Genetic analysis of radiation resistance in *Haloferax volcanii*

Callum Nicholas Hayes-Smith BSc (Hons)

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

~

September 2023

Contents

Academic Integrity Statement

I, Callum Nicholas Hayes-Smith hereby confirm that all work presented in this doctoral thesis, unless otherwise stated, is as the result of my own autonomous work based on my research and literature published, which is seen in the references used and can be found at the end of this thesis.

Callum Nicholas Hayes-Smith

Acknowledgements

First and foremost I'd like to thank my supervisors, Thorsten Allers and Alan Huett. Without your brilliant and stoic guidance I certainly wouldn't have made it through my annual nervous breakdowns, let alone four years of a PhD during a pandemic. I cannot express in words how grateful I am to you both for your constant support, empathy and humour – there was never a dull moment and I feel privileged to have worked with you both. Alan, thank you too for guiding me through the last seven years, to say it has flown by would be an understatement.

Nikita - we started together, and I will miss our weekly rambles and rants when things were not going to plan. I wouldn't have made it through without you. In addition, to all members of D119, thank you for your friendship over the last four years. Special mention to Laura and Ambika for all your help and support.

We stand on the shoulders of giants, and I would not be where I am today if it were not for the guidance and inspiration provided by my many teachers over the years. Special mentions must go to Ed March-Shawcross, Clifford Dammers and Stuart Ingleston-Orme who in their different ways inspired me and provided much needed encouragement over the years.

I'd also like to extend my thanks to those within the Combined Cadet Force, who were always on hand each week to provide a welcome break from PhD work and I'm very grateful to you all for reminding me that there was life beyond the lab. A big shout out to Brian, Brock, Tim and Susie – you guys are amazing and I remain privileged to count you amongst my closest friends.

Thank you to my amazing parents, Anton and Rachel, and siblings Ellie and William, for putting up with me when things were not going to plan. Thank you to my brilliant grandparents, Janet and Gordon for your endless love and support. To my parents especially, I will be forever grateful for the emotional (and financial!) support that you provided me over the last seven years at university. Thank you for always backing me – I will never be able to repay you for the love and attention you have afforded me. Thanks must also go to my beloved dogs, Pepper and Lola, for providing me with the best cuddles! A page of acknowledgements will never be anywhere near enough to explain how grateful I am to you all.

Most importantly however, I would like to thank myself. Without whom, this entire PhD would not have been possible. Well done me!

Anyway, enough waffle, let the magic begin…..

For Janet, Joyce and Bruce, who I know would have been proud…

Abbreviations

5-FOA	5-fluroorotic acid
ATP	Adenosine 5'triphosphate
BER	Base excision repair
BIR	Break-induced replication
BLAST	Basic local alignment search tool
Bp	Base pair
dHJ	Double Holliday junction
DNA	Deoxyribose nucleic acid
DSB	Double strand break
DSBR	Double strand break repair
dsDNA	Double-stranded DNA
EDTA	Ethylenediaminetetraacetic acid
HJ	Holliday junction
HR	Homologous recombination
ICL	Interstrand crosslink
Kb	Kilobase
Mb	Megabase
MMEJ	Microhomology mediated end joining
MMC	Mitomycin C
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
RNA	Ribonucleic acid
SDSA	Synthesis-dependant strand annealing
ssDNA	Single-stranded DNA
SSA	Single strand annealing
UV	Ultraviolet light
WT	Wild-type

Abstract

Archaea,considered as the third domain of life alongside bacteria and eukaryotes, represent a highly diverse group of organisms. Attention to archaeal DNA repair pathways has been considerable for a long time and many archaeal species inhabit extreme environmental niches where there is a higher rate of genomic insult. It is therefore thought that such archaea possess efficient and robust novel DNA repair pathways, allowing survival in such conditions. Such extreme conditions can also be found beyond Earth, such as on the surface of Mars and investigation into DNA repair in the archaea represents a pivotal stepping stone to understanding how organisms adapt to "Martian" conditions. In the work presented here, the gene encoding a XerCD-like integrase, found within the integrated prophage on the main chromosome of *Haloferax volcanii* was isolated from a genomic library and overexpression leads to increased resistance to genotoxic stress imposed by ultraviolet light and mitomycin C. Deletion of this gene does not impact the growth rate or sensitivity to DNA damaging agents, likely due to the presence of an additional eleven homologs within the genome. Deletion of the entire prophage region does not delete all XerCD homologs, yet shows an even larger increase in cell survival after UV and MMC treatment. Genetic analysis suggests that XerCD may interact, directly or indirectly, with UvrC as overexpression of XerCD somewhat mitigates the UV sensitivity seen in a UvrC deleted strain. The defect is not fully complemented, so further study is required. Previous data has shown that various XerCD-like integrases are upregulated in the presence of MMC. Real-time PCR carried out here indicates that phage induction may occur when cells are treated with DNA damaging agents, which may contribute to the cell death seen, and therefore strains deleted for the integrated prophage may be more beneficial to use for DNA damaging assays due to increased cellular survival. Replication Protein A transcripts are upregulated in response to MMC, aiding in interstrand crosslink repair. The interplay between XerCD-like integrases, that usually function in DNA replication, and DNA repair warrants further study.

Chapter 1: Introduction

1.1 Life beyond Earth: adaptation to the extreme

1.1.1 Introduction to life beyond Earth

Adaptation to extreme environments, specifically of those beyond planet Earth, will play a pivotal role in our understanding of how life can exist on planets more extreme than our own and will thus provide useful insights into how harsh planets could support human life. Over the last few decades, humans are gradually overcoming technological challenges of deep space missions and the possibility of colonization of extra-terrestrial outposts is being seriously considered by various space agencies (NASA report, 2015). Among various suggestions, and fuelled by potential naivety, Mars arises at the forefront of public minds as a potential colonization destination that could support human life (McKay, Toon and Kasting, 1991). Mars is thought to provide a much more habitable environment for the future colonists from Earth compared to other planets in proximity to Earth, recent positive results from rovers and radar equipment have shown a 20km-wide lake of sub-glacial liquid water may be present (Orosei *et al.*, 2018). Moreover, various mineral sources and metal ores have also been discovered, providing putative steps, despite the absence of practical proof of extraction, towards potential colonization (Brasser and Mojzsis, 2017). If we assume that life would be based on carbon, it is encouraging that 95% of the atmosphere at the Mars surface is CO_2 – this could be used for production of organic carbon. In contrast, oxygen and nitrogen levels are sparse, however anaerobic organisms evolved first on Earth, so the same could be true here.

It should, however, be noted that in spite of the positive aspects of potential colonization discussed above, the act of getting to Mars provides the first key

roadblock. A major obstacle to human space exploration is the limitations imposed by the environmental conditions beyond Earth, and the effects of long-term exposure to such conditions that can not be accurately predicted. Exposure to intense solar (ionizing) radiation is of great concern, as is the long exposure to cosmic rays in deep space, both of which lead to significant DNA damage, which will be discussed later. It is these conditions, along with a host of others that form the key to survival outside of Earth.

Aside from Mars, Europa has also been explored for potential to contain life, although by Earth standards this will still be an extreme environment. An ice-covered moon of Jupiter, Europa might at first seem incompatible for life. However there is evidence of the presence of a salty ocean (Pappalardo *et al.*, 1999). The ice layer provides the most likely habitat for life on Europa, along with the seafloor. The stress of lethal radiation and low temperature near the surface impede the possibility of life on the surface. Suitable temperatures and liquid water are only found at the base of the ice layer, however model simulations have demonstrated that the hypothetical oceans would be too cold to support biological activity ($T < 253K$). These simulations further demonstrate high salinity, suggesting if life was found, these organisms would be halophilic. Conditions on Europa, as well as Earth analogues are reviewed in (Marion *et al.*, 2003).

Biological building blocks, such as amino acids, are key to life. Recently glycine, the simplest of the amino acids, was shown to be present on a comet via samples from NASA's Stardust spacecraft, showing that amino acid formation can occur abiotically (Elsila, Glavin and Dworkin, 2009). Thus a mechanism for amino acid formation on icy planets such as Europa and Mars, would be that amino acids could be transported via interactions with comets on the planet surface, thereby transmitting amino acids such as glycine (Pierazzo and Chyba, 2002). In addition, various other biochemical compounds have been found in extra-terrestrial objects such as carbonaceous meteorites. These include purines and pyrimidines (Burton *et*

al., 2012). These provide the potential for seeding planets with building blocks, as well as complicating the detection of “native” life on planets.

At the crux of the matter is the need for resistance to harsh conditions such as radiation and polymerization of bio-molecular monomers. A recent study investigated the polymerization of amino acids in the solid ice crust of icy moons such as Europa, their calculations showed that low-temperature conditions such as those found on Europa, are thermodynamically favourable for dehydration synthesis of organic monomers into polymers. However, it does not follow that the reactions are kinetically possible in such extreme environments (Kimura and Kitadai, 2015). Overall, we are nowhere near considering the possibility of a potential human colonization of planets such as Mars or Europa, however various results have come to light, such as availability of amino acids and organic compounds, that indicate life, although not human, may well exist somewhere. If it does, such life must be able to cope with incredibly extreme conditions of temperature, salinity, pressure and radiation – such organisms, and their adaptions, will be discussed in section 1.1.3. Matters such as the means of arrival and obtaining of nutrients and water become trivial if the more significant problems of harsh environmental conditions can not be mitigated in some fashion. To this end, delving into the genetics of extremophiles could provide a small step in the right direction.

1.1.2 Environments as a proxy for Mars

Adaption to Martian conditions is of great importance as the selection and adaption of life to exist in such harsh conditions provides a good insight into life on Mars, if it exists – applications of this to life detection technology could assist further in this regard. Simulations of Martian conditions lie at the centre of such technology and scientific advancement.

Several possibilities for environments analogous to Mars exist. Firstly, the microbial communities that live in rocks, termed “endoliths”. These communities contain various prokaryotes, lichens and algae and exist in microhabitats that could also be present on Mars. Such endolithic communities have been found on Earth in deserts such as the Dry Valleys of Antarctica – owing to Mars’ cold environment, this provides a good proxy for the low-temperature of martian conditions and is used as a model system (Friedmann, 1982). It is notable however that it is not resistance to cold conditions that allows life in this case, it is the creation of a microenvironment between rock that buffers the extreme cold and rapid temperature fluctuations as well as retention of water and fixed nitrogen. Therefore, this habitat might provide a proxy to Mars. In addition, such microenvironments may protect such communities from the effects of UV radiation (Rothschild, 1990).

Moving from the extreme cold found on Mars, Rothschild proposed a new analogue – organisms found in high salt conditions such as halite and gypsum. Although now an old paper, this proves to still be part of current thinking. Namely, that the abundant sulphate salts found on Mars could contain viable microbial communities. Various halophilic and acidophilic prokaryotes and eukaryotes have been documented in Mars analogue gypsum (Benison and Karmanocky III, 2014). Again, it should be noted that although this seems positive, these organisms, if they did exist in this form on Mars, may well not be able to cope with other environmental stresses such as radiation.

Unlike Earth, there is a significant UV flux on Mars due to UV-B and UV-C radiation. Various work on the biological effects of UV radiation have shown that Martian UV flux would not prevent life per se, although prolonged exposure to such radiation would mean life could not be sustained for long, indeed there was 100% cell death after 30 minutes (Cockell *et al.*, 2005). Such analogues can be created with relative ease in the laboratory.

Recent work has led to the theory that salt water may periodically form on the surface of Mars, which, combined with previous studies looking into salt crystals, may provide a basis on which life could form. It is for this reason, adaption to high salinity is an important consideration as to how organisms might survive in Martian conditions, a prime example are the Haloarchaea, which will be discussed later, but may provide a useful proxy to early life on Mars (Schorghofer, 2020).

The Mars environment is one of extreme desiccation. NASA use the Atacama Desert (Chile) in order to find life and conduct research. If life exists on Mars, or ever did exist, the planet's surface dryness and extensive radiation exposure would drive it underground – thus locations like Atacama are good place to look for life on Mars (NASA's ARADS project).

A summary of important Martian proxies can be found in Table 1.1.

Table 1.1 | Martian proxies found on Earth

Organism	Earth microenvironment	Martian proxy	Reference
Endoliths	Porous / cracked rock	Martian rockbeds	(Friedmann, 1982)
Hypothermophilic microbes	North polar ice cap / Antarctica	Low temperature of Mars (average -60 degrees Celsius)	(Friedmann, 1982) (McKay, 2010)
Salt-tolerant microorganisms (e.g. Haloarchaea)	Gypsum salts, hypersaline lakes such as the Dead Sea	Salt crystals (halite), salt lakes/water	(Mullakhanbhai and Larsen, 1975, Jaakkola *et al.*, 2016)
N/A	Atacama desert, Chile	Desiccation of Mars	NASA (ARADS)

1.1.2.1 Microgravity and Hypergravity

Microgravity has significant effects on the growth and survival of various microorganisms, including effects on human health (Taylor, 2015). The effects of microgravity can be simulated using a variety of technology and devices, although this can also be achieved through carrying out experiments on the International Space Station (ISS) (Horneck, Klaus and Mancinelli, 2010). Interestingly, various work has indicated that microgravity stimulates the growth of microorganisms and alters the production of antibiotics. Furthermore, there has been evidence of increased pathogen virulence due to microgravity, including in *Salmonella typhimurium* that was grown aboard Space Shuttle mission STS-115 and proteomic analysis compared to a ground-based control, the RNA chaperone, Hfq, was identified as a possible regulator of this process (Wilson *et al.*, 2007).

Analogous environments on earth can be simulated. Deguchi and co-workers studied the effect of hyper gravity by using acceleration via centrifugal rotation, here *E. coli* was one of the bacterial species to survive and still perform binary fission (Deguchi *et al.*, 2011). This type of experiment could be reproduced using archaeal species.

1.1.3 Extremophiles and their environments

Extremophiles, by definition, are organisms that thrive in habitats classified as extreme, which for other life-forms are intolerable. Such environments include extreme heat such as volcanoes and hot springs, anoxic environments and areas of intense radiation and chemical stress, intense cold such as Antarctic ice caps all the way through to highly salty, acidic and alkaline environments such as the Dead Sea, which contains dissolved salts of approximately 322.6 gm/litre (Nissenbaum, 1975). Extremophiles have been found in all of these places and are able to adapt to survive in such environments due to incredible flexibility in cell metabolism and genetic regulation, especially in response to osmotic stress due to salinity indicating that Halophillic extremophiles would be interesting to look at (Ambily Nath and Loka Bharathi, 2011).

For the most-part, extremophiles are prokaryotes, namely the Bacteria and Archaea. Microorganisms are classified based on the harsh environments within which they grow, the predominant classifications are presented in Table 1.2.

Table 1.2 | Extremophile classification and their environments

Microorganism nomenclature	Environmental adaption	Reference
Acidophile	Optimum pH < 3	(Gómez, 2011)
Alkaliphile	Optimum pH > 10	(Horikoshi, 1999)
Halophile	At least 1-5.2M salt concentration	(Antón, 2011)
Hyperthermophile	Optimum growth > 80°C	(Berenguer, 2011)
Thermophile	60°C ≤ Optimum growth ≥85°C	
Piezophile / Barophile	High pressure (>40 Mpa)	(Prieur, 2014)
Endolithic	Grows inside rock	(Walker and Pace, 2007)
Radioresistant	High radiation tolerance	(Romanovskaia, Rokitko and Malashenko, 2000)
Metallotolerant	Tolerance to heavy metals	(Barman, Jha and Bhattacharjee, 2020)
Xerophile	Tolerance to low water availability, resistant to desiccation	(Grant, 2004)
Cryophiles / Psychrophiles	Tolerance to low temperatures	(Margesin and Miteva, 2011)

1.1.3.1 Thermophiles

Thermophiles are found in areas of extreme high temperature such as volcanic chambers and hydrothermal vents, the latter of which is of special relevance to molecular biology owing to the discovery of *Thermus aquaticus* and Taq DNA Polymerase (Chien, Edgar and Trela, 1976).

Thermophiles have several mechanisms to support their existence in extreme temperatures. Thermostable proteins and cellular machinery are of great importance, molecules such as ATP, amino acids and polypeptides are altered to increase thermo stability above 250°C. The number of charged residues on protein surfaces is increased, as well as increased flexibility at high temperatures to prevent non-specific chemical interactions (White, 1984, Goldstein, 2007). As would normally be predicted, given the hydrogen bonding between Guanine and Cytosine versus Adenine and Thymine, GC content of thermophilic genomes would be higher than usual. Despite this assumption, there appears to be no trivial relationship between growth temperature and whole-genome GC-content (Hurst and Merchant, 2001). However, the authors make note that the GC content of 16S and 23S structural RNA is higher at increased temperatures.

Thermophilic bacteria can change the fluidity of their membranes to resist increased temperature by increasing the number of branched chain iso-fatty acids and saturated fatty acids (Patel, Skerratt and Nichols, 1991). Archaeal membrane chemistry, and how this leads to adaption to harsh conditions, will be discussed in a later section.

1.1.3.2 Psychrophiles

Psychrophiles live in extremely cold environments such as glaciers, ice caps and deep oceans, even the upper atmosphere. These organisms have an approximate upper temperature tolerance of 20°C (Margesin and Miteva, 2011). Such organisms have been shown to carry out DNA replication and cellular metabolism at -15°C (Mykytczuk *et al.*, 2013). Various psychrophiles have been isolated from Antarctic environments and at the forefront of this are the Archaea, specifically euryarchaeotes (Kurosawa *et al.*, 2010).

Adaptions to such environments include Cold Shock proteins (CSPs), mainly studied in bacteria, that act as adaptive proteins in cold environments. These proteins are small and bind to RNA via their Cold Shock domain to preserve its ssRNA conformation. Cold Shock proteins also have additional roles besides that of a chaperone. Such proteins do exist in some Archaea, although are mainly absent from thermophiles (Jones and Inouye, 1994).

Although generally considered absent in Archaea, a putative DEAD-box hRNA helicase gene was characterised in the Antarctic Archaeal methanogen *Methanococcoides burtonii*. This protein may be involved in RNA binding and is regulated by low temperatures. Interestingly, analysis of the 5'-UTR lead to the discovery of an 11bp sequence that matches closely cold box genetic elements present in the 5'UTR of cold-shock induced genes in bacteria. This, along with genetic studies, provides evidence that gene expression from a cold adapted Archaeal species involves bacterial-like transcriptional machinery regulation. It also provides the prospect of further analysis on DEAD-box RNA helicases in Archaea and their role in cold-adaption, presumably by curing cold-induced secondary structures within mRNA molecules (Lim, Thomas and Cavicchioli, 2000). Exact details remain unclear.

1.1.3.3 Halophiles

Halophiles are microorganisms that require, or prefer, a highly saline environment for growth. Such organisms can be found in salt lakes, oceans and in space trapped in gypsum or halite (covered in section 1.1.2). Halophiles are found in each of the three domains of life: Archaea, Bacteria and Eukarya and their metabolic diversity is rather large: oxygenic, anaerobic, phototrophs, fermenters and methanogens to name a few (Oren, 2002).

Most halophiles require a salt concentration of 1.5M to grow and maintain cellular structural integrity (Antón, 2011). Different strategies are employed by halophiles as an antidote to osmotic stress caused by increased salinity. Some bacteria accumulate inorganic ions in their cytoplasm, the "salt-in" strategy as well as evolving proteins that are more salt-tolerant (Siglioccolo *et al.*, 2011). A "salt-out" strategy is employed by other archaeal species whereby small organic solutes are stored within the cytosol, in order to balance the extracellular salt concentration. These, and other mechanisms, will be discussed in more detail within the Haloarchaea section.

Comparative analyses of Archaeal halophile vs. non-halophile proteomes have been carried out and the amino acid composition of halo-tolerant enzymes contains an increased proportion of aspartic and glutamic acids. These major differences are on the surface of the proteins and allow cooperation with electrostatic interactions and increased salt bridge formation (Tadeo *et al.*, 2009). Acidic amino acids lead to the negative charge on the proteins surface and this determines stability. Furthermore, these negative charges are thought to be important for halophilic protein solvation in order to prevent denaturation. Nucleoside diphosphate kinase from *Halomonas* sp. exhibit halophilic characteristics; In a study in 2008, residues 134 and 135 in the carboxyl-terminal region were shown to be Glu-Glu, whilst its non-halophilic homologue in *Pseudomonas* were Ala-Ala. The double mutation E134A-E135A in this nucleoside diphosphate kinase resulted in loss of acidic residues and thus a loss of halophilic characteristics (Tokunaga,

Arakawa and Tokunaga, 2008). This indicates acidic residues in the C-terminal region are important for halophilicity.

In addition, reduction of hydrophobic residues is also common in halophillic proteins.

Lipid composition is also important in halophilic properties and will be discussed in detail in section 1.3.6.

1.1.3.4 Xerophiles

Xerophiles are tolerant to areas of low water activity, there are two major studied environments that provide habitats for xerophilic organisms: foods preserved by dehydration or enhanced sugar levels, and hypersaline sites (such organisms are therefore also halophiles) (Grant, 2004). In 2015 a microbial community was unearthed in the deep-sea brine lake Kryos in the Mediterranean Ridge, this lake is filled with salts such as $MgCl_2$ and has low water activity. Data from this paper suggests metabolism of halophilic bacteria and Archaea down to a water activity (a_w) of 0.611 (Yakimov *et al.*, 2015). Wooden owls, purchased from Thailand have been shown to contain xerophillic species of *Aspergillus* fungi, showing that such organisms can be found in unexpected places (Hallsworth, 2019).

The environmental conditions on Mars would supposedly favour xerophilic organisms owing to the lack of water and extreme desiccation. A halophilic and osmophilic form of *Bacillus megaterium* was isolated from the Nubian Desert. This particular strain was able to grow under extreme desiccation, along with an osmophilic strain of *Mycococcus ruber* isolated from Antarctica. Both strains have been shown to grow under simulated Martian conditions and that xerophily and halophily may well be linked, an assumption backed

up by the fact that there was a high incidence of xerophilic forms among isolated halophilic bacteria isolated from various salty muds and soils (Imshenetsky, Kouzyurina and Jakshina, 1973).

1.1.3.5 Radio-resistant bacteria and archaea

The bacterium *Deinococcus radiodurans* is a front runner of extreme radiation resistance, best known for its efficient DNA repair systems and resistance to gamma radiation (Ferreira *et al.*, 1997, Romanovskaia, Rokitko and Malashenko, 2000). *D. radiodurans* culture are resistant to various types of DNA damage such as ionizing radiation and hydrogen peroxide as well as desiccation (Mattimore and Battista, 1996). This bacterium can survive gamma radiation exceeding $1.5x10^4$ Gy. A RecA homolog in strain R1 has been identified to provide a radiosensitive phenotype when mutated in *D. radiodurans*, linking radioresistance to DNA repair mechanisms, specifically the strand exchange mechanism in homologous recombination, facilitated by the RecA protein (Daly *et al.*, 1994). Cultures exposed to 1 – 1.5 Mrad irradiation sustain ≥120 double strand breaks per chromosome, which is subsequently repaired over a number of hours with 100% survival and a significant lack of mutations (Daly *et al.*, 1994, Minton, 1994). Of note, N-methyl-N'nitro-N-nitrosoguanidine (MNNG) is a mutagen to which *D. radiodurans* is sensitive (Sweet and Moseley, 1976). As a result, MNNG-mutagenized strains that are ionizing radiation sensitive have been isolated and subsequent complementation assays have been performed to restore radiation resistance. Essential genes appear to be involved in homologous recombination, and include RecA (*recA*) and DNA Polymerase 1 (*polA*), which is predictable considering the nature of the double strand breaks formed from ionizing radiation (Nowosielska *et al*, 2006). Mechanistic details of

homologous recombination and double strand break repair will be covered later (Gutman *et al.*, 1993, 1994, Udupa *et al.*, 1994). Of note, the *E. coli polA* gene was expressed in *polA*$^{-}$ *D. radiodurans* and there appeared to be complete restoration of wild-type phenotype with regards to resistance to forms of DNA damage including ionizing radiation. This indicates that the DNA Polymerase 1 in *D. radiodurans* can be compensated for by *E. coli* DNA Polymerase 1 and thus does not possess any strain specific properties that allow resistance to radiation and other forms of DNA damage (Gutman, Fuchs and Minton, 1994). Conversely, expression of the *D. radiodurans recA* gene in *recA- E. coli* cells, results in death of the recipient cells (Gutman *et al.*, 1994).

The nuclear and genome organisation of *D. radiodurans* provides an exciting model for radio resistance studies, yet little is known about its chromosomal organisation. It is hypothesised that its chromosomal organisation may play a role in radiation resistance (Passot *et al.*, 2015). Although it is reported that *D. radiodurans* is polyploid, a key question is why other polyploid organisms do not exhibit the same extent of radio resistance. Needless to say, resistance to radiation is of key importance when assaying the potential of Mars colonization and provides one, incredibly important, adaption required for life beyond Earth. Radiation dose rates for Mars have been predicted by HZETRN, a NASA computer model, and range from 200-300 mSv/yr (up to 0.03 krads), yet *D. radiodurans* can survive upwards of 1,500 krads of gamma radiation and thus provides a good foundation for research into the viability of life on Mars (Daly and Minton, 1996, Zeitlin *et al.*, 2004).

In terms of Archaeal resistance to radiation, *Thermococcus gammatolerans* represents one of the most radioresistant organisms amongst the Archaeal domain. Its genome has been characterised and sequenced, and contains a circular chromosome of 2.045 Mbp with no extra-chromosomal elements and encodes for 2157 proteins (Zivanovic *et al.*, 2009). Belonging to the Euryarchaeota phylum, *Thermococcus* is accompanied by two other genera, *Pyrococcus* and *Palaeococcus* (Fiala and Stetter, 1986, Takai *et al.*, 2000). Members of the *Thermococcus* genus have been isolated from hot-springs

and deep oil reservoirs (Miroshnichenko *et al.*, 2001). *T. gammatolerans* was isolated from hydrothermal chimneys in the mid-Atlantic Ridge and strain EJ3 was obtained by culture enrichement after irradiation with gamma radiation at doses of 30kGy (~3000krads) and withstands up to 5kGy (500krads) of radiation without any detectable lethality (Jolivet *et al.*, 2004, Tapias, Leplat and Confalonieri, 2009) . Such radiation resistance cements this archaeon as one of the most radioresistant Archaea, considering Archaea and Eukarya share many genetic and biochemical processes, further analysis into strain EJ3 may provide useful insights into radiation resistance along with *D. radiodurans* (Jolivet *et al.*, 2003).

1.2 Introduction to Archaea

1.2.1 The discovery of Archaea

In the 1960s Carl Woese used 16s rRNA sequence comparisons to infer genealogical relationships of organisms. Woese proposed that a third domain of life existed, and thus the conventional two domain tree of life should now be updated to contain a third branch – the archaeabacteria (Woese and Fox, 1977). Towards the latter part of the 1980s, phylogenetic trees constructed from the knowledge of gene duplications showed that eukaryotes and archaeabacteria were in fact sister groups on the tree, and the archaeabacteria were subsequently renamed Archaea (Iwabe *et al.*, 1989, Woese, Kandler and Wheelis, 1990) (Figure 1.1).

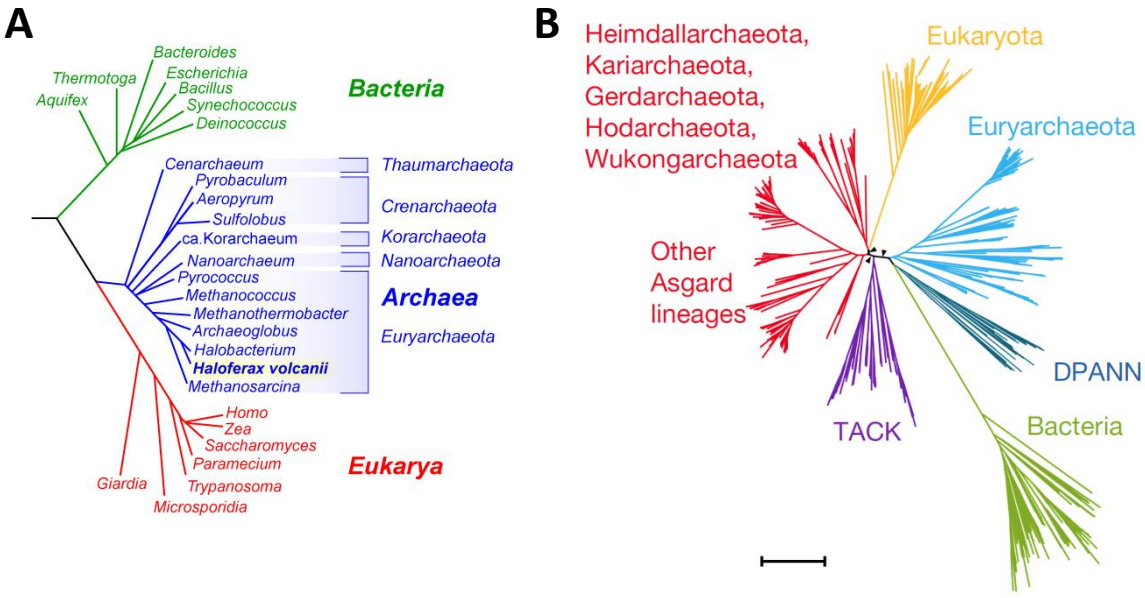


Figure 1.1 | **Depiction of the phylogenetic tree constructed via 16S rRNA sequencing.** A) Bacteria, Eukarya and Archaea now form three distinct domains of life. Figure adapted from (Allers and Mevarech, 2005). B) Updated model of two domain tree. Eukarya form a branch within the ASGARD clade (Liu *et al.*, 2021) .

The number of small subunit rRNA sequences became large enough that the three domains of life could be reliably analysed. Deeper analysis of rRNA sequences revealed two phyla – Crenarchaeota and Euryarchaeota (Winker and Woese, 1991) and advancement in sequencing technology and lineage sampling has subsequently defined various other phyla such as Korarchaeota, Thaumarchaeota and Aigarchaeota. The aforementioned groups were shown later to be part of a monophyletic group known as the TACK superphylum, along with Crenarchaeota from which it is hypothesised the eukaryotic cell emerged (Guy and Ettema, 2011), thus the most current phylogenetic tree displays eukaryotes as direct descendants from the Lokiarchaeota, lending weight to a two domain phylogenetic tree (Koonin and Yutin, 2014) (Figure 1.1) (Doolittle, 2020, Williams *et al.*, 2020). The debate between a two-domain and three-domain phylogenetic tree remains, especially since the discovery of the ASGARD superphylum pointing towards a two domain tree (Zhou *et al.*, 2018, Doolittle, 2020) (Nasir, Mughal and Caetano-Anollés, 2021).

That being said, however, recent analysis of eukaryotic signature proteins (ESPs) and their homologues in the ASGARD Archaea, along with additional phylogenetic analysis points to the eukaryotes having emerged from within the ASGARD Archaea, specifically as a sister group to the Hodarchaeales, a newly proposed order within the Heimdallarchaeia (Eme *et al.*, 2023) (figure 1.2).

A question emerging from the topic of eukaryogenesis is how eukaryotes are surrounded by a bacterial-like membrane (phospholipids) yet the nuclear contents is similar to that of ASGARD Archaea. Any eukaryogenesis with an ASGARD host would require a membrane replacement step, as the ether-linked archaeal membrane would be incompatible with extant eukaryotic membrane chemistry and the enzymology of membrane synthesis. The modern view is one of two separate endosymbiotic events, first the engulfing of an ASGARD archaeon (and subsequent membrane loss), followed by a

second to give rise to mitochondria (figure 1.3) (Krupovic, Dolja and Koonin, 2023).

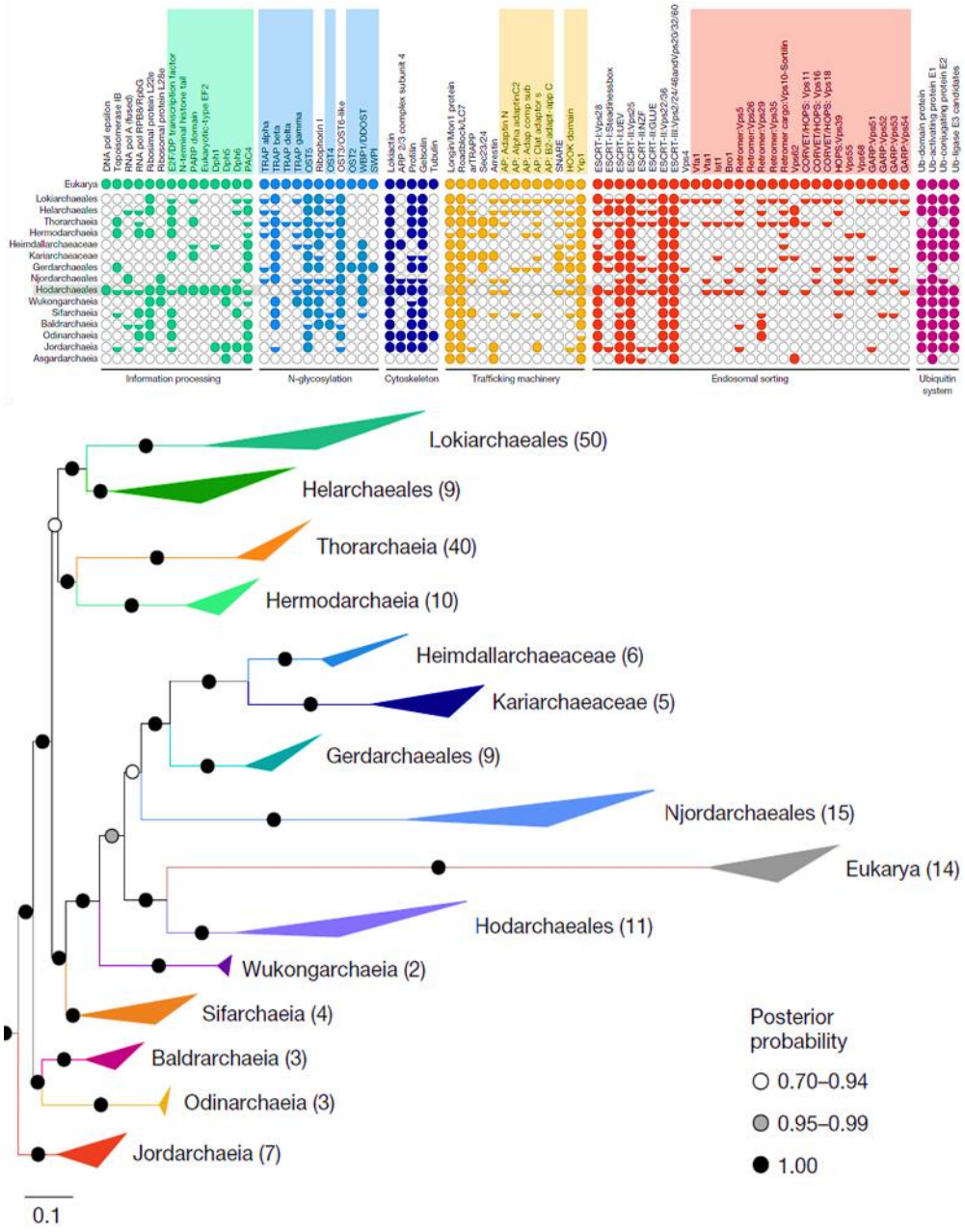


Figure 1.2 | **Overview of ESPs within the ASGARD Archaea, and the current position of the eukaryotes.** Eukarya (grey) are shown as a sister group to the Hodarchaeales. Figure taken from (Eme *et al.*, 2023)

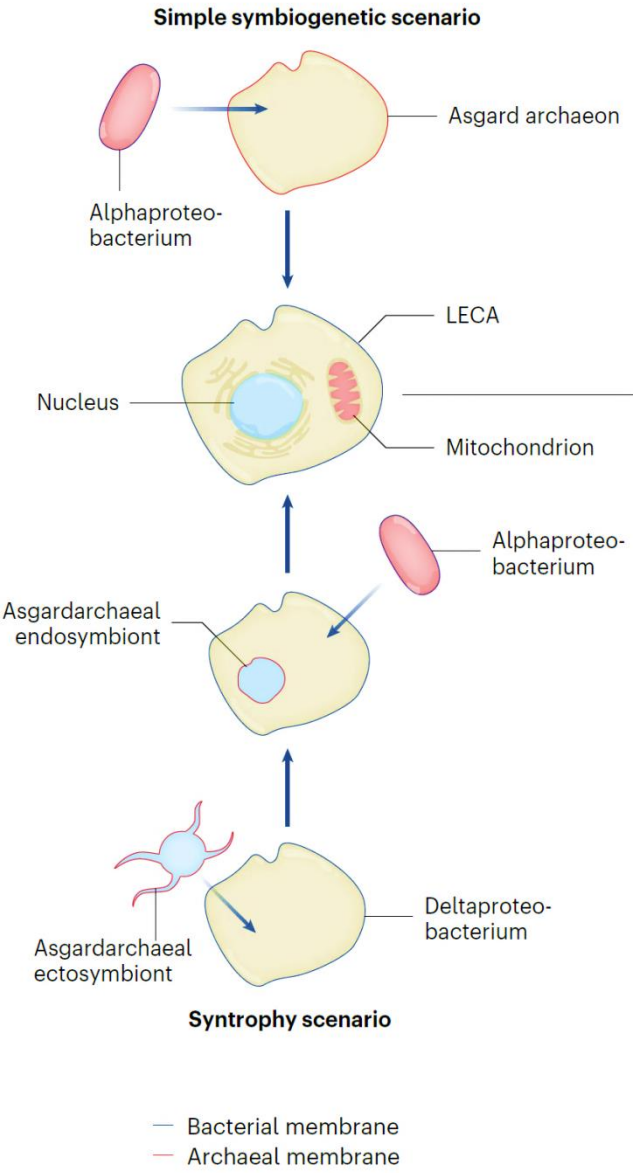


Figure 1.3 | **Models of eukaryogenesis.** Two alternative scenarios of eukaryogenesis that include either one (top) or two (bottom) endosymbiotic events. The two-step endosymbiotic model is consistent with not having a membrane replacement step to generate a bacterial-like eukaryotic cell membrane. Figure taken from (Krupovic, Dolja and Koonin, 2023)

1.2.2 Characteristics of Archaea

Although not all extremophiles are Archaea, the majority of known Archaea tend to be extremophiles and are some of the most hyperthermophilic, acidophilic, alkaliphilic and halophilic microorganisms known (Rampelotto, 2013). Archaeal cultures have been isolated from places such as hot springs, hydrothermal vents and salt lakes as well as more moderate habitats such as oceans and fresh water pools (Chaban, Ng and Jarrell, 2006). The overall cellular architecture of Archaea is similar to that of bacteria, hence their original classification as “Archaeabacteria”.

In spite of this, at a molecular level the two are far from similar, key is the unique lipid chemistry that helped to clearly delineate the Archaea from other organisms (De Rosa and Gambacorta, 1988). Archaea use isoprenoid chains linked via ether bonds to glycerol 1 – phosphate whereas in bacteria and eukaryotes, fatty acid chains are attached to glycerol – 3- phosphate via ester bonds. Recently the Archaeal G1P dehydrogenase and bacterial G3P dehydrogenase enzymes have been studied to try and resolve the “lipid divide” as the evolution of archaeal G1PDH will provide more information about how the transition from LUCA led to the archaeal domain (Carbone *et al.*, 2015). Some archaea use a lipid monolayer instead of a bilayer, fusing two phospholipid tails into a single chain with two polar head groups (Hanford and Peeples, 2002) (Figure 1.4). Most Archaeal cell membranes are also surrounded by an S-layer, which can be glycosylated. In the laboratory this can be problematic as it causes archaea to resist DNA uptake, however this can be circumvented during the transformation protocol if removed by chemicals such as EDTA (Cline *et al.*, 1989). Archaea also lack canonical peptidoglycan and use an unmodified initiator methionine in translation, similar to eukaryotes (Meyer and Albers, 2014). Furthermore, Archaeal DNA replication, transcription and translation resemble diluted forms of the eukaryotic system (Barry and Bell, 2006). Genome structure resembles both

bacterial and eukaryotic systems as they have circular chromosomes and arrange genes in operons, however chromosomes can possess a singular origin of replication or multiple origins as is the case in eukaryotes (Koonin and Wolf, 2008). Some archaea can also carry out DNA replication without an origin of replication, this will be discussed further in this thesis.

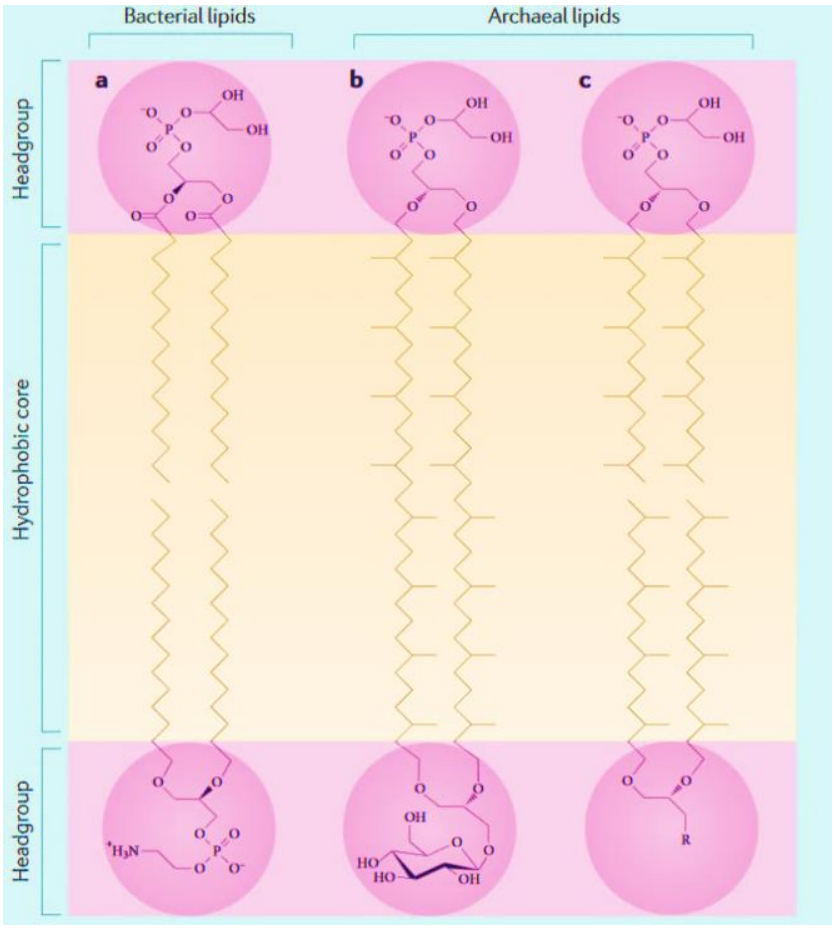


Figure 1.4 | **Outline of phospholipid structure between Archaea and Bacteria**. Figure taken from (Albers and Meyer, 2011).

1.2.3 Halophilic Archaea

Halophilic Archaea (known as Haloarchaea) inhabit hypersaline (0.5 – 4M) environments (up to saturation) such as alkaline salt lakes and the Dead Sea (Oren, 2008). To tolerate such osmotic stress, multiple mechanisms have arisen. A "salt-in" strategy is used where KCl and NaCl are accumulated in the cytoplasm until the overall concentration equals that of the outside. K^+ and Cl^- appear to be the key players as intracellular Na^+ concentration is kept fairly low, there is still no clear reason for this, however all halophilic microorganisms have transport systems based on Na^+/H^+ antiporters, to remove sodium from inside the cell (Oren, 1999, 2002, Grant, 2004). A theory is that K^+ ions hydrate less water molecules than Na^+, thus increasing the water activity of the cell cytosol in which chemical reactions can occur, in addition the water molecules within the hydration shell of K^+ are more disordered than those hydrating Na^+, leading to increased entropy of the system (Mancinelli *et al.*, 2007). To maintain cellular metabolism, intracellular proteins must adapt to function in hypersaline conditions, as a consequence most haloarchaeal proteins have an acidic *pI* and depend on high salt concentration for function (Siglioccolo *et al.*, 2011). Alternatively a "salt-out" strategy is employed, predominantly by Halophillic bacteria and eukaryotes, where the cell accumulates small organic molecules and solutes such as 2-Sulfotrehalose to balance the external salt concentration (Desmarais *et al.*, 1997). Haloarchaea are of specific note owing to their ability to tolerate hypersaline environments. Salt crystals such as Halite can be found on Mars and may provide a clue as to how life might exist in such conditions (Jaakkola *et al.*, 2016). In addition, recent analysis of five halophilic strains demonstrated that the proteomes of strains most dependant on $MgCl_2$, showed a higher tolerance towards salts abundant in Martian brines (Carré *et al.*, 2023).

1.3 Archaeal adaption to extreme environments

1.3.1 Polyploidy

Many eukaryotic organisms are polyploid, ranging from animals and plants to unicellular organisms. Polyploidy is even observed in the human liver (Wang *et al.*, 2017) and in various cancers (Zack *et al.*, 2013, Bielski *et al.*, 2018, Was *et al.*, 2022). Polyploidy has been used in the food industry and agriculture to increase the size of various fruits such as kiwifruit, naturally sourced polyploidy strawberries are also found, ranging in ploidy from diploid (2n) to decaploid (10n) (Wu *et al.*, 2012, Yang and Davis, 2017). Indeed polyploidy in crop plants is common, and increases seed and fruit yield.

In contrast to eukaryotes, bacteria are considered, in the main, to be monoploid, that is, they contain only a single chromosome set (Pecoraro *et al*., 2011). However a notable exception to this has already been covered in this review, namely *D. radiodurans*. This bacterium is resistant to radiation and desiccation, and most likely the former arises as a by-product of the latter (Mattimore and Battista, 1996). *D. radiodurans* has between 5-8 copies of its chromosome and can efficiently regenerate and repair entire chromosomes from overlapping fragments that have arisen from extreme chromosomal shattering. This process involves homologous recombination, which will be discussed in detail later, which is only possible due to homologous chromosomes being available for strand exchange, thus a monoploid organism would not survive in such conditions. Survival under such radiation stress provides a key evolutionary advantage to *D. radiodurans*, of which a high genome copy may facilitate recombinational repair of double-strand breaks (Hansen, 1978).

Various species of Haloarchaea have been shown to have polyploid genomes, and in fact a reversal of original thinking is needed here, as monoploid species appear to represent a small minority within the prokaryotic domain of

life. Currently, Haloarchaea appear to be exclusively polyploid, at least within the euryarchaea, suggesting that polyploidy represents a trait among all haloarchaea. Therefore Haloarchaea provide a suitable model genetic system for studying polyploidy as an adaption to extreme environments such as resistance to, or ability to quickly resolve, double strand breaks or re-start DNA replication (Soppa, 2011). *Haloferax volcanii* was observed to exhibit growth-phase-dependant copy number regulation, with between 15 and 20 genome copes during exponential phase and ten genome copes in stationary phase (Breuert *et al.*, 2006).

There are various reasons in the literature as to the evolutionary advantage of polyploidy in halophilic Archaea including:

(1) Firstly, the rate of genomic mutations remains low – the rate of spontaneous mutations in *H. volcanii* was assayed using a *pyrE2* reporter gene and it was discovered that genomic mutation rates were very low compared to other mesophilic genomes (~ 7.5 fold lower). It was indicated that this reduction in mutation rates was due to multiple genome copes in the cell (Mackwan *et al.*, 2007).

(2) *H. salinarum* is highly resistant to gamma irradiation and desiccation. Both genomic insults present as double strand DNA breaks, an incredibly toxic form of DNA damage. When irradiated with between 2.5-7.5 kGy of radiation lead to chromosomal fragmentation within *H. salinarum* which was repaired within 24hrs. However, at the time of the research, it was not known that *H. salinarum* was polyploid, however it does provide an interesting result given the mechanism of double strand break repair absolutely relies upon homologous chromosomes being present, something which only a polyploid organism would have (Kottemann *et al.*, 2005). Despite this being a 'stab in the dark' hypothesis, this has been experimentally justified for *D. radiodurans* and a recent paper noted that a mutant of *H. salinarum* (exhibiting a lethal dose expected to cause death of 50% of an exposed population (LD_{50}) of ≥11 kGy) has been found

and exhibits more resistance to radiation than *D. radiodurans*, indicating a similar mechanism might be in place that relies on homologous recombination-mediated repair. This is further substantiated by the fact that gene expression of single-stranded binding proteins (RPA) was increased (DeVeaux *et al.*, 2007).

(3) The ability to survive over long time periods (millions of years) in salt deposits has been disputed owing to the argument that DNA stability is not high enough to tolerate such environmental longevity, mainly due to cytosine deamination (Sawyer *et al.*, 2012, Lewis Jr *et al.*, 2016). This viewpoint was disrupted by the discovery that Haloarchaea are polyploid, meaning such mutations can be easily corrected by mechanisms involving chromosomal copies as templates for repair.

(4) Gene redundancy presents another positive reason for the existence of polyploidy within Archaea. Some advantages have already been discussed, such as the ability to use wild-type copies of the chromosome to repair damaged duplexes, such as in double strand break repair. This ability results in lower mutagenic load and radiation resistance. However, under unfavourable conditions, mutations in some gene copes arise, resulting in heterozygosity. Selection would therefore act upon individual cells, rather than populations. This has been demonstrated in *H. volcanii* where a heterozygous strain deleted of the *trpA* gene was constructed containing the *leuB* locus in one of two forms: wild-type *leuB* gene or a *leuB:trpA* gene. Growth is only present when both types of genome are present, i.e. heterozygous. Results showed that under growth-inhibiting conditions, genes could be mutated without the loss of wild-type information and thus heterozygous strains could evolve with greater growth capability than the homozygous parent strain (Lange *et al.*, 2011). A further line of enquiry would be to see if *H. volcanii* induces hyper-mutation under unfavourable conditions to survive.

Various other theories exist for the existence of polyploid genomes in Archaea however the common theme amongst all is that of the role of multiple genome copies in homologous recombination as a mechanism for DNA repair. Apart from the mechanism requiring multiple copies of the genome, the ability to repair DNA and the selection for polyploidy means that under limiting conditions, hyper-mutation could be tolerated in order to provide evolutionary advantage to individual heterozygous cells whilst reducing the risks associated with genome mutations.

1.3.3 DNA as a phosphate storage polymer

Of great importance to all extremophiles is the ability to inherit complete sets of genetic information without errors. However, in polyploid organisms such as Archaea, DNA may also serve as a storage molecule for carbon, nitrogen and/or phosphate. In the case of *H. volcanii*, evidence points towards the idea that the polyploid chromosomes are, aside from encoding the genetic information, also phosphate storage polymers. This endows with the ability to grow under an extreme lack of phosphate availability, indicating an intracellular source is used (Zerulla *et al.*, 2014). This was investigated by determining the ability of *H. volcanii* to use environmental gDNA as a source of carbon, nitrogen and/or phosphorous. Of note, significant growth occurred when phosphate was omitted from the environment, showing *H. volcanii* is able to use extracellular DNA as a phosphate source as well as intracellular phosphate stores (Figure 1.5) (Zerulla *et al.*, 2014).

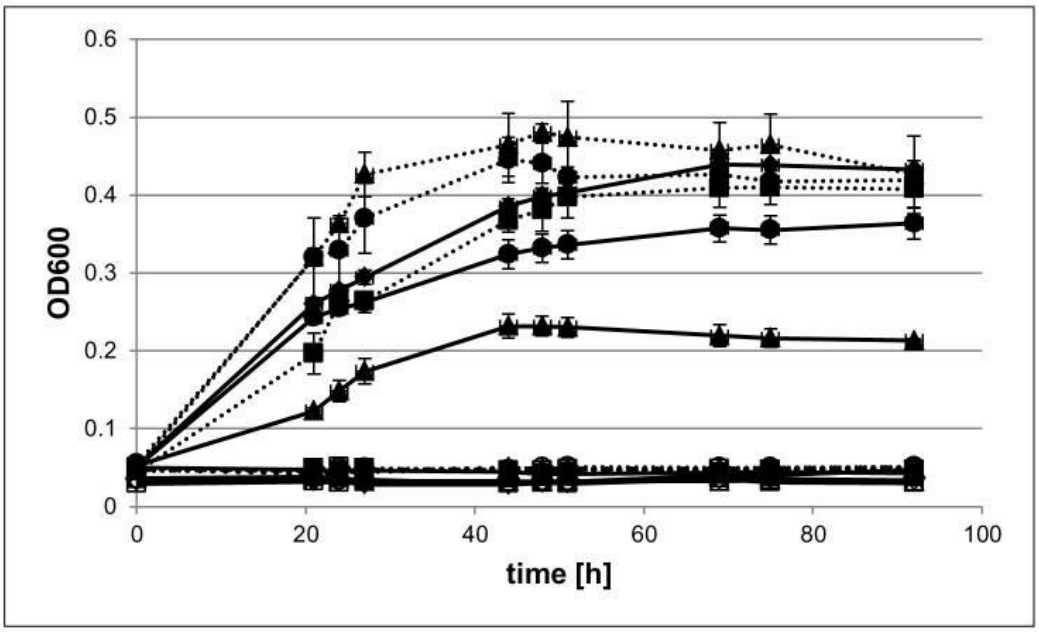


Figure 1.5 | ***H. volcanii* uses external DNA as a nutrient source and contains internal P and N stores.** *H. volcanii* was grown in microtiter plates in synthetic medium with added Carbon (C), Nitrogen (N) and Phosphate (P) as a positive control (diamonds). In additional cultures each one of the three nutrients was replaced with genomic DNA (dotted lines). C was replaced (squares), N was replaced (circles) and P was replaced (triangles). In further cultures each of the respective nutrients was omitted without replacement (solid lines, C=squares, N=circles and P=triangles). Note that DNA was able to complement the loss of phosphorus or nitrogen from the medium. Figure taken from (Zerulla *et al.*, 2014).

Moreover, the addition of external gDNA to cultures lacking any one of the nutrients enhanced the growth yield in all three cases, indicating a clear use of gDNA as a source for Carbon, Nitrogen and Phosphate in *H. volcanii*. The addition of gDNA to cultures lacking phosphate or ammonium resulted in faster growth compared to the control culture grown with all three nutrients present. Further analysis was carried out and can be viewed in (Zerulla *et al.*, 2014), the additional results lend more weight to the ability of *H. volcanii* to

use environmental gDNA, as well as internal DNA stores, as a source of phosphate.

The ability to obtain phosphate is of importance to extremophiles such as *H. volcanii*, specifically in this context, owing to the lack of phosphorous (relative to the abundance of silica, sulphur and iron) on the Martian surface (CLARK *et al.*, 1976, Foley *et al.*, 2008). The ability to use DNA as a self-contained store for phosphate represents yet another key adaption to survival in such conditions.

1.3.4 Halomucin

Halomucin is a large, 927 kDa protein, secreted by Archaea. Halomucin has been characterized in the haloarchaeon *Haloquadratum walsbyi*, Halomucin of strain HBSQ001 consists of 9159 amino acids and its sequence organization is similar to that of mucins found in animals, which protect against desiccation in areas of low water activity (Hollingsworth and Swanson, 2004, Bolhuis *et al.*, 2006). This is done by trapping a collection of water around the halophilic organisms living in low water activity environments. Halomucin resembles the mucous cocoon of lungfish that escape dehydration for several years out of water (Heimroth, Casadei and Salinas, 2018). Similar to that of animal homologues, halomucin contains domains that act as glycosylation and sulfation sites which contribute to its overall negative charge, thus creating an aqueous shield covering the cells. Furthermore, it has been shown that the gene for halomucin in *H. walsbyi* (*hmu1*) is transcribed in full and that *H. walsbyi* encodes two other proteins, which are smaller but analogous to halomucin. These proteins are termed Hmu2 and Hmu3 which are 2885 and 2079 amino acids in length respectively (Bolhuis *et al.*, 2006).

Halomucin has been stained by immunofluorescence which highlights the fact that halomucin is secreted, which was not an obvious theory due to its size being a proposed limit to secretion (Figure 1.6) (Zenke *et al.*, 2015).

H. walsbyi can also synthesize a poly-gamma-glutamate capsule, similar to that found in *Bacillus subtilis*, by means of the protein complex CapBCA which, besides protection against desiccation, forms a cross-linked matrix of poly-gamma-glutamate and contributes to the rigidity and maintenance of the square morphology of *H. walsbyi* (Ashiuchi and Misono, 2002, Bolhuis *et al.*, 2006). The genome organisation of *hmu1* and *capCB* is shown in Figure 1.7.

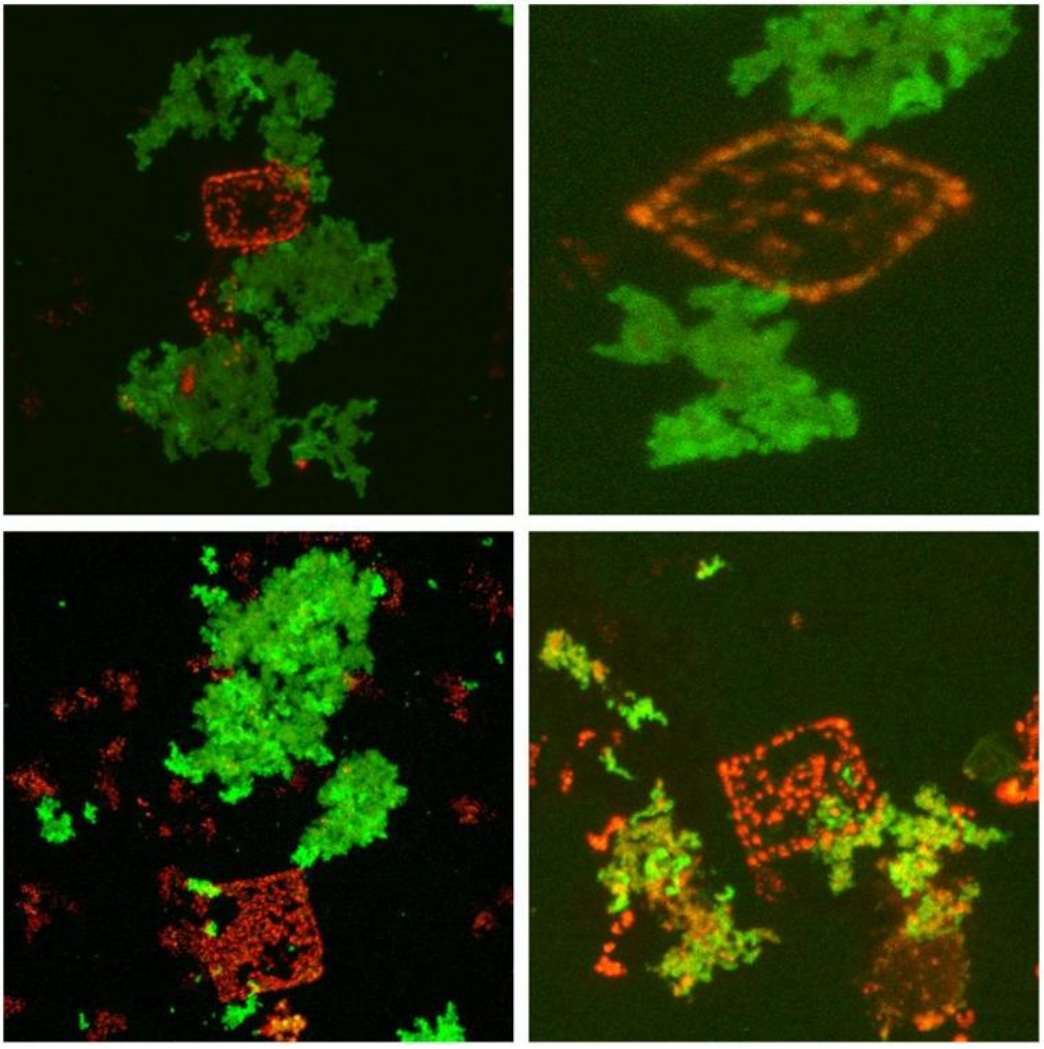


Figure 1.6 | **Immunofluorescence stained halomucin loosely associated with Nile Blue stained *H. walsbri* cells.** Immunofluorescence staining appears green while Nile

Blue staining of polyhydroxybutyrate granules appears red and marks *H. walsbyi* cells. Halomucin is clearly an extracellular protein. Figure taken from (Zenke *et al.*, 2015)

It should be noted that the immunofluorescence results show that halomucin forms in clusters and appears not to associate tightly with the cells, or indeed completely independently. The authors note that this is likely due to experimental disturbances and that in a natural environment devoid of such mechanical stress, loose association is sufficient to exert protective function.

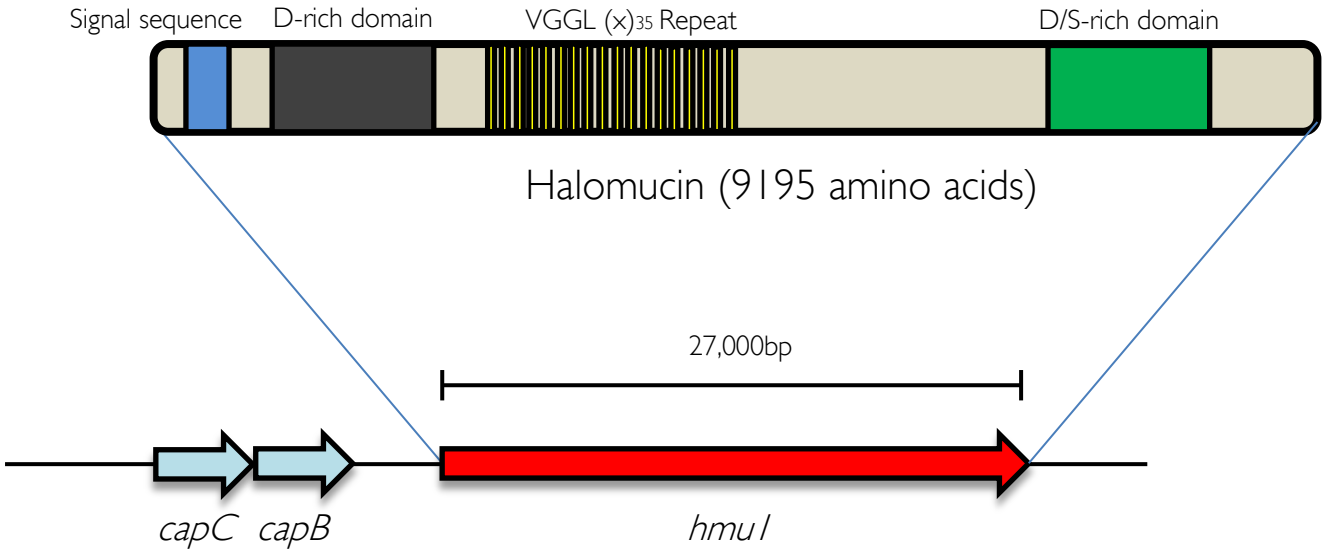


Figure 1.7 | **Halomucin domain organisation and genomic context**. Gene organisation of halomucin in context of *capCB* genes. The halomucin gene is 27kbp and adjacent to *capCB* genes. The protein structure of halomucin is shown. Figure adapted from (Bolhuis *et al.*, 2006).

An additional role of halomucin was suggested by (Zenke *et al.*, 2015), halomucin may be as a protection barrier against phages, however the literature on this is sparse.

1.3.5 Sphere formation in fluid inclusions

Exact, plausible, mechanisms for haloarchaeal survival in highly saline environments have yet to be confirmed. Compared to their bacterial relatives, when looking at their genome sequence, Archaea, in general, do not appear to produce spores, unlike the likes of *Bacillus subtilis* which forms endospores when in harsh environments as protection from extreme heat and radiation. Here Spo0A is a key transcriptional regulator (Errington, 2003, Piggot and Hilbert, 2004) and sporulation represents a key mechanism by which bacteria can resist simulated Martian conditions (Onyenwoke *et al.*, 2004). A strain of *Halobacterium salinarum* has been isolated from a fluid inclusion in a 97,000 year old halite crystal (Mormile *et al.*, 2003), the strain was identified by 16S rDNA sequencing. The presence of small spherical objects within such inclusions has also been reported, along with the signs of spherical cells shown to be within lab-grown halite in 1988 (NORTON and GRANT, 1988, Schubert, Lowenstein and Timofeeff, 2009).

In the absence of spores, it is easy to assume that spherical cells in ancient inclusions represent a haloarchaeal form of stasis. Most Haloarchaea have an S-layer which forms an envelope around the cell, however this does not contain peptidoglycan. Not much is known about the mechanism of sphere formation in Haloarchaea, however a study in 2012 indicated that sphere formation is in response to low external water activity (a_W) (Fendrihan *et al.*, 2012). Haloarchaeal cells (*H. salinarum,* strain NRC-1), stained with a green dye, were embedded in halite. Water activity in the solution surrounding the rod-shaped Haloarchaea was lowered by suspending the cells in Tris-buffered LiCl (4M) ,which lead to an immediate formation of spheres (Figure 1.8).

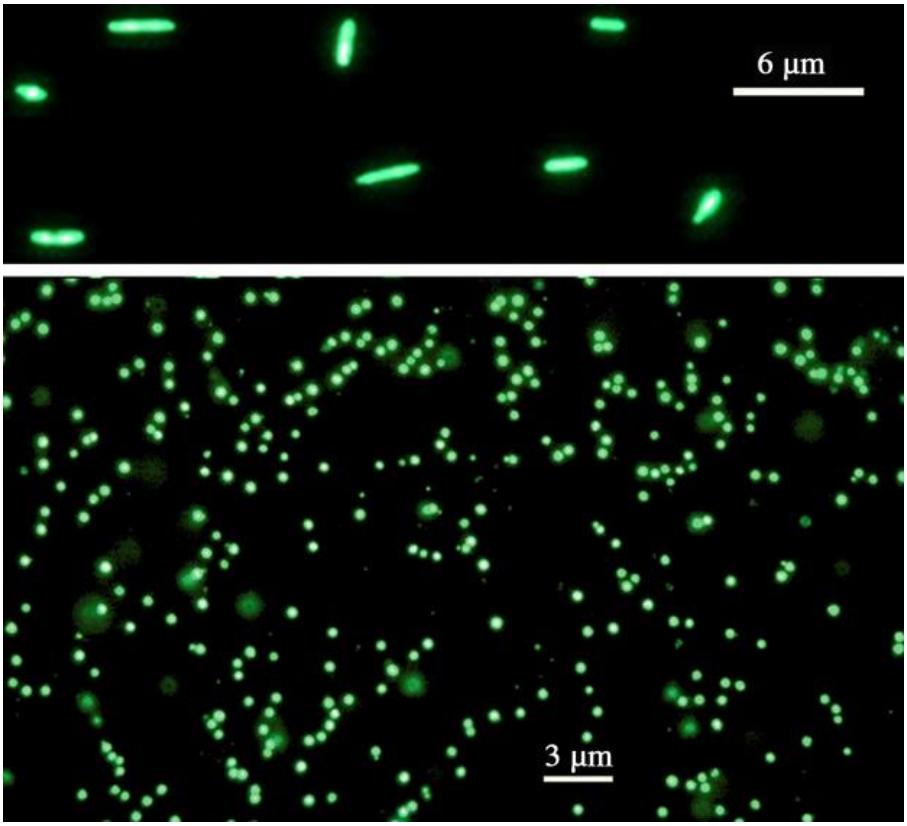


Figure 1.8 | **Haloarchaeal sphere formation after addition of 4M LiCl.** Rods (upper panel) and spheres (lower panel) of *H. salinarum* NRC-1, following staining with LIVE/DEAD kit. Spheres were produced by exposure of rods to Tris-buffered 4M LiCl. Figure taken from (Fendrihan *et al.*, 2012).

Several other haloarchaeal species have been examined by the above authors, in addition to *H. salinarum*. For example, *H. mediterrranei* also forms spheres when exposed to LiCl. This suggests that sphere formation is a crucial survival mechanism employed by halophilic Archaea to allow habitation of saline environments.

Interestingly, the ATP content within the spheres was measured, and was found to be 50-fold lower than that of rod shaped cells. Data from dormant bacteria and spores indicate similarly reduced ATP levels, thus demonstrating

that sphere formation in Haloarchaea is potentially analogous to bacterial spores; a protective mechanism allowing transition to a dormant state whilst the environmental conditions are harsh (Setlow and Kornberg, 1970). It would be of interest to determine alternative sphere promoting stimuli in haloarchaea as well as the genes responsible, as the formation of spore-like spheres represent a major mechanism of survival in extreme conditions. Resistance to ionizing radiation would also be interesting to investigate in such Haloarchaea, especially to see if sphere formation is affected.

1.3.6 Membrane chemistry

Archaeal membranes are composed of phospho-, glyco- and phosphoglycolipids, which are significantly different in structure to their bacterial counterparts. Archaeal membranes consist of phosphodiester-linked polar head groups on the *sn-1* position of the glycerol backbone – a glycerol-1-phosphate structure. Archaeal G-1-P is therefore an enantiomer of the glycerol-3-phosphate of bacterial phospholipids. Isoprenoid hydrocarbon side chains are linked to the *sn*-glycerol-1-phosphate backbone via an ether bond, whereas bacteria and eukarya use fatty acid side chains, linked to *sn*-glycerol-3-phosphate via an ester bond (De Rosa and Gambacorta, 1988, Oger and Cario, 2013). Furthermore, the hydrocarbon chain structure in archaeal membranes are often made up of 20 carbons (phytanyl chain) or a head to head dimer of two chains, bringing the total to 40 carbons (Figure 1.9).

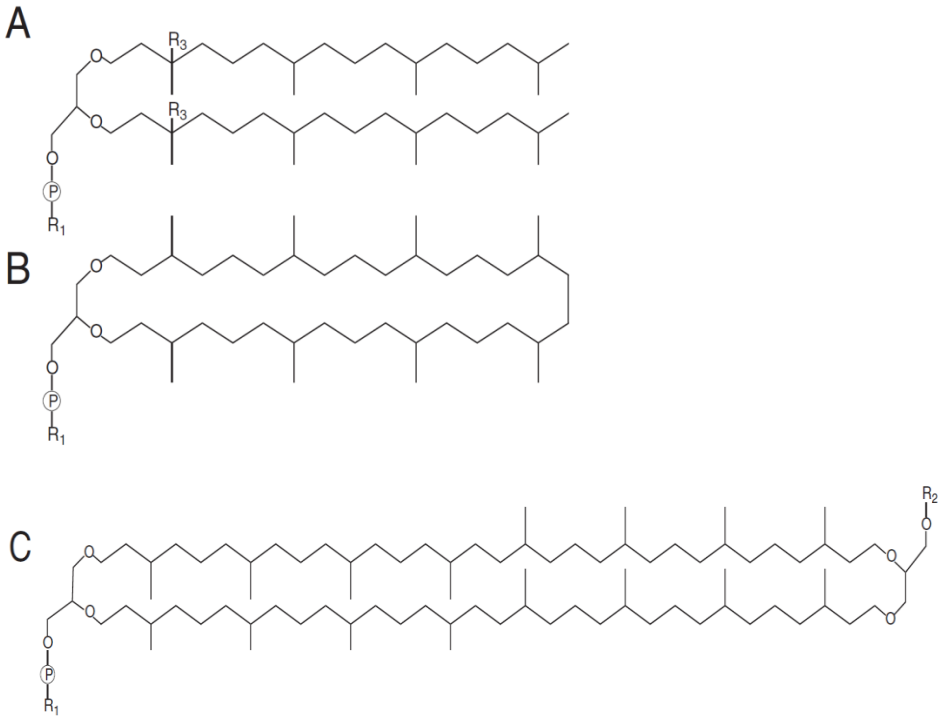


Figure 1.9 | **Structures of archaeal lipids**. (A) Diphytanyglycerol. (B) cyclic archaeal diether lipid. (C) Head to tail arrangement of two phytanyl chains. Figure adapted from (Oger and Cario, 2013)

Until recently, the stereo specificity of archaeal lipids represents the "lipid divide", where the G1P backbone is Archaea-exclusive. A recent discovery of a bacterial G1P dehydrogenase (homologous to that in Archaea) in *Bacillus subtilis* challenged this idea (Guldan, Sterner and Babinger, 2008).

Head to head condensation of two diether lipid molecules is functionally important and leads to the formation of a glycerol-dialkyglycerol tetraether lipid known as caldarcheol. The enzymatic mechanism underlying this process is not resolved, however it is noted that ethers possess greater chemical stability than esters in harsh environmental conditions. Saturated phytanyl lipid cores are much more rigid than fatty acid lipids and the branched chains

pack tighter than the straight-chain lipids outside of the Archaea (Hanford and Peeples, 2002). Tetraether lipids are considered thermophilic and link the leaflets of the bilayer via covalent bonds (Figure 1.10). This covalent linkage makes the membrane more rigid, allowing tolerance of extreme conditions and temperature fluctuations (Koga, 2012).

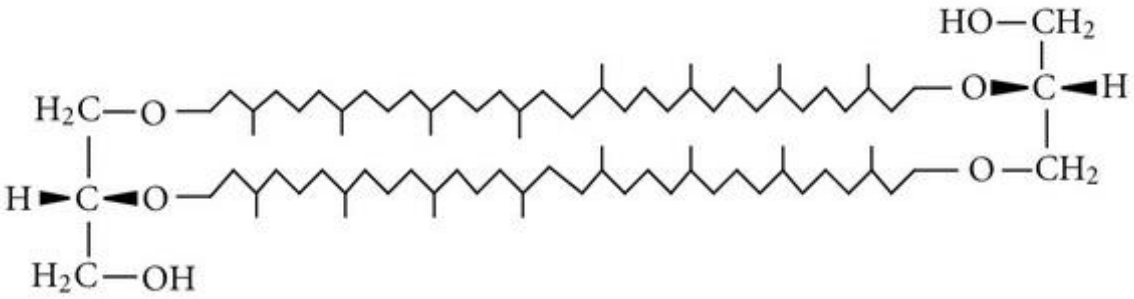


Figure 1.10 | **Structure of archaeal tetraether lipids, the two membrane leaflets are linked by covalent bonds providing structural rigidity to the cell membrane.** Figure adapted from (Koga, 2012).

Moreover, recent work on lipid composition of haloarchaeal membranes in *H. volcanii* and four other haloarchaeal species uncovered extraordinary high levels of menaquinone within the membrane. This ubiquity suggests menaquinones may function beyond their normal role of electron and proton carriers, acting also as ion permeability barriers and against oxidative stress (Kellermann *et al.*, 2016). In addition it is put forward that the difference in membrane chemistry of Archaea was key for integrating the respiratory machinery from bacteria into the archaeal domain.Furthermore, the archaeal phospholipids potentially hold two negative charges, previous evidence indicates that phosphatidylglycerol and methylated-phosphatidylglycerolphosphate represent the bulk of glycerolipids in Haloarchaea (Kellermann *et al.*, 2016). The dominance of such anionic lipids

among Haloarchaea has been suggested as an adaptive response to counter membrane stress in highly saline environments (Russell, 1989).

Saturation of lipids within the cell membrane is also an important adaption to extreme environments. A recent paper also showed that unsaturation levels of glycerolipids increased with Na^+ concentration (Dawson, Freeman and Macalady, 2012). The paper suggests that increased salinity leads to a decrease in hydration of the hydrophilic region of the membrane, resulting in increased strength and rigidity of the membrane via increased hydrogen bonding between lipid head groups. Kellermann obtained data in support of this idea, however they also suspect that under stressful conditions such as low or high levels of salt, the bulk glycerolipid composition displayed less unsaturation (i.e. decreased membrane motion and permeability). Data thus suggests that *H. volcanii* may modulate its membrane motion by changing the levels of unsaturation in lipid chains (Kellermann *et al.*, 2016).

Lastly, experiments inducing oxygen and UV stress prompted a consistent decrease in unsaturation levels of most glycerolipids. This decrease supports the idea that *H. volcanii* can adjust its membrane composition to protect itself from such damage (Kellermann *et al.*, 2016).

Overall, Haloarchaea increase membrane strength as a barrier to osmotic stress by fusing lipid tails together across membrane leaflets as well as increasing the number of anionic lipids in the membrane and altering the saturation levels of lipids to protect against osmotic stress and UV radiation/oxygen radical damage.

1.4 *Haloferax volcanii* as a model organism

1.4.1 Genome organisation

Haloferax volcanii was first isolated from the Dead Sea in 1975 (Mullakhanbhai and Larsen, 1975). Coccoid in morphology, it appears red in colour owing to the presence of carotenoid pigments. *H. volcanii* is of great use in the lab owing to its ease of culture and the availability of a wide scope of confirmed genetic techniques, a rare feature for an extremophile.

The complete genome of the *H. volcanii* strain DS2 consists of a main circular chromosome (2.85Mb), smaller chromosomes, pHV1 (85kb), pHV3 (438kb), pHV4 (636kb) and a small plasmid pHV2 (6.4kb). *H. volcanii* grows in optimal salt concentrations of 1.7 – 2.5M NaCl. The average genomic GC content is 65% and coding density is 86%. The genome encodes 6 rRNAs in two rRNA operons as well as 51 tRNAs. (Hartman *et al.*, 2010). Furthermore, the six origins of replication in *H. volcanii* (strain DS2) have been characterised fairly recently (Norais *et al.*, 2007), there are six that are ORC-dependant and a further origin on pHV2. Experimental work by Hartman et al indicated that the main chromosome and three smaller chromosomes all contain functional ARS sequences as well as two origins of replication on its main chromosome, denoted *oriC1* and *oriC2*. Ploidy is a key mechanism for survival in extreme environments and will be discussed in depth in chapter 1.3.1. However it should be noted that *H. volcanii* possesses a high level of ploidy, on average 18-20 copies per cell in exponential phase (Breuert *et al.*, 2006).

Deep sequencing techniques have been applied to *H. volcanii* (strains DS2, H26) to map the position of origins of replication. The H26 lab strain has a major chromosome with four origins of replication owing to the integration of pHV4 and the fact that pHV1 and pHV3 have their own origins. This is in comparison to the six origins present in the wild-type strain, DS2 (Hawkins *et al.*, 2013) (Figure 1.11).

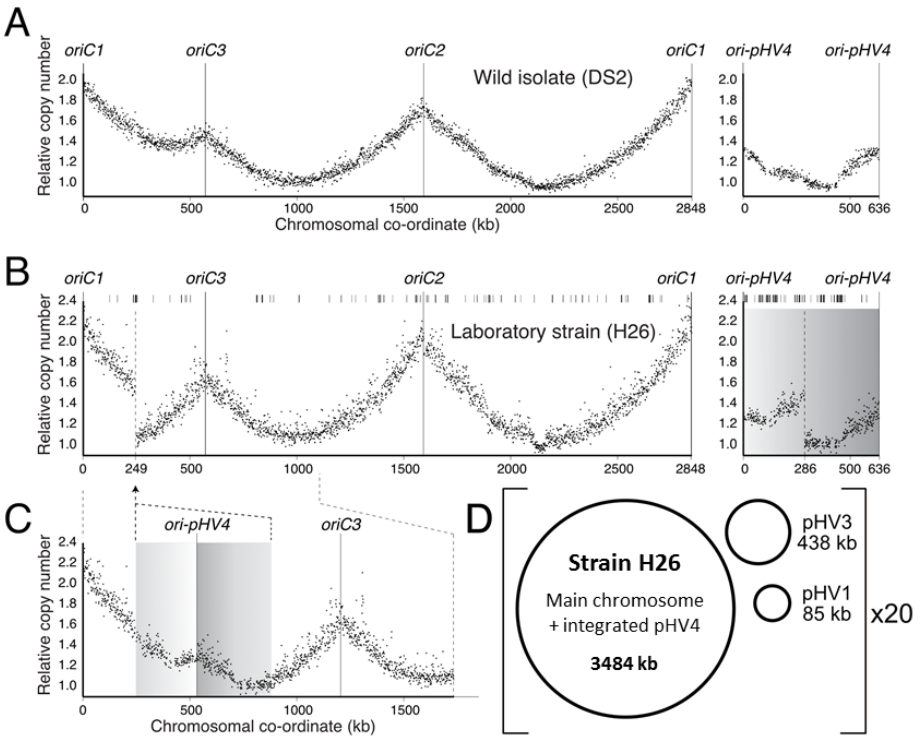


Figure 1.11 | **Replication profiles for *H. volcanii* wild isolate and laboratory strain.** (A) Relative copy number plotted against chromosomal coordinate for the main chromosome and pHV4 of wild-type strain DS2. Vertical lines represent origins of replication and circular chromosomes are displayed linearized at position 0. (B) Relative copy number plotted against chromosomal coordinate for the main chromosome and pHV4 of laboratory strain H26 mapped to the DS2 genome. pHV4 shading reflects chromosomal coordinate. (C) Reconstructed assembly of the main chromosome with pHV4 integrated at approximately 249kb. Grey shading from B indicates orientation of pHV4 integration. (D) *H. volcanii* copy number is approximately 18-20. Figure adapted from (Hawkins *et al.*, 2013).

1.4.2 *Haloferax volcanii* genetic techniques

Haloferax volcanii is easily cultured in the lab and a range of genetic techniques exist in this model organism. *H. volcanii* has genetic tools including a transformation system, various reporter genes and auxotrophic markers (Table 1.3). Drug selection is also well established (Table 1.4).

Table 1.3 | Common auxotrophic markers used in *H. volcanii*

Gene	Selectable Phenotype	Reference
pyrE2	Uracil biosynthesis	(Bitan-Banin, Ortenberg and Mevarech, 2003)
trpA	Tryptophan biosynthesis	(Allers, H. Ngo, *et al.*, 2004)
leuB	Leucine biosynthesis	
hdrB	Thymidine biosynthesis	(Ortenberg, Rozenblatt-Rosen and Mevarech, 2000)

Table 1.4 | Common selectable drug resistance markers in *H. volcanii*

Gene	Selectable phenotype	Reference
ShBle	Bleomycin resistance	(Nuttall *et al.*, 2000)
hmgA	Mevinolin resistance	(Wendoloski, Ferrer and Dyall-Smith, 2001)
gyrB	Novobiocin resistance	(Holmes and Dyall-Smith, 1990)

Reporter genes such as β-galactosidase can be utilized in *H. volcanii*, however it lacks a functional copy of the β-galactosidase gene (Delmas *et al.*, 2009). This can be resolved by introducing an active β-galactosidase gene from *Haloferax alicantei* to allow for canonical blue-white screening (Holmes and Dyall-Smith, 2000).

Transformation of *dam*$^-$ plasmid DNA can be accomplished with relative ease using EDTA to remove the S-layer and subsequent uptake of DNA with the aid of polyethylene glycol (PEG 600) (Cline *et al.*, 1989). *H. volcanii* strains deleted for the *mrr* restriction system can be directly transformed using *dam*$^+$ DNA owing to their inability to recognise and degrade the dam-methylated DNA, thus negating the need for passage through a *dam*$^-$ *E. coli* strain prior to transformation (Allers *et al.*, 2010a).

Transformation can be then combined with subsequent gene knock-out / knock-in methods, auxotrophic markers (described in Table 1) are key to this process. Of note, the *pyrE2* gene can be counter-selected using 5-fluoroorotic acid (5-FOA) as *ura*$^-$ cells are resistant to this compound due to their inability to convert 5-FOA into the toxic analogue, 5-flurouracil (Bitan-Banin, Ortenberg and Mevarech, 2003, Wang *et al.*, 2004). A plasmid containing a deletion construct for the desired gene (and generally a *trpA* marker in place of the gene) as well as the *pyrE2* marker are transformed into a host strain deleted for *pyrE2*. Should a "pop-in" event occur, successful transformants will grow on media lacking uracil, owing to the gain of the *pyrE2* gene. The "pop-out" event is then allowed to occur by the withdrawal of uracil selection. These transformants lose the *pyrE2* marker and are thus selected on media containing 5-FOA. The *trpA* marker allows for direct selection of wild-type cells (Figure 1.12).

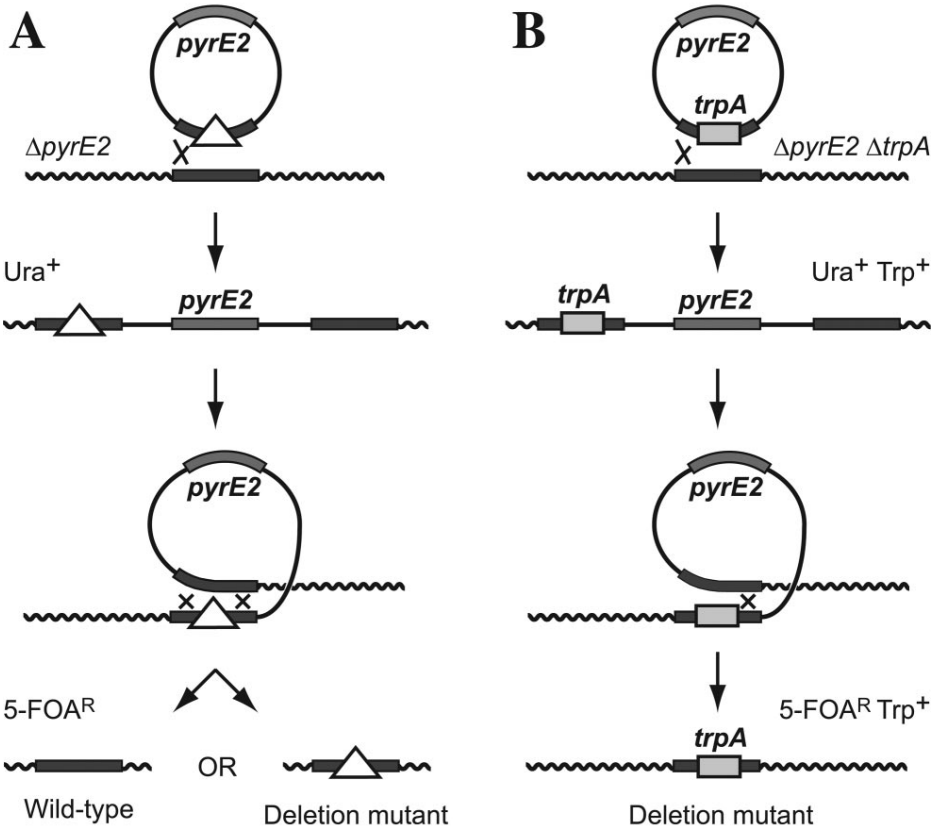


Figure 1.12 | **Overview of "pop-in" / "pop-out" method of gene deletion based on the *pyrE2* gene.** (A) A plasmid carrying the *pyrE2* marker and flanking sequences of the gene to be deleted is used to transform Δ*pyrE2 H. volcanii* strain to uracil prototrophy. The "pop-in" event has occurred to the left of the deletion. Loss of the plasmid by an intrachromosomal cross-over event can occur on the left of the deletion, restoring the gene to wild-type, or on the right of the deletion, resulting in the desired mutant. In both cases the cell is rendered auxotrophic for uracil and is therefore resistant to 5-FOA. (B) The gene is replaced with the *trpA* marker, and the plasmid is used to transform Δ*pyrE2 ΔtrpA H. volcanii* to prototrophy for uracil and tryptophan. Loss of the plasmid by crossing-over on the right of the deletion,

resulting in a *trpA*-marked mutant can be selected in one step. Figure taken from (Allers, H. Ngo, *et al.*, 2004).

Genes can be overexpressed in *H. volcanii* by introducing a plasmid expressing the desired gene downstream of the tryptophan-inducible promoter *p.tnaA*. This promoter is tightly regulated in *H. volcanii* (Allers, 2010).

Lastly, a modified GFP reporter gene can be used in *H. volcanii* and is adapted to function in hypersalinity. Three amino acid substitutions have occurred to enable this (Crameri *et al.*, 1996, Reuter and Maupin-Furlow, 2004).

1.5 DNA Damage

1.5.1 Introduction to DNA Damage

Damage to DNA can arise both endogenously (i.e. spontaneous) or exogenously (i.e. from environmental factors). Damage to DNA occurs continuously from these sources, and thus avoiding any damage to your genetic material is impossible. At the same time as needing to preserve genome fidelity and sequence information for the perpetuation of life, mutagenesis plays a critical role in driving evolution, therefore low mutational load is often tolerated whereas high mutational load is not compatible with life. Therefore, efficient repair of DNA lesions is critical to all life. In fact, the current reported figure of mutational load in *H. volcanii* is approximately $1.2x10^{-3}$ events per genome per generation (Kucukyildirim, Ozdemirel and Lynch,2023). Chemical reactions occur all the time, and certain lesions occur more often in extreme conditions, therefore Archaea are thought to possess highly efficient DNA repair systems to cope with harsh conditions, such as ionizing radiation or high salinity (Figure 1.13).

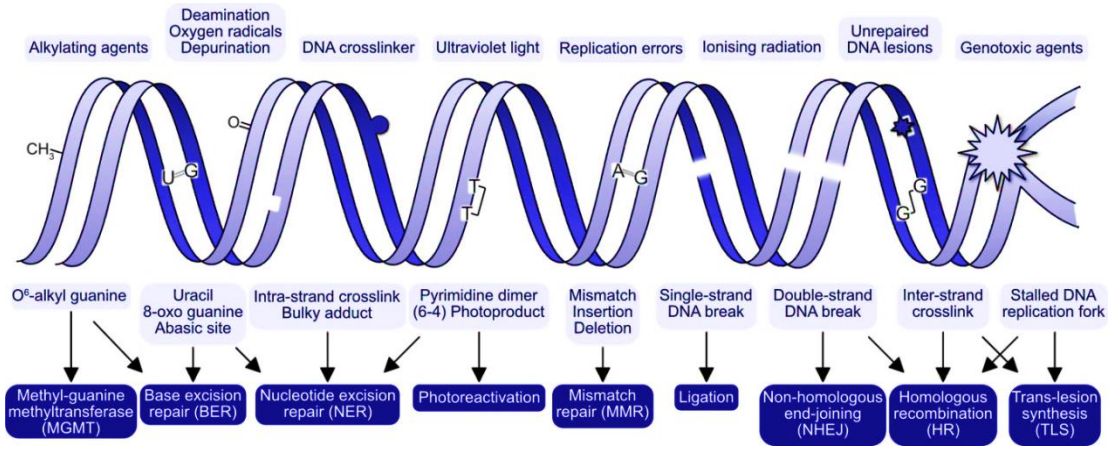


Figure 1.13 | **Summary of DNA damage causes, consequences and repair pathways.**

Figure taken from (White and Allers, 2018).

1.5.2 Endogenous DNA Damage

Normal cellular metabolism can generate lesions in DNA via generation of reactive chemical species within the cell.

Oxidative Damage

Oxidative free radicals are common within cells and generated as a by-product of canonical cellular metabolism. Such radicals are highly reactive and can interact with DNA resulting in significant damage implicated in mutagenesis, carcinogenesis and aging (Dizdaroglu *et al.*, 2002). Examples of free radicals include reactive oxygen species and reactive nitrogen species (ROS and RNS). Such radicals have the potential to cause oxidative base lesions such as 8-oxo-2'-deoxyguanosine, which as a result of oxidative modification, subsequently forms hydrogen bonds with adenine, leading to a G●C to T●A base pair transition (Kasai and Nishimura, 1984) (Figure 1.14).

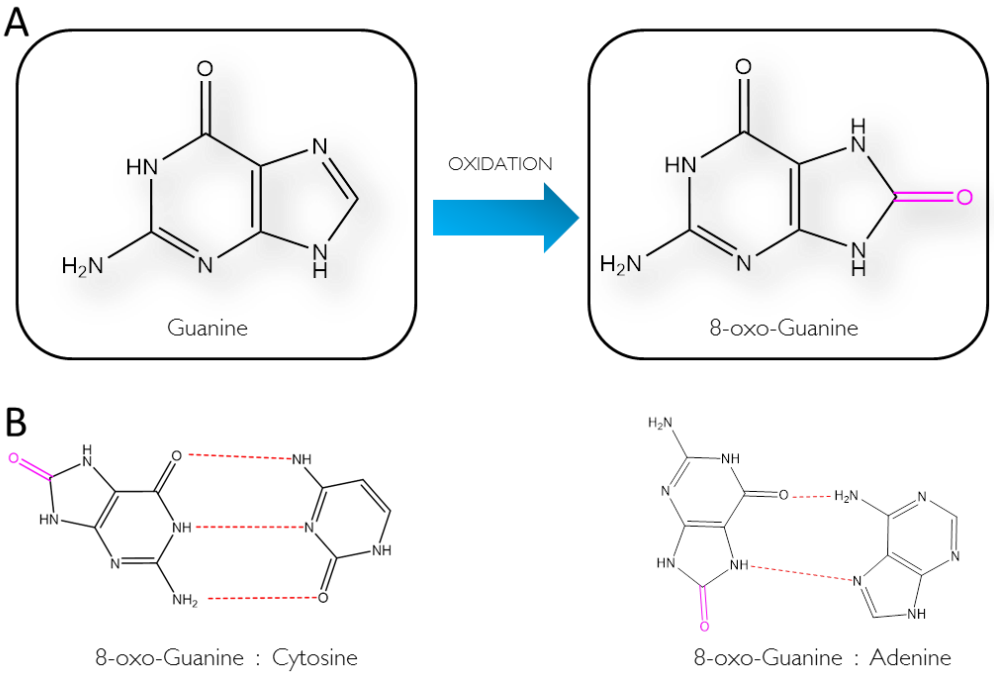


Figure 1.14 | **Overview of 8-oxo-Guanine DNA lesion and its effect on base pairing.** (A) Oxidative stress within the cell, as a by-product of metabolism and free radicals, leads to oxidation of guanine at the 8th position – shown in pink. (B) 8-oxo-guanine can base pair with both cytosine and adenine as a result of oxidation. DNA lesion shown in pink, hydrogen bonds in red. Figure adapted from (Nakabeppu, 2014).

Oxidative damage, as a result of free radical species, can also lead to nicks in the DNA backbone, replication of which could lead to a double strand break, one of the most toxic lesions known.

Another prominent example of a lesion caused by oxidative radicals is that of a Thymine glycol, which pairs successfully, but causes replication issues by acting as a block to replicative polymerases (Figure 1.15).

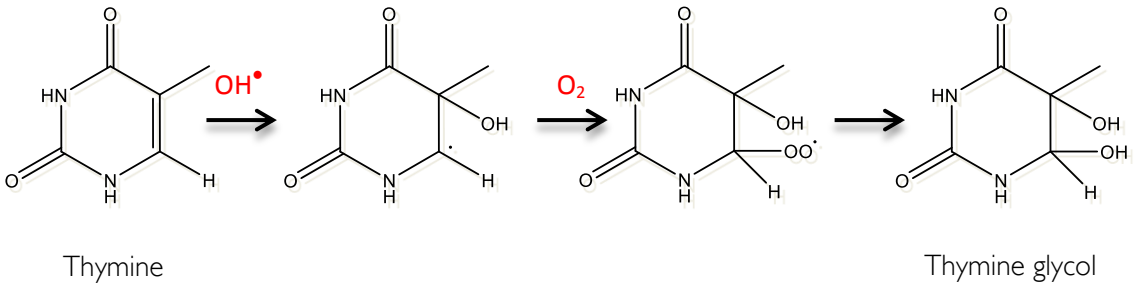


Figure 1.15 | **Formation of thymine glycol via oxidative radicals and oxygen**. Figure adapted from (Simic, 1994).

As seen above, a hydroxyl radical attacks the double bond of the thymine base at C5 or C6. The 6-hydroxythymine radical intermediate can react with O_2 to form thymine glycol (Demple and Linn, 1982). Thymine glycol is

structurally sound, however affects the ability of replicative polymerases to carry out DNA replication.

Further damage can result from a similar mechanism to that shown above, except the target is either adenine or guanine bases rather than thymine. In this case the products formed are 4,6-Diamino-5-formamidopyrimidine and 2,6-Diamino-4-hydroxy-5-formanidopyrimidine respectively, also known as FaPy products (Figure 1.16). Such products contain a broken ring structure and thus are unable to be read by the replicative DNA polymerases at all (Breimer, 1990, Greenberg, 2012).

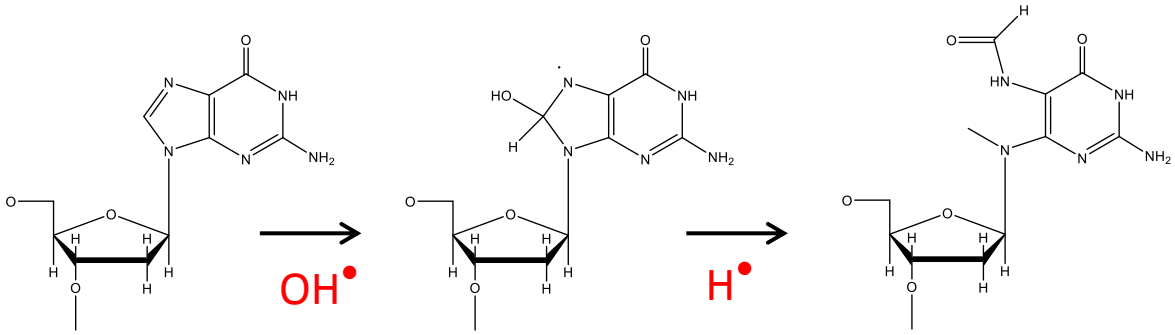


Figure 1.16 | **Formation of FaPy●dGuanine**. FaPy●dG is formed from reduction, compared to oxidation in the case of 8-oxo-Guanine. Figure adapted from (Greenberg, 2012).

