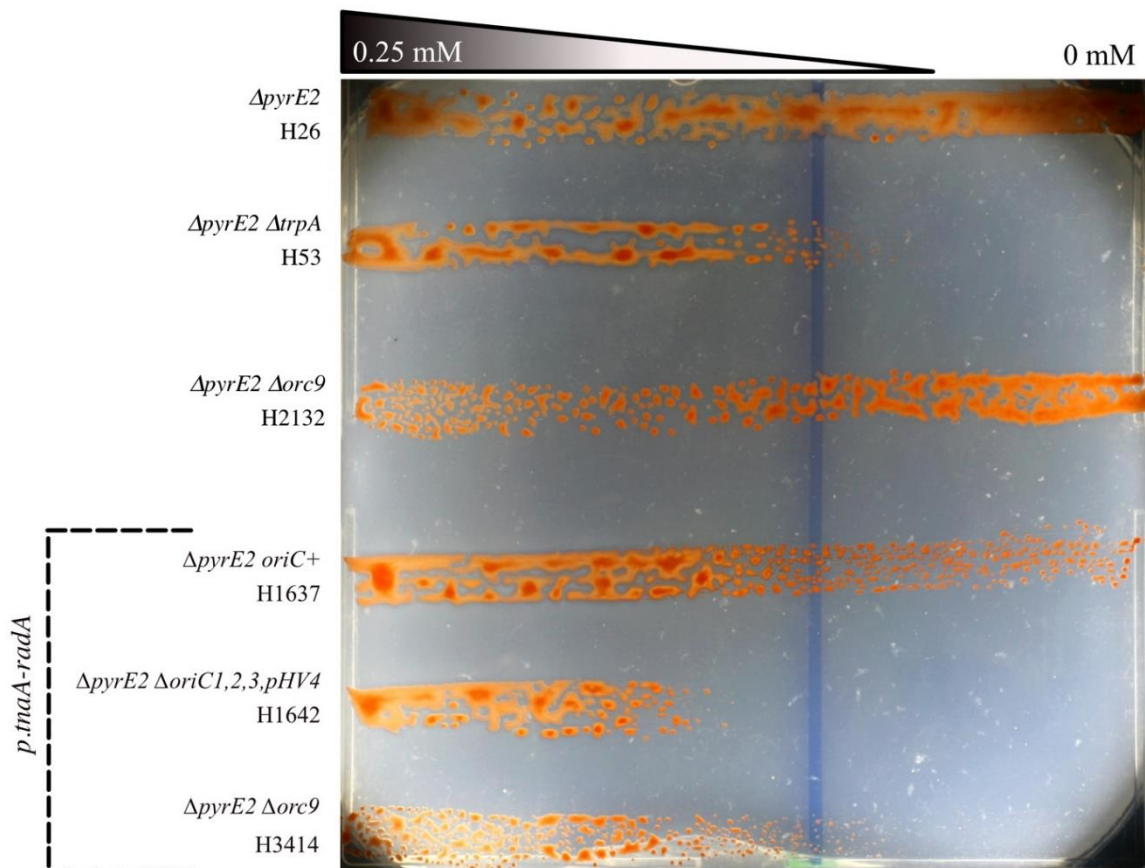


Figure 4.7: *p.tnaA-radA* strains on a tryptophan gradient plate. Strains H26, H53, H2132, H1637, H1642 and H3414 were all grown in YPC broth for 2 overnights and grown to A_{600} of ≈ 0.6 . All cultures were diluted to 10^{-4} in 18% SW before being painted onto plates. 0.25 mM was used as the highest tryptophan concentration on the gradient plates.

4.3.4 Orc9 and FEN1 relationship

The pop-in strain H4055 ($\Delta pyrE2$, $\Delta fen1$, $orc9+::[p.tnaM3-orc9+ pyrE2+]$) was constructed by Victoria Smith by transforming H938 ($\Delta pyrE2$, $\Delta fen1$) with pTA2074 (contains the *p.tnaM3-orc9* allele, unpublished data). H4055 was cultured for 4 overnights in YPC broth in attempt to delete the wild-type *orc9* gene and replace it with the *p.tnaM3-orc9* allele in a $\Delta fen1$ deletion background. The pop-out culture was plated on Hv-Ca + 5-FOA plates. Colonies grown were patched onto YPC plates and screened by colony PCR using the same primers: *orc9extF* and *orc9R*. Using the same screening strategy as before, 50 pop-out clones were screened, and all were wild-type (Figure 4.7).

The success rate of deletion of *radA* in a background where all but one DNA replication origin is deleted is 1/70 (Thorsten Allers, unpublished). This is a reference for how difficult a gene is to delete (since *radA* cannot be deleted in originless strains). The pop-out success rates for the pop-in strains H4054 (Δ *pyrE2*, *orc9*::[*p.tnaM3-orc9*+ *pyrE2*]) and H4056 (Δ *pyrE2*, Δ *orc3*, Δ *ori-pHV4-2*, Δ *orc2*, Δ *oriC3*, Δ *orc1*, Δ *oriC1*, Δ *orc5*, Δ *oriC2*, *orc9*::[*p.tnaM3-orc9*+ *pyrE2*]) were 5/40 and 5/20 respectively. Comparing this to 0/50, it is possible that replacing the native *orc9* promoter with *p.tnaM3-orc9* in a Δ *fen1* deletion background is difficult because a Δ *fen1* Δ *orc9* double deletion shows a synthetic lethality phenotype. Even repeating the pop-outs with increased tryptophan concentrations (0.5 and 0.75 mM), did not produce successful pop-out clones. A different strain construction strategy should be used in order to generate a Δ *fen1* strain where *orc9* is under the control of *p.tnaM3* promoter.

4.3.3 Epitope-tagged Orc9

The *orc9* genomic clone, pTA1432, was constructed by Thorsten Allers prior to this study (unpublished data) (Figure 4.8).

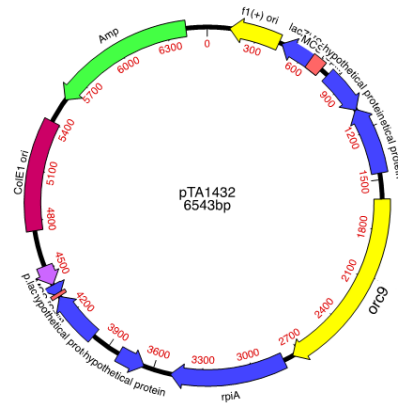
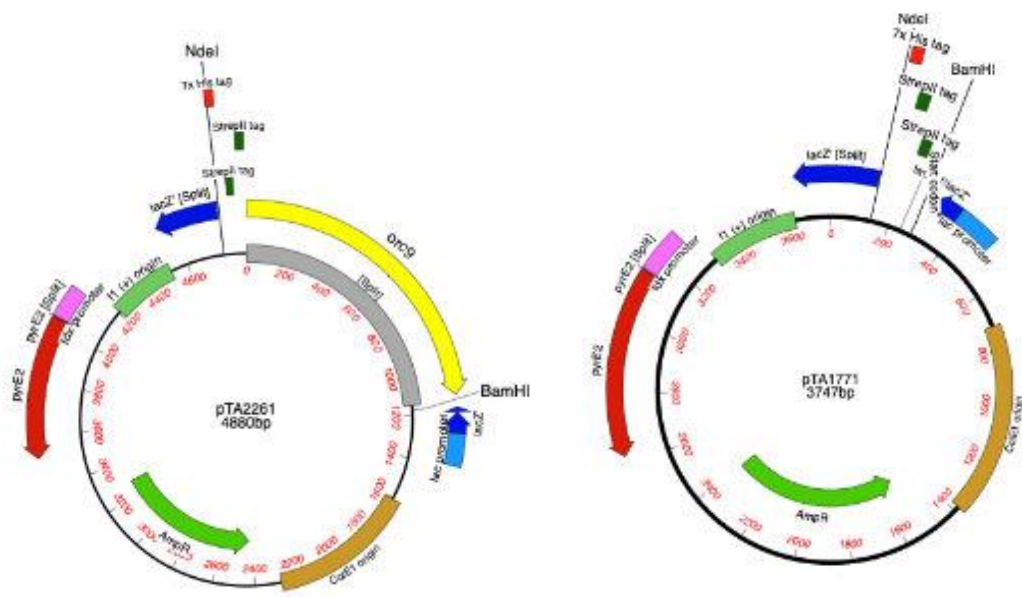


Figure 4.8: **pTA1432**. Genomic clone of *orc9* (HVO_0194). Previously constructed by Thorsten Allers (unpublished data).

pTA1432 was used as a template to PCR amplify the *orc9* gene using primers *orc9BspF* and *orc9EcoR* introducing *BspHI* and *EcoRI* restriction sites respectively. This PCR product was digested with *BspHI* and *EcoRI* and ligated into the *PciI* and *EcoRI* sites in pTA1771 to generate the plasmid pTA2261 (Figure 4.9). *BspHI* and *PciI* restriction enzymes produce compatible cohesive overhangs and ligation results in the inactivation of both restriction sites.