Deamination

Cytosine, guanine, adenine and 5-methylcytosine all contain exocyclic amino groups which can be spontaneously lost (via deamination) in pH and temperature-dependant reactions of DNA within the cell. This results in the conversion of bases into uracil, hypoxanthine, xanthine and thymine (Kow, 2002) (Figure 1.17).

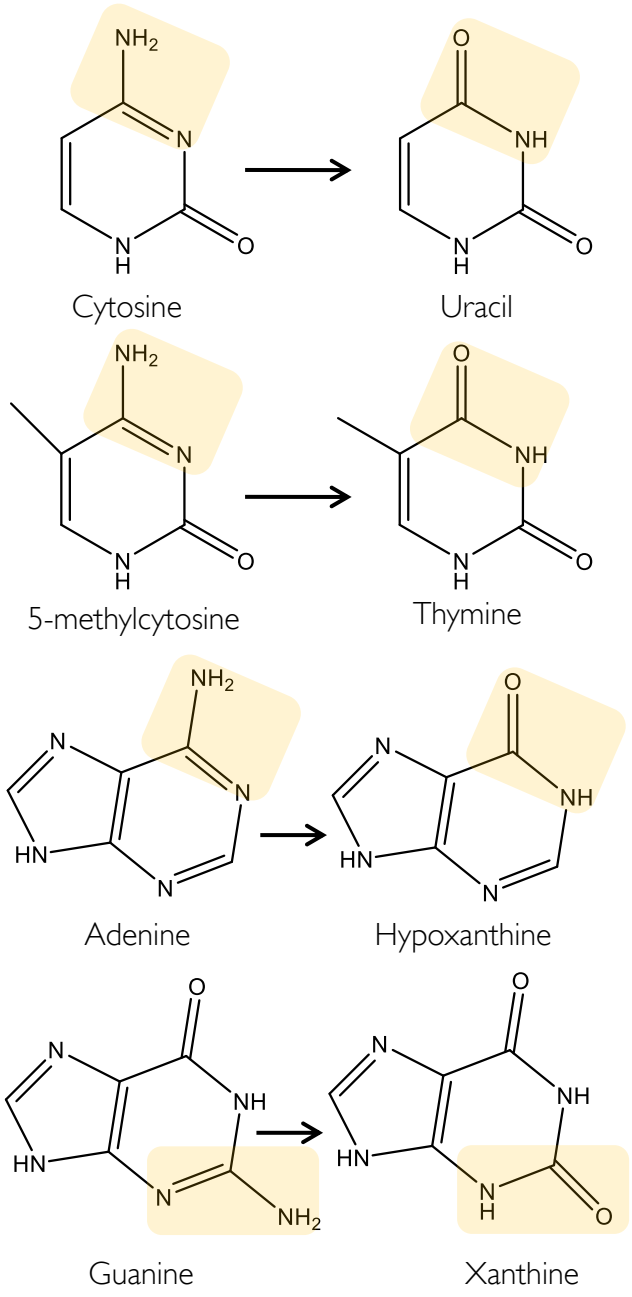
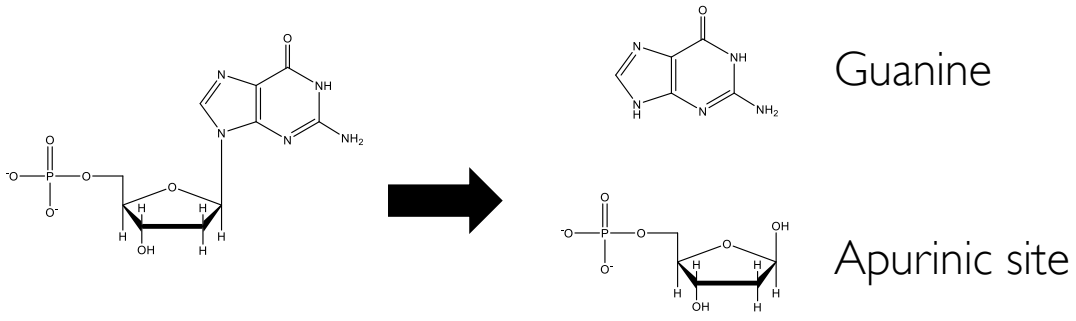


Figure 1.17 | **Products formed from the deamination of bases in DNA**.

5-methylcytosinecan be deaminated to thymine by the addition of water, resulting in a G●C to T●A transition (Lutsenko and Bhagwat, 1999). This is a problem given that thymine is present in DNA as the mistake can not easily be spotted and repaired. The process of cytosine deamination results in uracil within the DNA polynucleotide chain – a lesion easier to spot owing to the fact that uracil should be in RNA only, such a lesion can be easily repaired by a uracil-specific DNA glycosylase in the process of Base Excision Repair, which will be discussed later on (Schormann, Ricciardi and Chattopadhyay, 2014). In humans, cytosine deamination can be induced by an enzyme family of cytidine deaminases, such as Activation-induced Cytidine Deaminase(AID). AID is used to convert cytidine to uridine during somatic hypermutation of immunoglobulin maturation, the details of which will not be discussed here, although many reviews exist on this topic such as (Maul and Gearhart, 2010).

Depurination and Depyrimidination

The release of purine or pyrimidine bases from nucleic acids via N-glycosidic bond hydrolysis is known as depurination or depyrimidination respectively. As a result, an abasic site is left in the DNA (AP site) (Figure 1.18).



Guanine

Figure 1.18 | **Products formed from breakage of the glycosidic bond of a guanine nucleotide within DNA, leaving an abasic site.**

It has been estimated that approximately 2000-10,000 DNA purine bases are released in each human cell per day due to hydrolytic depurination (Lindahl and Nyberg, 1972).

Pyrimidine nucleosides are more stable than their purine counterparts with respect to the N-glycosyl linkage. The mechanism of depyrimidation is the same, however cytosine and thymine are lost at much lower rates (Lindahl and Karlström, 1973). Repair of such lesions is via Base Excision Repair.

Tautomerization of DNA

DNA base pairs are almost entirely discussed in terms of the keto-amine tautomers, which are thought to be the dominant form. However each base may be converted spontaneously to its minor form, the enol-imine tautomer, via a double proton transfer. Despite their potential importance in allowing duplex flexibility and accurate DNA replication, tautomerization can lead to inaccurate base pairing of DNA bases and thus mutagenic events (Abou-Zied, Jimenez and Romesberg, 2001, Singh, Fedeles and Essigmann, 2015).

1.5.3 Exogenous DNA Damage

Exogenous, or induced, DNA damage is caused by environmental agents such as chemicals and ionizing radiation. Exogenous damage can occur either directly, such as via UV light, or indirectly via the radiolysis of water to produce radical species. Key examples of induced DNA damage are outlined below.

Ionizing radiation

Ionizing radiation is a form of high-energy radiation that is able to remove electrons from atoms and molecules, generating ions which can break covalent bonds. One of the most toxic lesions caused by ionizing radiation is a double strand break (DSB) where both polynucleotide strands are broken, if unrepaired, such lesions can lead to extreme consequences within the cell, such as translocations and various cancers (Borrego-Soto, Ortiz-López and Rojas-Martínez, 2015).

Notable, however is the fact that single strand breaks predominate over double strand breaks. A major lesion caused by ionizing radiation mediated single chain breaks is a 3' phosphoglycolate (Henner *et al.*, 1983) (Figure 1.19). Such a break can lead to the inability of DNA replication to occur due to inhibition of DNA polymerase action.

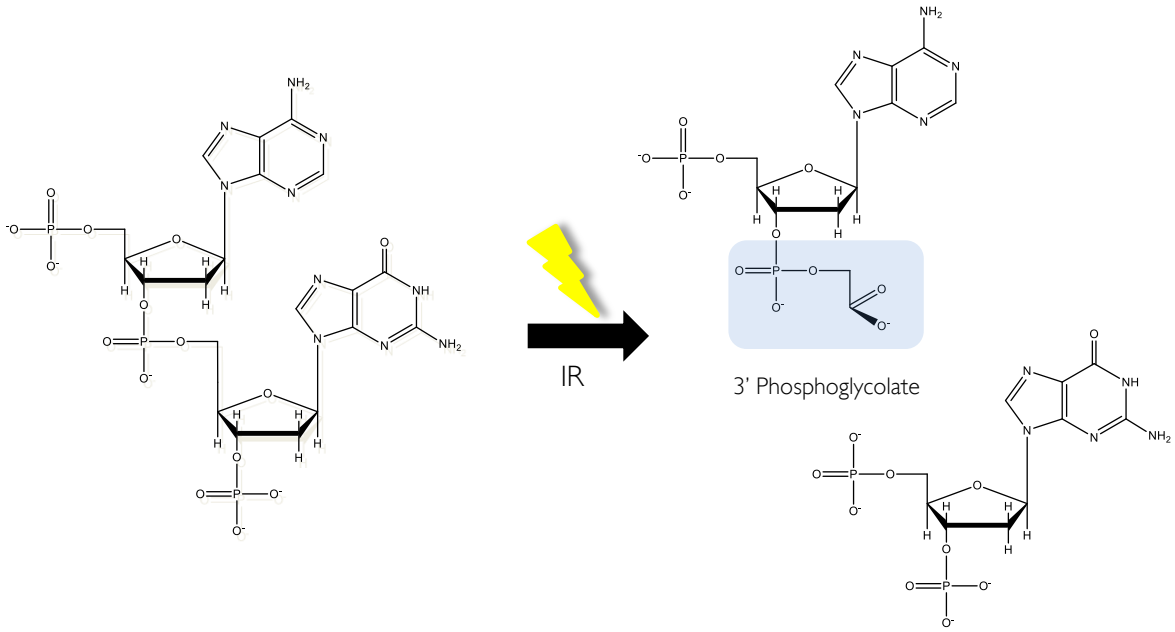


Figure 1.19 | **Schematic of ionizing radiation induced single strand break, forming a 3' phosphoglycolate lesion.** Own figure adapted from (Winters *et al.*, 1994).

Phosphoglycolate structures need to be dealt with quickly before DNA replication occurs, leading to a double strand break. Usually a break in the phosphodiester backbone of DNA can be fixed using DNA Ligase-mediated repair, however this requires an in tact phosphate and hydroxyl end – the latter of which is not present owing to the non-specific targeting of ionizing radiation-mediated breaks and breakage of the deoxyribose ring.

A problem with this is that the above product is not a substrate for the action of DNA ligase, however DNA ligase begins the process of ligation before checking the integrity of the 3'OH group and thus aborts the process half way through, after leaving an adenyl group on the 5' phosphate thereby making the issue worse due to making the 5' phosphate end a non-ligatable end. However, the gene product of the *APTX* gene, aprataxin, is a member of the histidine triad family of nucleotide hydrolases/transferases and can deal with the, now dirty, 5' end. This is carried out by removing the adenyl group from the 5' end (Figure 1.20) thus allowing more time for base excision repair machinery to deal with the 3' phosphoglycolate and insert a new, undamaged, deoxynucleoside triphosphate (dNTP) using the undamaged strand as a template. More details of base excision repair will be discussed later (Ahel *et al.*, 2006).

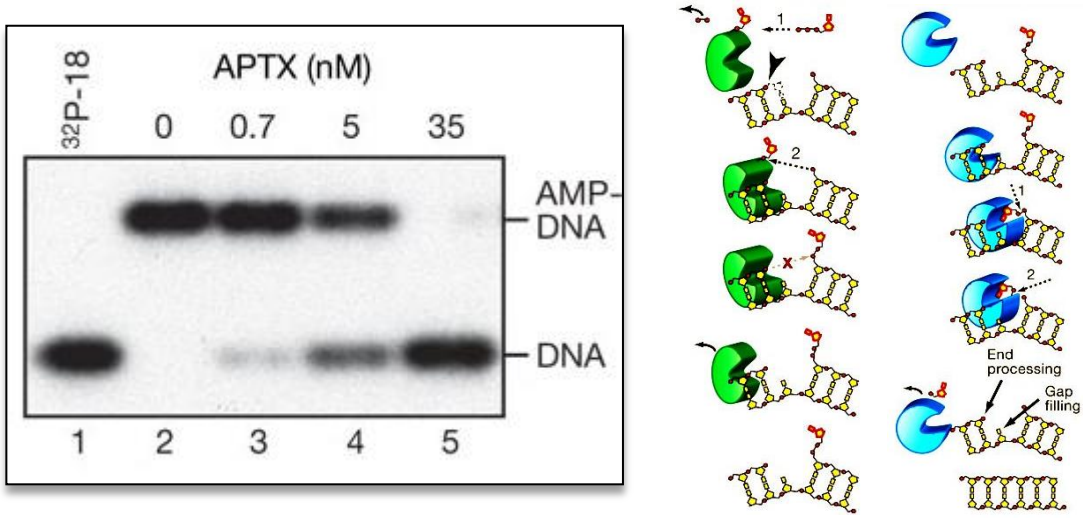


Figure 1.20 | **Aprataxin removes AMP from nicked DNA-adenylate**. (A) Reactions contained DNA-adenylate and the indicated amounts of $^{GST\text{-}His}$APTX. Figure taken from (Ahel *et al.*, 2006). (B) Abortive ligation schematic with DNA Ligase shown in green (left) compared to successful ligation assisted by the action of aprataxin acting to remove the 5' adenylate group (right). Aprataxin shown in blue and acts to remove the 5' adenyl group to generate a ligatable 5' phosphate. Figure taken from (Ahel *et al.*, 2006).

As indicated previously, ionizing radiation can also lead to a single-strand DNA break via an indirect method, namely the radiolysis of water. Major products formed from the radiolysis of water include oxygen radical species, such as ●OH and ●H, capable of producing lesions in DNA as seen in the previous chapter, a major lesion is 8-oxo-guanine. However, such free radicals can also initiate a single strand break of the polynucleotide strand, leading to 3'-phosphoglycolyate products via nucleophilic attack (Ward, 1988, Le Caer, 2011).

Ultraviolet light

Ultraviolet (UV) radiation is a major source of DNA damage, leading to a variety of DNA lesions such as cyclobutane pyrimidine dimers (CPDs), single and double-stranded DNA breaks and 6-4 photoproducts along with their photoisomerisation product, the Dewar valence isomer. Excitation from UV radiation leads to a covalent crosslink between carbons 5 and 6 on neighbouring pyrimidine bases (Figure 1.21) (Douki *et al.*, 2000, Douki and Sage, 2016).

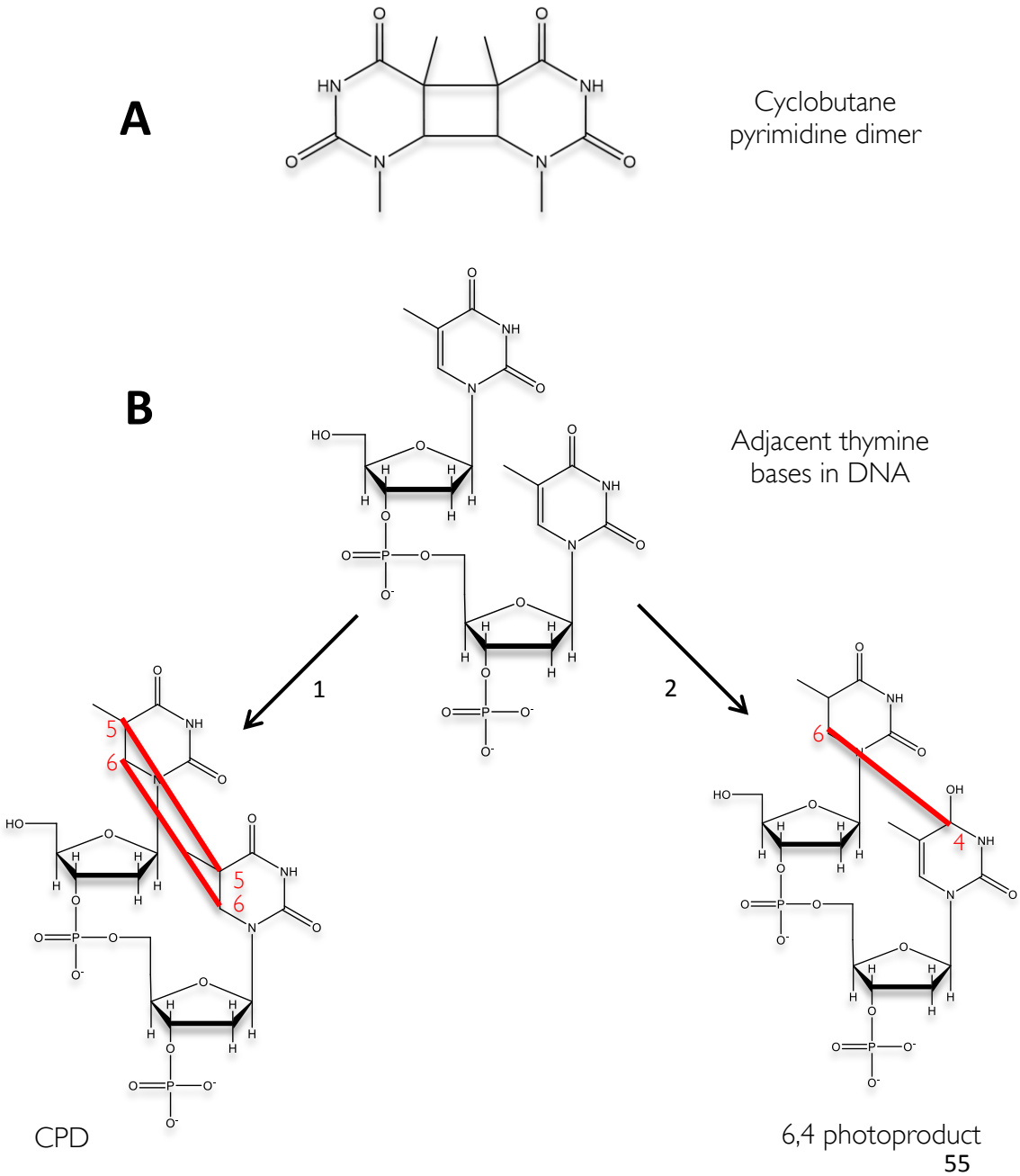


Figure 1.21 | **UV-irradiation products. Covalent bonds and numbered carbons shown in red.** (A) UV light excites electrons in the C=C double bonds, allowing them to react with the neighbouring pyrimidine, in this case two thymines. (B) Formation of CPDs and 6,4 photoproducts. (B1) Schematic of CPD formation between two adjacent thymine bases in DNA. (B2) Formation of a 6,4 photoproduct as a result of covalent cross-link bridge between C4 and C6 of adjacent thymine bases. Figure adapted from (Yokoyama and Mizutani, 2014).

It should be noted, however that purine photoproducts are also formed, such as A-T adducts, which have been shown to be mutagenic (Bose and Davies, 1984). Furthermore, a study was carried out to investigate the bias towards formation of T●T dimers over C●C and C●T dimers. Essentially, the formation of TT dimers proceeds via the smallest energy barrier (Durbeej and Eriksson, 2003).

Reactive oxygen species can also arise from UV radiation. Such radicals can act on DNA and lead to the formation of single-stranded breaks (Cadet, Sage and Douki, 2005). If two single strand breaks occur in close proximity to each other, or if a replication fork stumbles across a break in the leading strand, a double-strand break can occur, which is extremely toxic to the cell, although base pairing is not affected, the ability of the DNA polymerase to replicate the DNA past this 'bulky' lesion is compromised.

Generally, the products from UV radiation can be spotted quickly due to the distortion of the DNA double helix, owing to the formation of CPDs and/or 6,4 photoproducts. (Figure 1.22). Repair proceeds, in the absence of light, via nucleotide excision repair, which will be covered in the DNA repair section. The damaged bases can be repaired quickly via a more direct method using a photolyase enzyme. This is not the case in placental mammals, however as they lack photolyase enzymes (Mees *et al.*, 2004) (Eker *et al.,* 2009).

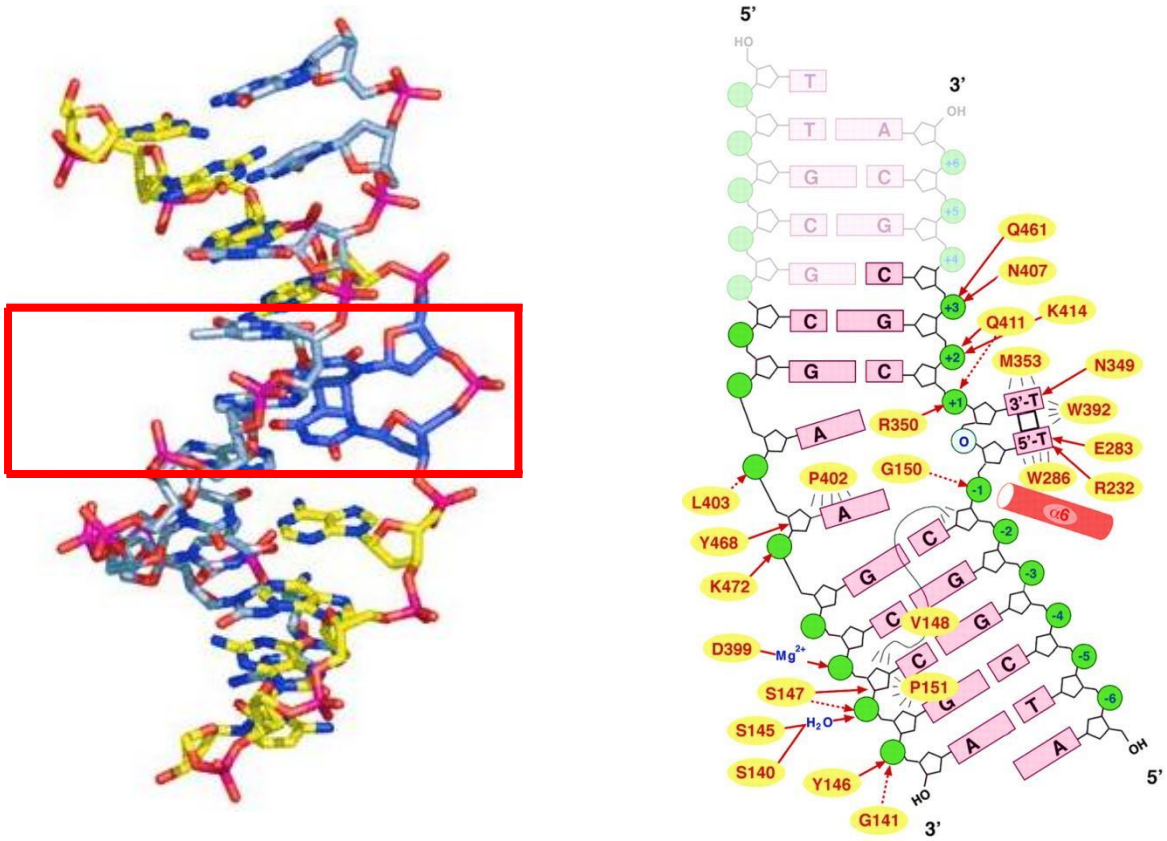


Figure 1.22 | **Backbone distortion caused by covalent linkage between two adjacent pyrimidine bases.** (A) Backbone distortion and CPD lesion are highlighted within the red box. (B) More detailed figure outlining CPD lesion and backbone distortion. Figure adapted from (Mees *et al.*, 2004).

However, although base pairing is not affected, the ability of the DNA polymerase to replicate the DNA past this 'bulky' lesion is compromised.

A recent paper used microorganisms living in hypersaline environments and identified new genes involved in UV radiation resistance via a metagenomics approach. Small-insert libraries were constructed with DNA isolated from microorganisms in hypersaline habitats and they were hosted in a UV-sensitive strain of *E. coli* (*recA* mutant) which were then exposed to UVB. Four clones were identified that conferred resistance to UV radiation in *E. coli*. Five genes were then identified, playing a part in DNA repair pathways. One gene product was a RecA-like protein, functionally complementing the host defect and thus restoring wild-type UV resistance. Two other genes encoded a TATA-box binding protein and an unknown protein – both conferring UV resistance to the host. Further resistance was observed to a compound considered a chemical analogue to UV radiation, 4-NQO. Further analysis will be needed on the gene products to confirm their function in DNA repair, however this functional metagenomics approach represents a novel approach to a canonical mutant screen (Lamprecht-Grandío *et al.*, 2020).

Chemical mutagens and DNA crosslinking agents

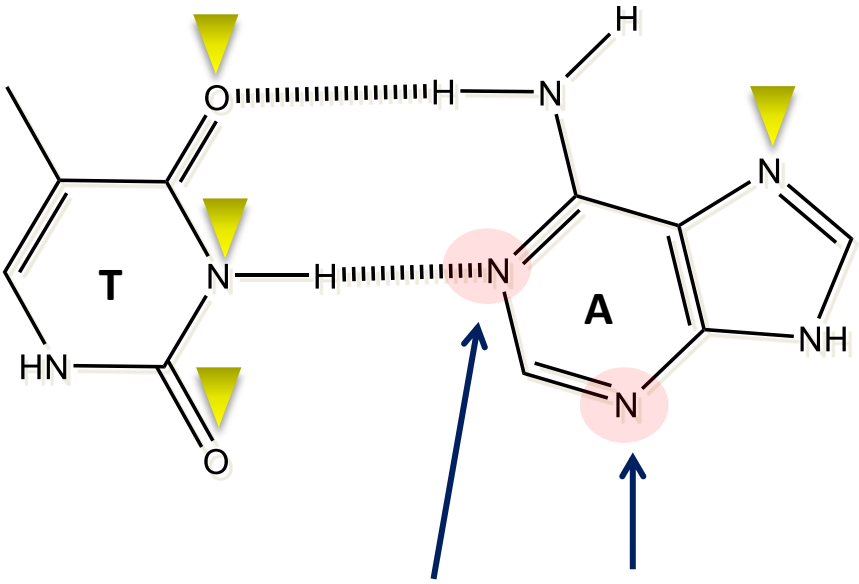
A variety of chemical agents have the ability to cause lesions to DNA. The main players are summarised below.

Alkylating agents

Alkylating agents add methyl or ethyl adducts to the DNA template via reacting with the ring nitrogens and extracyclic oxygen atoms. Key examples include Methyl methane sulphonate (MMS), which leads to either N^7-methylguanine or N^3-methyladenine. Ethyl methane sulphonate (EMS), which leads to O^6-ethylguanine, and Methyl-N-nitrosoguanidine (MNNG) which leads to either O^6-methylguanine or O^4-methylthymine. Most methylating agents induce formation of N^7-methylguanine owing to the high nucleophilic reactivity of the N^7 position within the guanine base, accounting for, on average, 70% of total alkylating lesions in DNA (Fu, Calvo and Samson, 2012).

Alone, N7-methylguanine does not have any mutagenic properties, although it is prone to spontaneous depurination and thus an AP site can be formed. However, N3-methyladenine is, in contrast, highly cytotoxic owing to its ability to block replicative DNA polymerases, thus inhibiting DNA synthesis (Engelward *et al.*, 1998, Fu, Calvo and Samson, 2012).

In ssDNA formed during replication or transcription, the N1 position of adenine and N3 position of cytosine are also targets for methylation by monofunctional methylating agents, resulting in replication-fork blocking and mispairing lesions, namely 1-methyladenine and 3-methylcytosine (Shrivastav, Li and Essigmann, 2010). Repair mechanisms will be featured in a future section, however a summary of alkylation sites on DNA bases can be seen in Figure 1.23.

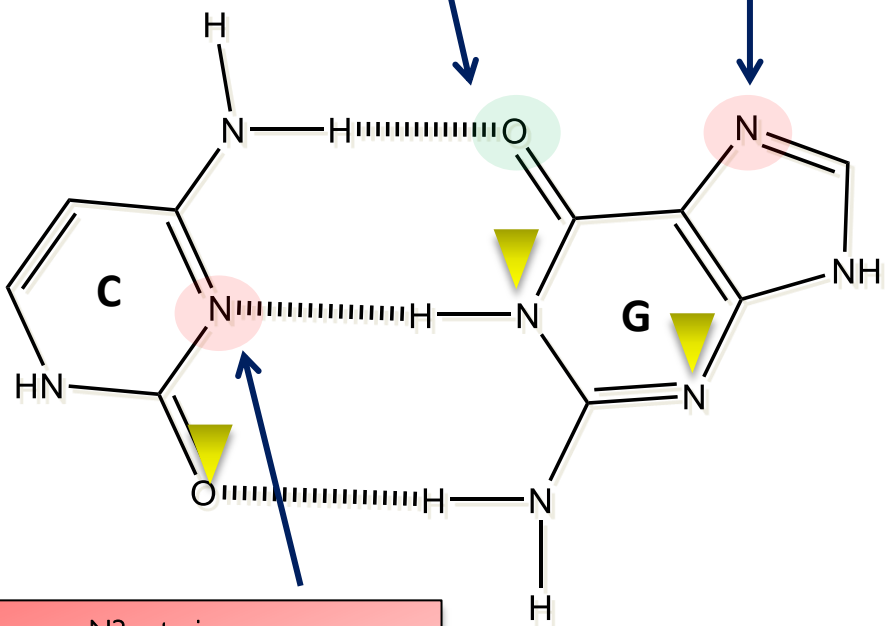


- N1-adenine
- Toxic (replication block)
- Mutagenic
- A → T transversion

- N3-adenine
- Toxic (replication block)
- Mutagenic
- A → T transversion

- O6-guanine
- Toxic (replication block)
- Mutagenic
- G → A transition

- N7-guanine
- Toxic (replication block) after depurination event
- Mutagenic



- N3-cytosine
- Toxic (replication block)
- Mutagenic
- C → T transition

Figure 1.23 | **Sites of alkylation damage on DNA bases.** Alkylating agents react with nitrogen and oxygen atoms on DNA bases to form covalent alkyl adducts. Major sites of alkylation damage are highlighted in red and green, with minor lesions denoted with yellow markers. Further detail is provided in the coloured boxes attached to each of the major lesions. Figure adapted from (Fu, Calvo and Samson, 2012).

In contrast to the monofunctional alkylating agents above, bifunctional alkylating agents can lead to interstrand crosslinks, one of the most toxic lesions owing to the strong covalent linkage created between polynucleotide strands, meaning replication can not occur at all. The nitrogen mustards are a prominent example of compounds within this group which react with N7-guanine to form monoadducts. These adducts can then react with another base to form a guanine-guanine interstrand crosslink (Figure 1.24) (Rink *et al.*, 1993).

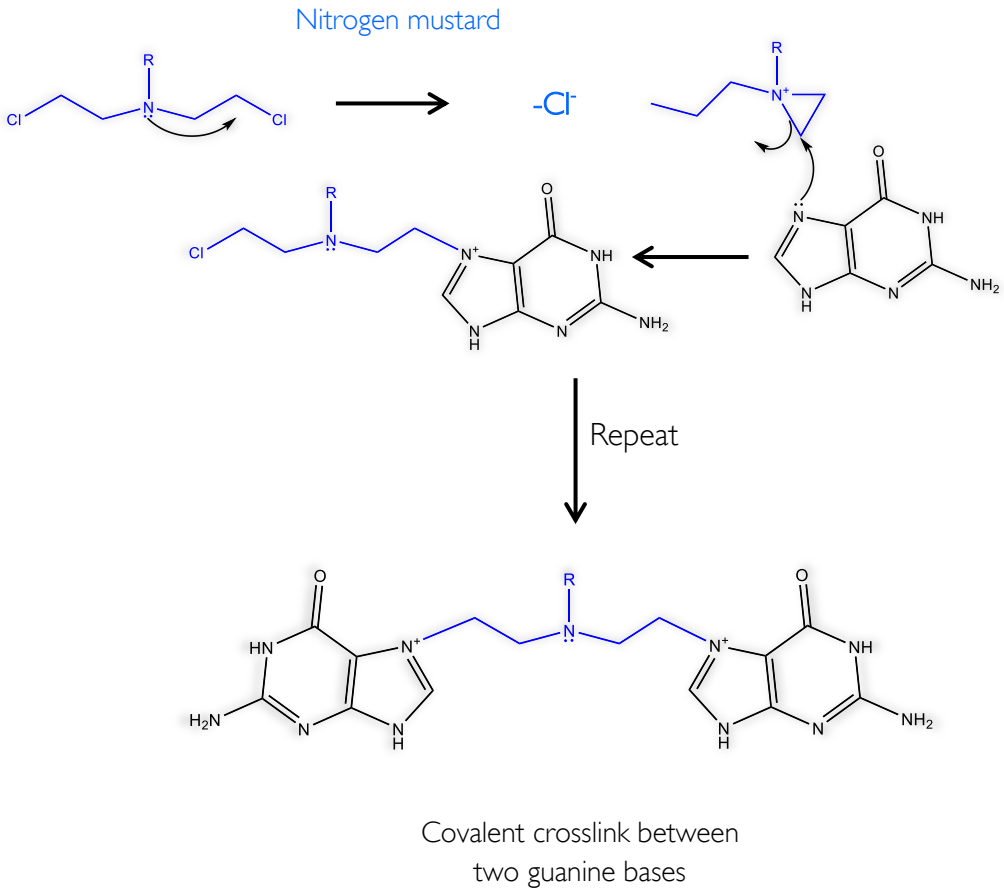


Figure 1.24 | **Mechanism of guanine-guanine crosslink formation mediated by Nitrogen mustard**. Note that only one side of the attachment is shown, the same mechanism is repeated on the other side. Figure adapted from (Polavarapu *et al.*, 2012).

Repair of interstrand covalent crosslinks (ICLs) requires sophisticated DNA repair machinery, from different sources, which can act together to repair the lesion. Both nucleotide excision repair and homologous recombination are involved in restoring the ability of the cell to replicate DNA. Endogenous sources of ICLs are mainly electrophilic aldehydes that react with amino groups of purine nucleotides (the N^2 position of Guanine and the N^6 position of Adenine). Exogenous ICL inducers are chemicals such as nitrogen mustard (above) and mitomycin C, the latter of which is often used as a chemotherapeutic agents and inhibits DNA synthesis by forming crosslinks between Adenine at the N^2 position, and Guanine at either O^6 or N^2 (Verweij and Pinedo, 1990, Weng *et al.*, 2010).

Mechanism of repair depends on whether the lesion is within a replicative structure or not, and uses both homologous recombination (HR), translesion synthesis (TLS) and nucleotide excision repair (NER) as shown by Li *et al.,* 2008 (figure 1.25). Many endonucleases have been shown experimentally to be involved in the repair of ICLs in eukaryotes, specifically XPF-ERCC1, MUS81-EME1, SLX1-4 and FAN1 (Zhang and Walter, 2014). A more detailed view of the process, showing the endonucleases involved in eukaryotic ICL repair is shown in figure 1.26.

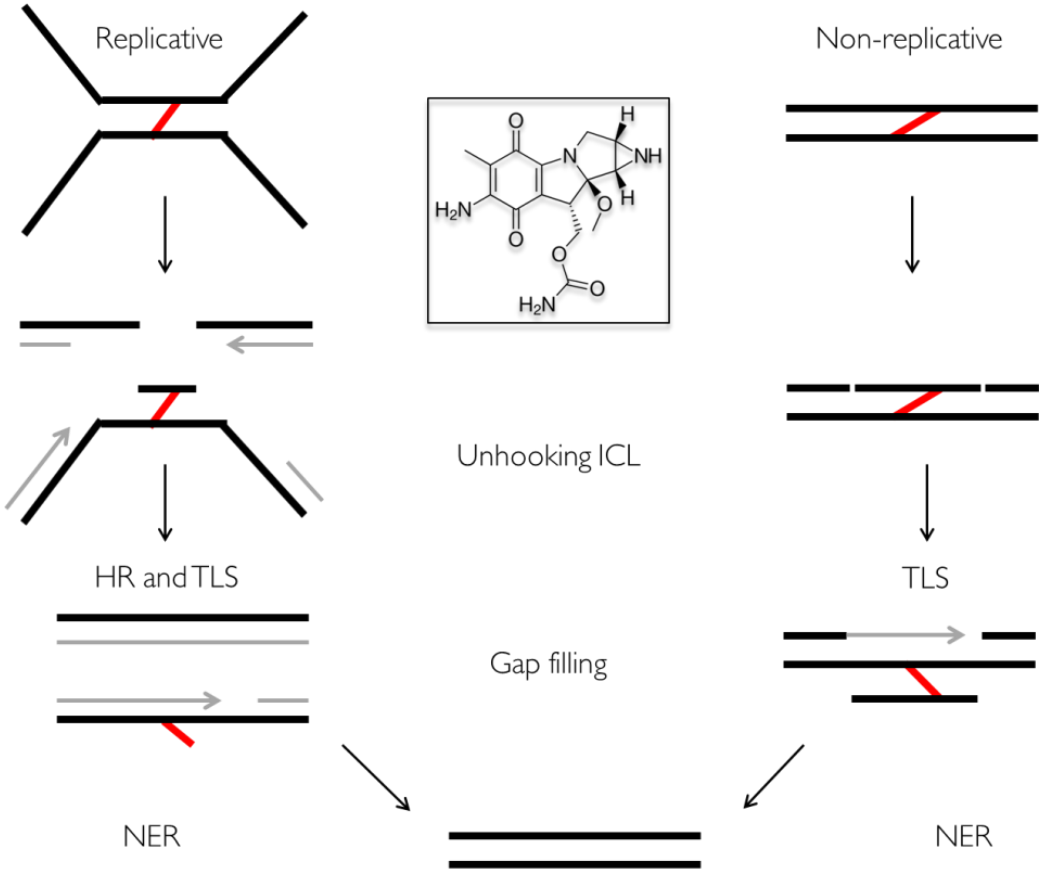


Figure 1.25 | **Overview of repair of ICLs**. Repair of replicative ICLs involves DNA nicking, unhooking and gap filling using either HR or TLS (left). Non-replicative structures use endonuclease mediated cleavage followed by TLS (right). Both structures involve a final NER step to deal with the DNA adduct. Figure adapted from (Rogers *et al.*, 2020).

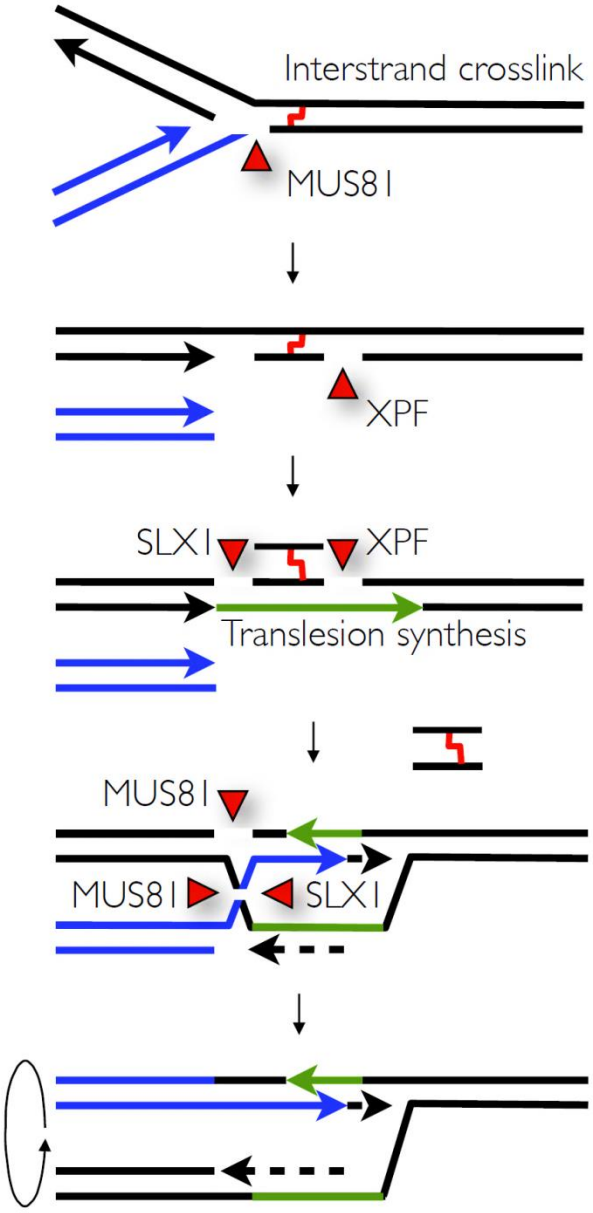


Figure 1.26 | **Endonucleases involved in ICL repair.** A DSB is generated via MUS81 on the leading strand template. The ICL is unhooked by proteins including XPF, SLX1, MMR and Fanconi anemia proteins. TL polymerases are recruited to fill in the gap. Replication is restarted using break-induced replication to create a Holliday junction that can be cleaved by either MUS81 or SLX1. Strand exchange proteins not shown, however Rad51 is used in eukaryotes and the homologue RadA is used in archaea. Figure adapted from (Brill, 2013)

Intercalating agents

Intercalating agents are capable of intercalating with DNA bases (Figure 1.27), which then affects base-stacking van der Waals interactions between the rungs of the nucleotide ladder, thereby lengthening and unwinding the double helix. A common example is the compound ethidium bromide (EtBr), which is often used in the laboratory during the process of gel electrophoresis. However, more generally, intercalating agents and small molecules are promising anti-cancer drugs (Braña *et al.*, 2001, Rescifina *et al.*, 2014).

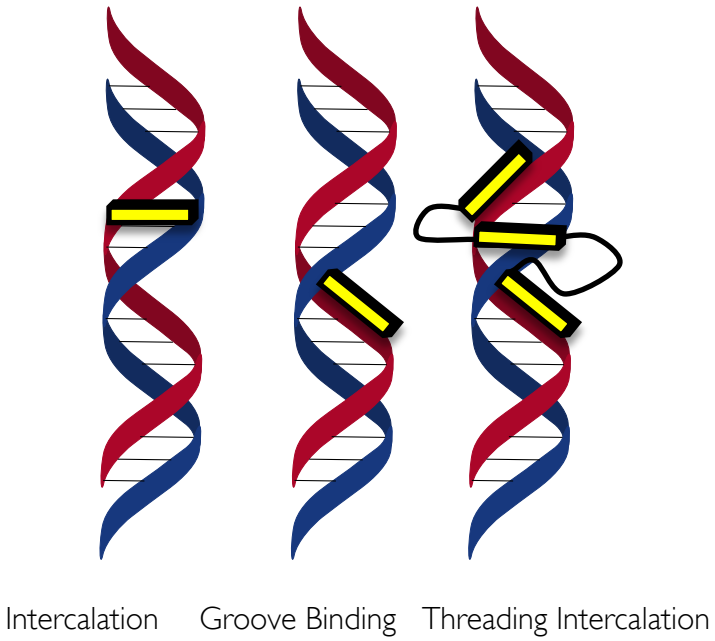


Figure 1.27 | **Schematic representation of classical intercalation, groove binding and threading intercalation mode of DNA.** Classical intercalating molecules are free from bulky substituents and can intercalate without any significant part of the bound molecule in either the major or minor groove. Groove binding intercalators have bulky substituents that must be in one groove or the other. Threading intercalators carry bulky substituents next to the intercalating moiety, and the substituents are placed in the major and minor grooves. Figure adapted from (Rescifina *et al.*, 2014)

1.5.4 DNA damage in space

The predominant form of DNA insult in space comes from solar radiation, both ionizing radiation such as gamma rays, as well as UV radiation, both of which cause significant damage to the DNA template directly or indirectly via reactive oxygen species (Jones *et al.*, 2007). Dosage estimates of space radiation exposure is ~1 mSv/day. This is almost 150 times higher than that on the surface of the Earth (Ohnishi, 2004). These examples of DNA damage are discussed at length in the previous section on exogenous DNA damage. Microgravity also leads to DNA damage and detrimental health defects (West, 2000) and, on a molecular level, has been shown to induce apoptosis in human cells (Dang *et al.*, 2014). Furthermore, DNA repair pathways themselves appear to be negatively affected by simulated microgravity, leading to a decrease in DNA repair capacity. Real-time PCR has shown that under such conditions, gene expression for DNA repair machinery is reduced, especially mismatch repair machinery (Kumari, Singh and Dumond, 2009).

Consideration of DNA repair mechanisms to deal with the above types of damage is critical to the logistics behind any space mission, including potential colonization.

1.6 DNA Repair

1.6.1 Direct repair of DNA damage

Most organisms are able to quickly and directly reverse various forms of chemical DNA damage, without the need for complex signalling cascades and repair proteins. There are three examples of how lesions in DNA are repaired directly: single-stranded breaks are ligated by DNA ligase, CPDs formed from UV light are corrected by a DNA photolyase and alkylation damage is removed, in the case of methyl groups, by methyltransferase enzymes.

Single-strand DNA breaks

Single-stranded DNA breaks are caused, as described in the previous section, by direct ionizing radiation or indirectly via the radiolysis of water producing reactive radical species. However, other, more subtle, mechanisms do exist such as the failure of Okazaki fragment ligation during DNA replication. In the presence of 'clean' 5' and 3' ends, repair is facilitated by the action of DNA ligase which catalyses the formation of a phosphodiester bond via a nucleophilic attack mechanism. In essence, after an adenylyl group is transferred from the enzyme to the 5' phosphate group at the ssDNA break, the 3'OH group acts as a nucleophile to the negatively charged pyrophosphate group and the adenylyl group is displaced, sealing the gap and producing AMP (Figure 1.28).

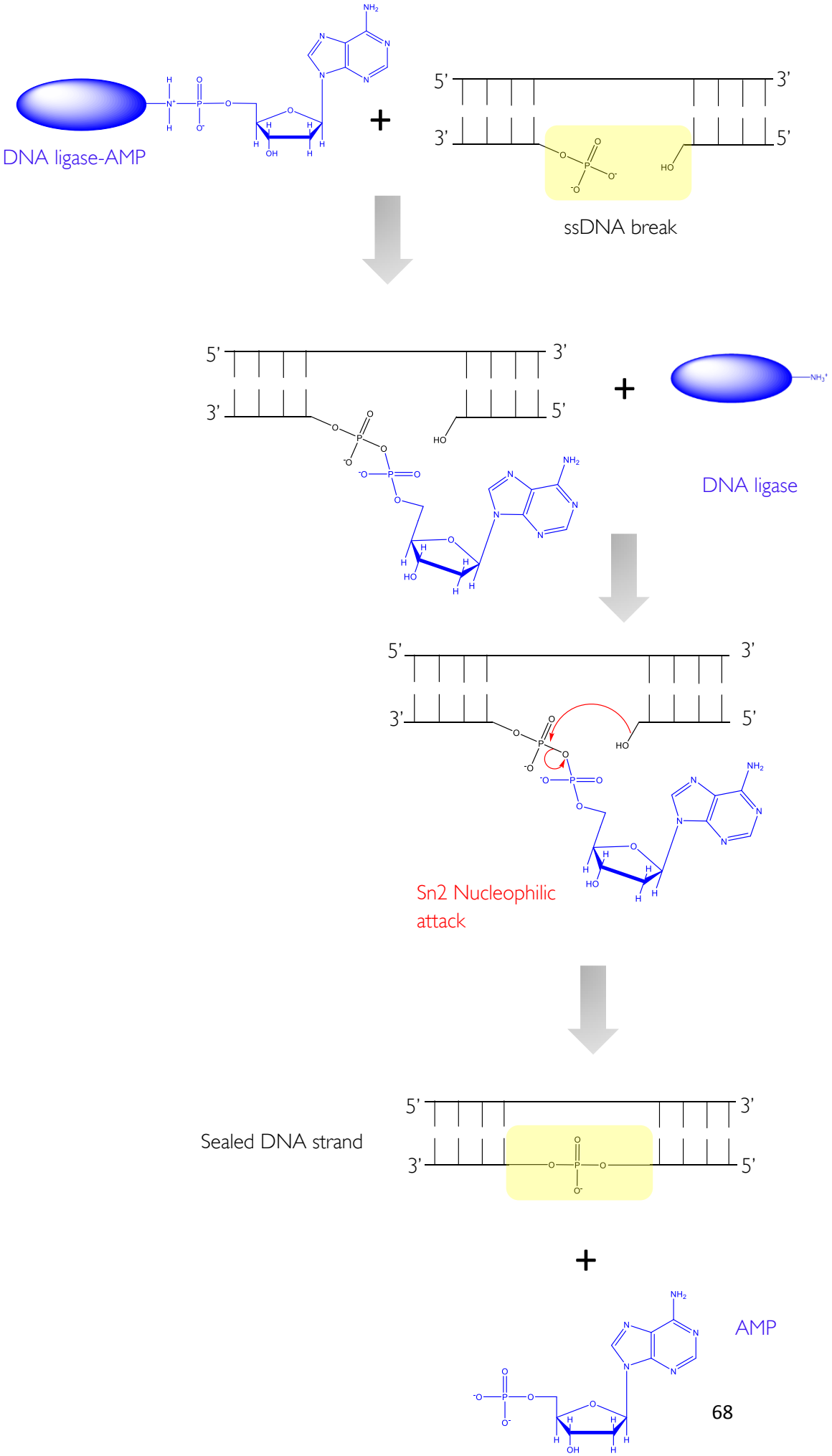

DNA ligase-AMP
5'
3'
3'
5'
ssDNA break
5'
3'
3'
5'
DNA ligase
5'
3'
3'
5'
Sn2 Nucleophilic attack
Sealed DNA strand
5'
3'
3'
5'
AMP

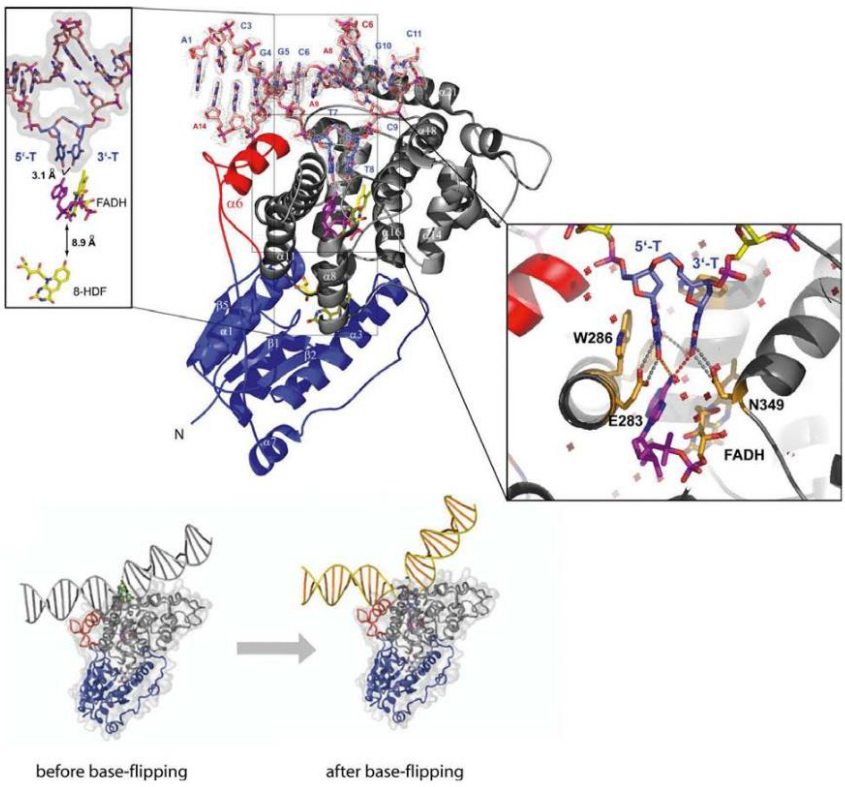
Figure 1.28 | **Mechanism of DNA ligase-mediated phosphodiester bond formation.** In the first step an adenylyl group is transferred from ATP or NAD to a conserved lysine residue on DNA ligase (not shown). The adenylyl group is then transferred to the 5' phosphate group at the DNA break. Subsequent nucleophilic attack by the 3'OH group facilitates closure of the DNA break and release of AMP. Figure adapted from (Lehman, 1974).

Pyrimidine dimers (photoreactivation)

Photoreactivation is the process by which damage caused by UV, such as cyclobutane pyrimidine dimers and 6,4-photoproducts,are repaired directly using lesion specific photolyase enzymes in the process of photoreactivation which wasdiscovered by Albert Kelner in 1949 (Kelner, 1949).

DNA photolyases are responsible for light-induced repair of pyrimidine dimers and contain light-absorbing chromophores. Each enzyme contains two chromophores: all contain flavin adenine dinucleotide ($FADH^-$), however the second cofactor can be one of two groups, either pterin folate (5,10-MTHF) or deazaflavin (8-HDF) (Sancar, 1994). These cofactors associate with the photolyase and use energy from visible light to break open bonds between the bases via electron transfer to the lesion and free radical attack. CPD photolyase uses a base-flipping mechanism to flip the thymine dimer into its active site (Figure 1.29).

A


5'-T
3'-T
3.1 Å
FADH
8.9 Å
8-HDF
α6
N
W286
N349
E283
FADH
before base-flipping
after base-flipping

B

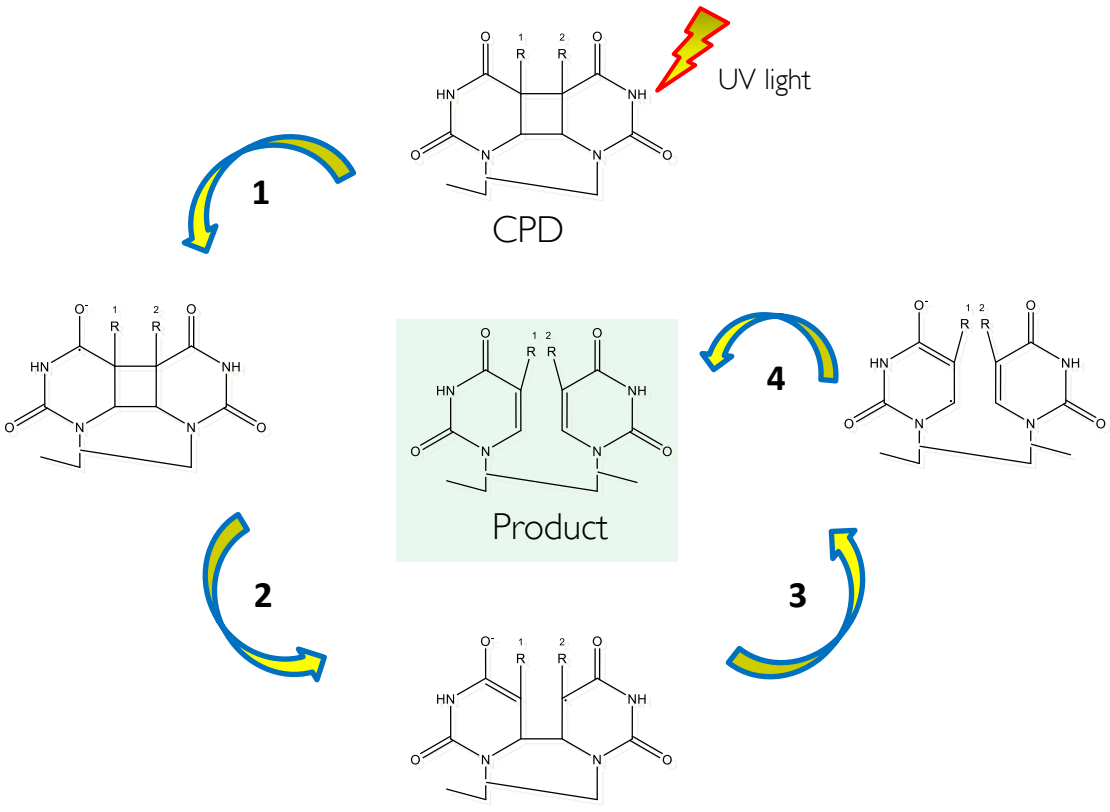

UV light
CPD
1
2
3
4
Product

Figure 1.29 | **Mechanism of CPD photolyase.** (A) CPD photolyase uses a base-flipping mechanism to flip out the damaged bases into its active site for electron transfer. Figure taken from (Essen and Klar, 2006). (B) Chemical mechanism of electron transfer and breakage of the CPD covalent linkage. Figure adapted from (Liu *et al.*, 2011).

Placental mammals lack the photolyase enzymes and therefore are unable to repair UV induced lesions directly and instead rely on excision repair.

Repair of alkylation damage: O^6-methylguanine

O^6-methylguanine is formed by chemical agents such as EMS or MMS, as discussed earlier. The resultant adduct allows guanine to pair with thymine (or cytosine still), leading to a G●C → A●T transition (Loechler, Green and Essigmann, 1984). Direct reversal of this damage is as a result of removal of the O^6-methyl group via methylguanine DNA methyltransferase enzymes (MGMT). MGMT flips out the damaged base and removes the methyl group. A cysteine (Cys38) located in the enzyme's active site accepts the methyl group leading to enzyme inactivation and dissociation from the DNA (Olsson and Lindahl, 1980). The specific enzyme that carries out de-methylation of the O^6-methyl group from guanine, in *E. coli*, is the Ada protein, which is activated to become a transcription factor which then activates the adaptive response in bacteria, inducing expression of genes including itself and *alkA, alkB* and *aidB* (Lindahl *et al.*, 1988, Takahashi *et al.*, 1988). MGMT enzymes are found across bacteria, budding and fission yeast and humans.

1.6.2 Excision repair

Excision repair embroils multiple pathways, each acting slightly differently to bring about the same outcome, depending on the extent of the damage across the DNA strand. The principle involved in all excision repair pathways is fairly simple, the damaged base(s) are removed and an endonuclease cuts the phosphodiester backbone. The backbone is removed, allowing DNA polymerase to add a new dNTP, replacing the damage. The undamaged strand acts as a template for polymerase mediated gap fill.

1.6.2.1 Base Excision Repair

The base excision repair pathway (BER) likely evolved to deal with natural endogenous chemicals, such as ROS. BER mainly deals with small, non-bulky lesions affecting a single (or very few) bases of the DNA, examples include replacing incorrectly incorporated uracil bases (either from replication error or cytosine deamination) or damaged bases such as 8-oxoguanine. Although exact mechanistic details vary between organisms and species, common themes are present (Figure 1.30):

(1) Damage recognition and removal of damaged base(s) by a lesion-specific DNA glycosylase, resulting in an abasic site.
(2) Cleavage of the phosphodiester backbone, to allow access to the damaged site, by an apurinic/apyrimidinic (AP) endonuclease or AP lyase.
(3) Removal of the remaining backbone fragment by a lyase or phosphodiesterase.
(4) DNA polymerase-mediated gap fill, using the undamaged strand as a template.
(5) Sealing of the phosphodiester backbone by DNA ligase.

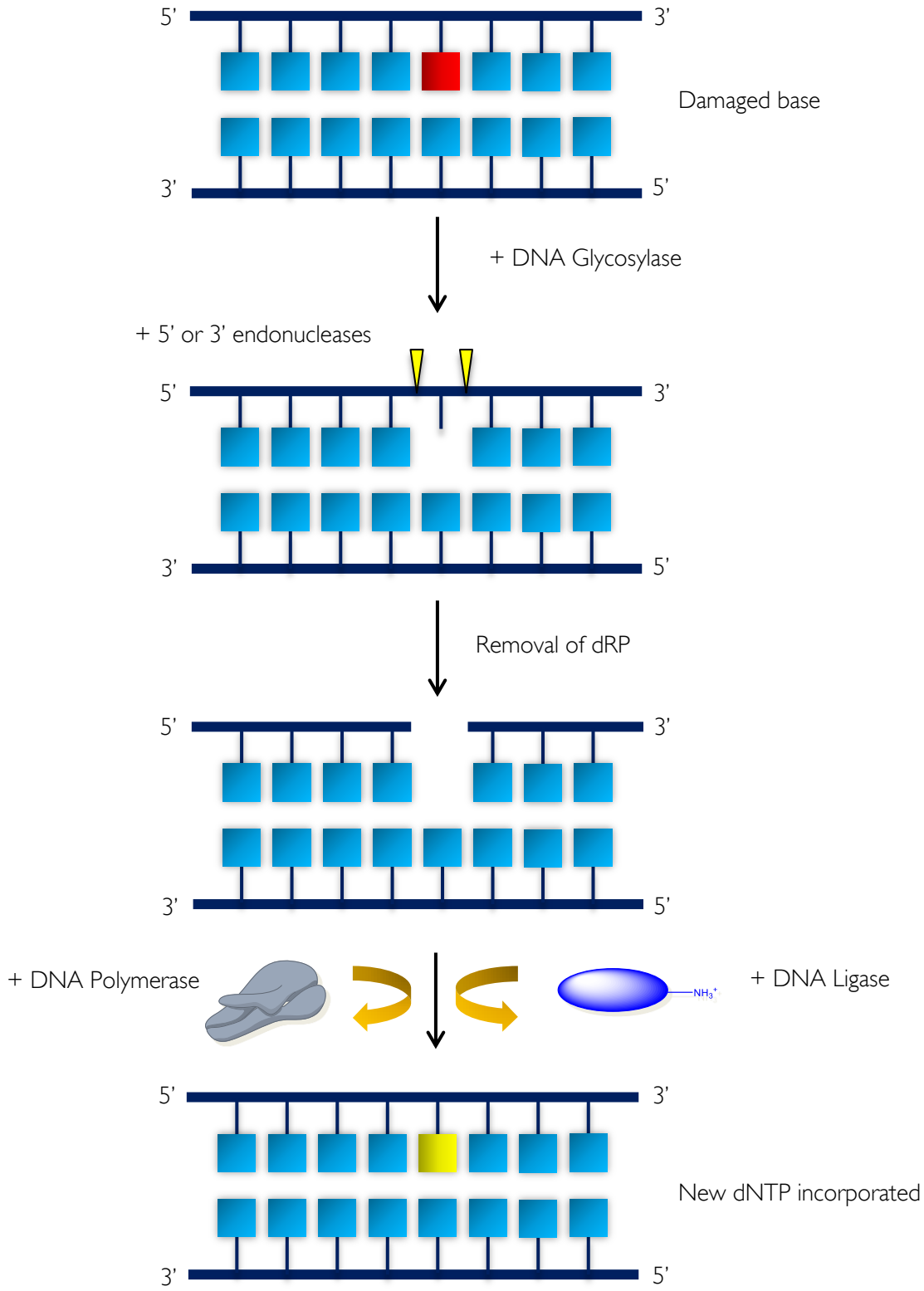


Figure 1.30 | **Base excision repair.** DNA lesion-specific glycosylases remove the damaged base leaving an AP site. AP endonucleases cleave the sugar phosphate backbone (either 5' or 3') to remove the baseless deoxyribose-phosphate (dRP). DNA polymerase fills the gap and DNA ligase seals the nick (Kim and Wilson , 2012).

Lesion specific DNA glycosylases initiate the BER pathway by recognising the DNA damage and using electrostatic orientation along the glycosylase active site along with insertion of the amino acid Arg272 into the DNA via the minor groove, the resultant pressure allows the damaged base to be flipped out (Slupphaug *et al.*, 1996). Subsequent catalysis of the cleavage of the N-glycosidic bond within the damaged nucleotide leaves an abasic site (Lindahl, 1980, Nilsen, Lindahl and Verreault, 2002). An AP endonuclease then generates a single strand break either upstream (5') or downstream (3') of the damaged base. AP endonucleases have been classified into two families depending on homology to either Exonuclease III, or Endonuclease IV of *E. coli* (Bonura, Schultz and Friedberg, 1982, Demple *et al.*, 1999). Class 1 AP endonucleases, also known as AP Lyases, cleave 3' of the AP site via a β-elimination reaction (and are also all DNA glycosylases) forming a 3'-α,β-unsaturated aldehyde and a 5'-phosphate. By contrast, class 2 AP Endonucleases, the most abundant type, cleave 5' of the AP site via a hydrolysis reaction, thus forming 5'deoxyribose phosphate and 3'-hydroxyl termini (Mosbaugh and Linn, 1980) (Figure 1.31). Cleavage 3' via a class 1 AP Endonuclease requires subsequent processing via a 3'→5' phosphodiesterase. Similarly, cleavage by a class 2 APEendonuclease requires processing via a 5'→3' phosphodiesterase.

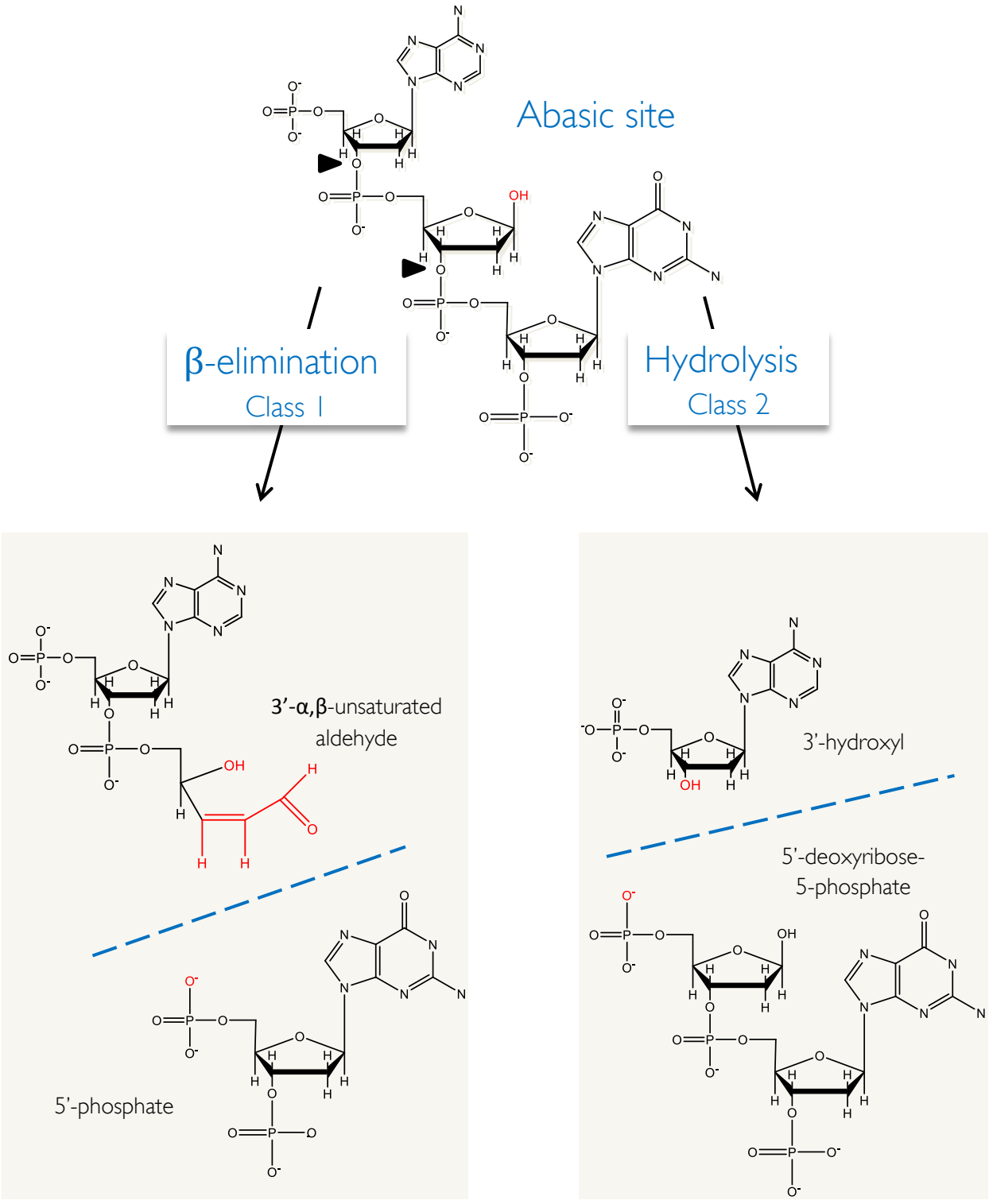


Figure 1.31 | **Chemical structures of abasic site and cleavage positions and products of the two types of AP endonuclease.** Cleavage positions indicated by black triangles. Figure adapted from (Li and Wilson 3rd, 2014).

In mammalian cells, BER can occur in one of two ways, either short-patch BER or long-patch BER, which result in the replacement of a single nucleotide and 2-6 nucleotides respectively. Short-patch repair is more common, however long patch repair can occur in places such as the mitochondria which encounters substantial oxidative damage. Long-patch BER generates a longer flap structure during the DNA synthesis phase and the overhang is cleaved by the flap-specific endonuclease FEN1. DNA ligase then seals the gap. However cells lacking FEN1 are still somewhat competent at long-patch BER and thus is was discovered that another flap removing enzyme, DNA2, was also involved in mammalian cells (Copeland and Longley, 2008, Robertson *et al.*, 2009). In contrast, short-patch repair involves a repair polymerase, such as DNA polymerase β in humans, to fill the gap. Ligation is carried out once the gap is filled (Matsumoto and Kim, 1995).

Returning momentarily to the lesion 8-oxoguanine as discussed above, in humans this is repaired by BER using 8-oxoguanine DNA glycosylase (hOGG1) which scans the DNA for 8-oxoguanine. Using cytosine on the opposite strand as a guide to specific guanine residues, upon encountering a damaged guanine, hOGG1 subsequently removes the damaged base by cleaving the N-glycosidic bond within the guanosine triphosphate (Bruner, Norman and Verdine, 2000, David, 2005). A summary of common DNA glycosylases that carry out the first step of BER can be found in Table 1.5.

Table 5.1. | Summary of common DNA glycosylases found in humans and *E. coli.*

Activity	*E. coli*	Humans	Substrate	References
Uracil DNA Glycosylase	Ung	UDG	Uracil	(Dusseau *et al.*, 2001)
3-meA DNA Glycosylase I	Tag		3-methyladenine	(Sakumi *et al.*, 1986)
3-meA DNA Glycosylase II	AlkA	MPG	3-methyladenine, 7-methyladenine, 3-methylguanine, 7-methylguanine	(O'Brien and Ellenberger, 2004), (Nakabeppu, Kondo and Sekiguchi, 1984), (Samson *et al.*, 1991), (Chakravarti *et al.*, 1991)
FaPy DNA Glycosylase	Fpg (MutM)		Formamido-pyrimidine	(Boiteux, O'Connor and Laval, 1987)
PD DNA Glycosylase	T4 DenV		Pyrimidine dimers	(Grafstrom, 1986)
Adenine DNA Glycosylase	MutY	MYH	Removes A when mispaired with G	(Luncsford *et al.*, 2013), (Zhang *et al.*, 1998)
Thymine DNA Glycosylase		TDG	Removes T when mispaired with G	(Neddermann and Jiricny, 1993)

1.6.2.2 Nucleotide Excision Repair

In humans, Nucleotide excision repair (NER) is the principle repair pathway for cyclobutane pyrimidine dimers formed by UV light, as placental mammals lack direct repair mechanisms such as a photolyase. Lesions repaired by NER tend to have distorted the DNA helix and NER machinery can be activated via global repair pathways or a specialised form of repair, namely, transcription coupled repair (TC-NER) which allows preferential repair of the transcribed strand, should transcription be halted due to a bulky lesion such as a thymine dimer. Defects in TC-NER can lead to premature ageing syndromes such as Cockayne syndrome and Trichothiodystrophy whereas defects in global NER leads to extreme photosensitivity and cancer predisposition syndromes such as Xeroderma Pigmentosum (XP) (Gillet and Schärer, 2006, Hanawalt and Spivak, 2008). Global NER will be the main focus of the following paragraphs.

As with base excision repair, the overall process is similar, however the mechanistic minutiae are very different between organisms. NER begins by damage recognition of the bulky DNA backbone and subsequent bimodal cleavage of the phosphodiester backbone by recruited endonucleases. Strand unwinding is then facilitated by the action of DNA helicases and then the DNA gap, which is larger than that in BER, is filled in by DNA polymerase and ligase, using the template strand as a guide (Figure 1.32).

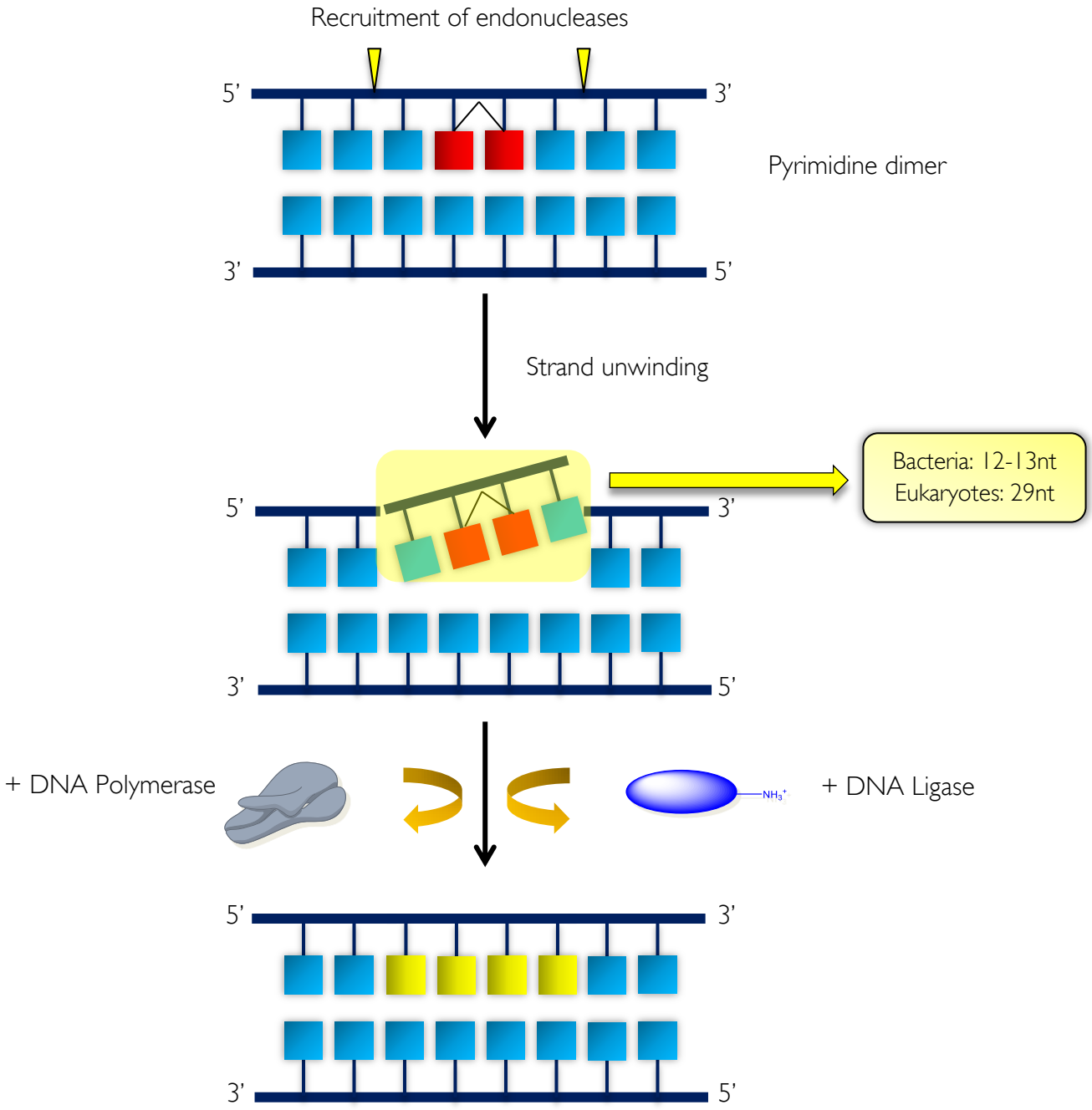


Figure 1.32 | **Nucleotide excision repair mechanism.** Endonucleases are recruited to the DNA lesion and make incisions 5' and 3' of the damage. Helicases displace the oligonucleotide strand, which is longer in eukaryotes than bacteria, and DNA polymerase, along with DNA ligase, fill in the gap and seal the nick. Adapted from (Sancar and Rupp, 1983).

NER begins, in *E. coli*, with the UvrAB complex (consisting of A_2B_1 stoichiometry) identifying the DNA lesion via backbone distortion. UvrA, the damage sensor, binds to the DNA lesion and the UvrB helicase is then loaded by UvrA and kinks the DNA to allow unwinding of DNA in the region of the damage (Jaciuk *et al.*, 2011). UvrB then takes over and inserts a β-hairpin between the two strands of the double helix, locking itself in place, the nucleotide directly behind the β-hairpin is flipped out and inserted into a small, conserved, pocket in UvrB. UvrA then leaves and plays no further part in the repair pathway. Once locked in place, and only then, UvrB recruits an exonuclease called UvrC, consisting of two nuclease sites which allows cleavage of the phosphodiester backbone 5' and 3' of the lesion (Verhoeven *et al.*, 2000). First UvrC relocates to the 3' side, where the first incision is made, then subsequent relocation to the 5' side occurs where a second incision is made. In bacteria incisions are ~12nt apart. Another protein, known as Cho (UvrC homologue) has been found and can incise the DNA at the 3' side of a lesion during NER. The incision site is located 4nt further away from the damage, possibly due to binding to a different domain of UvrB, compared to the cut site of UvrC (Moolenaar *et al.*, 2002). Cho may be used to deal with lesions that have been poorly cut by UvrC, for example if the lesion is too big, translocation by UvrC form the 3' to the 5' side of the lesion is not possible. Finally, UvrD, a superfamily 1A helicase, is recruited to the oligonucleotide via interaction with UvrB (Manelyte *et al.*, 2009). Subsequent unwinding of the oligonucleotide in a 3'→5' direction occurs via two motor domains in UvrD that use ATP hydrolysis to facilitate translocation along the DNA. The subsequent gap is filled by DNA polymerase 1 and the nick is sealed by DNA ligase, releasing UvrBC (Caron, Kushner and Grossman, 1985, Fischer, Maluf and Lohman, 2004). UvrB is thought to remain associated with the DNA until its displacement by DNA polymerase 1 and UvrC has been shown to be displaced by UvrD, moving it out of the way and allowing DNA polymerase to carry out DNA synthesis (Orren *et al.*, 1992).

In eukaroytes, NER is carried out by a series of factors: XPA, RPA, TFIIH, XPC-RAD23B (also known as HR23B in humans), XPG and ERCC1-XPF. In eukaryotic systems, DNA damage and helix distortion is sensed by XPC and RAD23B which binds opposite the lesion. XPC, similar to UvrB, inserts a β-hairpin between the DNA strands and the two damaged bases are flipped out, unlike in bacteria where UvrB flips out the base adjacent to the damage. A double β-hairpin domain binds two nucleotides opposite the lesion, which as a result of lesion-determined DNA instability, show increased ssDNA geometry. Thus, thermodynamic stability is probed by XPC before docking with the DNA (Gunz, Hess and Naegeli, 1996).The role of HR23B is not well understood, although it is thought to act to protect XPC from ubiquitin-mediated damage (Ng *et al.*, 2003). XPC, however, recognizes a wide range of diverse DNA structural changes and can recognize a 'bubble' DNA duplex distortion with and without an associated lesions allowing for an increased, non-lesion-specific, repertoire of DNA binding affinities (Sugasawa *et al.*, 2001). HR23B appears to lack damage recognition capabilities so its activity in NER is thought to be conferring stability to XPC as association of XPC with UV-induced lesions is impaired in the absence of RAD23/HR23 proteins (Sugasawa *et al.*, 1996, Bergink *et al.*, 2012). Of note is that XPC-HR23B is part of a larger, trimeric complex with Centrin-2, a calcium binding protein (Nishi *et al.*, 2005), the role of which within the NER pathway remains elusive. Further complexity is added owing to the lack of binding affinity of XPC-HR23B for cyclobutane pyrimidine dimers (Sugasawa *et al.*, 2001).

The missing link is that DDB2 is also involved and has a hydrophobic-binding pocket ideally suited to bind CPDs, however 6,4-photoproducts have been shown to not be excised via DDB-mediated stimulation, however XPC is able to carry out this function (Wakasugi *et al.*, 2001; Fitch *et al.*, 2003). The exact role of DDB2 appears to be in facilitating NER in the context of chromatin and handing over the UV-induced lesion to XPC via ubiquitylation (Wakasugi *et al.*, 2002; Sugasawa *et al.*, 2005).

The next factor to be recruited is the transcription factor TFIIH, which is recruited via direct interaction with XPC-HR23B and the helicase subunits XPB and XPD are responsible for unwinding the oligonucleotide (Yokoi *et al.*, 2000). Engagement of XPD with the lesion enables assembly of the pre-incision complex (PIC) where helicase action generates a local unwinding of the damaged microenvironment. Once this is formed the endonucleases XPA and XPG, along with RPA, are located to the site of damage independent of one another and XPC-HR23B dissociates at this point, stabilising the PIC. Eukaryal XPD has been shown to proofread for the presence of a DNA lesion in the translocated strand as a mechanism with which to increase fidelity of the NER reaction (Mathieu *et al.*, 2013). Subsequent recruitment of XPF/ERCCI nuclease, via interaction with XPA (Li *et al.*, 1994), forms the final pre-incision complex and duel incision, either side of the lesion, occurs 5', mediated by ERCC1-XPF, and 3', mediated by XPG (O'Donovan *et al.*, 1994, Evans *et al.*, 1997). Finally the DNA gap is processed via replicative polymerase-mediated fill (DNA Pol δ/ε), although DNA Pol ε appears to perform this task with higher fidelity than Pol δ, and the nick is ligated by DNA Ligase 1 (Shivji *et al.*, 1995).

In contrast, most Archaea contain homologous proteins of the eukaryotic NER machinery such as XPF, XPG, XPB and XPD. The detection of bulky lesions is the essential first step for all NER pathways. In transcription-coupled repair, the stalled RNA polymerase acts as a DNA damage sensor and it is thought that archaeal RNA polymerase could also function in this fashion, however this remains to be confirmed as Archaea lack homologues of TCR machinery found in bacteria such as Mfd (Dorazi *et al.*, 2007). A summary of putative NER genes is summarised in a selection of the Archaea in Figure 1.33.

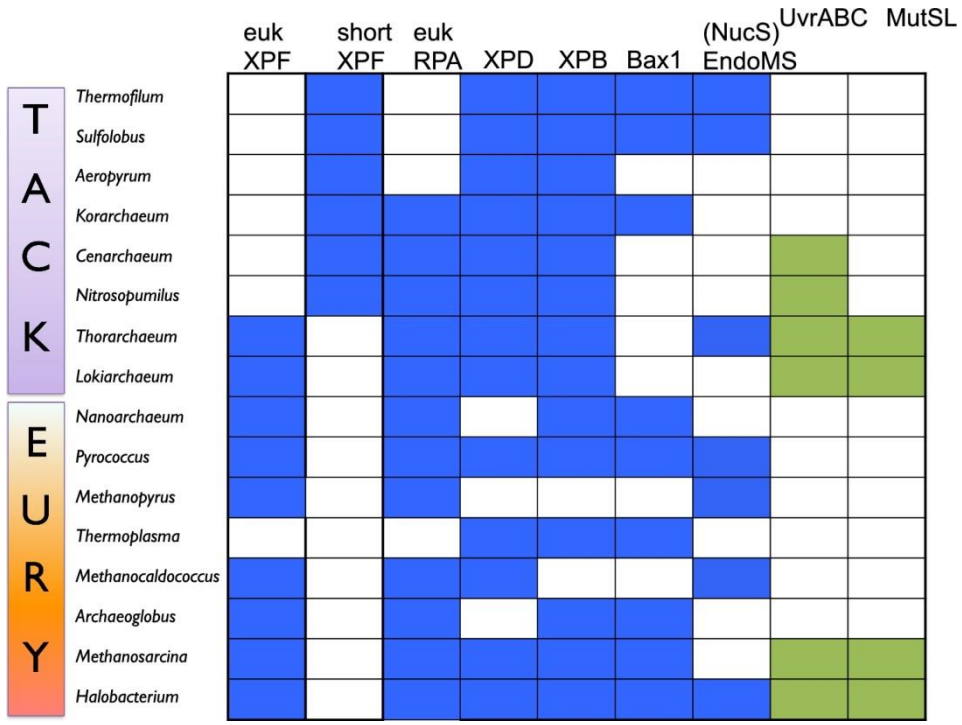


Figure 1.33 | **Distribution of DNA repair genes in Archaea.** Genus names on the left are organised as members of the TACK superphylum and Euryarchaea. For each genus, a shaded box indicates the presence of the relevant gene. Bacterial genes probably acquired by lateral gene transfer are shown in green, others in blue. Figure taken from (White and Allers, 2018).

No XPC-HR23B homologues have been identified in archaea, therefore it remains unclear how DNA damage is detected in archaea lacking the bacterial UvrABC system. The exception is that archaeal SSBs are implicated in the DNA damage response and mutants of *Halobacterium* NRC1 showing increased ionizing radiation resistance were shown to have upregulated expression of RPA proteins (DeVeaux *et al.*, 2007) and thus SSB proteins could be involved

in assisting with DNA damage sensing and recruitment of repair factors, this remains to be confirmed however SSB proteins are able to melt damaged DNA (Cubeddu and White, 2005). Deletion of the genes for UvrA, B or C in *Halobacterium* NRC1 resulted in UV sensitivity in spite of the homologs of XPF, XPB and XPD being present (Crowley *et al.*, 2006).

Most Archaea have eukaryotic-like NER proteins, although their function is far from solidified in archaeal genetics. Deletion of the XPF homologue, Hef, in the hyperthermophilic archaeon *Thermococcus kodakaraensis* resulted in increased sensitivity to mitomycin C (MMC) which is a cross-linking agent. This suggests a role for Hef in the repair of crosslinks (Fujikane *et al.*, 2010) as well as canonical NER, consistent with known roles of the eukaryotic XPF along with Mus81 in cross-link repair via homologous recombination (Kikuchi *et al.*, 2013). The structure of the XPD homologue from *Sulfolobus tokodaii* has been reported and the mutations leading to xeroderma pigmentosum in humans, could be mapped to the crystal structure of the archaeal enzyme. The archaeal XPD consists a four domain structure with two motor domains, an Arch and FeS domain (Constantinescu-Aruxandei *et al.*, 2016, White and Allers, 2018). A lesion recognition pocket adjacent to the pore through which XPD pulls ssDNA was identified in the eukaryotic XPD. Two amino acids, Tyr-192 and Arg-196 were identified as key players and mutations at these positions reduce activity of repair (Mathieu *et al.*, 2013). The same mutations in the XPD homologue in the archaeon *Ferroplasma acidophilum*, Tyr-171 and Lys-175, resulted in abolishment of the ability to deal with the lesion, as it usually stalls at CPD lesions (Mathieu *et al.*, 2013, White and Allers, 2018). XPB homologues are also present in archaea, although mechanistic details remain elusive but are thought to potentially involve Bax1, allowing helicase action and unwinding by XPB and cleavage by Bax1, acting in unison as a helicase-nuclease machine (Rouillon and White, 2010). In summary, Archaea lacking bacterial-like UvrABC systems remain elusive. The lack of UvrABC could be compensated for via recruitment of other factors which are, as yet, unknown. Furthermore, completely different excision pathways could exist as

well as the potential for leaving the lesions to stall the replication fork and dealing subsequently with the damage via replication fork restart (Grogan, 2015).

1.6.2.3 Mismatch Repair

Mismatch repair (MMR) functions to rectify mis-incorporated bases in DNA in the event that canonical fidelity mechanisms fail to spot the error before the template is replicated and a mutation is introduced. The MMR mechanism and proteins are conserved across most bacteria and eukaryotes, however only a limited number of archaeal species possess such MMR proteins, indicating that other mechanisms might be at play (Grogan, 2004). However, as mutational mismatches are fixed in the DNA, a critical task for the MMR machinery is to identify parental and daughter DNA strands in order to correctly rectify the mismatch (Figure 1.34).

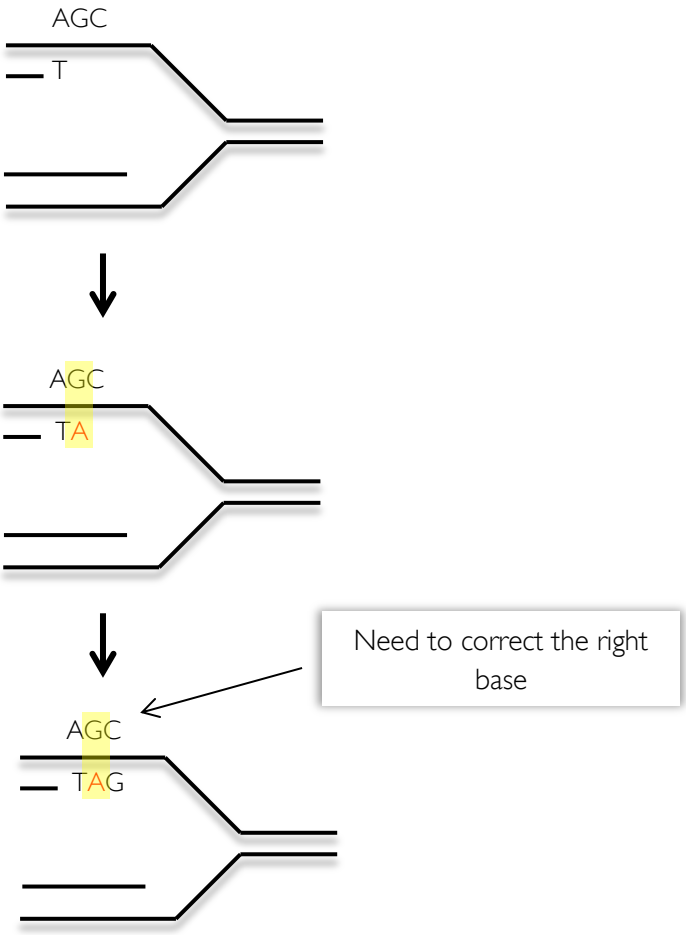


Figure 1.34 | **Strand discrimination is a key step in MMR.** Once the replication fork has passed and introduced the mutation, a G●A base pair now exists. In order to correctly deal with the mutation, MMR machinery must recognise that the adenine base has been mis-incorporated, and correct it, rather than the guanine.

Strand discrimination in *E. coli* is carried out via methylation of specific 5'-GATC-3' sites within DNA by DAM methylase. MMR proteins can sense the methylation status of each DNA strand owing to the hemi-methylated state of DNA during DNA replication. The nascent DNA strand is un-methylated for a time owing to the lag of the DAM methylase in methylating the adenine base in the GATC site of the nascent strand (Lu, 1987, Barras and Marinus, 1989).

Bacterial MMR proceeds, firstly, via damage recognition by the binding of MutS-ADP homodimer at the mismatch (Su *et al.*, 1988). MutS, once bound to the mismatch, initiates a 60° kink in the DNA with the mismatch at the vertex of the bend, leading to a widening of the major groove around the mismatch (Natrajan *et al.*, 2003). Mismatch binding by MutS involves a key amino acid, Phe36, which binds to one of the mismatched bases (Yamamoto *et al.*, 2000) in addition to ATPase activity, although the latter of which is poorly understood. It is thought that, upon binding to the mismatch, MutS-ADP ATPase activity is altered and ATP hydrolysis is suppressed in one of the two monomer subunits, becoming the rate-limiting step of the reaction (Antony and Hingorani, 2004). Upon binding to DNA, MutS has been shown to use a sliding clamp mechanism (figure 1.35) to scan the DNA duplex for mismatches, and upon finding one, stalls briefly, altering ATPase activity to increase ATP levels, thus forming a more stable complex with the mismatch and allowing recruitment of downstream MMR effectors (Jeong *et al.*, 2011) (Jiricny, 2006). The role of ATP hydrolysis in the general context of bacterial MMR is not well understood.

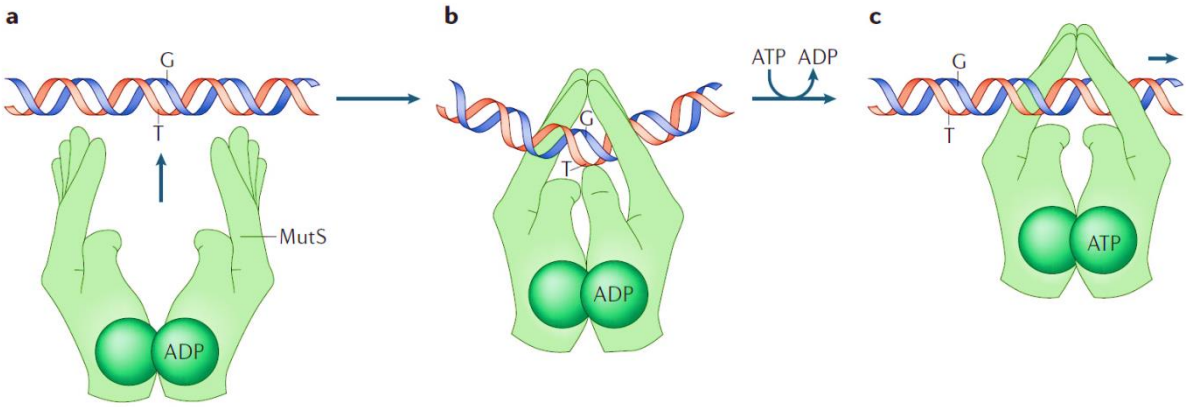


Figure 1.35 | **The MutS sliding clamp and its activation.** (A) The ADP-bound MutS homodimer binds to a G-T mismatch in duplex DNA. In the absence of DNA the finger domains are open and ATP-binding sites are dimerized. (B) In the presence of a mismatch, the ADP-bound MutS wraps around the DNA and is anchored at the mismatch by a Phe-X-Glu wedge that is intercalated into the minor groove. (C) Exchange of ADP for ATP brings about a conformational change, releasing the Phe-X-Glu wedge from the mismatch site, allowing the clamp to freely translocate along the DNA in either direction, allowing searching for a strand discrimination signal. Figure taken from (Jiricny, 2006).

Once MutS has bound DNA and altered its ATPase activity, it recruits MutL, which can only recognise the specific ATP-bound MutS in the sliding clamp conformation (Yang *et al.*, 2022) (Acharya *et al.*, 2003). MutL changes the properties of mismatch-bound MutS, lengthening the stall time, preventing it from sliding away too far from the mismatch, which his has been shown to do when isolated (Qiu *et al.*, 2015). MutL then activates the MutH endonuclease to make incisions in the unmethylated DNA strand at the nearest GATC site, either 5' or 3' of the mismatch (Hall and Matson, 1999). Via interaction with MutL, the 3'→5' helicase, UvrD, loads at the nick and unwinds the DNA duplex towards the mismatch, from the nick generated by MutH (Dao and Modrich, 1998, Hall, Jordan and Matson, 1998, Yamaguchi, Dao and Modrich,

1998). Downstream events depend on the location of the nick, either 5' or 3' to the mismatch. Cleavage 5' of the mismatch requires endonucleolytic degradation by 5'→3' endonucleases such as ExoVII and RecJ whereas a 3' nick requires nucleases such as Exo1 or ExoX which degrade in a 3'→5' direction (Cooper, Lahue and Modrich, 1993, Burdett *et al.*, 2001). The resultant gap is filled by the canonical replicative DNA Polymerase III and the gap sealed by DNA Ligase (Figure 1.36) (Lahue, Au and Modrich, 1989).

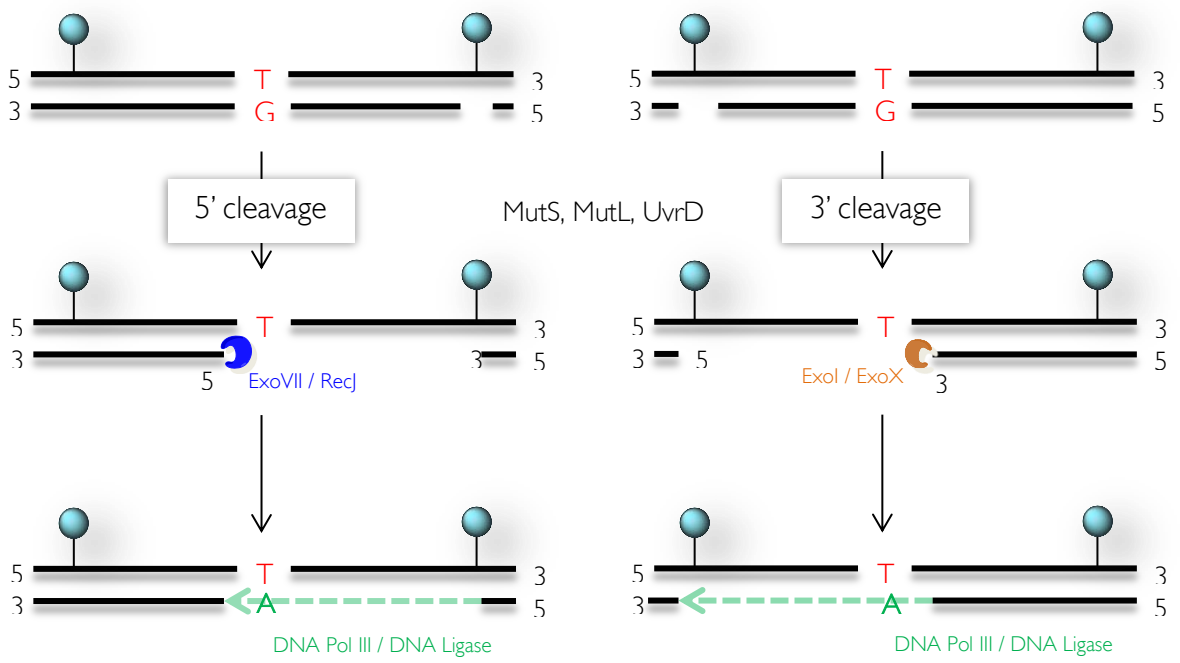


Figure 1.36 | **Overview of bacterial MMR nick processing.** Figure stages described in text. MutS is converted into a sliding clamp that allows movement along DNA to search for the methylation signal. MutL has been shown to trap MutS at the mismatch, so in this model, DNA looping occurs to facilitate the search.

In eukaryotic MMR, strand discrimination is carried out by gaps in the DNA strands. In the case of the lagging DNA strand, nicks are present before DNA ligase acts to join the Okazaki fragments and the PCNA sliding clamp is thus still associated with the DNA end which can act in a similar fashion to MutH and recruit the MMR machinery (Clark *et al.*, 2000, Iyer *et al.*, 2008). The leading strand is more of a contentious issue, and not clear cut, however current thinking appears to be that ribonucleotides act as signals for MMR on the leading strand. In this mechanism, ribonucleotide repair begins at the site of the ribonucleotide and RNaseH2 nicks the DNA, leading to a gap which can act as a signal for MMR (Ghodgaonkar *et al.*, 2013).

MMR is initiated via recognition of a mismatch by homologues of MutS and MutL, known as MutSα (and MutSβ) and MutLα respectively. There are no known homologues of MutH, owing to the different mechanism of strand discrimination employed by eukaryotes. The first step is binding of MutSα to the mismatch in DNA (MutSβ can act to recognise larger insertion deletion loops (Sharma *et al.*, 2014)). Similar to bacterial MMR, mismatch binding provokes an ADP→ATP exchange that converts the MutSα/MutSβ into sliding clamps that diffuse along the DNA, facilitating the mismatch search and binding by MutLα (Gradia *et al.*, 1999). Once bound, MutLα forms a complex with MutSα, MutLS which diffuses along the DNA until a PCNA sliding clamp is encountered, indicating the nascent strand. Interaction between PCNA and MSH6/3, subunits of MutSα and MutSβ respectively, has been shown and can occur either side of the mismatch (Kleczkowska *et al.*, 2001). The side on which the PCNA is encountered determines downstream events.

If the MutLS ternary complex slides in the 5’ direction and encounters a PCNA:RFC complex, it recruits the exonuclease, Exo1, to displace the RFC clamp loader. The activated Exo1 then degrades DNA from the nick in the 5’→3’ direction. Alternatively, if the MutLS complex slides in the 3’ direction, Exo1, being a 5-→3’ exonuclease can not directly act to degrade the DNA strand from the gap at the 3’ end of the mismatch. In spite of this, the human MMR process has been reconstituted *in vitro* and it has been shown that

bidirectional nick degradation can occur using Exo1 (Constantin *et al.*, 2005) (Fang and Modrich, 1993). This can occur owing to the latent endonuclease function of MutLα and the looping of DNA to facilitate sliding of MutLS in the 3' direction. MutLα can therefore make a nick 5' of the mismatch allowing Exo1 to carry out 5'→3' degradation (Kadyrov *et al.*, 2006). An alternative hypothesis is that Exo1 contains a cryptic 3'→5' exonuclease activity which could be triggered via interaction with PCNA:RFC. Replicative DNA polymerases (e.g. DNA Polymerase δ for the lagging strand) then load at the 3' terminus of the original discontinuity, carrying a bound PCNA moiety. The gap is then filled using the parental strand as a template and DNA ligase 1 seals the nick (Jiricny, 2006) (Figure 1.37).

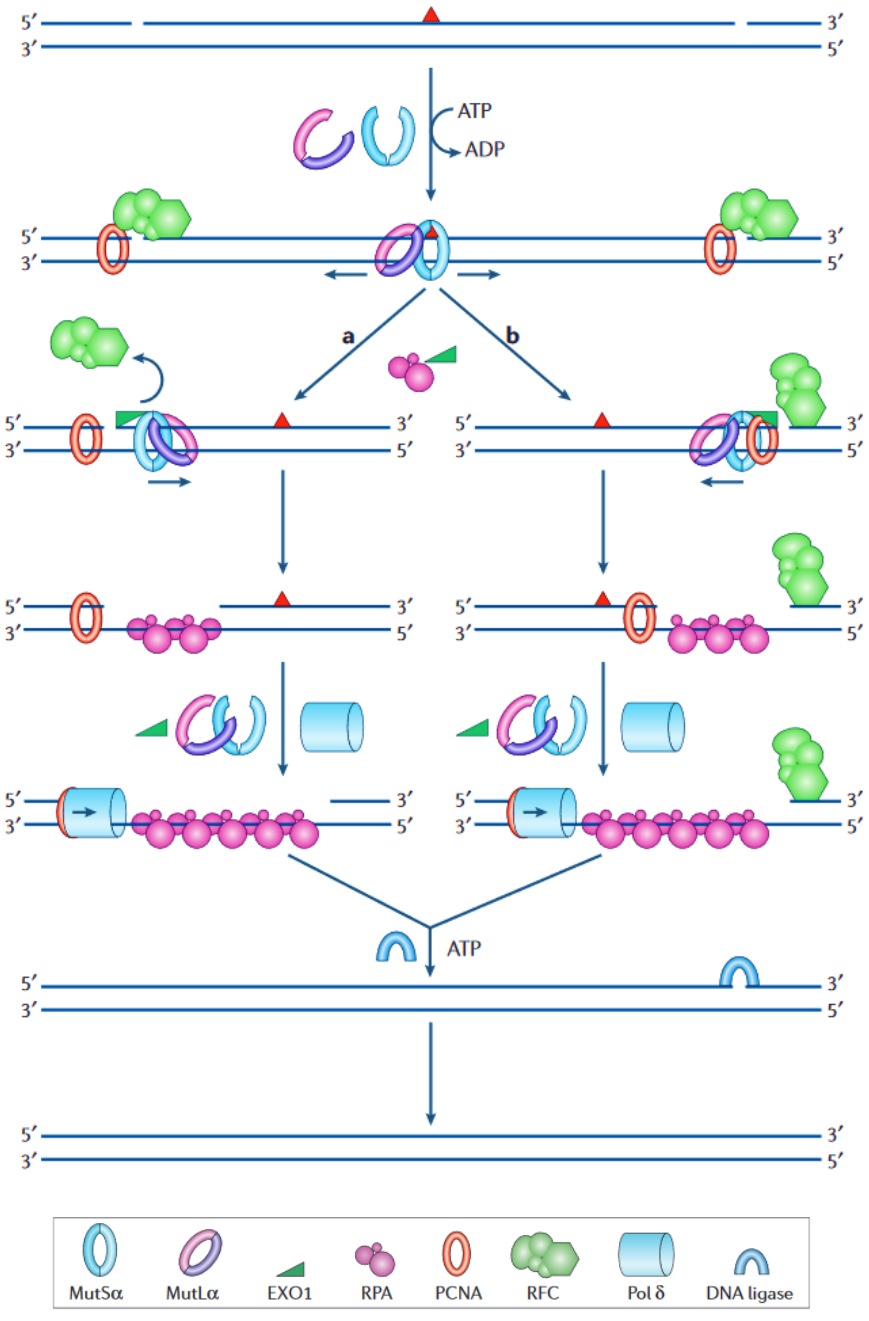


Figure 1.37 | **The reconstituted human mismatch-repair system.** The mismatch is bound by MutSβ (or MutSβ) and recruits MutLα. The ternary complex undergoes an ATP-driven conformational switch, releasing the sliding clamp from the site of the mismatch. Claps then diffuse either 5' or 3' of the mismatch, looping out DNA. (A) Those that migrate 5' will encounter PCNA:RFC and load Exo1, which will degrade the DNA strand in a 5'→3' direction. The ssDNA gap is bound by RPA. (B) Clamps that migrate 3' of the mismatch will also encounter PCNA:RFC from the opposite side and will also recruit Exo1. Once the mismatch is removed by Exo1, DNA polymerase δ fills in the gap and the nick is ligated via DNA ligase. Diagram shows lagging strand and is adapted from (Jiricny, 2006).

Archaea generally lack MMR machinery with any homology to MutS/MutL in either bacteria or eukaryotes (Ishino *et al.*, 2016). In spite of this, some such as *Sulfolobus spp* do exhibit mutational avoidance mechanisms, suggesting alternative means of dealing with mismatches (Grogan, 2004). EndoMS, a mismatch specific endonuclease has been found in archaeal species such as *Thermococcus kodakarensis* which uses a base flipping mechanism to deal with mismatched bases and has been shown to interact with PCNA. This suggests a novel MMR system exists in Archaea, starting with initiation of a dsDNA break by EndoMS (Ishino *et al.*, 2016, Nakae *et al.*, 2016). An advantage of using a dsDNA break to kick off the process has been suggested, being that there is no need to distinguish between parental and nascent DNA strands to pinpoint the incorrect base, as both strands are resected during double strand break repair (DSBR). Further evidence for this mechanism is provided due to the fact that EndoMS has been found in an operon with the recombinase, RadA which is critical in performing homology searching in DSBR (Ren *et al.*, 2009). However, many questions remain such as the necessity of PCNA/EndoMS interaction for activity and if it is possible to locate EndoMS at the replication fork, further indicating interaction with PCNA and a role in MMR and/or DNA replication.

1.6.3 Homologous recombination

Homologous recombination (HR) is a, generally, error-free process that can act to repair the most toxic lesions in a cell, double-stranded DNA breaks. Double stranded breaks are so toxic owing to the gap generated, with no intact strand to act as a template, the correct order of missing bases can not be determined, or indeed the number of intervening bases between the dsDNA ends generated via the initial insult. Homologous recombination requires the involvement of a homologous (near-identical) sequence of DNA, in eukaryotes this requires cells to be in either S or G2 phase of the cell cycle, providing an identical sister-chromatid for HR. If the cell is in G1, HR is inhibited in favour of Non-homologous end joining (NHEJ), a more error prone mechanism of DNA repair (Zhao *et al.*, 2017). In other organisms, such as bacteria and Archaea that are polyploid, this is less of an issue and archaeal survival in mutagenic environments is aided by this lack of a regulated cell cycle and constant presence of multiple genome copies. Aside from performing a key part in DSBR within the cell, HR also occurs deliberately during meiosis, resulting in either a cross-over event or a gene-conversion both increasing the genetic diversity of the gametes produced as well as ensuring the correct homologues have aligned in the correct fashion, providing a signal for the mitotic spindles to retract and ensuring correct chromosomal disjunction. Further to this, HR is also a multi-functional process and can contribute to restarting stalled replication forks and removing interstrand crosslinks, along with NER (Neale and Keeney, 2006). Homologous recombination plays a further role in human disease. Mutations in recombination mediators such as the BRCA1/BRCA2 genes can lead to a predisposition to breast and ovarian cancers (Venkitaraman, 2002). It should be noted, however, that although considered to be hypomutagenic, homologous recombination depends on searching out homology to other sequences in the genome, this can lead to issues in humans owing to the repeated nature of their genome, leading to potential homology searches outside of the chromatid microenvironment generated as a result of S phase,

and thus lead to large scale genomic rearrangements (Bishop and Schiestl, 2000, Hoang *et al.*, 2010).

Homologous recombination must therefore be highly regulated to ensure accurate and time-specific activation. The overall mechanism is conserved between the three domains of life and is carried out in three major steps: Pre-synapsis, Synapsis and Post-synapsis, all of which will be discussed in the next section (Figure 1.38). Although the SDSA pathway is noted, it will not form a major element of this chapter.

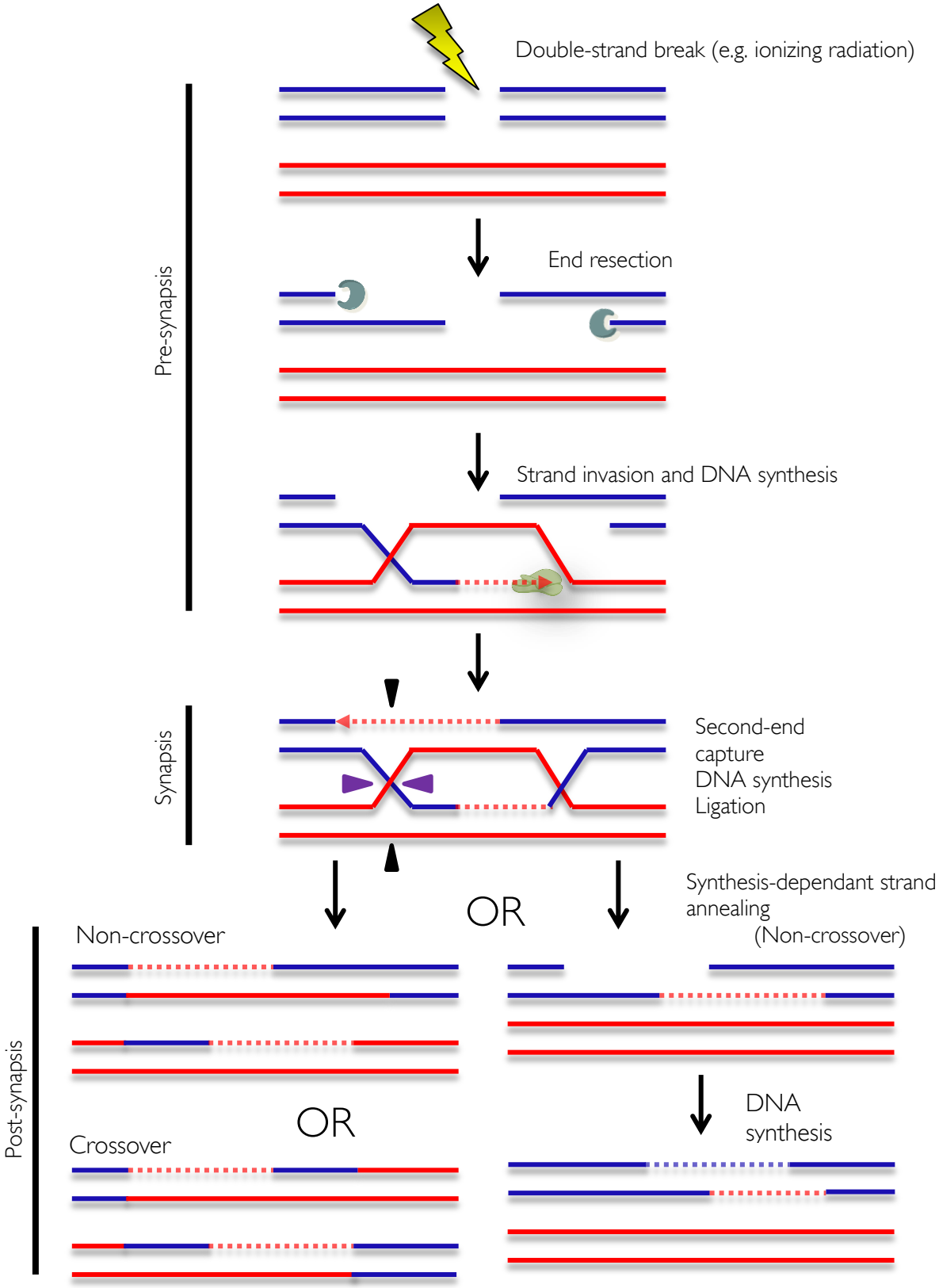

Double-strand break (e.g. ionizing radiation)
Pre-synapsis
End resection
Strand invasion and DNA synthesis
Synapsis
Second-end
capture
DNA synthesis
Ligation
OR
Synthesis-dependant strand
annealing
(Non-crossover)
Non-crossover
OR
Crossover
DNA
synthesis
Post-synapsis

Figure 1.38 | **Overview of homologous recombination and potential end products.** After a DSB is generated, either via damage, or deliberately during meiosis, the 5' ends of each duplex are resected to leave ssDNA 3' tails. One of these 3' tails invades a homologous duplex, searching for homology. Once homology is found the 3' tail binds to the bottom strand of the homologous duplex, thus creating a displacement loop from the top strand of the invaded duplex. The second 3' tail is then captured to pair with the D-loop and DNA synthesis occurs, followed by ligation, generating two points of strand exchange known as Holliday junctions. Resolution can occur one of two ways via endonuclease mediated cleavage, generating either a non-crossover (gene conversion) product or a crossover product. Alternatively, there need not be any recombination at all and via a mechanism known as synthesis dependant strand annealing, the first 3' tail can be displaced via helicase action once DNA synthesis has occurred. Subsequent polymerase mediated fill can then occur as the missing bases have been determined and ligation can seal the nick, resulting in an alternative gene conversion product and an unmodified duplex. SDSA is favoured in G1 phase of the cell cycle, when a sister chromatid is not present as recombination does not occur. Blue DNA is the damaged DNA, Red DNA is the homologous duplex. Polymerase and exonucleases are shown in cartoon form and filled triangles indicate cut sites for endonucleases. Dotted DNA indicated newly synthesised DNA and the colour indicates the template used.

1.6.3.1 Pre-synapsis

Pre-synapsis is the initial stage of homologous recombination and is instigated via a double stranded DNA break. During this process, dsDNA is resected in a 5'→3' direction, generating ssDNA 3' overhangs (tails). Recombinases are loaded onto the 3' tails, displacing single-stranded DNA binding proteins, and facilitate strand invasion of a homologous duplex. In bacteria, the recombinase is RecA, in eukaryotes it is Rad51, in the case of DNA repair and DMC1 specifically in meiosis. In Archaea the recombinase is RadA.

Pre-synapsis in *E. coli*

Pre-synapsis is mediated, in the main, by the RecBCD enzyme protein machinery, although the RecFOR pathway serves as a backup function. The principle function of RecBCD is to carry out end resection and load the recombinase, RecA, onto the 3' ssDNA tails.

The process begins by RecBCD encountering and binding to a double-strand break in the DNA, covering approximately 20 nucleotides. The 3'-ended strand is bound to the RecB helicase and the 5' end is threaded through RecC and bound to the RecD helicase (Figure 1.39) (Ganesan and Smith, 1993). RecB and RecD are two helicase motors that can act on the two strands of DNA simultaneously. RecB, a slow helicase, acts on the 5'→3' strand and RecD, a fast helicase, acts on the 3'→5' strand resulting in a ssDNA loop occurring on the 5'→3' strand (Taylor and Smith, 2003). The two DNA strands are separated either side of a pin structure within RecC, facilitating duplex separation by the helicase subunits (Singleton *et al.*, 2004). As this process occurs, RecB contains a second functional exonuclease domain and is able to digest both strands of DNA, for this reason the product of the RecBC loci is also known as Exonuclease V (Goldmark and Linn, 1972).

ssDNA tails are generated via modulation of the RecB nuclease in response to a specific sequence, known as a Chi (χ) site, which was shown to consist of the DNA sequence 5'-GCTGGTGG-3', which is recognised by the RecC subunit (Smith *et al.*, 1981). It has been proposed that χ sites are recognise by RecC as the ssDNA passes through a channel within the RecC subunit. This view has been challenged slightly by one group who also discovered that this signalling requires additional nucleotide involvement on the 3' side of the χ site, which extend into the RecB Nuclease domain and thus recognition may also depend on RecB (Amundsen, Sharp and Smith, 2016). It was also found that degradation was not uniform between the two strands, the strand 3' terminal at the entry site of RecBCD is degraded more vigorously than the 5' terminal strand. Once RecC interacts with a χ site, the 3' nuclease activity, although not the helicase activity, is attenuated (Dixon and Kowalczykowski, 1993). Further analysis has also shown that whilst the 3'→5' nuclease activity of RecB is attenuated, nuclease activity of opposite polarity is up-regulated, resulting in a 5'→3' degrading nuclease via a RecC-mediated signal (Anderson and Kowalczykowski, 1997a). It is not known exactly how this occurs, but one idea is that upon recognition of a chi site, RecC clamps down on the 3' tail, allowing the 5' tail to access the nuclease site more readily and thus be cleaved at a higher frequency (Singleton *et al.*, 2004).

However, a second hypothesis has emerged that involves signalling by RecC, once a χ site is encountered, to RecD to stop unwinding DNA which in turn signals to RecB to cut DNA four – six nucleotides to the 3' side of χ, which has historically been shown to occur (Taylor *et al.*, 1985), and to begin loading RecA, thus providing a mechanism by which the 5' terminal strand is not cut. The two mechanisms have been shown to occur under different conditions and concentrations of Mg^{2+} and ATP, the ratio of which seems to determine the method of choice (Amundsen *et al.*, 2007). Overall the exact biochemical nature of the RecBCD-Chi interaction is not conclusive.

In either case the result is formation of a ssDNA 3' tail onto which RecB then loads RecA, displacing SSB (Anderson and Kowalczykowski, 1997b). The point at which RecBCD dissociates from the DNA is unknown.

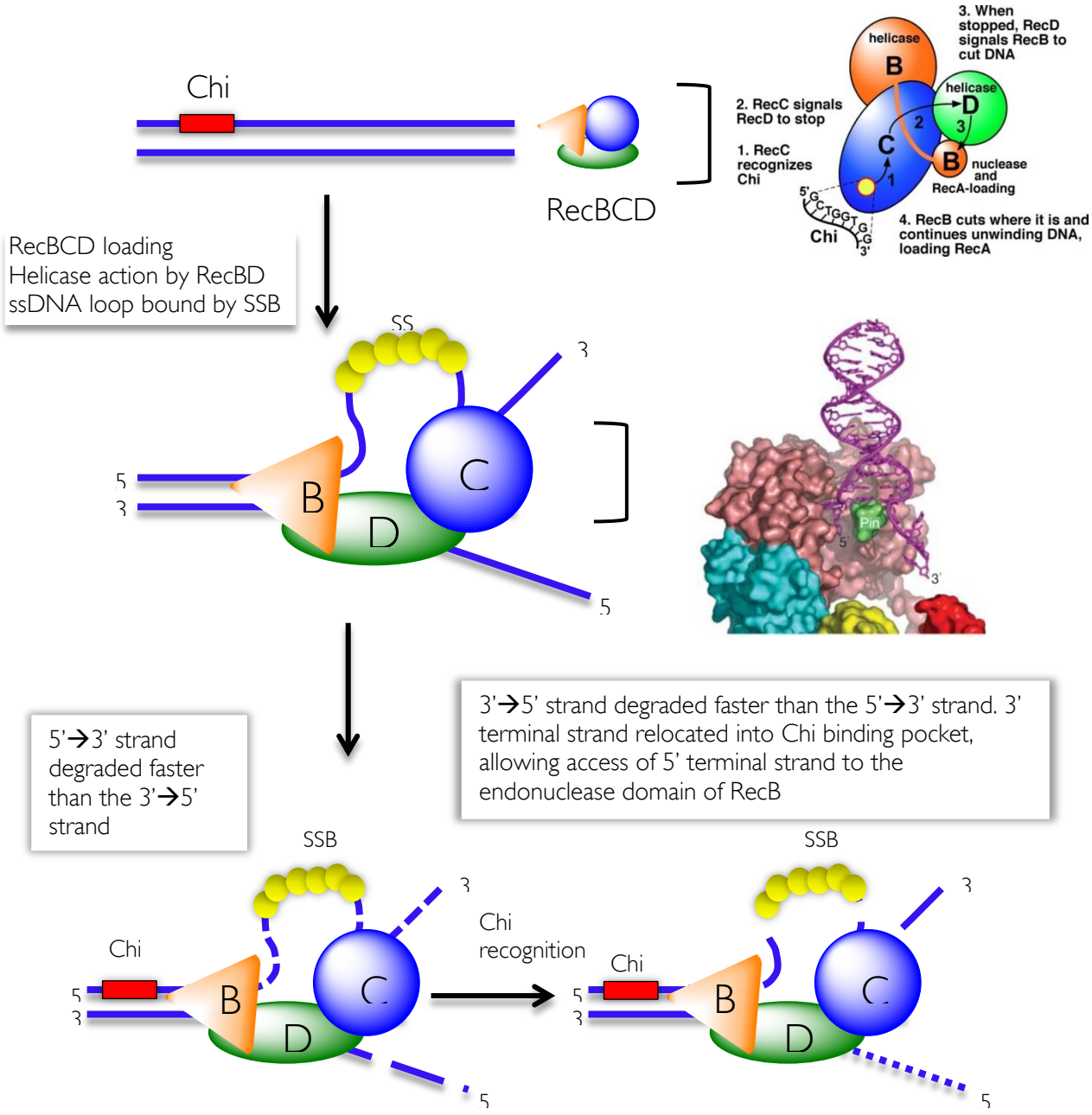


Figure 1.39 | **Overview of RecBCD biding to a DSB.** Detail discussed in text. Cartoon RecBCD image taken from (Amundsen *et al.*, 2007). Pymol structure of RecC pin separating DNA strands taken from (Singleton *et al.*, 2004).

RecA, the strand exchange protein in *E. coli*, is then loaded by RecB onto the ssDNA loop on the 3' terminal strand. A dimer of RecA-ATP is required and cycles on and off DNA using ATP hydrolysis, nucleating randomly on ssDNA, competing for binding to DNA with SSB, forming a right handed nucleoprotein filament in a 5'→3' direction. A total of six RecA dimers are found per turn of the helix, thus extending the length of the DNA filament by 1.5x relative to B form DNA thereby facilitating homology searching in the synapsis stage (Stasiak and Di Capua, 1982, Egelman and Stasiak, 1986, Bell *et al.*, 2012, Lovett, 2012, Wu, Lu and Li, 2017).

Lastly, it should be noted that RecBCD acts to carry out all steps and load RecA without needing to recruit additional factors. End resection can be carried out separately by recruited factors, in the absence of RecBCD. These include the RecQ 3'→5' helicase and a 5'→3' exonuclease, RecJ, generating a 3' ssDNA tail. However, they are unable to load RecA onto the ssDNA tail (Morimatsu and Kowalczykowski, 2014). Further suppressor screens have identified further factors that can function to carry out this process when RecQ and RecJ are unable to function, due to *recB* or *recC* mutants. RecQ and RecJ can only function if mutations are present in SbcB Exonuclease I and SbcCD nuclease that digest 3' tails and ds/ssDNA respectively, which can both act to degrade the DNA break before RecQ and J can act (Ivancic-Bace, Salaj-Smic and Brcic-Kostic, 2005). Further processing ability is provided by SbcA mutation which activated the nuclease, Exonuclease VIII (*recE*) which can degrade DNA in the same fashion as RecJ, generating a 3' ssDNA tail (Kushner, Nagaishi and Clark, 1974, Smith, 2012).

In order to then load RecA, in the absence of RecBCD, additional machinery must exist. Mutants for RecBCD have been isolated and the RecFOR genes

were discovered in suppressor screens and have the ability to load RecA, although this is not their primary function (Ivančić-Baće *et al.*, 2003).

Pre-synapsis in eukaryotes

In eukaryotes, end resection to generate 3' ssDNA tails is mediated by the Mre11-Rad50-Nbs1 (MRN) complex with CtIP (Sae2 in budding yeast), acting as a regulator of the resection process (Dolganov *et al.*, 1996, Trujillo *et al.*, 1998) (Sartori *et al.*, 2007). Mre11 interacts with both Rad50 and Nbs1 and contains both 3'→5' exonuclease activity and a separate endonuclease activity, thus is orthologous to SbcCD in bacteria (Paull and Gellert, 1998). Rad50 is a nucleotide binding protein that contains both walker A and B motifs, separated by a coiled coil structure. Overall, Rad50 consists of a coiled-coil super secondary structure, typical of the SMC proteins which are involved in chromatid cohesion and chromosome condensation. Rad50 acts as a dimer, mediated by a Zinc-Finger-like domain in the hinge region, binding both DNA ends and restricting diffusion, thus mediating correct repair downstream. Nbs1 is the most enigmatic component of the complex and associates with the complex as a monomer. Associating with Mre11 and containing a BRCA1 C-terminus domain as well as several phosphorylation sites for ATM and ATR Ser/Thr kinases, Nbs1 forms the flexible adaptor domain of MRN and acts as a regulator and protein recruiter facilitating DNA damage checkpoint, and subsequent p53 activation (D'Amours and Jackson, 2002, Stiff *et al.*, 2005). Upon a DSB, the MRN complex binds to each double strand end, linked together via the Rad50 coiled-coil zinc hinge, thus sensing the DSB (Petrini and Stracker, 2003). The MRN complex uses the nuclease activity of Mre11 to initiate resection by creating a nick that can act as an internal entry site for further nucleases capable of 5'→3' exonuclease degradation (Gobbini *et al.*, 2016). Rad50 binds ATP and has weak ATPase activity, allowing the ATP bound form of Rad50 to predominate. Rad50-ATP

has a closed conformation and in this state, Mre11 is an endonuclease and the 3'→5' exonuclease function is down-regulated, ATP can therefore act as a molecular switch between the endonuclease and exonuclease functions of Mre11, helping to prevent 3' end degradation (Majka *et al.*, 2012) (Cannavo, Reginato and Cejka, 2019) (Reginato and Cejka, 2020). Upon phosphorylation of CtIP at Thr847 (Anand *et al.*, 2016), Mre11 is activated to make an endonucleolytic cut on the 5' terminal strand and subsequent ATP hydrolysis by Rad50 activates the exonuclease function of Mre11 which can then resect the DNA strand in a 3'→5' direction. 5'→3' degradation can then occur via two redundant nucleases, Exo1 and Dna2 (which also requires the BLM helicase) (Zhu *et al.*, 2008). Overall this results in a ssDNA tail (Figure 1.40).

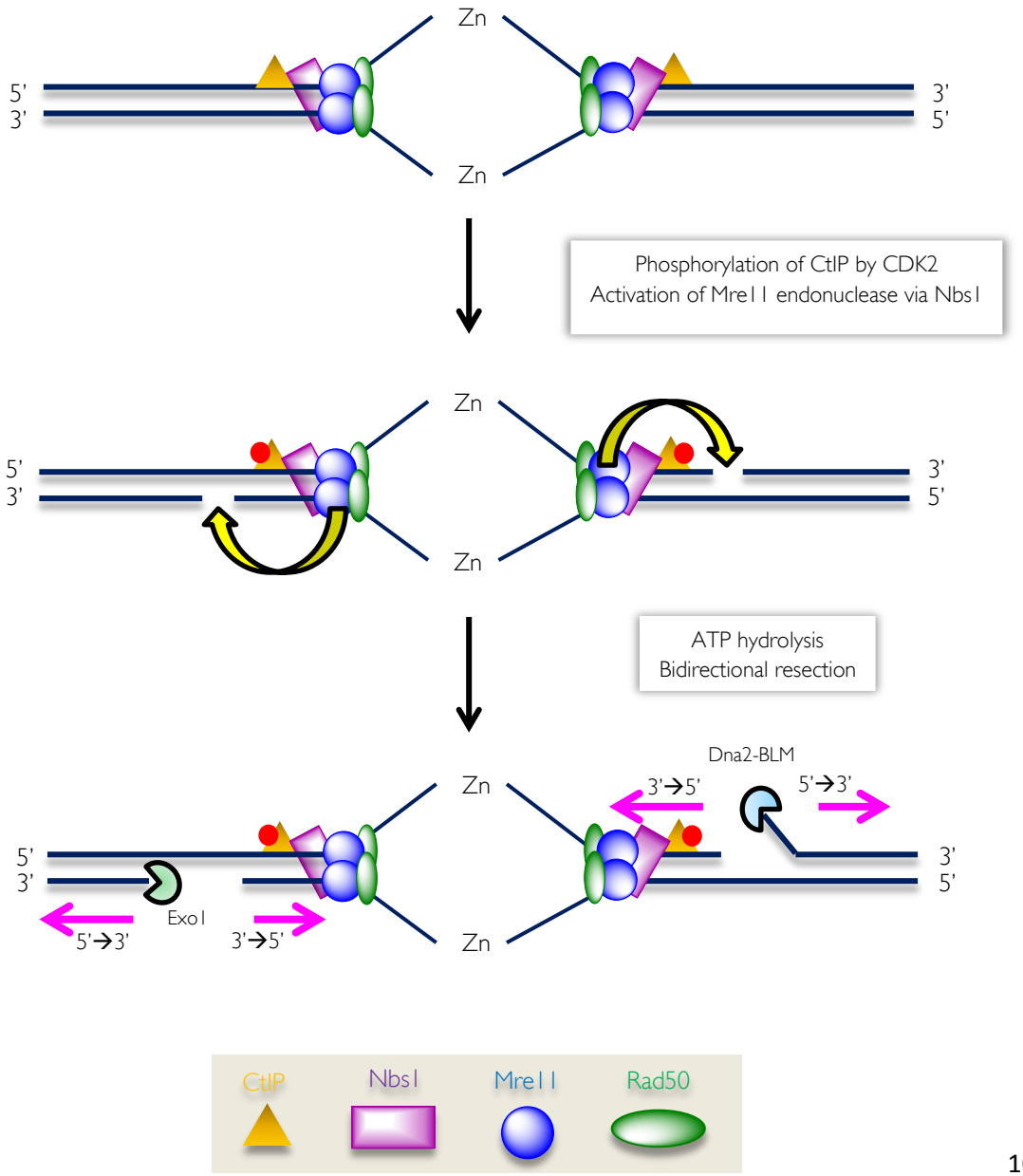


Figure 1.40 | **Model for eukaryotic (mammalian) DSB resection.** The MRN complex and CtIP are localised to double stranded DNA ends generated via a DSB. In the ATP bound state, Rad50 blocks the exonuclease function of Mre11. Phosphorylation of CtIP by CDK2 as a result of the DNA damage pathway mediated by ATM/ATR leads to activation of the endonuclease function of Mre11 which makes cuts in the DNA upstream of the break. At the same time, Rad50 hydrolyses ATP which allows Mre11 to act as a 3'→5' endonuclease, thus degrading the DNA oligonucleotide from the endonuclease-mediated incision, to the DSB, helping to generate a 3' ssDNA tail. Further to this, 5'→3' dsDNA exonucleases such as Exo1 can be recruited to resect DNA in the opposite direction, extending the length of the DNA tail. Dna2, a ssDNA nuclease, can also perform this function, however it needs a RecQ helicase such as BLM or WRN to release ssDNA. Figure adapted from (Gobbini *et al.*, 2016).

It has been shown that Nbs1 is involved in activating Mre11 endonuclease activity, along with CtIP, to boost the MRN complex activity. Nbs1 functions as a sensor of CtIP phosphorylation, and in response, can activate Mre11 endonuclease functions via physical interaction. Acting as a sensor of CtIP phosphorylation serves to also restrict resection to the latter stages of the cell cycle, when CtIP is phosphorylated (Anand *et al.*, 2019). Furthermore it was shown that Nbs1 can inhibit the 3'→5' exonuclease activity of Mre11 on ssDNA ends (Deshpande *et al.*, 2016). Overall, the exact nature of Mre11 nuclease regulation is unclear due to the *in vitro* nature of various data, dependant on exact experimental conditions and further analysis is required for a more clear-cut sequence of events. In meiosis, DSBs are generated deliberately via the action of the topoisomerase-like activity of Spo11. Once Spo11 cuts the DNA at a recombination hotspot, two Spo11 monomers remain attached to the DNA via a 5' Tyrosine linkage. In this case, MRN is recruited and nicks the strand associated with Spo11 using the endonuclease function of Mre11 (Liu, Wu and Lichten, 1995, Keeney, Giroux and Kleckner, 1997, Kumar and De Massy, 2010). Switching to the exonuclease function, Mre11 then degrades in a 3'→5' direction to generate a ssDNA tail along with the exonuclease function of proteins such as Exo1.

In either case, the RecA homologue, Rad51 is loaded onto the ssDNA tail, displacing the SSB homologue, RPA, with the assistance of recombination mediators such as BRCA2, Rad52 and Rad54. In the presence of ATP, Rad51 nucleates the ssDNA filament, forming a nucleoprotein filament and is loaded by Rad52 (and BRCA2 in humans), which also acts to displace RPA, similar to RecFOR in bacterial systems (Thorslund and West, 2007, Hilario *et al.*, 2009, Mazin *et al.*, 2010, Holloman, 2011). Dmc1 is the meiosis-specific recombinase used instead of Rad51 (Sehorn *et al.*, 2004).

Pre-synapsis in Archaea

The eukaryotic MRN-dependant homologous recombination system is generally conserved within Archaea including proteins such as Mre11, Rad50 and RadA (homologue of RecA, Rad51) whereas archaeal-specific factors include NurA (a 5'→3' nuclease) and HerA (a hexameric helicase) which are often found in an operon with Mre11/Rad50 (Hopkins and Paull, 2008).

The first step, pre-synapsis, is mediated by these proteins. Initial resectioning is catalysed by Rad50-Mre11-HerA-NurA complex. Mre11-Rad50 binds to the DSB and has been shown, in *Pyrococcus. furiosus*, to generate a short 3' overhang via degredation of the 5' terminal strand for a short time. HerA-NurA then initiate 5' strand resection (Hopkins and Paull, 2008). Mre11-Rad50 has also been shown to be present in *H. volcanii*, however Mre11-Rad50 appears to restrain homologous recombination in this case, suggesting a regulatory mechanism to prevent undesired recombinational events – a good strategy for polyploid organisms which contain many sites for strand exchange around their genomes (Delmas *et al.*, 2009). One end resection has taken place, the SSB/RPA homologue binds to the 3' ssDNA(Madru *et al.*, 2023). Displacement of such proteins and loading of the RecA/Rad51 recombinase homologue, RadA is then carried out. In *H. volcanii* RadB, a paralogue of RadA, has been reported to act as a recombinase mediator, assisting RadA in strand exchange and leading to a conformational change

promoting its polymerisation on DNA (Wardell *et al.*, 2017). RadA is proposed to polymerise via a ball and socket mechanism by inserting Phe residues into hydrophobic RadA binding pockets (Figure 1.41).

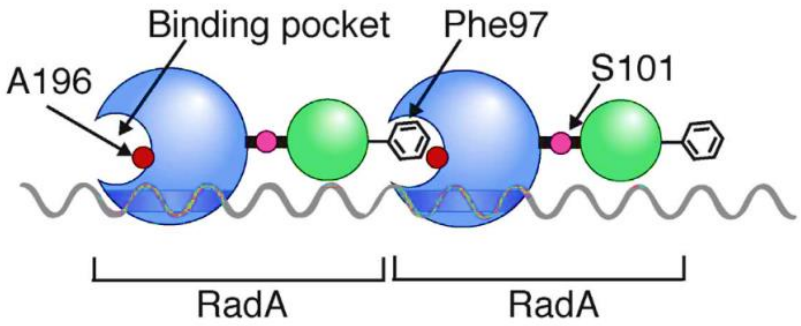


Figure 1.41 | **Schematic of RadA polymerisation via a ball and socket mechanism, using Phe residues.** Figure taken from (Wardell *et al.*, 2017).

1.6.3.2 Synapsis

Once the recombination mediator (RecA, Rad51, RadA) is loaded onto the ssDNA 3' tail via the mechanisms mentioned in the previous section, the core reaction, known as synapsis, occurs resulting in strand exchange, homology searching and displacement of the D-loop. In bacteria this is mediated by RecA, Rad51 (DMC1 in meiosis) in eukaryotes and RadA in Archaea. The recombinase, once it is polymerised and created a nucleofilament along the ssDNA stretches the DNA strand 1.5x relative to B form DNA, with each monomer binding 3 nucleotides, allowing a greater area to facilitate the search for a homologous duplex (Chen, Yang and Pavletich, 2008). For the purpose of simplicity, this section will focus on the RecA mediator. Homology searching is the first step in the process and is the most enigmatic. A recA nucleoprotein filament needs to search the genome rapidly to find a

sequence of homology. The multivalent nature of the RecA filament allows sustained contact with long dsDNA, while also allowing for frequent dissociation events, permitting rapid simultaneous searching of different parts of the genome, thus facilitating homology searching (Forget and Kowalczykowski, 2012). Furthermore, RecA filament sliding has been proposed as another mechanism by which RecA searches out homology via diffusing along a dsDNA track, it is estimated that upon encounter of a RecA filament, dsDNA would diffuse over 60-300bp prior to dissociation (Ragunathan, Liu and Ha, 2012). In spite of this, the degree of sliding was not significant and the observed sliding distances were equal to that of a RecA filament. It has also been suggested that it is the DSB event that triggers the broken chromosome to physically migrate to begin the homology search, exploring approximately 30% of the nuclear volume (Miné-Hattab and Rothstein, 2012). This seemingly 'active' activity contradicts the previous idea that homology pairing is an innate process, dependant on genome architecture, possibly occurring before the DSB is made (Barzel and Kupiec, 2008). Although this theory would hold whilst in G2 and a sister chromatid is physically tethered together, this does not seem to hold for ectopic recombination, unless the chromosomes are somehow aligned via an unknown mechanism. Although in a contradictory fashion, it has been observed that, in meiosis, there is a bias favouring allelic over ectopic recombination (Barzel and Kupiec, 2008).

In addition, homology searching occurs within the context of chromatin, and thus chromatin remodelling factors could potentially assist with this process by sliding nucleosomes along DNA. It has been shown in *Saccharomyces cerevisiae* that the Rad54, a SWI2/SNF2-related recombination mediator, interacts with Rad51 *in vitro* and can catalyse bidirectional nucleosome rearrangement, which is further stimulated by Rad51 (Alexeev, Mazin and Kowalczykowski, 2003). Rad54 has also been shown to act as a molecular motor, via ATP hydrolysis, to drive translocation of the pre-synaptic complex along the donor dsDNA and facilitate homology searching along with RPA,

which itself acts to allow simultaneous surveillance of both strands of donor DNA. An attractive mechanism, although it should be noted that Rad54 is absent from bacteria (Figure 1.42). In contrast, bacteria appear to conduct homology searching via a more passive mechanism. The emergence of an ATP-dependant mechanism in eukaryotes to search for homology may have resulted in order to adapt and facilitate unique challenges, such as longer chromosomes and chromatin. In eukaryotes, however, Rad54 has been assigned as a key player in remodelling the donor DNA for Rad51/RPA mediated homology search (Crickard *et al.*, 2020).

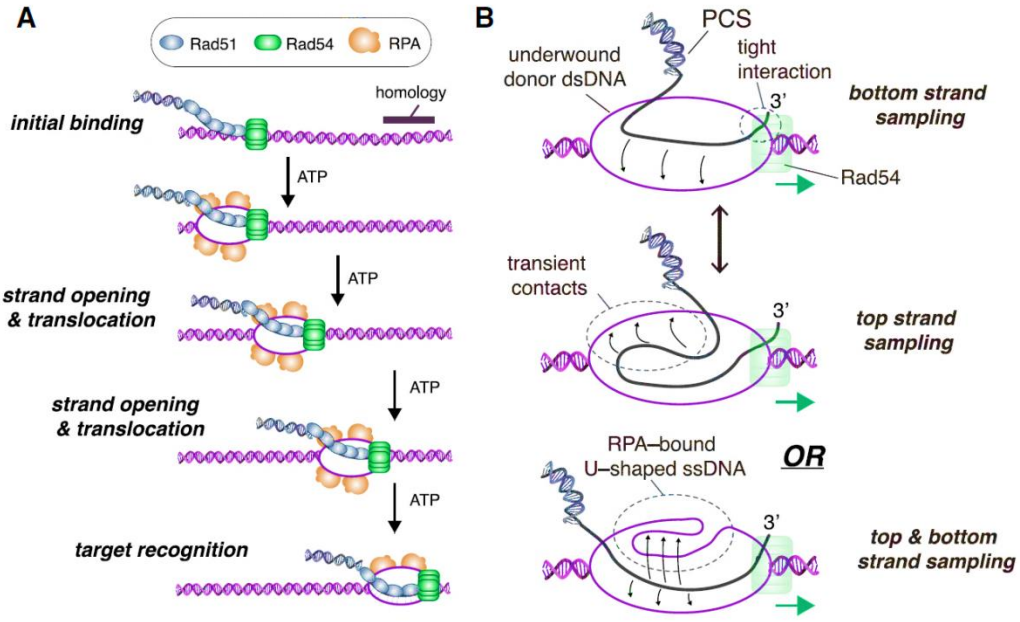


Figure 1.42 | **Model showing influence of Rad54 and RPA on homology searching.** (A) Depiction of the pre-synaptic complex (PCS) linked to donor dsDNA via the binding activity of Rad54, where ATP-dependant forward progression of the complex is coupled to deformation of the DNA duplex, enabling RPA association, promoting bivalent homology searching of the top and bottom strands. (B) Model depicting rapid sampling of donor DNA by Rad51 ssDNA within the translocating pre-synaptic complex. Rad51 is not depicted for clarity and Rad54 is shown bound at or near the 3' end of the pre-synaptic complex ssDNA. Figure adapted from (Crickard *et al.*, 2020).

The details of homology searching remain debated, however the end product is homology-mediated base pairing via strand invasion and the formation of the D-loop (Figure 1.43) catalysed by the recombinase, and eventually a double Holliday junction via second-end capture by Rad52, polymerase-mediated fill and ligation (McIlwraith and West, 2008). Although the process described above covers bacterial and eukaryotic systems, the archaeal recombinase, RadA also initiates strand exchange (Kil *et al.*, 2000).

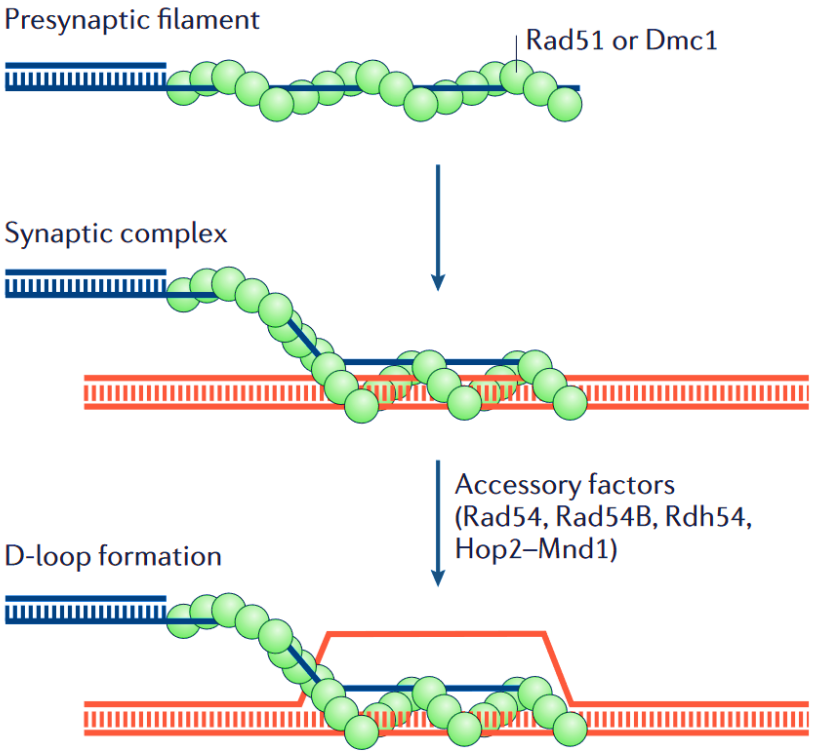


Figure 1.43 | **Recombinase filament homology search and D-loop formation in eukaryotes.** Rad51, or Dmc1, assemble onto the ssDNA tails to form a nucleoprotein filament, displacing RPA (not shown). Homology searching then occurs using donor dsDNA and a D-loop is formed. The mechanism by which Hop2-Mnd1 assist in the process is unknown. Figure adapted from (Sung and Klein, 2006)

1.6.3.3 Post-synapsis

Many mechanisms of Holliday junction resolution are present between species. Resolution of homologous recombination generates either a cross-over product, or a gene conversion product.

Synthesis-dependant strand annealing (SDSA)

Synthesis dependant strand annealing (SDSA, Figure 1.38) offers an alternative processing capacity in eukaryotes before a Holliday junction is formed, to allow a non-crossover product to be formed. Contrary to the idea that cross-over and non-crossover products are generated in competition with each other, it was shown that the SDSA pathway acts earlier in meiosis and crossover products are generated later, and in fact non-crossover products are mostly formed from SDSA, without the need for a Holliday junction (Allers and Lichten, 2001).

This process can occur in G1, whereas via DNA damage detection mechanisms, HR is restricted from occurring in G1. This process can also be useful in resolving an inappropriate recombination event, reducing genome rearrangements. During this process, helicases act to reverse the extended D-loop before a Holliday junction occurs. The nascent DNA strand anneals to the second end of the DSB. The process occurs via D-loop disruption by RecQ helicases such as BLM or WRN (Chu and Hickson, 2009). In *S. cerevisiae*, the Srs2 helicase, as well as Sgs1, has been shown to disrupt D-loops so could also function in this pathway in yeast (J. Liu *et al.*, 2017).

Single-strand annealing (SSA)

In eukaryotes, Rad52 also acts, aside from loading Rad51, second-end capture and progression through to a Holliday junction, in a single-strand annealing pathway – another alternative to recombinational repair (Figure 1.44). If a DSB occurs between two repeated sequences, resection exposes the repeated elements, which can then anneal, mediated by Rad52, leaving a flap structure either side of the DNA helix. The flaps are subsequently removed by endonuclease cleavage (Rad1/Rad10 in *S. cerevisiae* or XPF/ERCC1 in humans (Al-Minawi, Saleh-Gohari and Helleday, 2008)). This, however, is a mutagenic process although it does not require Rad51 (Lyndaker and Alani, 2009).

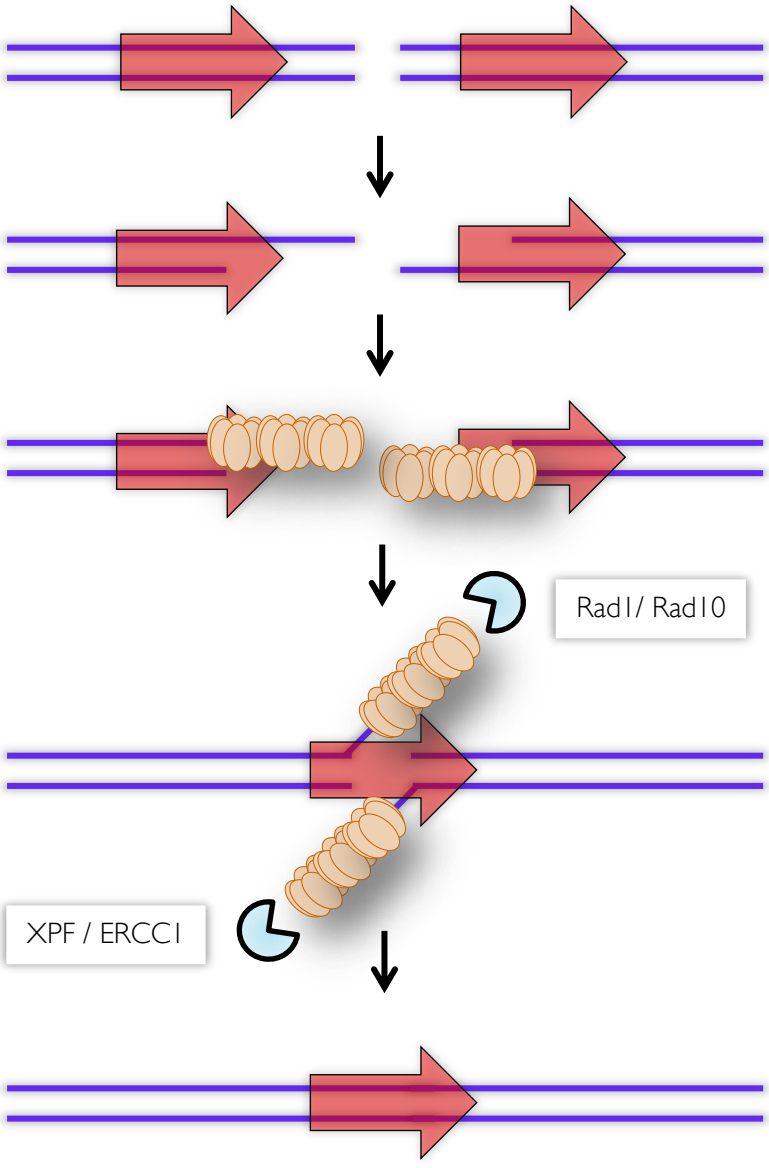


Figure 1.44 | **SSA pathway.** At a DSB, Rad52, once the 5' ends are resected, can load onto the 3' ssDNA tails and facilitate annealing of repeated DNA. Subsequent end processing via endonucleases and ligation allows repair of the DSB, however the intervening sequence is lost. Figure adapted from (Lyndaker and Alani, 2009).

Resolution of the Holliday junction

Holliday junctions are covalently linked complexes and involve two dsDNA molecules, as a result endonucleases, also known as resolvases, are required to resolve the structure. The specifics differ between organisms, however the principle is the same. If cleavage is mediated in a same sense fashion at both Holliday junctions, this results in a non-crossover product. Alternatively, if cleavage occurs in an opposite sense at each Holliday junction, a crossover product will arise (Figure 1.45).

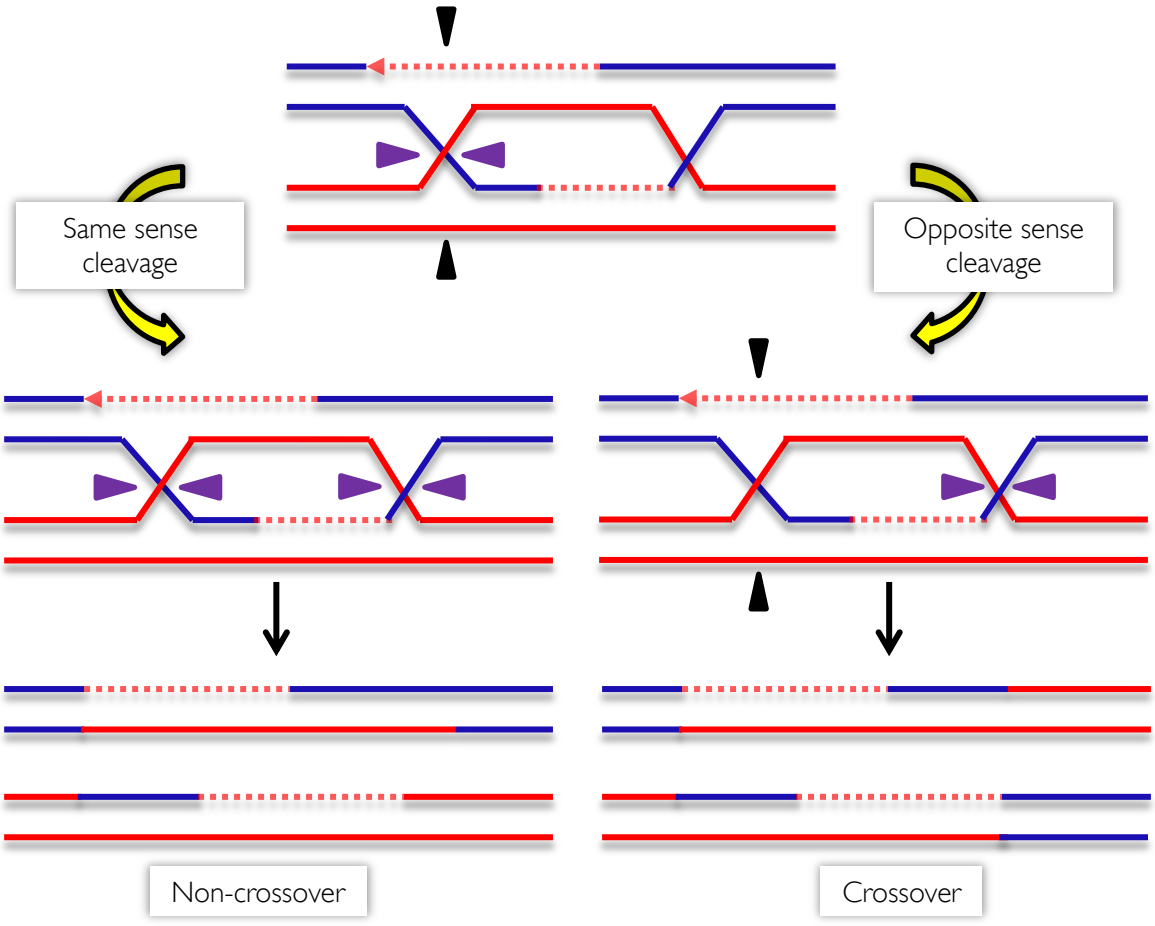


Figure 1.45 | **Double-Holliday junction resolution by endonucleases.** Same sense cleavage by endonucleases (resolvases) results in a non-crossover product. Opposite sense cleavage results in a crossover.

In *E. coli* Holliday junctions are resolved by the RuvABC complex (Sharples *et al.*, 1990). RuvA is a tetrameric DNA-binding protein that binds to the Holliday junction independent of DNA sequence, using two helix-hairpin-helix motifs, in each subunit, to contact the phosphodiester backbone (Rafferty *et al.*, 1998). Exhibiting a fourfold symmetry, RuvA has been shown to resemble a four-petal flower (Rafferty *et al.*, 1996) and, along with RuvB, facilitates branch migration activity and thus, strand-exchange (Iwasaki *et al.*, 1992). Maintenance of the four-way junction is facilitated via a central acidic pin (Glu^{55}, Asp^{56}) within the RuvA protein, mediating phosphate backbone repulsion of the four DNA strands involved (Rafferty *et al.*, 1996, Ingleston, Sharples and Lloyd, 2000). RuvB, a hexameric ATPase, exhibits DNA binding and helicase activity. RuvB binds either side of the DNA junction via contacts with RuvA, ATPase activity is enhanced by this binding. RuvB then catalyses branch migration using its helicase motor to pull the DNA strands through the complex, towards the RuvA-mediated junction. DNA is pulled from both sides towards RuvA, due to the opposing polarities of each RuvB complex either side of the central RuvA, and the strands are pushed out of the complex (Parsons *et al.*, 1995) (Figure 1.46).

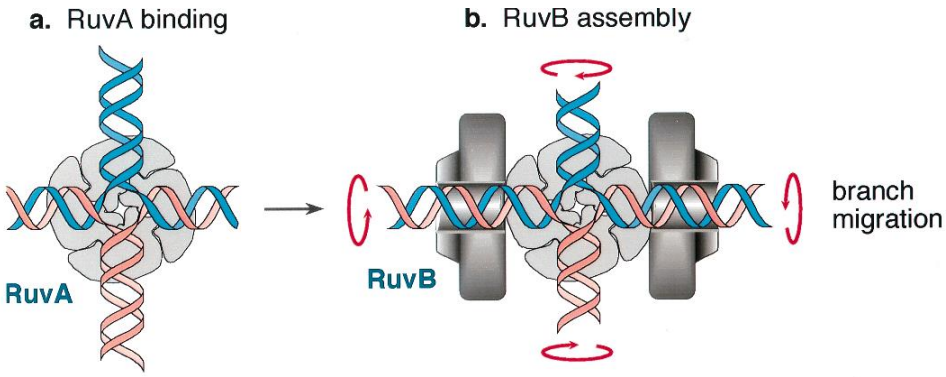


Figure 1.46 | **Model for action of RuvAB.** (A) RuvA tetramer binds to the Holliday junction and uses its acidic residues (not shown) in each monomer to facilitate strand separation. (B) RuvB binds to DNA via interactions with RuvA on either side of the junction. ATP hydrolysis is upregulated when contact is made with RuvA and subsequently, RuvB drives strand exchange by pulling the DNA strand through the complex. Figure adapted from (Whitby *et al.*, 1996).

The homodimeric RuvC endonuclease is then recruited and makes two symmetric 5'-phosphorylated nicks on two strands, near the centre of the Holliday junction, at the consensus 5'-$_{A/T}$TT↓$_{G/C}$-3' which it scans for during the branch migration process (Figure 1.47).

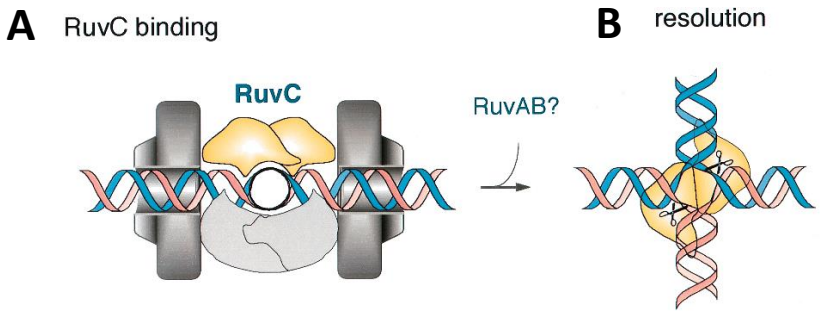


Figure 1.47 | **Model for RuvC binding and resolution of the Holliday junction.** (A) RuvC binds to the junction and (B) initiates single-strand cleavage at a consensus sequence. Figure taken from (Whitby *et al.*, 1996).

Such consensus sequences allow for productive bi-strand cleavage to occur exclusively in homologous regions and its proposed that the initial cleavage of one strand greatly accelerates cleavage of the second (Fogg and Lilley, 2000). Furthermore, increased resolution of structural studies has shown that a conformational change must occur within the DNA before RuvC can carry out the cleavage reaction. Binding of RuvC induces a distortion of base pair geometry around the exchange point and places the scissile phosphates too far from the RuvC active sites. Tension is relieved by RuvC by flipping out an adenine base 5' of the phosphate, allowing RuvC to cut, thus re-defining the model of RuvC cleavage (Figure 1.48).

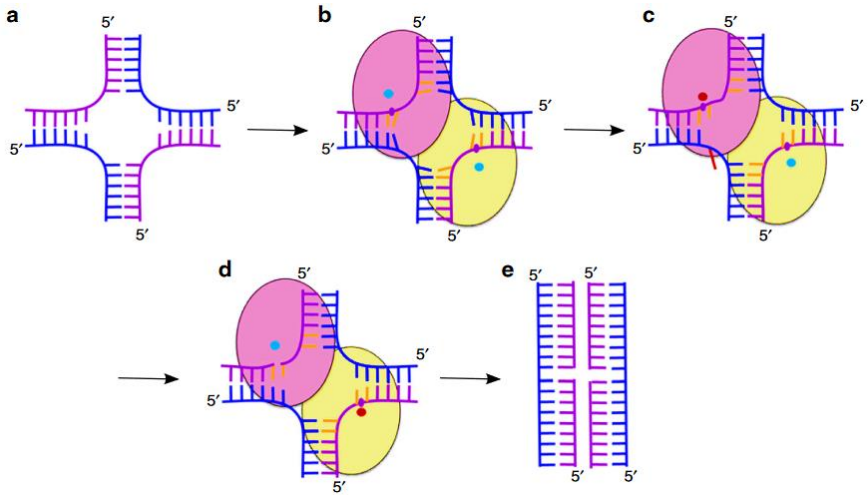


Figure 1.48 | **Mechanism of Holliday junction resolution by RuvC.** (A) Holliday junction. Cleaved and non-cleaved strands are shown in purple and blue respectively. (B) Binding of the HJ DNA. The subunits of the dimer are shown as yellow and pink ovals. The scissile phosphate is marked as a purple circle. Cyan circles show active sites. (C) Flipping of the adenine (red) opposite the scissile base. The active site is shown, active, as a red circle. (D) Second cut. (E) Resolution products. Figure taken from (Górecka *et al.*, 2019).

RuvC is not well conserved in bacteria, compared to RuvAB, however, bacteria lacking RuvC may resolve Holliday junctions using RusA instead (Sharples *et al.*, 1994) .

RecG, a RecQ-type helicase also possesses branch migration activity and has been shown to unwind Holliday junctions, as well as facilitating SDSA by dissociating D-loops and acting to process stalled replication forks. RecG, however, acts in a different way to RuvAB and is not a resolvase (Lloyd and Sharples, 1993). RecG has also been shown to direct DNA synthesis via interaction with PriA during recombination, allowing maintenance of genomic stability and inhibition of abnormal DNA synthesis at the break site (Azeroglu *et al.*, 2016).

Eukaryotic Holliday junction resolution is a far less trivial matter, and far more complex than the bacterial RuvABC pathway. Multiple, semi-redundant, pathways are present in eukaryotes and usage depends on the individual organism. Many proposed exonucleases exist that could resolve Holliday junctions and include MUS81-Mms4 (MUS81-EME1 (an XPF family endonuclease), however this complex cuts asymmetrically, unlike RuvC, suggesting a Holliday junction precursor (a nicked HJ) might be its preferred substrate. However, symmetrical endonucleases have been observed and include YEN1 (GEN1), and a ssDNA nuclease Slx1-Slx4 (SLX1-SLX4). The role of Sgs1/MutLγ/Exo1, especially their cleavage activity, in budding yeast is not clear (Schwartz and Heyer, 2011). Sequential cleavage by multiple endonucleases results in resolution of the Holliday junction via a complex, largely unknown process. It has been proposed that Mus81:Eme1 functions in meiosis first and Gen1 is activated as a chromosomal segregation safeguard nuclease if needed. SLX1-4 is thought not to be a key player in meiosis and is primarily concerned with the repair of interstrand crosslinks (with Mus81), although can function with Mus81:Eme1 at the G2/M transition to help resolve a junction (Figure 6.1) (Wyatt *et al.*, 2013), and can, in theory, step in during meiosis if Mus81 and GEN1 fail. SLX1-4 therefore can act as a backup function, however this is not its primary role. Both Mus81:Eme1 and GEN1 have been shown to function in mitotic cells as well as meiosis although their activity is downregulated until G2/M and thus the major product outside of M

phase is a non-crossover product via SDSA or Dissolution and suggests that crossovers are unwanted during DNA synthesis (Matos *et al.*, 2011).

The three major nucleases, SLX4 and Mus81-EME1 have been implicated in forming a tri-nuclease structure during prometaphase of the cell cycle and act to resolve branched and flap structures, predominantly as a result of replication and recombination intermediates (Wyatt *et al.*, 2017).

Double Holliday junction repair in meiosis, for unresolved reasons, is biased toward the formation of crossovers. Although MUS81-EME1 appears to be involved in resolving the Holliday junction, this represents a minor form of a resolution pathway, termed a Class 2 crossover pathway. The major crossover mechanism appears to be the Class 1 crossover pathway, involving MMR machinery, MutSγ (Msh4/Msh5) and MutLγ (Mlh1/Mlh3) along with Exo1 (Gray and Cohen, 2016). MutSγ has been shown in humans to form a sliding clamp capable of encircling a dHJ substrate, possessing the ability to translocate along the DNA, allowing recruitment of more MutSγ clamps at the DSB repair site (Snowden *et al.*, 2004). Resolution via MutLγ results in only crossover products and represents the major production of crossovers in meiosis, the mechanistic details are not clear however it appears that the bias involves ZMM proteins (Zakharyevich *et al.*, 2012).

In Archaea, the Holliday junction can be resolved by the endonuclease, Hjc, which has been shown to have analogous properties to the *E. coli* RuvC protein. Hjc can cleave in a symmetrical fashion, however does not show any strand specificity, unlike RuvC and RusA. Hjc from the thermophillic archaeon *Methanobacterium thermautotrophicum* has been shown to promote DNA repair in a resolvase-deficient (*ruv*) mutant of *E. coli* and depends on RecG for function (Bolt, Lloyd and Sharples, 2001). Furthermore, *Sulfolobus solfataricus* has been shown to contain a Hje protein, alongside Hjc, that is able to resolve a junction via introducing a pair of nicks within the 4-way junction, indicating redundancy of two separate mechanisms (Kvaratskhelia, Wardleworth and White, 2001). *H. volcanii* possesses both Hjc and a MUS81/XPF homologue,

Hef which functions redundantly with Hjc. Deletion of Hef in a Δ*radA* background is highly toxic, yet deletion of Hjc in the same background has no effect, suggesting Hjc acts exclusively in homologous recombination and Hef can act in alternative, non-homologous recombination dependant pathways (Lestini, Duan and Allers, 2010)

Asymmetric cleavage of the Holliday junction

In eukaryotes SLX1-SLX4 and MUS81:EME1 can also act on the branched structure to generate an asymmetric nick, subsequent rotation and a second cleavage reaction can occur, leading to disassembly of the Holliday junction, although not via a RuvC-type mechanism. Cleavage results in a crossover product with gaps that are filled by DNA polymerases and subsequently ligated. It is proposed that SLX1-SLX4 interacts with MUS81:EME1 to form a stable holoenzyme and that MUS81:EME1 resolves nicked, not intact, Holliday junctions. Therefore SLX1-SLX4 potentially introduces the initial, rate-limiting, nick in a single strand of DNA at the junction, MUS81:EME1 then acts to create a second dsDNA nick that resolves the junction (Svendsen *et al.*, 2009) (Figure 1.49).

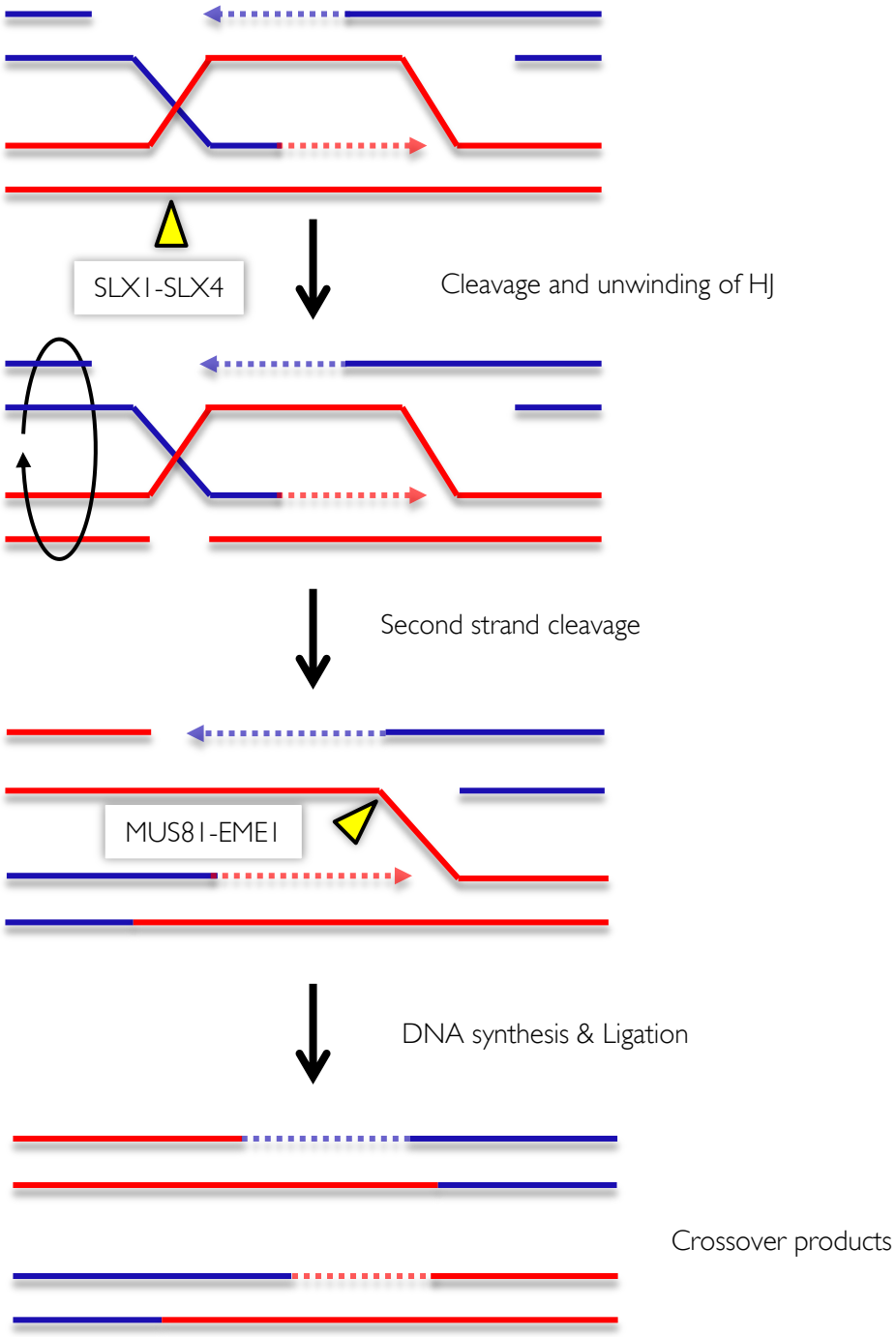


Figure 1.49 | **Mechanism of HJ resolution by co-ordinated action of SLX1-SLX4 and MUS81-EME1.** A crossover can result just after second-end capture by Rad52 if nucleases act on the structure before the dHJ is formed. SLX1/4 generates the first nick and, after unwinding, the second nick is generated by MUS81:EME1.

Holliday junction dissolution

In eukaryotes, outside of M phase, before the up regulation of Mus81:Eme1/GEN1, Holliday junctions can be processed via an alternative mechanism, that does not result in resolution by endonucleases and leads to a non-crossover product. Instead, the junction can be dealt with via a process called dissolution. In this process RecQ helicases are used to regress the branch point and the resultant DNA hemi-catenanes are processed in a decatenation reaction by topoisomerases. In addition to D-loop disrupting activity, RecQ helicases such as BML or WRN can act to push the branch point back towards the point of origin, a topoisomerase, such as TOPOIIIα can then catalyse a decatenation reaction in which the interlinked strands are uncoupled from each other, resulting in a non-crossover product. In addition to forming a non-crossover product, this provides a useful mechanism by which cells can rescue an inappropriate crossover event and re-direct the DNA down the SDSA pathway (Chu and Hickson, 2009)(Figure 1.50).

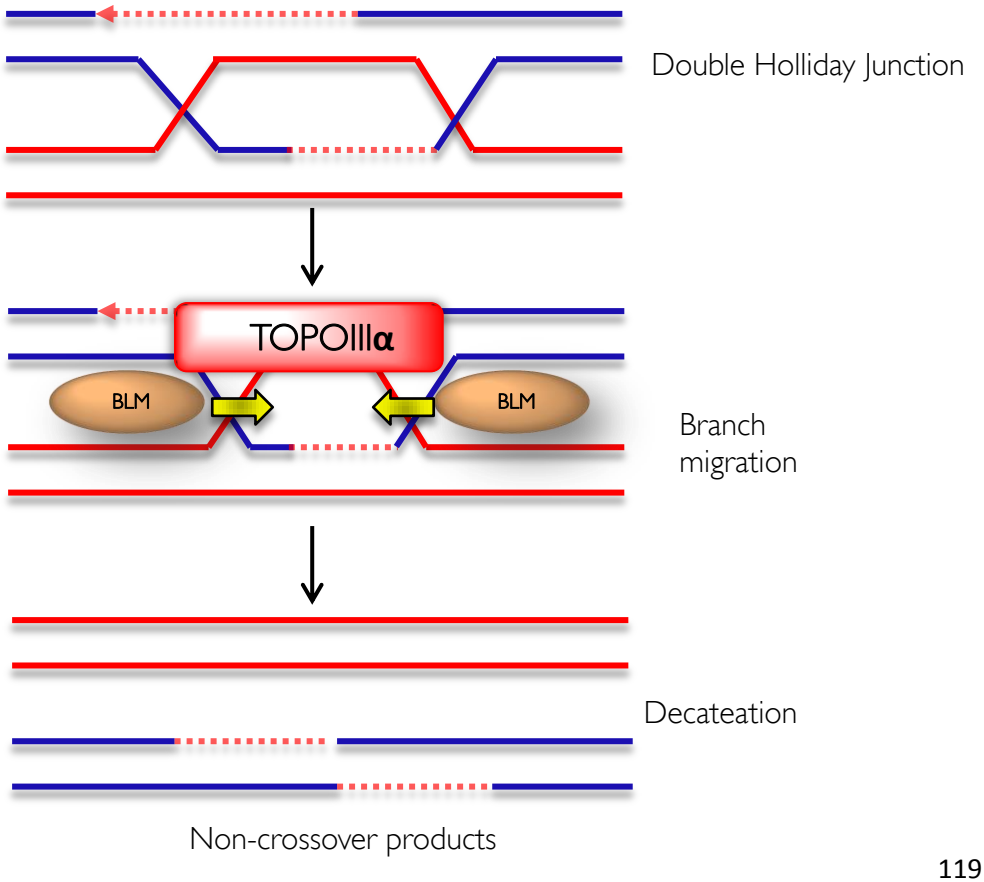


Figure 1.50 | **Dissolution of the Holliday Junction by BLM and TOPOIIIα.** Details in text. The RMI1-RMI2 complex (not shown) is involved in humans to stabilise the topoisomerase and favour decatenation by stabilising the BLM complex (Wang *et al.*, 2010). Figure adapted from (Chu and Hickson, 2009).

1.6.4 DNA replication

The fidelity of DNA replication is critical to reducing mutational load and fundamental to life across all domains. DNA replication is tightly regulated and must occur faithfully, prior to cell and nuclear division, to allow the genetic material to be passed on to the next generation. Three phases are found in all organisms: initiation, elongation and termination.

1.6.4.1 Replication initiation

DNA replication is initiated at defined sites within the genome known as origins of replication. Bacterial circular chromosomes contain a single origin from which replication proceeds, often in a concurrent fashion, with replication re-initiating before the previous strand has been fully replicated. The *E. coli oriC* contains five 9mer repeat elements, known as DnaA boxes, within the DNA unwinding element (DUE). DnaA boxes are sequence specific elements, spread across 260bp, that are bound by the initiator protein, DnaA, an AAA+ family ATPase which forms a protein filament at the *oriC* consisting of approximately 20 monomers (Messer, 2002) (Ozaki and Katayama, 2009). Following assembly of DnaA-ATP at the *oriC*, DnaA induces strand melting of the AT-rich 13mer repeats within the DUE to the left of *oriC* (Bramhill and

Kornberg, 1988, Chodavarapu and Kaguni, 2016). It is the ATP bound form of DnaA that allows this reaction to occur, rather than ATP hydrolysis (Sekimizu, Bramhill and Kornberg, 1987). Once the DNA has been melted with the local region of the DUE, this facilitates entry of the replicative helicase, DnaB and, via the AAA+ family ATPase helicase chaperone loader, DnaC, is loaded onto the ssDNA. DnaC acts in a passive capacity to inhibit DnaB ATPase activity, and therefore must dissociate before DnaB can unwind the parental duplex DNA in a 5'→3' direction. It has been shown that DnaC acts as a molecular switch that, via an ATP/ADP dependant mechanism at first inhibits DnaB function when bound to ATP, and once bound to DNA, ATP hydrolysis by DnaB then inhibits DnaC and facilitates helicase action (Davey *et al.*, 2002) (Kaguni, 2011). SSB then binds the ssDNA, via an oligosaccharide binding fold (OB fold), preventing secondary structure formation and nuclease degradation (Shereda *et al.*, 2008).

Eukaryotic genomes have multiple origins of replication along each chromosome, a good mechanism for replication of long chromosomes (O'Donnell, Langston and Stillman, 2013). Initiation begins by binding of the origin recognition complex (ORC), a heterohexameric protein complex, to the origin of replication. The mechanism of recognition is elusive, however basic Lys and Arg residues within the budding yeast Orc1 subunit have been shown to be important (Kawakami *et al.*, 2015). ORC recruits the replication factor Cdc6, an AAA+ family ATPase, and Cdt1, a helicase loader and stabiliser for the eukaryotic helicase, MCM. Together ORC, Cdt1 and MCM form the pre-replication complex and licence the origin for replication with regulation of this process brought about by CDK action with the context of the cell cycle (Sun and Kong, 2010, Frigola *et al.*, 2017). The pre-initiation complex is formed with the binding and formation of the CMG complex, consisting of Cdc45 and GINS (Moyer, Lewis and Botchan, 2006, Ilves *et al.*, 2010). As the helicase complex moves along DNA, the SSB equivalent, RPA, binds ssDNA to protect it from nucleases and secondary structure formation.

With the exception of three methanogenic species, all archaeal genomes appear to contain at least one gene with homology to Orc1 and Cdc6, with many species containing multiple proteins (Barry and Bell, 2006). *H. volcanii* has sixteen *orc1/cdc6* genes (Ludt and Soppa, 2018). Homologues of MCM are thought to function as the replicative helicases in archaea and most archaeal genomes encode a single MCM homologue. *In vitro* experiments have shown that archaeal enzymes possess robust helicase activity in their own right, and do not require additional proteins such as the GINS complex in eukaryotes. In *H. volcanii*, MCM and GINS homologues are known, however Cdc45 appears to be absent, RecJ in several species is proposed to be the Cdc45 homologue (Nagata *et al.*, 2017) (Makarova, Koonin and Kelman, 2012). In archaea, little is known about the loading of MCM, and with no homologues of DnaC or Cdt1, this suggests MCM interacts directly with the Orc proteins at the origin (Barry and Bell, 2006).

Archaea contain a variety of SSB proteins, overall showing more similarity to the eukaryotic RPA. The best studied is the SSB in *S. solfataricus*, showing domain structure similar to bacterial SSB (Wadsworth and White, 2001). Despite this, the crystal structure revealed the OB fold is more similar to that of human RPA (Kerr *et al.*, 2003). *H. volcanii* contains three genes encoding homologues of RPA (rpa1, rpa2, rpa3). Recent studies have shown that RPA2 is essential (Stroud, Liddell and Allers, 2012).

1.6.4.2 Replication elongation

Priming

DNA polymerases require a primer from which to extend DNA synthesis as they are unable to bind ssDNA and require a 3' tail from which to initiate synthesis via an Sn2 nucleophilic attack mechanism. As a result of the polarity of the double helix, the two DNA strands must be synthesised in different ways, known as continuous (leading) and discontinuous (lagging) strand synthesis.

In bacteria, the DnaG primase is recruited to the replication fork via interaction with DnaB, where it synthesises a primer in a 3'→5' direction (Rowen and Kornberg, 1978) (Corn and Berger, 2006). Generally speaking, DnaG primers are 11 residues in length, although they can vary from between 10 and 60 nucleotides (Frick and Richardson, 2001)

The eukaryotic primase is a heterodimer of PriS (Pri1) and PriL (Pri2), which associate with DNA Polymerase α and the Polymerase B subunit to form the primosome. Primase synthesises a primer of 8-12 nucleotides in length, which are then elongated to around 30 nucleotides by DNA polymerase α, before handing over to the replicative polymerases (Vaithiyalingam *et al.*, 2014) .

Archaeal homologues of PriS and PriL have been found but lack Polα and Subunit B accessory factors. Priming occurs over approximately 7 – 14 base pairs before handover to a DNA polymerase. Owing to the lack of a Polα subunit, PriS has DNA polymerase activity and thus it is possible could function in a similar way. PriL is proposed to have a regulatory role (Barry and Bell, 2006).

DNA synthesis

In bacteria, the RNA primers are extended by the DNA polymerase III holoenzyme, comprising of two polymerase subunits in a complex with a further nine protein subunits (Kelman and O'Donnell, 1995). The core of the enzyme consists of an α and ε subunit which are responsible for the polymerase and 3'→5' exonuclease proofreading function respectively, along with a θ subunit that aids in exonuclease function. The core of the enzyme drives catalytic activity. Processivity of the polymerase requires the β clamp, which is loaded onto DNA by the five-subunit γ complex (Kelman and O'Donnell, 1995).

Eukaryotes have two replicative polymerases, termed DNA Polymerase ε, which extends primers on the leading strand, and DNA polymerase δ, which extends primers on the lagging strand (Miyabe, Kunkel and Carr, 2011). The proliferating cell nuclear antigen (PCNA) is a sliding clamp with homology to the bacterial β-sliding clamp and aids Pol δ/ε in carrying out processive DNA replication by tethering them to the replication fork as well as having a role in recruitment of translesion polymerases for repair purposes. PCNA is loaded by a clamp loader known as RFC (Moldovan, Pfander and Jentsch, 2007).

Euryarchaea possess two families of DNA polymerase, known as B and D family polymerases and the crenarchaea only have a B family helicase. The crenarchaea do not carry D family polymerases. *In vitro* work has shown that PolB is targeted to the lagging strand, and PolD to the leading strand (Greenough, Kelman and Gardner, 2015) contrary to the previous theory that PolD operated on both strands in *Thermococcus*.

In a similar fashion to eukaryotes, archaeal DNA polymerases are assisted in their processivity by the PCNA sliding clamp, loaded by RFC (Matsumiya, Ishino and Morikawa, 2001). A summary of the players in the replisome of all three domains are summarised in Figure 1.51 using a eukaryotic fork. Bacterial forks differ in that they are unwound by DnaB which encircles the lagging strand template, rather than the leading strand template (Xu and Dixon, 2018).

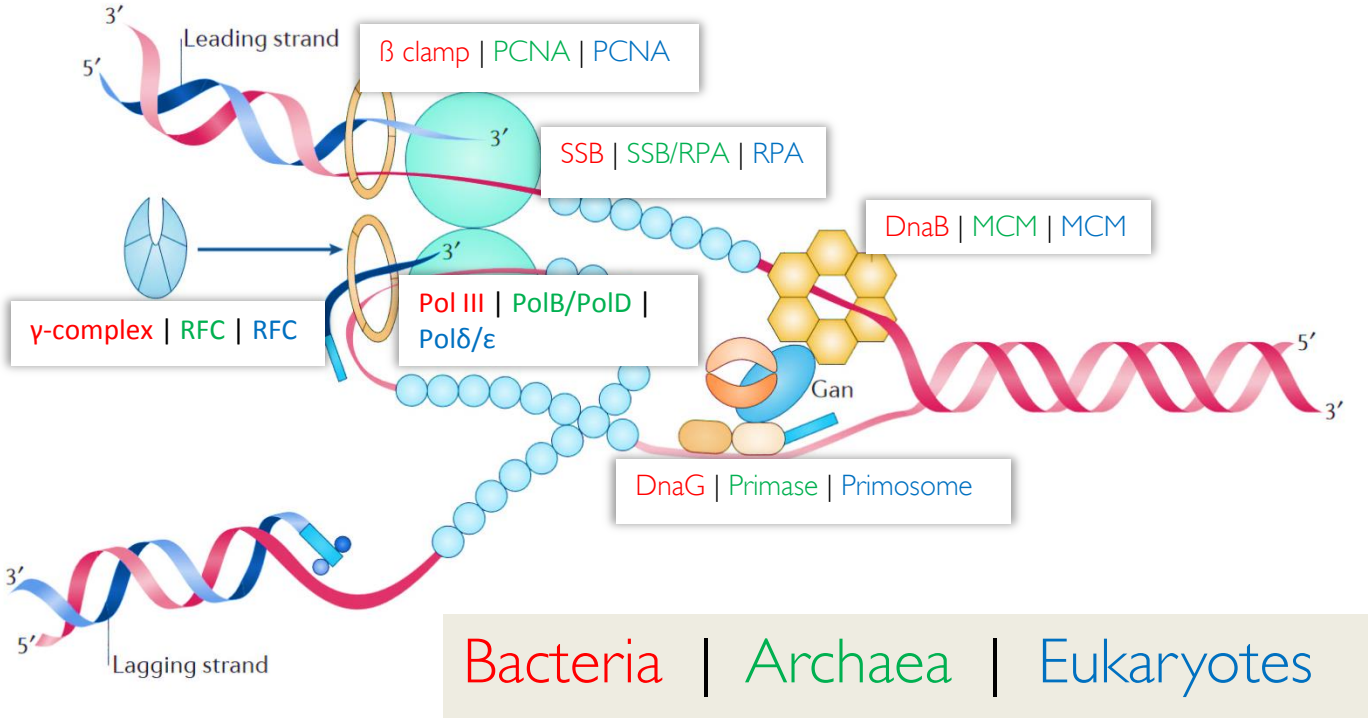


Figure 1.51 | **The replisome.** A eukaryotic replication fork indicating component parts from the three domains of life, bacteria, Archaea and eukaryotes. Note that DnaB in bacteria encircles the lagging strand template instead of the eukaryotic MCM complexes as shown above, which encircle the leading strand template. Figure adapted from (Lindås and Bernander, 2013).

Okazaki fragment maturation

In bacteria, primers are removed by the exonuclease function of DNA Polymerase I and replaced with canonical DNA synthesis and DNA Ligase can join the nick (Allen *et al.*, 2011). Eukaryotic systems are slightly more complex and various models exist. Primarily, using strand displacement synthesis from the upstream Okazaki fragment to displace the RNA primer, a flap-specific endonuclease (FEN1) can then degrade the flap structure, with DNA Ligase I sealing the nick. However, a long path mechanism has been proposed where cleavage first occurs via Dna2, with FEN1 then taking over. RNaseH has also been implicated in primer removal along with Exo1 (B. Liu *et al.*, 2017) (Zheng and Shen, 2011). *In vitro* work has instigated RNaseH and FEN1 in archaeal primer removal, where PolD (and PolB) facilitates displacement synthesis and flap generation, along with ligation via DNA Ligase 1 (LigA in *H. volcanii*) (Henneke, 2012) .

Termination

In bacteria, termination occurs at defined *Ter* sites, opposite to the *oriC*. Termination occurs when bidirectional replication forks converge. *Ter* sequences are in an opposite orientation to the approaching replication fork and form a replication block. Many *ter* sites exist and are named *TerA-J* (Duggin *et al.*, 2008). Such sites are active when bound to the companion termination protein, *Tus*, which interacts with DnaB and blocks its helicase action, ensuring termination occurs within the terminus region (Neylon *et al.*, 2005). Forks can also fuse together, which leads to termination.

In eukaryotes, the system is poorly characterised. It is thought that for termination to occur, replication forks must converge (owing to the lack of defined termination signals), and subsequent disassembly and decatenation by topoisomerases must occur. The removal of the CMG complex from chromatin is emerging as a key event in this process, although the details remain unknown (Dewar and Walter, 2017).

Archaeal termination is similar, in the sense that there are no defined termination signals within the DNA. The likely explanation is that termination occurs in a similar fashion to eukaryotic systems and is based on converging replication forks and disassembly of the replisome, followed by topoisomerase action. *H. volcanii* appears to have replication termination spread across broad zones (Hawkins *et al.*, 2013) .

1.6.4.4 XerCD Recombinase

During DNA replication, if unequal homologous recombination events occur, then the resultant chromosomes become linked at concatamers. This poses a problem for cell separation, as the linked chromosomes therefore need to be decatenated prior to cellular segregation.

DNA replication is not a continuous process, and is often interrupted and halted by various DNA lesions such as UV radiation, radical attack and other external genotoxic agents such as MMC. Homologous recombination represents a key method in which cells repair DNA damage that has occurred, and not been fixed whilst still classified as a lesion. If such a process is involved during DNA replication in organisms with circular chromosomes, and an unequal number of recombination events occurs, the two chromosomes in question will become linked together as a concatamer. This poses an issue for cell separation, as the linked chromosomes therefore need to be decatenated prior to cellular segregation. To ensure correct chromosomal disjunction, bacteria and Archaea have overcome this issue by resolving such dimer formation by using site-specific recombinases that act on such structures immediately prior to chromosomal segregation.

XerCD, in bacteria, act as site specific recombinases and deconcatentate chromosomal dimers formed via unequal recombination events during DNA replication, thus facilitating the termination stage (Castillo, Benmohamed and Szatmari, 2017, Farrokhi, Liu and Szatmari, 2019). XerC and XerD are tyrosine recombinases (Figure 1.52) that act specifically on a 28bp target *Dif* site (differential induced filamentation site). Such *Dif* sites are in close proximity to canonical termination sites (*ter* sites)(Graham et al., 2010)

XerCD complexes initiate the process by the activation of XerD by the gamma sub-domain of FtsK (Sherratt et al., 2010) (Grainge et al., 2013)(May et al., 2015), which translocates towards the XerCD/*dif* site to mediate initial strand exchange immediately prior to cell division, thus generating a transient Holliday junction

intermediate. XerC then mediates second strand cleavage and exchange to allow separation of the duplexes (Figure 1.52).

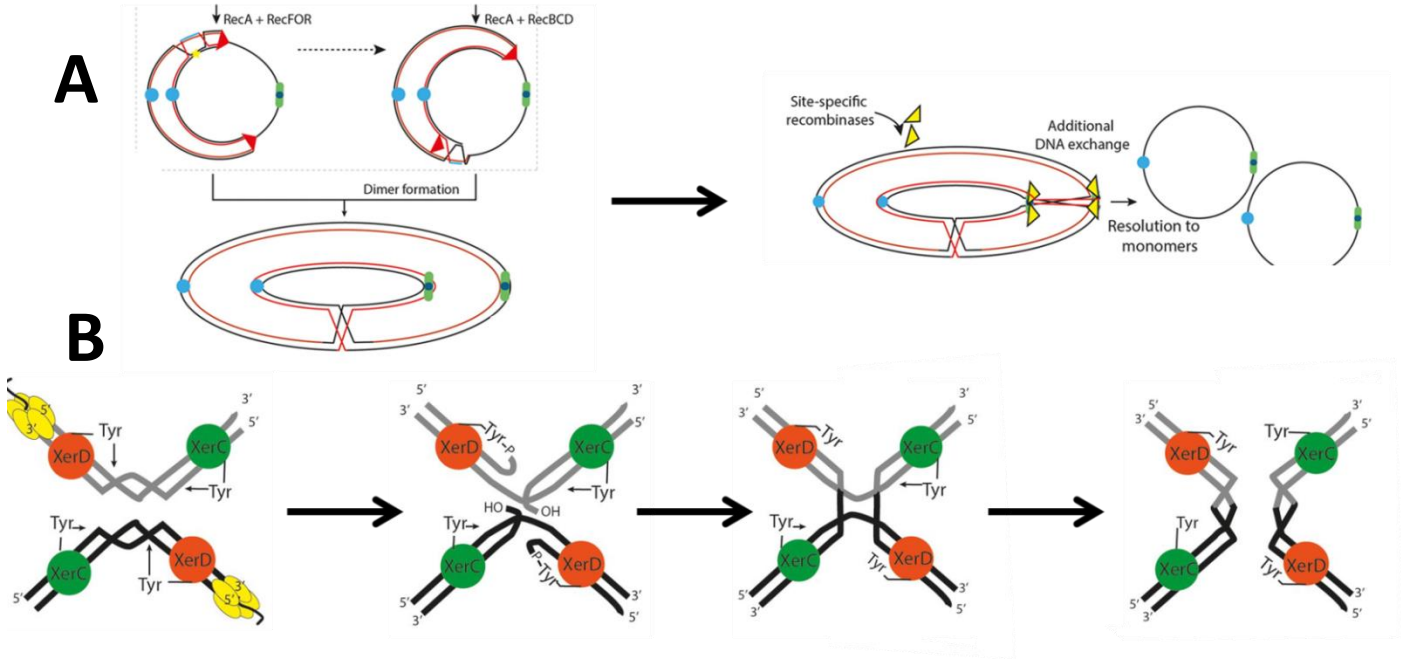


Figure 1.52 | **Mechanism of XerCD resolvase in *E. coli*.** A) Chromosomal dimers can form from unequal exchange events during double or single stranded break repair, using RecBCD or RecFOR respectively. XerCD (yellow triangles) is recruited to the terminus region (green circles) and acts on the *dif* site (blue dot). Origins of replication are shown as light blue dots. B) FtsK (yellow) translocates along the DNA duplex until XerD is reached. Upon activation XerD catalyses the first strand exchange, followed by XerC, resulting in resolution of the DNA dimer into separate DNA duplexes. Figures adapted from (Castillo, Benmohamed and Szatmari, 2017)

As such, XerCD represents a key step in bacteria for replication termination and correct chromosomal disjunction. However, the role of XerCD has yet to be confirmed at the replication fork, specifically in termination of the process.

Homologous sequences are found across the other domains of life, including archaea. In *H. volcanii* XerCD is located within the prophage region, but there are a total of twelve homologous seuqnces littered throughout the genome.

Interestingly, the process by which XerCD, in combination with FtsK, carries out unlinking of chromosomal dimers is via a stepwise unlinking model (Shimokawa et al., 2013) (Stolz et al., 2017) (figure 1.53).

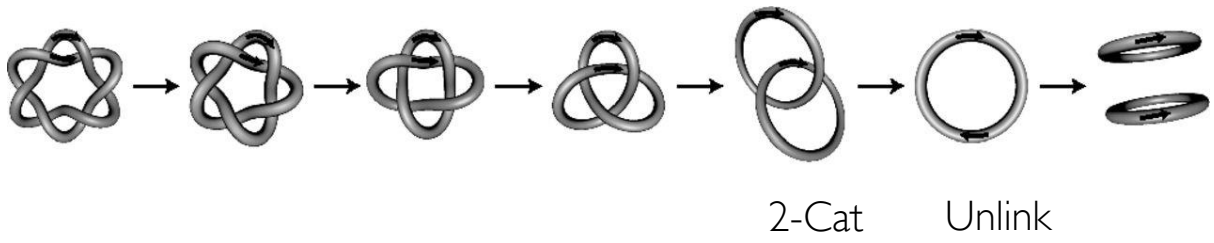


Figure 1.53 | **Proposed stepwise unlinking of chromosomal dimer by XerCD in combination with FtsK.** XerCD, when recruited to *dif* sites convert linked dimers to more simple three dimensional structures in a progressive manner. Eventually ending at a '2-cat' structure which is then unlinked to resolve the chromosomal dimer. Figure adapted from (Shimokawa *et al.*, 2013)

Furthermore, recent research has shown the involvement of XerCD, and its cognate *dif* sites in outermembrane vesicles. Specifically, OMVs are used to dispose of over-replicated DNA at the end of the cell-cycle. This was due to enrichment of DNA from the *dif* site within the vesicle lumen (Mansky *et al.*, 2023).

XerCD proteins are used in a variety of phages to allow integration into host chromosomes. A number of XerCD homologues are found within an integrated prophage region within the *H. volcanii* main chromosome (Figure 1.54).