(B) pTA1771



(C) *Bam*HI *Nde*I digest

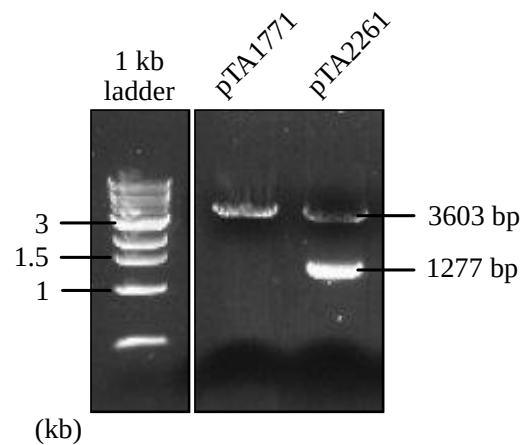


Figure 4.9: **pTA2261.** (A) Map of pTA2261 construct containing the *His7–StrepII tag* upstream of *orc9* gene. (B) Map of the pTA1771 construct containing the *His7–StrepII tag*. (C) Diagnostic *Bam*HI *Nde*I digest of pTA2261 showed bands at 3603 bp and 1277 bp, as expected.

The second step of the cloning procedure was to add the genomic sequence upstream of *orc9* in pTA2261. This increases the length of homology and the chance for integration of the construct into the chromosome. *KpnI* and *NdeI* digest of pTA1705 (constructed by Laura Mitchell, unpublished data) generated a fragment (783 bp) that was ligated into the *KpnI* and *NdeI* sites in pTA2261 to generate the plasmid pTA2292 (Figure 4.10).

(A) pTA2292

(B) pTA2261

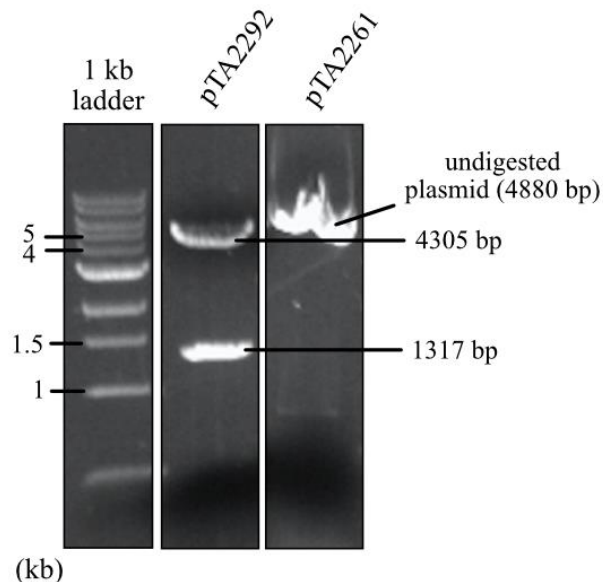
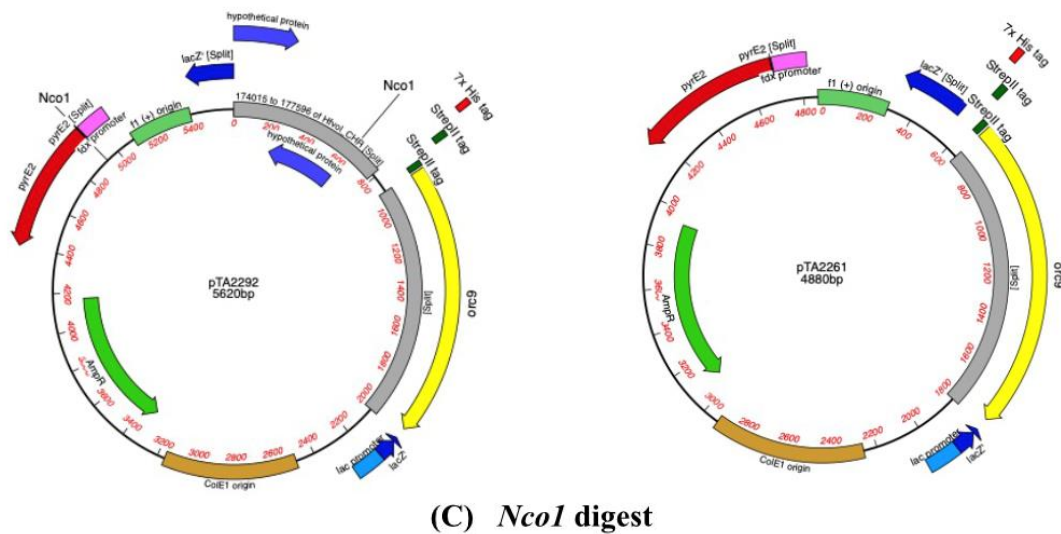


Figure 4.10: **pTA2292**. (A) Map of pTA2292 construct containing a fragment upstream of *His₇ – StreptII tag-*orc9**. (B) Map of the pTA2261 construct containing the *His₇ – StreptII tag* but not an upstream sequence. (C) Diagnostic *NcoI* digest of pTA2292 showed bands at 4305 bp and 1317 bp, as expected. pTA2261 showed an undigested band (no *NcoI* site).

Transformation of H4045 ($\Delta pyrE2 \Delta trpA \Delta mrr$) with pTA2292 gave rise to the pop-in strain H4589. Pop-out of this strain produced colonies on Hv-Ca + trp + 5-FOA plates, however the deletion success rate was very low. Colony PCR using primers *orc9extF* and *orc9R* only identified 2 merodiploid clones out of 110. Instead, H2135 ($\Delta pyrE2 \Delta trpA \Delta orc9::trpA$) was transformed with pTA2370 (*dam*- version of pTA2292). Prior to this transformation, pTA2292 was made *dam*- by passaging it through a *dam*- *Escherichia coli* strain (N2338, GM121), which does not methylate DNA. Unmethylated *dam*- DNA is safe from degradation by the *mrr* endonuclease in *H. volcanii* cells. The transformation yield was low and pop-outs plated on Hv-Ca + trp + 5-FOA plates, but did not give rise to any colonies. Next step should be deleting the *mrr* gene in H2135 ($\Delta pyrE2 \Delta trpA \Delta orc9::trpA$) and then transforming with pTA2292.

4.4 Discussion

Inducible Orc9

The strains used in the growth assays are tryptophan inducible in both WT and originless backgrounds. The work undertaken in this study shows that the level of *orc9* expression induced by the tryptophan in the media is responsible for the increase in growth rate and tolerance to MMC in a wildtype strain background. However, no increase in growth rates were observed in the inducible Orc9 strain in an originless background. Originless strains rely on RadA recombinase to initiate replication via a homologous recombination pathway (Hawkins *et al.*, 2013). The increased amount of homologous recombination taking place in originless strains might repair interstrand crosslinks (which require HR and NER) more efficiently than WT strains in an Orc9-independent fashion. This might be the reason why increased level of *orc9* expression does not result in the betterment of growth rates under MMC treatment. This is supported by the *orc9 radA* synthetic defect phenotype.

Orc9 RadA synthetic defect

$\Delta orc9$ mutants exhibit a severe growth defect when combined with depletion of *radA*, which is worse than depletion of *radA* in an *orc9*+ background. This suggests that Orc9 is involved in an alternative pathway of replication fork restart that avoids recombination (where RadA recombinase is used). Using recombination to restart

stalled replication forks runs the risk of producing genome rearrangements (Kowalczykowski, 2015). Orc9 might have evolved to carry out replication fork restart in a pathway that is independent from RadA, to ensure genomic stability by avoiding recombination. This is supported by the data from Chapter 3 where the $\Delta orc9$ mutant was shown to exhibit sensitivity to gamma radiation and phleomycin, which both induce DSBs and lead to replication fork stalling. Strains lacking *orc9* and *radA* both show hypersensitivity to DSBs when individually deleted. It is not wholly surprising that a synthetic growth defect is produced when they are both deleted in the same strain.

Using pulsed field gel electrophoresis, $\Delta orc9$ mutants were shown to have increased genomic instability when treated with MMC ($\approx 14\%$ genomic rearrangement rate) (Dasha Ausiannikava, Victoria Smith, unpublished data). Future work should include calculating the genomic rearrangement rate for wild type cells treated with MMC. If Orc9 functions in a repair pathway that avoids recombination, the wild type rearrangement rate should be lower than the $\Delta orc9$ rate.

Tagged Orc9

Successful construction of a *His₇ tag-StrepII tag-orc9* strain using pTA2292 could be used for a streptavidin-based protein pulldown to obtain biochemical information on which proteins Orc9 interacts with. This protein pulldown can be done with +/- MMC conditions to answer the question of whether Orc9 co-purifies with different proteins under MMC treatment compared to no MMC (control) conditions. Mass spectrometry could identify co-purified proteins and possibly any differences in the Orc9-dependent repair machinery between control and MMC treatment conditions.

Chapter 5: Transcriptome profiling of Orc9

5.1 Background

RNA sequencing

A transcriptome describes the complete set of transcripts expressed in a cell population and usually refers to mRNA. Transcriptome profiling of different cell types under different conditions generates quantitative data, which allows transcript levels to be compared within and between samples. Depending on the experiment, transcriptome profiling can be used to compare cells at different developmental stages or under different environmental conditions. Prior to the advancement of RNA sequencing technologies, hybridization-based techniques were used to catalogue and quantify the transcriptome of a cell type (Wang, Gerstein and Snyder, 2009). For instance, microarrays are based on measuring the hybridization of nucleotide sequences to oligonucleotide probes. This type of approach has limitations such as a high level of background noise due to cross-hybridization of probes (Okoniewski and Miller, 2006). The development of deep sequencing technologies has revolutionized methods of quantifying and mapping transcriptomes. RNA-sequencing (RNA-Seq) is a high throughput, sequence-based approach that directly determines cDNA sequences unlike previous hybridization methods. In general, the first step of this method is to construct a library by converting the total mRNA into cDNA fragments. These cDNA fragments are then sequenced by single-end or pair-end sequencing approaches. The last step is to map the sequence reads on to a reference genome (Holt and Jones, 2008). In summary, RNA-Seq provides a high throughput and coverage method to determine the mRNA abundances of each gene in a cell population.