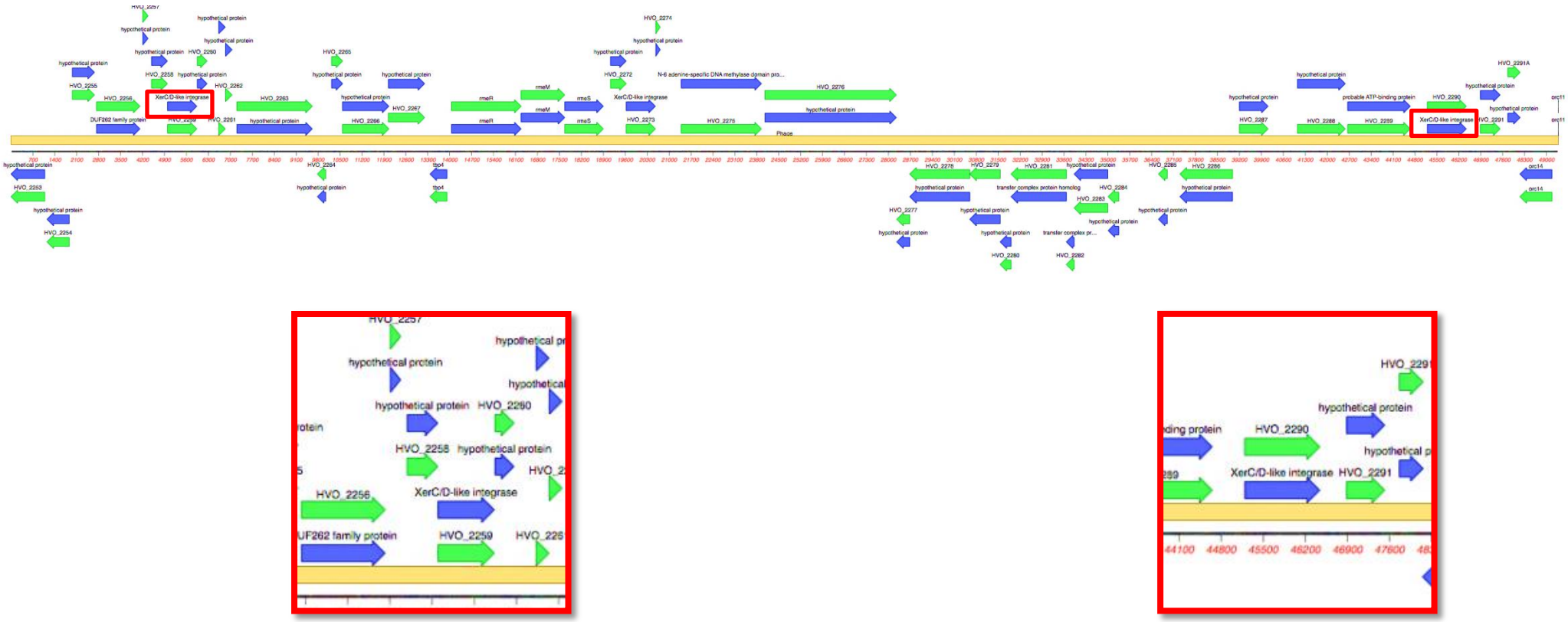
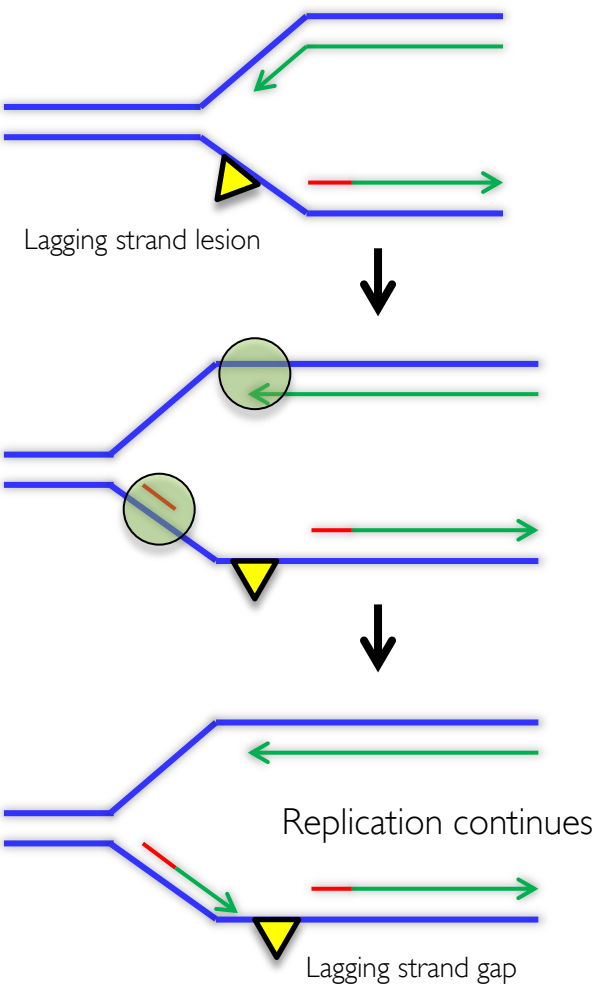
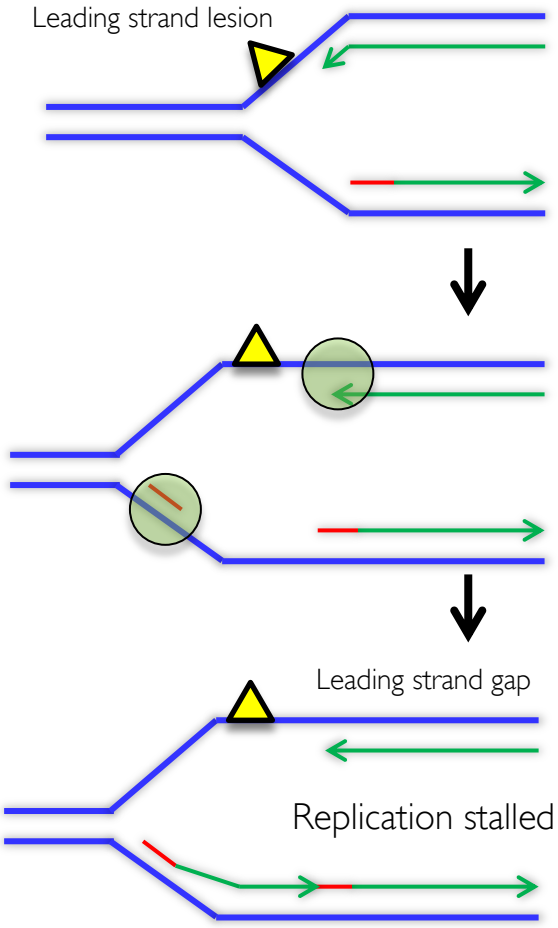


Figure 1.54 | **Overview of prophage region in *H. volcanii*.** XerCD homologues within this region are highlighted in red and shown in boxes.

1.6.4.3 Lesions in the DNA template

Non-coding lesions in the DNA template result in the replicative polymerase not being able to read the genetic code. If the lesion is in the lagging strand, owing to re-priming events downstream, replication can continue, although a gap will be left behind. A lesion in the leading strand is more problematic and the inability to re-prime DNA synthesis downstream of the lesion results in replication stopping completely, leading to a gap in the leading strand and a DSB. Furthermore an interstrand crosslink, similar to those formed from the mustard gases, also results in a stalled replication fork (Figure 1.55).


Lagging strand lesion
Replication continues
Lagging strand gap


Leading strand lesion
Leading strand gap
Replication stalled

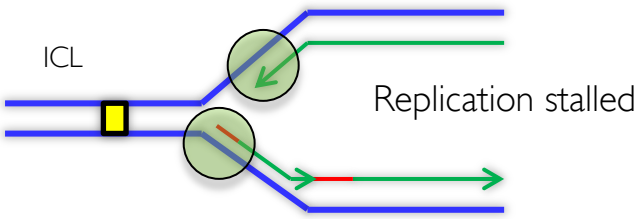

ICL
Replication stalled

Figure 1.55 | **Lesions in the DNA template.** Lagging strand lesions lead to a gap in the lagging strand, a trivial matter owing to the ability to re-prime easily. The lesion can be removed later. Leading strand lesions cause an issue as replication can not easily be re-primed and thus the replication fork stalls. Interstrand crosslinks also lead to a stalled fork.

1.6.4.4 Replication restart

Replication can be restarted, once the fork is stalled, via a number of means. The key mechanisms will be discussed in this section.

PriC-mediated leading strand restart

PriC, a replication restart mediator, is able to load the DNA replication machinery, including DnaG primase which can prime downstream synthesis on the leading strand. Although a gap will be left behind, DNA replication is able to continue. The gap left behind can be dealt with via homologous recombination once the lesion is removed via the appropriate machinery (Heller and Marians, 2006). In eukaryotes this method is not required as multiple origins of replication fire at the same time and can compensate due to the lack of termination sites.

Fork regression and HR-dependant restart (Break-induced replication)

Another method of dealing with a stalled replication fork is by regressing the fork into a Holliday junction. The resultant junction can be reset into a replication fork. The bacterial RecG helicase can regress the replication fork into a chicken foot structure, forming a Holliday junction along with RuvAB (Gupta, Yeeles and Marians, 2014). This allows endonucleases such as RuvC to access and cleave the junction. Subsequent processing by HR machinery such as RecBCD allows replication restart via PriA-dependant mechanisms once the D-loop is formed. PriA recruits the replicative helicase, DnaB, which begins to unwind the strands (Sandler and Marians, 2000). A second Holliday junction is formed which is dealt with by RuvABC (Figure 1.56).

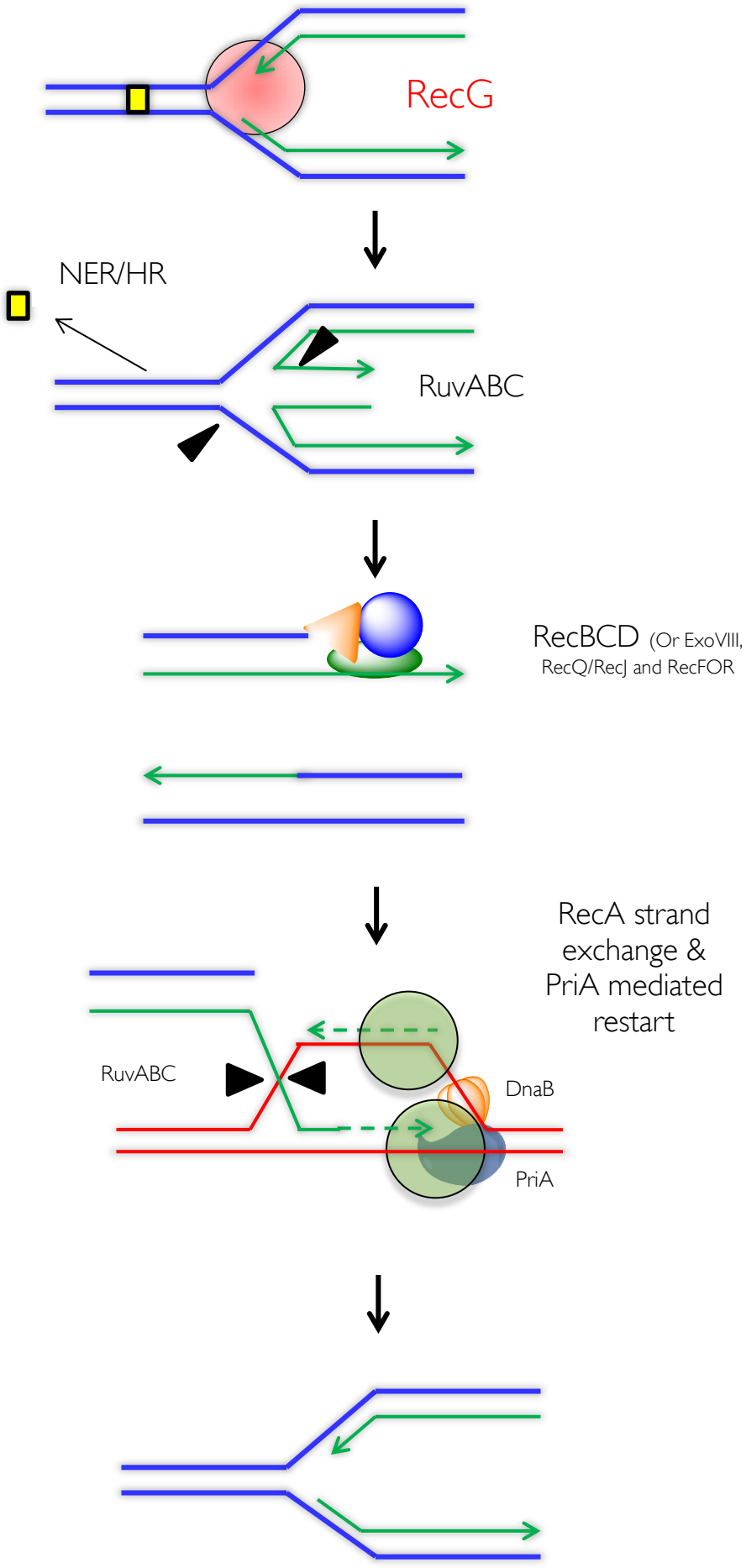

RecG
NER/HR
RuvABC
RecBCD (Or ExoVIII, RecQ/RecJ and RecFOR
RecA strand exchange & PriA mediated restart
RuvABC
DnaB
PriA

Figure 1.56 | **DNA replication restart of an ICL using RecG-mediated fork reversal, RecBCD and RuvABC.** Details in text.

In eukaryotes, fork regression is a debated issue and the importance of replication restart is considered lesser to that of bacteria owing to the ability of multiple origins to compensate for any fork stalling events. However, it has become clearer that fork restart mechanisms are present. Here, BLM, WRN and FANCM helicases are all capable of pushing back a fork to yield a Holliday junction (Gari *et al.*, 2008, Meng and Zhao, 2017). In eukaryotes, this process occurs in a similar fashion and is known as break induced replication (BIR). BIR in eukaryotes is then carried out by the canonical HJ proteins (Sakofsky and Malkova, 2017).

The archaeal nuclease Hef and the resolvase Hjc are thought to generate the DSB in this process (Lestini, Duan and Allers, 2010). Hjm may act analogously to the bacterial RecG in *Pyrococcus furiosus* (Fujikane, Shinagawa and Ishino, 2006)

Fork resetting

Following regression of the replication fork via the likes of RecG, the fork can be reversed in a recombination-independent fashion by the action of RecBCD.

In the case of an interstrand crosslink, RecG can push back the replication fork, generating a Holliday junction an allowing the crosslink to be a substrate for NER/HR repair. However, instead of cleaving the junction by RuvC, RecBCD can act to process the ends of the fork to degrade back to the branch

point (Chow and Courcelle, 2007). The replication fork has been reformed and replication can continue via PriA mediated restart (Figure 1.57).

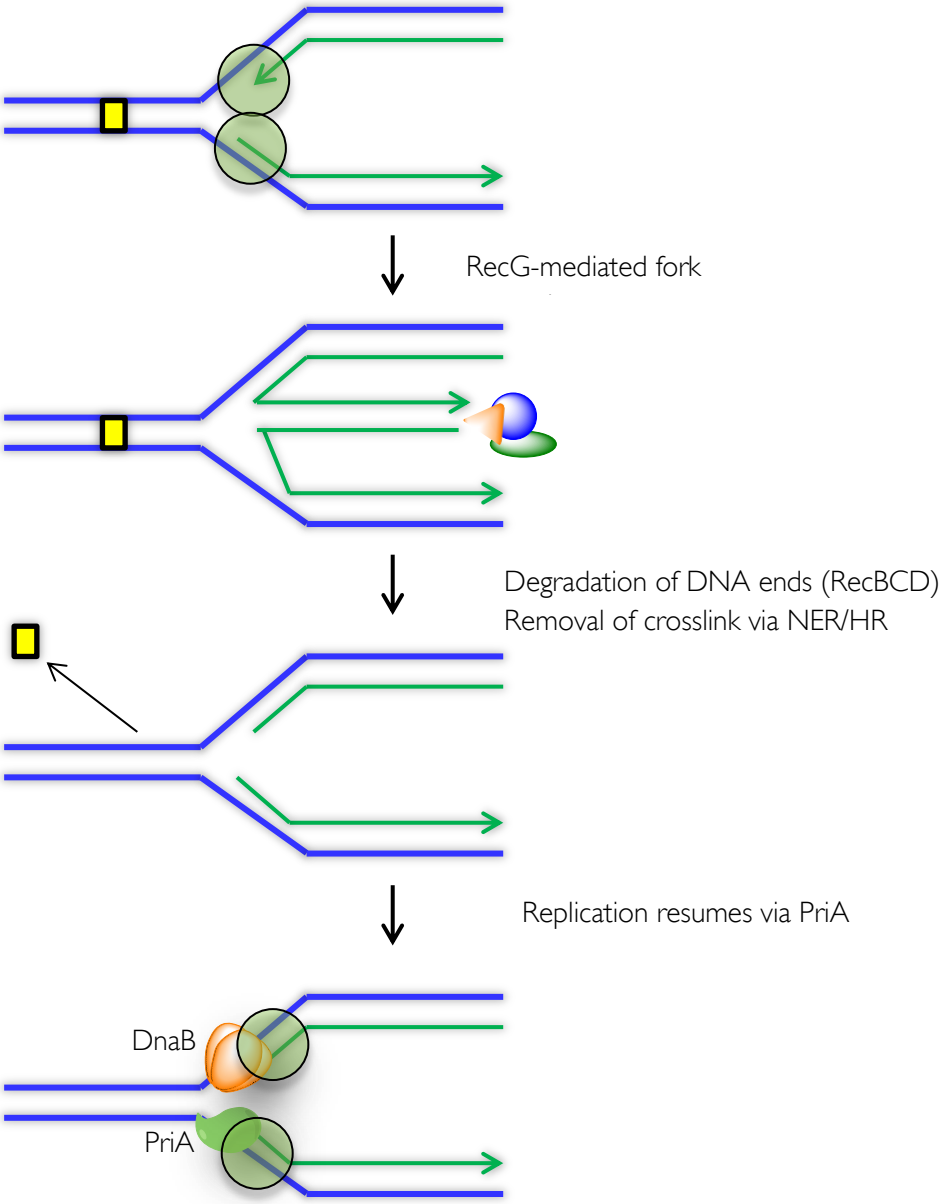


Figure 1.57 | **HR independent replication fork reversal and resetting by RecBCD.** In bacteria, once the Holliday junction is formed via the regression of the replication fork, mediated by RecG, RecBCD can act on the DNA ends produced to degrade them and reform a fork. The crosslink is removed via NER and HR and PriA can then act to restart DNA replication. Direct RecG mediated reversal can also occur after the ICL is removed, without the need for resection.

In humans, replication forks can be reversed using RecQ family helicases such as BLM and the DNA ends produced can be degraded by the Dna2 nuclease/helicase along with WRN helicase. RECQ1 limits Dna2 activity by preventing extensive nascent strand degradation (Thangavel *et al.*, 2015).

Lesion bypass (template switching)

An alternative method to deal with a lesion at the replication fork, that still allow replication (i.e. not a crosslink), for example a lesion on the leading strand, is to use a specialised mechanism to bypass the lesion (Lovett, 2017).

In bacteria, RecG can regress the junction as we've seen previously. At this point, a 3' ssDNA tail is created, owing to the lack of synthesis on the leading strand. This 3' tail can act as a primer for DNA synthesis in which the lagging strand is used as a template. RecG can then push the junction back to its original position and PriA can then recruit the replication machinery to restart replication on the leading strand. The lesion is thus bypassed and can be removed at a later point (Figure 1.58).

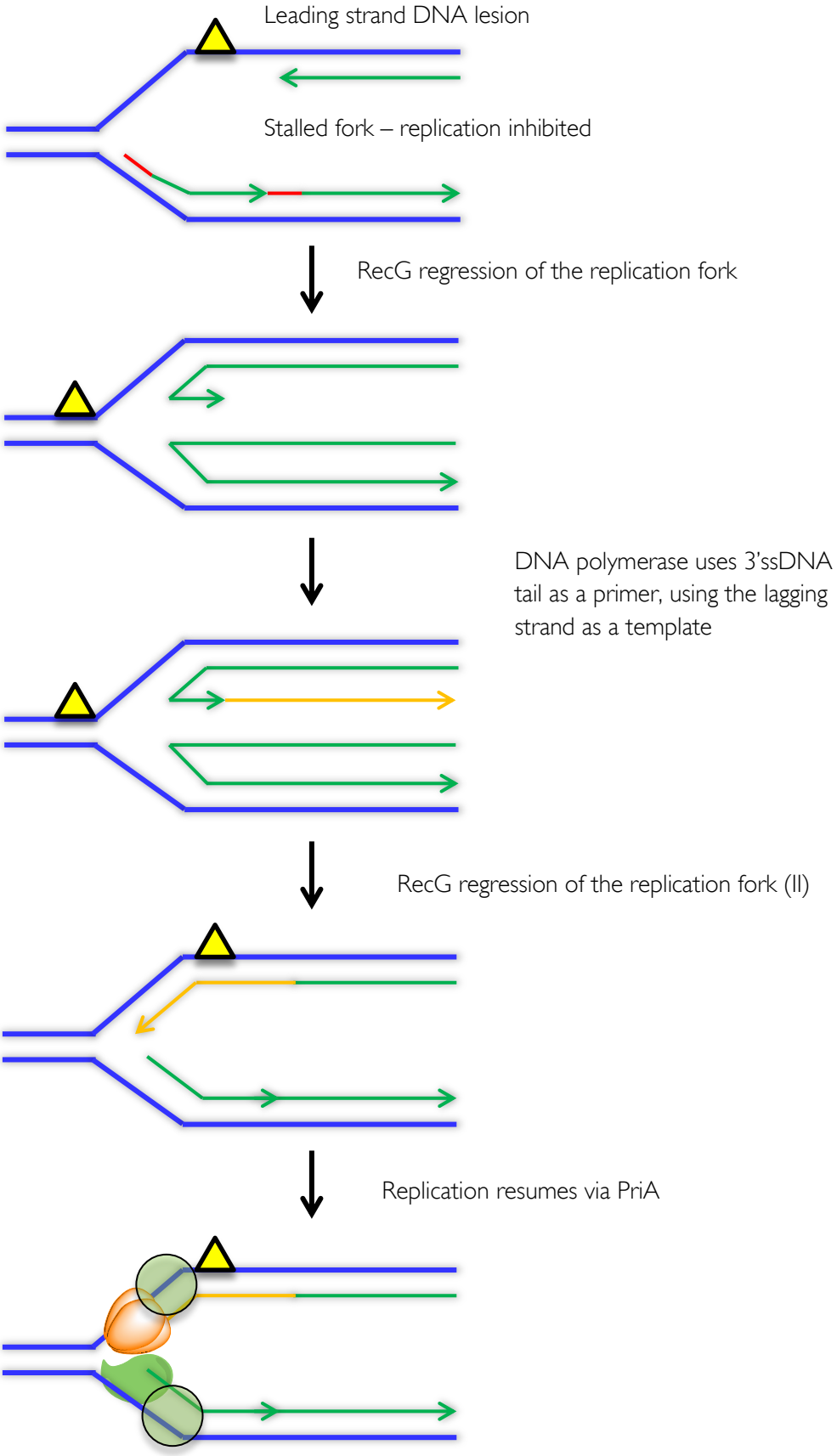

Leading strand DNA lesion
Stalled fork – replication inhibited
RecG regression of the replication fork
DNA polymerase uses 3'ssDNA tail as a primer, using the lagging strand as a template
RecG regression of the replication fork (II)
Replication resumes via PriA

Figure 1.58 | **Mechanism of template switching in bacteria.** RecG uses two round of replication fork pushing activity to allow DNA synthesis across the site of the lesion using the lagging strand as a template. The lesion is bypassed and can be removed later.

Translesion DNA polymerases

If a polymerase has been stalled at a replication fork, another option exists whereby via usage of the sliding clamp (β-clamp / PCNA), a less stringent DNA polymerase can be recruited to the fork and take over from the more stringent replicative polymerase. In *E. coli* PolV is a good example of this, as is Pol η in eukaryotes, which can easily bypass a CDP lesion (Goodman and Woodgate, 2013). In archaea translesion polymerases are generally absent, however it has been discovered that PriS, the small primase subunit, can function to traverse certain lesions in the DNA template (Jozwiakowski, Borazjani Gholami and Doherty, 2015).

Eukaryotes

Although the ability to process stalled replication forks in eukaryotes is possible, through the action of RecQ family helicases, such as BLM or WRN, eukaryotes and Archaea contain multiple origins of replication. Therefore, other replication forks can rescue stalled replication forks. Furthermore, it has been shown that excess levels of MCM helicase can licence previously dormant origins to assist should multiple replication origins stall (Ibarra, Schwob and Méndez, 2008) (Newman *et al.*, 2013).

1.6.4.5 Alternative methods of DNA replication

Replication in the absence of origins

Genetic studies in *H. volcanii* have shown that the four origins on the main chromosome are dispensable, cells without the origins remain viable and grow 7.5% faster than wild-type (Hawkins *et al.*, 2013). From this, two theories emerged, (1) In the absence of origins, ORC1 could bind non-specifically at random sites in the genome. (2) Dispersed initiation relies upon homologous recombination and can occur when D-loop intermediates are used to prime DNA replication. *H. volcanii* RadA deletion mutants were created and it was not possible to delete RadA in an originless background, indicating homologous recombination is essential to this process.

Originless replication could occur, at least in theory, within polyploid organisms, such as Archaea, owing to the saturated levels of available exchange partners with homology. The process occurs via break-induced replication, as seen in section 1.6.4.4. Needless to say, the fact that polyploidy is a strategy used by such Archaea, show the importance of this ability in survival against double-strand breaks, dependant on recombination and how this is critical to their extremophile properties.

Break-induced replication (BIR)

As seen in section 1.6.4.4, BIR can be used to restart a DNA replication fork from a double-strand DNA break, formed either from the collapse of a replication fork or from DNA damage.

BIR begins with a single DNA end, rather than typical HR which initiates with bistrand invasion of a homologous duplex. Once the 3’ end has invaded the duplex, homology search begins which eventually results in branch migration.

In BIR, one of two situations can occur. The first of which we've seen before in section 1.6.4.4, where the invading strand forms a Holliday junction structure which is cleaved by nucleases to re-establish the DNA replication fork. The second option is for the D-loop to act like the leading strand, with nascent DNA acting as the template for lagging strand synthesis (Figure 1.59). Dependant on the option elected, resultant synthesis leads to either conservative or semiconservative DNA synthesis.

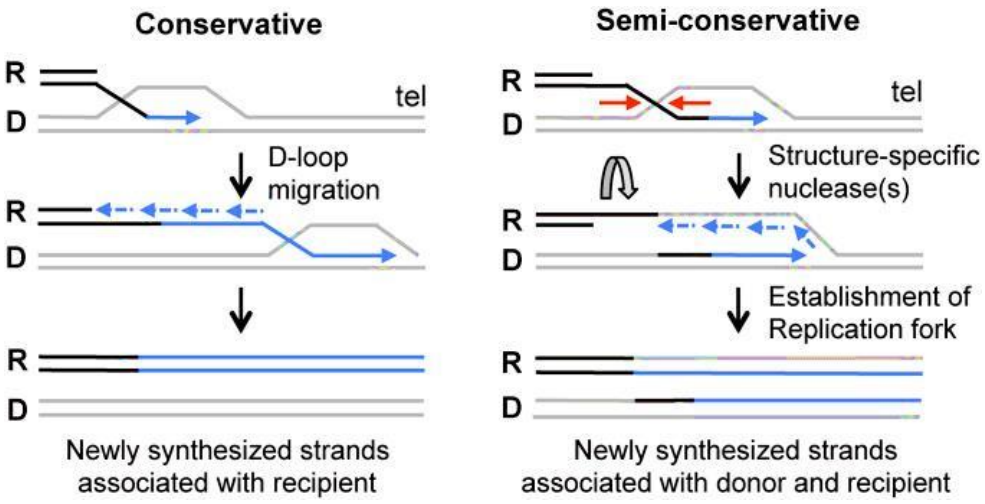


Figure 1.59 | **Model for conservative or semiconservative DNA synthesis in BIR.** If the Rad51-catalyzed D-loop migrated to the chromosome end (telomere) and lagging-strand synthesis initiates on the displaced nascent strand, both newly synthesised strand (blue) will segregate with the recipient (R) chromosome. Cleavage of the strand invasion intermediate by a nuclease establishes a replication fork that results in semi-conservative synthesis, detected by segregation of the newly synthesised DNA to the donor (D) and recipient (R) strands. Figure taken from (Donnianni and Symington, 2013).

The proteins involved in BIR, given its non-canonical mechanism, are indeed different to the normal processes of DNA replication. In eukaryotes, such as budding yeast, MCM helicase is not needed, instead a helicase known as Pif1, along with DNA Polδ is used. Pif1 helps in the recruitment of Polδ to the D-loop and is capable of promoting Polδ-mediated DNA synthesis (Malkova and Ira, 2013).

1.6.4.6 Double-strand break repair (DSBR)

In summary, repair in response to a DSB can occur through a variety of different mechanisms, each using a defined subset of specialised proteins and highly regulated by the cell cycle. The mechanisms are summarised in Figure 1.60 below and consist of homologous recombination (with or without a crossover), synthesis-dependant strand annealing, single-strand annealing and, a couple of other pathways not discussed in detail here, non-homologous end joining, a mutagenic pathway used in G1 where a homologous duplex is not available, and microhomology-mediated end joining, a subset of the NHEJ pathway (Wang and Xu, 2017). All pathways are regulated so Holliday junctions are only formed when absolutely needed, i.e. in the context of a sister chromatid or homologue. NHEJ is mainly used in G1, and also in immune cells to increase the diversity of human antibody repertoires (Soulas-Sprauel *et al.*, 2007). In meiosis, crossover formation is biased owing to the interplay between ZMM proteins (also known as the synapsis initiation complex) committing the Holliday junction to be resolved, in eukaryotes, by MutLγ and Exo1 (Lynn, Soucek and Börner, 2007). Junctions that escape this pathway are resolved by nucleases to yield either crossover or noncrossover products. Outside of meiosis, including in mitotic cells, noncrossover products are favoured and the major pathway in eukaryotes appears to be dissolution of the junction or, before the second end is captured by Rad52, SDSA. SSA occurs in a specific context where repeated dispersed elements are located either side of the DSB. Break induced repair is a separate pathway that repairs single DNA ends, such as those that arise from a stalled replication fork due to a lesion on the lagging strand, or an ICL. A similar mechanism is involved in the ability of certain Archaea (*H. volcanii*) to replicate without origins, using homologous recombination from a single end.

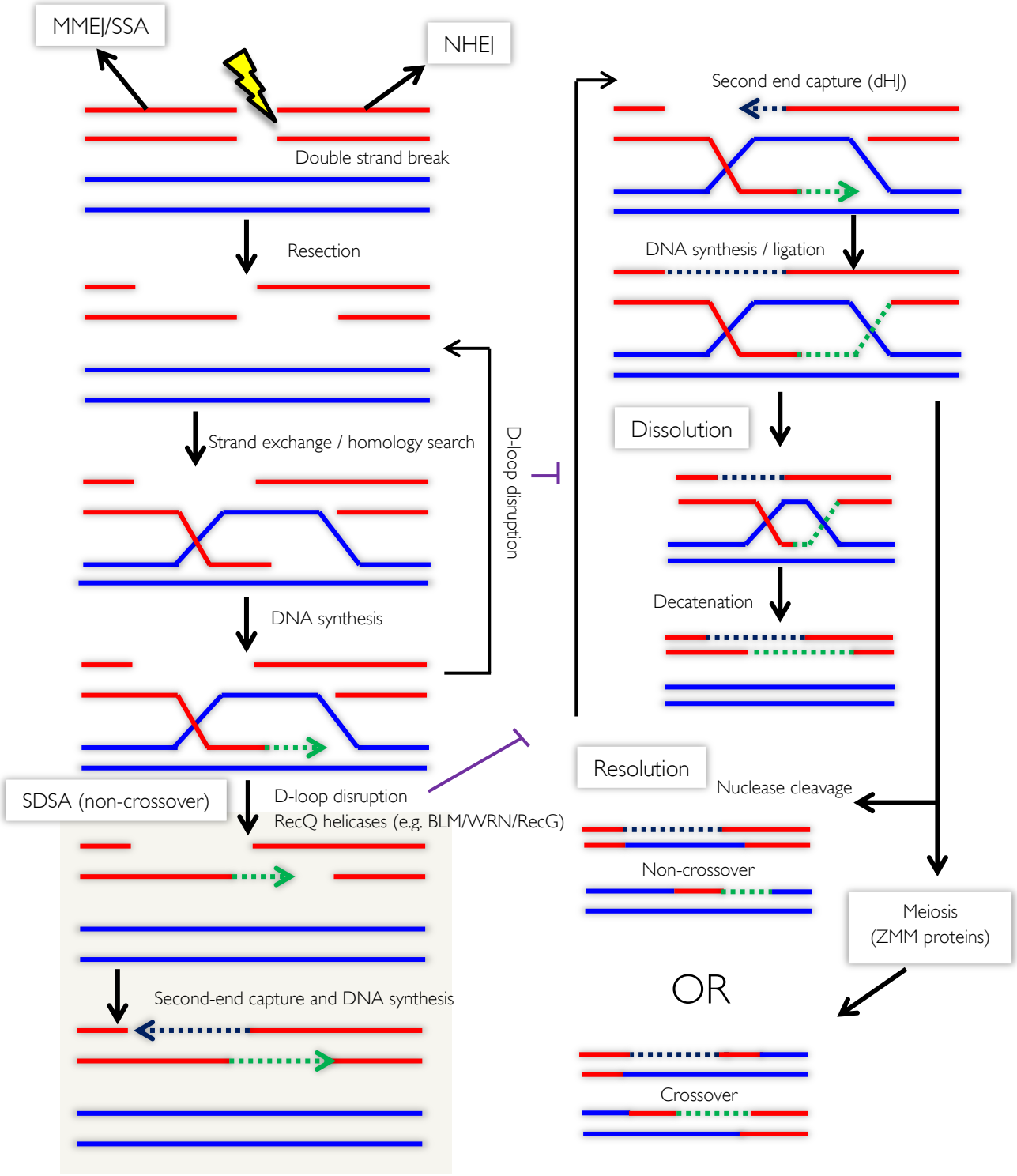

MMEJ/SSA
NHEJ
Double strand break
Resection
Strand exchange / homology search
DNA synthesis
D-loop disruption
SDSA (non-crossover)
D-loop disruption
RecQ helicases (e.g. BLM/WRN/RecG)
Second-end capture and DNA synthesis
Second end capture (dHJ)
DNA synthesis / ligation
Dissolution
Decatenation
Resolution
Nuclease cleavage
Non-crossover
Meiosis
(ZMM proteins)
OR
Crossover

Figure 1.60 | **Overview of DSBR mechanisms in somatic and meiotic cells.** Although a double strand break is shown, a break where there is only one end can be used by the BIR pathway (not shown). A double strand break occurring outside of meiosis can be resolved via multiple pathways. In G1 both NHEJ and MMEJ are favoured along with SSA, owing to the fact that a homologue is not needed. The regulation of this is carried out by various machinery, involving chromatin structure remodelling and variant histones. If a homologue or sister chromatid is present, homology search and strand invasion allows formation of a D loop. At this point the D loop can be disrupted by RecQ family helicases. Once initial DNA synthesis occurs, RecQ helicases can also disrupt the structure to lead to SDSA and a noncrossover product. Such structures are generally resolved prior to formation of the double Holliday junction. Second end capture facilitates formation of the dHJ which can then be formed into a non crossover product by RecQ helicases such as BLM along with topoisomerases to produce dissolution via a decatenation reaction. This mechanism is preferred in cells not undergoing meiosis. Nucleases are upregulated in M phase to ensure no Holliday junction structures are left before chromosome / chromatid separation. Nuclease cleavage in meiosis results in primarily crossover structures via an unknown mechanism, thought to involve MutLγ as part of the ZMM complex. Outside of meiosis resolution can occur via same or opposite sense cleavage resulting in either a crossover or noncrossover product. This occurs to a minority of junctions at this stage and involves MUS81:EME1 and SLX1-SLX4 in eukaryotes. In all cases GEN1 and its homologues, if present, can act as a safeguard nuclease to help resolve structures where previous nucleases have failed. Figure adapted from (Wright, Shah and Heyer, 2018).

Chapter 2: Materials and Methods

2.1 Materials

All strains and plasmids were constructed by the author unless stated otherwise.

2.1.1 Strains

[ ] Indicate integrated plasmid DNA

{ } Indicate episomal plasmid DNA

Table 2.1 ***H. volcanii*** **strains used in this study**

Strain	Parent	Genotype	Notes
H26	H18	Δ*pyrE2*	(Wendoloski, Ferrer and Dyall-Smith, 2001)
H645	H640	Δ*pyrE2, bgaHa,* Δ*mre11-rad50*	(Delmas *et al.*, 2009)
H1206	H1202	Δ*pyrE2,* Δ*mrr*	(Allers, H. Ngo, *et al.*, 2004)
H53	H47	Δ*pyrE2,* Δ*trpA*	(Allers, H.-P. Ngo, *et al.*, 2004)
H1895	H1892	Δ*pyrE2,* Δ*pillin*	(Strillinger *et al.*, 2016)
H509	H477	Δ*pyrE2,* Δ*uvrA*	(Lestini, Duan and Allers, 2010)
H1181	H1161	Δ*pyrE2,* Δ*uvrB*	(Lestini, Duan and Allers, 2010)
H1187	H1175	Δ*pyrE2,* Δ*uvrC*	(Lestini, Duan and Allers,

			2010)
H514	H482	Δ*pyrE2, Δuvr*D	(Lestini, Duan and Allers, 2010)
H358	H338	Δ*pyrE2, Δhef*	(Lestini, Duan and Allers, 2010)
H178	H158	Δ*pyrE2, Δhjc*	(Lestini, Duan and Allers, 2010)
H5562	H26	*ΔpyrE2 xerC/D+::[ΔxerC/D pyrE2+]*	This study
H5610	H5562	*ΔpyrE2, ΔXerCD*	This study
H5563	H53	*ΔpyrE2 ΔtrpA xerC/D+::[ΔxerC/D pyrE2+ trpA+]*	This study
H5630	H1206	*ΔpyrE2 Δmrr {p.syn::XerC/D pyrE2+ hdrB+}*	This study
H5629	H26	*ΔpyrE2 {p.syn::XerC/D pyrE2+ hdrB+}*	This study
H5628	H5610	*ΔpyrE2 ΔxerC/D {native::XerC/D pyrE2+}*	This study
H5627	H5610	*ΔpyrE2 ΔxerC/D {p.syn::XerC/D pyrE2+ hdrB+}*	This study
H5625	H26	*ΔpyrE2 {pyrE2+ hdrB+}*	This study
H5626	H26	*ΔpyrE2 {pyrE2+}*	This study
H1192	H1210	*ΔpyrE2 Δphage*	This study
H5635	H1192	*ΔpyrE2 Δphage {p.syn::XerC/D pyrE2+ hdrB+}*	This study
H5655	H1206	*ΔpyrE2 Δmrr {p.syn::xerCD_Cdn_Op*	This study

		pyrE2+ hdrB+}	

Table 2.2 ***H. mediterranei*** **strains used in this study**

Strain	Parent	Genotype	Notes
H826	H824	Δ*pyrE2*	Created by Moshe Mevarech

Table 2.3 ***E. coli*** **strains used in this study**

Strain	Genotype	Notes
XL1-Blue	*endA1, gyrA96 (Na1R), lac [F' proAB lacIqZΔM15 Tn10 (TetR)], Δ(mcrA)183, Δ(mcrCBhsdSMR-mrr)173, recA1, relA1, supE44, thi-1*	Standard cloning strain enabling blue/white selection in conjunction with pBluescriptII SK+ plasmid derivatives. Tetracycline resistant. Restriction endonuclease and recombination deficient, dam+. From Stratagene.
N2338 (GM121)	*F-, ara-14, dam-3, dcm-6, fhuA31, galK2, galT22, hsdR3, lacY1, leu-6, thi-1, thr-1, tsx-78*	For preparation of unmethylated DNA used in *H. volcanii* transformations. From Allers *et al. 2004*).

2.1.2 Plasmids

All plasmids contain the ampicillin resistance gene, AmpR, to allow for selection during cloning in *E. coli*.

Table 2.4 **Plasmids used in this study**

Name	Use	Notes
pTA908	*E. coli / H. volcanii* shuttle vector with 224bp fragment of p.*tnaA* promoter. Used for overexpression via tryptophan inducible gene expression. Contains pHV2 and ColE1 origins of replication, as well as AmpR and PyrE2 as selectable markers.	Created by Amy Stroud
pTA927	*P908* tryptophan inducible overexpression shuttle vector with insertion of annealed p.syn synthetic transcription terminator. Contains pHV2 and ColE1 origins of replication, as well as AmpR and PyrE2 as selectable markers.	(Allers *et al.*, 2010)
pTA1231	*Hel308a* overexpression construct with N terminal His tag.	Created by TA lab
pTA1992	Overexpression vector with constitutively expressed synthetic p.syn2 promoter.	(Haque, Paradisi and Allers, 2019)
pTA2773	XerC/D like integrase (HVO_2259) overexpression construct using p.syn promoter.	This study
pTA131	For deletion construct creation.	(Allers, H.-P. Ngo, *et al.*,

		2004)
pTA298	For deletion construct creation. Contains *p.fdx::trpA* to create tryptophan mediated deletion constructs. Digest with BamHI to liberate TrpA for inclusion in pTA131 suicide vector	(Allers, H.-P. Ngo, *et al.*, 2004)
pTA2781	XerCD under p.syn with C terminal strepII tag.	This study
pTA2782	XerCD under p.syn (Dam-)	This study
pTA2783	XerCD deletion construct. Non Trp marked.	This study
pTA2786	XerCD deletion construct. Non Trp marked. Dam-.	This study
pTA2784	XerCD deletion construct. Trp marked.	This study
pTA2785	XerCD deletion construct. Trp marked. Dam-.	This study
pTA2863	XerCD placed under native promoter with native pHV1 origin (low copy number). Shuttle vector used for complementation of XerCD deletion strain.	This study
pTA2864	XerCD placed under native promoter with native pHV1 origin (low copy number). Shuttle vector used for complementation of XerCD deletion strain. Dam -.	This study
pTA354	Used to generate pTA2863. Backbone vector with pHV1 origin.	Created by TA
pTA1992	Backbone for codon optomized XerCD gBlock cloning and p.syn::XerCD constructs	(Haque, Paradisi and Allers, 2019)
pTA2795	Backbone for codon optimised XerCD gBlock cloning and p.syn::XerCD constructs. Dam-.	This study
pTA912	*pyrE2* containing vector used	Created by TA lab

	for transformation viability assay	
pTA2868	Codon optimised XerCD placed under p.syn in pTA1992 backbone	This study

N.B. References provided where plasmid is published and reference is known.

2.1.3 Oligonucleotides

All oligonucleotides were ordered through Integrated DNA Technologies (IDT).

Table 2.5 **Oligonucleotides used in this study**

Name	Sequence 5'→3'	Use	Notes
pBSF2	5'-TTAAGTTGGGTAACGCCAGGG-3'	Forward primer for colony PCR of p1231 for Hel308 overexpression assay	Constructed by TA lab
Hel308R	5'-CGACGACGCAGGTGAGTTGG-3'	Reverse primer for colony PCR of p1231 for Hel308 overexpression assay	Constructed by TA lab
XerC_F	5'ATCTGACATGTTCAAGGAGGGGAACTAACATG3'	Forward primer for amplification of HVO_2259 from genome, incorporates Pci1 site. Used for cloning into p1992 to generate p2773	This study
XerC_R	5'GTACGGATCCACCAACTGTCTATCCTCGTTA3'	Reverse primer for amplification of HVO_2259	This study

		from genome, incorporates BamH1 site. Used for cloning into p1992 to generate p2773 with XerC_F.	
XerC_R_Nhe1	5'CAGATGCTAGCGAGTCCAATTTCCTTG3'	Reverse primer for amplification of HVO_2259 from genome, incorporates Nhe1 site. Used for cloning into p1992 to generate XXX. Places HVO_2259 under p.syn and includes C terminal strepII tag.	This study
XerDel_F_Cla1	5'-ATCTGATCGATGAGACATCTGTGGACAGCCCC-3'	Forward primer for amplification of 500bp upstream homology arm for XerC/D (HVO_2259) deletion construct. Includes Cla1 site for ligation into p131.	This study
XerDel_R_BamH1	5'-ATCTGGGATCCGTTAGTTCCCCTCCTTGAGCCG-3'	Forward primer for amplification of 500bp upstream homology arm for XerC/D (HVO_2259) deletion construct. Includes BamH1 site for ligation to downstream	This study

		homology arm.	
XerDel_F_BamH1	5'-ATCTGGGATCCGGTCGAAGTCATTATTTATCGAAACCC-3'	Forward primer for amplification of 500bp downstream homology arm for XerC/D (HVO_2259) deletion construct. Includes BamH1 site for ligation to upstream homology arm.	This study
XerDel_R_Not1	5'-ATCTGGCGGCCGCGAGCAACAATCCGTTCTCTGG-3'	Forward primer for amplification of 500bp downstream homology arm for XerC/D (HVO_2259) deletion construct. Includes Not1 site for ligation to p131	This study
PBSF	5'-GTAAAACGACGGCCAGT-3'	For Sanger DNA Sequencing reactions of pBluescript backbones (Forward)	Constructed by TA lab
pBSR	5'AACAGCTATGACCATG-3'	For Sanger DNA Sequencing reactions of pBluescript backbones (Reverse)	Constructed by TA lab

PtnaAFint	5’-GCCTGCCGATTACTTCACATTCGC-3’	For Sanger DNA sequencing of genomic library constructs (forward)	Constructed by TA lab
pBSR	AACAGCTATGACCATG	For Sanger DNA sequencing of genomic library constructs (reverse)	Constructed by TA lab
PhageProbeBII	5'-ACTCTTCTTTCGCTTCGTCACG-3'	Reverse primer for colony PCR of pTA2773 and pTA2780	Constructed by TA lab
pBSF2	5'-TTAAGTTGGGTAACGCCAGGG-3'	Forward primer for colony PCR of p2773 (dam -) in H1206 for p.syn::Xer survival assay	Constructed by TA lab
6264R2	5'-GGGACGCCGACCACTTC-3'	Forward primer for amplification of XerCD from gDNA of H26.	Constructed by TA lab
PhageProbeR3	5'-CAACGCCATTAGTCTGTCTGTAAGC-3'	Reverse primer for amplification of XerCD from gDNA of H26	Constructed by TA lab
Xer_CdnOp_V3	ATCTGACATGTTCAAGGAGGGGAACTACCATGGGCGCGGAGCCGGGCTCGTCGAAAATCTACGACAACAAGCACGACGAAGTCAACTACTTCATCACCCGGAAGCACGCGACCGGCCGGAGCGAACGGACGCTCAACTCCTACTCGCGCATCC	Synthetic double stranded DNA fragment (gBlock) of codon optimised XerCD. NcoI and BamHI sites used for cloning into p1992 cut with	This study

	TCCGCGAGTTCTTCCACGACCAGTTCCCGGA CCTCTCGCCCTCGGAGGTCGAAATCCGGCAC GTCGAGGACTACCTCATGGCGCTCACCGACC GCGGCGTCTCGCAGAACAGCAAGAAGAAGT ACCTGGAGGTGCTCTCGTCGTTCTACGGCTA CACCCTCAAGCGCCCGCAGTTCGAGGGGATT ACGTCGAACCCGGCCGCGGTCGTGATGGAG GAGATCCCCCGCGTCCGGCCCGACCGCCCG GACTGTGCCACCTGGGAGAACGCCTGCAAG CTCATCAACGCCATCCCCGACCCGCGGGACA AAACGGTGACGATCATCCTGGCGAAGACGG GCGCGCGCCTCCTGGAAGTCCTGTCGATTGA GGAGGACGACGTCGACCTGGAGAAGGGCTT CATTCGCCTCCGCGAACGCAAGGGCGGCAA ACAAACCGTCGTGCCGATCGACGACGAGAC GATCTACGCGATCAAGCGGTACCAGTTCGTC AACGCGGACCTCGACTCGCCCTACCTGTTCA CCTCGAACAAGGGGGGCCGGCTCTCCAAGG AGCGCATCCGCCGGGAAGTGAAAGCCGCGG CCGACCGCGCCGGCGTCGCGCCCAAGGAAG AACGCCGCTTCGAGAAGAAGTTCACGCCGC ACACGTTCCGCACCGTCTTCACGACGCTGAT GCGCAAGCAGGGCATGAAACCGTACATCCT CAAATACATTCGCGGCGACGCCAAGACGGA GACGATGGACATCTACACGCGGGTCGACCG CGACGAAGCCAAAGAGGAATACCTCAACTG	Pci1/BamH1 - generates N terminal His tagged CDS under p.syn.	

	CATTAAGGAGATCGGGCTGTGACGAGGATA GACAGTTGGTGGATCCGCTAGCACTCG		
PhageProbeF3	TTCACAAGCAACAAAGGAGGACGG	Primer for pilot PCR before qPCR phage assay	Constructed by TA lab
hel308R	CGACGACGCAGGTGAGTTGG	Primer for pilot PCR before qPCR phage assay	Constructed by TA lab
Hel308Fint	AGCGCTGGGAGGAGTACGGC	Primer for pilot PCR before qPCR phage assay	Constructed by TA lab
t.syn R	CCGAAAAATGCGATGGTCCAGAGGTGC	Primer for colony PCR of Codon optimised plasmid in H1206	Constructed by TA lab
p.syn F	CGAGAATCGAAACGCTTATAAGTGCCCCCCG GCTAGAGAGAT	Primer for colony PCR of Codon optimised plasmid in H1206	Constructed by TA lab
p.synR	CGATCTCTCTAGCCGGGGGGCACTTATAAGC GTTTCGATTCT	Primer for colony PCR	Constructed by TA lab

2.1.4 Chemicals and Enzymes

Unless otherwise stated, all enzymes were purchased from New England Biolabs (NEB) and chemicals from Sigma.

2.1.4.1 Media

Haloferax volcanii Media

Sterilisation of media

Sterilisation of all media is carried out via autoclaving for 15 minutes at 121°C.

Storage of media

Media are stored in the dark at room temperature unless otherwise stated.

30% salt water (30% SW)

4M NaCl

148 mM $MgCl_2.6H_2O$

122 mM $MgSO_4.7H_2O$

94 mM KCl

20 mM Tris.HCL pH 7.5

18% salt water (18% SW)

Dilution of 30% SW, autoclaved, then addition of $CaCl_2$ to a final concentration of 3mM once cool.

Trace elements

1.82 mM $MnCl_2.4H_2O$

1.53 mM $ZnSO_4.7H_2O$

8.3 mM $FeSO_4.7H_2O$

200 µM $CuSO_4.5H_2O$

Filter sterilised and stored at 4°C

10x YPC

5% yeast extract (Difco)

1% peptone (Oxoid)

1% Casamino acids

17.6 mM KOH

No autoclaving needed. Use immediately.

10x Ca

5% Casamino acids

17.6 mM KOH

No autoclaving needed. Use immediately.

Hv-Ca Salts

362 mM $CaCl_2$

8.3% v/v trace elements (above)

615 µg/ml thiamine

77µg/ml biotin

KPO4 Buffer

308 mM K_2HPO_4

192 mM KH_2PO_4

pH 7.0 with NaOH

Hv-YPC Agar

1.6% Agar (Bacto)

18% SW

1X YPC

3 mM $CaCl_2$

Microwaved prior to addition of 10X YPC and $CaCl_2$ to dissolve agar. 10X YPC added, autoclaved, then $CaCl_2$ added once cool, prior to pouring.

Hv-Ca Agar

1.6% Agar (Bacto)

18% SW

1X Ca

0.84% v/v Hv-Ca Salts

0.002% v/v KPO_4 buffer (pH 7.0)

Microwaved prior to addition of 10X Ca, Hv-Ca Salts and KPO_4 Buffer. 10X Hv-Ca added, autoclaved, then Hv-Ca Salts and KPO_4 Buffer added once cool, prior to pouring.

Hv-YPC Broth

18% SW

1X YPC

3 mM $CaCl_2$

$CaCl_2$ added after autoclaving, once cool.

Hv-Ca+ Broth

18% SW

1X Hv-Ca

30mM Tris.HCl (pH 7.0)

2.5% v/v HV-Min Carbon Source

1.2% v/v Hv-Min Salts

0.002% v/v KPO_4 Buffer (pH 7.0)

444 nM biotin

2.5 µM thiamine

18% SW, 1X Ca, 30 mM Tris.HCl (pH 7.0) added then autoclaved. Other components added once cool.

Haloferax volcanii Media Supplements

Table 2.6 **Table of commonly used media supplements and their working concentrations.** Solutions were sterilised via passing through a 0.2µm filter.

Name	Abbreviation	Final concentration
Leucine	Leu	50 µg/ml
Uracil	Ura	50 µg/ml
Thymidine	Thy	50 µg/ml (+ 50 µg/ml hypoxanthine for Hv-Ca and Hv-Min)
5-Fluoroorotic acid	5-FOA	50 µg/ml (+ 10 µg/ml uracil)
Tryptophan	Trp	50 µg/ml
Mevinolin	Mev	4 µg/ml

Table 2.7 ***H. volcanii*** **mutants and relative growth in different media.**

Genotype	Hv-YPC	Hv-Min	Hv-Ca
Δ*pyrE2*	+	Ura-	Ura-
Δ*trpA*	+	Trp-	Trp-
Δ*leuB*	+	Leu-	+
Δ*hdrB*	Thy-	Thy-	Thy-

Eschericha coli Media

Sterilisation of media

Sterilisation of all media is carried out via autoclaving for 15 minutes at 121°C.

Storage of media

Media is stored at room temperature.

Lysogeny Broth (LB)

1% tryptone (Bacto)

0.5% yeast extract (Difco)

170 mM NaCl

2 nM NaOH

pH to 7.0

LB Agar

300ml LB broth

1.5% agar

SOC Broth

2% tryptone (Bacto)

0.5% yeast extract (Difco)

10mM NaCl

2.5mM KCl

10mM $MgCl_2$

10mM $MgSO_4$

20mM glucose

Eschericha coli Media Supplements

Table 2.8 **Media supplements used with *E. coli***

Name	Abbreviation	Final Concentration
Ampicillin	Amp	50 µg/ml
5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside	X-Gal	40 µg/ml

2.1.4.2 Other Solutions

TE buffer

10mM Tris.HCl (pH 8.0)

1mM EDTA

Sodium Acetate

3M NaAc (pH 5.2, filter sterilised)

Mitomycin C stock and MMC agar

Mitomycin C (Sigma) was dissolved in sterile distilled water to a final concentration of 0.5mg/ml. This was then diluted as required by addition to either liquid culture or Hv-YPC/Hv-Ca agar once cooled, immediately prior to pouring.

2.2 Methods

2.2.1 *Escherichia coli* Microbiology

Growth and Storage

Cultures on solid media were grown overnight (12-16 hrs approx) in a static incubator (LEEC) at 37°C. Small liquid cultures (1-5ml) were grown for approximately 12-16 hrs in a shaking incubator (Innova 4330, New-Brunswick Scientific) at 37°C with 8rpm shaking. Larger cultures (100ml – 300ml) were grown shaking at 250rpm.

For long term storage, a 20% glycerol stock of the liquid cultures were made, then snap frozen on dry ice prior to storage at -80°C.

Preparation of Electrocompetent Cells

Electrocompetent cells, *dam+* and *dam-*, were prepared from the *E. coli* strains XL1-Blue (Tet^R) and N2338 respectively.

A 5ml culture was grown at 37°C with 8rpm rotation for ~16hrs with the appropriate antibiotic selection. Cells were diluted by 10^{-2} in LB broth, maintaining appropriate selection, and was grown at 37°C, 110rpm shaking until A_{650} = 0.5-0.8.Cells were pelleted at 6000 x *g* for 12 minutes at 4°C and the supernatant discarded. The pellet was resuspended in an equal volume of 1mM HEPES (pH 7.5) and kept on ice. Pelleting and resuspension steps were repeated using 0.5 volumes of 1mM HEPES, 0.25 volumes of 1mM HEPES + 10% glycerol, 0.1 volumes 1mM HEPES + 10% glycerol, and 0.001 volumes 1mM HEPES + 10% glycerol. Cells stored in 100µl aliquots at -80°C.

Transformation of *Escherichia coli* via Electroporation

Approximately 1µg of DNA, in a maximum volume of 4µl, was added to 40µl of Electrocompetent cells on ice and transferred to a pre-chilled electroporation cuvette (1mm electrode gap, GENEFLOW). The cuvette was pulsed at 1.8 kV in a gene pulser (BioRad). 1ml SOC media was added to the cells, which were then allowed to recover for up to an hour at 37°C with 8rpm rotation. Cells are subsequently plated onto selective plates and incubated at 37°C overnight.

2.2.2 *Haloferax volcanii* Microbiology

Growth and Storage

Cultures on solid media were grown in a static incubator (LEEC) at 45°C for 5-14 days in plastic bags to minimise desiccation and contamination. Small liquid cultures (1-5ml) were grown for 12-16hrs in a static incubator at 45°C with 8rpm rotation.

Cultures on solid media were stored at room temperature for up to one month. Though liquid cultures can be stored at room temperature, cells can pellet over the course of a few days, so resuspension is required.

Long-term storage is achieved through suspending the culture in 80% glycerol (with 6% SW) and storing at -80°C.

Transformation of *Haloferax volcanii using* PEG600

H. volcanii can be easily transformed with DNA via the addition of PEG600 (Cline *et al.*, 1989). For strains containing the restriction endonuclease, Mrr, which targets methylated DNA (Holmes, Nuttall and Dyall-Smith, 1991, Allers *et al.*, 2010a), *dam*- DNA must be used, therefore the relevant plasmids must first be passed through a *dam- E. coli* host. Alternatively, *dam+* DNA can be used if a Δ*mrr* host strain of *H. volcanii* is used.

Transformation Solutions

Buffered Spheroplasting Solution

1M NaCl

27mM KCl

50mM Tris.HCl (pH 8.5)

15% Sucrose

Unbuffered Spheroplasting Solution

1M NaCl

27mM KCl

15% Sucrose

pH 7.5

DNA to be transformed

83mM EDTA (pH 8.0)

15µl unbuffered spheroplasting solution

~1-2µg DNA in final volume of 30µl per transformation reaction

60% Polyethylene Glycol 600 (PEG600)

60% v/v PEG600 in unbuffered spheroplasting solution

Spheroplast Dilution Solution

23% SW

15% Sucrose

37.5 mM $CaCl_2$

Regeneration Solution

18% SW

1x YPC

15% Sucrose

30mM $CaCl_2$

Plating Solution

18% SW

15% Sucrose

30mM $CaCl_2$

A 5ml culture of *H. volcanii* was grown in Hv-YPC broth at 45°C until an A_{650} of 0.6-0.8. Cells were pelleted at 3300xg for 8 minutes in a 14ml round-bottom tube (Sarstedt), and the supernatant was removed subsequently. The cells were gently resuspended in 1ml of buffered spheroplasting solution and transferred to a round-bottomed 2ml tube (Eppendorf). Cells were pelleted at 3300 x *g* for 8 minutes and supernatant removed. The cells were resuspended in 200µl of buffered spheroplsating solution (200µl per transformation reaction). For each 200µl transformation reaction, a 20µl drop of 0.5M EDTA, pH 8.0, was pipetted onto the side of the tube, held horizontally. The tube was then inverted a couple of times and incubated at room temperature for ten minutes to allow formation of spheroplasts.

Transforming DNA was added in the same manner as the EDTA, having been prepared to the specification noted above. The mixed solution was then left at room temperature for a further five minutes.

60% PEG 600 was prepared as noted above, 250µl (equal volume) was added to each transformation reaction in the same manner as EDTA, however the tube was shaken horizontally approximately ten times. The solution was left at room temperature for 30 minutes. Spheroplasts were diluted by the addition of 1.5ml spheroplast dilution solution, inverted gently to mix, and was left at room temperature for two minutes. Spheroplasts were pelleted at 3300 x *g* for 8 minutes at room temperature, with subsequent supernatant removal.

The pellet was transferred to a sterile 4ml tube, whole, and 1ml of relevant regeneration solution (Hv-YPC or Hv-Ca) was added. The pellet was allowed to recover for 90 minutes at 45°C with 8rpm rotation. The pellet was resuspended thereafter via agitation of the tube and incubated for a further three hours, or overnight, at 45°C with 8rpm rotation. Cells were transferred to a fresh 2ml round-bottomed tube and pelleted at 3300 x *g* for 8 minutes with subsequent supernatant removal. The cells were resuspended gently in 1ml plating solution. Appropriate dilutions were made and 100µl was plated on appropriate media. Plates were incubated at 45°C for 5-14 days depending on experimental protocol.

2.2.3 Deoxyribonucleic Acid (DNA) Extraction

Escherichia coli plasmid purification

Plasmids were purified from *E. coli* via the Macherey-Nagel Nucleospin (Mini) and Nucleobond AX (Midi) kits. Protocol was followed as per the manufacturer's recommendations, unless otherwise stated. Minipreps and Midipreps used 1.5ml and 300ml of *E. coli* cell culture respectively. Plasmid DNA from minipreps was eluted into 30µl of elution buffer. Plasmid DNA from midipreps was isopropanol and ethanol precipitated, prior to resuspending in 200µl of TE. Plasmid DNA purity and concentration was checked via NanoDrop (NanoDrop 2000, Thermo Fisher) prior to being stored at -20°C.

Haloferax volcanii Genomic DNA Extraction

Genomic DNA Extraction via Spooling

A 5ml culture of *H. volcanii* was grown in Hv-YPC at 45°C until an A_{650} of 0.6-0.8. Cells were pelleted at 3300 x *g* for 8 mintues in a round-bottomed 2ml eppendorf. Supernatant was removed after spin. Cells were resuspended in 200µl ST buffer, followed by an equal volume addition of lysis solution. The tube was then inverted to mix the contents. The lysate was overlaid with 1ml of 100% ethanol (EtOH), after which the DNA was spooled at the interface onto a capillary tip until the liquid was clear. DNA was subsequently washed twice in 1ml 100% EtOH and the DNA was air-dried, finally being resuspended in 500µl of TE. A subsequent precipitation step was carried out via addition of 50µl of 3M NaAc (pH 5.2) and 400µl Isopropanol, after a quick spin at max rpm (~5min), 1ml of 70% EtOH was added to the pellet prior to an additional spin step as before, the DNA was resuspended in 100µl TE and incubated shaking at 45°C overnight to ensure DNA goes back fully into suspension.

ST Buffer

1M NaCl

20mM Tris.HCl, pH 7.5

Lysis Solution

100mM EDTA, pH 8.0

0.2% SDS

Plasmid Extraction from Haloferax volcanii

Plasmid miniprep extraction is carried out in *H. volcanii* in a similar fashion to *E. coli*. Macherey-Nagel Nucleospin (Mini) columns were used for the isolation of plamids, however due to the fact that *H. volcanii* plasmid copy-number is reduced, and the large amount of cellular debris produced in the process, the following alterations to the manufacturers guidelines were carried out:

(1) Starter Culture

30ml of starter culture was used

(2) Resuspension

Cell pellets were first resuspended in 100µl of ST buffer supplemented with 25µl of 500mM EDTA, then 125µl of the standard Macherey-Nagel resuspension buffer.

Supplemented ST buffer

1M NaCl, 20mM Tris.HCl pH 7.5, 50mM EDTA

(3) Lysis and Neutralisation

Lysis was carried out using 250µl of lysis buffer (Macherey-Nagel), followed by neutralisation via addition of 300µl neutralization buffer (Macherey-Nagel).

(4) Chloroform extraction

Chloroform extraction steps are required to separate DNA from the large amount of cellular debris to avoid column blockage. An equal volume of isoamyl alcohol:chloroform mixture was added to cell lysates and vortexed briefly to mix. Samples were then subjected to centrifugation for 1 minute at 3300 x *g*. The top aqueous layer was transferred to a fresh eppendorf tube, which was then loaded onto the miniprep column.

2.2.4 Total RNA Extraction

RNA Extraction from *H. volcanii* for RNA sequencing

Colonies were used to inoculate 5ml of Hv-YPC broth and incubated at 45°C overnight rotating. Cultures were diluted the following day to reach OD_{650} ~0.4 the following day. Cultures were subsequently diluted in 50ml Hv-YPC to reach an OD_{650} ~0.2 in sterile flasks. A this stage, cultures are treated if required. Cells were pelleted at 14,000 x *g* for 8 minutes at 4°C using a tabletop centrifuge. Cells were subsequently resuspended in 1ml acidic Trizol (Invitrogen) to lyse the cells. The mixture was transferred to a DNAse/RNAse free 1.5ml eppendorf tube and incubated at room temperature for 5 minutes. 200µl of chloroform was added and the mixture was shaken vigerously by hand for 15-20 seconds before being allowed to stand at room temperature for 5 minutes. Cell lysate was then centrifuged at 4°C for 15 minutes at 12,000 x *g*. The top aqeous layer was carefully removed and placed into a fresh, sterile, RNAse/DNase free 1.5ml eppendorf. RNA was further precipitated with the overlaying of 500 µl isopropanol (propan-2-ol) followed by inversion of the tube to mix. The sample was incubated at room temperature for 10 minutes. After centrifugation at 12,000 x *g* for 10 minutes at 4°C, the supernatant was removed and the pellet was washed with 1.2ml of ice cold 75% ethanol (RNAse free) by inverting the tube and spinning at 7,500 x g for 5 minutes at 4°C. This step was repeated and the supernatant was removed with the tube being left to dry inverted on paper towels. The pellet was then left to air dry for ~10 minutes. The pellet was resuspended in 50 µl of RNAse/DNAse free water. 5 µl of 10X Ambion TURBO DNase buffer and 1 µl of TURBO DNase (Ambion) were added before incubation for 30 minutes at 37°C. 5µl of DNAse inactivation reagent (Ambion) was then added and the sample was incubated for 3 minutes at room temperature. Samples were centrifuged for 1.5 minutes at 14000 x *g* at room temperature and the supernatant was transferred to a fresh DNAse/RNAse free tube for further analysis. The concentration and purity of the samples was measured using a Nanodrop ($A_{260/280}$ ~ 2).

cDNA synthesis, library preparation and RNA Sequencing was carried out by the Deep Seq facility (University of Nottingham). Sequencing was performed using Illumina NextSeq500 instrument with a 150 cycle mid output sequencing kit to generate ~ 10 million paired end reads per sample.

2.2.5 Nucleic Acid Manipulation

Restriction Digests

Restriction digests were carried out following manufacturer's instructions (NEB). Where possible CutSmart buffer and HF enzymes were used.

Dephosphorylation of Vector DNA

The addition of 5 units (U) of Shrimp alkaline phosphatase (rSAP) to relevant restriction digests allowed prevention of vector self-religation. Samples were incubated at 37°C for 30 minutes, prior to gel extraction.

Ligation of DNA ends

All ligation reactions were carried out using T4 DNA Ligase (NEB). Reactions generally contained a molar ratio of 3:1 insert to vector, unless otherwise stated in individual protocols. 5U of ligase per ug of DNA was added, as well as 1x T4 DNA Ligase buffer. Ligations were left overnight (~12-16hrs) at 15°C prior to ethanol precipitation.

Ethanol Precipitation

DNA was ethanol precipitated via the addition of 1:10 volume of 3M Sodium Acetate (NaAc) pH 5.3 and 2 volumes of 100% EtOH. Samples were incubated at -20°C for a minimum of one hour prior to pelleting the DNA at 4°C, 20000xg for 30 minutes. Pellets were subsequently washed in 70% EtOH, which was spun for 10 minutes, 4°C at 20000xg. The supernatant was removed and pellets left to air dry prior to resuspension in either sterile dH_2O or TE.

Nucleic Acid Quantification

DNA sample concentration and purity was assessed by measuring the absorbance at 260nm and the 260:280nm absorbance ratio respectively using the NanoDrop 2000 spectrophotometer (NanoDrop 2000, Thermo Fisher).

Agarose Gel Electrophoresis

TAE buffer and agarose powder (Sigma) was used for gel casting and running of DNA samples through an agarose gel. DNA samples were run with 1x loading dye, along with a DNA ladder (either 100bp or 1Kb, NEB). Visualisation of bands was carried out via staining with SYBR safe. Gels were run at 100V for 45 minutes unless otherwise stated.

TAE (Tris/Acetic acid/EDTA)

40mM Tris.HCl

20mM Acetic acid

1mM EDTA

5X Loading Dye

50mM Tris.HCl

100mM EDTA

15% w/v Ficoll

0.25% w/v Bromophenol Blue

0.25% w/v Xylene Cyanol FF

Agarose Gel Extraction of DNA

DNA bands were visualised using a Blue GelPic LED Box (FastGene). Bands were excised and subjected to the Macherey-Nagel DNA purification kits. Manufacturer's guidelines were followed.

PCR Amplification

Amplification of DNA via PCR was performed using Q5 HotStart Polymerase (NEB) unless stated otherwise. This polymerase is suitable for amplification of DNA templates with high G:C content, especially when used in combination with the GC enhancer (NEB). Colony PCR was also carried out using Q5 polymerase, unless stated otherwise.

Q5 PCR Reaction

200uM of each dNTP

0.5uM of each primer

1ng-1ug genomic DNA (1pg-1ng plasmid DNA)

1x Q5 reaction buffer (NEB)

1x Q5 High GC enhancer (NEB)

0.02U/µl Q5 HotStart Polymerase (NEB)

Table 2.9 **PCR Reaction Conditions**

Step	Q5 Conditions
Initial Denaturation	98°C, 30 seconds
Denaturation	98°C, 10 seconds
Annealing	Tm°C, 20 seconds
Extension	72°C, 20/30 seconds/kb for plasmid and genomic DNA respectively
Final Extension	72°C, 5 minutes

Tm Value Calculations

Tm Values were calculated using the SantaLucia (John, 1998) method, with salt correction applied as per Owczarzy (Owczarzy *et al.*, 2004). This mathematical calculation is provided as part of the NEB Tm Calculator.

Colony PCR

Colony PCR was used to screen large numbers of colonies, or provide diagnostic results of cloning without the need to extract plasmid DNA from *H. volcanii*. *H. volcanii* growing on solid agar media were lightly touched with a sterile pipette tip, ensuring only a small number of cells were transferred without disturbing the colony. The cells were resuspended in 100µl dH_2O by pipetting up and down. Subsequent boiling at 100°C lysed the cells, which were then cooled on ice. 1µl of this DNA prep was used as the template DNA in a 50µl PCR reaction with Q5 polymerase.

Dideoxy DNA Sequencing Reactions

DNA sequencing was carried out by the dideoxy chain termination method published by Fred Sanger (Sanger, Nicklen and Coulson, 1977). Analysis was performed by Deep Seq, University of Nottingham.

Oligonucleotide Synthesis

Oligonucleotides were synthesised by Integrated DNA Technologies (IDT) or Eurofins (MWG).

2.2.6 Genetic Manipulation of *Haloferax volcanii*

Generating a Genomic Clone

Either PCR or a restriction digest was used to isolate genomic clones of genes of interest from genomic DNA prepared by spooling. PCR primers were designed to contain restriction sites to facilitate downstream cloning. After PCR, the product was ran on a TAE gel and then digested with appropriate enzymes overnight. DNA was gel extracted and ligated into a suitable vector. Diagnostic restriction digests were used to confirm cloning success, as well as subsequent sequencing.

Generating a Deletion Construct via PCR

Primers were designed to flank approximately 0.5-1kb upstream and downstream of the gene to be deleted in order to generate an upstream and downstream homology arm. BamH1 restriction sites were introduced between the upstream and downstream fragments, to allow ligation of the two without the intervening gene, and relevant restriction sites were chosen for the 5' and 3' homology arms in order to ligate into pTA131. The presence of BamH1 within the middle of the homology arms allows for insertion of a *p.fdx::trpA* construct (from pTA298) to facilitate tryptophan mediated selection. Use of an Nde1 site allows for insertion of a tagged version of the gene in place of the native gene, doubling up as the new start codon.

Gene deletion / replacement

Strains deleted for *pyrE2* (Δ*pyrE2*) are transformed with a plasmid containing a deletion construct (see above) for the desired gene, either with or without *trpA*. Selection is provided via *pryE2* marker on the deletion construct, generated in pTA131, via plating on media lacking uracil. Inclusion of a *trpA* marker at the site of the gene deletion provides a more rigorous selection criteria. Typical selection media is Hv-Ca (+ additives as required) to select for the integrated plasmid, known as a "pop-in". A single colony is selected and grown, non-selectively, overnight in 5ml Hv-YPC to an A_{650} of 1.0. The culture is diluted 1:500 into fresh Hv-YPC and growth and subsequent dilution is repeated to facilitate integration across all genome copies. Removal of uracil selection by growth in Hv-YPC during this stage allows the loss of the integrated plasmid and native gene, either by recombination upstream or downstream of the deletion, this is known as "pop-out". Sucessful pop-outs are selected on Hv-Ca + 5-FOA media (+ additives as required). Addition of the *trpA* marker allows direct selection. Colonies are then restreaked and tested for the desired genotype by colony hybridization and Southern blot (figure 2.1).

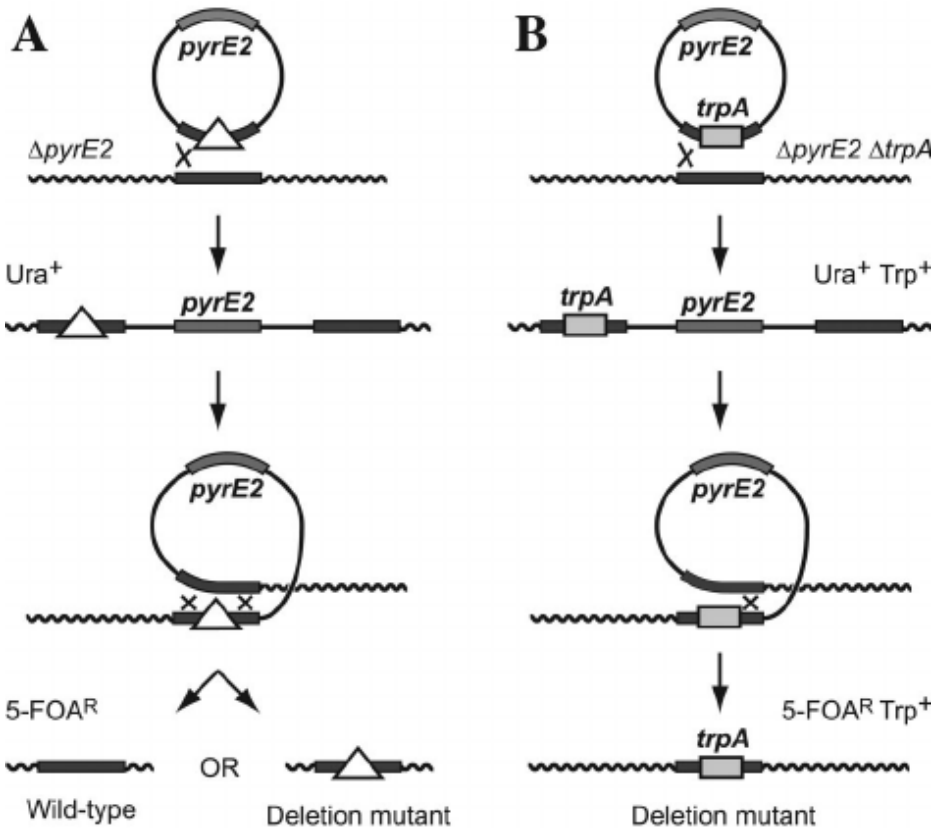


Figure 2.1| **Overview of "pop-in" and "pop-out" method of gene deletion**. (A) A plasmid carrying the *pyrE2* marker and flanking sequences of the gene to be deleted is used to transform a Δ*pyrE2 H. volcanii* strain to uracil prototrophy. Here, the crossover used to integrate the plasmid (pop-in) has occurred to the left of the deletion. Subsequent loss of the plasmid by intrachromosomal crossing over can occur on the left of the deletion, restoring the gene to wild type, or on the right of the deletion, resulting in the desired mutant. In either case the cell is rendered auxotrophic for uracil and is therefore resistant to 5-FOA by virtue of its inability to convert this compound to the toxic analog 5-fluorouracil. (B) The gene is replaced with the *trpA* marker, and the plasmid is used to transform a Δ*pyrE2* Δ*trpA H. volcanii* strain to prototrophy for uracil and tryptophan. Loss of the plasmid by crossing over on the right of the deletion, resulting in a *trpA*-marked mutant, can be selected in one step (Allers, H.-P. Ngo, *et al.*, 2004).

2.2.5.1 Screening of Genotypes

Screening for relevant pop-in and pop-out steps during gene deletion is done in a variety of ways. If a selectable marker was used in generation of the strain (such as *trpA*), the colonies can be plated directly onto selective media. The main problem is that alone, this method is not sufficient to confirm if a specific gene has been deleted in *H. volcanii*, this is predominantly due to the ploidy of *H. volcanii* being so high (approximately 20 genome copies in exponential phase). Therefore, as mentioned above, strains could be merodiploid and contain both wild-type and deletion alleles on different chromosomal copies. Even a single copy of the wild-type allele will cause problems. Selection of homozygous strains therefore requires additional methods of confirmation with high levels of sensitivity. The methods used are colony hybridisation and Southern blotting.

Colony Lift

Potential candidates for gene deletions are chosen and patched onto Hv-YPC plates using sterile toothpicks. Patches are grown at 45 $^{\circ}$C for three days. Colonies are then lifted from the plate be placing a cut-out of GE Healthcare Amersham Hybond N^{+} charged membrane into the surface for 1 minute. Cell lysis is achieved by transferring the membrane colony side up onto Whatman paper (soaked in 10% SDS) and incubating for approximately 5 minbutes. The membrane is transferred to Whatman paper soaked in denaturing solution for the same length of time to denature proteins and DNA. Neutralisation is carried out by transferring the membrane to Whatman paper soaked in neutralising solution for the same length of time, this step was repeated twice. The membrane is washed for 30 seconds in 2X SSPE and air dried.

20X SSPE

3M NaCl

230mM NaH_2PO_4

32mM EDTA

pH 7.4

Denaturing Solution

1.5M Nacl

0.5M NaOH

Neutralising Buffer

1.5M NaCl

0.5M Tris.HCl

1mM EDTA

DNA Crosslinking to Membrane

Crosslinking of DNA to a membrane is achieved using a UV Crosslinker and irradiating with 3000J/m^2 UV.

Southern blot by Vacuum Transfer

Genomic DNA (*H. volcanii*) was purified and digested. Digested DNA was resolved on a 0.75% TAE gel (200ml) at 50V for 16hrs with buffer circulation. The gel was post-stained with 0.5ug/ml EtBr for 30 minutes. The DNA was acid-nicked in 0.25M HCl for 15 minutes, washed in dH_2O for 10 minutes and denatured in denaturing solution for 45 minutes. A 15x25cm GE healthcare Amersham Hybond-XL membrane was soaked in dH_2O before soaking in denaturing solution. Vacuum transfer was carried out using a Vacugene XL gel blotter and pump (Pharmacia Biotech) for 1hr at 40mBar. The membrane was briefly washed for 30 seconds in 2X SSPE and air dried before crosslinking by UV irradiation at 3000J/m^2.

Denaturing Solution

1.5M Nacl

0.5M NaOH

20X SSPE

3M Nacl

230mM NaH_2PO_4

32mM EDTA

pH 7.4

Hybridisation

Membranes were soaked in prehybridisation solution at 65°C with rotation for 3 hours. Radiolabelled DNA probes were made with 50ng of DNA and 0.74 mBq of α-^{32}P dCTP (Perkin Elmer). DNA was denatured at 100°C for 5 minutes, incubated with the radioisotope and HiPrime random primer hexanucleotides (Roche) at 37°C for 15 minutes. Probes were purified on a BioRad P-30 column and was added to 10mg/ml salmon sperm DNA to a final volume of 500µl. The DNA mix was denatured at 100°C for 5 minutes followed by quenching on ice.

For Southern blots, 1µl of 1ug/ml 1kb ladder was also included in the radiolabelling reactions for visualisation. The prehybridisation solution was replaced with 30ml hybridisation solution at 65°C with rotation overnight. Membranes were then washed twice in low stringency wash solution for 10 and 30 minutes respectively. Two further washes of 30 minutes were carried out using high stringency wash solution. Membranes were air dried before being wrapped in Saran wrap and exposed to a phosphoimager screen (Fujifilm BAS Cassette 2325) for 24 hours. The screen was then scanned using a GE Healthcare Typhoon.

Real-time PCR (qPCR)

Real time quantitative PCR was carried out using Applied Biosystems (ABI) 7500 Fast qPCR machine. Samples were diluted (10^{-1} – 10^{-4}) before being added to SYBR Green master mix (Thermo Fisher) with the relevant primer combinations in a 96-well plate. Appropriate cycling conditions were selected.

ΔΔCt calculations (qPCR)

Fold gene expression levels were calculated using the ΔΔCt method. First, the mean Ct values for both the endogenous control and the target were calculated, after which ΔCt was calculated by subtracting the relevant endogenous sample Ct from each respective target Ct. ΔΔCt values were then calculated by substracting the control ΔCt value from each ΔCt value for each sample. Finally, fold change in expression was calculated using the following equation:

$$Fold\ change\ in\ expression = 2^{-\Delta\Delta \mathrm{Ct}}$$

2.2.7 Phenotypic characterization of *Haloferax volcanii*

Ultraviolet (UV) Irradiation Sensitivity Assay

5ml of Hv-YPC was inoculated with a single colony and grown at 45°C overnight with 8rpm rotation. The culture was then diluted into fresh 5ml Hv-YPC broth and grown to an A_{650} of 0.4 (approximately 10^8 cells/ml). A range of serial dilutions was then carried out in 18% SW according to expected killing (Delmas *et al.*, 2009), after which duplicate 20µl aliquots were spotted into relevant agar media and allowed to dry. Plates were then exposed to UV light (1J/m^2/sec) for different lengths of time. Plates were then protected from visible light to prevent photo-reactivated DNA repair via incubation in a black plastic bag. Plates were incubated for 3-6 days at 45°C. Colonies were then counted and survival fractions were calculated with respect to a no UV irradiated control plate.

Gamma Irradiation Sensitivity Assay

5ml of Hv-YPC was inoculated with a single colony and grown at 45°C overnight with 8rpm rotation. Cultures were then diluted in fresh Hv-YPC broth and left to grow to an A_{650} of 0.4 (approximately 10^8 cells/ml). A range of serial dilutions were carried out in 18% SW according to expected killing (Delmas *et al.*, 2009) and 20µl aliquots were spotted onto relevant agar media, left to dry, and irradiated using the Caesium-137 gamma source (March 2021 output – 4.516 Gy/min). Plates were then incubated at 45°C for 3-6 days in a black plastic bag. Colonies were then counted and survival fractions were calculated with respect to a non-irradiated control plate.

Pairwise Competition Assay [Gamma radiation]

Single colonies of relevant strains were grown in 5ml Hv-Ca+Ura overnight at 45°C, rotating at 8rpm. Cultures were then diluted and left to grow in fresh media until an A_{650} of 0.4 was reached. Cultures were then diluted 10^{-4} fold, liberating

approximately 10^3 cells/ml. Fresh Hv-Ca+Ura+Trp media was inoculated with 100µl of diluted cell culture, generating a 1:1 ratio of strains. This mixed culture was then incubated until an A_{650} of 0.4. Equal volumes of mixed culture were then aliquoted into four separate 4ml tubes and treated with a dose of gamma radiation as per table 2.10.

Tubes were then removed after the appropriate time had lapsed and each culture was then diluted as per table 3.1 and 100ul plated on relevant media. Some of the remaining culture was then diluted 1:2 into fresh Hv-Ca+Ura+Trp media and placed in the incubator to grow at 45°C, rotating at 8rpm for two hours, after which plating was repeated. Cultures were then diluted further, as per before, and left for another two hours, after which a final 1:1000 dilution was carried out and the culture left for approximately 20hrs. Relevant plates were then selected and ratio of strains determined via X-Gal.

Table 2.10. **Gamma radiation dosage**

Gamma Dosage	Length of irradiation required
No irradiation	None
750 Gy	166 minutes
1000 Gy	221 minutes
1500 Gy	332 minutes

Table 2.11. **Gamma radiation plate dilutions**

Gamma Dosage	Dilutions plated	Reference
No irradiation	10^{-5}, 10^{-6}, 10^{-7}	(Delmas *et al.*, 2009)
750 Gy	10^{-4}, 10^{-5}, 10^{-6}	
1000 Gy	10^{-3}, 10^{-4}, 10^{-5}	
1500 Gy	10^{-2}, 10^{-3}, 10^{-4}	

Pairwise Competition Assay [UV radiation]

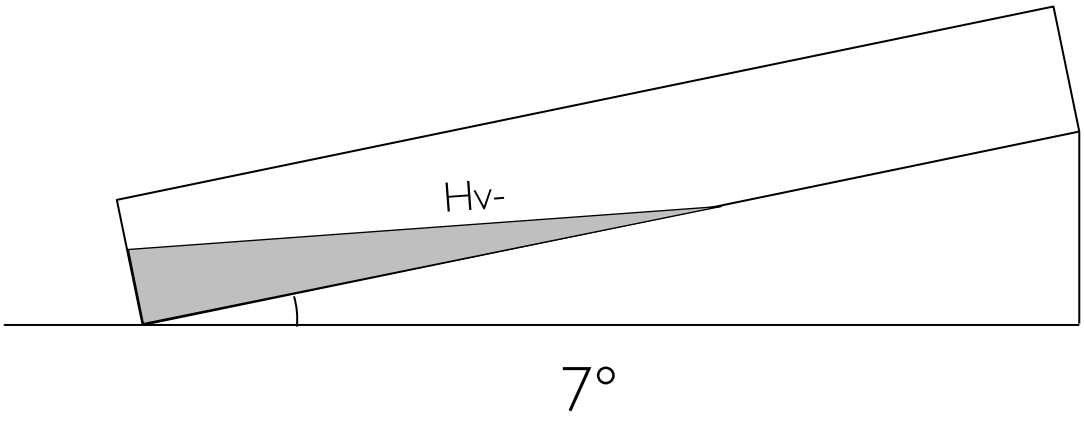
Single colonies of relevant strains were grown in 5ml Hv-Ca+Ura overnight at 45°C, rotating at 8rpm. Cultures were then diluted and left to grow in fresh media until an A_{650} of 0.4 was reached. Mixed culture was then set up using 100µl of a 10^{-4} dilution of each culture in Hv-Ca+Ura+Trp and left to reach A_{650} of 0.4. 1ml aliquots were placed in hydrophilic cell culture dishes and irradiated using 300J/m^2 UV. Some culture was left unirradiated. Serial dilutions of each culture was then made in 18% SW and 100ul was plated of each appropriate dilution on Hv-Ca+Ura+Trp plates (table 2.12). After plating, cultures were diluted appropriately to maintain the A_{650} of 0.4 across the different time points. Cultures, after dilution, were then placed in the 45°C incubator, rotating at 8rpm for varying amounts of time, up to 20 hrs. Cultures after each regeneration time were appropriately diluted as per table 3.2, left to grow for 4-5 days in a black plastic bag and then sprayed with X-Gal to determine the ratio of cultures.

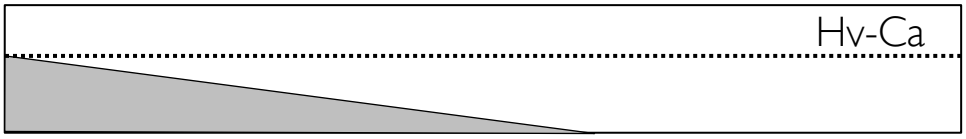
Table 2.12. **UV Dose and dilutions for UV competition assay.**

UV Dosage	Dilutions plated	Reference
No irradiation	10^{-5}, 10^{-6}, 10^{-7}	(Delmas *et al.*, 2009)
300 J/m^2	10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}	

Tryptophan Gradient Plates

In order to create media containing a gradient of Tryptophan across the plate, fading to no tryptophan on one side of the plate, plates were poured with 17ml Hv-Ca+Trp (to maintain episomal selection of genomic library constructs). If required, uracil can be added. Tryptophan was added to a final concentration of 50ug/ml, 0.25mM. Tryptophan wedges were created by pouring on a 7° slant to form a tapered wedge (figure 2.2). Once set, the plate was placed on a flat surface and the wedge was covered with 50ml Hv-Ca agar (+ uracil if required). Tryptophan will therefore diffuse through the wedge into the upper layer of media lacking Tryptophan generating a gradient across the plate (SZYBALSKI and BRYSON, 1952, Hawkins *et al.*, 2013).





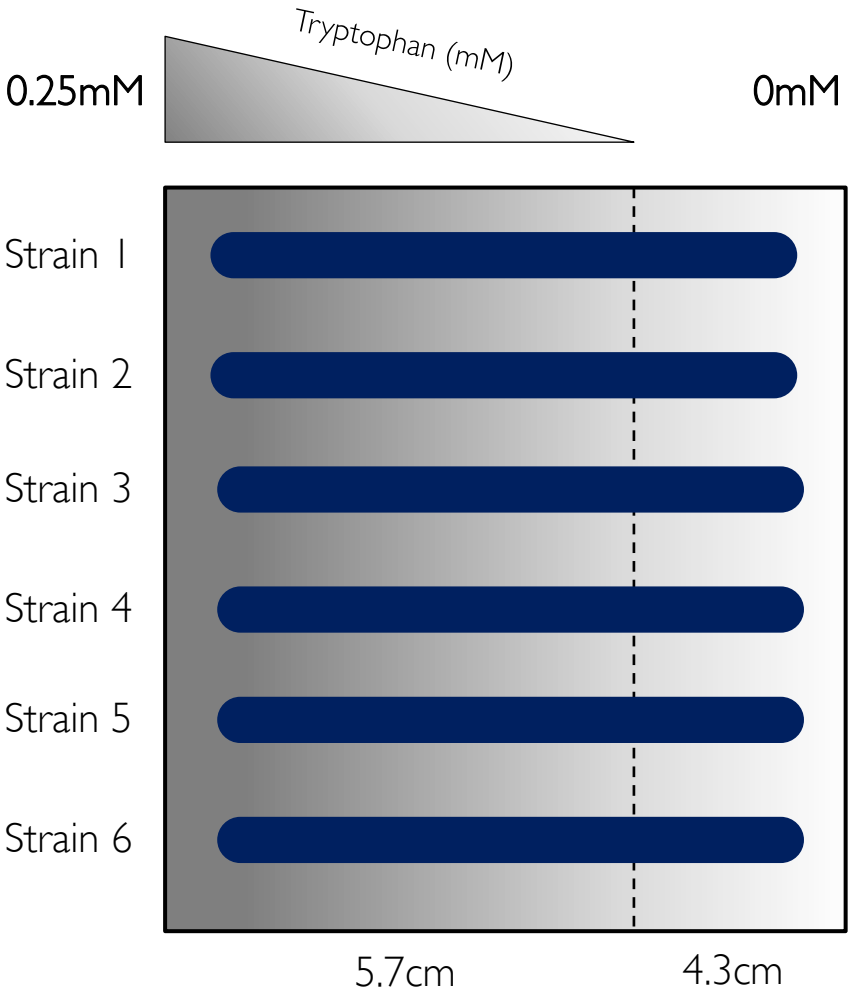


Figure 2.2| **Tryptophan Gradient Plate Preparation.** Plates were poured with 17ml Hv-Ca+Trp at 0.25mM on a 7° slant to form a wedge. Once set, 50ml Hv-Ca was poured over the wedge to create an even top. Strains were painted in 'lanes' across the gradient.

5ml of respective cultures of *H. volcanii* strains to be analysed were grown overnight in Hv-YPC or Hv-Ca (+ relevant additives) at 45°C, rotating at 8rpm. Cultures were then diluted and grown for a subsequent overnight under the same conditions until they reached an optical density of approximately A_{650} 1.0. Serial dilutions of cultures were then made in 18% salt water until a final dilution of 10^{-4} was reached. Autoclaved paintbrushes (The Range) were lubricated first using 18% salt water, after which it was dipped into the diluted culture and painted in one direction evenly across the Trp gradient plate. The paintbrush was then dipped in the same culture and painted across the plate, along the same lane, but in the opposite direction to ensure an even and equal streak across the plate.

Once dry, the plates were incubated at 45°C for 5 days. If irradiation by UV was required, plates were left to grow for a couple of hours, after which they were irradiated before being placed back into the incubator in opaque black bags to minimise photolyase activation and DNA repair via photoreactivation.

Growth Assay

The standard growth rate of cultures can be determined using an Epoch 2 Microplate Spectrophotometer (BioTek). Cultures were prepared in 5ml Hv-YPC or Hv-Ca broth (with required additives) and grown to mid-late exponential phase, corresponding to an A_{650} of between 0.2 and 0.8. Cultures were diluted and grown again to mid-late exponential phase. Serial dilutions of cultures were then made before loading 150ul of culture into a 96-well plate (Corning), in triplicate, alongside appropriate blanks. Blank samples were loaded in the wells around the entire perimeter of the 96-well plate to mitigate any evaporation of sample and formation of salt crystals. The plate was sealed around the edge using two layers of microporous tape (Boots) and incubated at 45°C with double orbital shaking at 1000rpm for 72 hours in the Epoch plate reader. Readings at A_{600} were taken every 15 minutes and converted to a 1cm pathlength by dividng the raw A_{600} value by 0.14. Blank readings were taken into account by subtracting the average blank reading from all raw A_{600} values prior to dividing by 0.14. If the generation time was required, it is calculated by plotting the growth on a $\log_2$ scale and using the following equations:

$$G = \frac{t}{n} \qquad n = \frac{logx - logy}{log2}$$

G = generation time
t = time
n = number of generations
x = end A_{650}
y = start A_{650}

Glass Bead Plating Assay

In order to minimise ring formation as an artefact of spotting cultures onto solid media and irradiating with UV (see results), a more specialised approach was attempted to generate a more high-throughput method of culture plating. For this, a three-dimensional bead-guide was designed based on (Hamilton *et al.*, 2002). This had been used as part of a previous PhD student's thesis (Adam Collins and Alan Huett). The bead-guide was manufactured by Medical Manufacturing (University of Nottingham) from polytetrafluroethylene (PTFE). The bead-guide is placed onto pre-set agar plates contained within OmniTray plates (Nunc). Sterile 5mm glass beads (Sigma) were then dispensed into each lane via use of a 5mm bead dispenser (TissueLyser). 20µl samples of culture were then dispensed onto each bead and beads were then rolled along the lanes created by the bead-guide, resulting in homogenous spreading of the 20µl sample agross the agar. Bead-guides were washed with 70% ethanol between uses.

Archaeal Microscopy Sample Preparation and Imaging

Single colonies of relevant strains were inoculated into 5ml of relevant media (Hv-YPC) and left at 45°C, rotating at 8rpm overnight. Cultures were then diluted and left to grow to an A_{650} of 0.6. Glass slides were cleaned with ethanol and Kimwipes, dried and placed on a flat, clean surface. Agarose pads were placed onto the centre of the slides and left for 45 minutes to set. 1ml of culture was then spun at 3300xg for 8 minutes in a 2ml round-bottomed tube and resuspended in the same volume of 18% SW. Staining was carried out via addition of Acridine Orange (Invitrogen), or Concanavalin A (ThermoFisher). 20µl of 100µg/ml AO stain, or 100µl of 1mg/ml ConA, was added to 1ml of culture. Cultures were left to stain for approximately two minutes at room temperature. 10µl of culture was then spotted onto the centre of the agarose pad. Imaging was carried out via the University of Nottingham School of Life Sciences Imaging Service (SLIM) microscopes, specifically the Zeim 200M using a FITC filter.

Archaeal Fixing - Agarose Pad Preparation

Agarose pads were prepared by dissolving 250mg of electrophoresis grade agarose in 10ml of dH_2O. Whilst hot a P-1000 pipette was used to transfer 400µl of molten agarose to 600µl of 30% buffered salt water, which has been pre-heated to 70°C. Agarose pad solution was left to shake at 200rpm in an Eppendorf table-top shaker (Eppendorf 5355 Thermomixer). 100µl of agarose was then pipetted onto the centre of a clean glass microscope slide. Two support slides, wrapped in masking tape, were placed either side of the agarose pad, with another clean slide placed over the top of the agarose pad to ensure it set flat (figure 2.3).

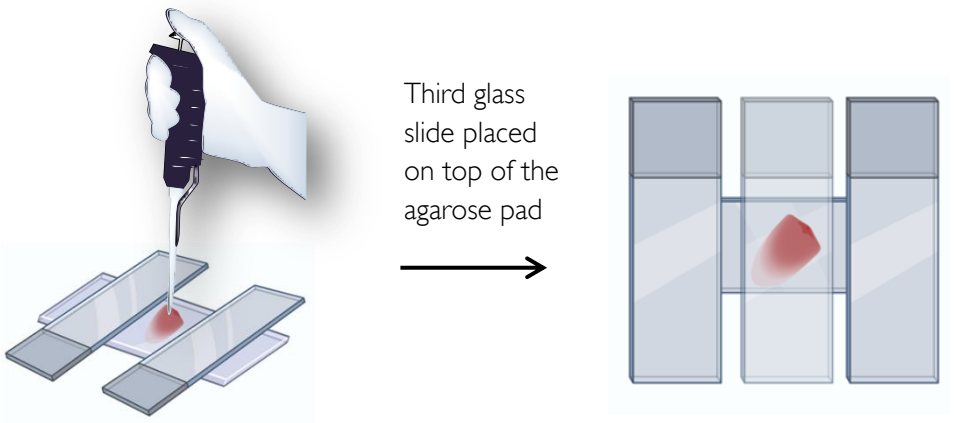


Figure 2.3| **Method for agarose pad preparation**. Image generated using BioRender.

Archaeal Sphere Formation

Relevant strains were used to inoculate 5ml of relevant growth media with a single colony and left to grow at 45°C, overnight (12-16hrs), rotating at 8rpm. Cultures were diluted into fresh media and allowed to grow to A_{650} of 0.6. The cultures were pelleted at 3300xg, 8 minutes at room temperature and the supernatant discarded. Pellets were resuspended in either fresh growth media, or the same volume of TL buffer (Fendrihan *et al.*, 2012). Cultures were left undisturbed at 37°C for approximately two hours to form spheres. Cultures were spun as before, and pellets resuspended in fresh media. This culture was then used as per the microscopy protocol, resuspended in 18% SW, stained with Acridine Orange and imaged using a FITC filter (School of Life Sciences Imaging Services, SLIM).

Mitomycin C (MMC) Sensitivity Assay

5ml of Hv-YPC was inoculated with a single colony and grown overnight at 45°C with 8 rpm rotation. The culture was diluted into 5ml of fresh Hv-YPC and grown to an A_{650} of 0.4. Cells were then serially diluted (100 – 10-8) in 18% salt water and 20µl of each dilution was spotted onto Hv-YPC agar containing 0 – 0.1µg/ml MMC. Plates were air dried for approximately 1hr before incubation at 45°C for 4-5 days. During this period, colonies were counted and survival fractions were calculated relative to an untreated control.

Chapter 3: Construction and Screening of Overexpression Genomic Library

3.1 Background

Effect of radiation on the DNA template

In order to investigate the feasibility of Archaeal life, either in years gone by, or currently, on different planets, one must consider how Archaea are able to adapt to survive under such harsh conditions. A key theme underpinning such resistance is the ability to repair DNA lesions, which are ever-present both terrestrially and extra-terrestrially. Typical insults to the viability of life beyond Earth arise from DNA damage inflicted primarily, although not exclusively, from radiation stress, the predominant forms of which are Ultra-Violet (UV) radiation and ionizing radiation, such as gamma radiation. UV induced lesions tend to be in the form of cyclobutane pyrimidine dimers (and to a lesser extent, 6,4-photoproducts) in the DNA template, whereas ionizing radiation is generally more harsh and leads to formation of single and double-stranded DNA breaks and oxidative damage. Both forms of damage lead to significant issues for DNA replication and transcription, either directly, or indirectly via reactive oxygen and nitrogen species (free radicals) (White and Allers, 2018). It is therefore a necessity for Archaeal species to contain hyper-efficient DNA repair systems as well as lesion avoidance mechanisms tailored to radiation stress.

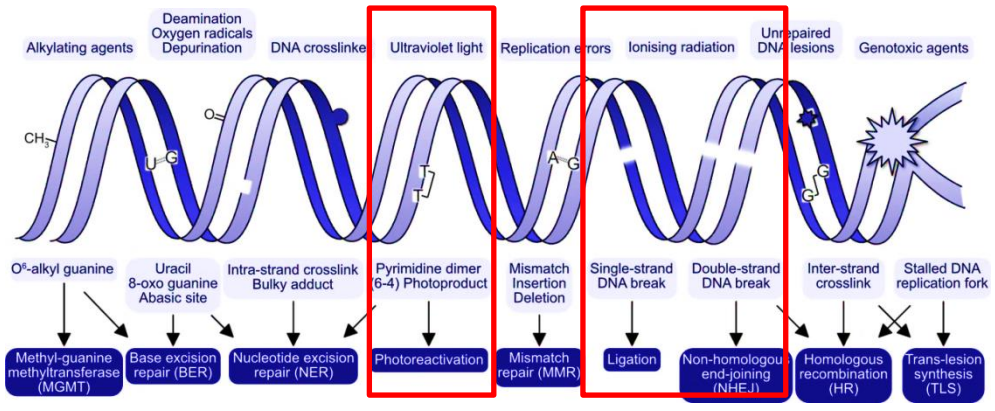


Figure 3.1 | **Overview of DNA lesions and repair mechanisms**. Relevant lesions highlighted. Adapted from (White and Allers, 2018).

3.2 Aims and Objectives

It will be of significant interest to determine the genetic mechanisms underpinning resistance to radiation stress in *Haloferax volcanii*. Furthermore to identify any specific genes or biochemical pathways involved in resistance to ionizing and UV radiation, as well as genes underpinning resistance to both individually.

The objectives of this chapter were to:

- Determine cell survival and viability under both UV and gamma radiation comparing a hyper-resistant strain to a wild-type strain to ensure genome library is screened with the correct dosage of radiation.

- Carry out a pairwise competition assay using resistant and wild-type strains to map respective growth of strains after radiation is applied.

- Generate an overexpression genome library using *Haloferax volcanii* strain H26.

- Place each genomic fragment downstream of a strong, inducible, promoter, the promoter used here is the tryptophan inducible promoter (PtnaA).

- Transform the library into *H. volcanii*.

- Screen the library under both UV and gamma radiation.

- Determine if Hel308 overexpression could act as a positive control

3.3 Materials

3.3.1 Strains

[ ] Indicate integrated plasmid DNA

{ } Indicate episomal plasmid DNA

Table 3.1 ***Haloferax volcanii*** **strains used in this chapter**

Strain	Parent	Genotype
H26	H18	Δ*pyrE2*
H1206	H1202	Δ*pyrE2*, Δ*mrr*
H645	H640	Δ*pyrE2*, bgaHa, Δ*mre11-rad50*
H1895	H1892	Δ*pyrE2*, Δ*hdrB,*Δ*mrr,*Nph-pitA, cdc48d-Ct, Δ*pilB3C3*

3.3.2 Plasmids

Table 3.2 **Plasmids used in this chapter**

Name	Use
pTA908	Overexpression vector with pHV2 origin and tryptophan inducible promoter. Shuttle vector containing ColE1 origin and Amp^R/pyrE2 selectable markers.
pTA927	Overexpression vector with pHV2 orgin and tryptophan inducible promoter. Shuttle vector containing ColE1 origin and AmpR/pyrE2 selectable markers. t.Syn terminator sequence added.
pTA1231	Overexpression vector for Hel308

3.3.3 Oligonucleotides

Table 3.3 **Oligonucleotides used in this chapter**

Name	Use	Sequence
pBSF2	Forward primer for colony PCR of Hel308 overexpression vector, pTA1231.	5'-TTAAGTTGGGTAACGCCAGGG-3'
Hel308R	Reverse primer for colony PCR of Hel308 overexpression vector, pTA1231.	5'CGACGACGCAGGTGAGTTGG-3'

pTA908 (tryptophan inducible overexpression shuttle vector)

pTA908 is a tryptophan inducible overexpression vector capable of replication in both *H. volcanii* and *E. coli* owing to both pHV2 and ColE1 origins of replication. ApaI and ClaI sites have been inserted into the MCS. This vector was created by Thorsten Allers in 2008.

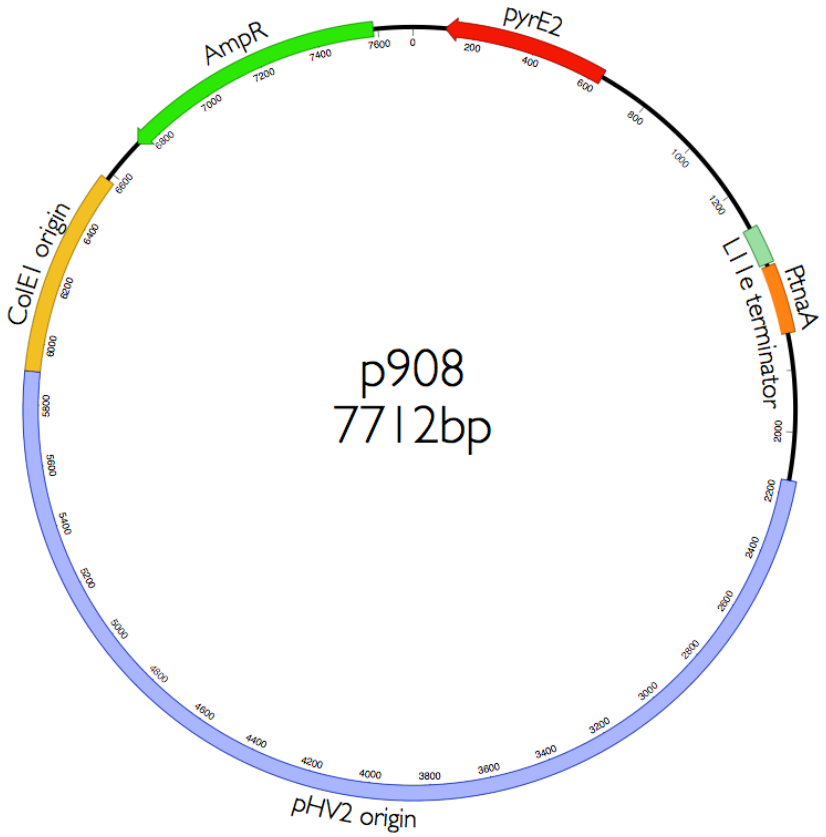


Figure 3.2 | **pTA908**. Constructed by Thorsten Allers in 2008 for overexpression of genes downstream of tryptophan inducible promoter.

pTA927 (tryptophan inducible overexpression vector with t.Syn terminator sequence)

pTA927 is a tryptophan inducible overexpression vector capable of replication in both *H. volcanii* and *E. coli* owing to both pHV2 and ColE1 origins of replication. The t.Syn transcriptional terminator sequence has been added. Created by Thorsten Allers in 2008 using pTA908 backbone (Allers *et al.*, 2010a).

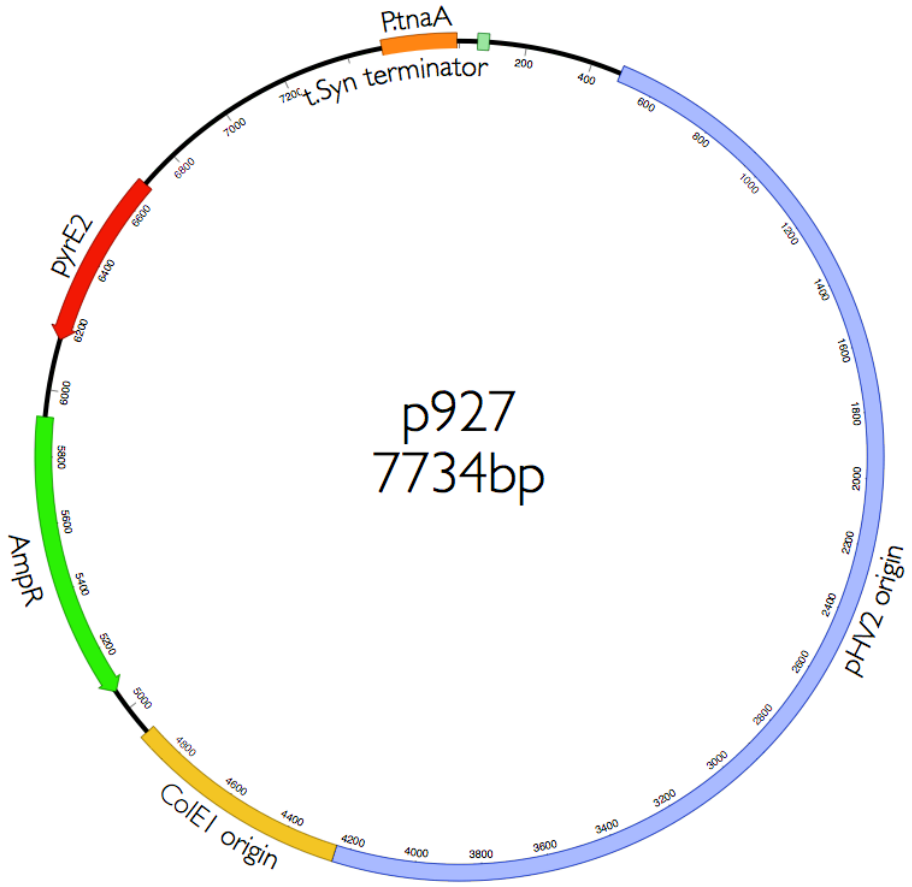


Figure 3.3 | **pTA927**. Constructed by Thorsten Allers in 2008 for overexpression of genes downstream of tryptophan inducible promoter. Constructed from p908 backbone.

pTA1231 (overexpression vector for hel308)

pTA1231 is an overexpression vector for the Hel308a gene. The Hel308a coding sequence is placed immediately downstream of the p.*tnaA* promoter, allowing tryptophan-mediated expression, useful as Hel308 is toxic in *E. coli*, and high expression levels are toxic in *H. volcanii*. Both ColE1 and pHV2 origins of replication are present, allowing for transfer of this construct between hosts as applicable. An N-terminal His tag is present, allowing for protein purification as required.

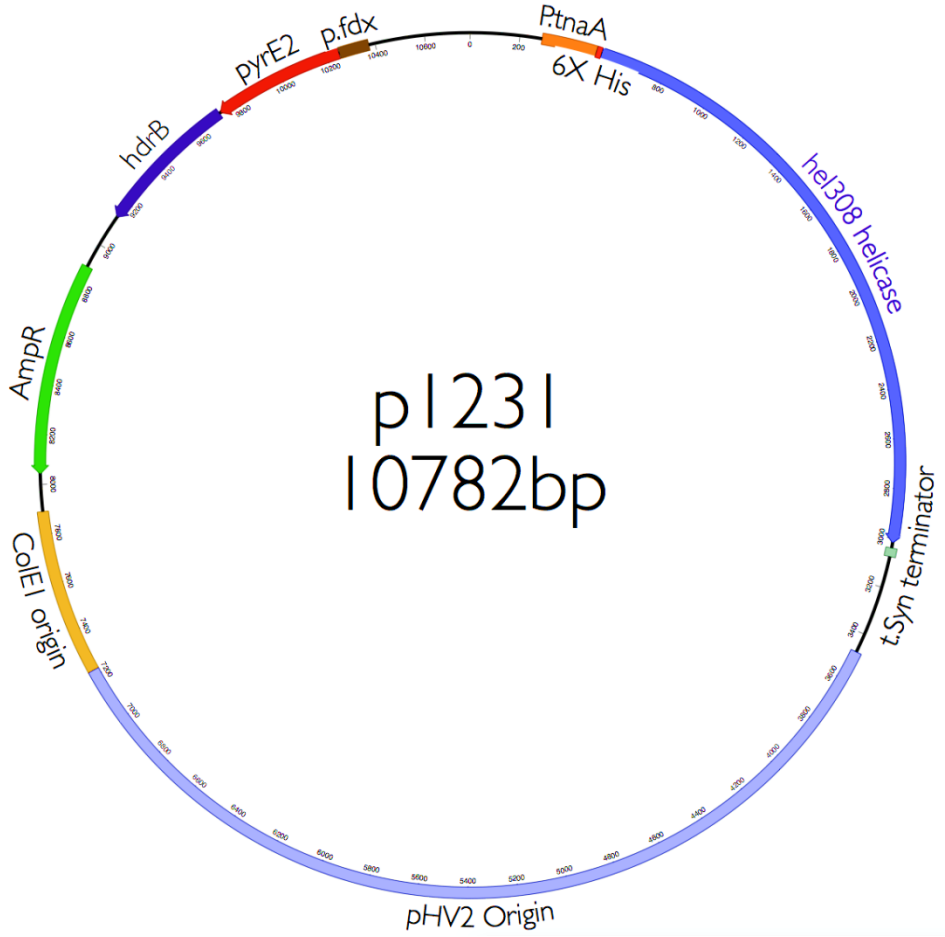


Figure 3.4 | **pTA1231**. Constructed by Thorsten Allers in 2010. Hel308a overexpression vector used for inducible overexpression of N-terminal His tagged Hel308.

3.4 Results

The objectives of this chapter are to carry out initial calibration assays using both UV and gamma radiation to determine the optimal dosage for screening a genomic library, after which a genomic library can be generated and screened. Competition assays were carried out in liquid culture to confirm the timescale over which further rounds of irradiation need to be carried out, if at all. Analysis of the colonies resulting from this screen via survival assays under UV stress confirms potential candidates.

3.4.1 Generation of genomic overexpression library

Haloferax volcanii genomic DNA (strain H26) was first isolated via spooling. A pilot partial digest was carried out using 1µg of gDNA with varying enzyme units, produced via dilution. Digests were carried out in CutSmart buffer for 30 minutes. Owing to frequent cut sites within the G/C rich genome of *H. volcanii*, Aci1 (C^CGC) was used. Additional Bovine Serum Albumin (BSA) was added to the reactions to maximise enzyme efficiency at low concentrations.

Table 3.3.1. **Dilution of Aci1**

		Use 1U/µL and 1X BSA/buffer to set up		
	1U/µL	**0.5U/µL**	**0.2U/µL**	**0.1U/µL**
Enzyme (10U/µL stock)	5	25	10	5
CutSmart	5	2.5	4	4.5
BSA (100x)	0.5	0.25	0.4	0.45
SDW	39.5	22.25	35.6	40.05
Total	50	50	50	50

Pilot digest was set up as per table 3.4 and run on a 0.8% TAE agarose gel for 2.5 hours at 45V (Figure 3.5a).

Table 3.4. **Pilot restriction digest**

Amount of enzyme (µL)	**1U/µg**	**0.7U/µg**	**0.5U/µg**	**0.2U/µg**	**0.1U/µg**
AciI (diluted above)	1	1	1	1	1
gDNA (461ng/µl)	2.1	2.1	2.1	2.1	2.1
CutSmart (10X)	5	5	5	5	5
Extra BSA (100X)	0.5	0.5	0.5	0.5	0.5
SDW	41.4	41.4	41.4	41.4	41.4
Total	50		50	50	50

This titration showed the best DNA fragment distribution at 0.5U/µg, so an up-scaled reaction was then carried out, using 0.5U/µg of Aci1 and 50ug of genomic DNA. Bands ranging between 3kb and 5kb were excised from the gel and extracted. The average fragment length is therefore 4kb.

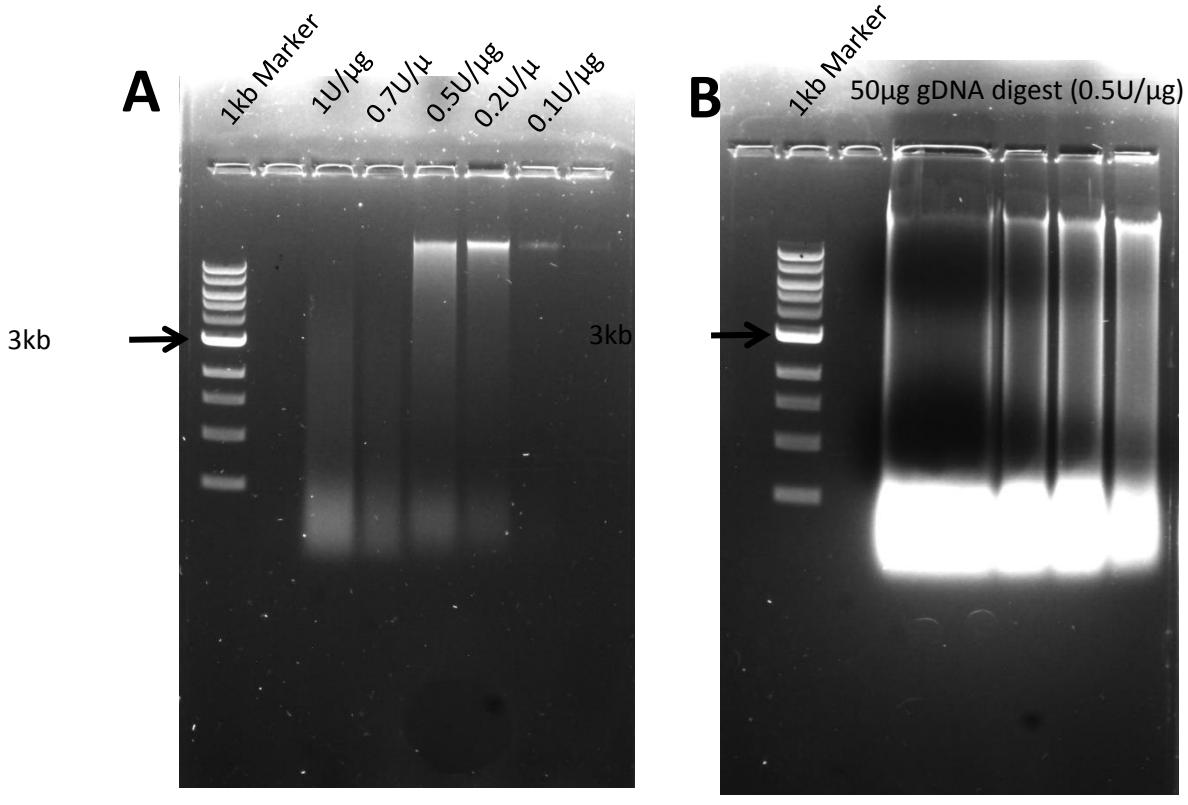


Figure 3.5 | **A) Agarose gel electrophoresis of pilot partial digest of *H. volcanii* genomic DNA.** 1ug of genomic DNA was digested with varying concentrations of enzyme units and ran for 2.5hrs at 45V. In this case strain H26 was used. **B) Upscaled digest of H26 genomic DNA.** 3kb-5kb region of gel extracted as per protocol listed above.

Once extracted, genomic DNA fragments were then ligated to vector pTA908, cut with compatible enzyme Cla1 (NEB). pTA908 failed to carry out ligation, both with fragments inserted and to itself as a control, therefore vector pTA927 was used instead. Looking into this further demonstrated a lack of self-ligation in pTA908. pTA927 was cut to completion overnight at 37°C and the 5' ends were phosphatase treated. The cut vector was ran on a 1% TAE agarose gel and extracted (figure 3.6).

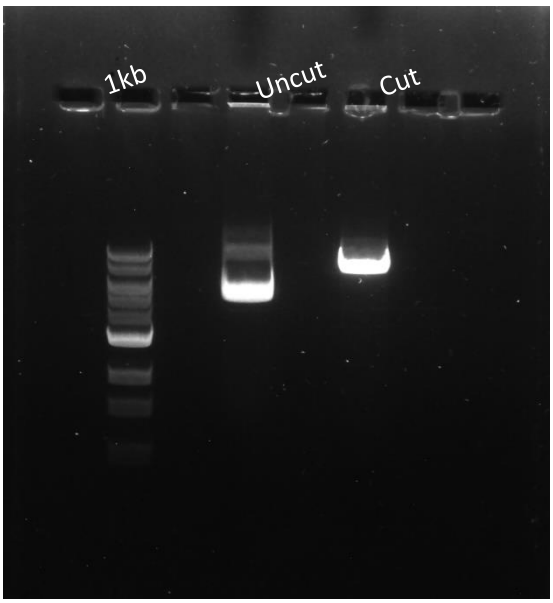


Figure 3.6| **Agarose gel electrophoresis of pTA927 cut with Cla1**. Band extraction conducted as per manufacturer's protocol.

After genomic library DNA and vector DNA had been purified, a ligation reaction was set up using T4 DNA Ligase in T4 DNA Ligase buffer. The molar ratio of insert to vector was 18.5:1 to ensure maximum genomic library fragment coverage. The reaction was left at 15°C for 48hrs to fully ligate, after which the sample was ethanol precipitated.

Firstly, the genome library was transformed into *E. coli* (Xl1-Blue) using standard electroporation procedure, for which over 6000 clones were produced. This number is well over the 4,602 clones required by the Carbon Clarke formula (Clarke and Carbon, 1976) , thus achieving a 99% chance of isolating any particular 4kb fragment from the genome in the library.

The cloning above was carried out for a second time, after discovery of an updated calculation which is specific for genomic library creation (Zilsel, Ma and Beatty, 1992). Probability was divided by two to take into account lack of directional cloning:

$$n = \frac{\ln(1 - \Phi^{f})}{\ln(1 - f)}$$

$$n = Number\ of\ recombinants$$

$$f = \frac{fragment\ size\ (kb)}{Size\ of\ genome\ (kb)}$$

$$\Phi = Probability$$

Given the updated equation above, ~25,000 *E. coli* cfu were required to be obtained in order to have 99% probability of obtaining at least one copy of each gene. Just over 25,000 *E. coli* colonies were obtained.

Overnight cultures were prepared in selective LB broth, after which glycerol stocks were then prepared from each plate and stored at -80°C as well as isolating plasmid DNA via Midiprep column purification (Macherey-Nagel).

3.4.2 UV Radiation Survival Assays

Haloferax volcanii strains H1206 and H645 were grown in 5ml of Hv-YPC overnight at 45°C, rotating at 8rpm. Strain H1206 acted as a wild-type control strain and H645 as a UV resistant strain, mimicking resistant constructs within the genome library to be screened at a later date. Strain H645 is deleted for mre11/rad50 which renders this strain hyper-resistant to UV radiation stress (Delmas *et al.*, 2009). In spite of this being counter-intuitive, this is thought to occur via increasing levels of homologous recombination, as mre11/rad50 appears to restrain this. As Archaea are heavily recombinogenic, this was surprising. However, due to their extensive genome copy number, the deletion of mre11/rad50 genes may act to maintain acceptable levels of genomic stability in *Haloferax volcanii*, by increasing the amount of genome-wide homologous recombination events (Delmas *et al.*, 2009).

Strains were independently diluted and grown to an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Strains were serial diluted according to expected cell death and 20µl aliquots were spotted in duplicate on Hv-YPC agar plates. Plates were then immediately subjected to varying amounts of UV radiation (0J/m^2, 60J/m^2, 90 J/m^2, 120 J/m^2, 180 J/m^2, 300 J/m^2). After irradiation, plates were placed in a black bag to avoid visible light mediated DNA repair mechanisms becoming activated and grown for 4 days, after which the survival fraction was calculated as a percentage of respective strain on the no UV plate (Figure 3.7a).

3.4.3 Gamma Radiation Survival Assays

Haloferax volcanii strains H1206 and H645, representing wild-type control and a hyper-resistant strain respectively (Delmas *et al.*, 2009), were grown in 5ml Hv-YPC overnight at 45°C, rotating at 8rpm. Cultures were subsequently diluted and grown to an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Strains were serially diluted according to expected cell death and 20µl aliquots were spotted in duplicate on Hv-YPC agar plates. Plates were then immediately subjected to varying amounts of gamma radiation (0 Gy, 750 Gy, 1000 Gy, 1500 Gy). After irradiation, plates were placed in a black bag to avoid visible light DNA repair system activation. Colonies were grown for 5 days, after which they were counted, survival fractions were generated as a percentage of respective strain on the no irradiation control plate (Figure 3.7b).

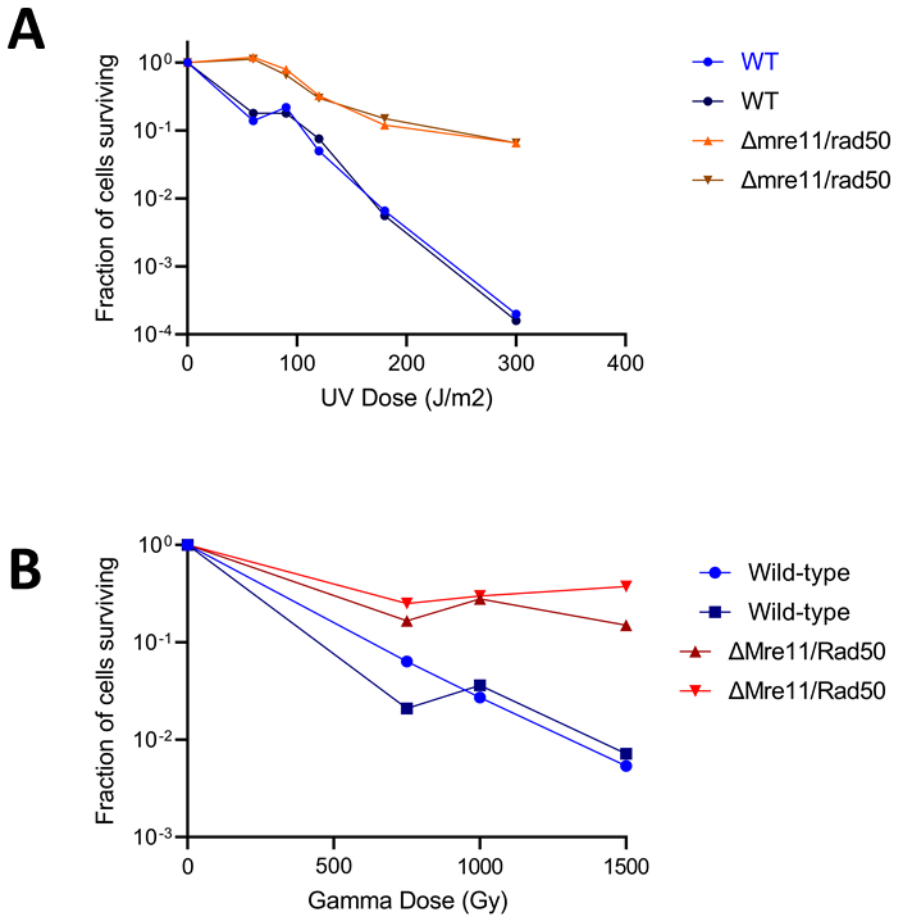


Figure 3.7 | **A) Cell survival assay under UV stress**. Wild-type (H1206) and UV resistant (H645) strains were grown in culture and dilutions were spotted as 20µl aliquots into YPC plates. Plates were then subjected to varying doses of UV radiation, combining UV and plating stresses. Survival fractions were calculated as a percentage of cells on the relevant strains' no radiation control plate. **B) Cell survival assay under gamma radiation stress.** As above, wild-type and resistant strains, H1206 and H645 respectively were grown in culture, diluted and 20µl aliquots were spotted onto Hv-YPC plates. Plates were placed under the gamma source for designated amounts of time to get varying doses of radiation. Survival fractions were calculated as a percentage of relevant strains no radiation control plate. In both cases, colour shading indicates technical replicates.

The result of this assay was to determine the correct dosage of UV and gamma to apply to cells to allow resistant cells to flourish, yet enough wild-type cells to die. In the UV assay, approximately 99.99% of the wild-type cells are killed. In the gamma assay, 99% of the wild-type cells are killed. However, in both cases, the level of radiation dosage is not so high that resistant cells are also killed and thus not retrieved from the screen.

Therefore the UV assay demonstrates that a UV dose of 300J/m^2-400 J/m^2 is appropriate, the gamma assay shows that 1500 Gy can also be employed to screen the library effectively. For both methods of irradiation, after the initial 'pilot' experiment, the assay was repeated using three biological replicates (figure 3.8).

The second calibration assay concurred with the pilot data, showing the robustness of the assay, and that 300J/m^2 UV is the most appropriate dosage to use for the library screen, as approximately only 1/10^5 wild-type cells survive, whilst 1/10^3 of the resistant cells survive (figure 3.8a).

The fraction of cells surviving under gamma radiation poses a more interesting dilemma. Even at 1500 Gy (approximately 8hrs of irradiation), the dosage does not discriminate well enough between resistant cells and background, therefore it does not make sense to continue with gamma radiation as a method of screening the library, as such, gamma irradiation will be used as a downstream method of analysis on constructs isolated under UV. This will provide an interesting alternative perspective and allow isolation of constructs that aid in both resistance pathways (to UV and gamma radiation) as well as those that act independently in each pathway. Furthermore, given that even 'resistant' cells show a survival rate of 1/10, there is not much scope for isolation of genes that provide logarithmic increases in resistance to gamma irradiation.

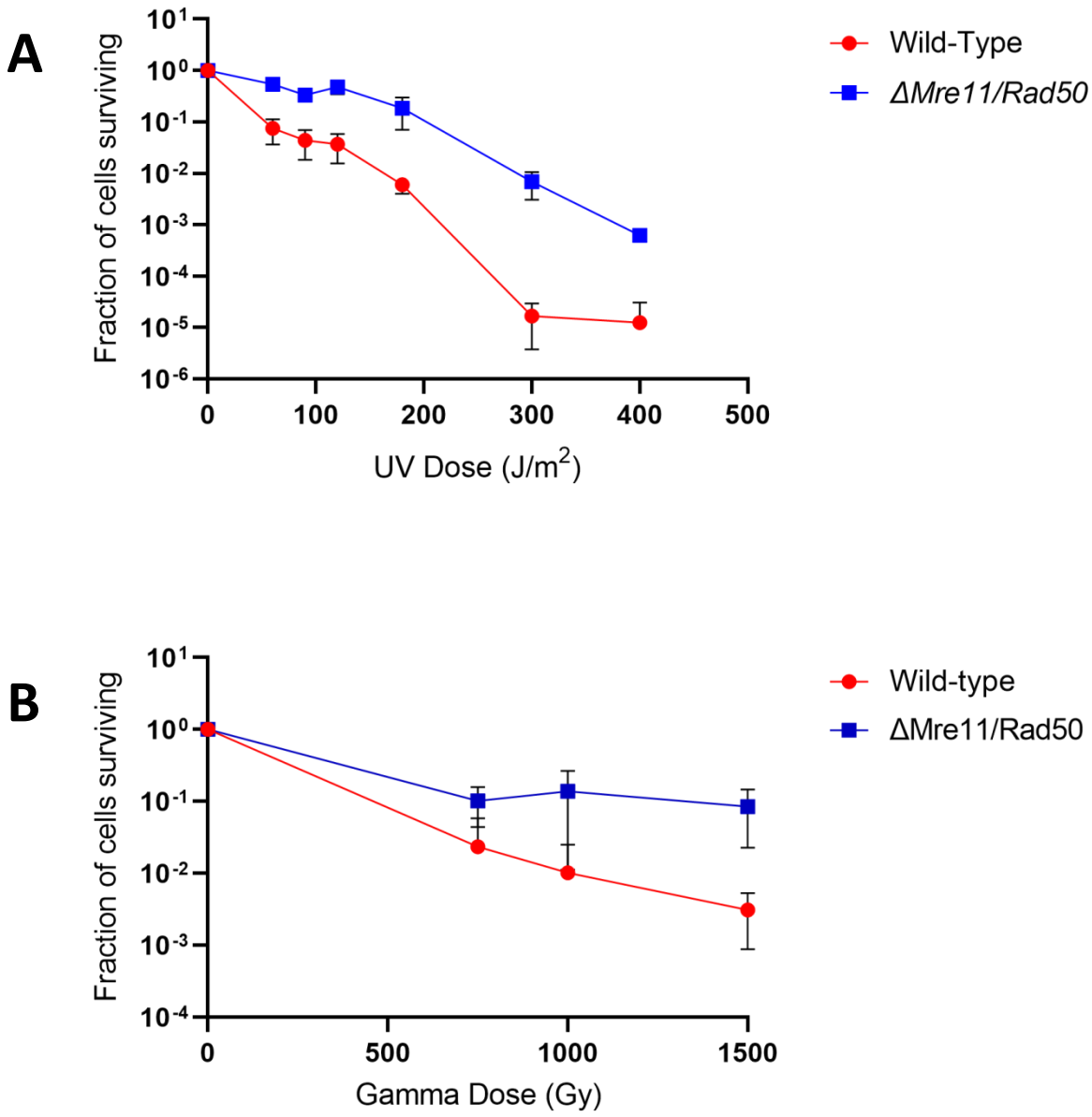


Figure 3.8 | **Cell survival assay**. A) Wild-type (H1206) and UV resistant (H645) strains were subjected to varying doses of UV irradiation, and cell viability was plotted (n=3). A non irradiated plate was used as a control. B) Wild-type (H1206) and UV resistant (H645) strains were subjected to varying doses of gamma irradiation and viability was plotted (n=3). A non irradiated plate was used as a control. In both cases the assay was carried out by spotting 20µl dilutions on the plates, and then allowing to dry prior to irradiation. Plates were grown in the dark for ~5 days before counting.

Prior to genomic library transformation and screening (using 300J/m^2 UV), the UV lamp used during all calibration assays thus far, including the pairwise competition assays, failed. Given the age of the lamp and container, it became difficult to source a replacement lamp, and even so, the intensity would be changed. Therefore it became necessary to repeat all calibration assays using another source of UV light. As such, a UV crosslinker, canonically used for crosslinking of DNA to membranes, was employed to carry out irradiation of the genomic library. The calibration survival assays were repeated, in the same fashion, but irradiated using the crosslinker. Biological replicates (n=3) were carried out and the data shown in figure 3.9. It is clear that the crosslinker initially provided a much higher dose of UV than the original lamp (figure 3.9a), so the assay was repeated with intervening values to find the most appropriate dose, and to ensure that a lower dosage didn't produce a similar survival rate (figure 3.9b).

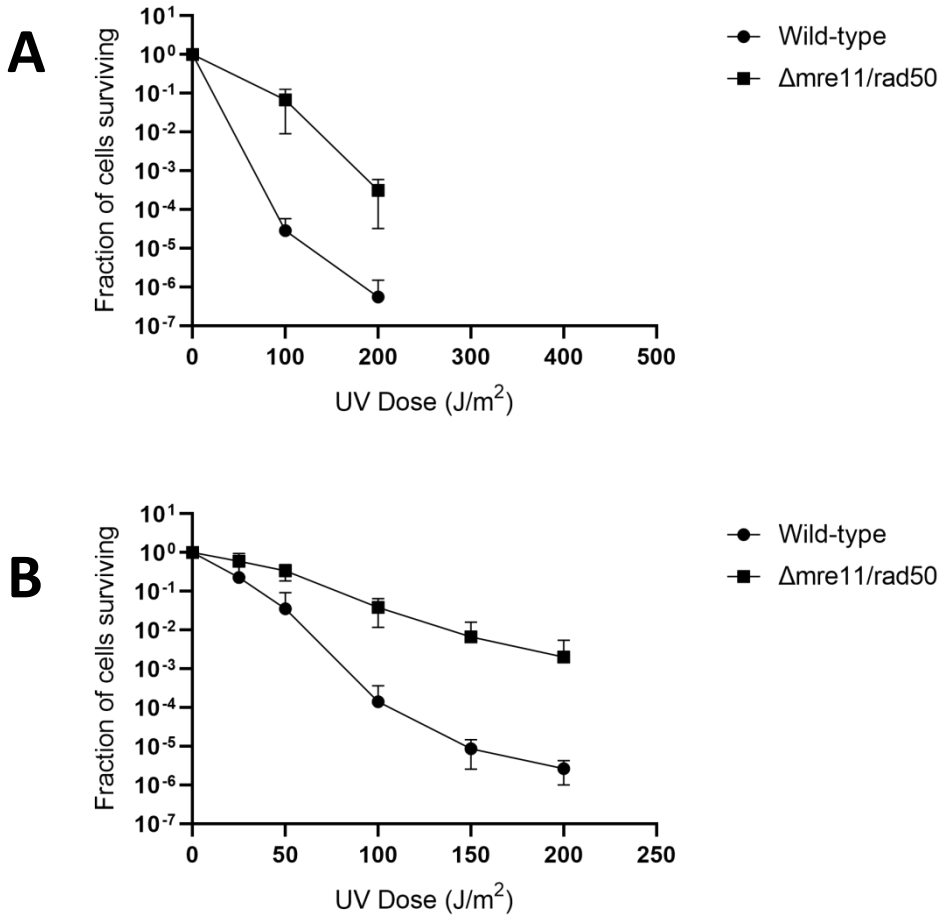


Figure 3.9 | **Cell survival assay using UV crosslinker**. (A) UV survival assay using 100-400J/m^2 UV. All cells died at 300J/m^2. (B) UV survival assay with intermittent values, showing a graded response to increasing UV dose. In both cases n=3.

3.4.4 Gamma Radiation Pairwise competition assay

Wild-type and radiation resistant strains H1206 and H645 respectively were grown overnight in 5ml of Hv-Ca+Ura to mimic the genome library conditions that will be laid out later on. Cultures were grown overnight at 45°C, rotating at 8rpm. Cultures were diluted the following day in fresh media and allowed to grow and reach an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Cultures were independently serially diluted to 10^{-4} fold, after which 100µl of each culture were mixed into a fresh tube of 5ml Hv-Ca+Trp+Ura (equating to approximately 10^3 cells of each culture). The mixed culture was allowed to reach an A_{650} of 0.4, after which 1ml aliquots were placed into four 4ml tubes. One tube acted as a no irradiation control and was left otherwise untreated. The other three tubes were irradiated at 750 Gy, 1000 Gy and 1500 Gy. After irradiation had finished, various dilutions were made of each culture and 100ul of different dilutions were plated on Hv-Ca+Ura+Trp plates. Some of the remaining culture was then diluted 1:2 into fresh media and left to grow in the dark at 45°C, rotating at 8rpm for two hours. After the time had elapsed, the culture dilutions and plating steps were repeated as above. Another dilution was made, and culture left for another two hours to grow at 45°C. Another dilution and plating step was carried out, after which a final 1:1000 dilution was made into fresh media, which was left to grow overnight for 20hrs. Plates were left for five days to grow in the dark, after which appropriate plates were selected from each time point after irradiation, as well as the no irradiation control.

Plates were then sprayed with X-Gal, which allowed H1206 and H645 strains to be visually distinguished, as H645 contains the bgaHa gene, rendering X-Gal subject to cleavage and a blue molecule to be formed. Wild-type colonies (H1206) therefore remain red in colour due to the lack of the ability to cleave X-Gal colourless precursor molecule into the blue end-product (Holmes and Dyall-Smith, 2000, Delmas *et al.*, 2009) . The ratio of wild-type to Δ*mre11-rad50* cells were calculated for each time point under different radiation dosages and a growth curve plotted (figure 3.10). Data was normalised using the counts from the no irradiation control plate at time 0hrs.

Aberrant results lead to numerous repetitions of this assay, with similar results each time. Eventually the move to a UV based competition assay was decided once the decision to screen the library with UV was made, based on the survival assays (see above).

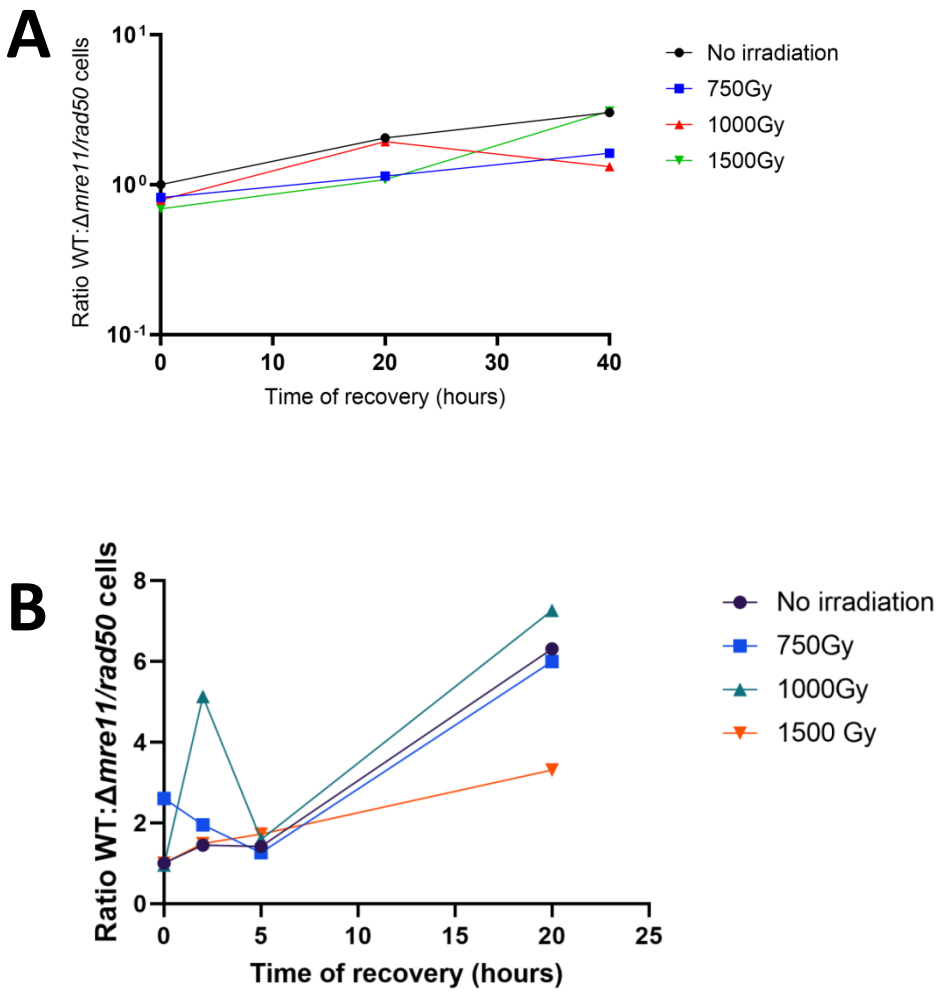


Figure 3.10 | **Gamma radiation pairwise growth competition assay. A) Gamma competition assay with time points 0hrs, 20hrs and 40hrs after irradiation**. A mixed culture was set up between wild-type and resistant strains in a 1:1 ratio. Cells were irradiated, or left untreated and allowed to grow for 0hrs, 20hrs and 40hrs. After each time point, dilutions of cells were plated and cells were incubated for 5 days at 45°C. Cells were sprayed with X-Gal and the ratio of red (wild-type) to blue (resistant) colonies counted. Data was normalised to the counts from the no irradiation control plate at time point 0hrs. **B) Gamma competition assay with additional time points**. The gamma competition assay was repeated as per graph A, however additional time points were added (0hrs, 2hrs, 5hrs, 20hrs) to see cell ratios in initial stages of irradiation.

Given the erroneous results from this initial pilot competition assay using gamma radiation, and the subsequent data from the survival assays (figure 3.8b), the decision was made to irradiate using UV instead of gamma radiation. The main reasoning for the switch was that the cell death provided by irradiation in liquid culture is not sufficient for screening, therefore irradiation on plates must be used. This is evidenced by unpublished data showing that plating cultures on solid media provides an independent stress to the cells, which when combined with irradiation, leads to greater cell death. Furthermore, irradiation on solid media allows for homogeneity with respect to the irradiation across the plate, and across individual cells, whereas in liquid culture, some cells may not be irradiated fully, or even at all, thus leading to spurious results.

3.4.5 UV Pairwise competition assay

Wild-type and radiation resistant strains H1206 and H645 respectively were grown overnight in 5ml of Hv-Ca+Ura to mimic the genome library conditions that will be laid out later on. Cultures were grown overnight at 45°C, rotating at 8rpm. Cultures were diluted the following day in fresh media and allowed to grow and reach an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Cultures were independently serially diluted to 10^{-4} fold, after which 100µl of each culture were mixed into a fresh tube of 5ml Hv-Ca+Trp+Ura (equating to approximately 10^3 cells of each culture). The mixed culture was allowed to reach an A_{650} of 0.4, after which 1ml was aliquot into a 5ml tube to act as a no irradiation control. A subsequent aliquot of 1ml culture was placed into a hydrophilic cell culture dish and irradiated using 300J/m^2 of UV radiation. Once irradiated, both cultures were spun down and resuspended in the same volume of fresh media. Serial dilutions were prepared and 100µl of each appropriate dilution was plated on Hv-Ca+Ura+Trp plates. After plating varying dilutions, some of the remaining cultures were diluted 1:2 into fresh media and wrapped in tin foil, they were allowed to continue growing for a further two hours at 45°C, rotating as before. A subsequent repeating of the previous dilutions and plating was then carried out. Another 1:2 dilution was then set up, allowing further growth for another three hours. After another round of dilutions and plating, a final 1:1000 dilution was set up in fresh media, and the culture returned to 45°C overnight (approximately 20hrs). After 20hrs had lapsed, cells were diluted and plated as per previous details.

Plates were left for five days to grow in the dark, after which appropriate plates were selected from each time point after irradiation, as well as the no irradiation control. Plates were then sprayed with X-Gal, which allowed H1206 and H645 strains to be visually distinguished, as H645 contains the bgaHa gene, rendering X-Gal subject to cleavage and a blue molecule to be formed. Wild-type colonies (H1206) therefore remain red in colour due to the lack of the ability to cleave X-Gal colourless precursor molecule into the blue end-product (Holmes and Dyall-Smith, 2000, Delmas *et al.*, 2009) . The ratio of wild-type to Δ*mre11-rad50* cells were calculated for each time point under different radiation dosages and a growth curve plotted (figure 3.11). Data was normalised using the counts from the no irradiation control plate at time 0hrs.

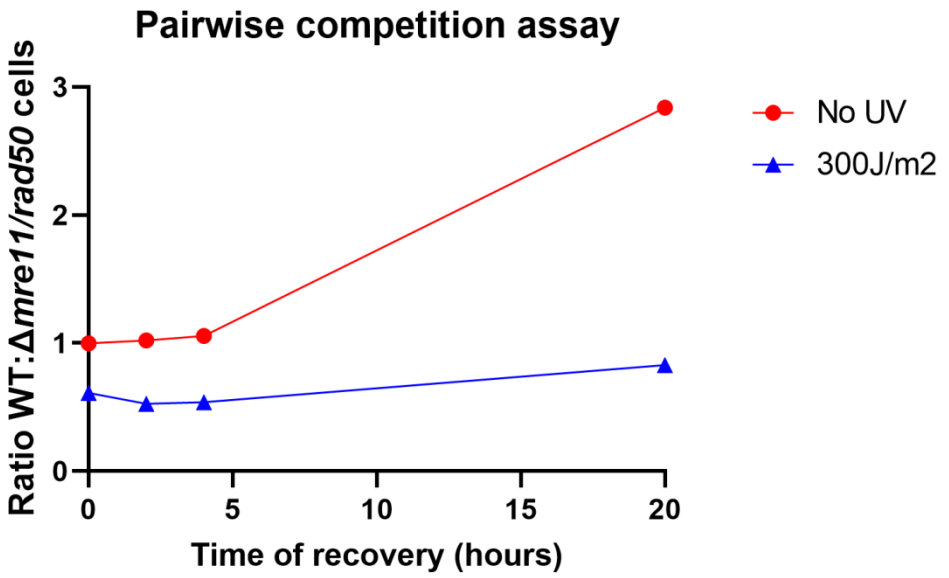


Figure 3.11 | **UV radiation pairwise competition assay between strains H1206 and H645**. The ratio of wild-type (H1206) to radiation resistant (H645) strains was calculated at each time point after irradiation, for both the 300J/m^2 culture, and the no UV control. Samples were plated at each time point indicated after which the ratios were calculated based on the X-Gal screening method detailed in the text. Colony numbers were normalised to the no UV control plate at time point 0hrs.

Expected killing was sub-optimal, therefore a new method of irradiation has been developed, detailed in section 3.5, along with relevant rationale behind new strategy. In spite of the above, however, the data show that after UV treatment, wild-type cells do not ever reach the saturation level of resistant colonies, even after 20hrs of regeneration. Therefore the genome library, if resuspended for a period of time in optimal growth conditions, will not suffer a competition event between the two strains where wild-type will regain momentum above that of the resistant colonies.

Although the data above are somewhat helpful, the ratio of wild-type:resistant strains was not as low as expected, based on the calibration survival curves. Therefore this pilot experiment was repeated, this time with three biological replicates.

However, the methodology was changed somewhat to increase the killing seen and therefore provide a more robust assay for how a single wild-type cell recovers over time.

1ml of culture was probably too dense to ensure equal irradiation across all cells in the culture, in addition the culture remained static whilst being irradiated. The new method was to repeat the assay with 500µl of culture, with constant agitation of the petri plate whilst irradiation occurred. This would allow all cells within the culture to be irradiated equally, and therefore allow more cell death. To provide agitation to the plate without having to place hands under the UV lamp, the petri plate (without the lid) was placed onto a thermomixer using double-sided sticky tape (Eppendorf Thermomixer R 0.5ml Shaking Heater Block) (Figure 3.12).

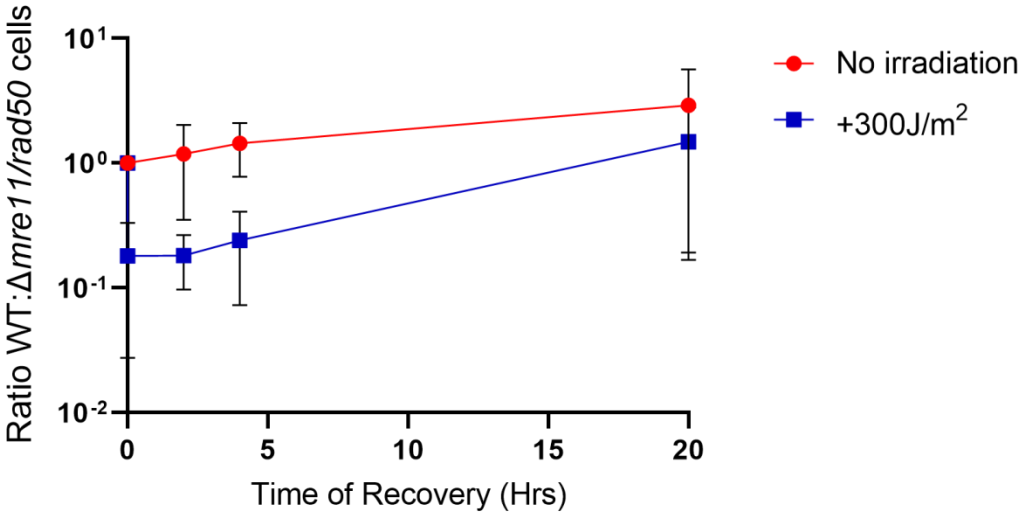


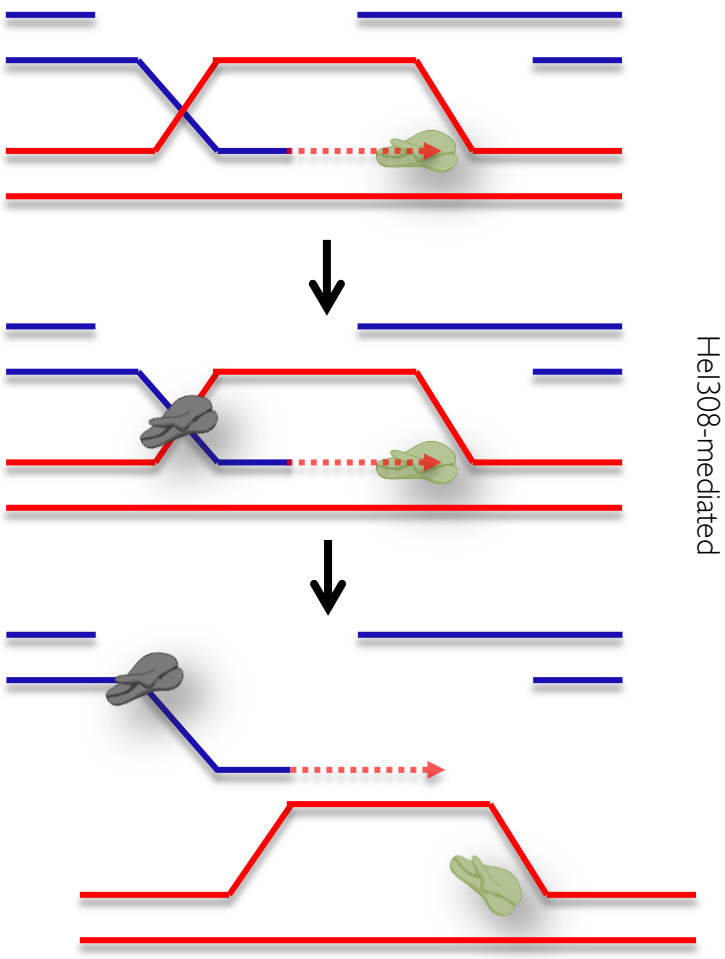
Figure 3.12 | **UV radiation pairwise competition assay between strains H1206 and H645 (Updated method)**. Pilot assay was repeated using three biological replicates, however irradiation was carried out on 500µl of culture, as opposed to 1ml, to provide maximum coverage and cell death. Furthermore, the culture was placed under the UV lamp and agitated using an Eppendorf thermomixer for the entirety of the irradiation.The ratio of wild-type (H1206) to radiation resistant (H645) strains was calculated at each time point after irradiation via blue/pink screening mediated by XGal staining.

As shown in figure 3.12, a single 'wild-type' cell, after irradiation with 300J/m^2 UV, will take approximately 20hrs to recover back to pre-irradiated levels. This becomes relevant for the screening process, as one dose of UV will be rendered useless after 20hrs, bringing all cells back to pre-irradiated levels. Therefore, screening must make use of sequential 'hits' of UV irradiation, less than 20hrs apart from each other to ensure a lone survivor will not show up in the final outgrowth of resistant colonies.

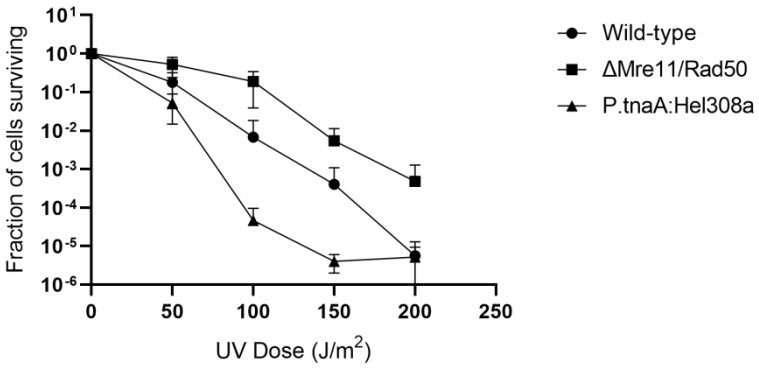
Of note, this assay was carried out using the original UV lamp which subsequently broke. The assay was not repeated using the crosslinker owing to the reduced spatial capacity of the instrument, and therefore the lack of ability to provide culture agitation during irradiation. However, the results still stand and provide a solid backbone of evidence which will dictate the exact method of screening.

After the calibration assays were carried out, it became interesting to see if, given the nature of the protein, if Hel308 would provide a potential positive control for the genome library screen and represent a 'resistant' construct that one could aim to isolate from the screen. The theory behind this was the nature of Hel308 being an anti-recombinogenic monomeric 3'→5' helicase, conserved throughout Archaea and metazoans. Hel308 is part of the helicase superfamily 2 and mutational analysis has shown that mutants in *Drosophila melanogaster* exhibit a hypersensitive phenotype to DNA crosslinking agents (Boyd *et al.*, 1981). Hel308 is thought to act in the destabilization of D-loops (figure 3.13a), thus reducing the amount of recombinational repair occurring within cells. Given that the strain deleted for mre11/rad50 shows increased resistance to UV radiation (Delmas *et al.*, 2009), owing to reduced levels of recombinational repair, thus mediating genomic stability given the polyploid nature of *H. volcanii*, it is thought that an overexpression of an anti-recombinase, such as Hel308, could provoke the same phenotypic effect and therefore be a viable candidate to arise from the screen. Therefore, a survival assay was carried out using a Hel308 overexpression construct, under the same, Tryptophan inducible promoter, as the initial calibration curves and those used in the construction of the library (see later). The data are shown in figure 3.13b and show that Hel308 is indeed toxic, even at expression levels imposed by p.*tnaA*, and therefore can not act as a control construct which can be expected to come from the screen. Validation of the presence of the Hel308 overexpression vector was carried out using colony PCR, using primers pBSF2 and Hel308R (figure 3.13c).

A


Strand invasion and DNA synthesis
Hel308-mediated

B


Fraction of cells surviving
UV Dose (J/m2)
Wild-type
ΔMre11/Rad50
P.tnaA:Hel308a

C

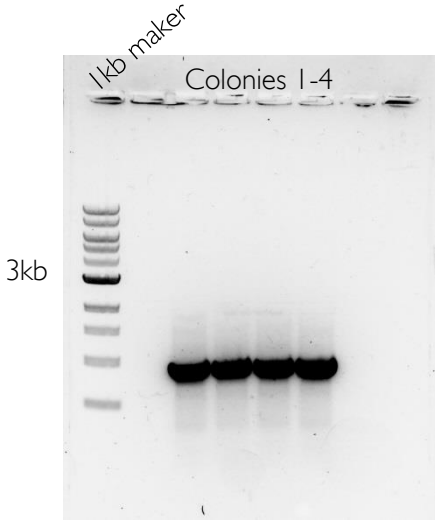
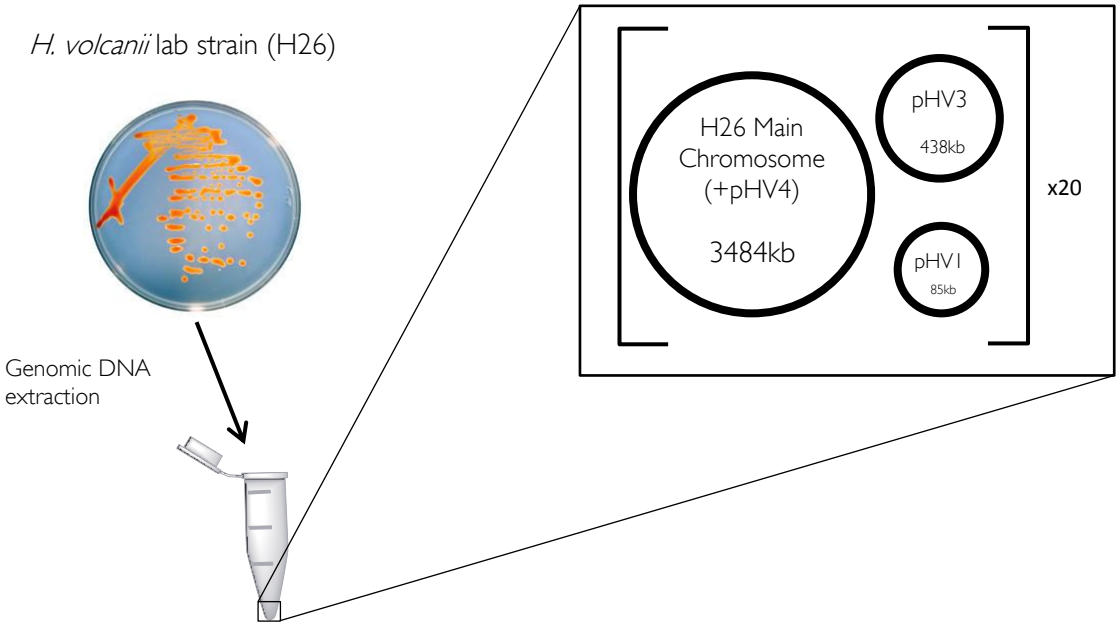

1kb maker
Colonies 1-4
3kb

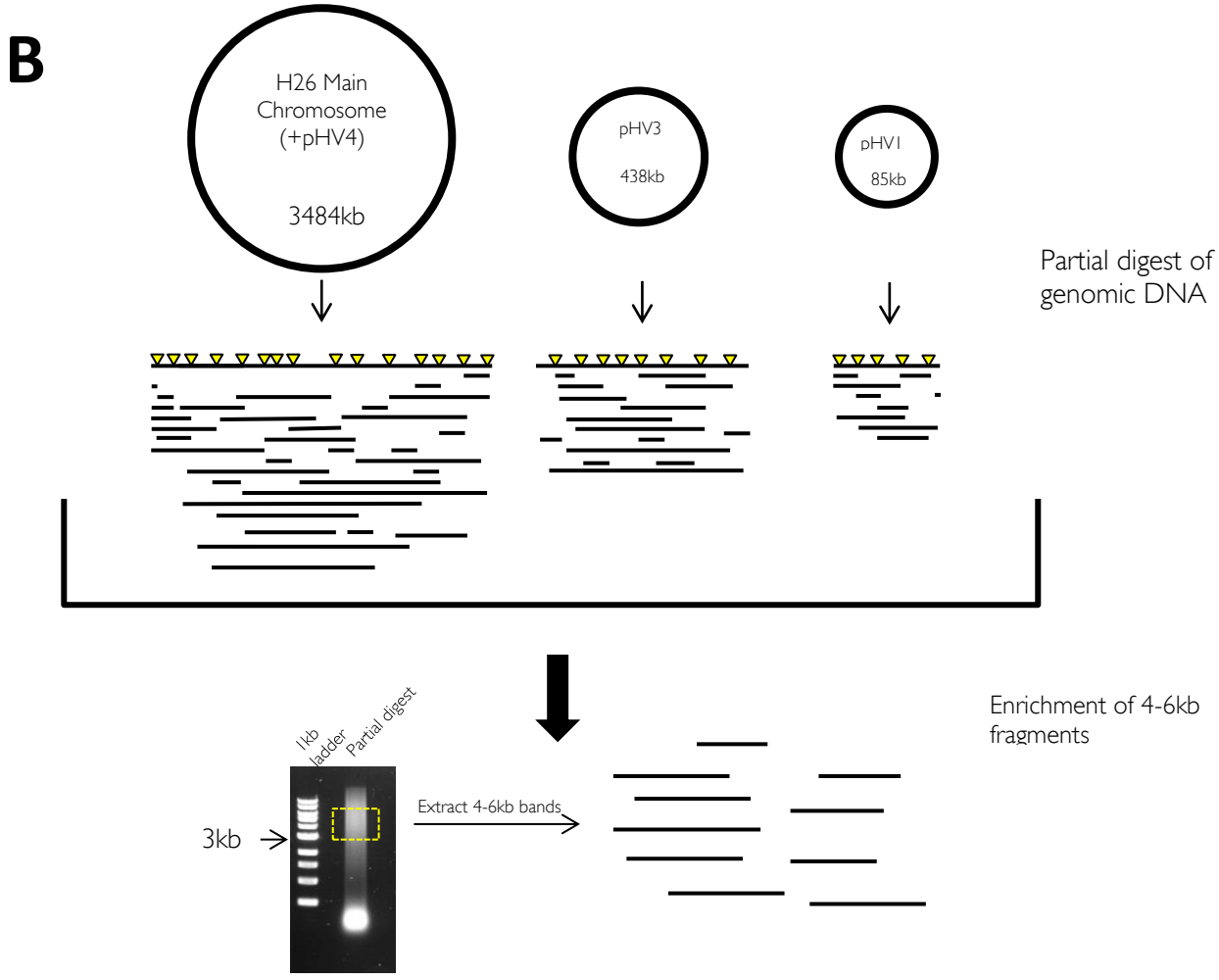
Figure 3.13 | **Hel308 Mechanism and Overexpression Survival Assay with Colony PCR**. A) Proposed mechanism of potential Hel308 action at D-loop. Homologous chromosomes shown in red and blue, with DNA polymerase and Hel308 shown in green and black, respectively. B) Construct pTA1231, containing an overexpression construct for Hel308a under the p.*tnaA* promoter was used to determine the viability of cells compared to wild-type (no hel308, strain H1206). Assay carried out as previous, using spotting of 20µl cultures. C) Colony PCR results of transformed strain, prior to survival assay.

3.4.6 Genomic library construction

Haloferax volcanii strain H26 genomic DNA was first spooled (figure 3.14a) and purified, after which a partial genomic digest was carried out using AciI (NEB), which cuts frequently in G/C rich genomes such as *H. volcanii*. The digest was run with 0.5U/ug of AciI, for 2.5 hrs at 45V on a 0.8% TAE agarose gel. Bands were size enriched, averaging at 4kb (figure 3.14b). Genomic library fragments were then ligated directly downstream of a tryptophan inducible promoter in the shuttle vector pTA927 (figure 3.14c).

A




B
H26 Main Chromosome (+pHV4)
3484kb
pHV3
438kb
pHV1
85kb
Partial digest of genomic DNA
1kb ladder
Partial digest
3kb
Extract 4-6kb bands
Enrichment of 4-6kb fragments

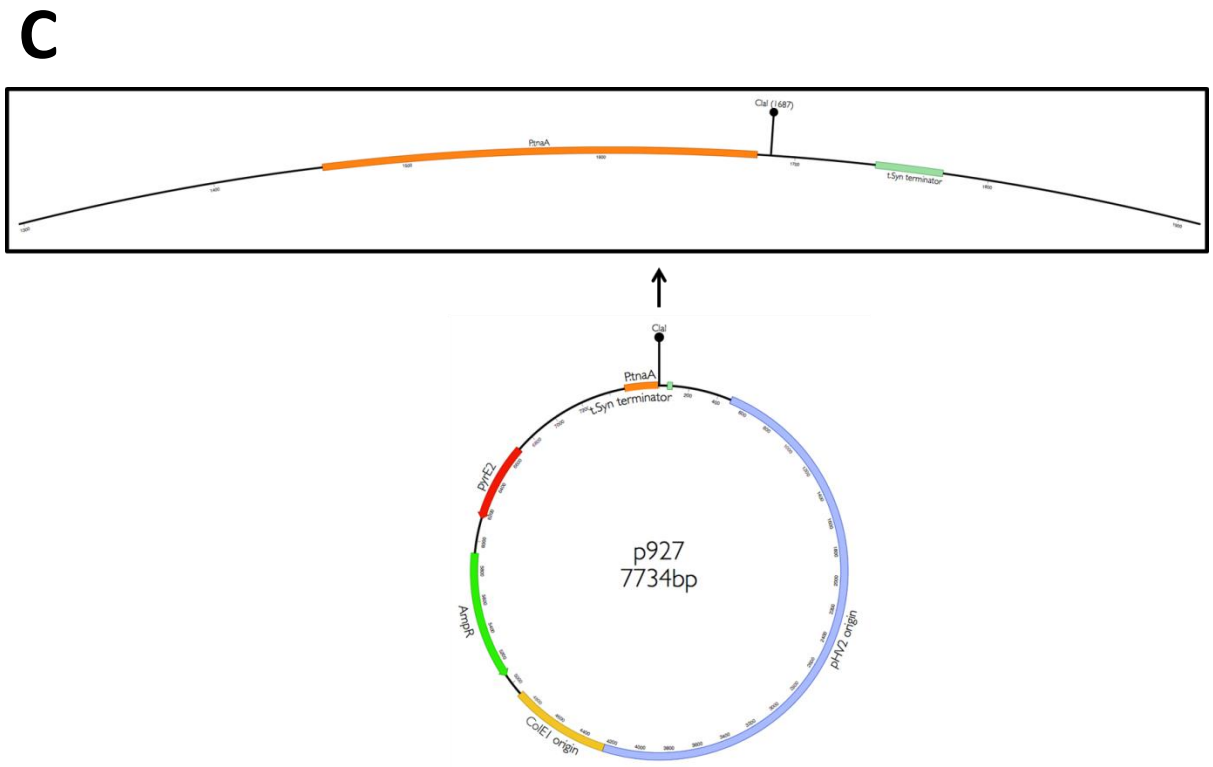

C
ClaI (1687)
PtnaA
t.Syn terminator
ClaI
PtnaA
t.Syn terminator
PhrE2
p927
7734bp
AmpR
ColE1 origin
pHV2 origin

Figure 3.14| **Genomic library cloning schematic. A) Genomic DNA extraction.** Strain H26 was grown overnight in 5ml Hv-YPC, after which genomic DNA was spooled using standard procedure after cell lysis and DNA precipitation. Genomic DNA was then washed in EtOH and left to air dry prior to resuspension in TE. **B) Partial Restriction Digest.** 50µg of genomic DNA was then subjected to a partial restriction digest using 0.5U/µg of AciI (NEB) for 30 minutes in CutSmart buffer. Fragments were then size enriched prior to gel purification. **C) pTA927 shuttle overexpression vector ligation.** Fragments were then ligated in molar excess into the ClaI site within the vector pTA927, thus placing fragments downstream of the strong tryptophan-inducible promoter p.*tnaA*.

The genomic library constructs were then transformed by electroporation into *E. coli* strain XL1-Blue and ≥25000 recombinant colonies were generated (figure 3.15). All clones were pooled and constructs were then isolated via Midiprep protocol after overnight incubation at 37°C, shaking at 200rpm. Glycerol stocks of each plate were prepared and stored at -80°C. Midiprep DNA was then used to transform *H. volcanii* strain H1206, which is deleted for the restriction enzyme, Mrr, therefore susceptible to direct transformation with *dam+* DNA from *E. coli*. Over 250,000 *H. volcanii* transformants were then achieved and subsequently screened.

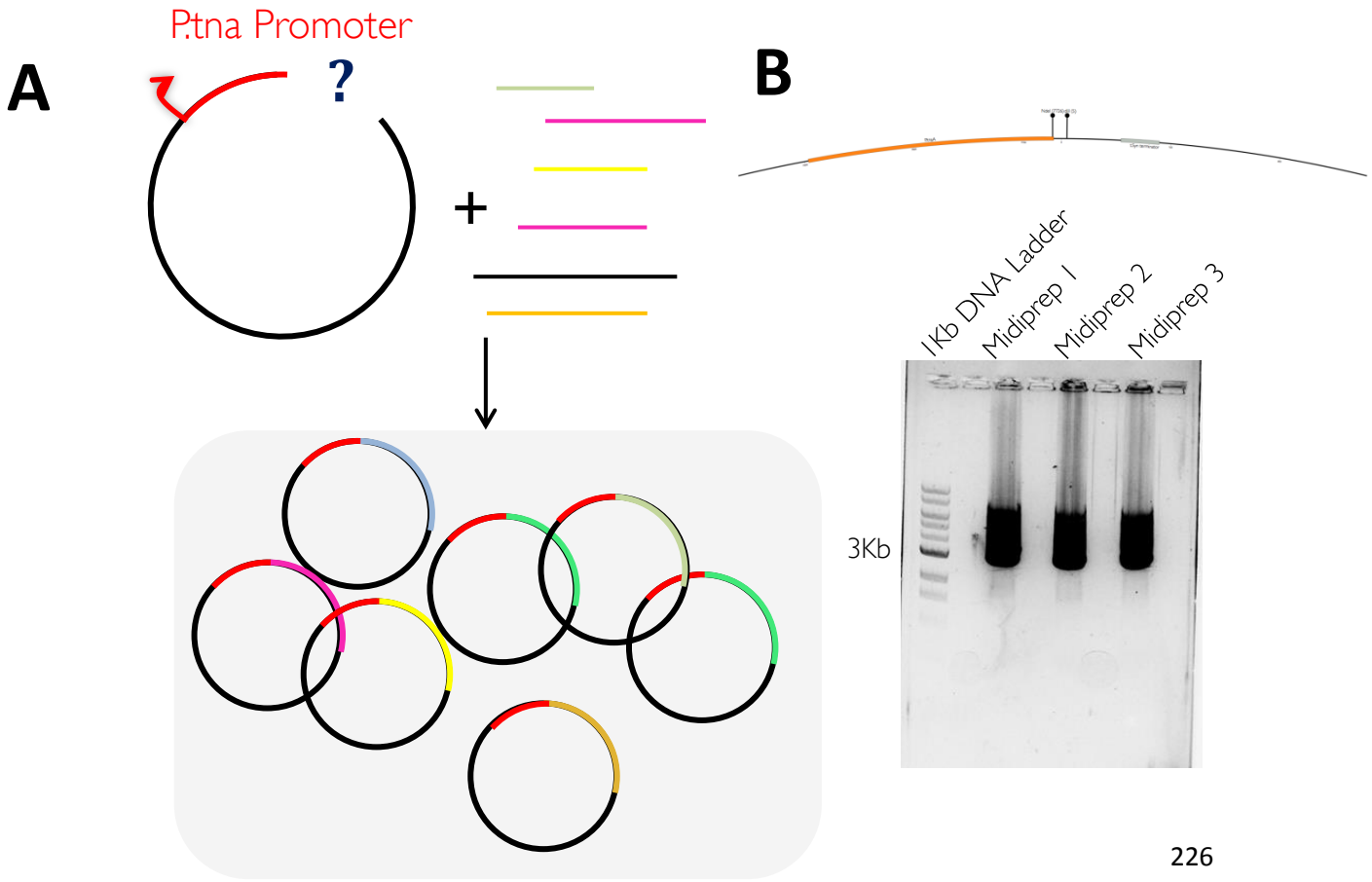


Figure 3.15| **Constructs generated from library cloning.** (A)Schematic depicting the variation produced from insertion of partially digested genomic library fragments into pTA927, downstream of the p.*tnaA* promoter. (B) Results of diagnostic digest of midiprepped DNA from genomic library, confirming size selection and cloning to be successful. Cut sites on pTA927 shown.

In order to validate that genomic library construction had worked, a digest of the midiprepped DNA was carried out using restriction endonucleases flanking the insertion site (Nde1/HindIII) at 37°C overnight, after which the resulting fragments were run on a 0.8% TAE gel for 1.5hrs at 80V. The results can be found in figure 3.15b and show that size selection of fragments had been successful and fragments produced result in the majority of bands being seen between 4kb and 6kb

Algorithms, based on the Poisson distribution, have been produced in order to assist with probability calculations when transforming a limited pool of constructs into an organism. For the library transformation into the *H volcanii* host, the GLUE algorithm was used, which stated that at least 37,000 transformants must be obtained in order to achieve 99% coverage (Patrick, Firth and Blackburn, 2003, Firth and Patrick, 2005). With ~250,000 transformants screened, this threshold was more than surpassed.

3.4.7 Genome library UV screens and validation assays

Immediately after transformation of the genome library constructs into *Haloferax volcanii*, plates were subjected to 300J/m^2 of ultraviolet radiation using the UV lamp. Plates were then kept in a black plastic bag and allowed to grow for five days at 45°C. Thousands of colonies came from this initial pilot screening attempt, therefore the process was repeated and irradiation conditions changed. Furthermore, the UV lamp broke at this point, therefore the crosslinker was now used for subsequent screening attempts. Despite this, prior to the UV lamp breaking, downstream survival assays were carried out on a random selection of colonies isolated from this screen (figure 3.16). In both cases, no colonies appear to be particularly resistant, and control strain does not land where expected.

The second attempt at screening the library was more successful and used two separate doses of 200J/m^2 UV. Two doses were given, 48hrs apart and resulted in many less colonies appearing five days later. However, over twenty colonies were present and when tested for a UV resistance phenotype, most didn't appear to have significantly increased resistance levels compared to wild-type. This is not surprising given the competition assay data, showing that a single colony can survive and return to pre-irradiated levels after 20hrs.

The third and final screening attempt was entirely successful and whittled down the transformants to a mere 21 colonies. Screening was carried out by irradiation with 200J/m^2 UV, thrice, exactly 20hrs apart, thus not allowing for "out-growth" of stochastic colonies prior to the second irradiation event. Once screening had taken place, the 21 colonies were stored in glycerol at -80°C, and single colonies restreaked for downstream analysis and phenotypic confirmation.

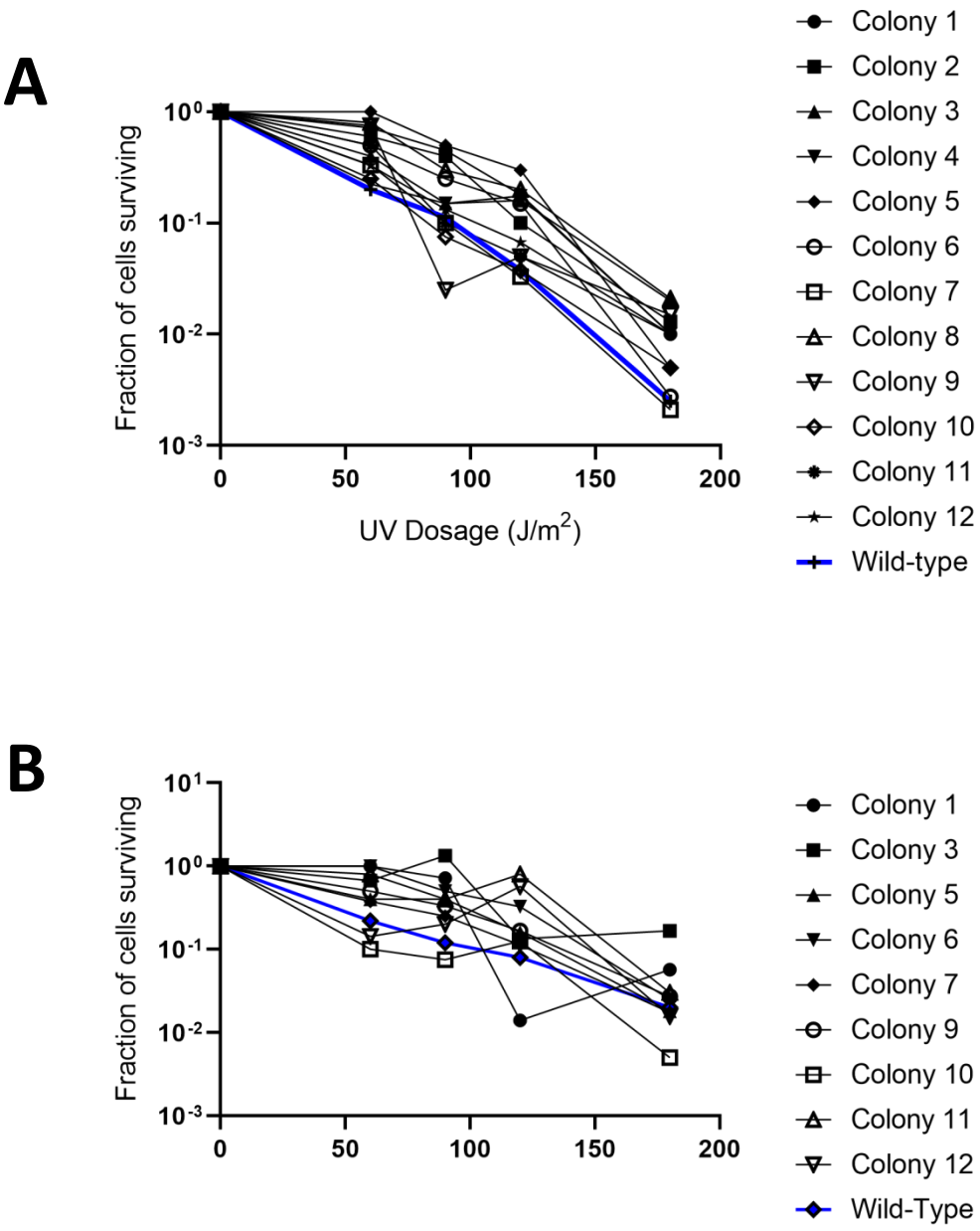


Figure 3.16 | **Survival validation assay for initial screening attempt.** (A)Survival assay with 12 selected colonies from the initial screen under 300J/m^2 UV. (B) Repeat of survival assay using selected colonies as above, less for colony 2, 4 and 8 as they failed to grow in overnight culture.

The second screen, using two separate temporally distinct doses of UV, 48hrs apart, yielded better results. Candidates from this screen were also subjected to the same validation survival assay alongside the wild-type host (figure 3.17).

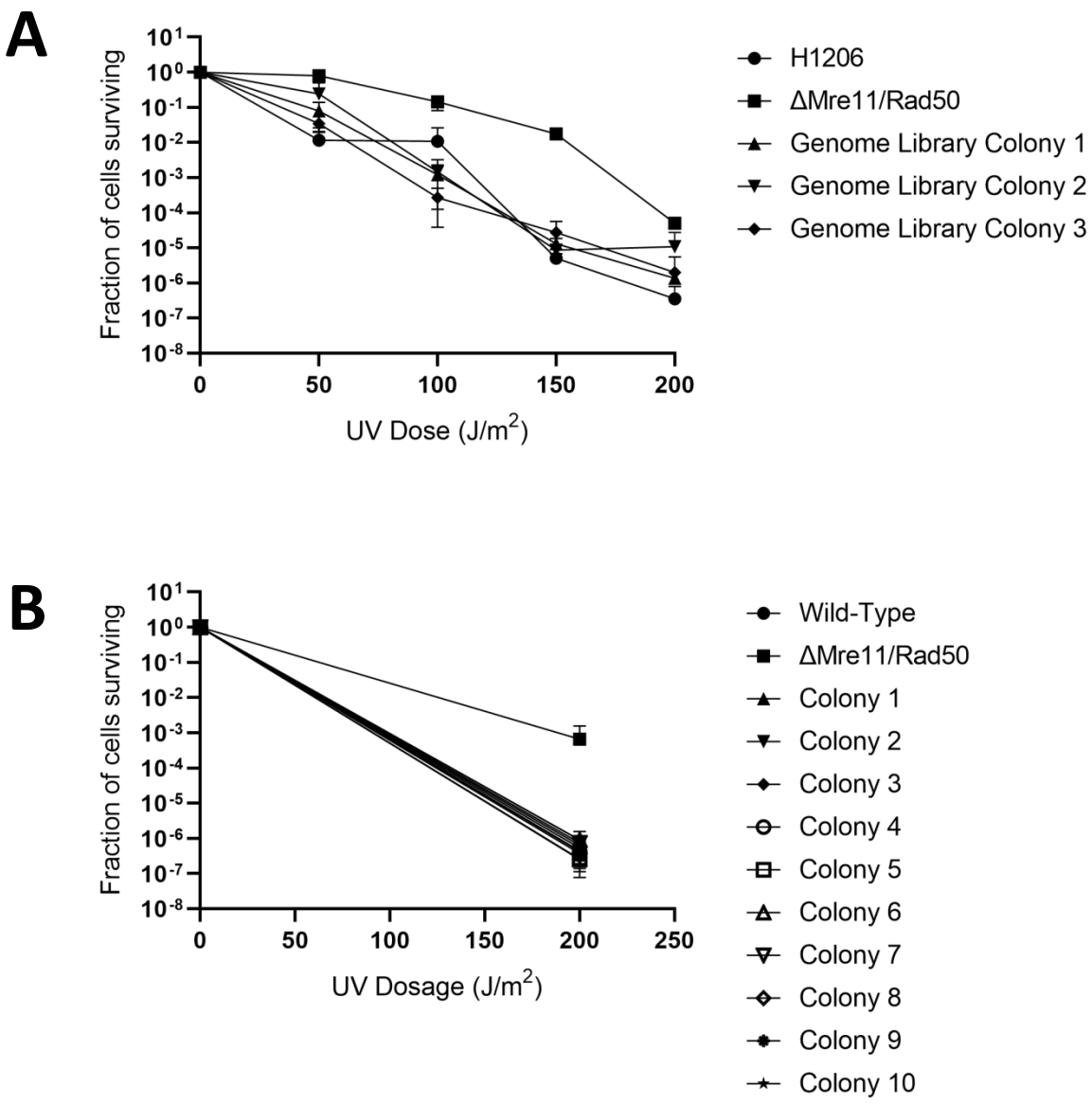


Figure 3.17| **Survival validation assay for second screening attempt.** (A) Three colonies selected and validated alongside wild-type and Δmre11/rad50 strain. (B) Additional colonies selected and validated as before.

In this case, the control colonies did fall where expected in terms of survival rate compared to the initial calibration assays, so the data are more trustworthy than the previous attempts. However, there still appears to be very little that looks to be significantly resistant especially compared to the already resistant H645 strain (mre11/rad50 deletion strain).

The third and final screening attempt was successful in reducing the number of colonies obtained from the screen and providing sufficient killing without allowing for outgrowth of stochastic survivors. The screening was carried out using three rounds of 200J/m^2 20hrs apart, then allowing survivors to grow for 5 days. Although a good number (approximately 100 colonies) grew, many were very small colonies, a subset of these were restreaked and didn't grow, therefore can be discounted. The larger, pink, colonies were selected and taken for further phenotypic analysis via a validation survival assay (figure 3.18).

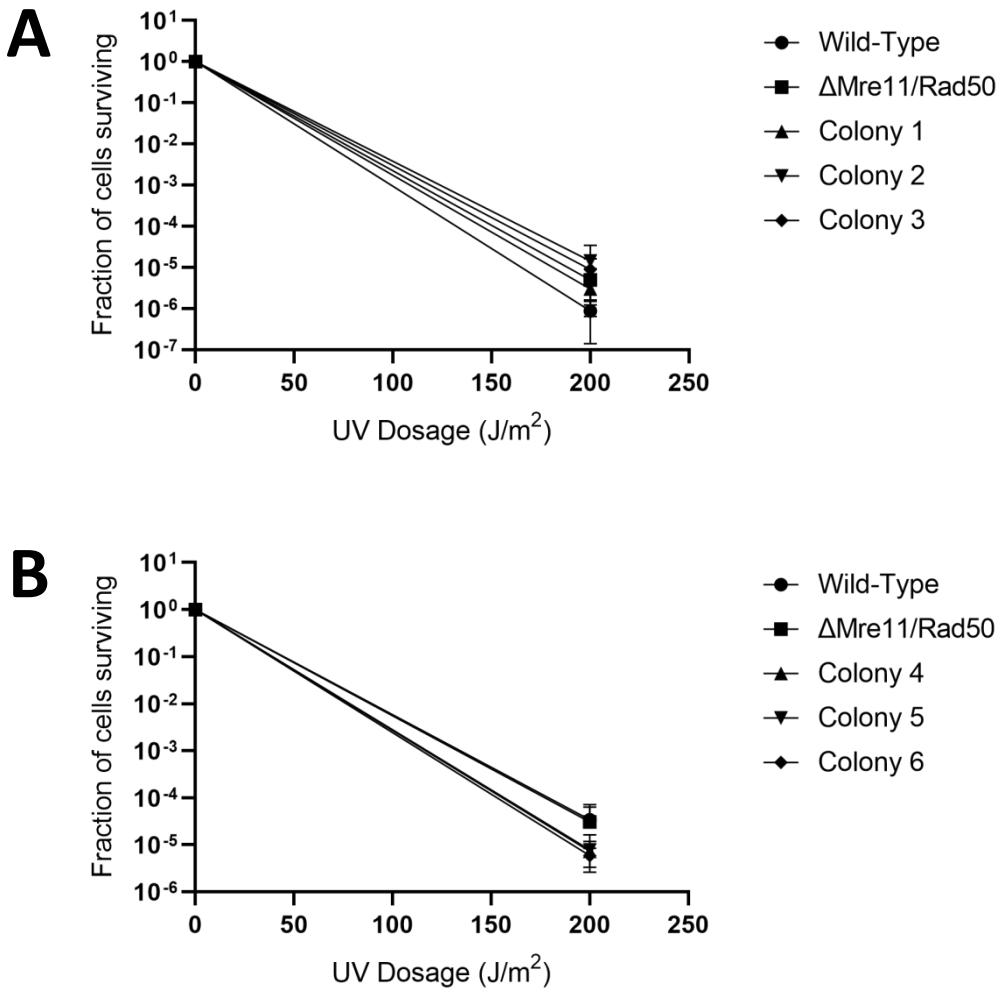


Figure 3.18 | **Survival validation assay for final screening attempt.** Initially six colonies were selected for the first assay (Colonies 1-3 (A), Colonies 4-6(B)), however resistance did not appear to be greater than wild-type.

Unfortunately, none of these initial colonies look to be resistant in any way. The reason behind this is the method used to determine viability. The method used is as described in the materials and methods section above. Essentially 20µl spots are plated for different dilutions of culture, which are then allowed to dry and are then subjected to varying doses of UV irradiation. During these assays, as the UV dosage increased, the appearance of colonies appeared to change morphologically, from individual colonies being visible for counting, to a doughnut-like ring around the circumference of the spot in question. This has been termed “Haloferax ring formation” and an example of which is shown in figure 3.19.

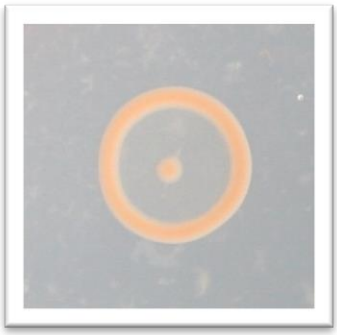


Figure 3.19 | **Haloferax Ring.** Example of a ring formed from UV irradiation of a spotted culture (20µl). Such formation hinders counting for cell viability assays and appears to correlate with increasing UV dosage, especially 100J/m^2 upwards.

Given the archaeal ring formation problem, which means spotted cultures are hard, if not impossible, to count accurately, alternative methods were then attempted to try and reduce this problem.

In order to minimise ring formation as an artefact of spotting cultures onto solid media and irradiating with UV, a more specialised approach was attempted to generate a more high-throughput method of culture plating. For this, a three-dimensional bead-guide was designed based on (Hamilton *et al.*, 2002).

This had been used as part of a previous PhD student's thesis (Adam Collins and Alan Huett). The bead-guide was manufactured by Medical Manufacturing (University of Nottingham) from polytetrafluroethylene (PTFE). The bead-guide is placed onto pre-set agar plates contained within OmniTray plates (Nunc).

Sterile 5mm glass beads (Sigma) were then dispensed into each lane via use of a 5mm bead dispenser (TissueLyser). 20µl samples of culture were then dispensed onto each bead and beads were then rolled along the lanes created by the bead-guide, resulting in homogenous spreading of the 20µl sample agross the agar (figure 3.20). This was only attempted using a single colony, Colony 2, which seemed to be the most promising colony so far in the previous screen (figure 3.18a).

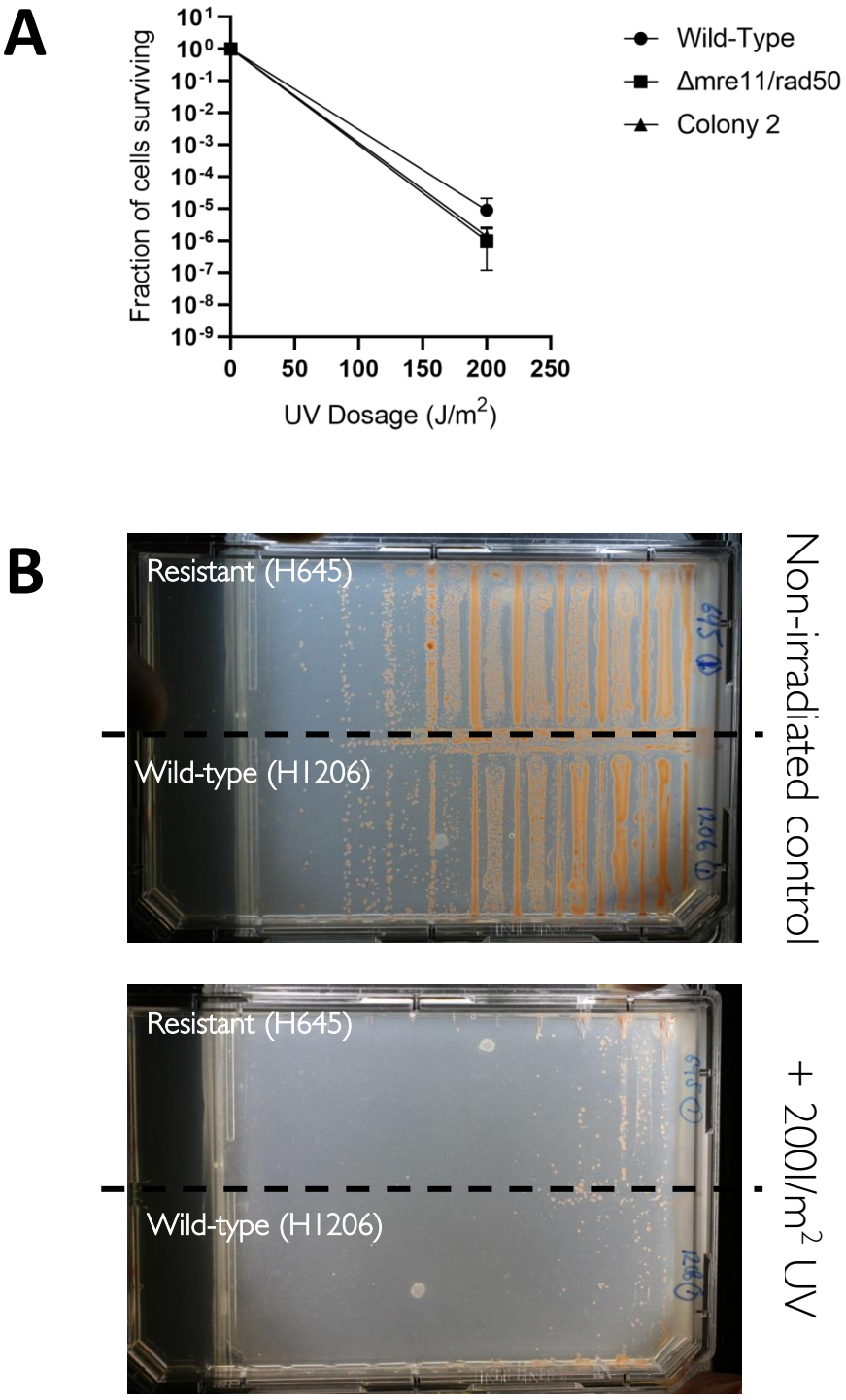


Figure 3.20 | **Bead plating assay.** (A) Graph depicting survival assay results for colony 2 using bead assay. (B) Images of bead plates showing substantial culture leakage between lanes, however the ring formation problem has been mitigated as a result.

Given that the plating method using glass beads and guided lanes did not work as effectively as desired, the results from this assay were not substantial enough to use as a method for validating the resistance of colonies obtained from the library screen. Therefore another method was used, this time it was to directly plate the 20µl, not as a spot, but across an entire petri plate of agar. The downside to this is reducing the number of dilutions used, but the assay did appear to work (figure 3.21).