Orc9 as a key regulator of DNA damage response

A recent paper used RNA sequencing to investigate the role of the Orc9 homologue (Cdc6-2) in *Sulfolobus islandicus*, which belongs to the archaeal phylum Crenarchaeota (Sun *et al.*, 2018). Wild type and $\Delta cdc6-2$ samples were treated with 4-nitroquinoline 1-oxide (4-NQO) prior to RNA extraction. Analysis of differentially expressed genes (DEGs) between samples revealed that the $\Delta cdc6-2$ strain could not trigger the 4-NQO-induced DNA damage response. Among the upregulated DEGs were genes involved in homologous recombination and among the downregulated DEGs were genes involved in cell division and genome segregation. This study was

the first one to identify a key regulator of DNA damage response in an archaeal organism.

5.2 Aims

Our hypothesis is that Orc9 acts as a transcription factor that regulates the transcript levels of genes involved in DNA damage response in *Haloferax volcanii*, and that the DNA damage response will be triggered by mitomycin C (MMC). RNA sequencing will identify the downstream targets of Orc9 and allow for comparison of mRNA abundances between treated and untreated samples. This data will allow the definition of the MMC-induced damage response in *H. volcanii* and determine the role Orc9 plays in executing these responses.

Dose optimisation of MMC for such an experiment is essential. A high dose of MMC would result in too much cell death and total RNA extracted would not be of good quality for subsequent sequencing. However, a low MMC dose risks having little or no DNA damage response in cells.

The aims of this chapter are:

- To perform a growth assay using WT and $\Delta orc9$ strains to optimize the dose of MMC. This data will inform the ideal concentration at which MMC induces a DNA damage response of sufficient magnitude to influence the transcriptome profile, while ensuring reasonable cell viability.
- To use this concentration of MMC to treat both the WT and $\Delta orc9$ cells prior to RNA extraction for subsequent RNA sequencing.
- To test cell viability by taking cell samples after MMC treatment and spotting on Hv-YPC plates for determination of CFUs (Colony Forming Units).

5.3 Results

Dose Optimization of MMC

Cultures of strains H53 (WT) and H2655 ($\Delta orc9$) were set up by inoculating fresh colonies into 5 ml Hv-YPC broth and were grown at 45 °C overnight. Fresh cultures were diluted (1/500) and allowed to grow overnight at 45 °C. Overnight cultures were diluted into fresh Hv-YPC broth and grown until they reached $A_{650} \approx 0.05$ for setting up the growth assay. The same 2nd overnight cultures were used to inoculate cultures in 100 ml Hv-YPC broth for treatment and RNA extraction the next day. To match the 4-NQO incubation time used in Sun *et al.*, 2018, 6 hours was picked as the duration of

MMC treatment before RNA extraction. The generation time of *H. volcanii* is ~2.5 hours and therefore 6 hours is ~2 - 3 generations. Growth of strains H53 (WT) and H2655 ($\Delta orc9$) was measured with different concentrations of MMC in the media for dose optimization (Figure 5.1). The aim was to determine a concentration of MMC at which a greater change in growth rate is induced in H2655 ($\Delta orc9$) cells compared to H53 (WT). The concentration of 0.03 $\mu\text{g/ml}$ MMC was chosen because at 6 hours of incubation, it retards the growth of H2655 ($\Delta orc9$) more than it does H53 (WT). This step was needed to confirm the concentration of MMC required for incubation with *H. volcanii* prior to RNA extraction.

MMC treatment

100 ml cultures of H53 (WT) and H2655 ($\Delta orc9$) at $A_{650} \approx 0.2$ were split into two 50 ml cultures. 1.5 μl of MMC (stock concentration of 1000 $\mu\text{g/ml}$) was added to one of the 50 ml cultures for both strains to give the final concentration of 0.03 $\mu\text{g/ml}$ MMC. 1.5 μl of sterile distilled water was added to the other 50 ml cultures as controls. Cell viability assay showed that after 6 hours of incubation with 0.03 $\mu\text{g/ml}$ MMC, H53 (WT) and H2655 ($\Delta orc9$) cells had 71% and 50% survival respectively (ratio of 1.42).

RNA extraction was carried out as described in *Chapter 2: Materials & Methods*.

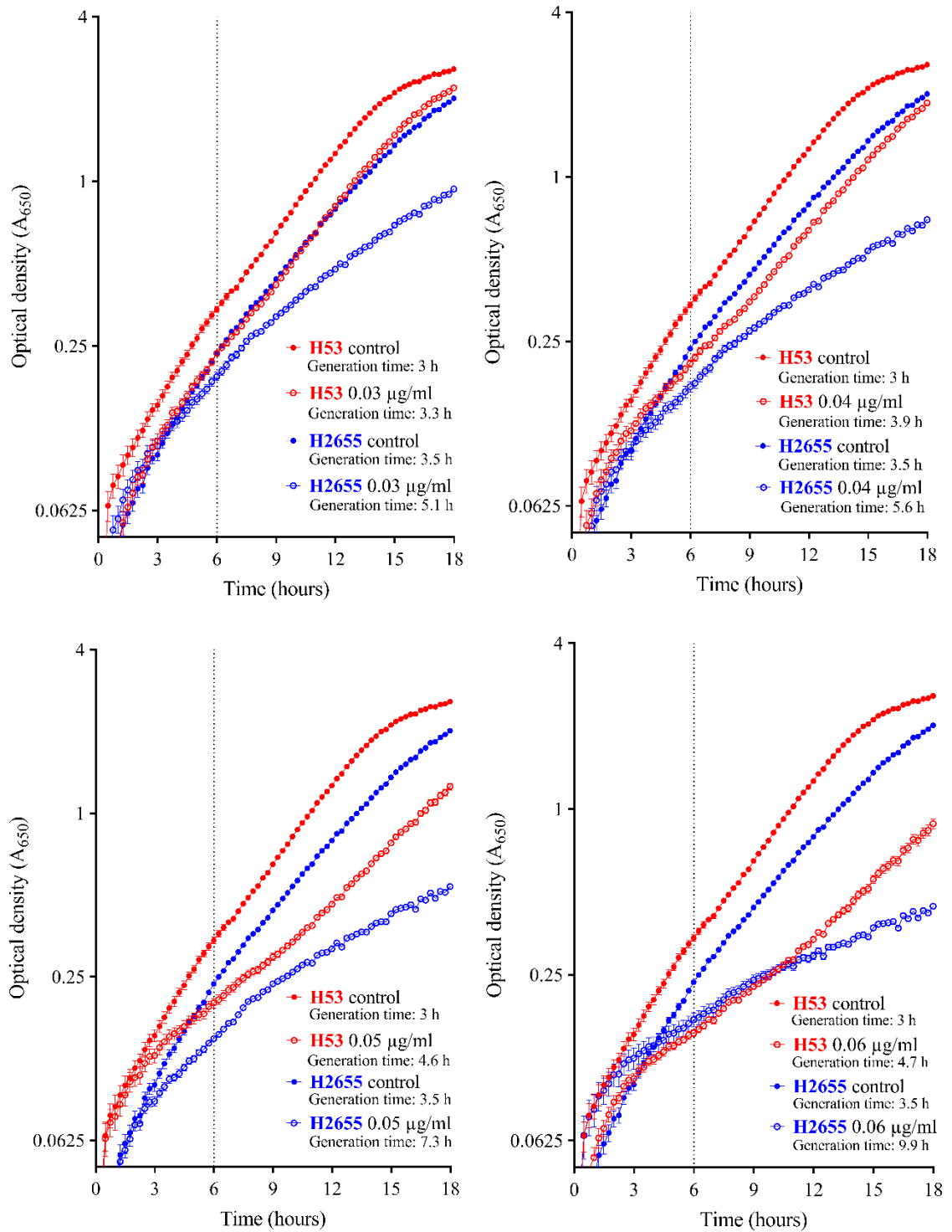


Figure 5.1: Dose optimisation of MMC for RNA-Sequencing. Growth of strains H53 (*orc9+*) and H2655 (*Δ orc9*) was measured with different levels of MMC (0, 0.03, 0.04, 0.05 and 0.06) in the media. Both strains ($n=5$) were incubated on the same 96 well plate and measured simultaneously using an Epoch2 Microplate Spectrophotometer (BioTek). Error bars represent standard error of the mean (SEM). Optical density (A_{650}) is given on a $\log_2$ scale. A vertical line is drawn on each graph at 6 hours to represent the incubation time. A dose of 0.03 $\mu\text{g/ml}$ was picked for treatment before total RNA extraction.

RNA-Seq analysis

The RNA-Sequencing determined that there were in total 527 DEGs between any of the contrast conditions. Figure 5.2 shows an overall expression profile between the different conditions.