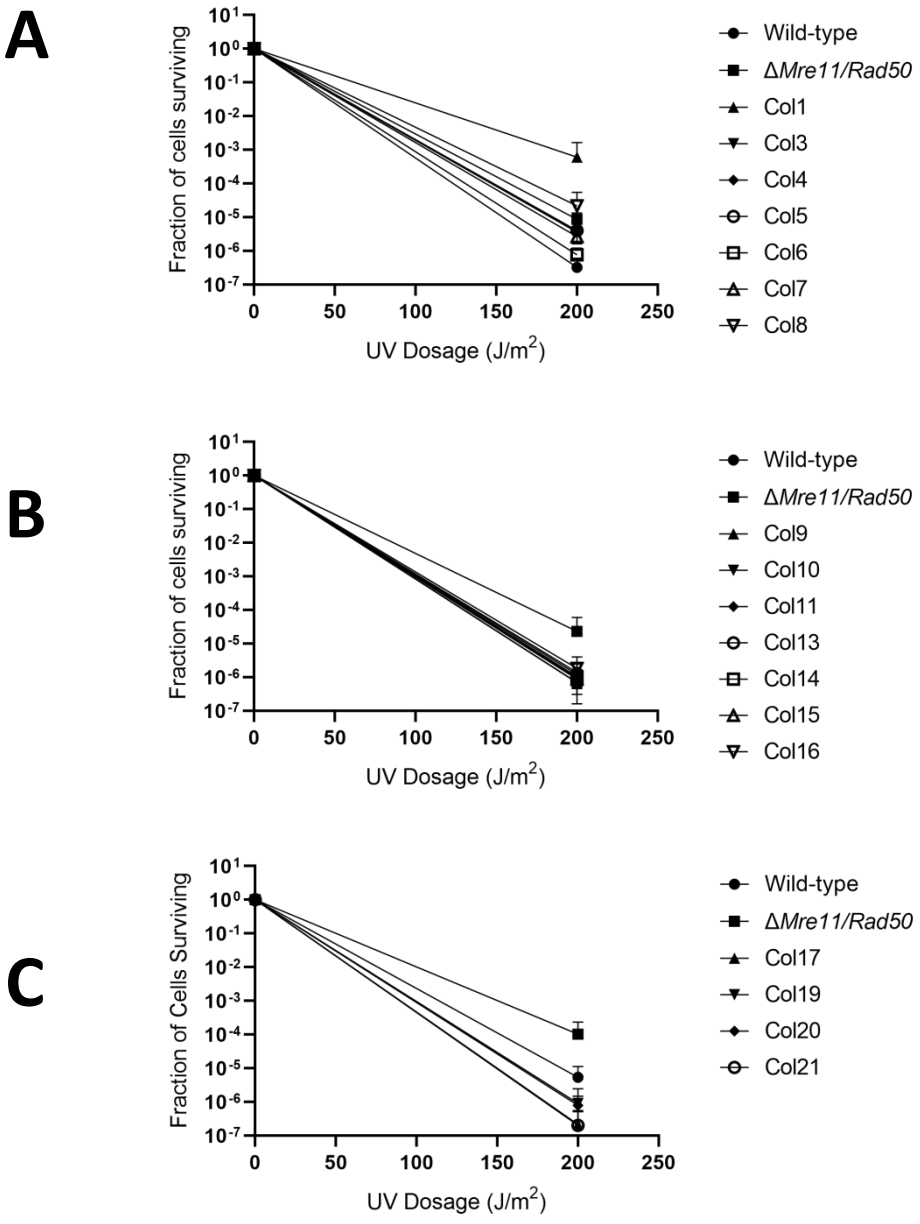


Figure 3.21 | **UV survival validation assay using direct plating**. Direct plating method was used to generate survival data for the 21 colonies from the final screen. Colonies 1-8 (A), 9-16 (B) and 17-21 (C).

Although ring formation didn't occur, counting still represented a hurdle due to the numbers of colonies on the plates and the UV dosage leading many colonies to appear very small indeed and almost translucent, making counting very difficult indeed. Further to this, many colonies appeared as tiny and colourless – reminiscent of those encountered during the screening stages, which were then discounted. Attempts were made to streak these out, and none grew out. However some of the intermediate colonies did (those that looked slightly more healthy, but still very small). It was therefore very difficult to distinguish between these two types of colony, and so counting survivors was not as accurate as it could be. In a final attempt to use this method, however, the direct plating method was used with lowering the UV dosage to 150J/m^2 (figure 3.22). In spite of this, the results were still inconclusive and the controls couldn't be counted due to saturation on plates and the inability to accurately and consistently distinguish between colonies that may or may not grow out when streaked.

Therefore this method was also abandoned. However, figure 3.21a shows colony 1 as a potentially interesting candidate for future analysis. Should this agree with future data with new validation methods, this colony should be studied further.

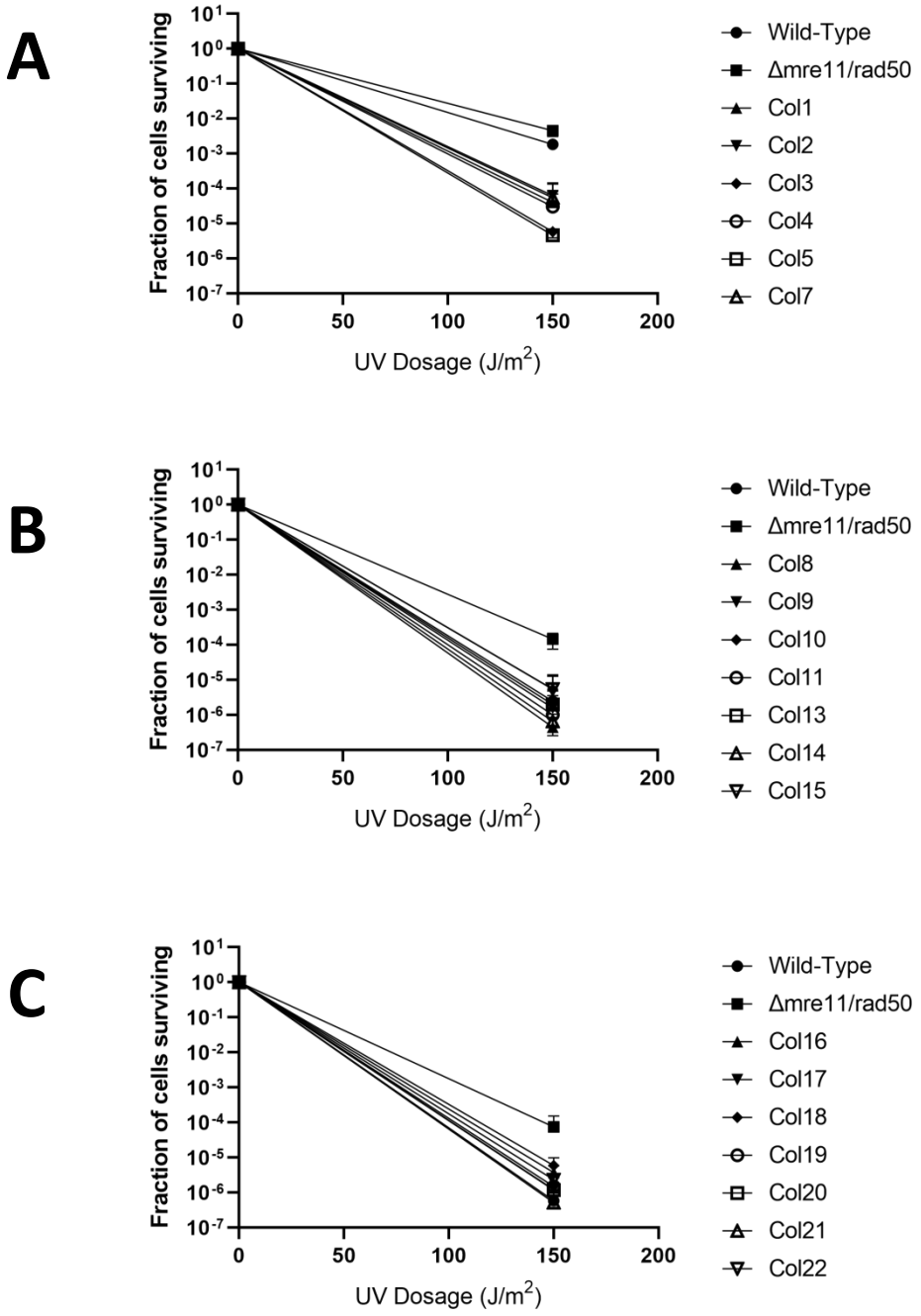
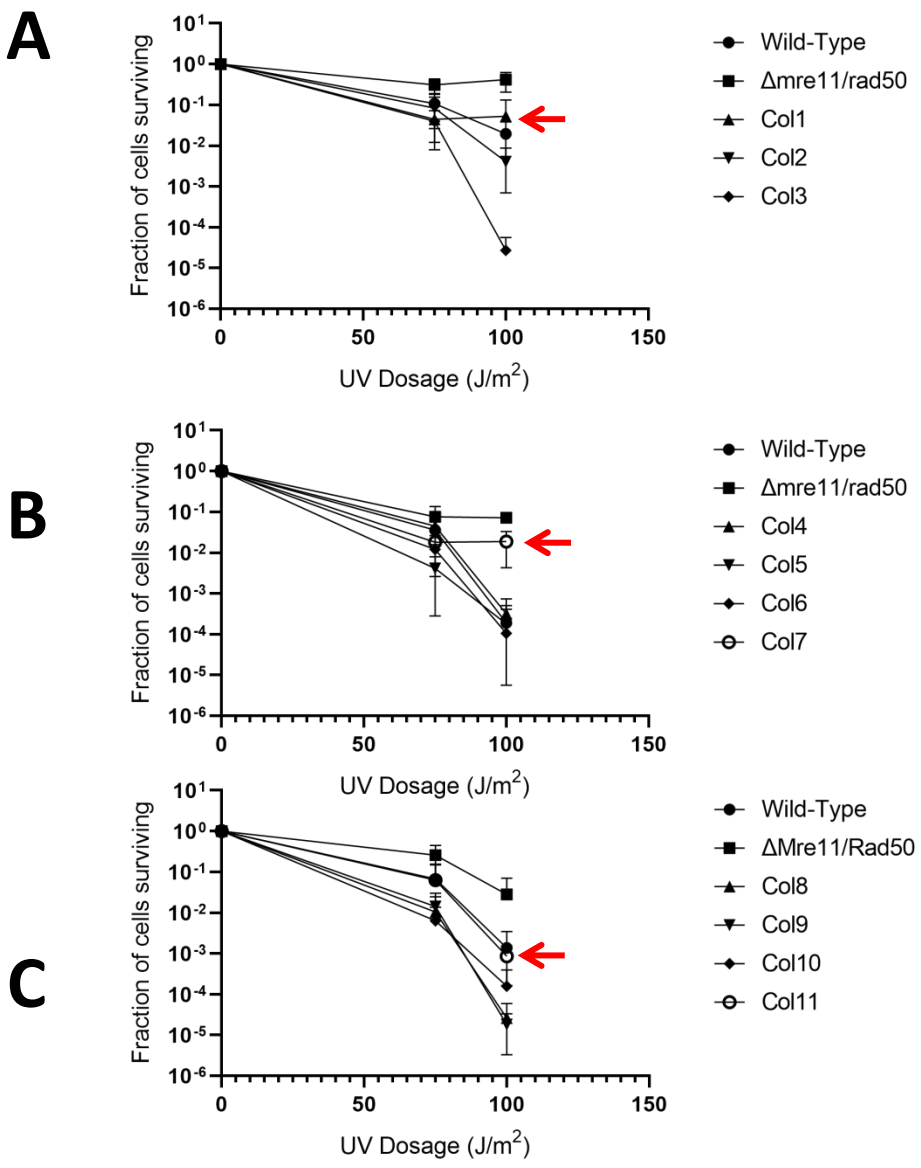


Figure 3.22 | **UV survival validation assay using direct plating and lowering UV dosage**. Direct plating method used again but UV dosage lowered to 150J/m2 in an attempt to reduce 'sick' colonies and allow more accurate quantification of surviving colonies. Colonies 1-7 (A), 8-15 (B) and 16-22(C).

Given that direct plating yielded no fewer issues compared to previous methods, the decision to revert back to spotting was made. This technique works, and was the original technique used for all calibration assays, so is robust and well-used. The main issue was mitigating ring formation, thus allowing for accurate counting of different dilutions of culture on UV-treated plates. This was achieved through lowering the UV dosage so that it remains high enough to still show a difference compared to the control cultures, but not high enough to cause ring formation to occur. The results can be seen in Figure 3.23 and appear to work well, as the control colonies generally sit in the correct location based on previous data, and some putative colonies can be selected from this assay as a result. Selected colonies are indicated and were chosen as the levels of resistance were higher than other colonies in the assay and showed similar resistance levels to the Mre11/Rad50 deletion strain.



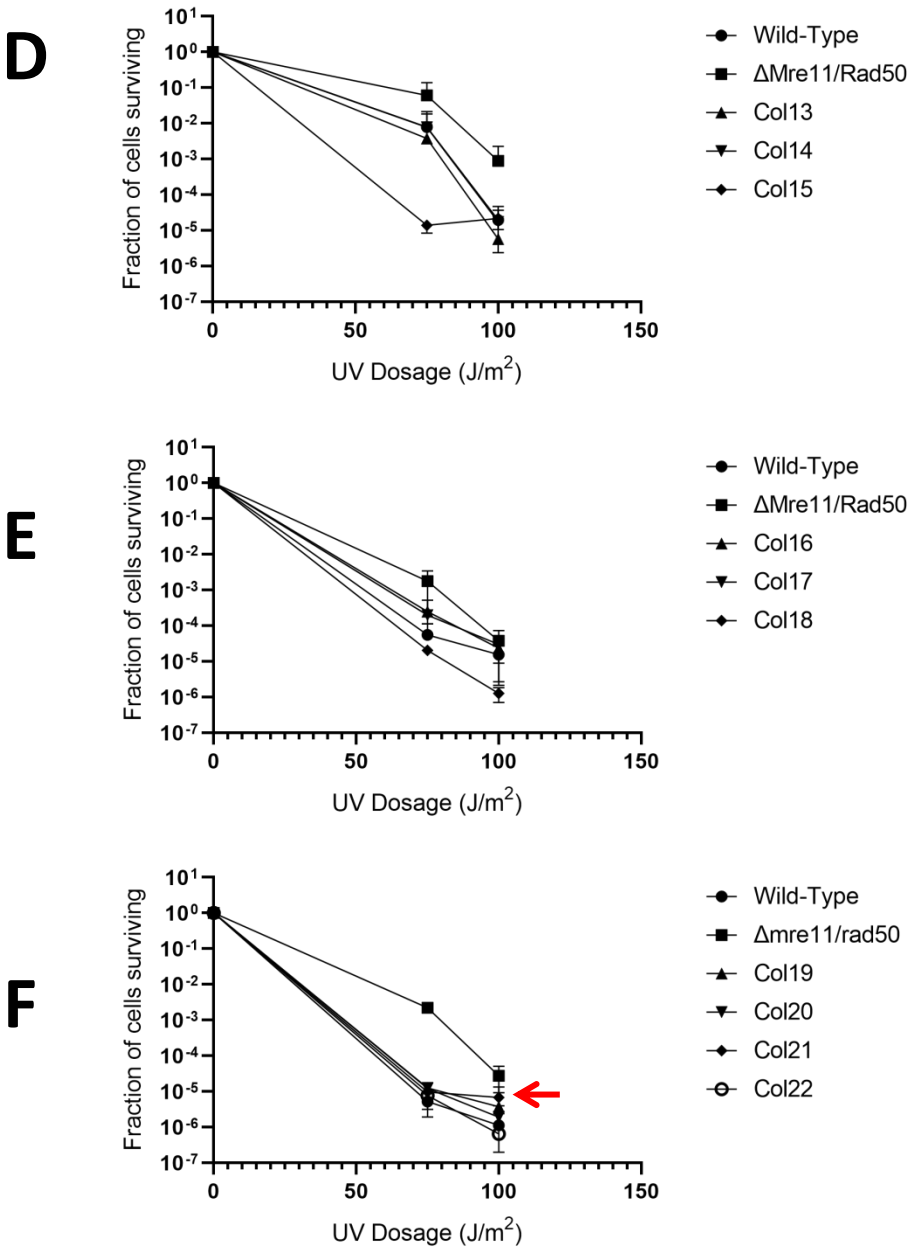


Figure 3.23 | **UV survival validation assay using spotting and lowering UV dosage**. Colonies 1-3(A), 4-7(B), 8-11(C), 13-15(D), 16-18(E) and 19-22(F).

Given the above, this assay represents the final validation attempt on the 21 colonies isolated from the screen, and the assay from which putative colonies of interest are selected for downstream analysis in subsequent chapters. These colonies are colonies 1, 7, 11 and 21. Of note, colony 1 appears to be resistant, as also shown by the previous survival curve in figure 3.21a.

3.4.8 Analysis of *Haloferax* ring formation

One hypothesis to come from the ring formation issue during the validation of candidates from the genome library screen is that, similar to *Sulfolobus acidocaldarius*, UV light may induce expression of gene(s) involved in pili formation (van Wolferen *et al.*, 2013). Therefore, to test this, strain H1895, which is deleted for the gene encoding the adhesion pillins PilB and PilC (Δ*pyrE2*, Δ*hdrB,*Δ*mrr,*Nph-pitA, cdc48d-Ct, Δ*pilB3C3)* was spotted alongside wild-type and resistant strains, H1206 and H645 respectively. The prediction being that a pilli mutant will not be able to form rings.

Upon irradiation with UV, UV dosage above 75J/m^2 appeared to form rings in all strains, showing that it is not the result of UV-induced pili, and is therefore some other biological artefact.

A second potential cause could be UV-induced phage activity. Notably, strains lacking the prophage region have been shown in this study to be resistant to UV irradiation, and ring formation seemingly occurs at a far reduced frequency. See chapter 5 for more details.

3.5 Discussion

3.5.1 Radiation survival assay

Radiation survival assays with both ultraviolet and gamma radiation indicate that the levels of radiation to be used for screening the genome library should be 300J/m^2 and 1500 Gy respectively. Whilst UV radiation provides substantial killing efficacy and good discrimination between wild-type and resistant cells, the nature of the gamma source and the long incubation times for irradiation mean that the killing seen via this method is much lower than would be required to screen the library. Therefore a revision of strategy was carried out and gamma no longer provided a means to screen the library, it will now only be used as a tool on candidates selected via UV initially. The benefit to UV for screening is also that irradiation will be combined with plating stress on the cells, allowing further killing on solid media rather than in liquid culture. Liquid culture would not be sufficient due to the 'shielding' provided to cells at the centre of this culture, meaning that cells on the periphery receive the majority of the UV light. Any cells in the centre of the culture will get a far smaller dosage.

3.5.2 Pairwise competition assay with gamma radiation required a revision of strategy

The original strategy for screening the genome library was purely to look at resistance under ionizing radiation, such as gamma radiation. Initial competition assays were carried out (section 3.4.5) however initial results provided inconclusive results due to a lack of intermediate time points within the data set. A subsequent assay was carried out with smaller time points, however the data didn't demonstrate that resistant cells had much of an advantage over wild-type. A red flag was that the amount of killing didn't appear to match the data provided by the initial survival assay in section 3.4.2. It was thought that the nature of the gamma source caused this apparently aberrant dataset to come into being for a number of reasons:

I. The nature of the competition assay and gamma source meant that it would be a requirement to irradiate liquid cultures, rather than cultures on solid media as per the initial survival assays in section 3.4.2. It has been shown that plating for Archaeal cells represents a stress factor in its own right, therefore any phenotype screened for in terms of survival, is actually a combination of resistance to both radiation stress and plating stress combined, in the case of section 3.4.2. Liquid cultures being irradiated in the gamma source therefore have multiple hours before being plated, therefore the plating and irradiation stresses are temporally separated, unlike the initial survival assay where plates are irradiated as soon as cultures are placed on the agar [unpublished data, Thorsten Allers].

II. Irradiation in liquid culture, especially of large volumes such as 1ml, allows irradiation of only some of the cells within the culture. Therefore not all cells get irradiated and those that do will not be irradiated with the same dosage. Therefore cells within the culture remain unirradiated and therefore incubation time after irradiation, in which cells can begin to recover, begins, for those cells, as soon as the culture enters the gamma source, leading to their outgrowth at time 0hrs after irradiation actually being approximately 5hrs. Ideally, all cells should be irradiated with the same dosage of radiation, after which recovery can occur. This explains a possible scenario underpinning the rapid growth of cultures, and lack of initial cell death at time point 0hrs.

III. Another potential issue using the gamma source is the nature of the radiation dosage. As the gamma source irradiates over a long period of time (up to 5.5hrs in this case), the issues above are increased due to chronic exposure of lower doses of radiation. Furthermore, cells may be more tolerant to prolonged dosage of irradiation at lower doses, rather than higher doses delivered in a shorter time frame due to ongoing DNA repair mechanisms that will be activated in response to the genotoxic insults provided by the gamma rays. The dosage may not be strong enough for the rate of damage to overcome the rate of repair to a sufficient degree to see substantial cell death.

IV. Irradiating the genome library, which are on plates, would be impractical in the spatially restricted gamma source. The genome library would therefore have to be irradiated in liquid culture, which lends itself to the issues described above.

As a result of the above and data retrieved from the initial competition assay using gamma radiation, a revision of the strategy was required. It was decided to switch to ultraviolet radiation initially, owing to the ability of delivering higher doses of UV in much shorter time windows, as compared to gamma radiation. This would minimize the opportunity of cultures to begin to regenerate whilst being irradiated. Furthermore, liquid cultures can be agitated whilst being irradiated, to ensure equal cell irradiation. Therefore the new strategy was to first irradiate the genome library on plates, using 300J/m^2 UV. The nature of the UV lamp, compared to the gamma source, meant this was a much more pragmatic scenario with which to screen thousands of colonies. Ultraviolet radiation was decided upon for both ease, as well as the nature of the DNA lesions imposed upon the library. Although, as discussed previously, UV light leads to CPDs and 6,4-photoproducts as predominant lesions, single-stranded DNA breaks also occur, both directly, and indirectly via oxidative damage.

Single-stranded DNA breaks have been shown to lead to replisome disassembly and thus, more toxic, double-stranded DNA breaks are therefore formed, a very similar end-result as is caused by ionizing radiation. (Tyrrell, 1995, Vrtis *et al.*, 2021). Overall single stranded DNA breaks are more common than double stranded breaks and it is estimated that in mornal human cells, 1% of ssDNA lesions are converted to ~50 double stranded DNA breaks per cell, per cell cycle (Vilenchik and Knudson, 2003). Once the genome library has been screened via UV irradiation, perhaps through multiple iterations, gamma radiation will be employed once cell numbers are appropriate enough to transfer to plates for irradiation. The benefits of using this approach are not only that both types of radiation will be used, and therefore colonies arising from the screen will be resistant to both types of radiation which depicts the reality of extra-terrestrial conditions, but also genetic factors underpinning resistance to either radiation types, but not both, can be deciphered. This perhaps will shed light on novel mechanisms used in order to adapt to both radiation stresses together, and individually.

3.5.3 Pairwise competition assay with UV

The pairwise competition assay was repeated with UV irradiation, in liquid culture. Although data suggest that wild-type cells recover more quickly from irradiation, there is no evidence through this assay that points to a potential issue for out-competition of resistant cells by wild-type that remain after irradiation. In spite of this, and the fact that the genome library will not spend much time as resuspended

culture (≤1hr), cell death was still not as expected through the survival assay in section 3. Reasons for this appear to be the fact that too much culture was irradiated in a petri plate, without agitation of the culture throughout. A revision of this method was carried out. Changes were that two 500µl cultures were irradiated, and then pooled together, instead of a single 1ml culture. Furthermore, the cultures will be placed on a rocker under the UV lamp in order to ensure maximum and equal irradiation of all cells within each aliquot of cell culture. This revised method provided a much better set of results and showed that after approximately 20-24hrs, a single colony that survives stochastically would return to previous levels prior to irradiation. This allowed the final screening attempt to take advantage of this and produce a harsher and more appropriate screen.

3.5.4 Genome library screen with Ultraviolet radiation

The initial stages of the screen posed a number of issues, with many more colonies returning from the screen than planned. This was due to the nature of the screen and the fact the UV lamp, which was used for the first screening attempt, using 300J/m^2, was actually much weaker than thought. This resulted in the UV lamp breaking just as the second screen was about to occur (after a repeat of the cloning efforts, details not shown here). As a result all calibration assays needed to be repeated using the UV DNA crosslinker, which provided a more powerful dosage of UV irradiation to the cells, however meant that the pairwise competition assay couldn't be repeated, due to the size limitations imposed by the crosslinker. Eventually, all creases were ironed out and the library was transformed, using the GLUE algorithm for colony numbers, and screened using the appropriate UV dosage, which for the crosslinker was 200J/m^2. Three iterations were carried out, 20hrs apart, to stop any stochastic outgrowth, as proven by the competition assays.

The main issues came with validation of colonies obtained from the screen. These issues were mainly the inability to count the surviving colonies accurately, mainly due to ring formation under UV stress. Interestingly, rings do not form under non-UV conditions and do not appear to be pili related, so the situation remains unclear. Ring formation certainly is dosage dependant with UV, so could be some alternative biological response to UV damage, or, more likely, an artefact of heating, or some other change, caused by UV irradiation. Overall, these problems

were eventually limited such that appropriate survival statistics could be generated so that putative resistant colonies of interest could be chosen.

Furthermore, Hel308 was assayed for the feasibility of providing a positive control for the library, however it was, as expected, toxic when overexpressed therefore can not act as a control.

As a result of the screen, eventually 21 colonies were analysed for UV resistance compared to wild-type and resistant backgrounds, and as a result, four colonies were taken further for downstream phenotypic and biochemical analysis.

Chapter 4: Genetic and Phenotypic analysis of Overexpression Library Screens in *Haloferax volcanii*

4.1 Background

Genetic and phenotypic analysis of constructs identified must be carried out to confirm what coding sequences are contributing to the resistance phenotype seen. It is further prudent to confirm the resistance phenotype is not due to variation present on the vector backbone, and is directly a result of the coding sequences downstream of the tryptophan inducible promoter. Once this has been verified, and genes identified by sequencing, it is important to narrow down the choice of coding sequence to identify the exact gene responsible, before then reconstituting the resistance phenotype, using the coding sequence for such genes only, by amplifying separately from the genome via PCR. If such a phenotype is seen, especially if the effect is larger when the promoter is altered to a more constitutive or more active version, such as p.*tnaA*, then a deletion construct must be made to assay the importance of this gene to DNA repair systems. Also importantly, any inference made to explain the resistance phenotype seen must be tested to tease out any potential mechanisms by which a given gene is acting to increase levels of UV resistance.

4.2 Aims and Objectives

The aim of this chapter is to take colonies displaying a potentially UV resistant phenotype, isolated from the genome-wide screens in chapters 3, and to carry out phenotypic studies, clarifying the nature and extent of their apparent resistance to UV and/or 4-NQO. Genetic and bioinformatic analysis is also carried out in this chapter for the colonies isolated to further investigate potential mechanistic reasons behind the phenotype shown. Therefore, the aims for this chapter are as follows:

- Sequence the four colonies isolated to determine the nature of the genes contained within each fragment.

- Map each fragment to the genome of the laboratory strain, H26, from which the library fragments were derived.

- Carry out tryptophan dependence assays to confirm the resistance phenotypes show are as a result of genetic material downstream of the p.*tnaA* promoter.

- Re-insert each fragment into a H26 background, to confirm the resistance phenotype is due to the genome library vector, and not mutations in the background host.

- Carry out growth assays to determine if overexpression constructs have an impact on out-growth after UV irradiation

- Generate a deletion construct for any genes of interest and confirm phenotype.

- Perform DNA damaging assays using UV and MMC

- Place genes of interest under a constituitive promoter, p.syn, to confirm resistance phenotype

- Complement defects in Nucloetide Excision Repair and Homologous recombination resolvase activity to determine if constructs can complement the defect and therefore could act in those pathways.

- Generate an expression construct for candidate gene(s) to allow for protein overexpression and purification.

- Place a StrepII tag at the C terminus of candidate genes to allow for future purification of the protein product via use of a Strep-Tactin column.

4.3 Materials

4.3.1 Strains

[ ] Indicate integrated plasmid DNA

{ } Indicate episomal plasmid DNA

Table 4.1 ***Haloferax volcanii*** **strains used in this chapter**

Strain	Parent	Genotype
H1206	H18	Δ*pyrE2,* Δ*mrr*
H645	H640	Δ*pyrE2*, bgaHa, Δ*mre11-rad50*
H53	H47	Δ*pyrE2,* Δ*trpA*
H509	H477	Δ*pyrE2,* Δ*uvrA*
H1181	H1161	Δ*pyrE2,* Δ*uvrB*
H1187	H1175	Δ*pyrE2,* Δ*uvrC*
H514	H482	Δ*pyrE2,* Δ*uvrD*
H358	H338	Δ*pyrE2,* Δ*hef*
H178	H158	Δ*pyrE2,* Δ*hjc*
H26	H18	Δ*pyrE2*

4.3.2 Plasmids

Table 4.2 **Plasmids used in this chapter**

Name	Use
pTA908	Overexpression vector with pHV2 origin and tryptophan inducible promoter. Shuttle vector containing ColE1 origin and Amp^R/pyrE2 selectable markers.
pTA927	Overexpression vector with pHV2 orgin and tryptophan inducible promoter. Shuttle vector containing ColE1 origin and AmpR/pyrE2 selectable markers. t.Syn terminator sequence added.
pTA1231	Overexpression vector for Hel308
pTA1992	Overexpression vector with p.syn promoter. Used to place XerCD like integrase under p.syn.
pTA2773	p1992 with XerCD-like integrase (HVO_2259) inserted using Pci1/BamH1 sites under p.syn promoter.
pTA131	Destination vector for deletion constructs.
pTA2783	Xer deletion construct. Non-trp marked.
pTA2784	Xer deletion construct. Trp marked.
pTA2786	Xer deletion construct. Non-trp marked. Dam -
pTA2785	Xer deletion construct. Trp marked. Dam -
pTA2863	XerCD placed under native promoter with native pHV1 origin (not high copy number). Shuttle vector. Used for complementation of XerCD deletion strain.
pTA2864	XerCD placed under native promoter with native pHV1 origin (not high copy number). Shuttle vector. Used for complementation of XerCD deletion strain. Dam-
pTA354	Used to generate p2863. Backbone vector with pHV1 origin.
pTA2780	p1992 with XerCD-like integrase (HVO_2259) inserted using Pci1/BamH1 sites under p.syn promoter. Dam -
pTA2781	p.syn::Xer::Strep construct.
pTA2782	p.syn::Xer::Strep construct. Dam -

pTA298	For generation of TrpA marker.

4.3.3 Oligonucleotides

Table 4.3 **Oligonucleotides used in this chapter**

Name	Use	Sequence
PtnaAFint	Sequencing primer	5’-GCCTGCCGATTACTTCACATTCGC-3’
t.synR2	Sequencing primer	5’-GATCACGCCGAAAAATGCGATGGTCCAGAGGTGGATCCG-3’
XerC_F	Amplification of HVO_2259 including a 5’ Pci1 site.	5’ATCTGACATGTTCAAGGAGGGGAACTAACATG3’
XerC_R	Amplification of HVO_2259 including a 3’ BamH1 site	5’GTACGGATCCACCAACTGTCTATCCTCGTTA3’
XerC_R_Nhe1	Amplification of HVO_2259 including a 3’ Nhe1 site for addition of C-terminal StrepII tag	5’CAGATGCTAGCGAGTCCAATTTCCTTG3’
XerDel_F_Cla1	HVO_2259 deletion cloning	5’-ATCTGATCGATGAGACATCTGTGGACAGCCCC-3’
XerDel_R_BamH1	HVO_2259 deletion cloning	5’-ATCTGGGATCCGTTAGTTCCCCTCCTTGAGCCG-3’
XerDel_F_BamH1	HVO_2259 deletion cloning	5’-ATCTGGGATCCGGTCGAAGTCATTATTTATCGAAACCC-3’
XerDel_R_Not1	HVO_2259 deletion cloning	5’-ATCTGGCGGCCGCGAGCAACAATCCGTTCTCTGG-3’
pBSF2	Forward primer for colony PCR	5’-TTAAGTTGGGTAACGCCAGGG-3’
pBSR3	Reverse primer for colony PCR	5’-ACCCCAGGCTTTACACTTTATGC-3’
PhageProbeFII	Forward primer for colony PCR and probe generation	5’-CACCATCATCCTCGCCAAAAC-3’
6264R2	For amplification of XerCD from genome	5'-GGGACGCCGACCACTTC-3'
PhageProbeBII	Reverse primer for colony PCR and probe generation	5’-ACTCTTCTTTCGCTTCGTCACG-3’

PhageProbeR3	For amplification of XerCD from genome	5'-CAACGCCATTAGTCTGTCTGTAAGC-3'
XerC_F	Forward primer for amplification of genomic HVO_2259 with Pci1 site	5’ATCTGACATGTTCAAGGAGGGGAACTAACATG3’
XerC_R_Nhe1	Reverse primer for amplification of genomic HVO_2259 with Nhe1 site for ligation into pTA131 to insert StrepII tag at C terminus	5’CAGATGCTAGCGAGTCCAATTTCCTTG3’

pTA927:gLib fragments (pTA927 with genome library fragment 1, 7, 11 and 21 under p.*tnaA*)

pTA927 is a tryptophan inducible overexpression vector capable of replication in both *H. volcanii* and *E. coli*. Genome library fragments 1,7, 11 and 21 isolated from the initial screen under UV stress is under the control of the p.*tnaA* promoter. Genome library fragments represented by black arrow.

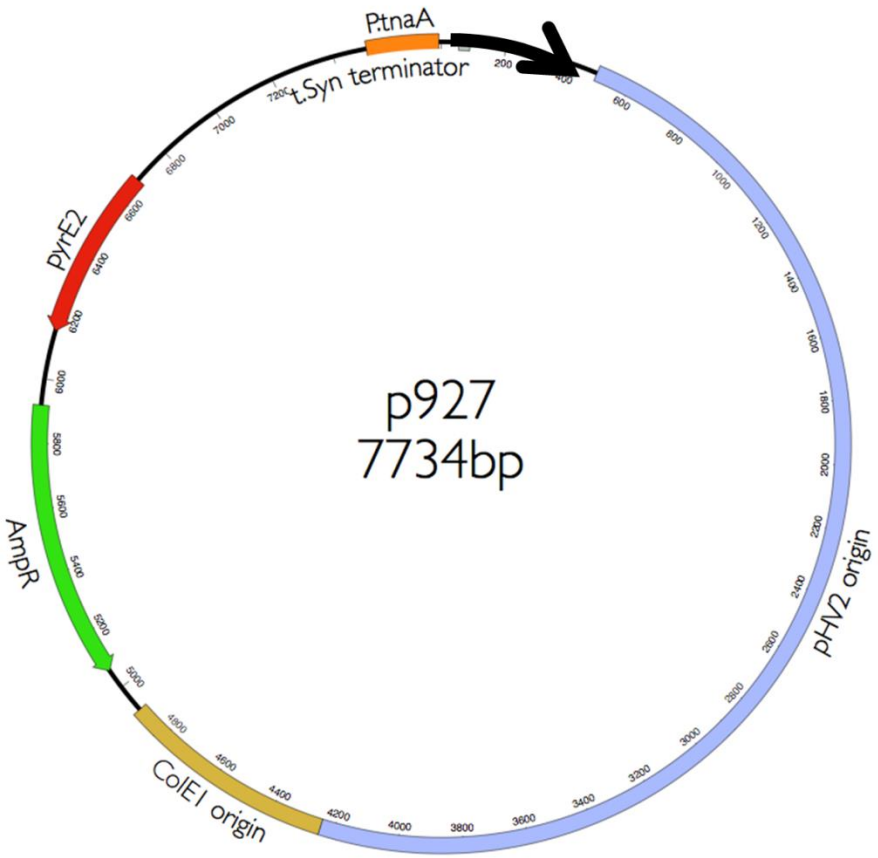


Figure 4.1 | **pTA927 gLib construct**. Constructed in this study.

pTA1992 (expression vector for p.syn-mediated expression and StrepII tagging)

pTA1992 is a shuttle vector that contains the constitutively expressed p.syn promoter sequence directly upstream of a MCS. HVO_2259 is placed in this vector to allow high levels of transcription, in the case of pTA2773, and a C-terminal strep II tag, in the case of pTA2781.

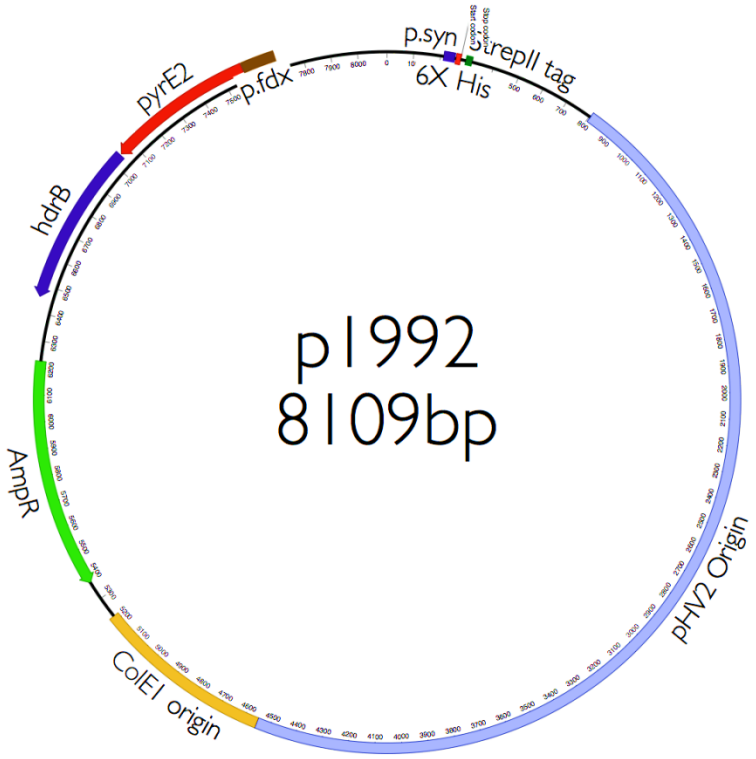


Figure 4.2 | **pTA1992**. Constructed by TA lab.

pTA2773 (p.syn:HVO_2259 expression construct)

pTA2773 was made to place XerC/D-like integrase, HVO_2259, under the constitutive p.syn promoter, driving high levels of expression. The construct is un-tagged.

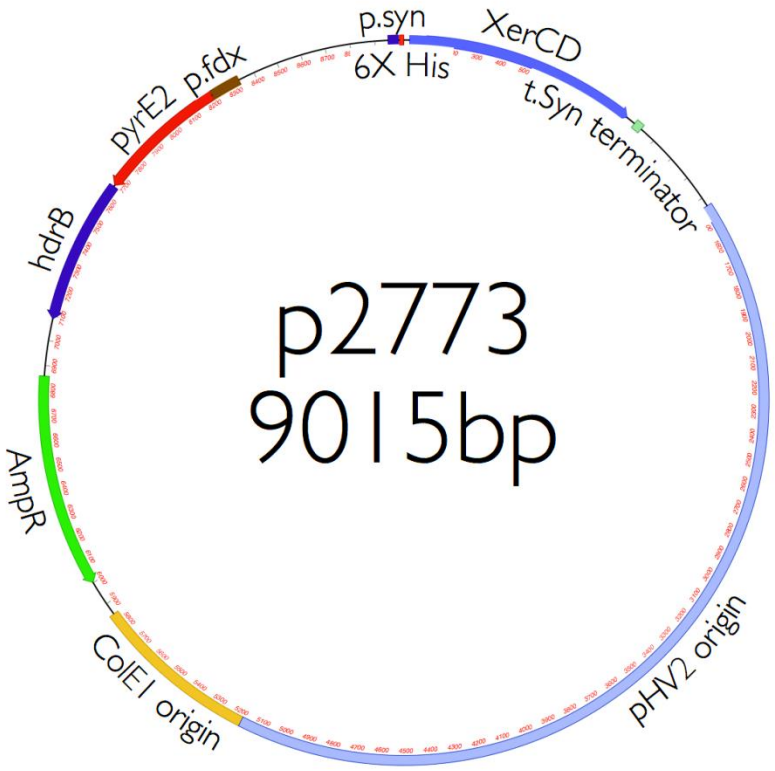


Figure 4.3| **pTA2773**. Constructed in this study.

pTA131 (suicide vector used for deletion construct synthesis)

pTA131 is an integrative vector, used for generation of deletion constructs to facilitate ‘pop-in’ and ‘pop-out’ mediated gene deletion in *H. volcanii*. The *pryE2* marker allows for auxotrophic selection.

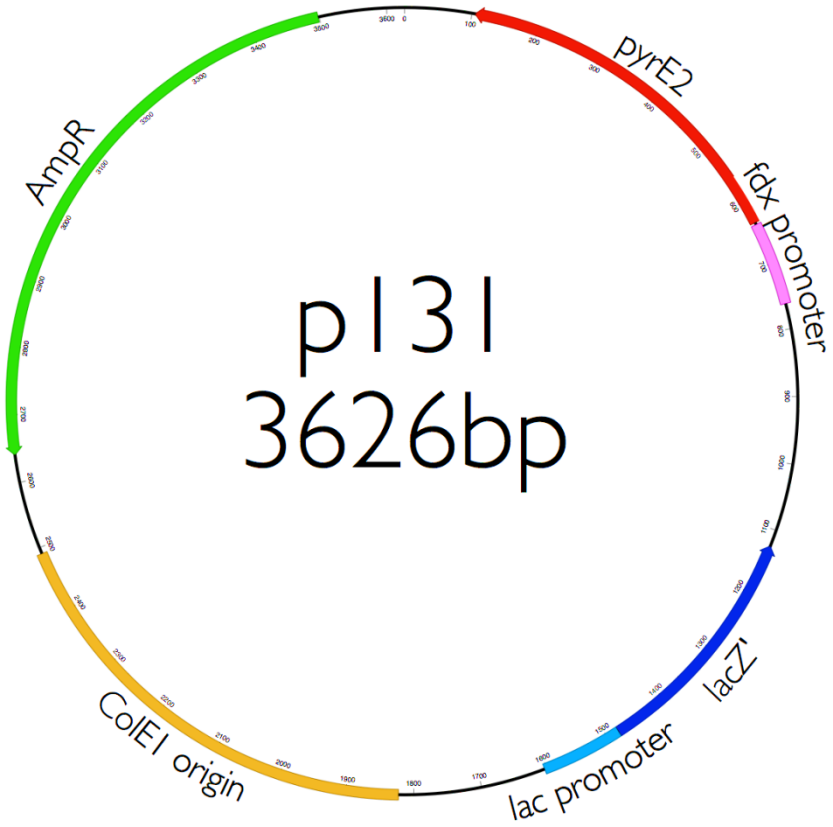


Figure 4.4| **pTA131**. Constructed by TA lab.

pTA2781

pTA2781 was generated to add a C-terminal strepII tag to the XerCD coding sequence. This allows purification in the future, should it be required via use of a Strep-Tactin column.

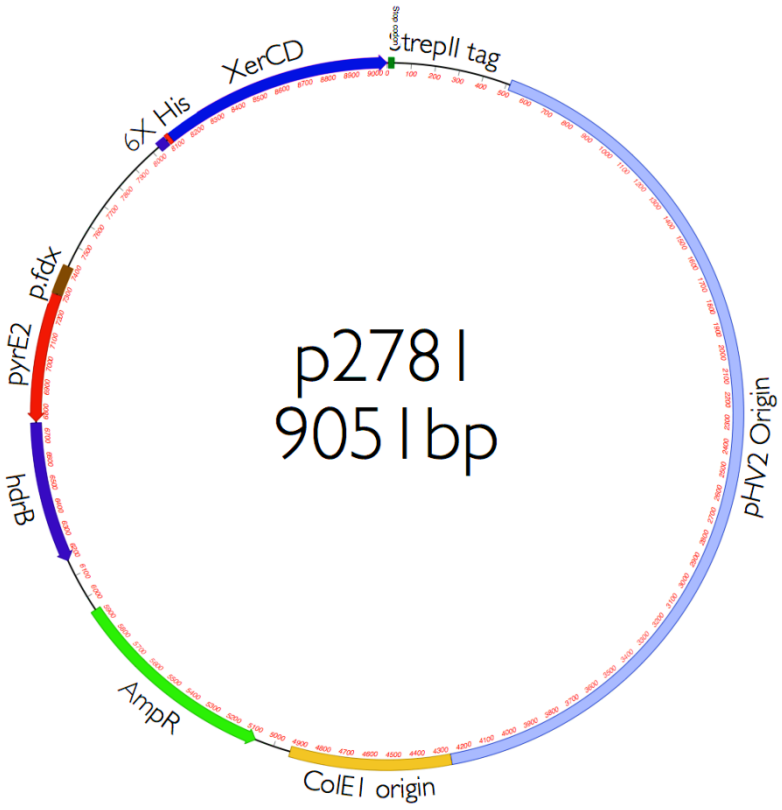


Figure 4.5 | **pTA2781**. Constructed in this study.

pTA2783

pTA2783 is the deletion construct generated for HVO_2259 (XerCD). Upstream and downstream homology arms have been ligated into pTA131 to allow for a subsequent “pop-in” in a Δ*pyrE2* strain.

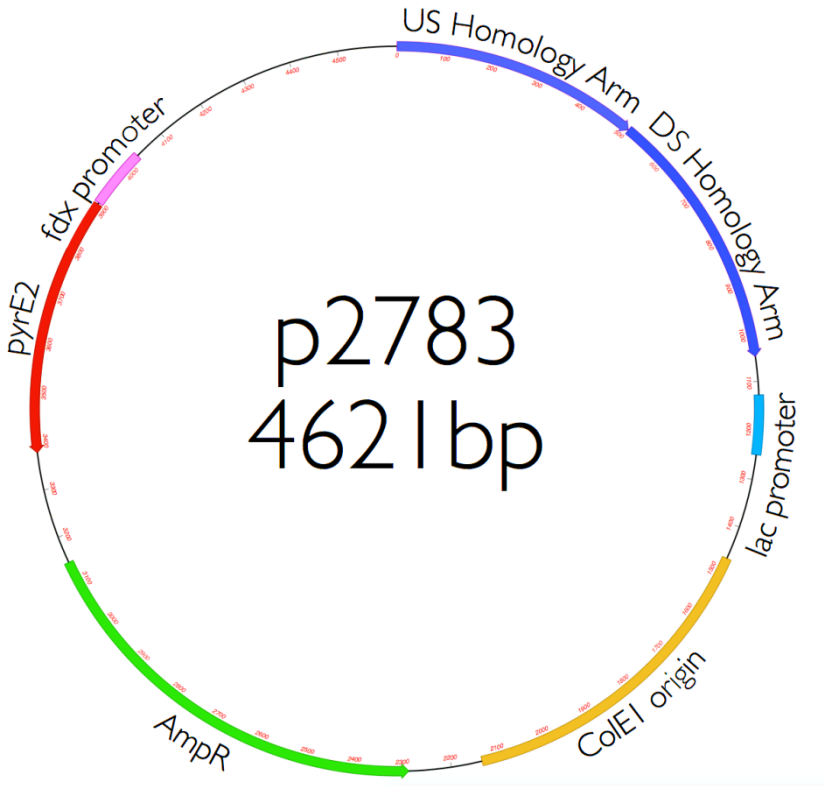


Figure 4.6 | **pTA2783**. Constructed in this study.

pTA2784

pTA2784 is a further XerCD (HVO_2259) deletion construct generated to be TrpA+ marked to allow for direct selection and to increase the efficiency of pop-out selection. Homology arms remain, as does the presence of *pyrE2*.

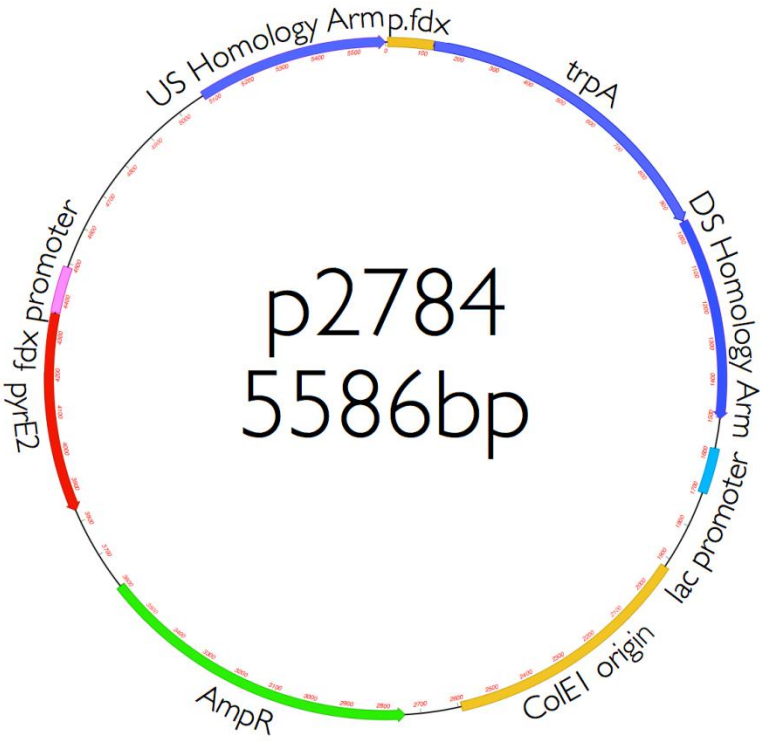


Figure 4.7 | **pTA2784**. Constructed in this study.

pTA2785

pTA2785 is a dam- version of pTA2784 above. This was generated by passing through a dam- strain of *E. coli* and allows for direct transformation into archeal strains that are not *Δmrr*.

pTA2786

pTA2786 is a dam- version of pTA2783 above. This was generated by passing through a dam- strain of *E. coli* and allows for direct transformation into archeal strains that are not *Δmrr*.

pTA2863

pTA2863 contains XerCD under its native promoter in a low-copy number origin background (pHV1). This construct is used to facilitate complementation of the Δ*xerCD* strain.

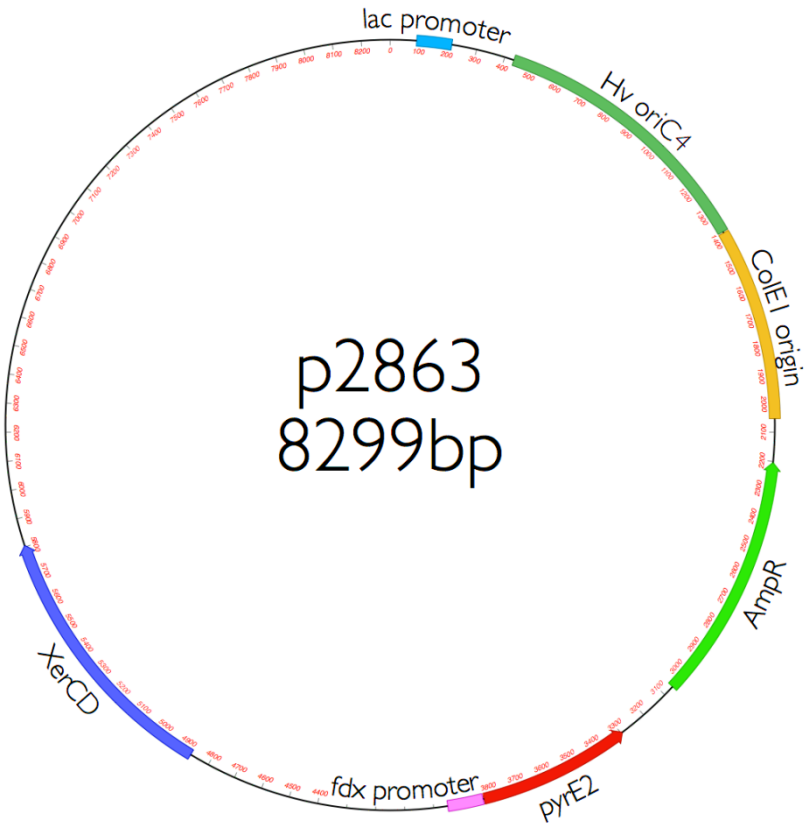


Figure 4.8 | **pTA2863**. Constructed in this study.

pTA354

pTA354 is used as a backbone for cloning to generate pTA2863. It contains the pHV1 origin of replication.

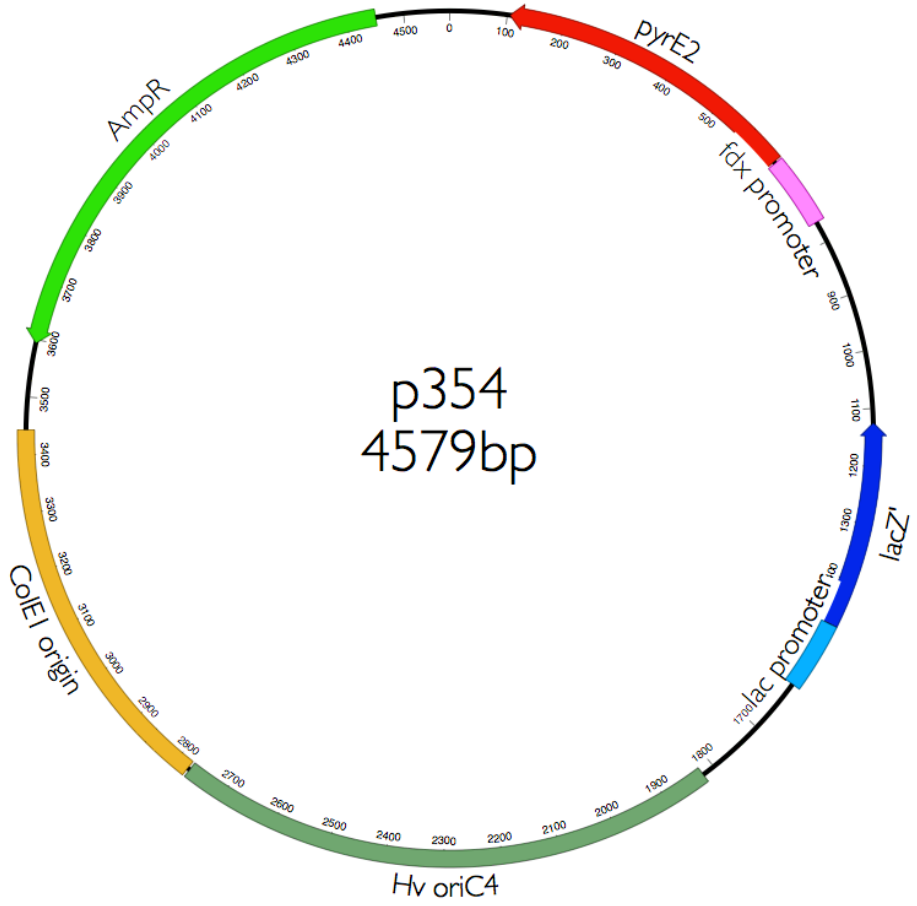


Figure 4.9 | **pTA354**. Constructed by TA lab.

pTA298

Construct containing trpA cassette. TrpA marker used in generation of TrpA marked deletion strain.

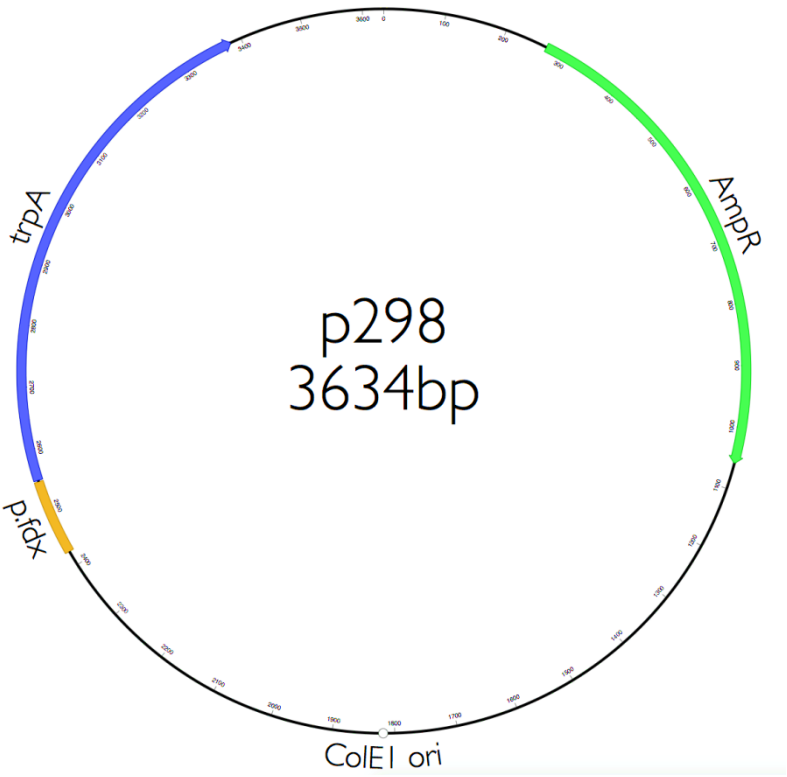


Figure 4.10 | **pTA298**. Constructed by TA lab.

4.4 Results

The aims of this chapter were to sequence all constructs obtained from the genome library screen and identify candidate genes of interest for further analysis. Tryptophan gradient analysis will confirm any UV resistance seen is due to elements downstream of the promoter, and transformation of constructs into a different host strain confirms host genotype was not a contributing factor to the UV resistance seen in the screen. Cloning of candidate genes of interest separately via PCR, and placement under a strong constitutive promoter will further confirm if future work is required and if that indeed is a gene of interest. Further experiments will aim to dissect how candidate genes are contributing to UV resistance. Any candidate genes will also be tagged at either the N or C terminus to facilitate future purification if required.

4.4.1 Sanger Sequencing of Genome Library Constructs

Colonies selected from the genomic library screens were then subjected to plasmid purification and subsequently were prepared for Sanger sequencing reactions (Sanger, Nicklen and Coulson, 1977). Plasmids were sequenced, using both sense and antisense strands, with respect to the p.*tnaA* promoter, by Deep Seq, University of Nottingham, using the following primers:

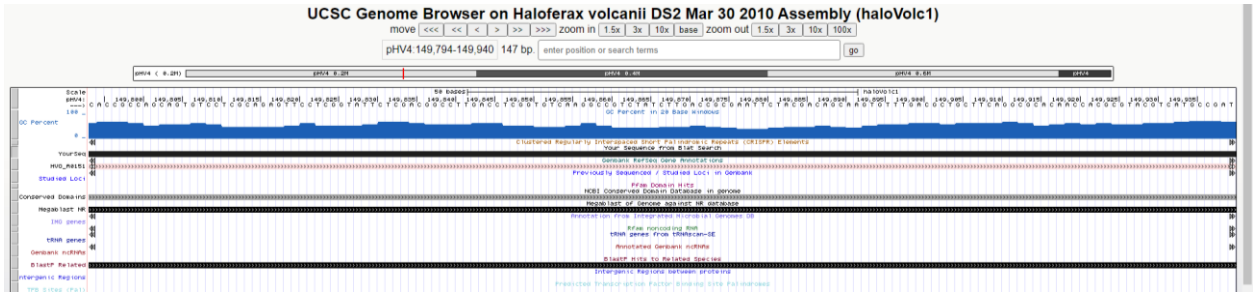
PtnaAFint (forward)

t.synR2 (reverse)

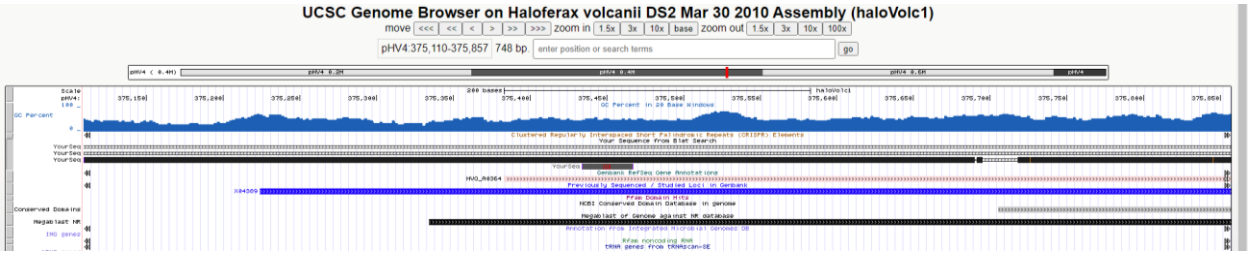
Fragments were read in both sense and antisense directions, and compared to the sequence trace to confirm accuracy of sequence. After sequences had been checked, genomic library fragments were then blasted to the *H. volcanii* DS2 genome and potential ORFs identified. Once this had been done, the sequence was then mapped to the genome to confirm the BLAST results (figure 4.4).

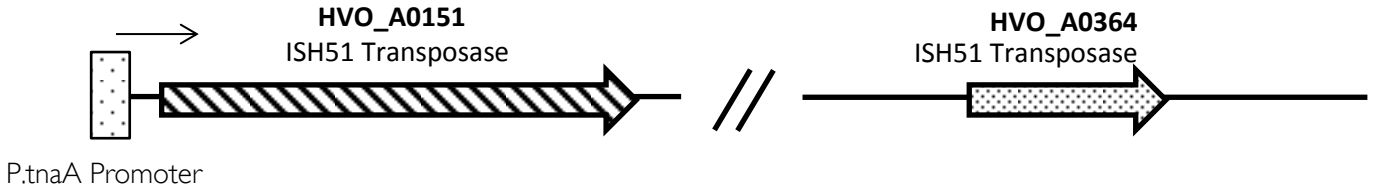
A

Genome Library Colony I
(Forward Sequence)



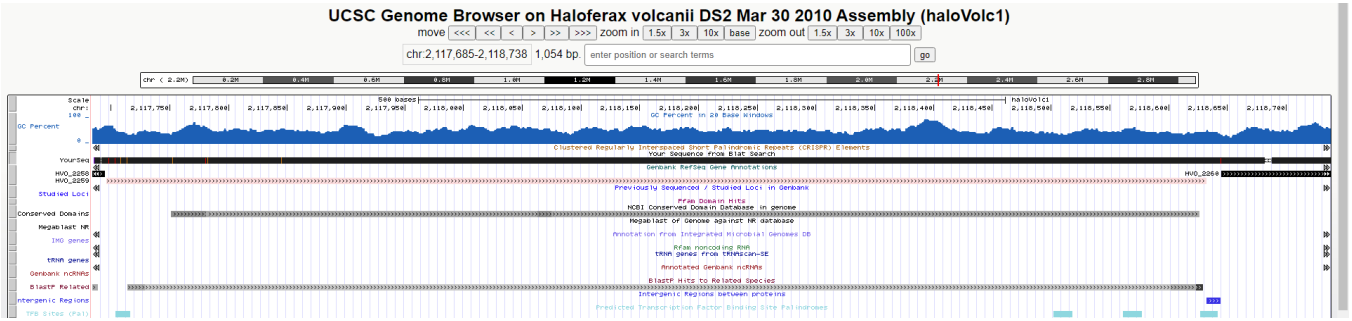
Genome Library Colony I
(Reverse Sequence)



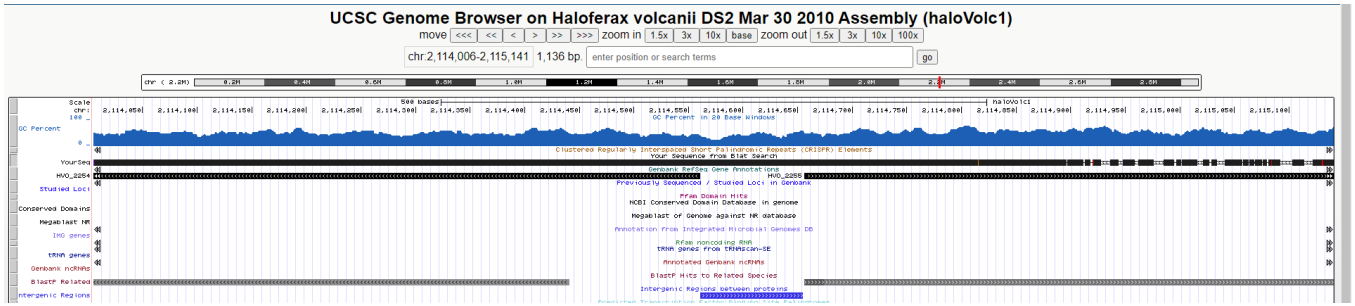


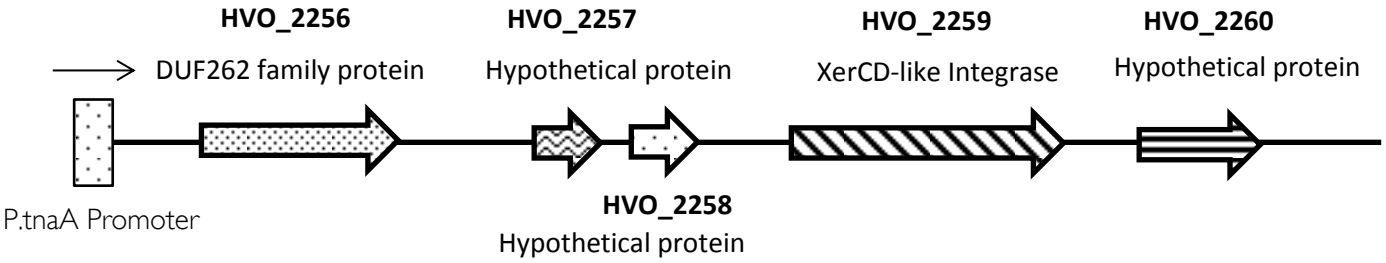
B

Genome Library Colony 7 (Forward Sequence)



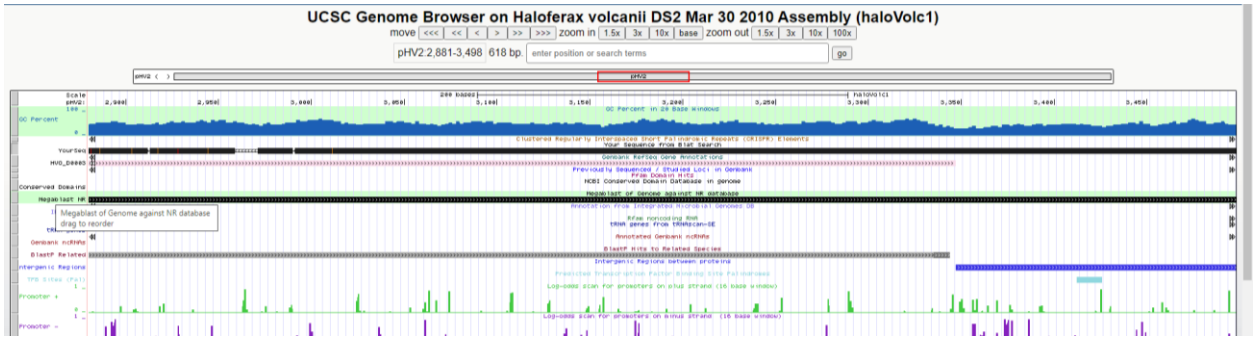
Genome Library Colony 7 (Reverse Sequence)



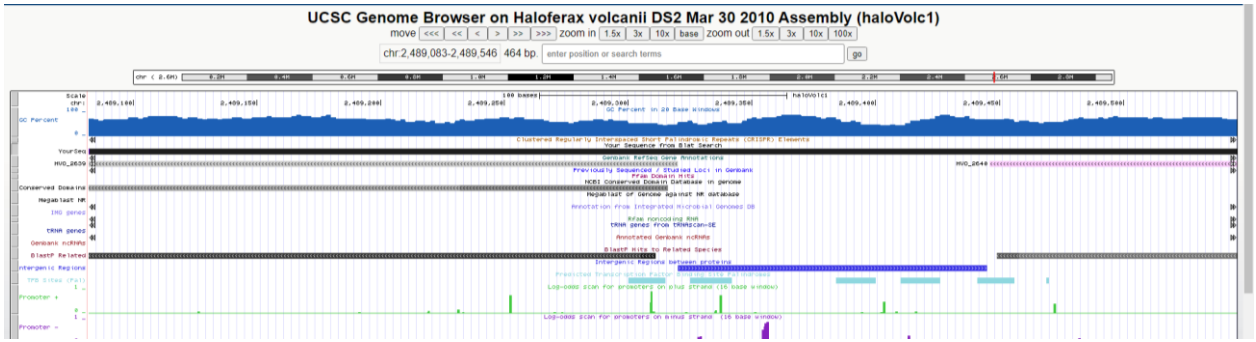


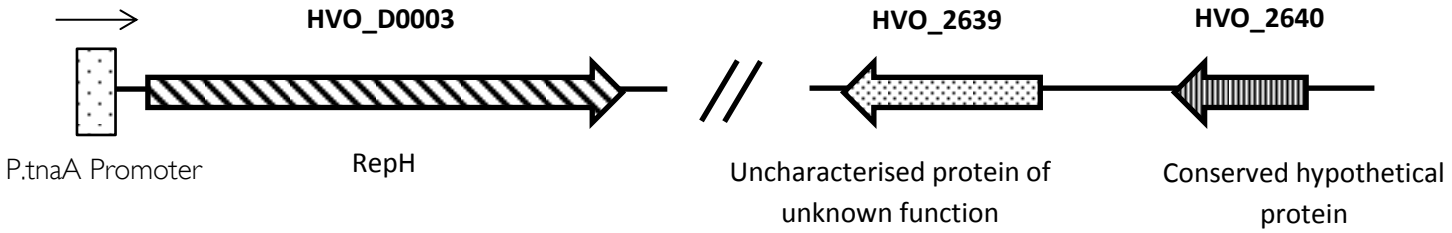
C

Genome Library Colony 11 (Forward Sequence)



Genome Library Colony 11 (Reverse Sequence)





D

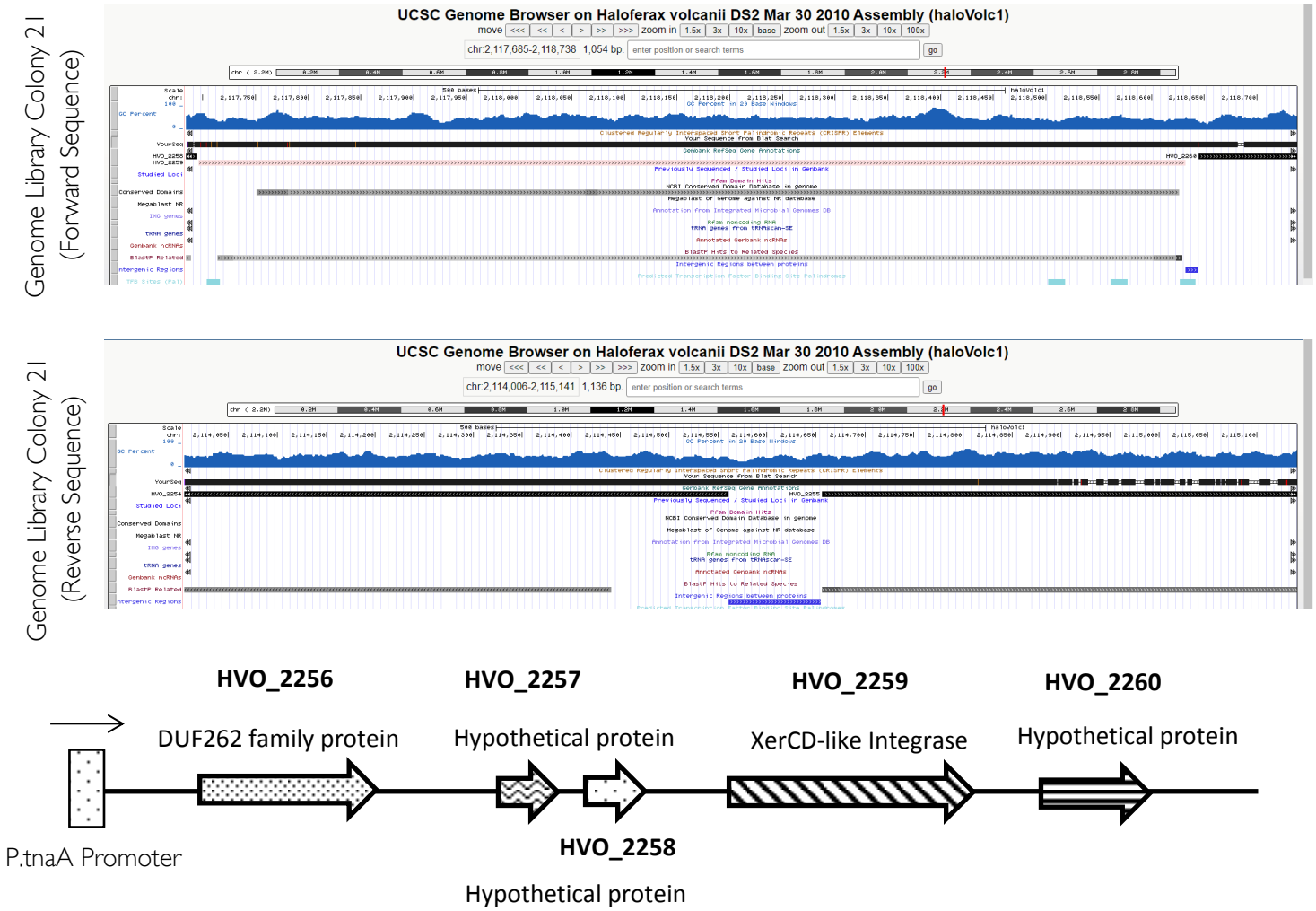
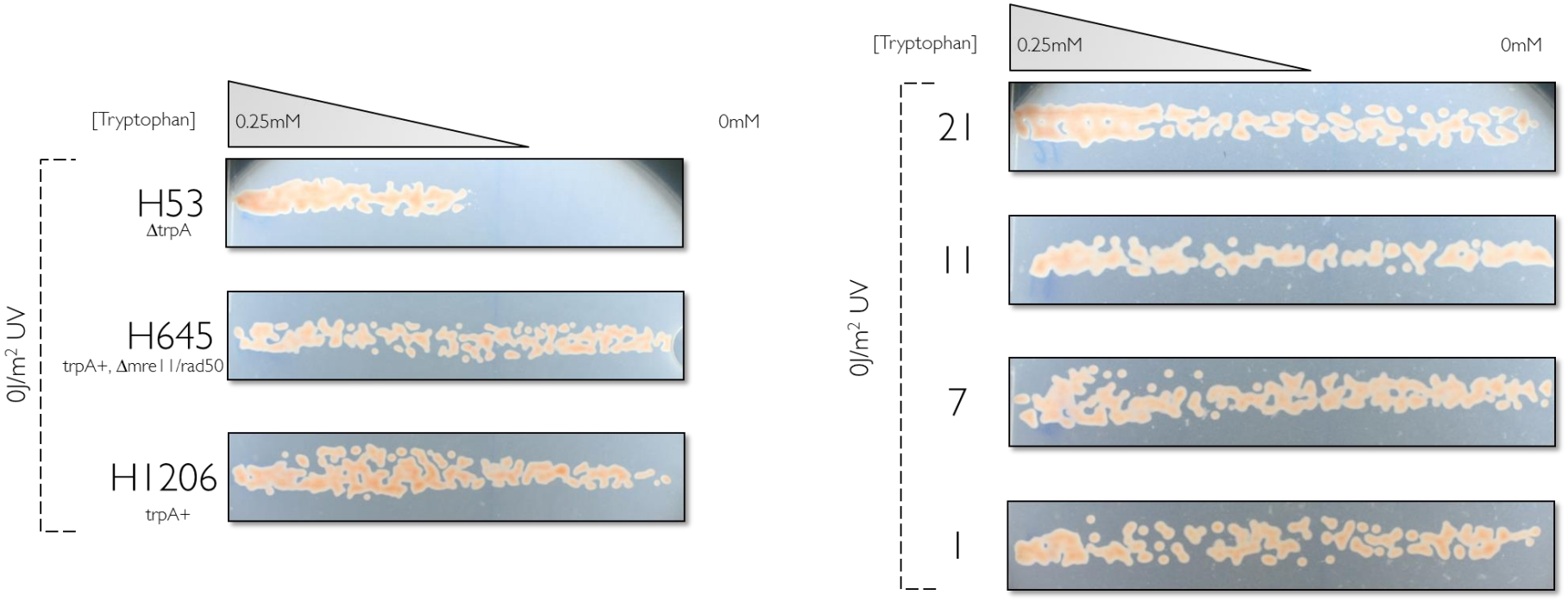


Figure 4.4 | **Genome Library Sequencing Maps to DS2 Genome**. Sequencing mapping to the genome of DS2 are presented in figures A-D for colonies 1, 7, 11 and 21 respectively. Of note, Colonies 7 and 21 contain the same fragment and colonies 1 and 11 contain more than one fragment(see text for discussion).

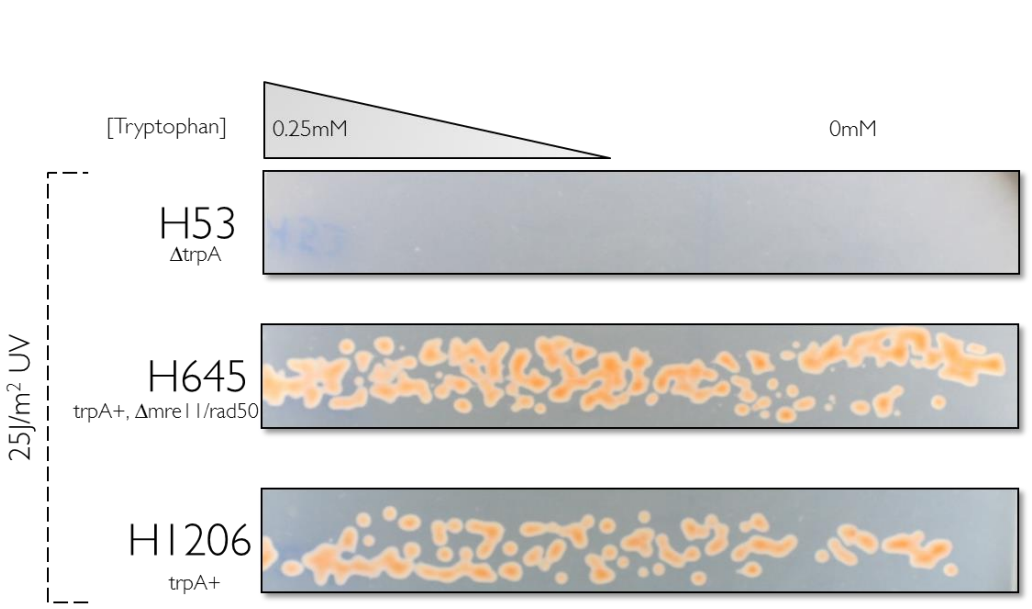
4.4.2 Tryptophan Gradient Assays

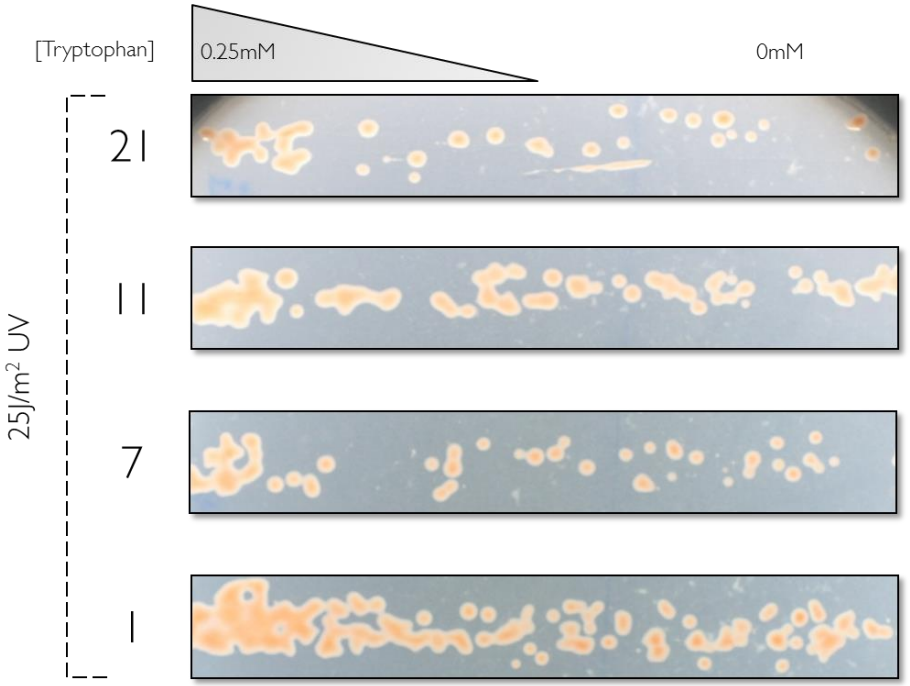
Haloferax volcanii strain H1206(Δ*pyrE2,* Δ*mrr*) was transformed with genome library episomes, gLib1, gLib7, gLib11 and gLib21. Cultures were then grown overnight and diluted for a subsequent overnight to reach A_{650} of 1.0 the next day. At each stage, selection was maintained by growing in Hv-Ca broth. The second overnight was grown in Hv-Ca supplemented with 0.25mM tryptophan to induce expression of the genome library fragments under the p.*tnaA* promoter. Control strains, H645(Δ*pyrE2, bgaHa,* Δ*mre11-rad50*) and H53 (Δ*pyrE2,* Δ*trpA*) were also grown under the same conditions. Overnights were serially diluted and then appropriate dilutions were plated using autoclaved paintbrushes along tryptophan gradient plates after being first dipped in 18% SW, and then left to dry for approximately 20 minutes. Gradient plates were constructed by first pouring a layer of Hv-Ca+Trp to form a tapered wedge, followed by covering with Hv+Ca to form a flat surface and a tryptophan gradient across the plate (SZYBALSKI and BRYSON, 1952).After a subsequent incubation to allow growth and expression, plates were subjected to varying doses of UV, including a no UV control (Figure 4.5).

A

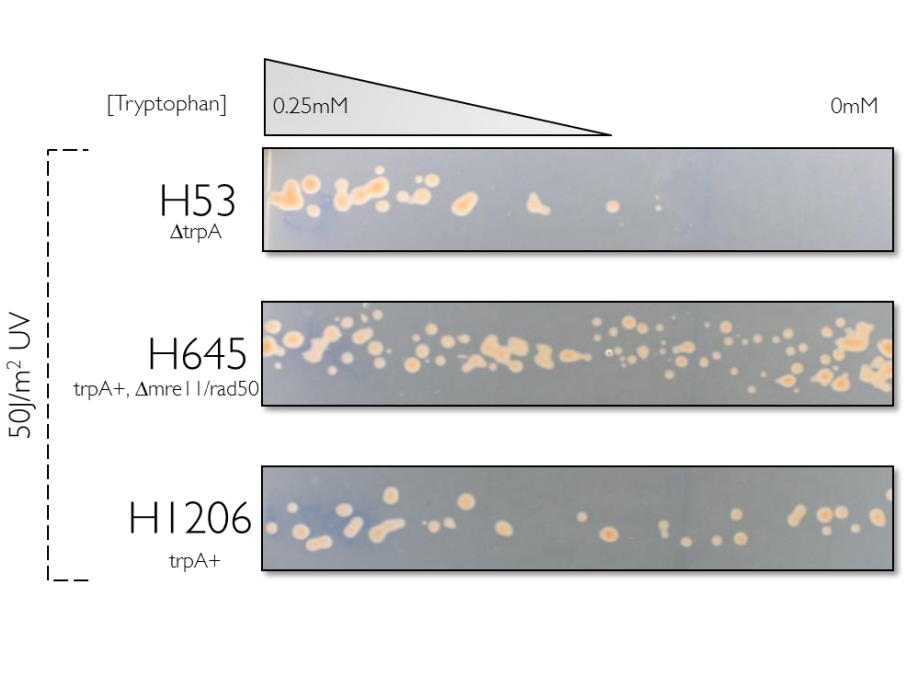

[Tryptophan]
0.25mM
0mM
H53
ΔtrpA
H645
trpA+, Δmre11/rad50
H1206
trpA+
0J/m2 UV
[Tryptophan]
0.25mM
0mM
21
11
7
1
0J/m2 UV

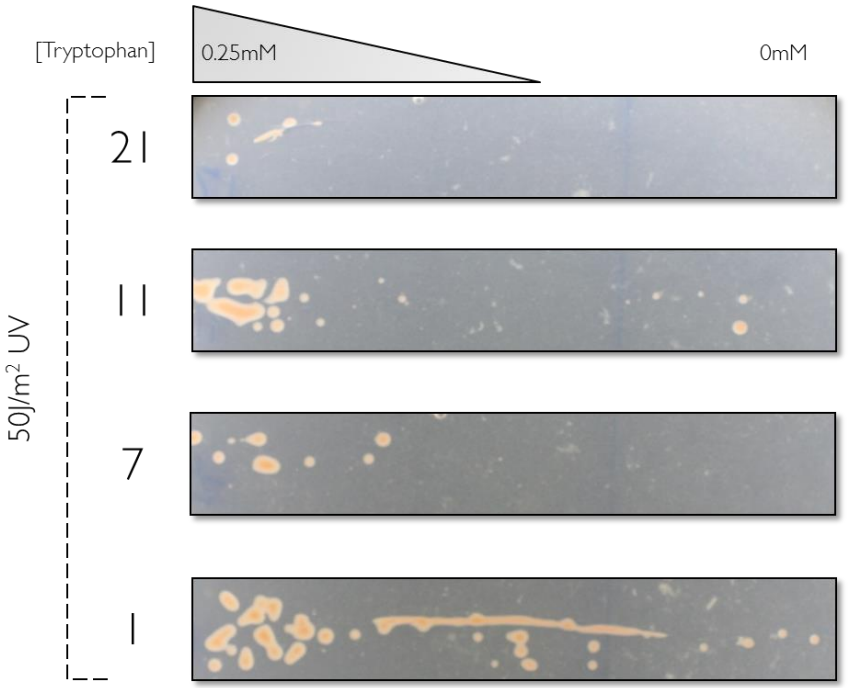
B


[Tryptophan]
0.25mM
0mM
25J/m² UV
H53
ΔtrpA
H645
trpA+, Δmre11/rad50
H1206
trpA+


[Tryptophan]
0.25mM
0mM
25J/m² UV
21
11
7
1

C


[Tryptophan]
0.25mM
0mM
50J/m2 UV
H53
ΔtrpA
H645
trpA+, Δmre11/rad50
H1206
trpA+


[Tryptophan]
0.25mM
0mM
50J/m2 UV
21
11
7
1

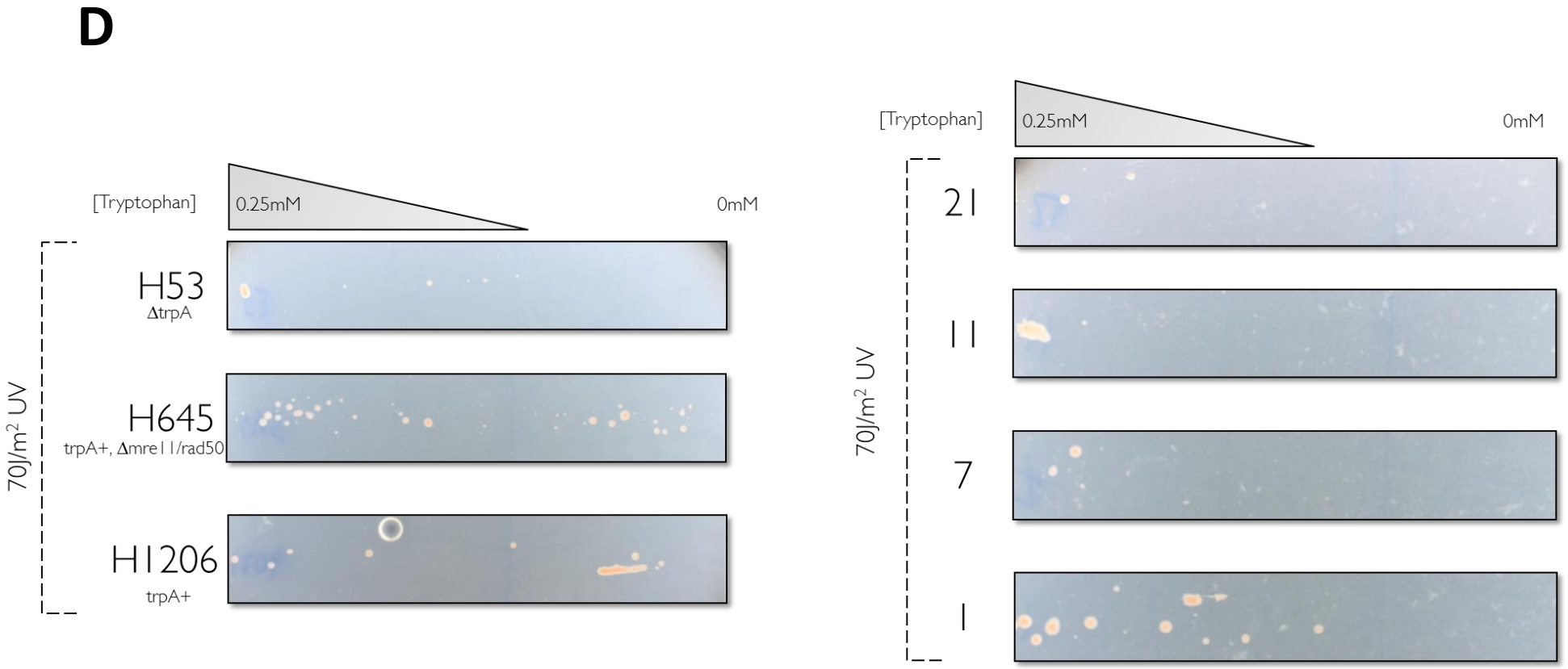


Figure 4.5 | **Genome library fragments and controls on tryptophan gradient plates.** gLib1, gLib7, gLib11 and gLib21 in a H1206 background were painted onto tryptophan gradient plates with a tryptophan concentration of 0.25mM. H1206 (+trpA*)* and H53 (ΔtrpA*)* acted as controls for visualisation of the gradient. UV doses used were 0J/m^2 (A), 25J/m^2 (B), 50J/m^2 (C) and 70J/m^2 (D).

4.4.3 UV Resistance Assay in H26 Background

Genomic library fragments, under the p.*tnaA* promoter, were prepared from their current host, H1206, via standard miniprep extraction using chloroform (see methods). After plasmid purifications had been carried out, they were then transformed into an alternative background, strain H26. Genotypically these strains are very similar. Therefore, a subsequent UV resistance assay was carried out to validate the phenotype picked up from the initial screening results in H1206. This demonstrates that the UV resistance phenotype seen is not due to a mutation, or other intrinsic property, of the host strain, but instead is as a result of the episomes (Figure 4.6). Colony 1 was discounted from further anaylsis.

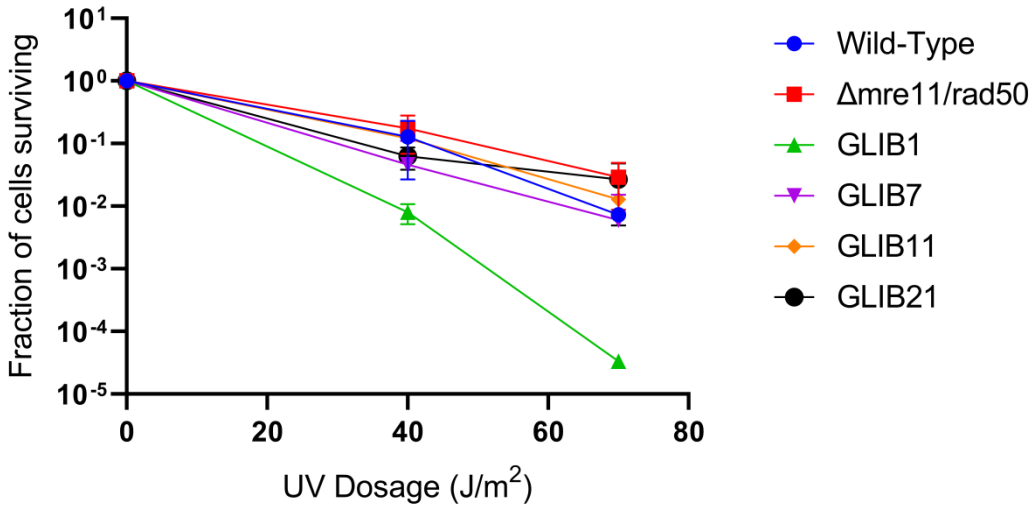


Figure 4.6| **Genome library UV survival assay in H26 background**. H26 (ΔpyrE2) was transformed with the genome library episomes (gLib1, 7, 11 and 21). Survival assays were then carried out on the transformed strains to confirm UV resistance. The spotting procedure was carried out as mentioned in the methods section. H1206 (ΔpyrE2, Δmrr) and H645 (ΔpyrE2, bgaHa, Δmre11-rad50) were taken along, as normal, to represent control colonies.

4.4.4 UV Growth Assay of library constructs

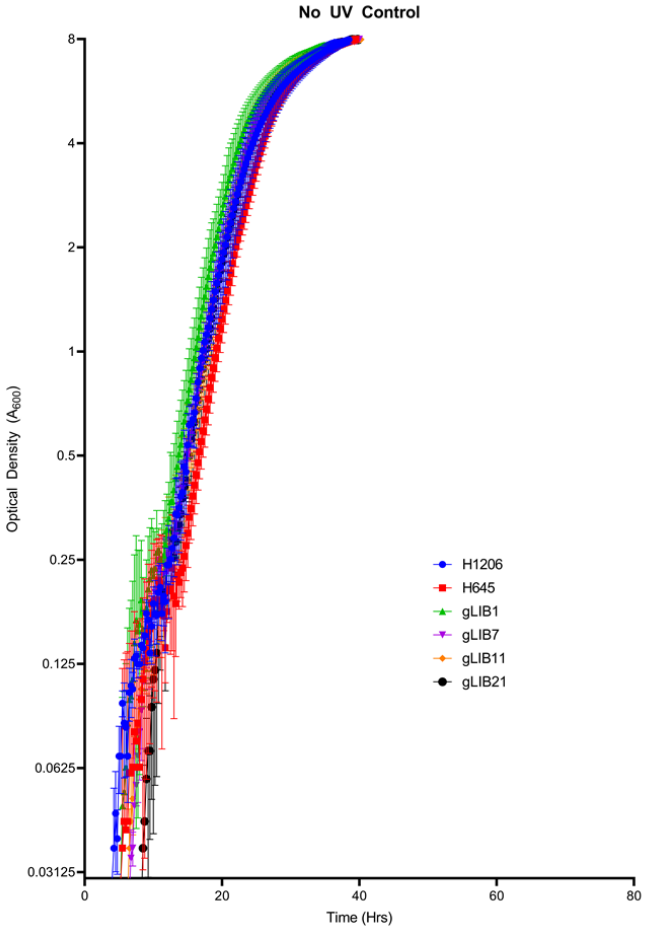
Genomic library constructs were subjected to a growth assay to measure the difference in growth rate as a result of overexpression of fragments downstream of the p.*tnaA* promoter, should there be one. The standard growth rate of cultures can be determined using an Epoch 2 Microplate Spectrophotometer (BioTek). Cultures were prepared in 5ml Hv-YPC or Hv-Ca broth (with required additives) and grown to mid-late exponential phase, corresponding to an A_{650} of between 0.2 and 0.8. Cultures were diluted and grown again to mid-late exponential phase. Serial dilutions of cultures were then made before loading 150µl of culture into a 96-well plate (Corning), in triplicate, alongside appropriate blanks. Genome library colonies were irradiated with 300J/m^2 before non-irradiated samples and blanks were loaded. Blank samples were loaded in the wells around the entire perimeter of the 96-well plate to mitigate any evaporation of sample and formation of salt crystals. The plate was sealed around the edge using two layers of microporous tape (Boots) and incubated at 45°C with double orbital shaking at 1000rpm for 72 hours in the Epoch plate reader. Readings at A_{600} were taken every 15 minutes and converted to a 1cm pathlength by dividing the raw A600 value by 0.14. Blank readings were taken into account by subtracting the average blank reading from all raw A_{600} values prior to dividing by 0.14. If the generation time was required, it is calculated by plotting the growth on a $\log_2$ scale and using the following equations:

$$G = \frac{t}{n} \qquad n = \frac{logx - logy}{log2}$$

G = generation time
t = time
n = number of generations
x = end A650
y = start A650

The results can be seen in figure 4.7.

A


No UV Control
Optical Density (A600)
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
0
20
40
60
80
Time (Hrs)
H1206
H645
gLIB1
gLIB7
gLIB11
gLIB21

B

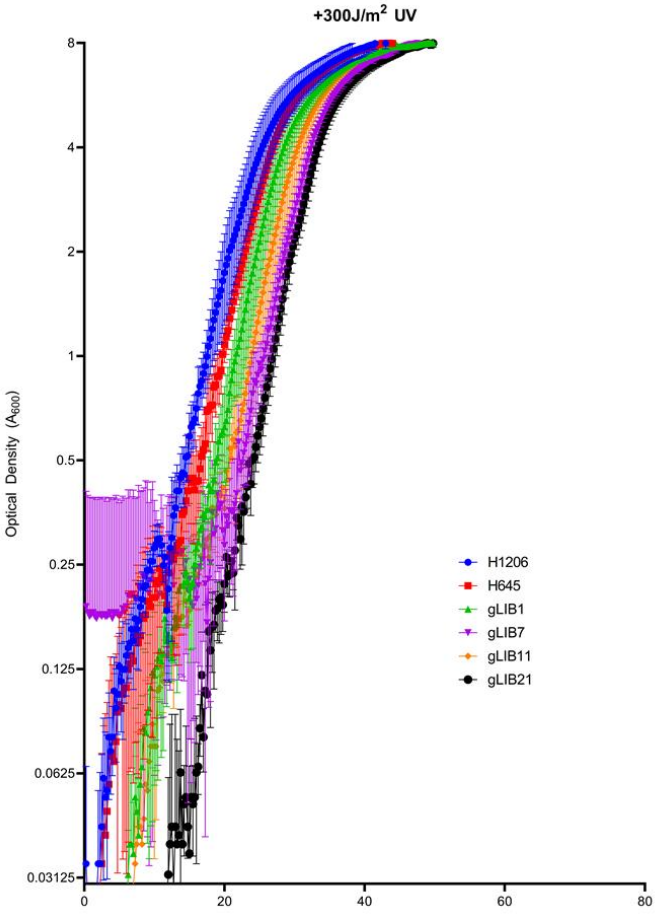

+300J/m2 UV
Optical Density (A600)
8
4
2
1
0.5
0.25
0.125
0.0625
0.03125
0
20
40
60
80
H1206
H645
gLIB1
gLIB7
gLIB11
gLIB21

Figure 4.7 | **Optical Density Growth Assay of Genomic Library Constructs.** Dilutions of genomic library construct (gLib1, 7, 11 and 21) in a H1206 background were diluted and loaded into 96-well plates in triplicate. Cultures were irradiated, prior to loading of unirradiated controls and blanks before the OD was measured over 72hrs.

4.4.5 Generation of HVO_2259 deletion constructs

Within the most resistant gLib fragment, which was also obtained twice from the screen (colonies 7 and 21) was the gene encoding a XerCD-like integrase. The XerCD-like integrase coding sequence (HVO_2259) was therefore chosen for downstream analysis due to its role in resolving chromosomal dimers. In order to generate a deletion of the XerC/D-like integrase (HVO_2259) for further analysis, genomic DNA from H26 was prepared via spooling and left to resuspend for five days in the fridge, with occasional mixing to ensure a homogeneous suspension. Upstream and downstream homology arms were amplified (0.5kb each) either side of the HVO_2259 genomic locus using primers incorporating various restriction sistes to facilitate downstream cloning. The upstream homology arm was constructed using primers XerDel_F_Cla1 and XerDel_R_BamH1, to insert a 5' Cla1 site and a 3' BamH1 site. The downstream homology arm was amplified using XerDel_F_BamH1 and XerDel_R_Not1 to add a 5' BamH1 site, compatible with ligation to the upstream homology arm, and a 3' Not1 site. After PCR amplification, bands were extracted using standard procedure (Macherey-Nagel) and ligated to pTA131, which was also cut with Cla1 and Not1. An overnight ligation took place at 15°C, allowing for ligation of the homology arms to each other, via the BamH1 sites, and then subsequent ligation of Cla1 and Not1 sites to the pTA131 deletion vector, generating pTA2783. Five colonies were miniprepped and a diagnostic digest, designed to liberate the ligated homology arms (~1kb), was carried out using ClaI and NotI (Figure 4.8). Colony 2 was selected and sequenced to confirm correct cloning.

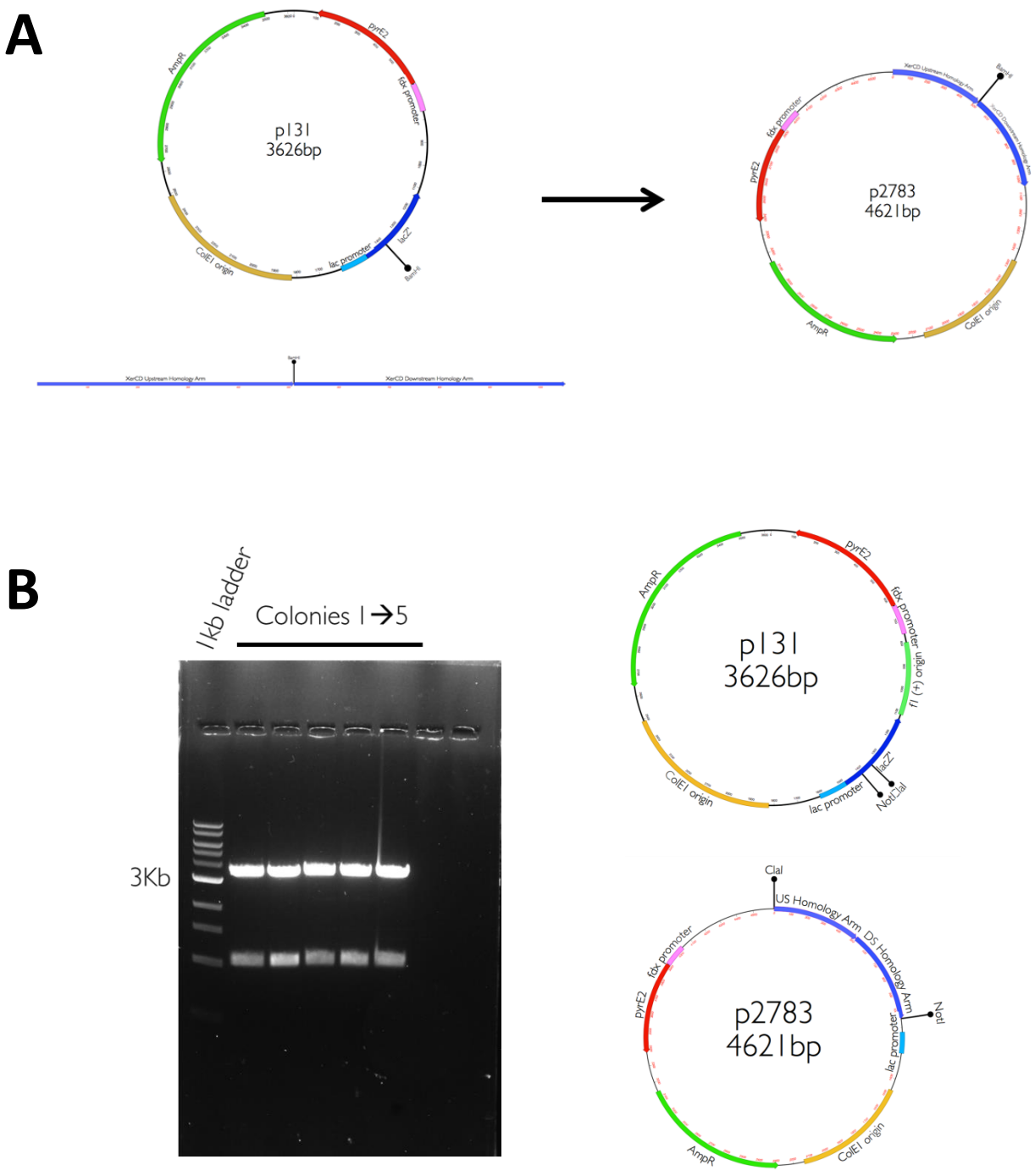
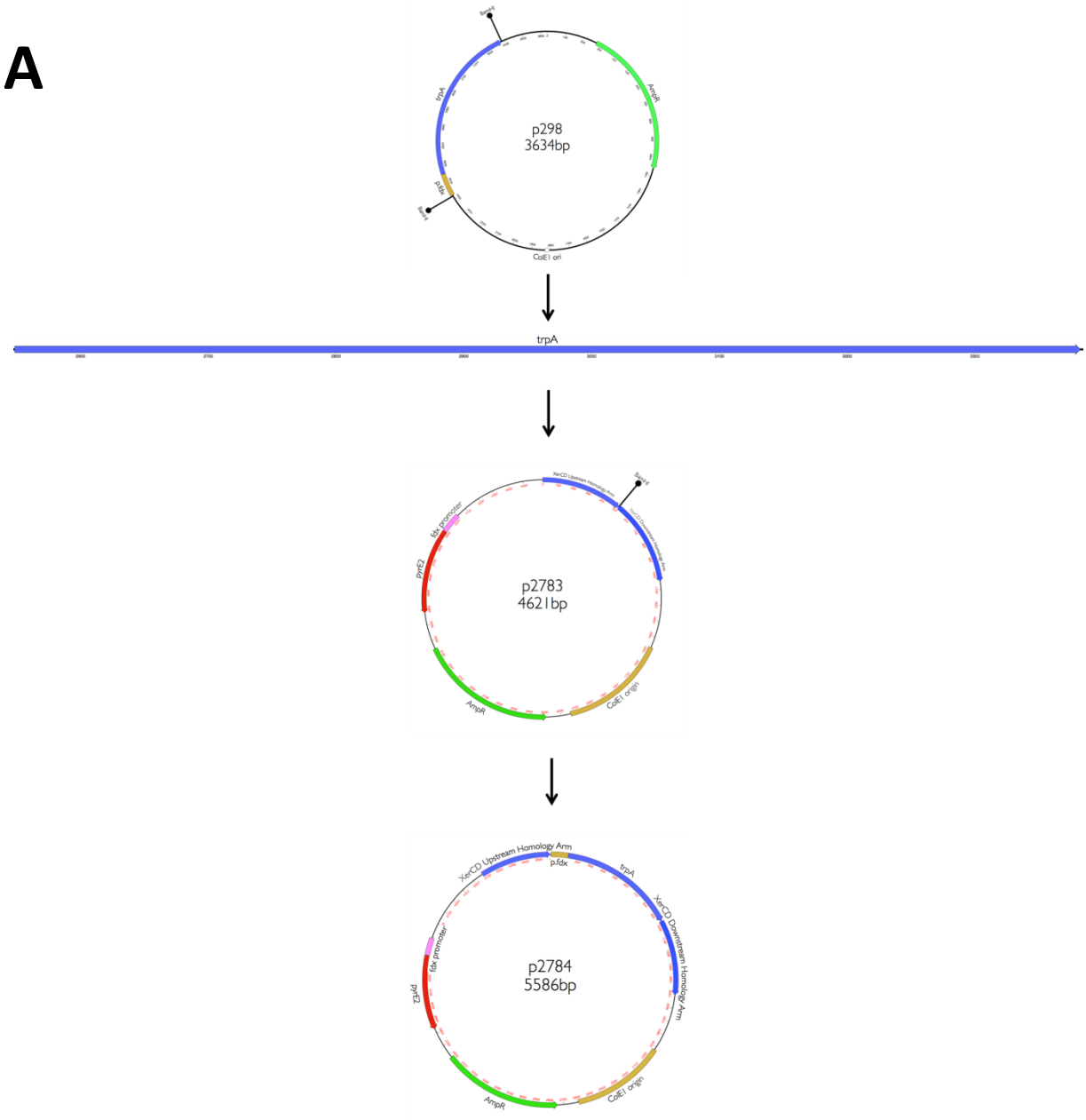


Figure 4.8 | **Cloning strategy and diagnostic digest of pTA131 HVO_2259 deletion cloning.** A) Cloning strategy for homology arm cloning into pTA131 to generate deletion construct pTA2783. B) Diagnostic digest of five selected minipreps confirming correct cloning. Expected banding pattern of ~1kb and 3.6kb.

To make a TrpA+ marked deletion construct, pTA2783 was digested with BamHI, linearizing the plasmid in the middle of the two homology arms. The vector was then 5' phosphatase treated with 1µl rSAP (NEB) for 30 minutes at 37°C prior to gel extraction. The TrpA insert including its p.fdx promoter was liberated from pTA298 using BamHI and subsequently ligated into the digested pTA2783 vector in a 3:1 molar ratio, generating plasmid pTA2784. Colonies were selected for DNA miniprep purification and a diagnostic digest was carried out on six colonies to confirm correct insert orientation and cloning (Figure 4.9).



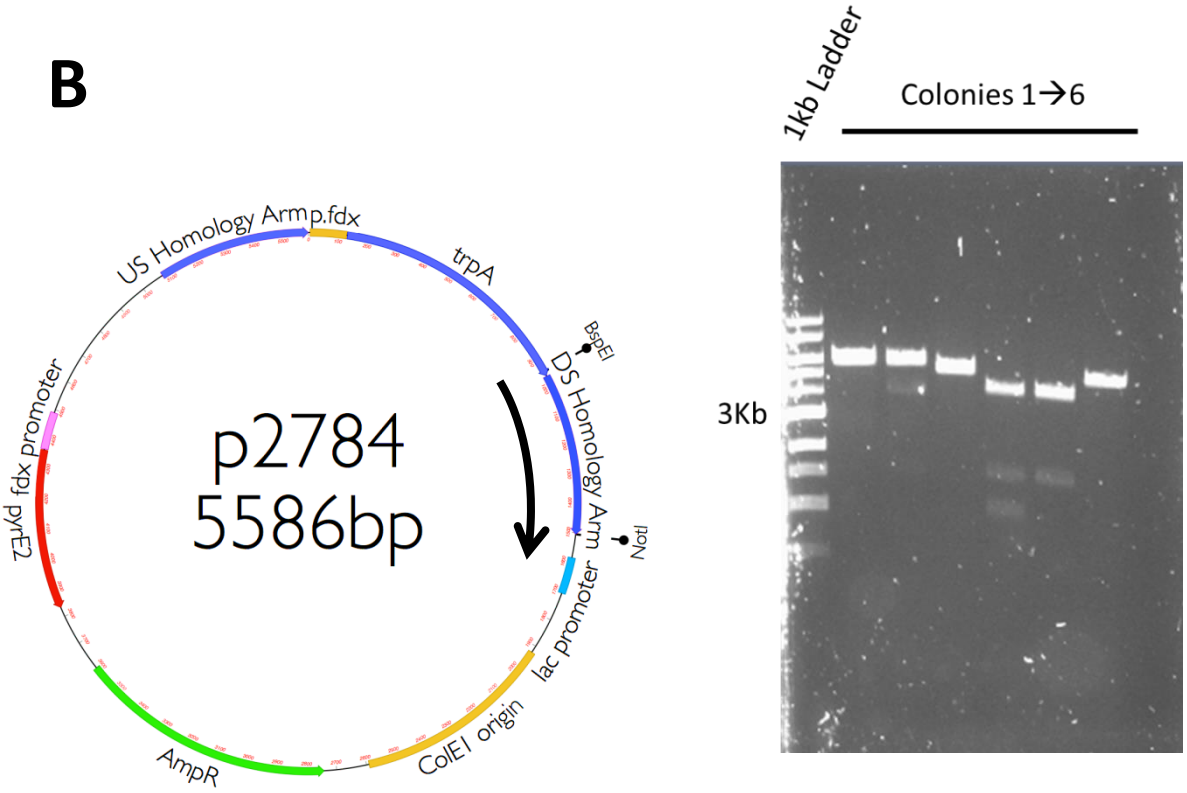


Figure 4.9 | **Cloning stragety to create TrpA+ marked deletion construct.** A) Cloning strategy to ligate TrpA marker into pTA2783. B) Miniprep diagnostic digests to confirm correct orientation of insert and cloning using BspEI and NotI.

4.4.6 Generation of HVO_2259 deletion strain

In order to generate a deletion strain for XerC/D-like integrase, HVO_2259, H26 (Δ*pyrE2*) was first transformed with construct pTA2786. Transformations were plated on Hv-Ca plates to select for uracil prototrophy and generation of a "pop in" strain via homologous recombination using the upstream homology arm, generating strain H5562 (Δ*pyrE2 ΔtrpA xerC/D+::[ΔxerC/D pyrE2+]*). Colonies were selected based on their ability to grow on media lacking uracil and were then grown

overnight in Hv-YPC at 45°C overnight. Cultures were then diluted in fresh Hv-YPC broth for a second overnight incubation. This process was repeated for a total of three successive overnight incubations, allowing the 'pop out' step to occur via intraspecific homologous recombination, which can happen either using the upstream or downstream homology arms. Cultures were then plated using appropriate dilutions (10^0 – 10^{-3}) on Hv-Ca+5FOA plates and left to grow for ~5 days, thus generating the pop out strain H5610 (*ΔpyrE2 ΔxerC/D*) to be confirmed by colony hybridization and Southern blot.

This process was repeated to make a TrpA+ marked deletion strain, to increase chances of successful deletion generation and allowing selection via one step (see chapter 1 for details). For this to work, strain H53 (*ΔpyrE2 ΔtrpA*) was used as the parent strain to allow for trpA mediated selection. H53 was transformed with pTA2785 and subsequent plating was carried out in the same way, selecting for uracil and tryptophan prototrophy on Hv-Ca media, generating a Trp marked pop in strain H5563 (*ΔpyrE2 ΔtrpA xerC/D+::[ΔxerC/D pyrE2+ trpA+]*). Overnight dilutions were made to generate the pop out (N.B. this was not given a strain number as the non-Trp marked deletion was completed successfully) in Hv-YPC broth as before.

Colony Hybridization

A probe for XerC/D was generated from a PCR of XerC/D coding sequence in pTA2773 using primers:

PhageProbeFII

PhageProbeBII

Pop outs were patched in an ordered array along with H26 (*ΔpyrE2 xerC/D+*) and H1192 (*ΔpyrE2 Δphage*) as positive and negative controls, respectively for the presence of XerC/D, which occurs in the prophage region of the main chromosome. The XerC/D probe was radiolabelled and then hybridized to the array, confirming a number of deletions for the non-trp marked construct, and all colonies deleted from the trp marked deletion, outlining the increased efficiency of the latter (figure 4.10).

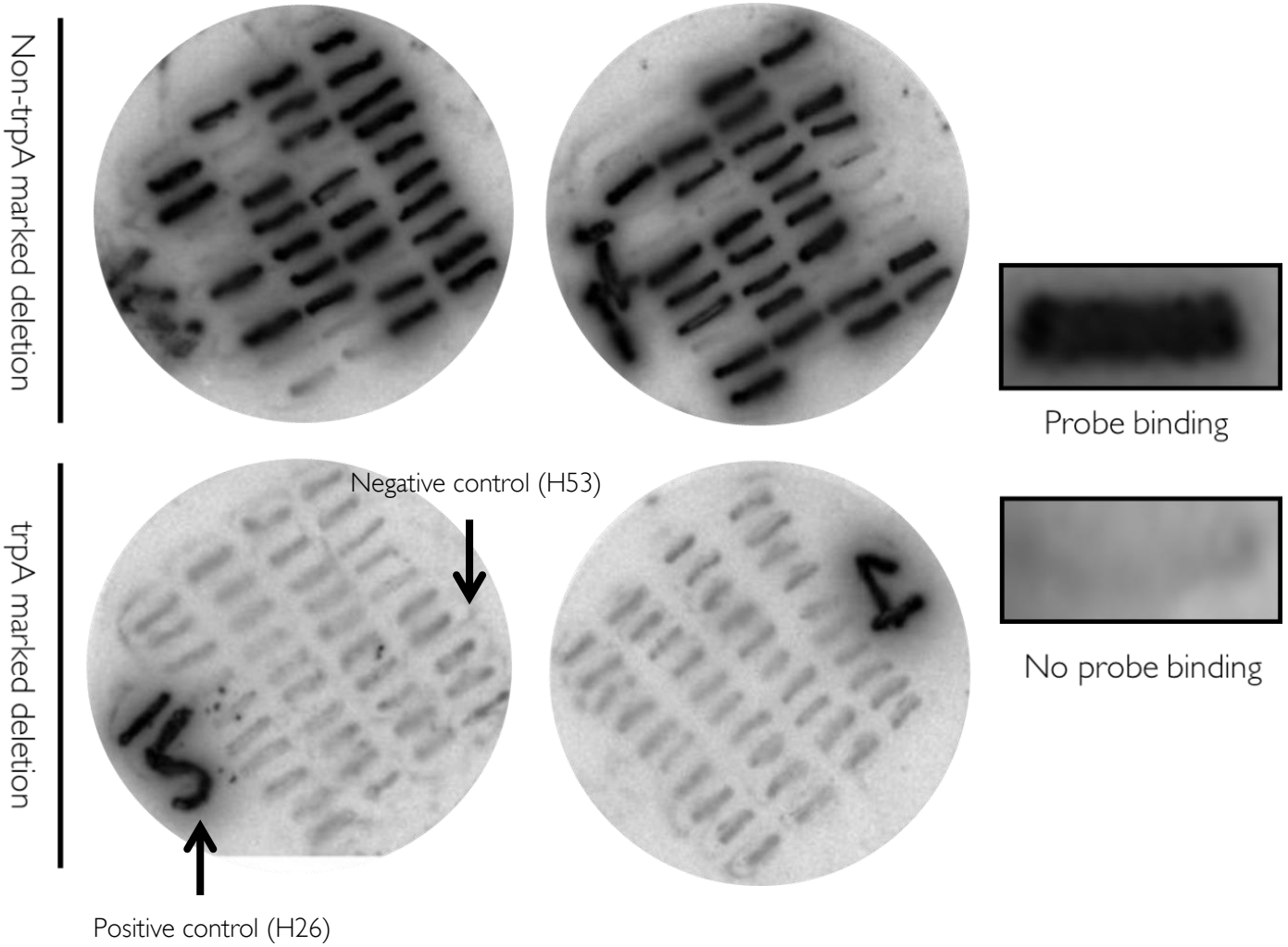


Figure 4.10 | **Colony hybridization of XerC/D probe to array of 'pop out' strains.** Negative controls were streaks of H53. Positive control was strain H26 and was used to number each plate.

Southern blot

Due to the potential for any given strain to still be merodiploid at this point, meaning that not all genomic copies of the gene are deleted, a diagnostic digest, in conjunction with a Southern blot must be employed at this stage to confirm that the deletion is genuine. First, genomic DNA was prepared from four potential 'pop outs' via spooling.

A suitable probe was generated from construct pTA2783 via a restriction endonuclease digest with KpnI and NheI. The probe generated from the backbone is 804bp in length, suitable to cover either side of the deletion site (figure 4.11).

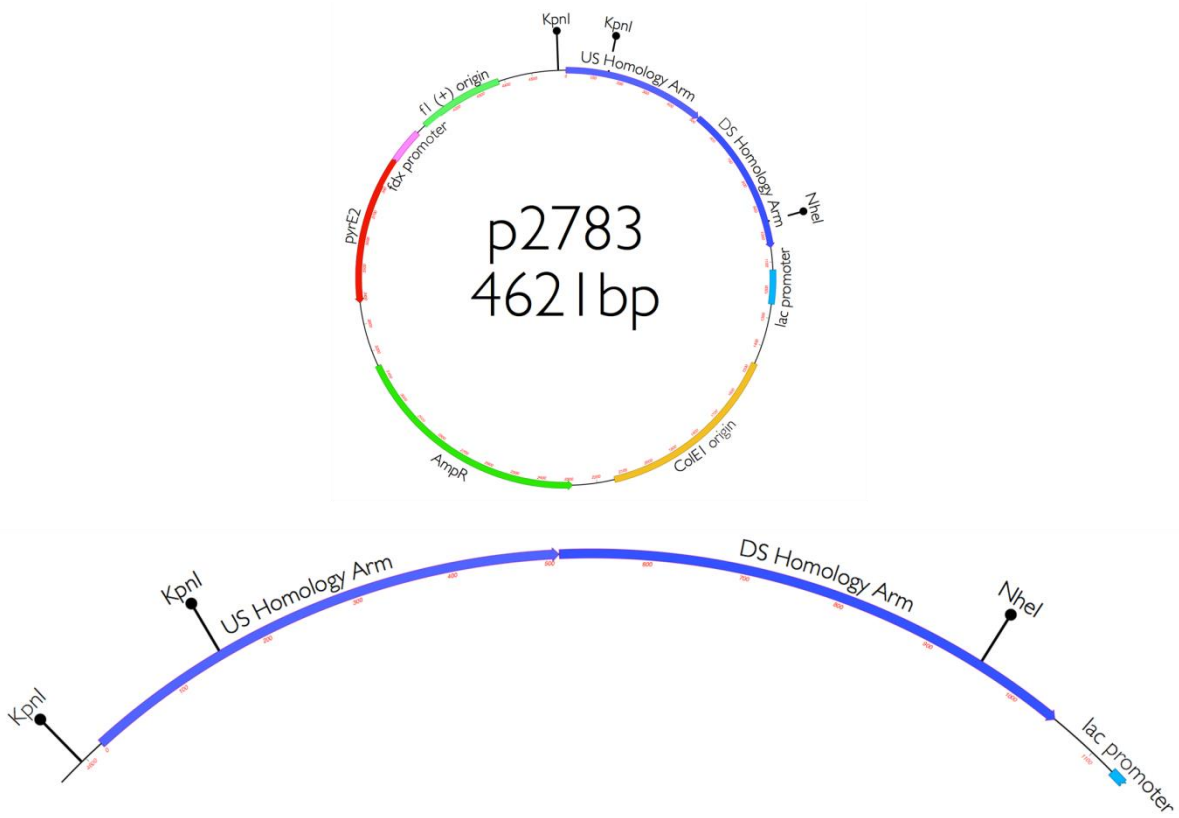
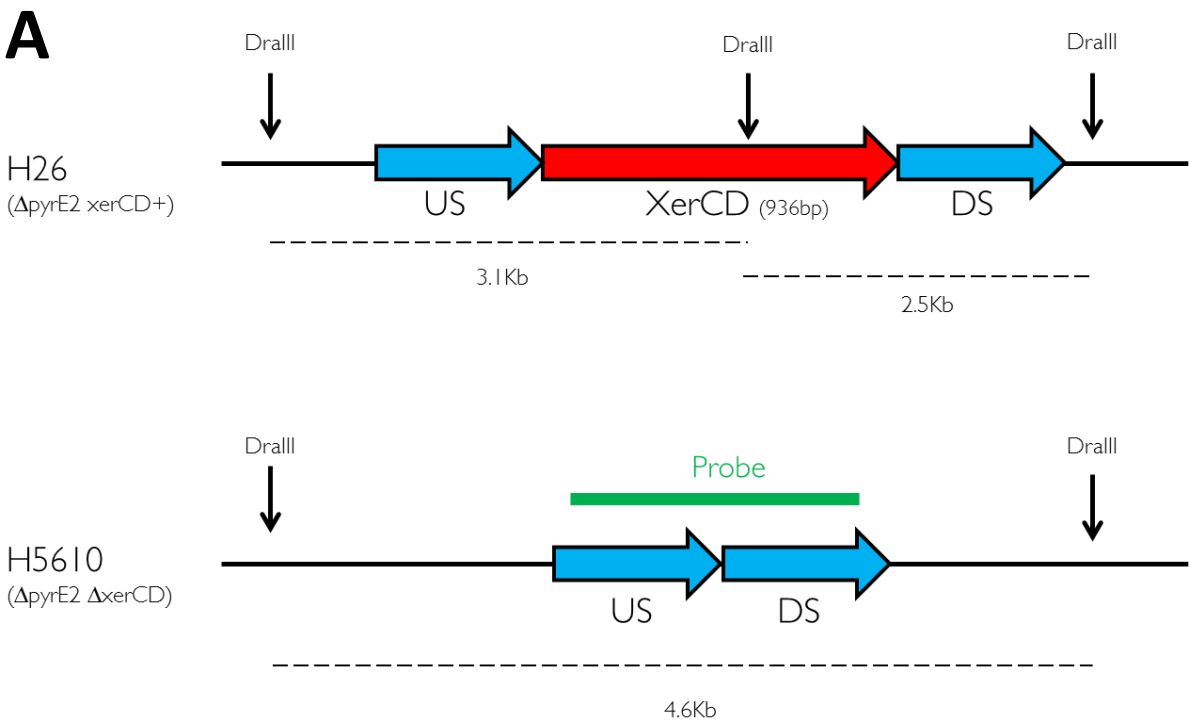


Figure 4.11 | **Restriction digest of pTA2783 to generate probe for XerCD deletion.** Digestion of pTA2783 with KpnI and NheI generates a deletion probe for XerCD.

A Southern blot was then carried out (see chapter 2 for details) after radiolabelling the probe (figure 4.12) and digesting genomic DNA preps from colonies 1→4, as well as H26 (+ve control) with DraIII.



B

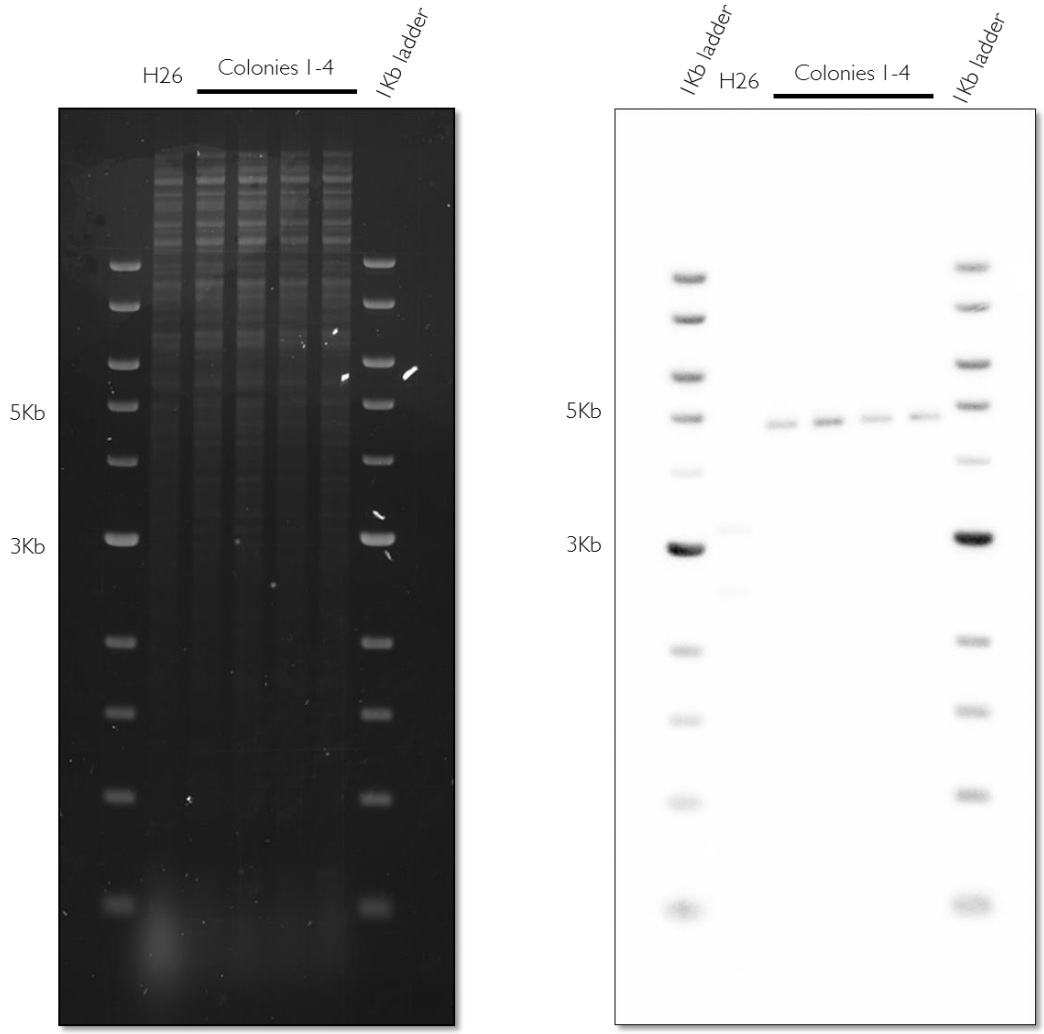


Figure 4.12 | **Southern blot of XerCD deletion strain.** A) Southern blot strategy with DraIII. Expected fragment sizes for the respective outcomes are shown as dashed lines. B) Agarose gel post-stained with EtBr depicting the four gDNA preps having been digested with DraIII. H26 is included as a positive control. Southern blot also shown after probe hybridization and visualisation demonstrating successful XerCD deletion in all colonies selected.

Colony 2 was selected for subsequent allocation to strain number H5610 (*ΔpyrE2 ΔXerCD*), which now can act as a strain deleted for XerCD for downstream assays.

4.4.7 Phenotypic Analysis of XerCD deletion strain

Placing XerCD under native promoter with low copy number origin of replication

Construct pTA2863 was generated to place the XerCD coding sequence under the control of its native promoter in a low copy number origin background. This facilitates its use for complementation of the deletion strain H5610 in downstream DNA damaging assays. In order to generate this construct, pTA354 was used as the backbone, due to the presence of the pHV1 low copy origin. XerCD was isolated from genomic DNA of the wild-type strain H26 via spooling and was subsequently subjected to PCR using primers:

6264R2

PhageProbeR3

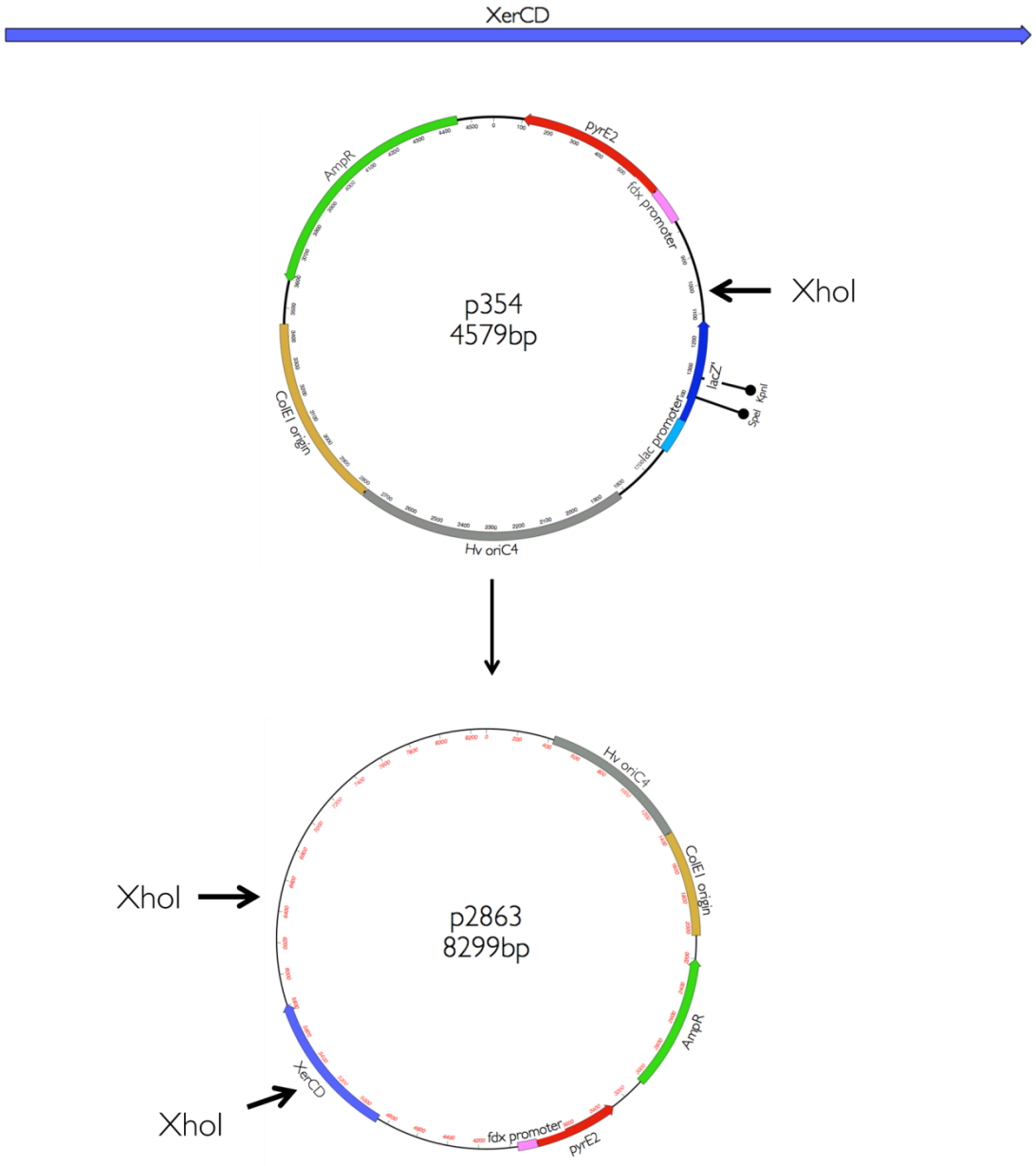
The XerCD coding sequence, with its native promoter was then digested and ligated into the pTA354 backbone using KpnI and SpeI restriction endonucleases (figure 4.13).

Cloning confirmation was obtained via a diagnostic digest with XhoI. In total, six colonies were chosen. Of the six colonies, only colony 1 was correct, which was subsequently sequenced using primers:

pBSF2

PhageProbeBII

The construct was verified for downstream use as pTA2863.


XerCD
p354
4579bp
AmpR
PyrE2
fdx promoter
XhoI
lacZ'
XhoI
SpeI
lac promoter
ColE1 origin
Hv oriC4
p2863
8299bp
XhoI
XhoI
XerCD
fdx promoter
PyrE2
AmpR
ColE1 origin
Hv oriC4

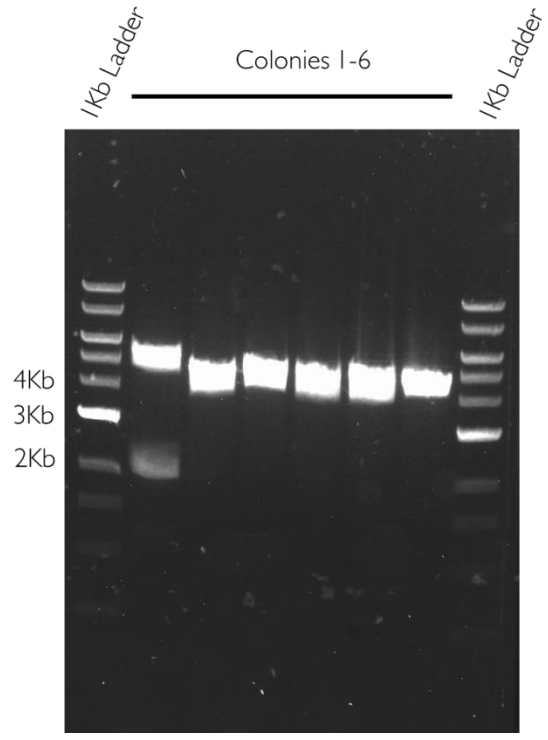

1Kb Ladder
Colonies 1-6
1Kb Ladder
4Kb
3Kb
2Kb

Figure 4.13 | **Cloning strategy for pTA2863 construction.** A) Cloning stragegy. XerCD is PCR amplified from the genome and ligated into pTA354 using SpeI and KpnI. B) Diagnostic digest confirming colony 1 as the correct construct, which was later verified by sequencing. Correct constructs produce two fragments (2.4Kb and 5.9Kb) whilst incorrect fragment insertion yields a single fragment (4.5Kb).

DNA damaging assay with UV radiation

In order to test survival fractions after exposure to UV light, firstly the following constructs were transformed into wild-type *H. volcanii*, H26 (Δ*pyrE2*) and H1206 (Δ*pyrE2* Δ*mrr*). Strain H1206 was used to facilitate transformations of dam+ DNA. The following strains were used:

[ ] Indicate integrated plasmid DNA

- { } Indicate episomal plasmid DNAH26 *(ΔpyrE2)*
- H5610 *(ΔpyrE2 ΔxerCD)*
- H5625 *(ΔpyrE2 {pyrE2+ hdrB+})*
- H5626 *(ΔpyrE2 {pyrE2+})*
- H5628 *(ΔpyrE2 ΔxerCD {native::XerC/D pyrE2+})*
- H1192 *(ΔpyrE2 Δphage)*
- H5629 *(ΔpyrE2 {p.syn::XerC/D pyrE2+ hdrB+})*
- H5635 *(ΔpyrE2 Δphage {p.syn::XerC/D pyrE2+ hdrB+})*

Colonies were confirmed using colony PCR with the following primers:

For empty vector controls:

pBSF2

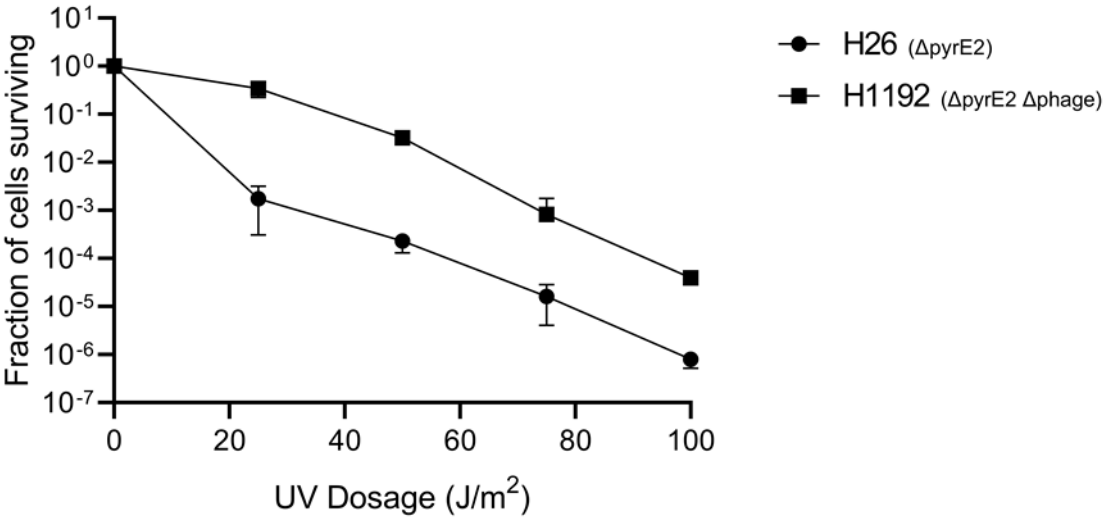
pBSR3

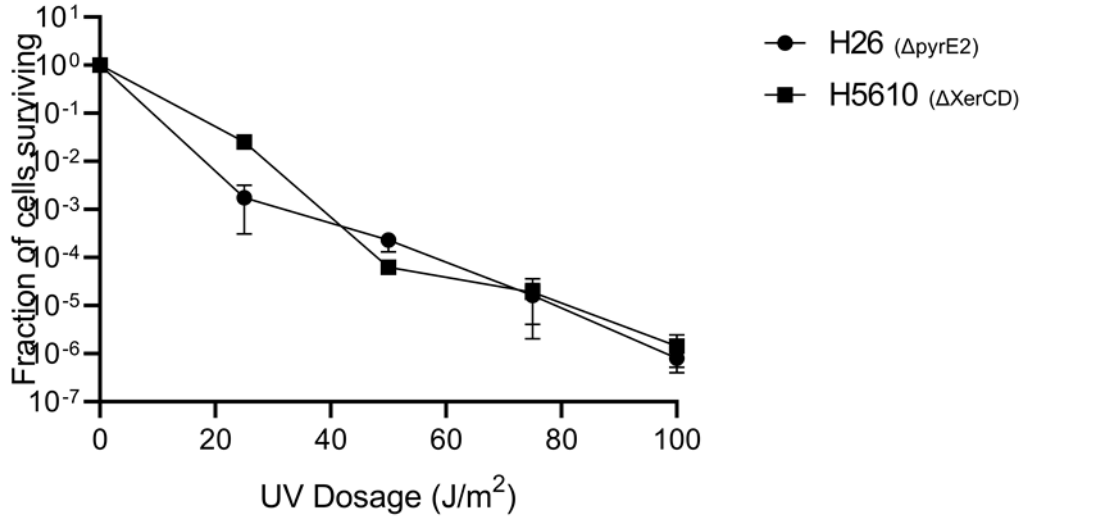
For other constructs:

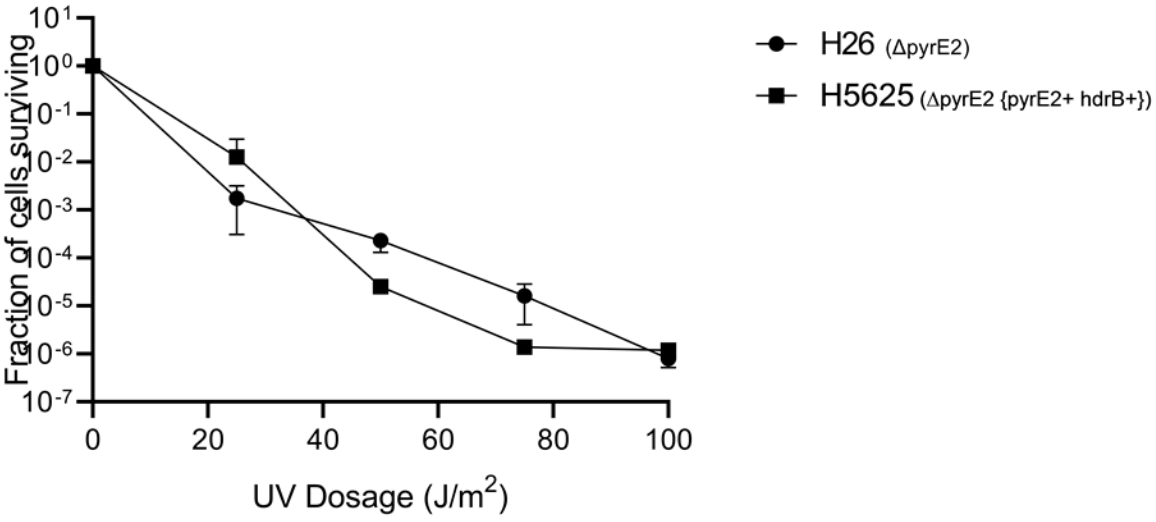
pBSF2

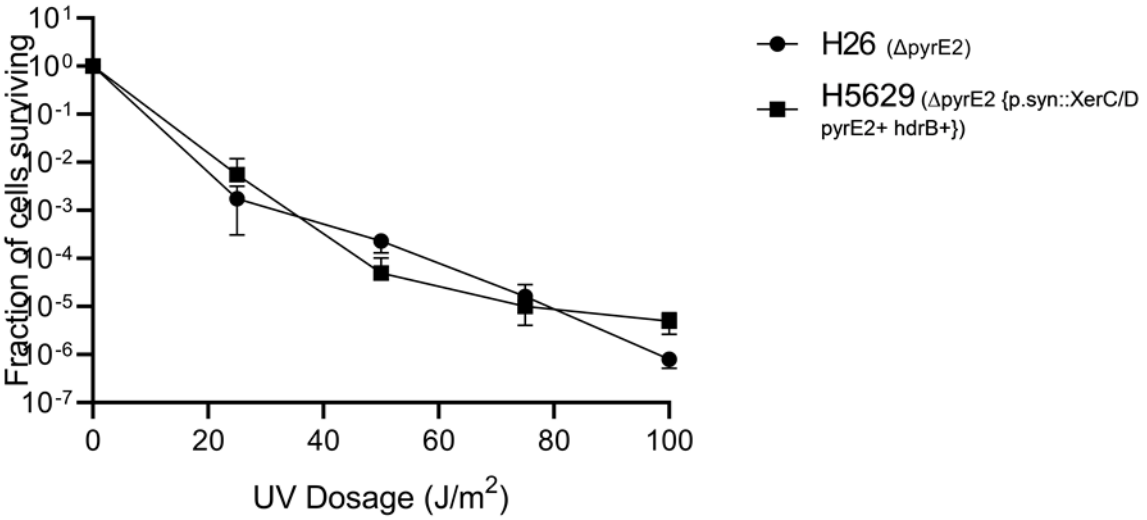
PhageProbeBII

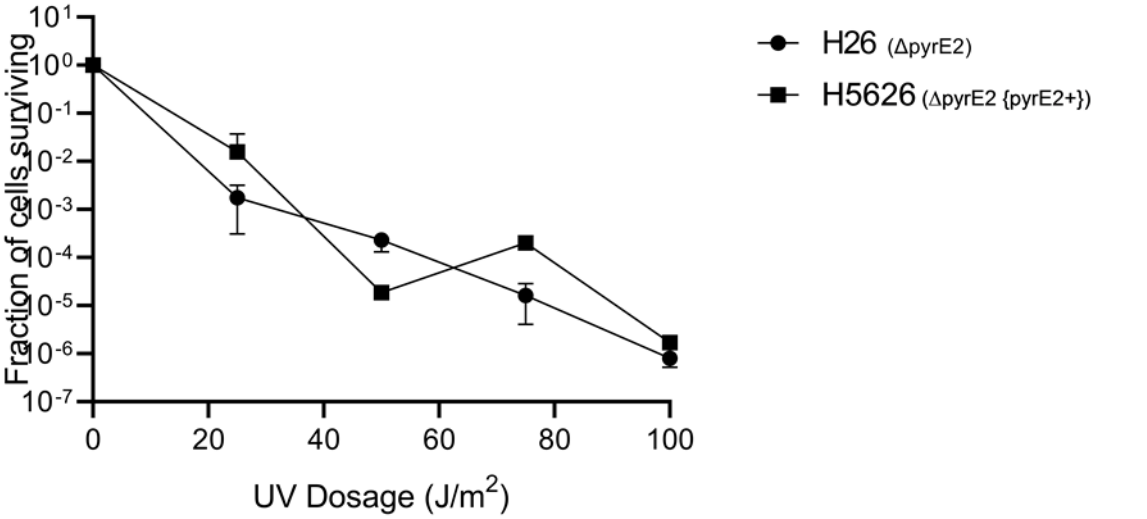
Once strains had been constructed and checked, each was grown overnight in either Hv-Ca or Hv-Ca+Ura. Strains were grown to an A_{650} of 0.4 and were then diluted in 18% salt water. Dilutions were spotted and plates subjected to UV irradiation of varying dosages. A non irradiated control for each strain was used to normalise survival fraction. After four days, spots were counted and survival fractions calculated. Results can be seen in figure 4.14.

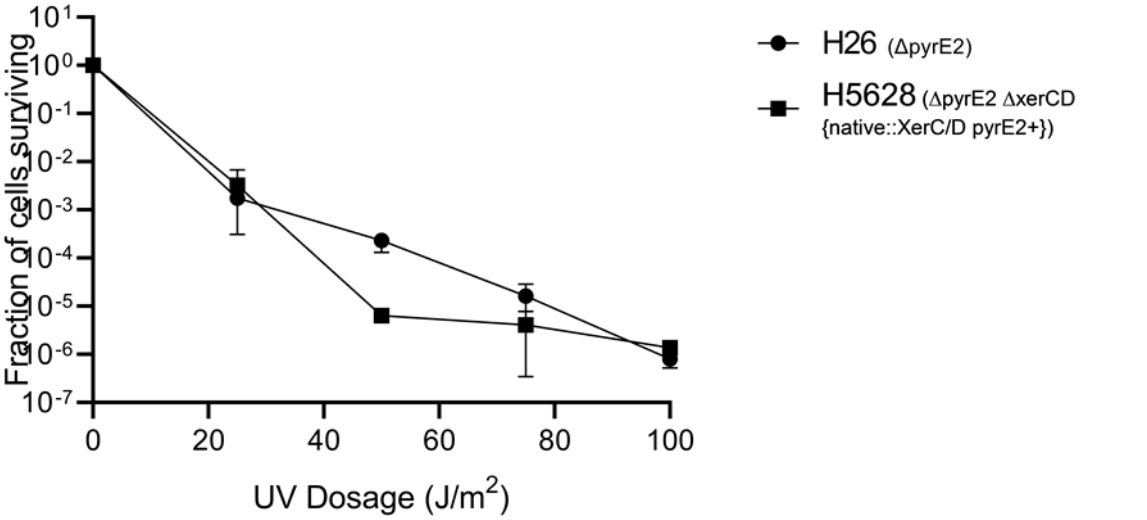

Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H1192 (ΔpyrE2 Δphage)


Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H5610 (ΔXerCD)


Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H5625 (ΔpyrE2 {pyrE2+ hdrB+})


Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H5629 (ΔpyrE2 {p.syn::XerC/D pyrE2+ hdrB+})


Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H5626 (ΔpyrE2 {pyrE2+})


Fraction of cells surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
0
20
40
60
80
100
UV Dosage (J/m^2)
H26 (ΔpyrE2)
H5628 (ΔpyrE2 ΔxerCD {native::XerC/D pyrE2+})

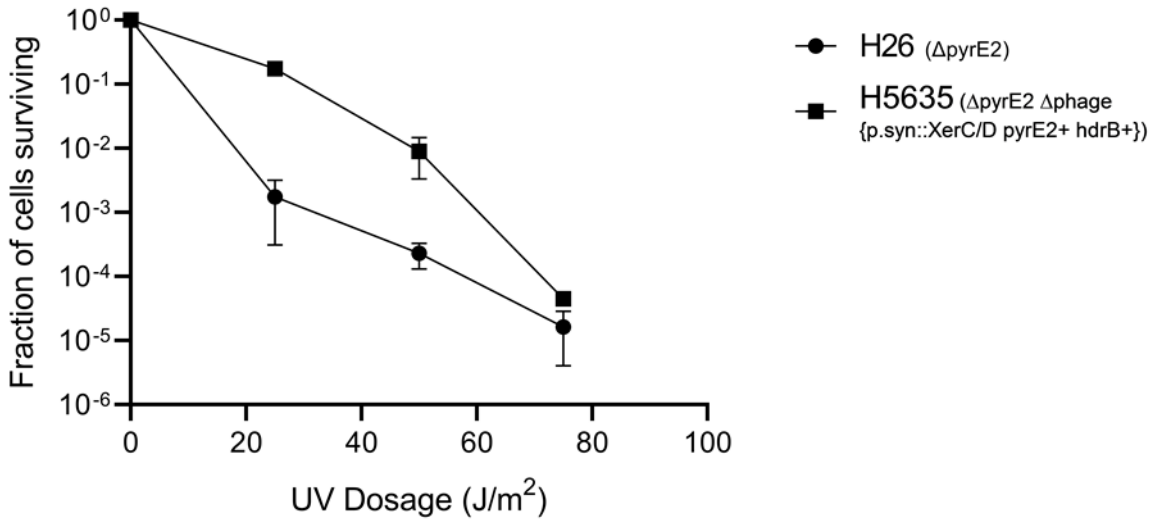
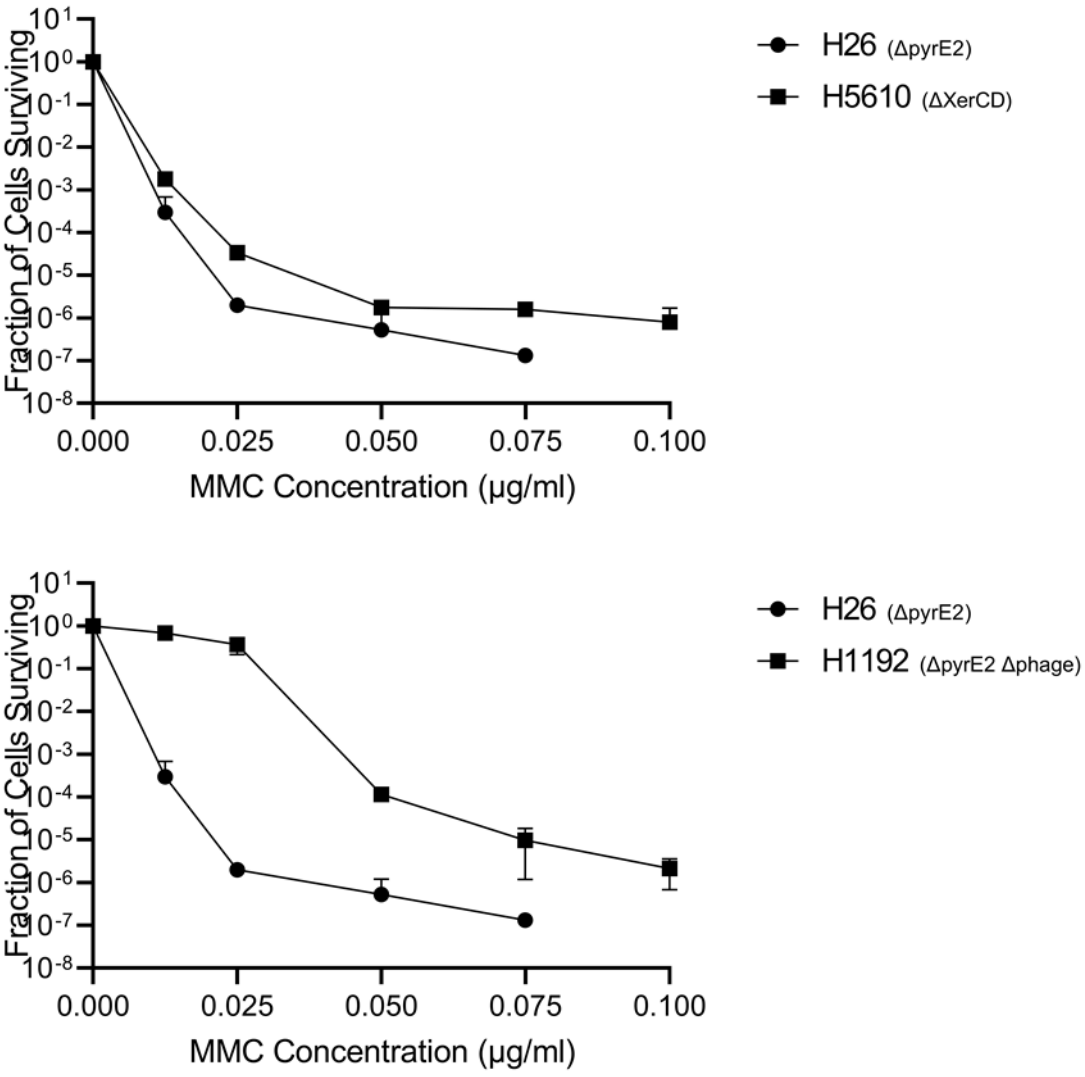
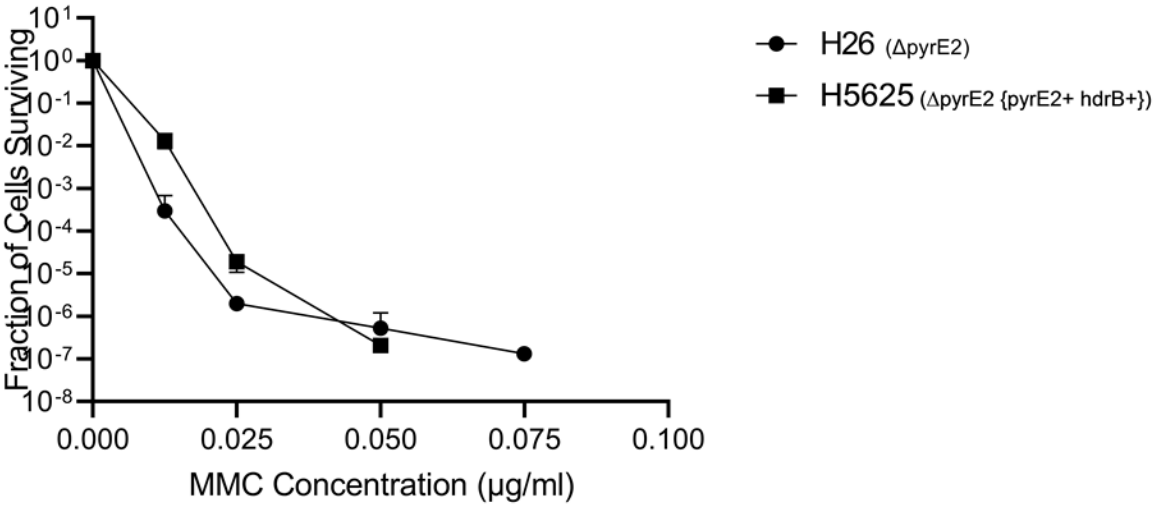


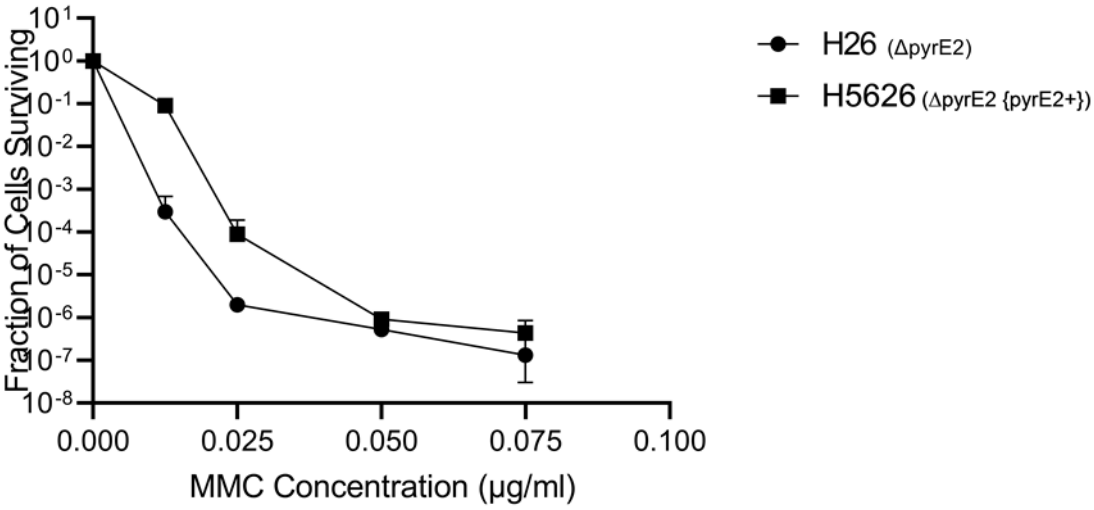
Figure 4.14 | **UV DNA damaging assay of XerCD deletion strain and associated controls.** Respective strains were grown overnight, diluted to reach exponential phase and then dilted and spotted before being subjected to varying dosages of UV radiation. Survival fractions were normalised to a non irradiated control and data plotted against the survival fraction for wild-type strain H26. No data point shows 100% cell death.(n=3).

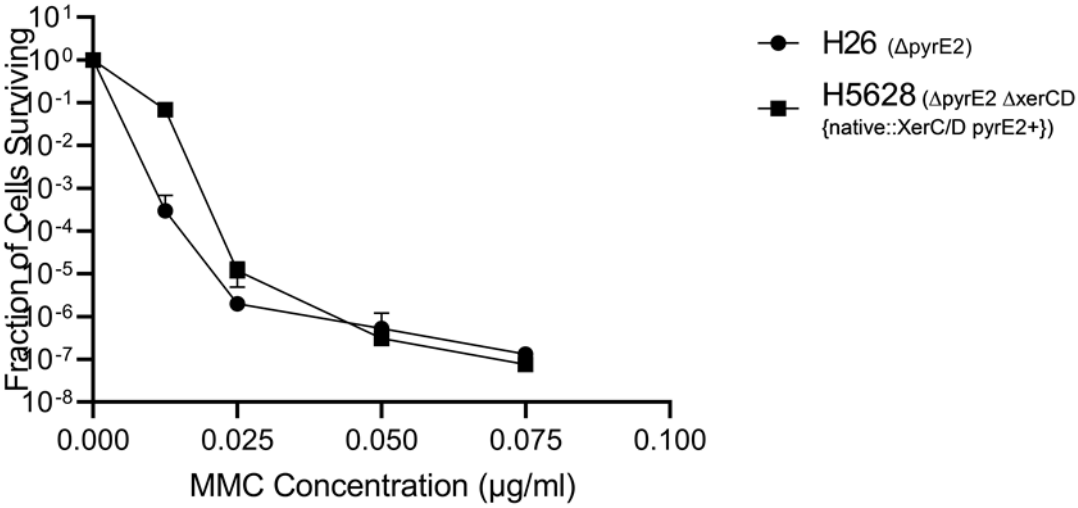
DNA damaging assay with MMC

The same strains were then subjected to treatement with MMC. The results can be seen in figure 4.15.




Fraction of Cells Surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
0.000
0.025
0.050
0.075
0.100
MMC Concentration (µg/ml)
H26 (ΔpyrE2)
H5625 (ΔpyrE2 {pyrE2+ hdrB+})


Fraction of Cells Surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
0.000
0.025
0.050
0.075
0.100
MMC Concentration (µg/ml)
H26 (ΔpyrE2)
H5626 (ΔpyrE2 {pyrE2+})


Fraction of Cells Surviving
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
0.000
0.025
0.050
0.075
0.100
MMC Concentration (µg/ml)
H26 (ΔpyrE2)
H5628 (ΔpyrE2 ΔxerCD
{native::XerC/D pyrE2+})

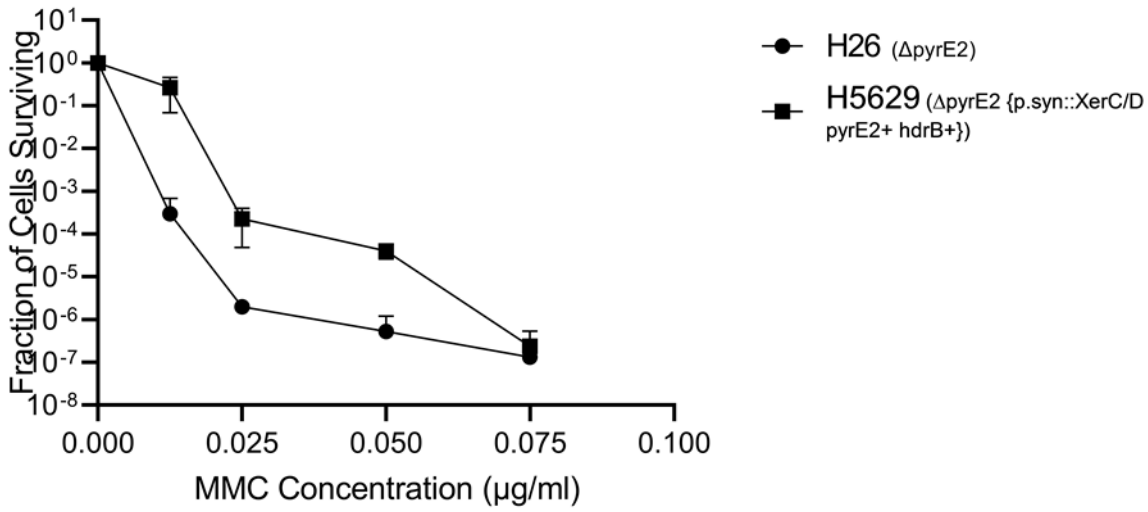


Figure 4.15 |**MMC DNA damaging assay of XerCD deletion strain and associated controls.** Respective strains were grown overnight, diluted to reach exponential phase and then diluted and spotted on plates containing relevant doses of MMC. Each experiment was carried out in biological triplicate (n=3). No data point shows 100% cell death.

4.4.8 Phenotypic Analysis of HVO_2259 Deletion using growth assay

In order to determine if and how XerCD influences growth rate, the same strains were subjected to 72 hours of growth with optical density monitoring every 15 minutes. Strains were grown for two consecutive overnights, ensuring that only actively dividing cells were used. Cells were diluted (see chapter 2 for details) in a 96 well plate in Hv-Ca or Hv-Ca+Ura broth. Optical density (OD, A650) was continuously measured, allowing the plotting of growth curves for each strain (figure 4.16).

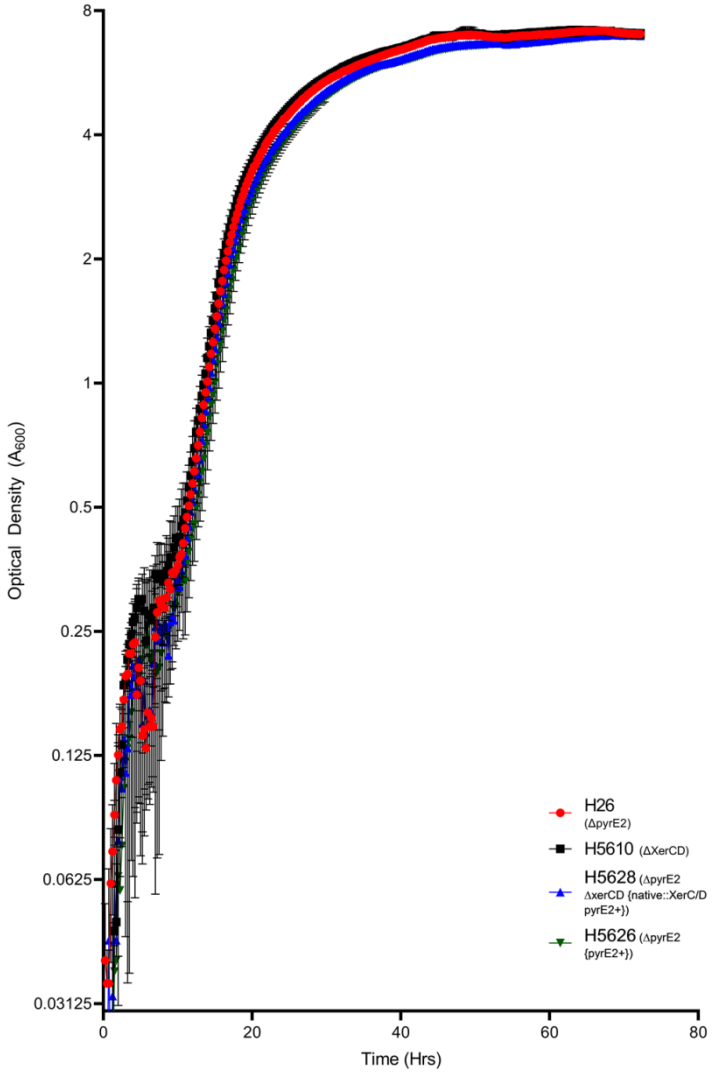


Figure 4.16 |**Growth assay of XerCD and respective deletion strains.** Wild type (H26) and XerCD deletion strainswere monitored for growth via measurement every 15 minutes of optical density (A_{650}). This was carried out over a period of 72hrs. An empty vector control was also used.

No perturbation in growth rate is seen when the coding sequence for HVO_2259 (XerC/D) is deleted (strain H5610). Therefore the complementation strain equally shows no appreciable difference in growth rate compared to wild-type strain H26.

4.4.9 p.syn:HVO_2259 cloning and phenotypic analysis

To validate HVO_2259 as the gene imparting UV resistance phenotypes on colonies 7 and 21, it was of interest to amplify the genomic CDS and place it under a stronger, constitutive promoter, p.syn. In order to do this, firstly pTA1992 was digested with Pci1 and BamH1. PCR amplification was carried out using primers:

XerC_F

XerC_R

to incorporate Pci1 and BamH1 sites at the 5' and 3' ends, respectively. After the bands were extracted, a ligation was set up in which a molar ratio of 5:1 (insert:vector) was used overnight at 15°C. After the ligations were cleaned via ethanol precipitation, and transformed into *E. coli* strain *XL1-Blue* via electroporation, miniprepped DNA was used to set up a diagnostic digest to confirm colonies of interest. As a result, a colony was prepped and sequenced by dideoxy sequencing (UoN, Deep Seq) to confirm cloning success. The resultant plasmid is known as pTA2773 (Figure 4.17).

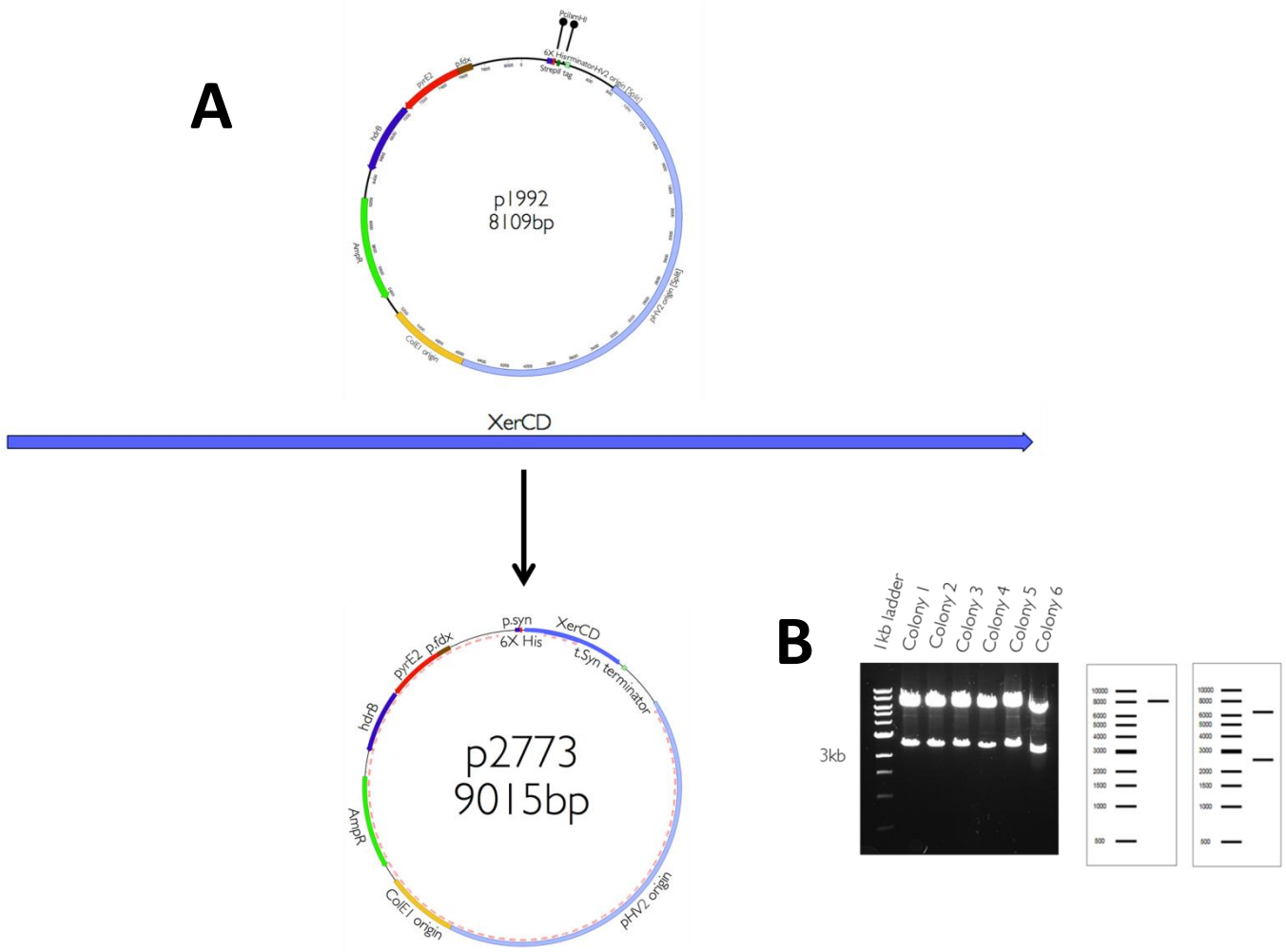


Figure 4.17 | **Cloning strategy for generation of pTA2773.** (A) Overview of cloning employed to generate pTA2773. pTA1992 was digested with Pci1 and BamH1, along with a PCR amplified XerC/D-like integrase (HVO_2259) with the same restriction sites. Ligations were carried out with a 5 molar excess of insert compared to vector. (B) In silico prediction of products of digestion of pTA1992 and pTA2773 (right) with ClaI and HindIII. Gel image of diagnostic restriction digest of pTA2773 (left) using six randomly selected colonies. Colony 1 was sequenced and confirmed correct.

4.4.10 p.syn:XerCD UV Resistance Assay

Once pTA2773 was generated, the plasmid was passed through a Dam- strain of *E. coli* (N2338 (GM121)) in order to allow subsequent transformation into background strain H1206, the same strain used for genome library construction and screening. After transformation, colony PCR was carried out to confirm the presence of the vector (figure 4.18) using primers:

pBSF2

PhageProbeBII

The correctly transformed strain was then subjected, along with the usual controls (H1206 and H645) to a UV resistance assay to confirm the UV resistance phenotype when HVO_2259 was independently amplified and placed under a stronger promoter. Results can be seen in figure 8.7.

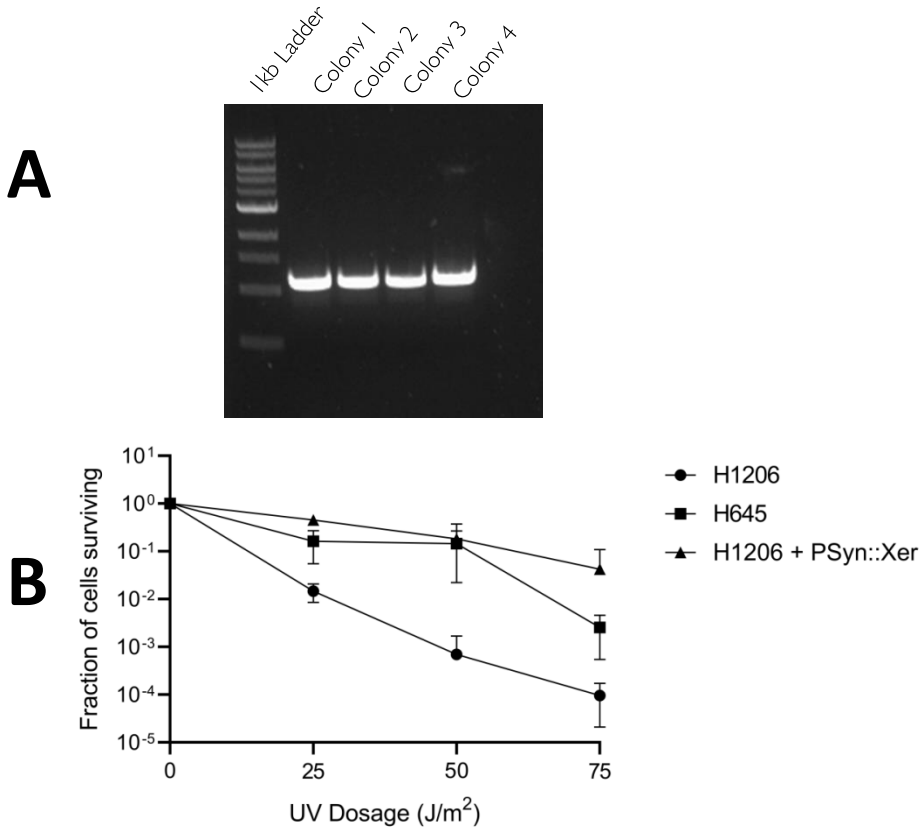


Figure 4.18 | **Generation of p.syn::Xer strain and UV resistance assay.** (A) Colony PCR of strain transformed with pTA2773 to generate overexpression of XerC/D-like integrase, HVO_2259. (B) UV resistance assay using transformed strain (H1206+p.syn::Xer).

4.4.11 UvrABCD Complementation Analysis

In order to attempt to dissect the mechanism by which XerCD overexpression provides resistance to UV radiation, complementation analysis was carried out. Key candidates of interest are the UvrABCD proteins, given their prominent role in nucleotide excision repair – a key DNA repair process in the repair of UV induced lesions, such as cyclobutane pyrimidine dimers. One main theory would be that, due to the recombinase function of XerCD in *E. coli*, in order to assist with deconcatenation of chromosomal dimers during DNA replication termination, XerCD may be able to assist, and therefore complement defects in, UvrC. UvrC being the canonical nuclease function within the nucleotide excision repair pathway in *E. coli*.

Strains deleted for UvrABC (H509, H1181 and H1187) were all transformed with the pTA2773 construct, restreaked and confirmed by colony PCR using the primers:

pBSF2

PhageProbeBII

Strains were then subjected to a UV damage assay using standard protocol. Strains were grown overnight in Hv-Ca or Hv-Hv-Ca+Ura broth and were subsequently diluted to an A_{650} of 0.4. Cultures were diluted in 18% salt water and spotted before being left to dry and then being subjected to UV irradiation. Plates were left to grow for 4-5 days, in the dark to minimise photolyase activity, after which survival fractions were counted (figure 4.19)

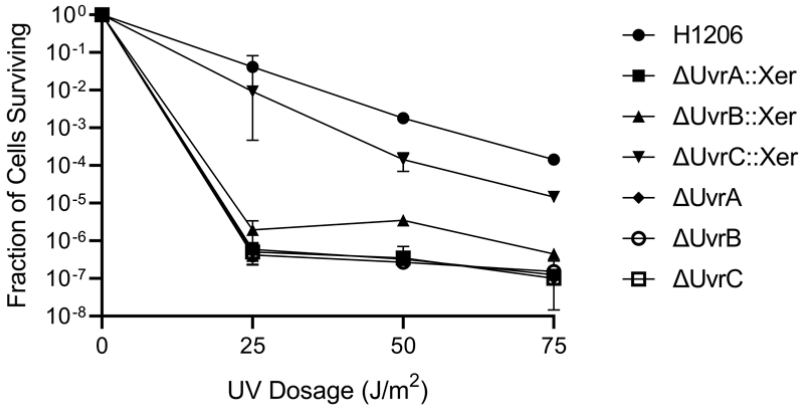


Figure 4.19 | **UV resistance assay of Δ*UvrABC* complementation strains.**

Of note, the Δ*uvrC* strain appears to have its growth defect complemented slightly compared to the other Uvr deletion strains, with UvrB coming runner up. The UvrB complementation is not as large as the UvrC, however subsequent testing and further analysis will be required to confirm this as the genuine reason that XerCD is assisting with UV resistance. It does, however indicate the ability of XerCD to act as a nuclease, further adding weight to the notion that it acts in DNA replication termination, specifically to resolve chromosomal dimers that arise through unequal numbers of crossover events.

4.4.12 Mre11/Rad50 epistatic analysis

To test the hypothesis that XerCD acts in the same pathway, or assists in some way with Mre11/Rad50, which would implicate the role of XerCD within the pathway of homologous recombination, the strain deleted for Mre11 and Rad50 (H645), which already shows a level of UV resistance compared to wild-type, was transformed with pTA2773 and confirmed using colony PCR using the primers:

pBSF2

PhageProbeBII

The resultant strains were then subjected to the same UV resistance assay to determine if XerCD acts in either the same, or a different pathway (figure 4.20).

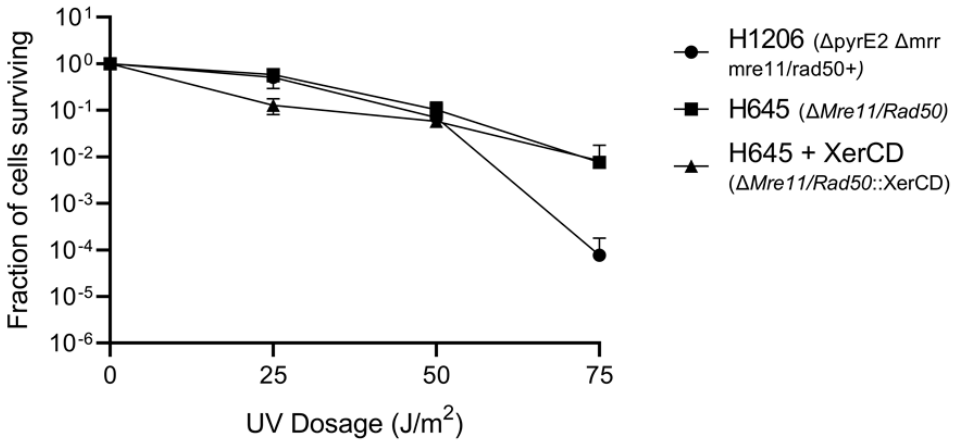


Figure 4.20 | **Epistatic analysis of Δ*mre11/rad50* strain, H645.**

As seen above, wild-type cells act as expected according to the variation seen during each UV experiment. Therefore, complementing H645 with XerCD seems to have no effect on the overall level of UV resistance seen from this strain, indicating that XerCD could be acting in the same pathway of UV resistance, namely homologous recombination. It could also be the case that the endogenous resistance of the strain is too high to see any further difference. Again, more experiments would be required to further tease out the exact nature of what XerCD is capable of doing, and under what conditions. However, initial results show that XerCD could be acting in multiple pathways and therefore could be considered a scaffold protein complex with roles outside of DNA replication termination.

4.4.13 Hjc/Hef complementation analysis

To further implicate XerCD in the pathway of homologous recombination, perhaps lending weight to a putative resolvase function, deletions of the resolvases in *H. volcanii* were complemented with XerCD.

Strains deleted for both Hef and Hjc (H358 and H178, respectively) were transformed with pTA2773. Colonies were confirmed by colony PCR using primers:

PhageProbeFII

PhageProbeBII

Colonies were selected in biological triplicate, grown to an A_{650} of 0.4 and diluted as in previous spotting assays. Spots were allowed to dry on either Hv-Ca or Hv-Hv-Ca+Ura plates before being subjected to varying doses of UV, or left untreated to allow for normalisation (figure 4.21).

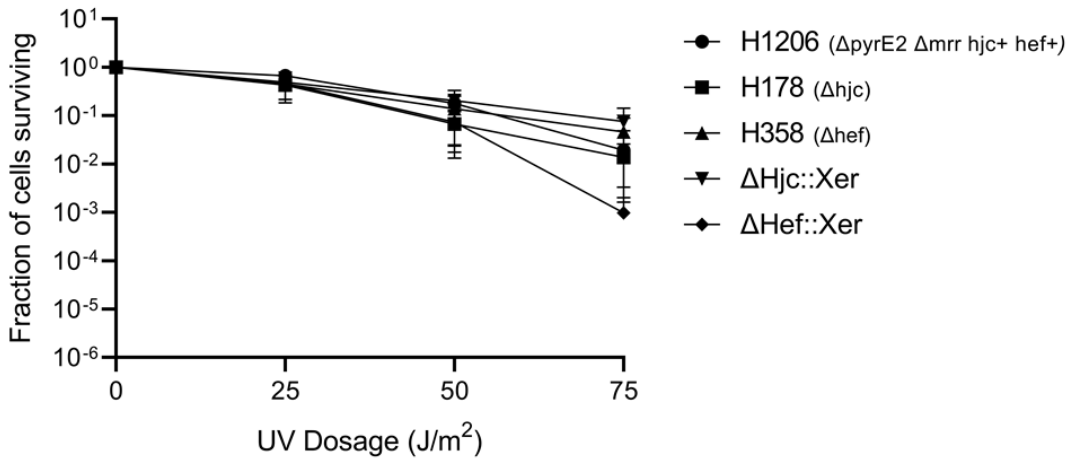


Figure 4.21 | **Complementation analysis of Hjc and Hef deletion strains.**

Unfortunately, the wild-type strain appears to rest higher on the Y axis than is expected from all other survival assay data using this strain. It is therefore assumed that all data points are shifted to reflect a higher UV resistance level than is reality. However, the relative differences should not be affected, assuming this is the case. Of course, it could only be the wild-type control that is shifted, therefore additional work would be required to repeat this experiment again should this line of enquiry prove interesting. The data however currently suggests that there is no significant phenotype upon expression of XerCD. However, looking at the data one could argue that the resistance levels of Δ*Hef* decrease when XerCD is overexpressed, and Δ*Hjc* shows increased levels of resistance when XerCD is overexpressed. This is perhaps not surprising given that Hef has not been shown to act in nucleotide excision repair, but Hjc is a known resolvase, again implicating XerCD as having a resolvase function. Data on this is however weaker than hoped, and was not furthered or repeated due to time constraints. Further work is advised, including repeating the assay above.

4.4.14 Addition of C-terminal StrepII tag to XerCD under p.syn promoter

The Strep-tag II binds specifically to engineered Streptavidins, known as Strep-Tactin by occupying the binding site of the natural ligand biotin. Elution is achieved by competitive binding of Desthiobiotin to the binding pocket on Strep-Tactin (Schmidt et al., 2013). In order to add a StrepII tag to HVO_2259, it was first of importance to determine if either the N or C terminus would suit best. In order to do this, crystal structures of *E. coli* homologues were analysed, and the protein sequence for HVO_2259 was put through AlphaFold to determine a potential structure, given no crystal structures from Archaea are available (figure 4.22). As a result, and given that C terminal tags appear to function better in *H. volcanii*, compared to N terminal tags in *H. volcanii*, a C terminal tag was chosen. C terminus shown in red, N terminus in blue. Functional analysis has shown that the C-terminal domains of both XerC and XerD in *E. coli* are important for coordinating catalysis within the heterotetramer (Ferreira et al., 2003). The tag may therefore interfere with catalysis, but this would need to be experimentally verified. If so, an N terminal tag must be used.

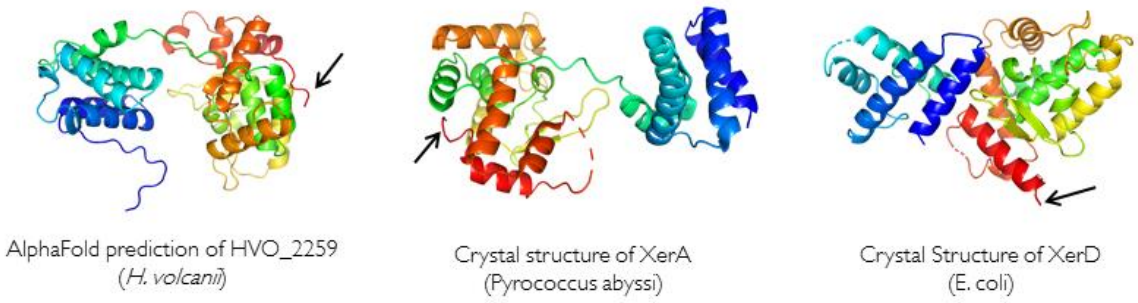


Figure 4.22 | **Protein structres and comparisons of XerC/D-like integrase (*H. volcanii*), XerD homologue (*E. coli*) and XerA (*P. abyssi*).** Although the N terminus does protrude from the structure enough to facciliate a tag, the C terminus also does so, and functions better with tagging in *H. volcanii*. C terminus is red, and indicated with an arrow in each case.

After determination of which terminus to apply the tag, cloning to add a C terminal Strep II tag was done by first carrying out PCR from genomic DNA (strain H26) using primers:

XerC_F

XerC_R_Nhe1

XerC_F adds a 5’ Pci1 site, whilst XerC_R_Nhe1 adds a 3’ Nhe1 site. Both sites facilitate in-frame cloning into destination vector pTA1992, which already contains a C terminal strep II tag downstream of the polylinker. The PCR product was purified and ligated, using a 3:1 molar excess of insert:vector, to pTA1992 digested with the same restriction enzymes. After which, an ethanol precipitation was carried out and the construct was ligated, along with a vector only control, into *E. coli* strain *XL1-Blue*. Six colonies were selected and miniprepped, after which a diagnostic digest was carried out to confirm cloning success. The new vector is known as pTA2781, and was also checked by Sanger sequencing (figure 4.23).

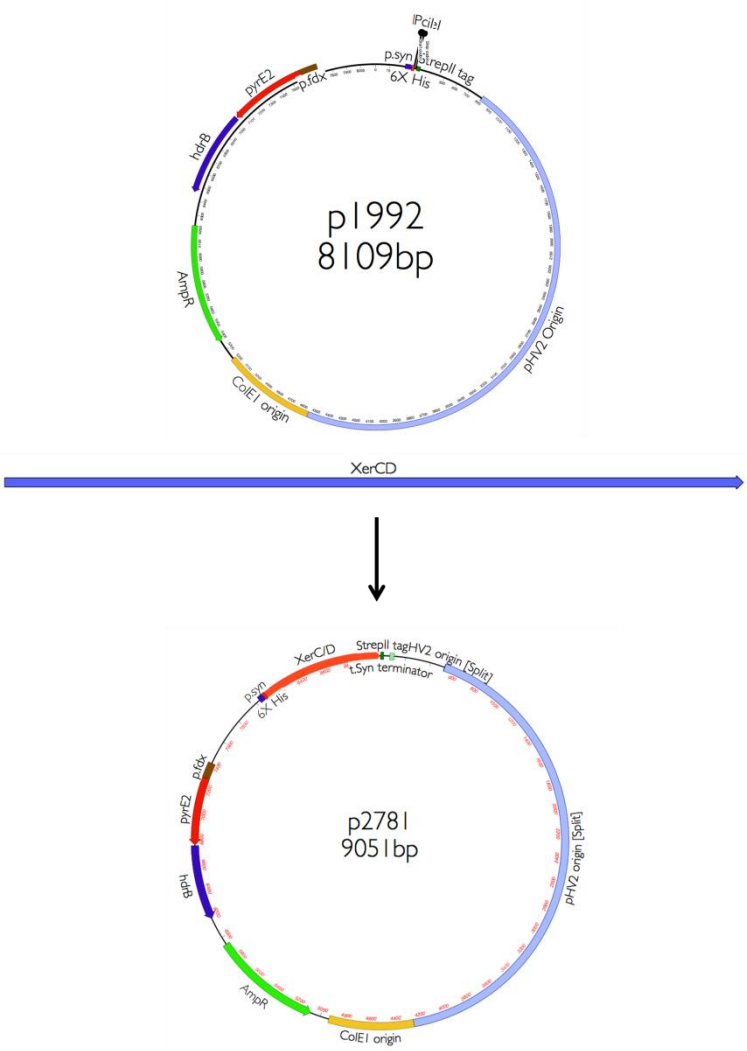


Figure 4.23 | **Cloning strategy for generation of C terminally tagged XerCD (pTA2781).**

4.5 Discussion

4.5.1 Sequencing indicates XerCD as gene of interest

Of all the four colonies chosen, sequencing of the constructs yielded two identical fragments. Each fragment has therefore been isolated independently from the same screen, so attention is immediately drawn to these constructs. Interestingly, the remaining construct (gLib1 was discounted due to potential backbone interference, figure 4.6) was BLASTed to the genome and either end of the sequence indicated that multiple fragments had been cloned into the vector, leading to this also being discounted. It was considered that resolution of this would require long-read sequencing to confirm the exact sequences present. This would have taken a prohibitive amount of time at the stage of the project, and so the construct was de-prioritised. The remaining fragment was interrogated, and yielded mostly hypothetical proteins, less for HVO_2259, a XerC/D-like integrase within the prophage region in the *H. volcanii* genome. This therefore was the first gene on the list to sub-clone and test.

4.5.2 UV resistance correlates with tryptophan dependence and is not due to host genotype

The results seen in section 4.4.3 indicate that, apart from colony 1 which was then discounted from the study, all other constructs showed an increased level of UV resistance compared to wild-type, when transformed into a different background strain, in this case – H26. All cloning to this point has been carried out using H1206 as the host strain, for its ease conferred by Δ*mrr*. This result indicates that the resistant phenotype seen, is not due to a genetic alteration that has occurred in the host strain through the cloning workflow and library generation and is instead due to something present on the episome.

Section 4.4.2 places each genomic library construct onto plates laced with a gradient of tryptophan, before being subjected to UV irradiation. Along with the relevant controls, both non-irradiated plates, and streaking of a known trp-

dependant strain, the results indicate that the UV resistance phenotype observed is correlated with a dependence on tryptophan concentration in the media.

This result shows, along with the results of section 4.4.3, that not only is the resistance phenotype due to the episomes, but specifically rules out any genetic alteration in the vector backbone that may confer this resistance, resistance is due to the fragment immediately downstream of the tryptophan inducible promoter. Despite this, the result is not as stark as expected, perhaps due to the promoter, p.*tnaA*, not being particularly active, or indeed the fact that the XerCD coding sequence is not the first sequence immediately downstream of the promoter, so RNA copy number will be diminished compared to that of other coding sequences at any given time. Furthermore, each fragment is assumed to retain all native promoters for each coding sequence, therefore any given gene will not be entirely reliant on the level of tryptophan present.

This result is however, a good one and shows that downstream analysis can now occur.

4.5.3 Overexpression of genome library fragments has no impact on growth

The impact on growth was assayed for each of the four genomic fragments in section 4.4.4, whilst under p.*tnaA* and their native promoters. Overexpressing these constructs has no effect on growth. This is however not entirely surprising given the discussion in section 4.5.2, in addition to the fact that irradiation of liquid culture is very difficult to do in any significant quantity. 150µl needed to be irradiated in 96 well plates, which means that not every cell in each well will be irradiated to the same extent, perhaps skewing the results somewhat. In either case, it was still hoped that a relative difference would still be seen between the fragments, if one does indeed exist. Lastly, the lamp used was the original lamp, before it broke and the UV crosslinker was used, therefore the UV dosage, although calibrated for 300J/m^2, was in fact most likely far less than this.

4.5.4 XerCD overexpression increases UV resistance phenotype

XerCD coding sequence was PCR amplified from the genome and placed under p.syn, a constitutive and strong promoter in *H. volcanii*. This both provided increased evidence for XerCD being the gene of interest, by virtue of the fact this is now an independent amplification from the genome, and showed that the resistance phenotype seen can be reconstituted and resistance levels increased. As a result of this, seen in section 4.4.9.1, XerCD was now pinned as the gene of interest within the genome library fragments. Resistance levels are increased compared to both wild-type and the Δ*mre11/rad50* strain, which already shows a level of UV resistance above that of wild-type. This result also indicates that the level of transcription in both wild-type cells, and in the genome library fragments, was a limiting factor. Placing the construct directly under the control of p.syn shows an increase in resistance level seen.

This result, and given that XerCD is part of an integrated prophage, indicates that under normal conditions, XerCD, and by inference the rest of the prophage region, is not actively transcribed to biologically significant levels. This is understandable, given the recent acquisition of the prophage region into the *H. volcanii* genome (see chapter 1 for details) and the usage of rare codons in the region. Placing XerCD under p.syn therefore 'activates' it to bring levels of transcription up to, and exceeding that of normal levels, allowing the protein product to act to increase levels of UV resistance. One should note, that for future work, a Cas3 deletion strain has been shown to have increased levels of UV resistance. Complementing this defect with XerCD may prove a useful future experiment (Miezner *et al.*, 2023).

4.5.5 XerCD deletion has no impact on survival fraction under UV stress but prophage deletion does

The next step was to generate a deletion strain for XerCD only, as an entire prophage deleted strain already existed. The results of DNA damaging assays, carried out under UV and MMC genotoxic stress, can be seen in section 4.4.7. Results from this are rather interesting, and both UV and MMC display the same, or similar, results in survival fraction, therefore both will be considered as one for the purposes of discussion. MMC is a crosslinking agent, therefore the lesions imposed are similar to that of UV light, namely crosslinks in the DNA duplex.