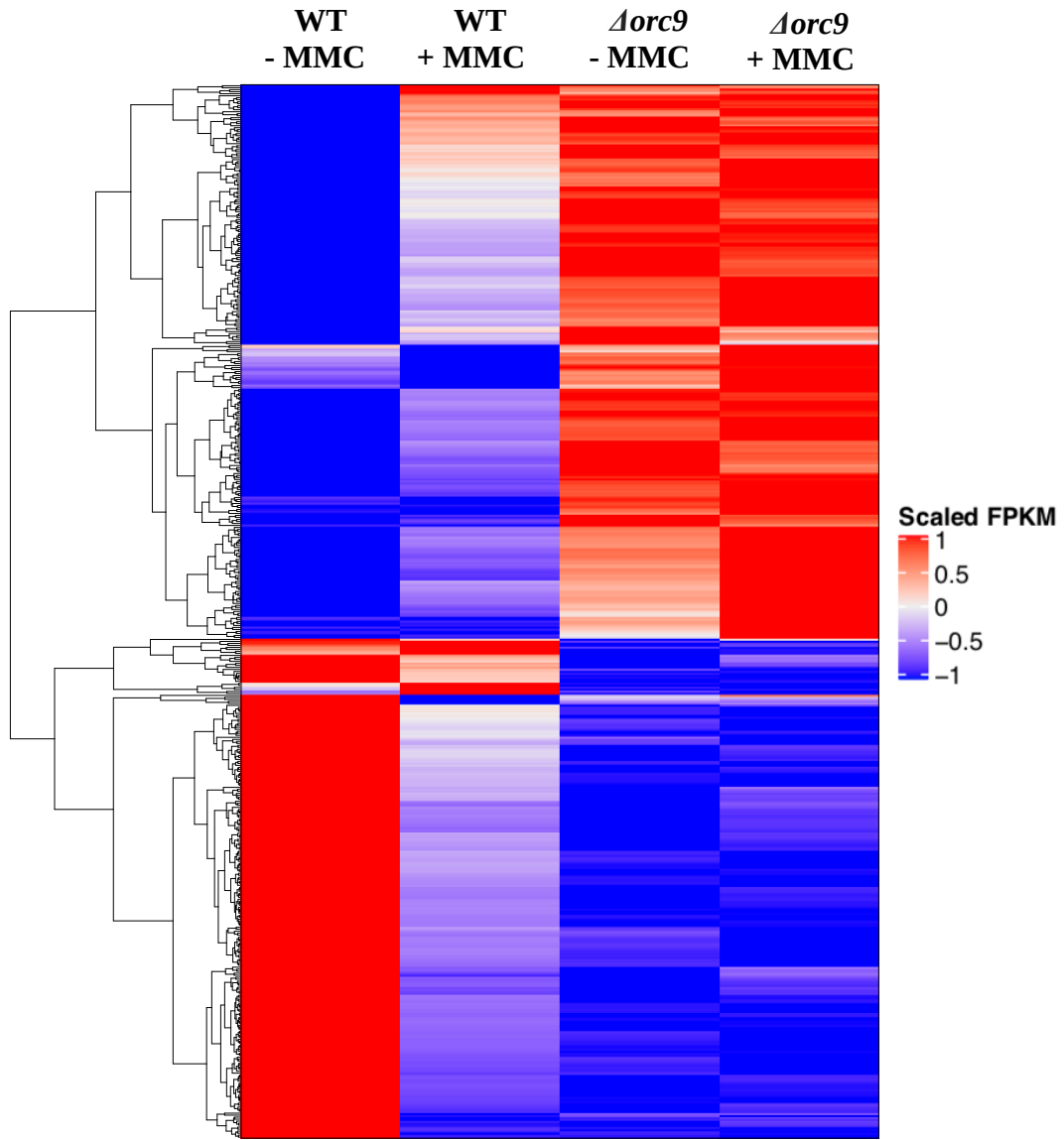


Figure 5.2: **Heatmap of the genome expression of *H. volcanii* H53 (WT) and H2655 ($\Delta orc9$).** Both strains were grown in the presence or absence of 0.03 $\mu\text{g/ml}$ MMC (indicated as + MMC and - MMC, respectively) for 6 h. Cell mass was collected from which total RNAs were prepared and used for RNA-Seq analysis. The heat map displays the scaled Fragment Per Kilo Million (FPKM) of the 527 differentially expressed genes that are significant in at least one of the multiple condition comparisons. Expression levels are illustrated with different colours with red colours representing the highest levels of expression whereas blue ones indicating the lowest levels of expression.

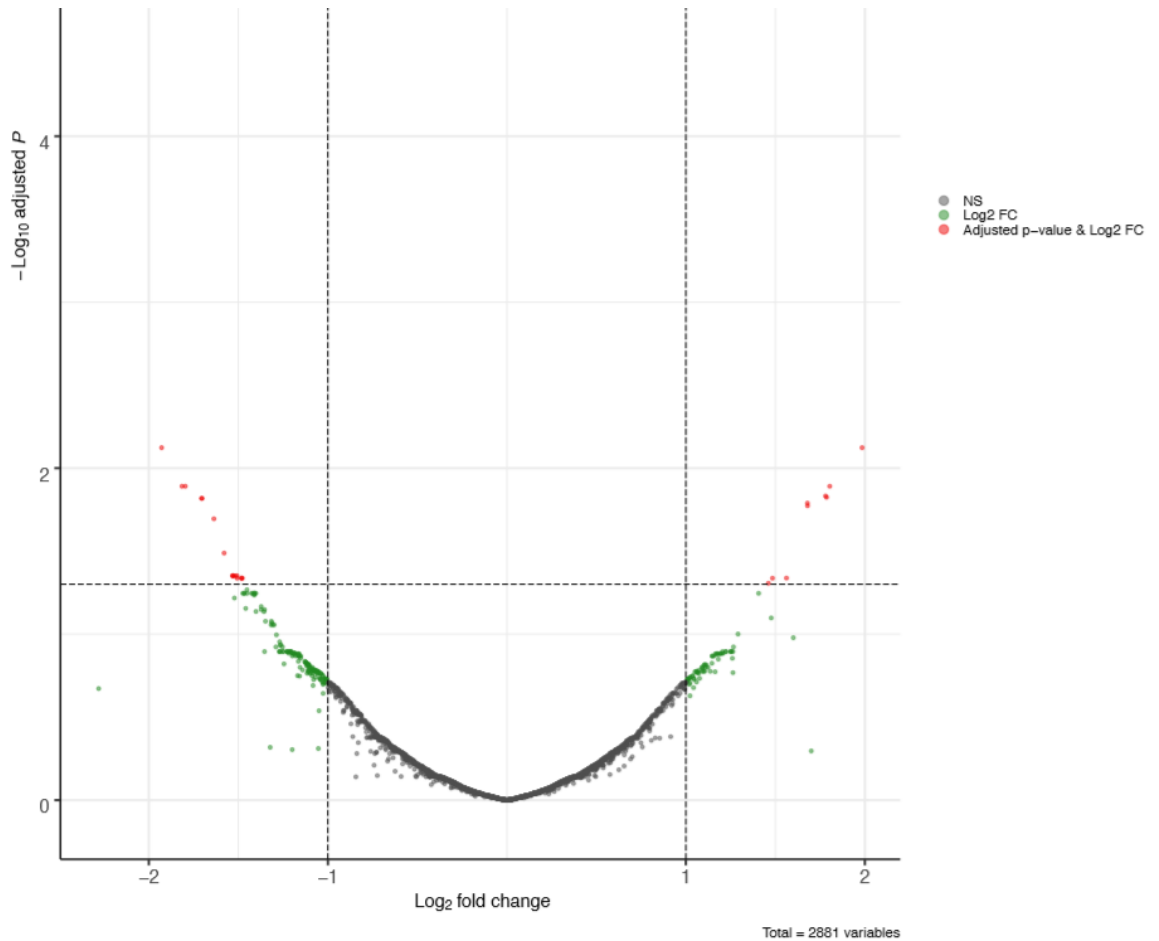


Figure 5.3: **Volcano plot showing the differentially expressed genes between WT – MMC and WT + MMC samples.** X-axis shows the fold change in gene expression. Y-axis shows the statistical significance of the fold change. Genes that are up or downregulated more than 2-fold (i.e. $-1 > \log_2 > +1$) are shown in red and green respectively. Changes in gene expression that are not significant and less than 2-fold are shown in grey.

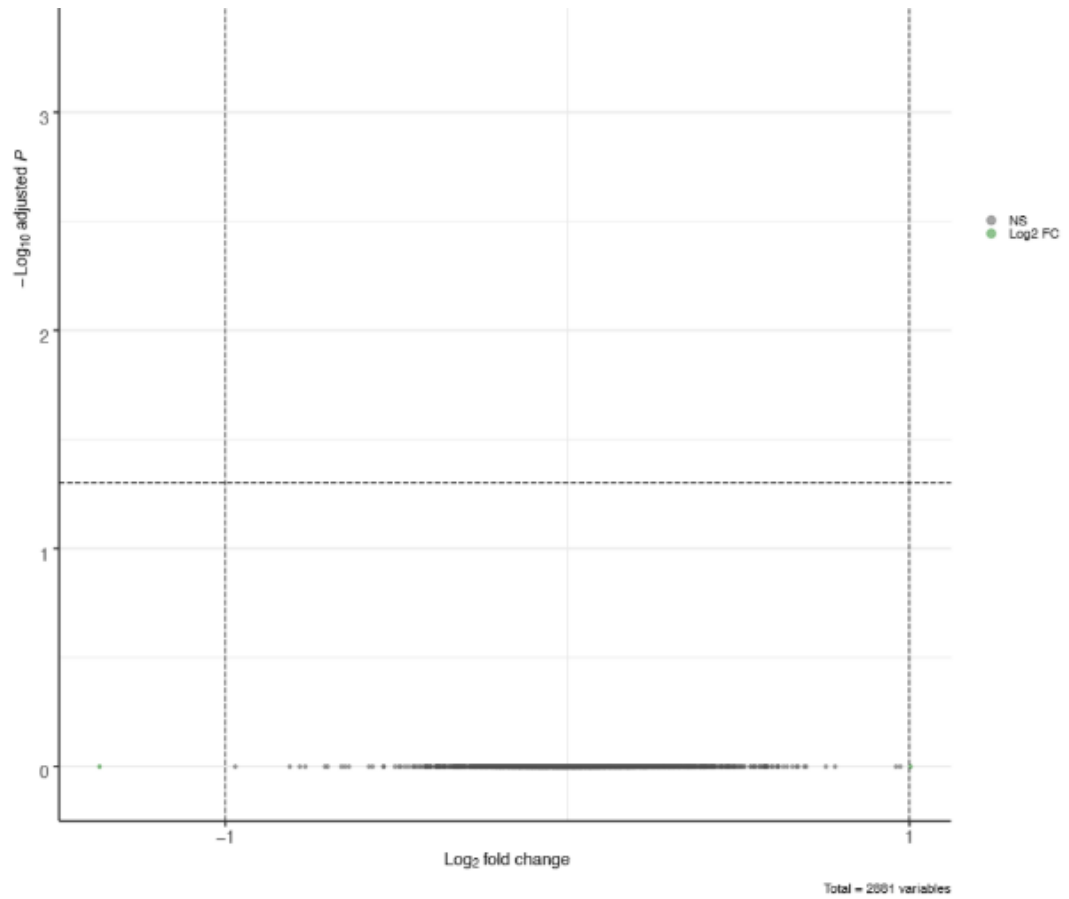


Figure 5.4: **Volcano plot showing the differentially expressed genes between $\Delta orc9$ – MMC and $\Delta orc9$ + MMC samples.** X-axis shows the fold change in gene expression. Y-axis shows the statistical significance of the fold change. Genes that are up or downregulated more than 2-fold (i.e. $-1 > \log_2 > +1$) are shown in red and green respectively. Changes in gene expression that are not significant and less than 2-fold are shown in grey.

There were 25 significant DEGs between WT + MMC and WT – MMC conditions. A selected list is shown in Table 5.1. The abundance of the *orc9* transcript is halved in the WT + MMC sample compared to WT – MMC (highlighted in yellow in Table 5.1). This is unexpected because there was a 6.65-fold increase in *cdc6-2* transcript levels upon 4-NQO treatment in *S. islandicus*.

Table 5.1 **Selected list of DEGs between WT + MMC and WT – MMC samples.**

Proteins	Gene Identity	mRNA abundance ratio (calculated based on FPKM)				Rare codon usage (%)
		WT - MMC	WT + MMC	$\Delta orc9$ - MMC	$\Delta orc9$ + MMC	
<i>DNA Replication, Transcription or Translation</i>						
Cdc6 (Orc9)	HVO_0194	1	0.516	0.024	0.035	4.53
DNA-directed RNA polymerase subunit B	HVO_0348	1	0.346	0.300	0.335	4.10
Orc1-type DNA replication protein (Orc2)	HVO_0634	1	0.358	0.464	0.523	4.16
30S ribosomal protein S8	HVO_2549	1	0.359	0.262	0.255	9.92
<i>Cellular respiration</i>						
ba3-type terminal oxidase subunit II	HVO_0944	1	0.306	0.265	0.292	6.67
terminal quinol oxidase	HVO_0763	1	0.349	0.629	0.587	7.95
succinyl-CoA synthase subunit alpha	HVO_2464	1	0.352	0.215	0.199	6.21
NADH dehydrogenase-like complex subunit H	HVO_0981	1	0.322	0.168	0.167	7.71
<i>Membrane Proteins</i>						
cell surface glycoprotein	HVO_2072	1	0.358	0.145	0.142	14.49
copper-translocating P-type ATPase	HVO_1751	1	0.334	0.423	0.420	3.48
sugar ABC transporter permease	HVO_1398	1	0.263	0.179	0.178	3.26
<i>Putative proteins associated with foreign genetic elements</i>						
hypothetical protein	NA	1	3.957	7.377	7.631	
hypothetical protein	HVO_1844	1	3.203	4.617	5.296	44.90
hypothetical protein	HVO_1863	1	0.352	0.602	0.898	7.32
transposase ISH51	HVO_0263	1	3.450	4.608	4.308	19.02
XerC/D-like integrase	HVO_1422	1	3.202	0.327	0.311	14.04
<i>Protein Transport</i>						
preprotein translocase Sec61 subunit gamma	HVO_0718	1	0.284	0.453	0.420	10.34

5.4 Discussion

The *orc9* transcript level was expected to increase upon MMC treatment in WT cells but the opposite result was observed. Unlike the 6.65-fold increase in *cdc6-2* transcript levels upon 4-NQO treatment in *S. islandicus*, MMC treatment in *H. volcanii* resulted in a decrease of the *orc9* transcript to 51.6%. This suggests a fundamental difference between the two archaeal species in regard to how the Cdc6-2/Orc9 protein works in DDR. As shown in the volcano plots and the heatmap, there is no difference in the gene expression profiles between $\Delta orc9$ samples (+ and – MMC). This suggests that the difference in gene expression profiles of WT – MMC and WT + MMC is mediated by Orc9.

There was an Orc9-dependent upregulation of XerC/D-like integrase protein gene in the MMC treated WT sample (highlighted in blue in Table 5.1). XerC/D-like integrase protein is involved in phage DNA integration using DNA recombination. Most of the hypothetical protein genes picked up by the RNA-seq have a high percentage of rare codon usage as seen on Table 5.1. This suggests that these are foreign genomic elements integrated into *H. volcanii* genome that are associated with integrated phage or plasmids. *H. volcanii* deleted for *tfeB* (a basal transcription factor) have been shown to induce a similar type of response where foreign genes are affected (Blombach *et al.*, 2018). This induction is a part of the DNA damage response induced by Orc9 and resembles the induction of lambda phage upon UV treatment in *E. coli*.

Similarly, there was an Orc9-dependent repression of a hypothetical protein gene (highlighted in green) in the MMC treated WT sample. In both treatment conditions, the $\Delta orc9$ samples have increased levels of this transcript. The gene for this hypothetical protein does not exist on our genome database and requires further investigation.

Some genes involved in DNA replication, transcription or translation were downregulated upon treatment with MMC in an Orc9-independent fashion. This is a part of the DNA damage response whereby cells stop DNA replication and transcription to ensure the repair of any lesions that might lead to mutations. The reason why several cellular respiration proteins were downregulated might be due to the same reason. Cell division requires energy in the form of ATP molecules and repressing these genes might slow down cell metabolism enough for cells to deal with DNA lesions before progressing with cell division.

The RNA-Seq in this study did not detect the downregulation of genes involved in cell cycle regulation or the upregulation of genes involved in known DNA repair pathways such as HR or NER. In Sun et al, 2018, several cell cycle regulation genes were significantly downregulated (such as *CdvA*) and several homologous recombination genes (such as *Mre11*) were upregulated. Future work should involve repeating the experiment with biological repeat samples (n=3). The MMC concentration chosen was ideal for the experiment so it does not need to change. However, different DNA damaging agents (such as MMS, phleomycin and 4-NQO) can be dose optimised and used to treat WT and $\Delta orc9$ samples for RNA-sequencing.

Chapter 6: Conclusions and Future Perspectives

The DNA damage sensitivity tests performed in this study provided evidence that Orc9 has a substantial role in the repair of damage induced by a range of mutagens. The *Δorc9* strain is hypersensitive to acute treatments with MMS, UV, 4-NQO, MMC, gamma radiation and phleomycin. In contrast, no increased sensitivity was detected when *Δorc9* strain was treated with hydrogen peroxide. This might be due to an Orc9-independent pathway of preventing or repairing oxidative DNA damage. For example, enzymatic inactivation of H₂O₂ in *Haloferax volcanii* might be efficient.

NER is the predominant pathway used to repair UV and 4-NQO induced damage. Therefore, the hypersensitivity of *Δorc9* strain to UV and 4-NQO suggests that Orc9 is involved in NER. However, this might be due to the indirect effects of UV and 4-NQO which leads to DSBs that require homologous recombination for repair.

Involvement of Orc9 in DSB repair by homologous recombination is evident in the results of the gamma radiation and phleomycin sensitivity assays. Orc9 is also involved in inter-strand crosslink repair since the *Δorc9* strain showed chronic sensitivity to MMC shown by the cell viability and growth assays. The unifying theme of these mutagens is that they can all directly (in the case of MMC) or indirectly (in the case of UV) cause replication fork stalling or collapse. Orc9 is most likely involved in accurate repair of DSBs and a pathway that re-establishes replication forks.

Hypersensitivity of *Δorc9* to MMS does not necessarily mean that Orc9 is directly involved in BER. If replication forks encounter methyl damage or BER intermediates (such as AP sites), strand breaks might be indirectly produced. Type of methyl adduct (and the choice of DNA glycosylase) has an influence on the decision of carrying out the short-patch or the long-patch BER pathway. Other alkylating agents with different alkylation properties can be used to treat WT and *Δorc9* strains to determine whether a hypersensitivity phenotype will be produced similar to MMS treatment.

Δorc9 ΔFEN1 double deletion shows a synthetic growth phenotype. This might be due to the involvement of FEN1 in the long-patch pathway of BER. Orc9 might have a role in predominantly the short-patch BER pathway. This would explain the synthetic

growth defect phenotype of the $\Delta orc9 \Delta FEN1$ double deletion strain. Another explanation for this might be due to FEN1 being an essential protein for microhomology dependent alternative end joining (inaccurate DSB repair pathway) (Sharma *et al.*, 2015). This is supported by the finding that FEN1 is overexpressed in majority of cancers such as breast, lung and pancreatic (Singh *et al.*, 2008). Results of this thesis showed that Orc9 has a major role in the accurate repair of DSBs via homologous recombination. FEN1 is involved in an alternative pathway to HR that repairs DSBs but in a highly mutagenic and inaccurate way. FEN1 and Orc9 are likely two proteins that act on DSBs in alternative pathways. Therefore, when both are deleted in the same strain, growth is severely retarded. This is also true for RadA and Orc9. Chapter 4 demonstrates the synergistic relationship between Orc9 and RadA. RadA is a key component of DSB repair by homologous recombination. This further supports the role of Orc9 in homologous recombination. Orc9 might have evolved to carry out replication for re-start in a pathway that is independent from RadA, to ensure genomic stability by avoiding recombination.

The tryptophan inducible *orc9* strains constructed in this study can be used to perform DNA damage sensitivity assays with several mutagens. This would show whether inducible *orc9* strains grown with increasing concentrations of tryptophan shows increasing survival to mutagens. This would directly link the level of *orc9* transcript with the amount of DNA repair that is carried out in cells. Obviously, level of transcript might not always correlate with level of the Orc9 protein. Western blots might be used to quantify the changing levels of the Orc9 protein under different tryptophan concentrations and +/- MMC. Once optimised to detect the *orc9* transcript, reverse transcriptase PCR might be used to compare levels of *orc9* transcript in different samples (changing tryptophan concentrations). This would confirm if the induction by tryptophan is occurring at the gene level. Another use of the inducible *orc9* strains might be checking for synthetic growth defect phenotypes. Deleted genes that might have synergistic relationships with *orc9* (just like *radA*) from the inducible *orc9* strains can be phenotyped using tryptophan gradient plates (See Chapter 2: Materials and Methods).

Using the tagged-*orc9* construct made in this study (pTA2292), a tagged-*orc9* strain can be made. A protein pulldown experiment would identify proteins that co-purify with Orc9 and elucidate potential interactors. Chromatin Immunoprecipitation

Sequencing (ChIP-seq) might be performed to identify the genome wide DNA binding sites of Orc9. This would elucidate any downstream targets of Orc9 directly through crosslinking Orc9 to its DNA binding sites and subsequent sequencing.

The RNA-seq data revealed that Orc9 changes the transcriptome profile of *H. volcanii* substantially. Evidently, $\Delta orc9$ samples (+ or – MMC) cannot mediate Orc9 induced response to DNA damage. The decrease in the *orc9* transcript level in the WT + MMC sample compared to WT – MMC (control) sample was surprising since the opposite was expected. However, this might be due to a fundamental difference between *H. volcanii* and *S. islandicus* in regard to the mechanism of Orc9 in DDR. Orc9 is most likely a transcriptional regulator that changes the expression of many genes. This supported by structural protein information. Like Orc1, Orc9 possesses a C-terminal winged helix domain which binds DNA. Further biochemical analysis can be carried out to determine the affinity of Orc9 to DNA that might give us an idea of the specificity of Orc9 compared to other Orcs.

The RNA-seq experiment can be repeated with more biological repeat samples (n=3) to get better p-values and more confidence in our data. Different DNA damaging agents (such as MMS, phleomycin, 4-NQO) can be optimised for this experiment and used to treat WT and $\Delta orc9$ samples for subsequent RNA-sequencing.

Overall, this study has provided insights to the role of Orc9 in DNA damage response. Orc9 is likely to be a key regulator that orchestrates a network of event that lead to repair of damage repair. If this is the case, this means that DDR is conserved in all domains of life including Archaea.

References

- Alberts, B. *et al.* (2014) *Molecular Biology of the Cell*. Sixth. Oxford: Taylor and Francis. Available at: <https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=5320520> (Accessed: 10 March 2020).
- Allers, T. and Mevarech, M. (2005a) 'Archaeal genetics — the third way', *Nature Reviews Genetics*, 6(1), pp. 58–73. doi: 10.1038/nrg1504.
- Allers, T. and Mevarech, M. (2005b) 'Archaeal genetics — the third way', *Nature Reviews Genetics*, 6(1), pp. 58–73. doi: 10.1038/nrg1504.
- Andoh, T. and Ide, T. (1972) 'Strand scission and rejoining of DNA in cultured mammalian cells induced by 4-nitroquinoline 1-oxide.', *Cancer research*, 32(6), pp. 1230–5. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/5030823> (Accessed: 29 September 2019).
- Arima, Y. *et al.* (2006) '4-Nitroquinoline 1-Oxide Forms 8-Hydroxydeoxyguanosine in Human Fibroblasts through Reactive Oxygen Species', *Toxicological Sciences*, 91(2), pp. 382–392. doi: 10.1093/toxsci/kfj161.
- Ausiannikava, D. and Allers, T. (2017) 'Diversity of DNA Replication in the Archaea.', *Genes*. Multidisciplinary Digital Publishing Institute (MDPI), 8(2). doi: 10.3390/genes8020056.
- Barry, E. R. and Bell, S. D. (2006) 'DNA Replication in the Archaea', *Microbiology and Molecular Biology Reviews*, 70(4), pp. 876–887. doi: 10.1128/MMBR.00029-06.
- Beranek, D. T. (1990) 'Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. Elsevier, 231(1), pp. 11–30. doi: 10.1016/0027-5107(90)90173-2.
- Blackford, A. N. and Jackson, S. P. (2017) 'ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response', *Molecular Cell*, 66(6), pp. 801–817. doi: 10.1016/j.molcel.2017.05.015.
- Bleichert, F., Botchan, M. R. and Berger, J. M. (2015) 'Crystal structure of the eukaryotic origin recognition complex', *Nature*, 519(7543), pp. 321–326. doi: 10.1038/nature14239.
- Blombach, F. *et al.* (2018) 'Structural and functional adaptation of *Haloferax volcanii* TFE α/β ', *Nucleic Acids Research*. Narnia, 46(5), pp. 2308–2320. doi: 10.1093/nar/gkx1302.
- de Boer, J. and Hoeijmakers, J. H. J. (2000) 'Nucleotide excision repair and human syndromes', *Carcinogenesis*. Narnia, 21(3), pp. 453–460. doi: 10.1093/carcin/21.3.453.
- Boiteux, S. and Laval, J. (1983) 'Imidazole open ring 7-methylguanine : An inhibitor of DNA synthesis', *Biochemical and Biophysical Research Communications*, 110(2), pp. 552–558. doi: 10.1016/0006-291X(83)91185-3.
- Bowater, R. and Doherty, A. J. (2006) 'Making Ends Meet: Repairing Breaks in Bacterial DNA by Non-Homologous End-Joining', *PLoS Genetics*. Public Library of Science, 2(2), p. e8. doi: 10.1371/journal.pgen.0020008.
- Brissett, N. C. and Doherty, A. J. (2009) 'Repairing DNA double-strand breaks by the prokaryotic non-homologous end-joining pathway', *Biochemical Society Transactions*, 37(3), pp. 539–545. doi: 10.1042/BST0370539.
- Busch, C. R. and DiRuggiero, J. (2010) 'MutS and MutL Are Dispensable for Maintenance of the Genomic Mutation Rate in the Halophilic Archaeon *Halobacterium salinarum* NRC-1', *PLoS ONE*. Edited by R. Aramayo, 5(2), p. e9045. doi: 10.1371/journal.pone.0009045.
- Butala, M., Žgur-Bertok, D. and Busby, S. J. W. (2009) 'The bacterial LexA transcriptional repressor', *Cellular and Molecular Life Sciences*, 66(1), pp. 82–93. doi: 10.1007/s00018-

008-8378-6.

DeLong, E. F. and Pace, N. R. (2001) 'Environmental diversity of bacteria and archaea.', *Systematic biology*, 50(4), pp. 470–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12116647> (Accessed: 29 September 2019).

Doetsch, P. W. and Cunningham, R. P. (no date) 'The enzymology of apurinic/apyrimidinic endonucleases.', *Mutation research*, 236(2–3), pp. 173–201. doi: 10.1016/0921-8777(90)90004-o.

Essen, L. O. and Klar, T. (2006) 'Light-driven DNA repair by photolyases', *Cellular and Molecular Life Sciences*. Birkhäuser-Verlag, 63(11), pp. 1266–1277. doi: 10.1007/s00018-005-5447-y.

Geier, G. E. and Modrich, P. (1979) 'Recognition sequence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease.', *The Journal of biological chemistry*, 254(4), pp. 1408–13. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/368070> (Accessed: 30 September 2019).

Goldstein, S., Meyerstein, D. and Czapski, G. (1993) 'The Fenton reagents', *Free Radical Biology and Medicine*. Pergamon, 15(4), pp. 435–445. doi: 10.1016/0891-5849(93)90043-T.

Götz, D. *et al.* (2007) 'Responses of hyperthermophilic crenarchaea to UV irradiation', *Genome Biology*, 8(10), p. R220. doi: 10.1186/gb-2007-8-10-r220.

Grogan, D. W. (2004) 'Stability and repair of DNA in hyperthermophilic Archaea.', *Current issues in molecular biology*, 6(2), pp. 137–44. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15119824> (Accessed: 30 September 2019).

Guy, L. and Ettema, T. J. G. (2011) 'The archaeal "TACK" superphylum and the origin of eukaryotes', *Trends in Microbiology*, 19(12), pp. 580–587. doi: 10.1016/j.tim.2011.09.002.

Hartman, A. L. *et al.* (2010) 'The Complete Genome Sequence of Haloferax volcanii DS2, a Model Archaeon', *PLoS ONE*. Edited by I. Friedberg. Public Library of Science, 5(3), p. e9605. doi: 10.1371/journal.pone.0009605.

Hawkins, M. *et al.* (2013) 'Accelerated growth in the absence of DNA replication origins', *Nature*, 503(7477), pp. 544–547. doi: 10.1038/nature12650.

Henle, E. S. and Linn, S. (1997) 'Formation, Prevention, and Repair of DNA Damage by Iron/Hydrogen Peroxide', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 272(31), pp. 19095–19098. doi: 10.1074/JBC.272.31.19095.

Holt, R. A. and Jones, S. J. M. (2008) 'The new paradigm of flow cell sequencing.', *Genome research*. Cold Spring Harbor Laboratory Press, 18(6), pp. 839–46. doi: 10.1101/gr.073262.107.

Hwang, D. S. and Kornberg, A. (1992) 'Opening of the replication origin of Escherichia coli by DnaA protein with protein HU or IHF.', *The Journal of biological chemistry*, 267(32), pp. 23083–6. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/1429655> (Accessed: 29 September 2019).

Ishino, S. *et al.* (2013) 'The archaeal DNA replication machinery: past, present and future.', *Genes & genetic systems*, 88(6), pp. 315–9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24789967> (Accessed: 20 September 2019).

Ishino, S. *et al.* (2016) 'Identification of a mismatch-specific endonuclease in hyperthermophilic Archaea', *Nucleic Acids Research*, 44(7), pp. 2977–2986. doi: 10.1093/nar/gkw153.

Jacob, F., Brenner, S. and Cuzin, F. (1963) 'On the Regulation of DNA Replication in Bacteria', *Cold Spring Harbor Symposia on Quantitative Biology*. Cold Spring Harbor

- Laboratory Press, 28(0), pp. 329–348. doi: 10.1101/SQB.1963.028.01.048.
- Jasin, M. and Rothstein, R. (2013) ‘Repair of strand breaks by homologous recombination.’, *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 5(11), p. a012740. doi: 10.1101/cshperspect.a012740.
- Karran, P. *et al.* (1980) ‘Escherichia coli mutants deficient in 3-methyladenine-DNA glycosylase’, *Journal of Molecular Biology*. Academic Press, 140(1), pp. 101–127. doi: 10.1016/0022-2836(80)90358-7.
- Kelman, Z. and White, M. F. (2005) ‘Archaeal DNA replication and repair’, *Current Opinion in Microbiology*, 8(6), pp. 669–676. doi: 10.1016/j.mib.2005.10.001.
- Kowalczykowski, S. C. (2015) ‘An Overview of the Molecular Mechanisms of Recombinational DNA Repair’, *Cold Spring Harbor Perspectives in Biology*, 7(11), p. a016410. doi: 10.1101/cshperspect.a016410.
- Krogh, B. O. and Symington, L. S. (2004) ‘Recombination Proteins in Yeast’, *Annual Review of Genetics*, 38(1), pp. 233–271. doi: 10.1146/annurev.genet.38.072902.091500.
- Kunkel, T. A. and Erie, D. A. (2015) ‘Eukaryotic Mismatch Repair in Relation to DNA Replication’, *Annual Review of Genetics*, 49(1), pp. 291–313. doi: 10.1146/annurev-genet-112414-054722.
- Lankinen, M. H., Vilpo, L. M. and Vilpo, J. A. (1996) ‘UV- and γ -irradiation-induced DNA single-strand breaks and their repair in human blood granulocytes and lymphocytes’, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 352(1–2), pp. 31–38. doi: 10.1016/0027-5107(95)00172-7.
- Large, A. *et al.* (2007) ‘Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene’, *Molecular Microbiology*, 66(5), pp. 1092–1106. doi: 10.1111/j.1365-2958.2007.05980.x.
- Larson, K. *et al.* (1985) ‘Methylation-induced blocks to in vitro DNA replication’, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 150(1–2), pp. 77–84. doi: 10.1016/0027-5107(85)90103-4.
- Leonard, A. C. and Méchali, M. (2013) ‘DNA replication origins.’, *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 5(10), p. a010116. doi: 10.1101/cshperspect.a010116.
- Lestini, R. *et al.* (2011) ‘Archaeal DNA Repair Nucleases’, in *DNA Repair - On the Pathways to Fixing DNA Damage and Errors*. InTech. doi: 10.5772/23361.
- Liu, Y. *et al.* (2005) ‘DNA Polymerase β and Flap Endonuclease 1 Enzymatic Specificities Sustain DNA Synthesis for Long Patch Base Excision Repair’, *Journal of Biological Chemistry*, 280(5), pp. 3665–3674. doi: 10.1074/jbc.M412922200.
- Loechler, E. L., Green, C. L. and Essigmann, J. M. (1984) ‘In vivo mutagenesis by O6-methylguanine built into a unique site in a viral genome.’, *Proceedings of the National Academy of Sciences*, 81(20), pp. 6271–6275. doi: 10.1073/pnas.81.20.6271.
- Macleod, F. *et al.* (2019) ‘Asgard archaea: Diversity, function, and evolutionary implications in a range of microbiomes’, *AIMS Microbiology*, 5(1), pp. 48–61. doi: 10.3934/microbiol.2019.1.48.
- Makarova, K. S., Koonin, E. V and Kelman, Z. (2012) ‘The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes’, *Biology Direct*, 7(1), p. 7. doi: 10.1186/1745-6150-7-7.
- Marteijn, J. A. *et al.* (2014) ‘Understanding nucleotide excision repair and its roles in cancer and ageing’, *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 15(7), pp. 465–481. doi: 10.1038/nrm3822.

- Méchal, M. (2010) 'Eukaryotic DNA replication origins: many choices for appropriate answers', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 11(10), pp. 728–738. doi: 10.1038/nrm2976.
- Müller, C. A. *et al.* (no date) 'The dynamics of genome replication using deep sequencing'. doi: 10.1093/nar/gkt878.
- Mullakhanbhai, M. F. and Larsen, H. (1975) 'Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement', *Archives of Microbiology*. Springer-Verlag, 104(1), pp. 207–214. doi: 10.1007/BF00447326.
- Müller, C. A. and Nieduszynski, C. A. (2012) 'Conservation of replication timing reveals global and local regulation of replication origin activity', *Genome Research*, 22(10), pp. 1953–1962. doi: 10.1101/gr.139477.112.
- O'Donnell, M., Langston, L. and Stillman, B. (2013) 'Principles and concepts of DNA replication in bacteria, archaea, and eukarya.', *Cold Spring Harbor perspectives in biology*. Cold Spring Harbor Laboratory Press, 5(7). doi: 10.1101/cshperspect.a010108.
- Okoniewski, M. J. and Miller, C. J. (2006) 'Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations.', *BMC bioinformatics*. BioMed Central, 7, p. 276. doi: 10.1186/1471-2105-7-276.
- Oren, A. (1999) 'Bioenergetic aspects of halophilism.', *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology (ASM), 63(2), pp. 334–48. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10357854> (Accessed: 29 September 2019).
- Oren, A. (2008) 'Microbial life at high salt concentrations: phylogenetic and metabolic diversity', *Saline Systems*, 4(1), p. 2. doi: 10.1186/1746-1448-4-2.
- Povirk, L. F. (1996) 'DNA damage and mutagenesis by radiomimetic DNA-cleaving agents: bleomycin, neocarzinostatin and other enediynes', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 355(1–2), pp. 71–89. doi: 10.1016/0027-5107(96)00023-1.
- Raymann, K. *et al.* (2014) 'Global Phylogenomic Analysis Disentangles the Complex Evolutionary History of DNA Replication in Archaea', *Genome Biology and Evolution*, 6(1), pp. 192–212. doi: 10.1093/gbe/evu004.
- Robinson, N. P. and Bell, S. D. (2005) 'Origins of DNA replication in the three domains of life', *FEBS Journal*, pp. 3757–3766. doi: 10.1111/j.1742-4658.2005.04768.x.
- Salerno, V. *et al.* (2003) 'Transcriptional response to DNA damage in the archaeon *Sulfolobus solfataricus*', *Nucleic Acids Research*, 31(21), pp. 6127–6138. doi: 10.1093/nar/gkg831.
- Sartori, A. A. and Jiricny, J. (2003) 'Enzymology of base excision repair in the hyperthermophilic archaeon *Pyrobaculum aerophilum*.'', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 278(27), pp. 24563–76. doi: 10.1074/jbc.M302397200.
- Savva, R. *et al.* (1995) 'The structural basis of specific base-excision repair by uracil–DNA glycosylase', *Nature*, 373(6514), pp. 487–493. doi: 10.1038/373487a0.
- Sawa, M. and Masai, H. (2009) 'Drug design with Cdc7 kinase: a potential novel cancer therapy target.', *Drug design, development and therapy*, 2, pp. 255–64. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19920912> (Accessed: 29 September 2019).
- Schaper, S. and Messer, W. (1995) 'Interaction of the Initiator Protein DnaA of *Escherichia coli* with Its DNA Target', *Journal of Biological Chemistry*, 270(29), pp. 17622–17626. doi: 10.1074/jbc.270.29.17622.
- Schärer, O. D. (2013) 'Nucleotide excision repair in eukaryotes.', *Cold Spring Harbor*

- perspectives in biology*. Cold Spring Harbor Laboratory Press, 5(10), p. a012609. doi: 10.1101/cshperspect.a012609.
- Sedgwick, B. (2004) 'Repairing DNA-methylation damage', *Nature Reviews Molecular Cell Biology*, 5(2), pp. 148–157. doi: 10.1038/nrm1312.
- Sharma, S. *et al.* (2015) 'Homology and enzymatic requirements of microhomology-dependent alternative end joining', *Cell death & disease*. Nature Publishing Group, 6(3), p. e1697. doi: 10.1038/cddis.2015.58.
- Singh, P. *et al.* (2008) 'Overexpression and hypomethylation of Flap endonuclease 1 gene in breast and other cancers', *Molecular Cancer Research*. NIH Public Access, 6(11), pp. 1710–1717. doi: 10.1158/1541-7786.MCR-08-0269.
- Sinha, R. P. and Häder, D.-P. (2002) 'UV-induced DNA damage and repair: a review', *Photochemical & Photobiological Sciences*. Royal Society of Chemistry, 1(4), pp. 225–236. doi: 10.1039/b201230h.
- Sirbu, B. M. and Cortez, D. (2013) 'DNA Damage Response: Three Levels of DNA Repair Regulation', *Cold Spring Harbor Perspectives in Biology*, 5(8), pp. a012724–a012724. doi: 10.1101/cshperspect.a012724.
- Sleigh, M. J. (1976) 'The mechanism of DNA breakage by phleomycin in vitro.', *Nucleic Acids Research*. Oxford University Press, 3(4), p. 891. doi: 10.1093/NAR/3.4.891.
- Sleytr, U. B. and Messner, P. (1983) 'Crystalline Surface Layers on Bacteria', *Annual Review of Microbiology*, 37(1), pp. 311–339. doi: 10.1146/annurev.mi.37.100183.001523.
- Slupphaug, G., Kavli, B. and Krokan, H. E. (2003) 'The interacting pathways for prevention and repair of oxidative DNA damage', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. Elsevier, 531(1–2), pp. 231–251. doi: 10.1016/J.MRFMMM.2003.06.002.
- Sun, M. *et al.* (2018) 'An Orc1/Cdc6 ortholog functions as a key regulator in the DNA damage response in Archaea', *Nucleic Acids Research*, 46(13), pp. 6697–6711. doi: 10.1093/nar/gky487.
- Tomasz, M. *et al.* (1987) 'Isolation and structure of a covalent cross-link adduct between mitomycin C and DNA', *Science*, 235(4793), pp. 1204–1208. doi: 10.1126/science.3103215.
- Tomasz, M. (1995) 'Mitomycin C: small, fast and deadly (but very selective).', *Chemistry & biology*. Elsevier, 2(9), pp. 575–9. doi: 10.1016/1074-5521(95)90120-5.
- Wallace, S. S. (2014) 'Base excision repair: a critical player in many games.', *DNA repair*. NIH Public Access, 19, pp. 14–26. doi: 10.1016/j.dnarep.2014.03.030.
- Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature Reviews Genetics*. Nature Publishing Group, 10(1), pp. 57–63. doi: 10.1038/nrg2484.
- Wegrzyn, K. E. *et al.* (2016) 'Replisome Assembly at Bacterial Chromosomes and Iteron Plasmids.', *Frontiers in molecular biosciences*. Frontiers Media SA, 3, p. 39. doi: 10.3389/fmolb.2016.00039.
- White, M. F. and Allers, T. (2018) 'DNA repair in the archaea-an emerging picture', *FEMS Microbiology Reviews*, 020, pp. 514–526. doi: 10.1093/femsre/fuy020.
- WILKANSKY, B. (1936) 'Life in the Dead Sea', *Nature*. Nature Publishing Group, 138(3489), pp. 467–467. doi: 10.1038/138467a0.
- Woese, C. and Fox, G. (no date) 'Phylogenetic structure of the prokaryotic domain: The primary kingdoms'. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC432104/pdf/pnas00033-0388.pdf> (Accessed: 19 September 2019).

Wright, W. D., Shah, S. S. and Heyer, W.-D. (2018) 'Homologous recombination and the repair of DNA double-strand breaks.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 293(27), pp. 10524–10535. doi: 10.1074/jbc.TM118.000372.

Wyatt, M. D. and Pittman, D. L. (2006) 'Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks.', *Chemical research in toxicology*. NIH Public Access, 19(12), pp. 1580–94. doi: 10.1021/tx060164e.

Xiao, W., Chow, B. L. and Rathgeber, L. (1996) 'The repair of DNA methylation damage in *Saccharomyces cerevisiae*', *Current Genetics*. Springer-Verlag, 30(6), pp. 461–468. doi: 10.1007/s002940050157.

Zaremba-Niedzwiedzka, K. *et al.* (2017) 'Asgard archaea illuminate the origin of eukaryotic cellular complexity', *Nature*. Nature Publishing Group, 541(7637), pp. 353–358. doi: 10.1038/nature21031.

Appendix: FPKM values for all significant differentially expressed genes

Geneid	H2655-Control	H2655-MMC	H53-Control	H53-MMC	seqid	start	end	strand	old_locus_tag	db_gene_id
gene46	2.930318386	3.443679	1	3.436241953	NC_013967.1	46537	46815	+	NA	31787462
gene256	4.608328338	4.308207	1	3.450267431	NC_013967.1	235110	236276	-	HVO_0263	8926832
gene339	0.299505457	0.334762	1	0.345889053	NC_013967.1	313886	315715	+	HVO_0348	8926659
gene615	0.463870183	0.523178	1	0.357716039	NC_013967.1	569203	570429	-	HVO_0634	8924378
gene694	0.453246045	0.420293	1	0.284101008	NC_013967.1	644717	644890	+	HVO_0718	8924344
gene736	0.62897606	0.587409	1	0.349073515	NC_013967.1	684671	685159	-	HVO_0763	8924155
gene910	0.265233645	0.292181	1	0.306349706	NC_013967.1	855836	856375	+	HVO_0944	8926237
gene946	0.167521073	0.16736	1	0.321501434	NC_013967.1	892036	893085	+	HVO_0981	8925814
gene1079	0.268191208	0.251315	1	0.346323358	NC_013967.1	1019071	1019271	-	HVO_1115	8925952
gene1353	0.178991746	0.178427	1	0.262544329	NC_013967.1	1272027	1273130	-	HVO_1398	8926476
gene1354	0.197314445	0.207039	1	0.287784058	NC_013967.1	1273127	1274395	-	HVO_1399	8926187
gene1377	0.327371523	0.311423	1	3.201571771	NC_013967.1	1295035	1296081	-	HVO_1422	8924905
gene1505	0.237539175	0.303431	1	0.345535342	NC_013967.1	1425648	1425884	-	HVO_1556	8925439
gene1598	2.616442229	3.260438	1	3.490843255	NC_013967.1	1516953	1517207	+	HVO_1652	8924505
gene1696	0.422745479	0.419936	1	0.334471977	NC_013967.1	1620826	1623411	-	HVO_1751	8925461
gene1723	2.370600442	2.599141	1	2.752711768	NC_013967.1	1647441	1647956	+	HVO_1780	8924555
gene1778	7.377223478	7.630837	1	3.957342381	NC_013967.1	1698203	1698400	-	NA	31787481
gene1785	4.616712198	5.295525	1	3.203309025	NC_013967.1	1704453	1704632	-	HVO_1844	8925736
gene1802	0.601774835	0.897657	1	0.351663188	NC_013967.1	1720493	1720738	-	HVO_1863	8926396
gene1958	2.52544998	2.402219	1	2.796953799	NC_013967.1	1866807	1867007	-	HVO_2021	8926070
gene1973	0.555191135	0.48698	1	0.307245109	NC_013967.1	1882056	1882409	-	HVO_2037	8925921
gene2009	0.145403505	0.142156	1	0.35783494	NC_013967.1	1923913	1926396	+	HVO_2072	8924493
gene2398	0.215189062	0.199359	1	0.351740383	NC_013967.1	2332664	2333533	-	HVO_2464	8924017
gene2484	0.262257213	0.255478	1	0.358995059	NC_013967.1	2411343	2411735	-	HVO_2549	8924207
gene2891	3.824131211	4.249011	1	2.954028994	NC_013967.1	2800044	2800235	-	HVO_2964	8925785