The deletion for HVO_2259 (XerCD) only, yielded no difference in survival fraction compared to that of the wild-type. This is not surprising given that twelve homologoues of XerCD exist within the genome (Pérez-Arnaiz *et al.*, 2020), and are not all contained within the prophage region. Therefore, it is assumed that other homologues can stand in for HVO_2259 (XerCD) when this is deleted from the genome. If indeed a cell 'requires' twelve homologues, then it is not too difficult to imagine that overexpressing one under UV stress could yield a benefit to the host. It is also possible that, due to the natural levels of XerCD assumed to be so low, or even not active at all, (when not placed directly under p.syn) deletion therefore does not yield a significant UV sensitivity phenotype compared to wild-type. Complementation of the deletion, with XerCD under its native promoter equally does not alter the survival fraction under UV or MMC stress compared to wild-type.

Of real interest, however, was the result of the entire prophage deletion strain. Here, UV and MMC resistance appears to be increased significantly compared to that of wild-type. Cells appeared much healthier on the plates, too - growing quicker and larger when under genotoxic stress. It is important to consider why this would be, when XerCD alone has no effect on levels of UV or MMC resistance. Given this result, it is hypothesised that either:

1. Gene products within the prophage region inhibit novel or canonical DNA repair pathways.

2. Some phage possess Xer recombination activators, such as those from the DUF3653 family. These drive recombination in the absence of FtsZ binding (which XerCD usually required). It is possible that the hypothetical proteins, or indeed HVO_2259 itself, could be one of these and could thus activate other Xer homologues in the genome simultaneously (Midonet *et al.*, 2019).

3. Upon genotoxic stress, either UV or MMC treatment, the phage enters a lytic cycle, to excise itself from the genome, produce more phage and then exit the host cell, lysing it in the process – resulting in cell death. The lack of a prophage region, would render this process impossible and lead to increased cell survival.

The most likely scenario, and easiest to dissect further, is hypothesis three – UV/MMC treatment induces phage lysis and cell death. Indeed, this would mean that any survival assay with UV or MMC, the resulting survival curve is not a result of treatement alone, and in fact the cellular death seen is a combination of both irradiation, leading to irreversible DNA damage, and cell death from phage induction and lysis of cells. Of course, this, if true, would have an impact for all previous and future survival assays done using UV or MMC – they should be done in a phage deleted strain.

Further evidence for this hypothesis comes from unpublished data from our lab indicating through RNASeq of wild-type cultures exposed to MMC, that there were 527 differentially expressed genes (Cagla Touson MRes thesis, 2019). A selected list contained the XerC/D-like integrase, HVO1422, which is enriched 3-fold under MMC treatment compared to an untreated control sample. Whilst not in the prophage region itself, it is a homologue of HVO_2259, which has arisen from this study, therefore it is expected that MMC treatement (and possiblty UV treatment, although this is difficult to do for reasons discussed in section 5.4.3) should induce HVO_2259 XerC/D-like integrase, too. RNASeq of MMC treated samples of H26 and H1192 (Δ*prophage*) will therefore be carried out and will represent a key step in this study, along with a qPCR to validate any phage induction after MMC treatment.

4.5.6 XerCD deletion has no impact on growth

The XerCD deletion strain was used, along with an empty vector control and the complemented deletion strain with XerCD under its native promoter, for a standard growth assay measuring A_{650} over ~ 72 hours. Results indicate that the deletion has no impact on the growth of cells compared to wild-type, so XerCD when absent has no impact on either growth or UV/MMC sensitivity. This, again, is perhaps not surprising given the idea that the entire prophage region, and indeed homologues outside of this region, are not active under normal conditions. Clearly, however there is little energetic cost to maintaining XerCD, given that no growth defect is seen when deleted. That said, this is not confirmed for the entire prophage region, so using the entire prophage deletion strain in a growth assay would be of interest to determine how energetically costly this is to maintain within the genome. It is likely that the cost is small as even plasmid costs are rapidly lost due to fixation of compentatory mutations in other genes.

4.5.7 XerCD nuclease activity may complement UvrC defect

A significant result obtained from this chapter is the potential for XerCD to act, using its nuclease activity, in place of UvrC – the canonical nuclease in the nucleotide excision repair pathway, a major repair pathway for UV induced lesions. The result can be seen in section 4.4.9.2 and indicates that XerCD, under p.syn, can go some way to rescuing the phenotype of a single deletion of UvrC, which normally renders the cell more sensitive to UV induced DNA damage. Interestingly, the phenotype is not rescued back to 100% of its initial level, indicating that other factors are at play and that XerCD alone cannot complete all tasks assigned to that of UvrC. This is also initial data, so further study is warranted to ensure that this is indeed what XerCD is doing. Given the function of XerCD as a nuclease and recombinase, it is not suprising that it may have multiple functions and can therefore interact with numerous pathways and lend assistance where necessary. It is possible that the decatenation function accounts for this partial recovery by allowing more homologous recombination events to repair these lesions rather than excision.

Chapter 5: Genetic analysis of prophage activity in *H. volcanii* in relation to DNA damaging assays

5.1 Background

The genome of *H. volcanii* contains a putative 50kb phage region integrated into the main chromosome as a prophage. The phage is lysogenic and contains many hypothetical proteins alongside known coding sequences. For example, genes encoding Orc11 and Orc14 are both located within the prophage region (Hartman *et al.*, 2010) (Pérez-Arnaiz *et al.*, 2020).

DNA replication is not a continuous process, and is often interrupted and halted by various DNA lesions such as UV radiation, radical attack and other genotoxic agents such as MMC. Homologous recombination represents a key method in which cells repair DNA damage that has occurred, and not been fixed whilst still classified as a lesion. If such a process is involved during DNA replication in organisms with circular chromosomes, and an unequal number of recombination events occurs, the two chromosomes in question will become linked together as a concatemer. To ensure correct chromosomal disjunction, bacteria and Archaea have overcome this issue by resolving such dimer formation by using site-specific recombinases that act on such structures immediately prior to chromosomal segregation.

Importantly, XerCD is located within the prophage region and in total consists of twelve homologous sequences littered throughout the genome. XerCD, in bacteria, act as site specific recombinases and deconcatentate chromosomal dimers formed via unequal recombination events during DNA replication, thus facilitating the termination stage (Castillo, Benmohamed and Szatmari, 2017, Farrokhi, Liu and Szatmari, 2019). XerC and XerD are tyrosine recombinases (figure 5.1) that act specifically on a 28bp target *Dif* site (differential induced filamentation site). Such *Dif* sites are in close proximity to canonical termination sites (*ter* sites).

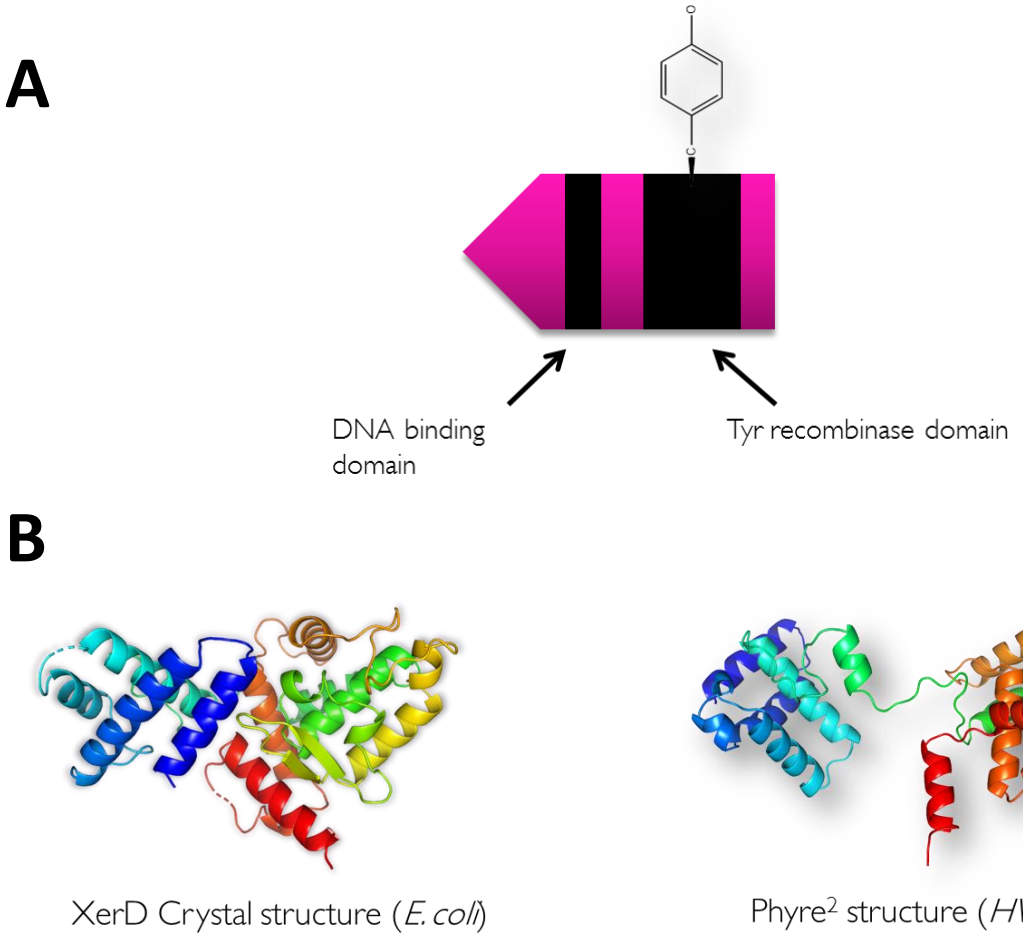


Figure 5.1 | **Domain architecture of XerCD tyrosine recombinase and proposed crystal structure prediction**. A) Domain architecture of XerCD recombinase. B) Comparison of solved crystal structure of XerD from *E. coli* compared with Phyre[2] prediction of HVO_2259 from *H. volcanii*.

The essential mechanism by which XerCD carries out this process is by targeting *Dif* sites on chromosomal dimers. XerCD complexes initiate the process by the activation of XerD by the gamma sub-domain of FtsK, which translocates towards the XerCD/*dif* site to mediate initial strand exchange immediately prior to cell division, thus generating a transient Holliday junction intermediate. XerC then mediates second strand cleavage and exchange to allow separation of the duplexes (figure 5.2). The catalytic mechanism by which this occurs is similar to that of other recombinases. The hydroxyl group on the active tyrosine residue attacks the scissile phosphate thus mediating formation of a 3' covalent phosphotryosyl enzyme-DNA covalent intermediate and a free 5'-hydroxyl end. This intermediate is attached by the other 5'-end to obtain the recombinant product (figure 5.2) (Castillo, Benmohamed and Szatmari, 2017).

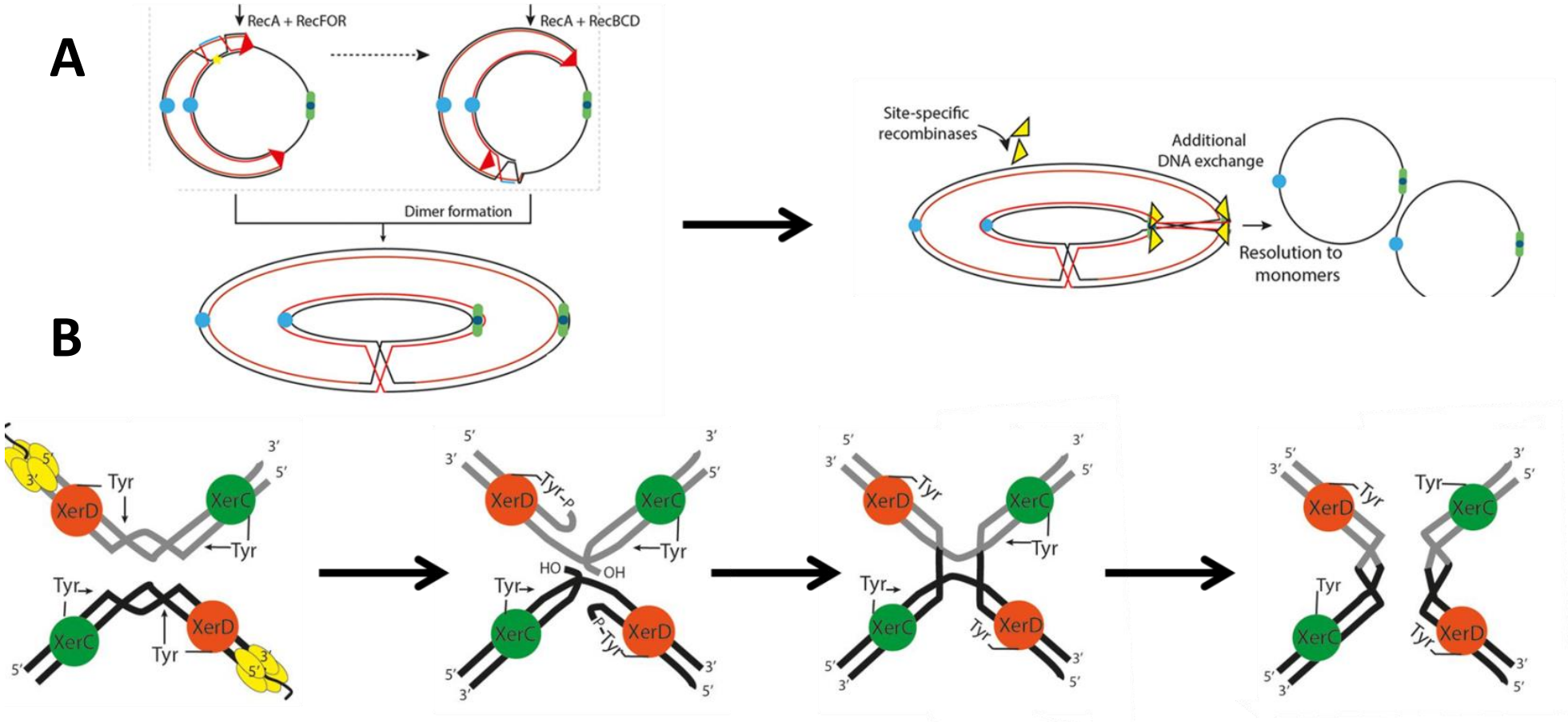


Figure 5.2| **Mechanism of XerCD resolvase in *E. coli*.** A) Chromosomal dimers can form from unequal exchange events during double or single stranded break repair, using RecBCD or RecFOR respectively. XerCD (yellow triangles) is recruited to the terminus region (green circles) and acts on the *dif* site (blue dot). Origins of replication are shown as light blue dots. B) FtsK (yellow) translocates along the DNA duplex until XerD is reached. Upon activation XerD catalyses the first strand exchange, followed by XerC, resulting in resolution of the DNA dimer into separate DNA duplexes. Figures adapted from (Castillo, Benmohamed and Szatmari, 2017)

In spite of the knowledge of XerCD systems in bacteria, knowledge of if and how XerCD homologues participate in DNA replication termination in Archaea is substantially lacking. and research into XerCD specifically as an agent of DNA repair is negligible (Alarcón-Schumacher *et al.*, 2022).

It has been reported by our lab previously (unpublished data) that certain XerCD homologues are upregulated in response to MMC treatment. It is therefore interesting to confirm this is true by repeating such an experiment using RNA Seq, but also prudent to include the phage deletion strain alongside to tease apart the differential responses – death due to phage activity and MMC compared to MMC alone.

Of note, it is hypothesised through this study that the prophage region within the main chromosome of *H. volcanii* undergoes a lytic pathway as a result of either irradiation with UV light, or treatment with MMC and thus contributes to cellular death observed when conducting DNA damaging assays.

5.2 Aims and Objectives

The aims of this chapter are to investigate the potential of prophage induction after treatment with DNA damaging agent MMC to assay the possibility that the phage becomes lytic and contributes to the cellular death seen during such experiments. This, one would hope, would contribute to the result seen in chapter 4 where deletion of XerCD (HVO_2259) only has no effect on UV resistance, whereas deletion of the entire prophage region shows increased resistance. Therefore the specific aims are as follows:

- Generate a codon optimised version of the XerCD coding sequence to see if UV resistance can be further increased.

- Carry out qRT-PCR (ΔΔCt) on the prophage region against an endogenous control under MMC-treated conditions. This will aim to determine if phage is induced under such conditions.

- Carry out RNA-Seq with wild-type and phage deleted strain under MMC to look for transcriptomic responses.

5.3 Materials

5.3.1 Strains

[ ] Indicate integrated plasmid DNA

{ } Indicate episomal plasmid DNA

Table 5.1 ***Haloferax volcanii*** **strains used in this chapter**

Strain	Parent	Genotype
H1206	H1202	ΔPyrE2, Δmrr
H26	H18	ΔPyrE2
H1192	H1210	ΔpyrE2 Δphage
H5630	H1206	ΔpyrE2 Δmrr {p.syn::XerC/D pyrE2+ hdrB+}
H5655	H1206	ΔpyrE2 Δmrr {p.syn::XerCD_Cdn_Op pyrE2+ hdrB+}

5.3.2 Plasmids

Table 5.2 **Plasmids used in this chapter**

Name	Use
pTA1992	Backbone for cloning of pTA2868
pTA2868	XerCD codon optimised construct. XerCD codon optimised CDS placed under P.Syn for DNA damaging assay.

5.3.3 Oligonucleotides

Table 5.3 **Oligonucleotides used in this chapter**

Name	Sequence
Xer_CdnOp_v3	ATCTGACATGTTCAAGGAGGGGAACTACCATGGGCGCGGAGCCGGGCTCGTCGAAAATC TACGACAACAAGCACGACGAAGTCAACTACTTCATCACCCGGAAGCACGCGACCGGCCG GAGCGAACGGACGCTCAACTCCTACTCGCGCATCCTCCGCGAGTTCTTCCACGACCAGTTC CCGGACCTCTCGCCCTCGGAGGTCGAAATCCGGCACGTCGAGGACTACCTCATGGCGCTC ACCGACCGCGGCGTCTCGCAGAACAGCAAGAAGAAGTACCTGGAGGTGCTCTCGTCGTTC TACGGCTACACCCTCAAGCGCCCGCAGTTCGAGGGGATTACGTCGAACCCGGCCGCGGTC GTGATGGAGGAGATCCCCCGCGTCCGGCCCGACCGCCCGGACTGTGCCACCTGGGAGAA CGCCTGCAAGCTCATCAACGCCATCCCCGACCCGCGGGACAAAACGGTGACGATCATCCT GGCGAAGACGGGCGCGCGCCTCCTGGAAGTCCTGTCGATTGAGGAGGACGACGTCGACC TGGAGAAGGGCTTCATTCGCCTCCGCGAACGCAAGGGCGGCAAACAAACCGTCGTGCCG ATCGACGACGAGACGATCTACGCGATCAAGCGGTACCAGTTCGTCAACGCGGACCTCGAC TCGCCCTACCTGTTCACCTCGAACAAGGGGGGCCGGCTCTCCAAGGAGCGCATCCGCCGG GAAGTGAAAGCCGCGGCCGACCGCGCCGGCGTCGCGCCCAAGGAAGAACGCCGCTTCG AGAAGAAGTTCACGCCGCACACGTTCCGCACCGTCTTCACGACGCTGATGCGCAAGCAGG GCATGAAACCGTACATCCTCAAATACATTCGCGGCGACGCCAAGACGGAGACGATGGAC ATCTACACGCGGGTCGACCGCGACGAAGCCAAAGAGGAATACCTCAACTGCATTAAGGA GATCGGGCTGTGACGAGGATAGACAGTTGGTGGATCCGCTAGCACTCG

pBSF	5’-GTAAAACGACGGCCAGT-3’
pBSR	5’-AACAGCTATGACCATG-3’
p.synF	5’-CGAGAATCGAAACGCTTATAAGTGCCCCCCGGCTAGAGAGAT-3’
p.synR	5’-CGATCTCTCTAGCCGGGGGGCACTTATAAGCGTTTCGATTCT-3'
PhageProbeF3	5’-TTCACAAGCAACAAAGGAGGACGG-3’
PhageProbeBII	5'-ACTCTTCTTTCGCTTCGTCACG-3'
Hel308R	5’-CGACGACGCAGGTGAGTTGG-3’
Hel308Fint	5’-AGCGCTGGGAGGAGTACGGC-3’

pTA1992

pTA1992 is a shuttle vector that contains the constitutively expressed p.syn promoter sequence directly upstream of a MCS. HVO_2259 is placed in this vector to allow high levels of transcription, in the case of pTA2773, and a C-terminal strep II tag, in the case of pTA2781.

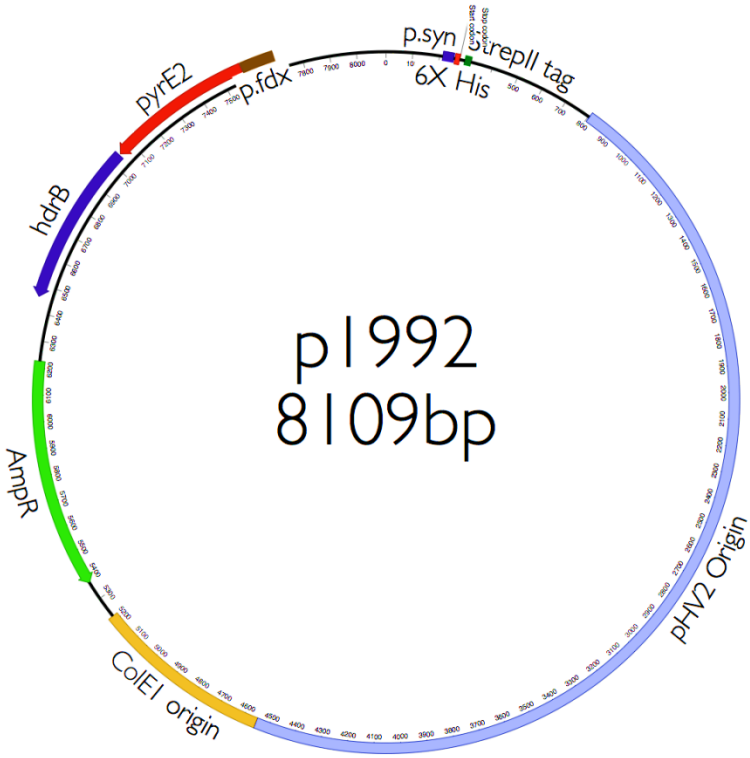


Figure 5.3 | **pTA1992**. Constructed by TA lab.

pTA2868

XerCD codon optimised coding sequence placed under the strong promoter, p.syn. Shuttle vector used for DNA damaging assays with codon optimised XerCD.

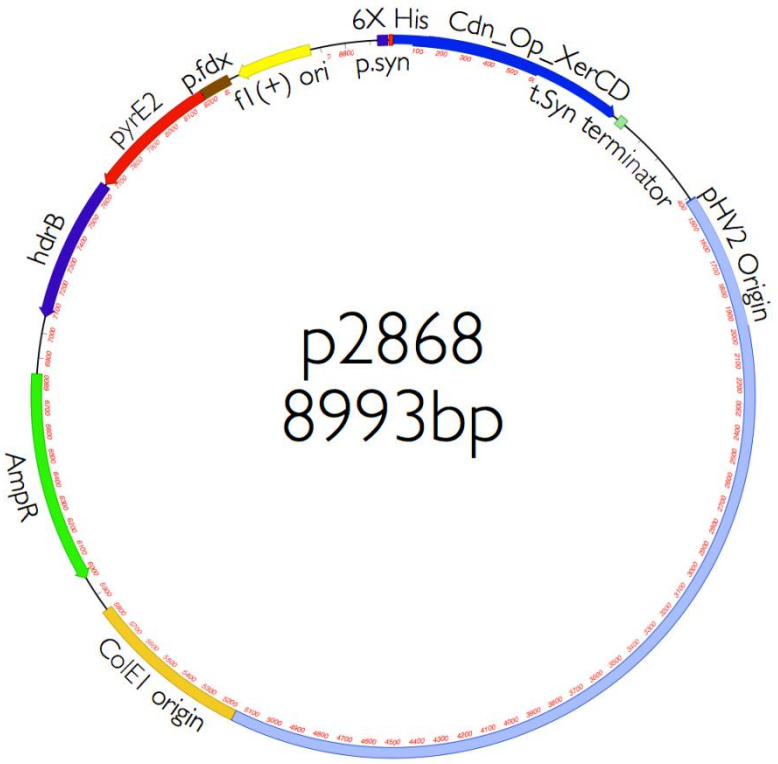


Figure 5.4 | **pTA2868**. Constructed in this study.

5.4 Results

The aim of this chapter was to generate a codon optimised XerCD coding sequence and see if this can contribute further to UV resistance. Results from this will determine if transcription or translation is the rate limiting step. Real-time PCR will aim to determine if the prophage is becoming lytic as a result of MMC exposure, and results of this will go some way to supporting the theory that phage induction is contributing to cell death. RNA sequencing also should confirm what genes are upregulated in response to MMC compared to an untreated control. The prophage deleted strain will be used as a background control.

5.4.1 Codon usage in the prophage region

In order to interrogate the function and role of XerCD further, the genomic region was scanned, yielding majority hypothetical proteins. It was of interest to scan for codon usage across this region as XerCD lies within a 52kb prophage region within the *H. volcanii* genome.

The region in question is A/T rich, indicating recent acquisition into the genome and that the region has not had time to bring GC content up to the typical level seen within the rest of the genome of *H. volcanii* (approx 65% GC). Codon usage was therefore predicted to be rare within this region, perhaps hindering translation efficiency of coding sequences within this region.

The GC content of the genome was plotted against chromosomal location, clearly showing the prophage region with reduced levels of guanine and cytosine bases compared to background (figure 5.5). The region is littered with rare codons, so it would be of use for future work to optimise this region.

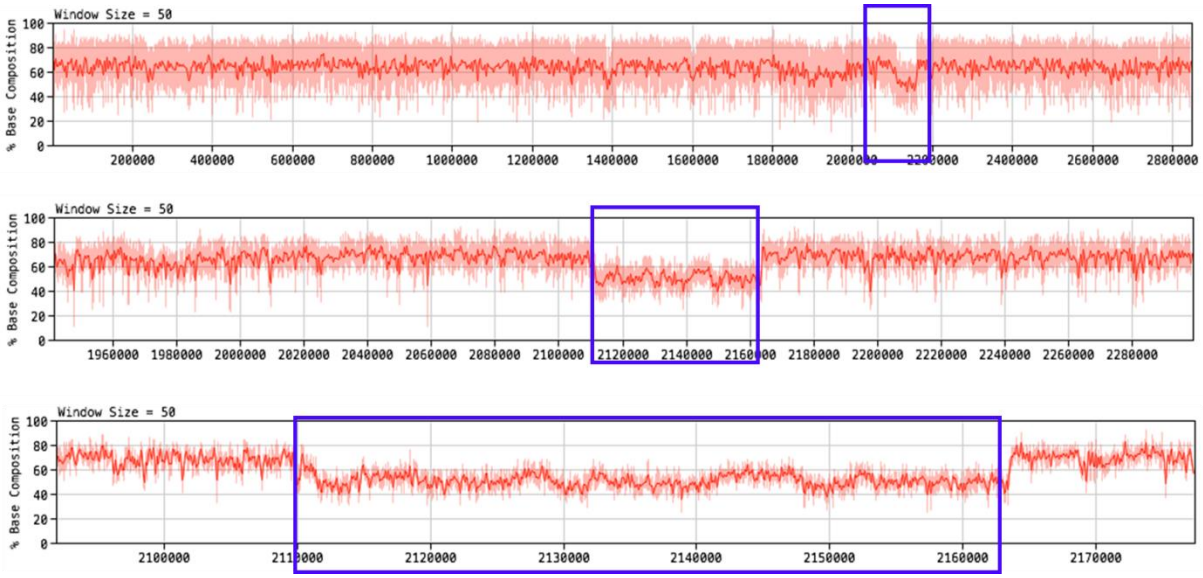


Figure 5.5 | **GC content of *H. volcanii* genome.** Prophage region is indicated by blue boxes, showing reduced G/C content compared to the rest of the genome. This region spans c. 52kb. Plotted using MacVector.

Further analysis of codon usage was carried out solely on the prophage region (figure 5.6). Rare codons are shown as red dots and canonical codons are not indicated.

Codon optimisation, carried out in section 5.4.2, and the result is shown in figure 5.6.

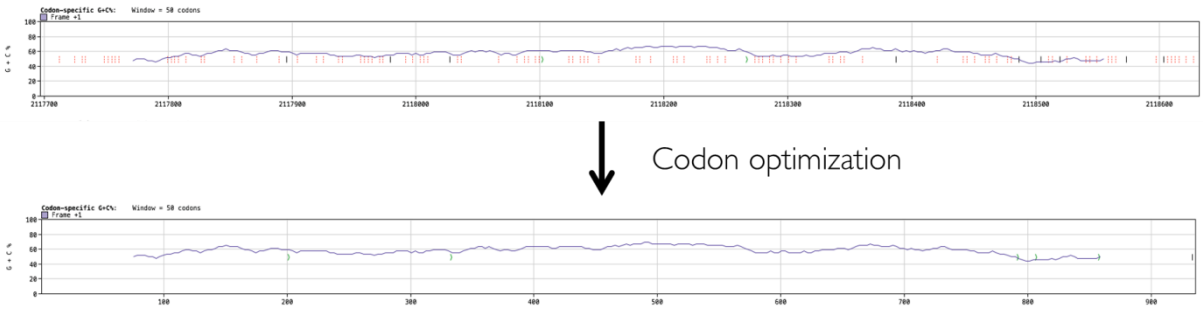


Figure 5.6 |**Codon usage across prophage region in *H. volcanii*.** Prophage region is shown with rare codons depicted as red lines across the blue genome. After optimisation via back translation, all codons are canonical.

5.4.2 Codon optimisation of XerCD

Codon optomisation was carried out using reverse translation (MacVector) according to the following optimisation matrix (figure 5.7, also an appendix).

Codon Bias for *Haloferax volcanii* main chromosome

Relative percentage per amino acid

	T				C				A				G			
T	TTT	Phe	F	7.75%	TCT	Ser	S	3.55%	TAT	Tyr	Y	8.04%	TGT	Cys	C	36.72%	T
	TTC	Phe	F	92.25%	TCC	Ser	S	22.78%	TAC	Tyr	Y	91.96%	TGC	Cys	C	63.28%	C
	TTA	Leu	L	1.07%	TCA	Ser	S	2.73%	TAA	***	*	22.95%	TGA	***	*	59.22%	A
	TTG	Leu	L	3.97%	TCG	Ser	S	44.59%	TAG	***	*	17.83%	TGG	Trp	W	100.00%	G
C	CTT	Leu	L	4.16%	CCT	Pro	P	1.94%	CAT	His	H	6.22%	CGT	Arg	R	4.42%	T
	CTC	Leu	L	67.15%	CCC	Pro	P	41.31%	CAC	His	H	93.78%	CGC	Arg	R	54.34%	C
	CTA	Leu	L	1.21%	CCA	Pro	P	3.48%	CAA	Gln	Q	17.69%	CGA	Arg	R	11.85%	A
	CTG	Leu	L	22.45%	CCG	Pro	P	53.27%	CAG	Gln	Q	82.31%	CGG	Arg	R	25.34%	G
A	ATT	Ile	I	14.50%	ACT	Thr	T	3.39%	AAT	Asn	N	6.51%	AGT	Ser	S	4.34%	T
	ATC	Ile	I	80.56%	ACC	Thr	T	46.04%	AAC	Asn	N	93.49%	AGC	Ser	S	22.01%	C
	ATA	Ile	I	4.93%	ACA	Thr	T	3.43%	AAA	Lys	K	26.99%	AGA	Arg	R	2.40%	A
	ATG	Met	M	100.00%	ACG	Thr	T	47.15%	AAG	Lys	K	73.01%	AGG	Arg	R	1.64%	G
G	GTT	Val	V	4.50%	GCT	Ala	A	3.41%	GAT	Asp	D	5.55%	GGT	Gly	G	9.73%	T
	GTC	Val	V	70.73%	GCC	Ala	A	45.39%	GAC	Asp	D	94.45%	GGC	Gly	G	64.57%	C
	GTA	Val	V	1.89%	GCA	Ala	A	4.67%	GAA	Glu	E	31.45%	GGA	Gly	G	7.54%	A
	GTG	Val	V	22.88	GCG	Ala	A	46.53%	GAG	Glu	E	68.55%	GGG	Gly	G	18.16%	G

Figure 5.7 |**Codon usage table for *H. volcanii*.**

The following construct was then ordered as a gBlock (Integrated DNA Technologies) for downstream processing. Of note, cleavage sites for NcoI and BamHI were included at the 5' and 3' ends of the coding sequence, respectively, to facilitate cloning. Restriction sites shown in yellow, start and stop codons in bold.

5’-
ATCTGACATGTTCAAGGAGGGGAACTACCA**TG**GGCGCGGAGCCGGGC
TCGTCGAAAATCTACGACAACAAGCACGACGAAGTCAACTACTTCAT
CACCCGGAAGCACGCGACCGGCCGGAGCGAACGGACGCTCAACTCCT
ACTCGCGCATCCTCCGCGAGTTCTTCCACGACCAGTTCCCGGACCTCT
CGCCCTCGGAGGTCGAAATCCGGCACGTCGAGGACTACCTCATGGCG
CTCACCGACCGCGGCGTCTCGCAGAACAGCAAGAAGAAGTACCTGGA
GGTGCTCTCGTCGTTCTACGGCTACACCCTCAAGCGCCCGCAGTTCGA
GGGGATTACGTCGAACCCGGCCGCGGTCGTGATGGAGGAGATCCCCC
GCGTCCGGCCCGACCGCCCGGACTGTGCCACCTGGGAGAACGCCTGC
AAGCTCATCAACGCCATCCCCGACCCGCGGGACAAAACGGTGACGAT
CATCCTGGCGAAGACGGGCGCGCCTCCTGGAAGTCCTGTCGATTG
AGGAGGACGACGTCGACCTGGAGAAGGGCTTCATTCGCCTCCGCGAA
CGCAAGGGCGGCAAACAAACCGTCGTGCCGATCGACGACGAGACGAT
CTACGCGATCAAGCGGTACCAGTTCGTCAACGCGGACCTCGACTCGCC
CTACCTGTTCACCTCGAACAAGGGGGGCCGGCTCTCCAAGGAGCGCA
TCCGCCGGGAAGTGAAAGCCGCGGCCGACCGCGCCGGCGTCGCGCCC
AAGGAAGAACGCCGCTTCGAGAAGAAGTTCACGCCGCACACGTTCCG
CACCGTCTTCACGACGCTGATGCGCAAGCAGGGCATGAAACCGTACA
TCCTCAAATACATTCGCGGCGACGCCAAGACGGAGACGATGGACATC
TACACGCGGGTCGACCGCGACGAAGCCAAAGAGGAATACCTCAACTG
CATTAAGGAGATCGGGCTG**TGA**CGAGGATAGACAGTTGGTGGATCCG
CTAGCACTCG-3’

In order to construct an overexpression strain, for the codon optimised sequence above, pTA1992 was digested with PciI (compatible with NcoI) and BamHI. The codon optimised XerCD sequence was ligated into pTA1992, to place the XerCD coding sequence under the P.Syn promoter. Cloning was confirmed by a diagnostic digest and Sanger sequencing using primers:

pBSF

pBSR

The final construct is pTA2868 (Figure 5.8).

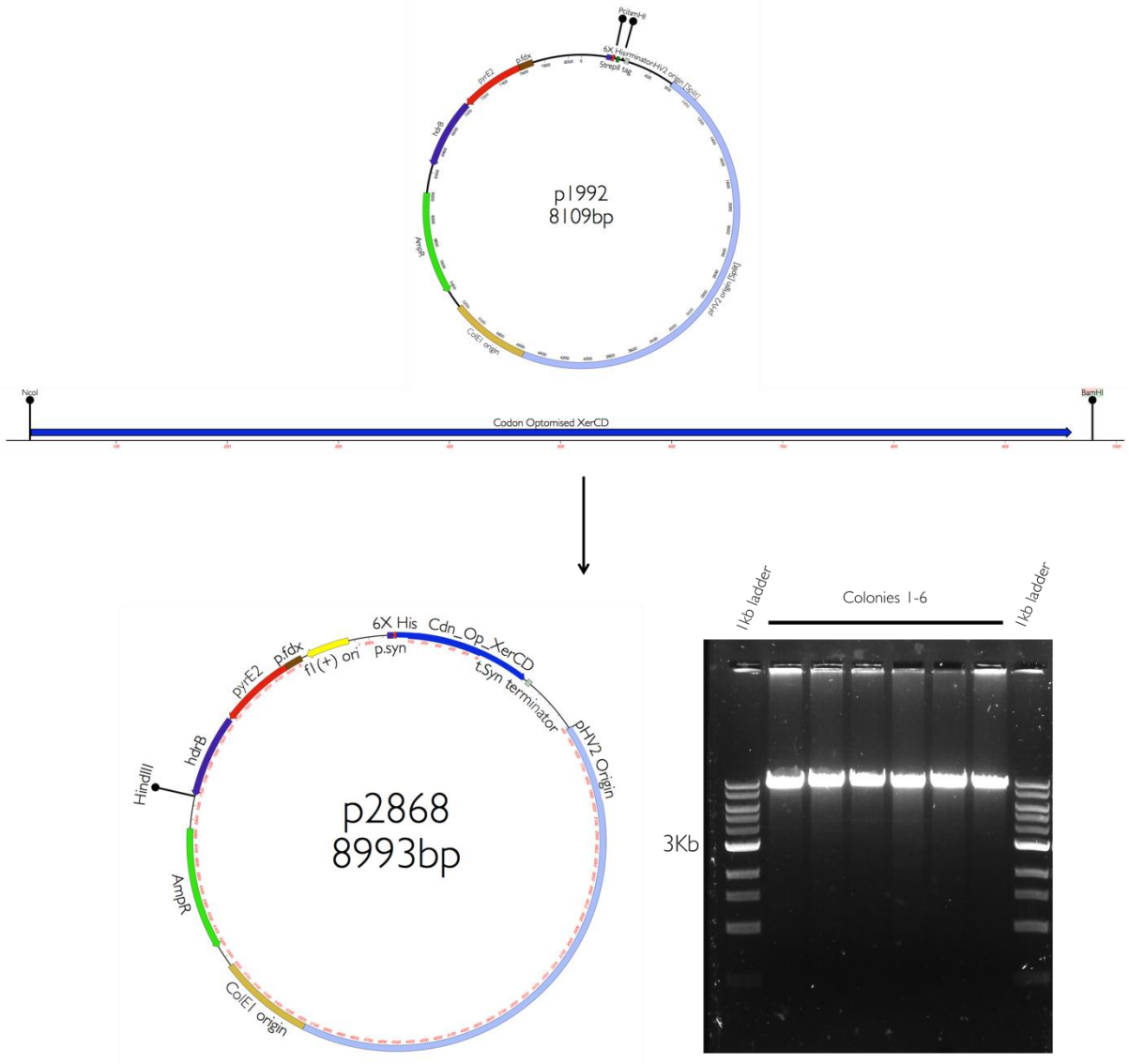


Figure 5.8 |**Cloning strategy for generation of pTA2868.** pTA1992 was digested, along with the Xer_CdnOp_v3 gBlock and subsequent ligation generates pTA2868. Cloning was confirmed by a restriction digest with HindIII and EcoRI. The EcoRI site in pTA1992 is destroyed during cloning, therefore only a single band should be produced. Construct was sequenced.

pTA2868 was then used to transform H1206 (*ΔpyrE2 Δmrr*), which was then checked using colony PCR with the following primers:

p.synF

t.synR

The resultant strain was given the number H5655 and was used in a standard UV DNA damaging assay (figure 5.9).

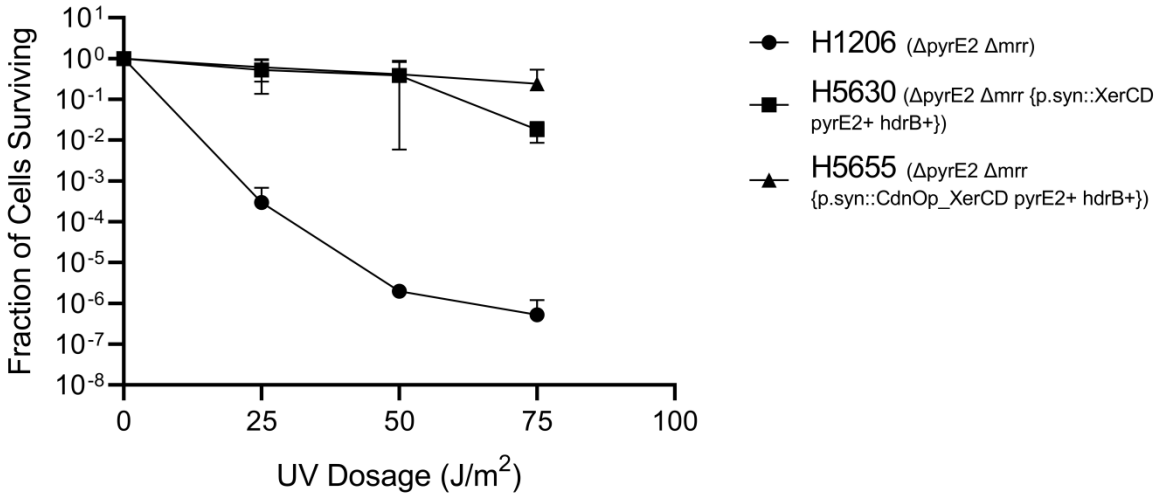


Figure 5.9 |**DNA damaging assay with codon optimised XerCD**. UV survival assay with codon optimized XerCD (H5655) compared to non-optimised coding sequence (H5655) and a wild-type control (H1206).

Results show that the codon optimised sequence performs not dissimilarly to the standard coding sequence, indicating that transcription under p.syn is the rate limiting step and that translation efficiency does not provide increased radiation resistance. However, there is an increased survival fraction seen in strains with the codon optimised sequence at 75J/m^2, although this is not different enough from the non-optimised sequence to warrant further investigation, the levels are noteworthy compared to that seen in wild-type cells (H1206).

5.4.3 Real-time qPCR to detect phage activity during DNA damage response

Firstly, a pilot PCR was carried out using primers for both the prophage region and the endogenous Hel308 control.

The primers used were the same as for the main qPCR assay, as were the cycling conditions (see below). The results are seen in figure 5.10.

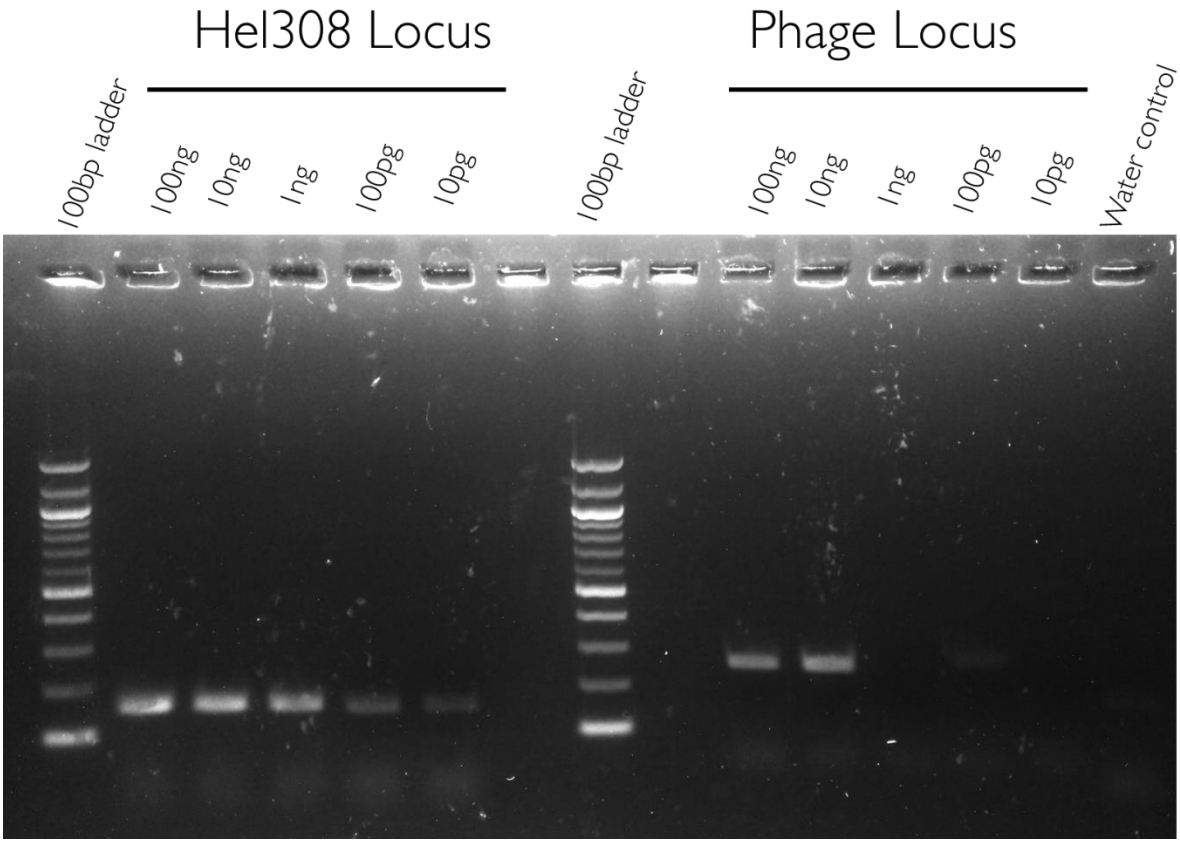


Figure 5.10 |**Pilot PCR of phage and hel308 loci.** PCR reactions were carried out for both the Hel308 endogenous control and the prophage loci using their respective primers. Products were resolved on a 1% TAE gel. PCR confirmed to work, with relative efficiencies not too dissimilar. Range of detection also clarified (10pg for Hel308 and 10ng for the phage locus) Relative amplification efficiencies are slightly different, with Hel308 being amplified more efficiently than phage. However, the bands are not so dissimilar that this would warrant stopping the experiment.

Individual cultures of H26 were grown overnight at 45°C, rotating at 8rpm. Each strain acted as a biological duplicate. Strains were diluted the following day to reach an A_{650} of 0.4 the following day. Once strains had reached this O.D., a 1ml aliquot was taken from each before addition of 0.05µg/ml of MMC. Strains were left to rotate at 45°C and aliquots of 1ml were taken from each sample at the 30 minute, 1 hour, 2 hour and 4 hour timepoints after MMC treatment. This was done as it is unknown exactly how long it takes phage to mobilise, if indeed this is happening at all.

Each 1ml sample was spun down at 3300 x *g* for eight minutes in a 14ml round bottomed tube. The samples were then treated in one of two ways, in order to either carry out qPCR on the pellet or the supernatant.

Preparation of the pellet

Each 1ml sample was spun as above and resuspended in 200µl ST buffer, followed by an equal volume of lysis solution. The tube was inverted to mix and the lysatewas overlaid with 1ml of 100% ice cold ethanol. The DNA was then spooled at the interface, washed twice with 100% ethanol and air dried. Finally the DNA was resuspended in 500µl TE and a subsequent precipitation step was carried out via addition of 50µl 3M NaAc(pH 5.2) and 400µl isopropanol. After a spin at max rpm for five minutes, 1ml of 70% ethanol was added prior to an additional spin step as before. The DNA was resuspended in 100µl TE.

Preparation of the supernatant

Each 1ml sample was spun as above and the supernatant was removed and placed into a fresh 14ml round bottomed tube. To each sample, an equal volume (1ml) of lysis solution was added and mixed, before the overlay of 4ml of ice cold 100% ethanol. The contents were then mixed and placed overnight at -20°C to further precipitate. Each tube was then spun at 20,000 x g in a Sorval SS-34 rotor and the pellet was then washed with 2ml of 70% ethanol and spun again. The supernatant was then removed, the pellets were left to dry before being resuspended in 500µl of TE buffer. Pellets were heated for five minutes at 37°C to increase the ease at which they went back into solution.

DNA from either the pellet or the supernatant was then diluted (10^{-1} – 10^{-4}) in sterile distilled water before being used to set up qPCR reactions as follows:

DNA……5µl

2X Sybr Green Master Mix (ThermoFisher)……..12.5µl

Forward Primer ……..1µl

Phage (PhageProbeF3)

Hel308 (Hel308Fint)

Reverse Primer……1µl

Phage (PhageProbeBII)

Hel408 (Hel308R)

SDW……5.5µl

<u>Cycling conditions were as follows:</u>

Initial denaturation……95°C, 10 minutes

Denaturation…….98°C, 10 seconds

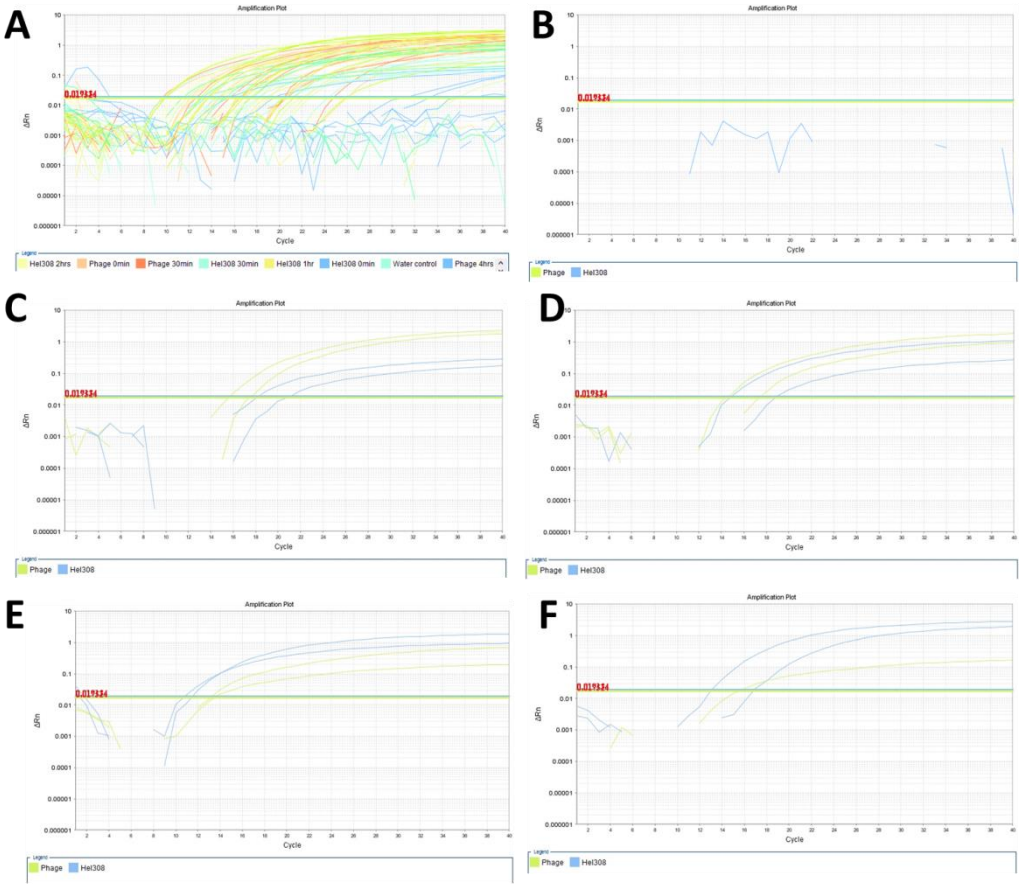
Annealing……..62°C, 30 seconds

Extension ……72°C, 10 seconds

Final Extension……72°C, 10 minutes

40 Cycles

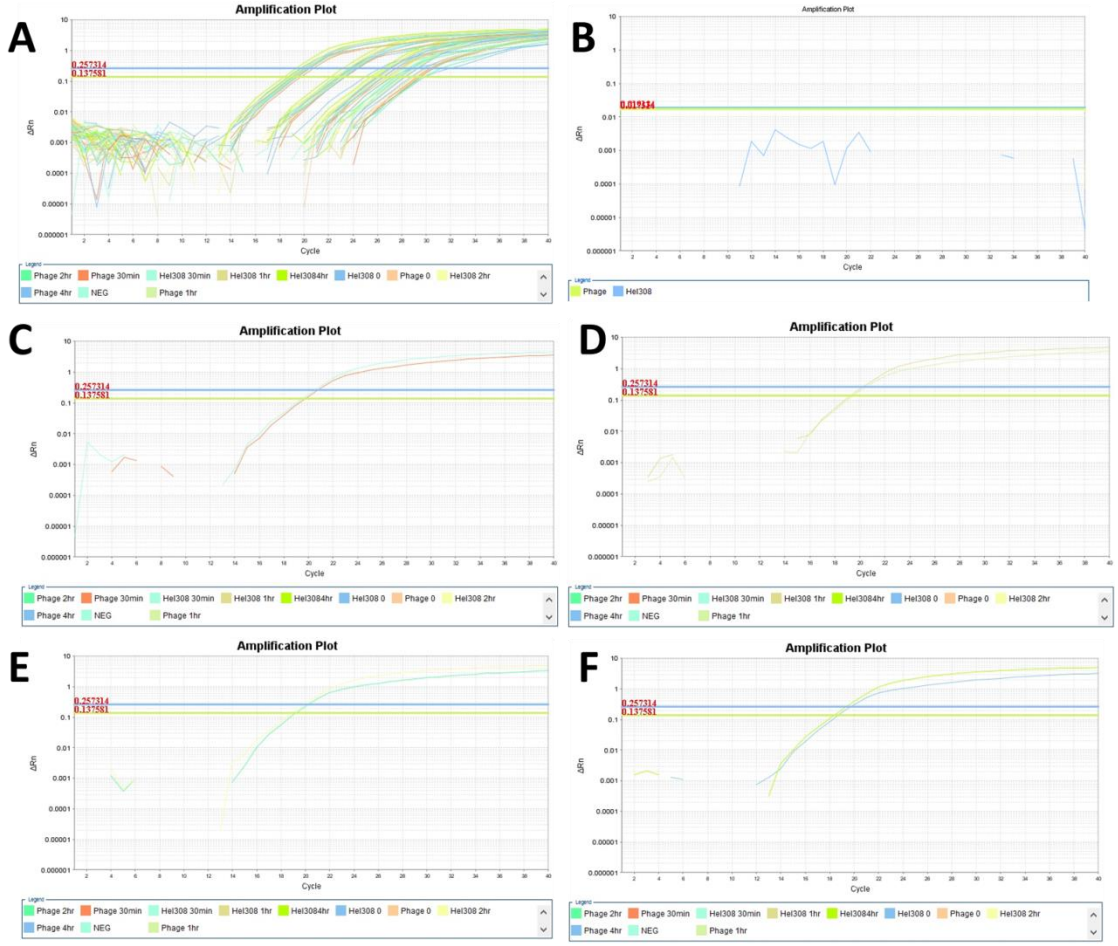
Both the results for the pellet and the supernatant can be seen below, along with the calculation of fold-change in copy number. Data used was averaged across two independent experimental runs for both the Hel308 endogenous control and the prophage region (Figure 5.11 and 5.12).



G

$2^{-\Delta\Delta Ct}$ (fold change)	
0hrs	1
30min	0.000013689
1hr	0.000003577
2hrs	0.000000394
4hrs	0.000001313

Figure 5.11 | **Amplification data from comparative qPCR of prophage region and endogenous control of cellular pellet.** A) Overall amplification data. B) Water control amplification curve. C) 30 minutes after MMC treatment. D) 1 hour after MMC treatment. E) 2 hours after MMC treatment. F) 4 hours after MMC treatment. G) Fold-change in copy number of prophage region compared to endogenous control, Hel308 locus.



G

$2^{-\Delta\Delta Ct}$ (fold change)	
0hrs	1
30min	0.88824
1hr	0.940746
2hrs	0.85204
4hrs	0.820644

Figure 5.12 | **Amplification data from comparative qPCR of prophage region and endogenous control of supernatant.** A) Overall amplification data. B) Water control amplification curve. C) 30 minutes after MMC treatment. D) 1 hour after MMC treatment. E) 2 hours after MMC treatment. F) 4 hours after MMC treatment. G) Fold-change in copy number of prophage region compared to endogenous control, Hel308 locus.

5.4.4 RNA Sequencing of MMC treated cultures

Total RNA was extracted from four biological replicates of H26 (Δ*pyrE2*) and H1192 (Δ*pyrE2* Δ*phage*) as detailed in the methods section (chapter 2). However, samples were also treated with 0.05µg/ml MMC and left for two hours shaking at 45°C before subsequent extraction. A non-treated control was used for all four samples of each strain.

Samples were then sent to DeepSeq, University of Nottingham for processing.

The Δ*phage* strain usage in this assay will shed light, and will allow the difference between death from MMC treatment alone, and death from a combination of phage activation and MMC treatment to be seen in terms of transcriptomic data.

Furthermore, we will be able to see if phage activation causes an increase in transcription of phage activation proteins in the host genome.

Selected genes are shown below in figure 5.13 and full data is contained within appendix I.

Protein	HVO number	H26 (Δ*pyrE2*)	H26 (Δ*pyrE2*) + MMC
RPA family protein	HVO_0291		0.91
ABC transporter ATP-binding protein	HVO_B0125		1.27
Na+/H+ antiporter NhaC family protein	HVO_1394	0.602	
Inositol monophosphatase	HVO_A0274		0.53

Protein	HVO number	H1192 (ΔpyrE2, Δphage)	H1192 (ΔpyrE2, Δphage) + MMC
Carbohydrate ABC transporter permease	HVO_B0294	1.3	
Cell division protein SepF	HVO_0392	1.29	
ABC transporter permease	HVO_B0186	0.98	
	HVO_2084	0.98	
ScpA family protein	HVO_0690	0.95	
Universal stress protein	HVO_3009		0.43
	HVO_2788		0.5

These genes represent part of the prophage region. Therefore inclusion in the RNA Seq data confirms they are absent from strain H1192. Inclusion in this table does not represent a transcriptional response for these genes.

Protein	HVO number	H26 (Δ*pyrE2*)	H1192 (ΔpyrE2, Δphage)
Winged helix-turn-helix transcriptional regulator	HVO_2263	15.4	
Type I restriction-modification system subunit M	HVO_2270	14.5	
Type I restriction-modification enzyme R subunit C-terminal domain-containing protein	HVO_2269	14.3	
Restriction endonuclease subunit S	HVO_2271	14.1	
TATA-box binding protein	HVO_2268	12.6	
Tyrosine-type recombinase/integrase	HVO_2273	12.0	
	HVO_2259	11.9	
	HVO_2290	11.8	
Endonuclease III	HVO_0878		0.33
DNA polymerase sliding clamp	HVO_0175		0.35

Protein	HVO number	H26 (Δ*pyrE2*) + MMC	H1192 (ΔpyrE2, Δphage) + MMC
Winged helix-turn-helix transcriptional regulator	HVO_2263	15.6	
Type I restriction-modification system subunit M	HVO_2270	14.5	
Type I restriction-modification enzyme R subunit C-terminal domain-containing protein	HVO_2269	14.4	
Restriction endonuclease subunit S	HVO_2271	14.3	
Tyrosine-type recombinase/integrase	HVO_2273	12.4	
	HVO_2259	12.8	
	HVO_2290	12.1	
Endonuclease III	HVO_0878		0.31
Flap endonuclease I	HVO_2873		0.34

Figure 5.13| **RNA Seq data of manually selected differentially expressed genes (DEGs).** Genes are shown in separate tables. Each table contains DEGs for that specific comparison. Upregulated genes are shown for each comparison.

RPA family protein

Replication protein A, or RPA, is a homologous protein to Single-stranded DNA binding protein in *E. coli*. The function of this protein in *H. volcanii* is to bind to and protect ssDNA from damage by external agents during DNA replication, transcription and repair processes. There are three RPA proteins in *H. volcanii*: RPA1 which is non-essential, along with RPA3, and RPA2 which is essential.

Inositol monophosphatase

Inositol monophosphatases (IMPases) are from a family of Mg^{2+} dependant phosphatases which act on an inositol monophosphate substrate. They dephosphorylate inositol phosphate to generate inositol. In eukaryotes, IMPase plays a vital role in intracellular signalling.

SepF

The archaeal SepF homologue has been demonstrated to play a key role in cell division in *H. volcanii*. SepF co-localizes with the known cell division proteins FtsZ1 and FtsZ2, homologoues of the bacterial FtsZ protein. This occurs at the midcell. Deletions of *sepF* have not been possible thus far, demonstrating its role may well be critical to the ability of cells to divide (Nußbaum *et al.*, 2021)

ScpA family protein

ScpA, along with ScpB are proteins found within archaeal condensing as partner proteins to the SMC subunit. Little is known about the physiological roles of the archaeal SMC family proteins. (Bell, 2022)

Tyrosine-type recombinase/integrase

Tyrosine-like recombinases have been seen already in this thesis, and explored in detail. Homologues of the bacterial XerCD system have been shown in this thesis to contribute to UV and MMC resistance. XerCD acts as a dimer to catalyse decatenation of chromosomal dimers at *dif* sites. This occurs during DNA replication termination to resolve chromosomal dimers that may have arisen from unequal homologous recombination events.

5.5 Discussion

5.5.1 Transcription is the rate-limiting step for XerCD activity

The codon optimised XerCD sequence performs in a similar way to the normal coding sequence when treated with UV radiation. This hints that transcription is the rate-limiting step, and that translation efficiency, if altered at all by the codon optimisation step, does not impact on the levels of UV resistance observed. Clearly, placing a coding sequence such as XerCD under the control of p.syn (when initial levels are presumed to be very low, as the phage region is deemed fairly inactive), increases transcriptomic levels in order to become the rate limiting step.

5.5.2 Real-Time qPCR demonstrates no increase phage activity upon MMC treatment

The results from the qPCR assay do not demonstrate phage replication and a subsequent increase in copy number comparative to the endogenous control region, Hel308. This is the case in both the cell pellet and the supernatant, which were both independently tested in separate qPCR runs, under the same conditions. The increased UV and MMC resistance shown by the prophage deleted strain, H1192, is therefore not explained by phage induction. That said, however, the experiment here does not show that phage is not induced per se, more that there is a lack of phage detected in both cases. This could be due to incomplete lysis of phage particles, if indeed they are packaged, and therefore the method of detection via qPCR does not show phage induction. Assuming phage is induced, and lysis during both SDS/EDTA and heating steps does indeed break apart packaged phage, then there could be another explanation, namely that phage activator proteins within the prophage region are transcribed at a higher level and thus contribute to phage activation, either at this locus, or elsewhere within the genome. If this is the case, then the RNA-Seq data should lend weight to this idea.

The data above does however show a fair difference between the amount of phage DNA detected in the pellet compared to the supernatant. Specifically, there is a

greater fold-change in copy number seen within the pellet, compared to the results in the supernatant (Figure 5.14).

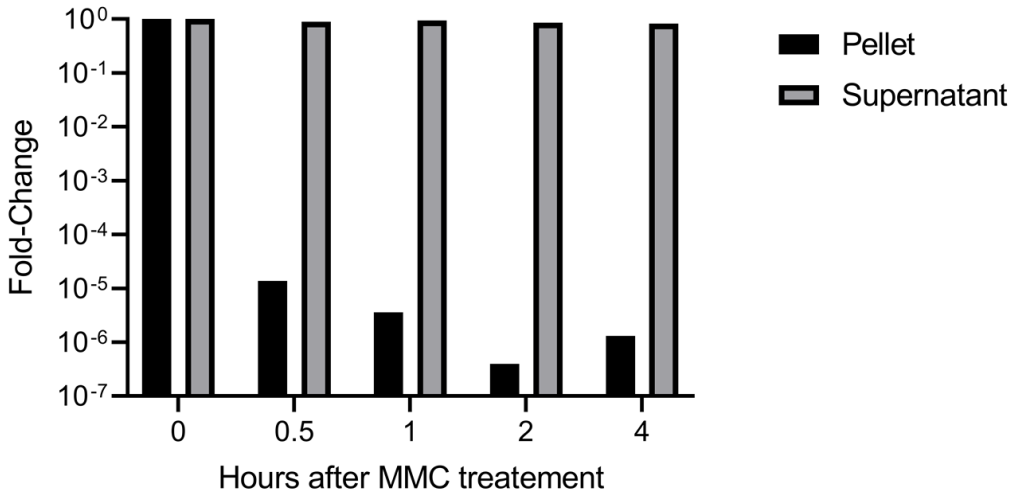


Figure 5.14 | **Fold-change plotted against time after MMC treatement.** Pellet and supernatant values included.

The data is plotted on a Log_{10} scale, in order to show the values for the pellet as they are so small in comparison to the supernatant. This shows that, although the relative amount of phage does not appear to go up, along with the fact that both loci amplified we can say that phage is present in both samples. However, there is far more phage detected in the supernatant, outlining that perhaps even without MMC treatment, phage is somewhat active.

Of note, the pellet:supernatant ratio is 1:1 without MMC treatment. The fold change of phage then drops dramatically after just 30 minutes of MMC treatment. The supernatant on the other hand remains the same. This could be explained due to loss of phage during the ethanol precipitation step. The other explanation is that there is background 'noise' of phage detection and chromosomal DNA in the supernatant that provides the 1:1 ratio seen. If the phage component decreases post MMC treatment, either through ethanol precipitation then this will be lost. Another line of thought is that if the phage is lytic, it may be rendered unsuitable for downstream detection and analysis as a result of MMC treatment.

More analysis is however required in order to determine why the prophage deleted strain is more resistant to genotoxic stress.

5.5.3 RNA Seq analysis

H26 vs. H26 +MMC

RNA sequencing has shown that under MMC induced genotoxic stress, wild-type cells (H26, Δ*pyrE2*) show higher levels of transcription of genes such as RPA, Inositol monophosphetase and ABC transporter ATP binding proteins. RPA acts to stabilise ssDNA formed during DNA replication, repair and transcription. This is perhaps not surprising given levels of transcription are increasing. In addition, RPA may be required to assist with stabilising any ssDNA formation during the repair of inter-strand crosslinks formed by MMC.

Inositol monophosphetase has a speculative role in repair of MMC induced DNA damage but inisitol signalling pathways have been linked to stress responses in various organisms as well as DNA repair pathways such as the ATM and ATR pathways in eukaryotic cells. Exact details are yet to be determined.

ABC transporter proteins are a class of membrane protein found in all domains of life. They play a crucial role in transporting a wide variety of molecules across cell membranes using the free energy derived from ATP hydrolysis. It is not surprising to see such a protein come from RNA sequencing data under MMC stress, as this protein may allow transport of chemicals such as MMC from the cell. Active transport of MMC from within cells will reduce the intracellular concentration of mitomycin C, making it less effective at introducing DNA crosslinks (Lee *et al.*, 2007).

RPA3 represents a key gene upregulated after MMC treatment. This represents one of the RPA homologues found in *H. volcanii*, alongside RPA2 and RPA 1. RPA2 is essential whereas the other two are not. RPA is required for efficient formation of a presynaptic complex during interstrand crosslink repair, demonstrating a function in recombinational processes *in vivo*. Furthermore, RPA is needed for the excision step during NER via activation of the XPF-ERCC1 nuclease (Abdullah et al, 2017) and may play a role in damage recognition (Jang et al, 2022). A recent paper has shown that depletion of RPA induces the activation of the Fanconi anemia (FA) repair pathway required for ICL repair. Furthermore, that RPA inhibition by an RPA inhibitor (HAMNO) significantly impairs maintenance and survival of FA-deficient cells. In summary, this data indicates that RPA depletion resulted in DNA damage signalling and activates the FA pathway of ICL repair.

Of note, no XerCD homologues were shown in the significant hits when MMC was added. This was unexpected and does not fit in to the current theory that MMC induces XerCD, and thus resistance to MMC. This also contradicts previous RNA seq data (Cagla Tosun MRes Thesis, 2019) which shows a XerCD-like integrase as three-fold more highly expressed compared to a control strain with no MMC treatment.

H1192 vs. H1192+MMC

RNA sequencing has shown that under MMC induced genotoxic stress, phage deleted strains show an increase in expression of RNA coding for universal stress proteins, USPs. Such proteins have been implicated in the response to DNA damaging agents in Archaea. Of note, without MMC treatment cells show an increase in transcription of cell division genes *SepF* and *ScpA*, showing that once damaged, cells prioritise repair before carrying out cell division.

H26 vs. H1192

Without MMC treatment, it is important to note that wild-type cells show increased levels of XerCD RNA. This confirms that the phage deleted strain does not contain this locus, but also shows that XerCD is transcribed prior to any genotoxic insult in H26. This is consistent with RNA sequencing data (appendix I figure 2) showing transcription of the prophage region in untreated cells. Although it has been considered that the phage region is relatively quiescent, perhaps at the protein level this remains correct if not at the transcriptional level. We are unable to compare data to the rest of the genome, so it is not possible to confirm this. Upregulation of transcripts for Endonuclease III in H1192 is interesting as deletion has been shown to decrease survival following UV irradiation in *Sulfolobales acidocaldarius* (van Wolferen, Ma and Albers 2015).

H26+MMC vs. H1192+MMC

When both strains are compared with MMC treatment, it is notable that transcription of XerCD homologues are increased in wild-type cells more than when not treated with MMC. Although not a large difference, this does go some way to showing that XerCD may be induced via MMC. However, it could also be argued due to the small effect size, that MMC perhaps causes other cellular responses which eventually then trigger XerCD transcriptional or other regulatory responses. This is consistent with the notion that XerCD is simply a 'backup' pathway that assists with repair of damaged DNA by utilising its nuclease activity.

There is a transcriptional response of DNA repair genes in response to MMC

It is worth mentioning that there is a lack of expected transcripts shown from the RNA seq data under MMC stress. Given the need for interstrand crosslink repair, and the potential for ssDNA and dsDNA breaks, loci such as UvrA, UvrB, UvrC as well as Holliday junction resolvases such as Hef and Hjc were expected to come out of this dataset as being induced under MMC treatment. No such increases were apparent, perhaps showing that a longer induction time was required before RNA extraction, or a higher concentration of MMC was required. This is especially confusing given that ICL repair uses so many components of different DNA repair pathways. However, RPA associated protein 3, *rpa3ap* (HVO_0291) has been shown to be upregulated when wild-type cells are treated with MMC. This is interesting as *H. volcanii* contains three RPA homologues: RPA2 (acts with replisome), RPA1/3 (non-essential). Δ*rpa3* and Δ*rpa3ap* strains are hypersensitive to DNA damaging agents, which can be seen in figure 5.15 (Stroud et al, 2012).

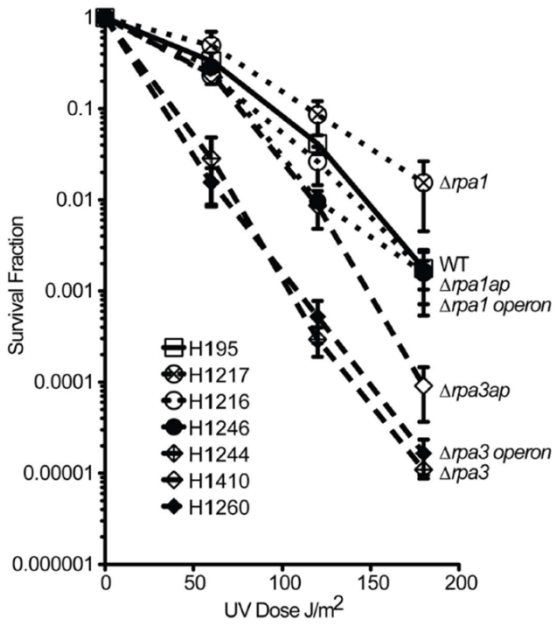


Figure 5.15| **Survival curve with various deletions of *rpa* genes.** Of interest in this thesis is the deletion of *rpa3* and *rpa3ap* showing hypersensitivity to UV irradiation. Figure taken from Stroud et al, 2012.

Genetic analysis has shown that RPA3 and RPAP3 function in the same DNA repair pathway(s), and function together as a specific RPA:RPAP complex. Pull down data suggests that these proteins interact directly with each other in a specific manner (Stroud et al, 2012).

This data, along with the RNA seq analysis indicates the importance of RPA and its associated proteins in the DNA damage response to MMC. This may indicate that lots of ssDNA is produced in response to MMC stress, through the various DNA damaging responses activated.

Chapter 6: Genetic and Biochemical analysis of Sphere formation in *Haloferax volcanii*

6.1 Background

Compared to their bacterial counterparts, Archaea do not appear to produce anything resembling endospores (spores), unlike bacterial species such as *Bacillus subtilis* (Nicholson *et al.*, 2000, Errington, 2003). Sporulation represents an important genetic and biochemically programmed event that provides resistance to simulated extra-terrestrial conditions, including radiation and desiccation stress (Onyenwoke *et al.*, 2004). In the absence of any known spores in Archaeal species, it is quick and convenient to assume that any spherical cells in Archaea represent a form of stasis, similar to that found in bacteria. If this is not the case it is also possible to assume they could represent more eukaryotic-like mechanisms of quiescence, such as a cell in G0 phase of the eukaryotic cell cycle. A study in 2012 has indicated formation of such spherical particles in haloarchaeal species in response to desiccation stress, specifically low water activity induced via resuspension in a lithium salt buffer. Immediate formation of spheres were reported, however, of note, *Haloferax volcanii* was not used in their system (Norton and Grant, 1988, Fendrihan *et al.*, 2012). Interestingly, ATP levels were shown to be reduced in these spherical cells, up to 50-fold lower than in non-spherical cells. Taken with the fact that dormant bacterial cells show similar biochemical changes, sphere formation potentially could be analogous to bacterial spores, or at least, a form of cellular dormancy that could assist in protection from harsh and unfavourable conditions. Furthermore, data has shown that, alongside ATP reduction, various proteins are also reduced in amount, specifically these include an S-layer protein. Rods of *Halobacterium salinarum* were also shown to produce up to 3-4 spheres, each of which can re-form a single rod shaped cell (Fendrihan *et al.*, 2012). Not much is known yet about sphere formation in *Haloferax volcanii*, or if they truly are dormant states.

6.2 Aims and Objectives

It will be of interest to look at potential sphere formation in *Haloferax volcanii* owing to the ease and amount of genetic tools available in this model organism. These tools can be used to uncover genetic changes that occur during sphere formation which will allow a deeper look into the potential for spheres to be dormant cellular states. Properties of spheres in *H. volcanii* will be interesting to determine, specifically if they provide resistance to any particular stresses, such as radiation stress. Furthermore it will be useful to decipher the exact nature of sphere formation and determine exactly what happens, if anything, to the outer coat of the Archaeal cells, specifically the S-layer, and if the entire S-layer is in fact removed, forming spheroplasts. This would counter the current data indicating low water activity is a trigger for sphere formation. Lastly, it will be useful to try and characterise any additional triggers for sphere formation, and the genetic and biochemical principles underpinning this.

The aims of this chapter were to:

- Determine if *Haloferax volcanii* also forms spheres under the conditions reported by (Fendrihan *et al.*, 2012) using *Haloferax mediterranei* as a positive control.

- Carry out a sphere reversion assay to determine sphere viability compared to wild-type control.

- Carry out a sphere UV resistance assay to test the hypothesis that spheres are more resistance to UV stress than wild-type cells.

- Form spheroplasts via an independent method, using EDTA, and image compared to spheres formed via lowering of water activity. This aims to lend weight, or detract, from the alternative hypothesis that spheres reported under low water activity, simply look that way owing to the LiCl buffer they are resuspended in, which perhaps removes the entire S-layer, forming spheroplasts. This would not indicate factually either way, however would provide supporting evidence along with further analysis.

- Determine if EDTA-mediated S layer removal is enough for a negative staining control with convavalin A.

- Test S-layer staining abilities of Concanavalin A, to determine if S layer is removed during sphere formation

- Carry out a transformation assay with titrations of EDTA, to determine if and how transformation efficiency is altered.

6.3 Materials

6.3.1 Strains

[ ] Indicate integrated plasmid DNA

{ } Indicate episomal plasmid DNA

Table 6.1. **Strains used in this chapter**

Strain	Parent	Genotype
H1206	H1202	Δ*pyrE2*, Δ*mrr*
H826	H824	Δ*pyrE2*

6.3.2 Plasmids

Table 6.2. **Plasmids used in this chapter**

Name	Use	Notes
pTA912	*pyrE2* containing vector used for transformation viability assay	Episomal supply of RadA marked with *pyrE2*.

pTA912 (*pyrE2* marker)

pTA912 is normally used for the episomal supply of *radA* marked with *pyrE2*. It was used as a plasmid for transformation during the EDTA transformation assay, to provide a *pyrE2* marker.

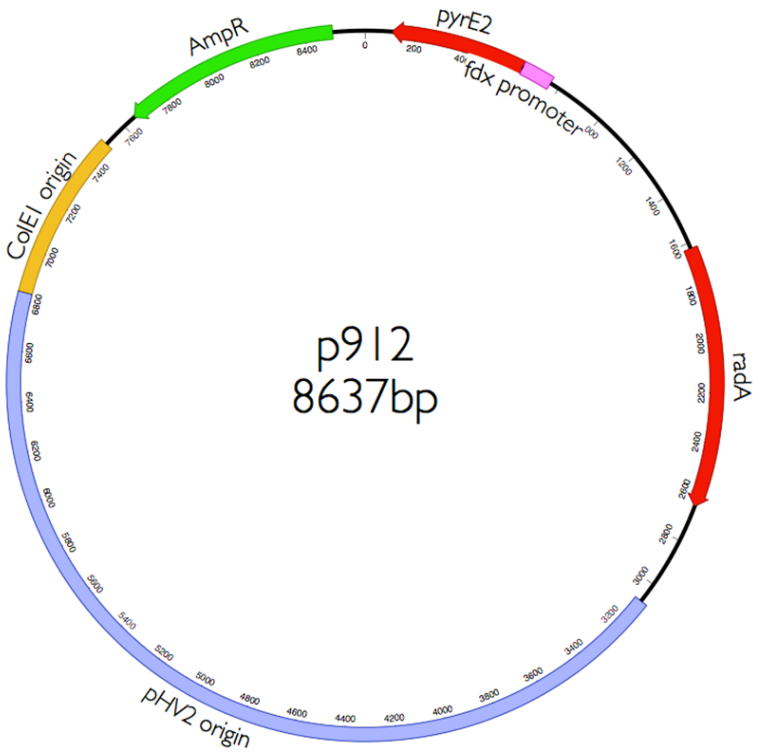


Figure 6.1 | **pTA912.** Constructed by Allers lab.

6.4 Results

The aims of this chapter were to confirm if spheres can be induced under low water activity in *H. volcanii* as this has not been shown before. The ability of spheres to 'revert' back to non-spheres was tested, including their viability. To see if spheres are a more stable state, similar to that of endospores in bacteria, UV resistance assays were carried out. Before any additional research is done, it was prudent to determine if the S-layer was simply being removed as a result of low water activity, forming spheroplasts with a more circular and compact morphology. Staining of the S-layer was attempted using a lectin, Concanavalin A. A suitable negative control was needed, and removing the S-layer via titration with EDTA seemed a good choice. This never yielded a good enough result to warrant further study. Transformation assays were carried out to determine if forming spheres impacted their ability to be transformed with exogenous DNA.

6.4.1 Sphere formation in *Haloferax volcanii*

5ml cultures of *H. volcanii* (H1206) and *H. mediterranei* (H826) were set up in Hv-YPC broth and left to grow overnight at 45°C, rotating at 8rpm. The following day, cultures were diluted and left to grow to an A_{650} of 0.6. Cultures were independently pelleted and resuspended in either fresh Hv-YPC, or 4M LiCl buffer, the latter of which aims to induce sphere formation. Cultures were then left for approximately two hours at 37°C, rotating at 8rpm in order for spheres to form. Cultures were resuspended in fresh Hv-YPC. 1ml of culture was spun at 3300xg for 8 minutes, and resuspended in 18% SW, followed by staining with Acridine Orange, after which cells were imaged under a FITC filter to determine phenotype. *H. volcanii* cells were compared to *H. mediterranei* cells, which are known to form spheres under such conditions, and therefore acted as a positive control for this assay (Fendrihan *et al.*, 2012). Of significant note, *H. volcanii* appeared to form spheres in a similar way to *H. mediterranei*. The morphological difference between non-spheres and spheres was striking (figure 6.2).

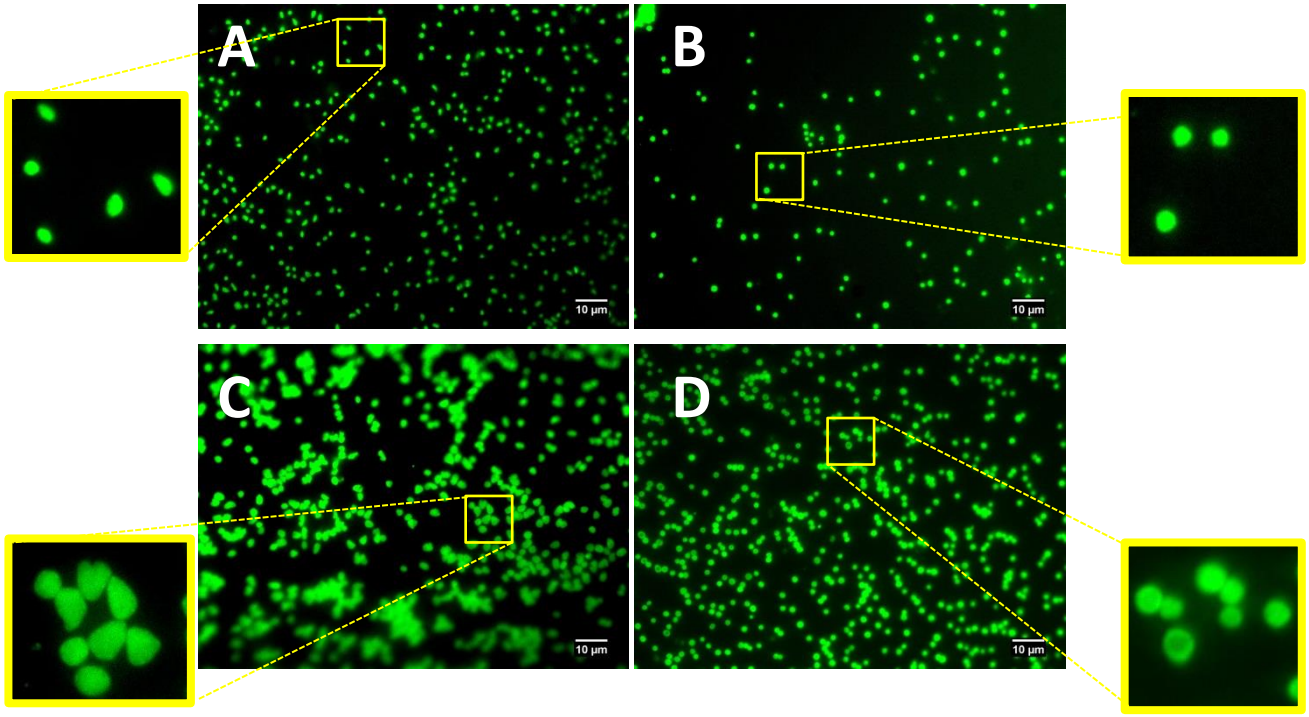


Figure 6.2| **Pilot sphere formation assay in *Haloferax volcanii*. A)** *H. mediterranei* cells incubated in Hv-YPC broth. These are wild-type, non-spherical cells. **B)** *H. mediterranei* cells resuspended in 4M LiCl buffer, cells have formed spheres. **C)** *H. volcanii* cells resuspended in Hv-YPC, thus representing wild-type, non-sphere cells. **D)** *H. volcanii* cells resuspended in 4M LiCl buffer, showing sphere formation with similar morphology to the *H. mediterranei* positive control.

6.4.2 Sphere Reversion and cell viability

Haloferax volcanii strain H1206 was inoculated in a 5ml culture of Hv-YPC overnight and diluted the following day. Culture was allowed to reach an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Cultures were spun and resuspended in either Hv-YPC, or LiCl buffer to form spheres. Cultures were grown for approximately two hours at 37°C to form spheres. Cultures were resuspended in 18% SW and aliquots were imaged via a FITC filter to confirm spheres had indeed formed. The culture was serially diluted and spotted in 20ul aliquots onto YPC agar plates to allow for growth over 4-5 days prior to counting.

A separate culture, treated the same way, was then resuspended in Hv-YPC broth and allowed to regenerate at 45°C for two hours, after which the dilution and plating procedure was carried out as above.

The process was repeated, using the standard transformation regeneration (regen) solution to compare the results.

Survival fractions for both spheres and non-sphere controls under both non-regen and regen conditions was carried out and calculated, results are shown in figure 6.3.

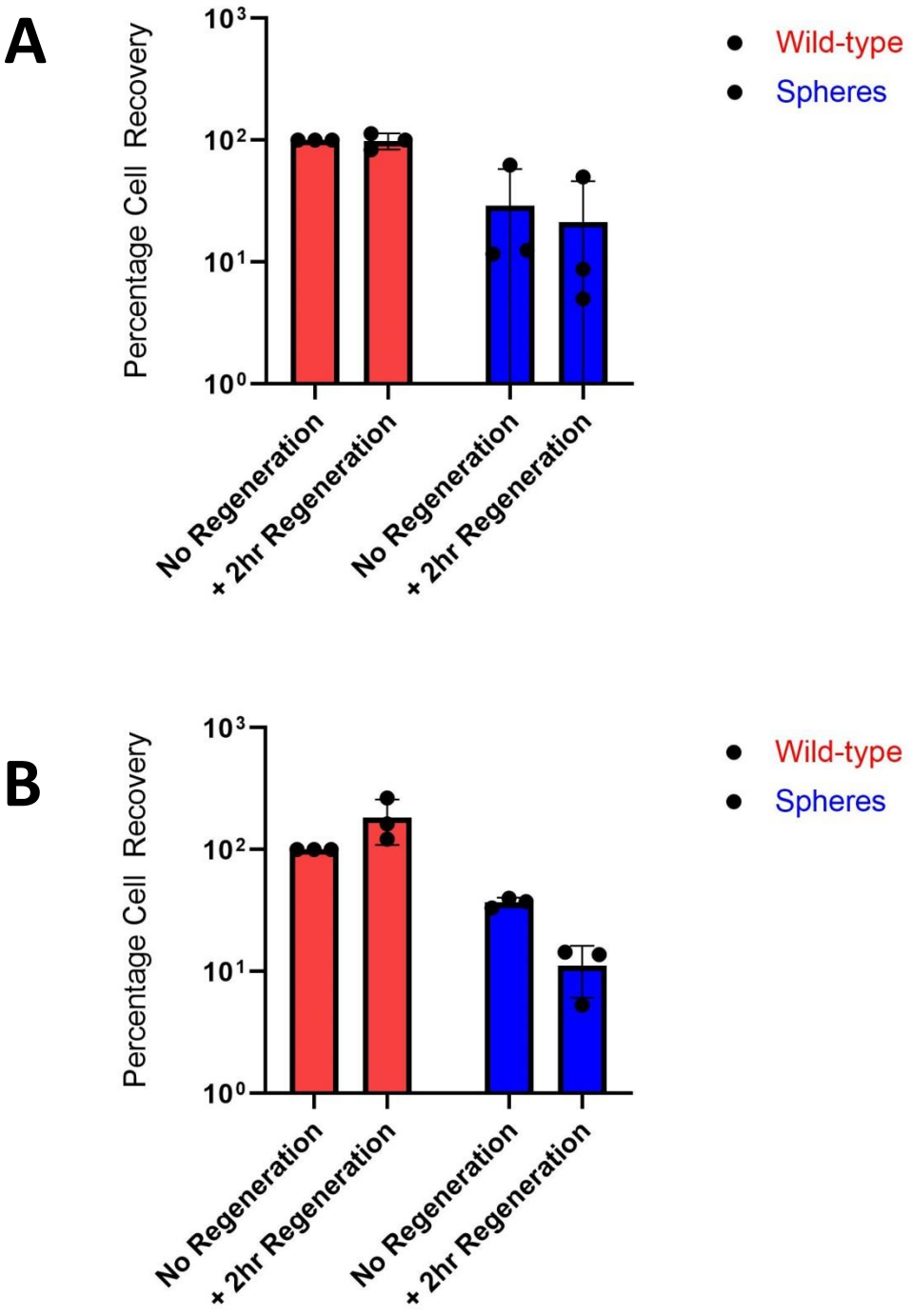


Figure 6.3 | **Sphere viability assay**. Cells were left to form spheres, or left untreated to act as a wild-type control. Both sets of cultures within each group were then allowed to either regenerate for two hours, or plated immediately. 20µl spots were plated at a range of dilutions in order for cell survival fraction to be calculated. Replicates are shown in similar colour schemes (n=3). (A) Regenerated in YPC, (B) regenerated in regeneration solution.

6.4.3 Sphere UV Resistance

5ml cultures of *H. volcanii* H1206 were set up in Hv-YPC broth overnight. Cultures were diluted and left to reach an A_{650} of 0.4, equating to approximately 10^8 cells per ml of culture. Cultures were spun down and resuspended in either YPC broth or LiCl buffer, the latter of which will induce sphere formation which will be checked via fluorescent microscopy after staining with Acridine Orange. Cultures were spun down and resuspended in fresh Hv-YPC, after which serial dilutions were carried out in 18% SW. Aliquots of each dilution were spotted in 20µl volumes onto Hv-YPC agar plates, plates were then treated with varying amounts of UV radiation. Plates were incubated in a black bag for 5 days prior to cell survival fraction being counted. Results are shown in figure 6.4.

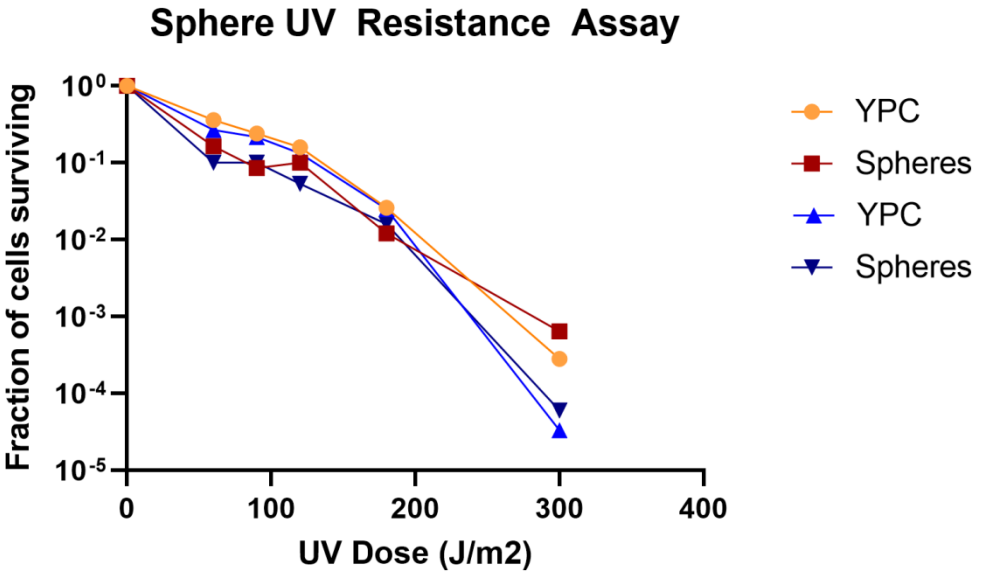


Figure 6.4 | **Sphere UV resistance assay**. Dilutions of either spheres or wild-type cells were plated and subjected to varying amounts of UV radiation doses. Fractions of cells surviving are shown on a logarithmic scale. Replicates shown in similar colour schemes (n=2).

6.4.4 Spheroplast formation in *Haloferax volcanii*

Haloferax volcanii cells (H1206) were grown overnight in 5ml YPC broth. Cells were then diluted to reach an A_{650} of 0.6 the following day. Cells were then pelleted at 3300xg for 8 minutes, prior to being resuspended in 1ml of buffered spheroplasting solution. Cells were spun again as before and resuspended in 200µl buffered spheroplasting solution, after which a 20µl drop of 0.5M EDTA (pH 8.0) was added to the side of the tube, and mixed by inverting. Cells were left for 10 minutes to form spheroplasts, after which cells were stained with Acridine Orange and imaged under a FITC filter (figure 6.5). Cells resemble those formed under LiCl buffer incubation and therefore this opens the door to the possibility of 'sphere' formation being nothing more than formation of spheroplasts, however more data is required, such as visualisation of the S-layer, which will be carried out later.

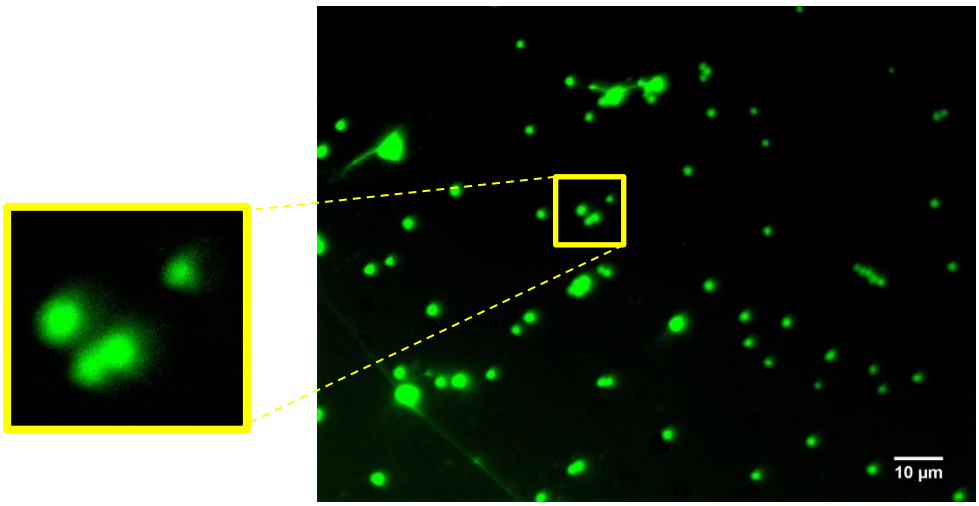


Figure 6.5 | ***Haloferax volcanii* spheroplast microscopy**. *H. volcanii* strain H1206 was allowed to form spheroplasts via addition of EDTA. Staining was carried out using Acridine Orange and imaging was done via a FITC filter on the Zeiss 200M microscope. Clumps of cells and individual cells can be seen.

6.4.5 Pilot S-layer Sphere staining in *H. volcanii*

If staining of the S-layer is to be the method of choice for determination of S-layer removal during sphere formation, then a suitable stain and negative control should be used. Concanavalin A was chosen, due to its ability as a lectin to bind to mannose and glucose residues on the S layer of *H. volcanii* (Gilboa-Garber, Mymon and Oren, 1998). Concanavalin A conjugated to Alexa Fluor 647 (Thermo Fisher) was used.

In order to determine the ability of ConA to stain the S-layer of *H. volcanii*, an initial staining assay was carried out. 5ml of relevant culture (H1206) was grown overnight in Hv-YPC broth and left to grow overnight at 45°C, rotating at 8rpm. The following day, the culture was diluted and left to grow to an A_{650} of 0.6. The culture was then pelleted, 1ml of culture was spun at 3300xg for 8 minutes, and resuspended in 18% SW, followed by staining with Acridine Orange and Concanavalin A, after which cells were imaged under a FITC filter (AO) or Cy5 filter (ConA) to determine phenotype (figure 6.6).

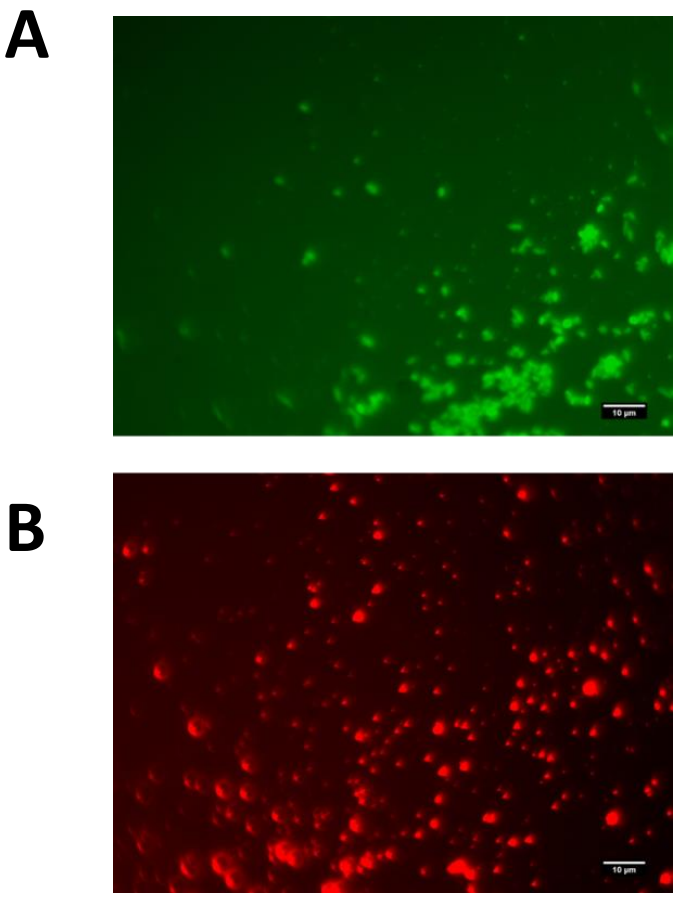
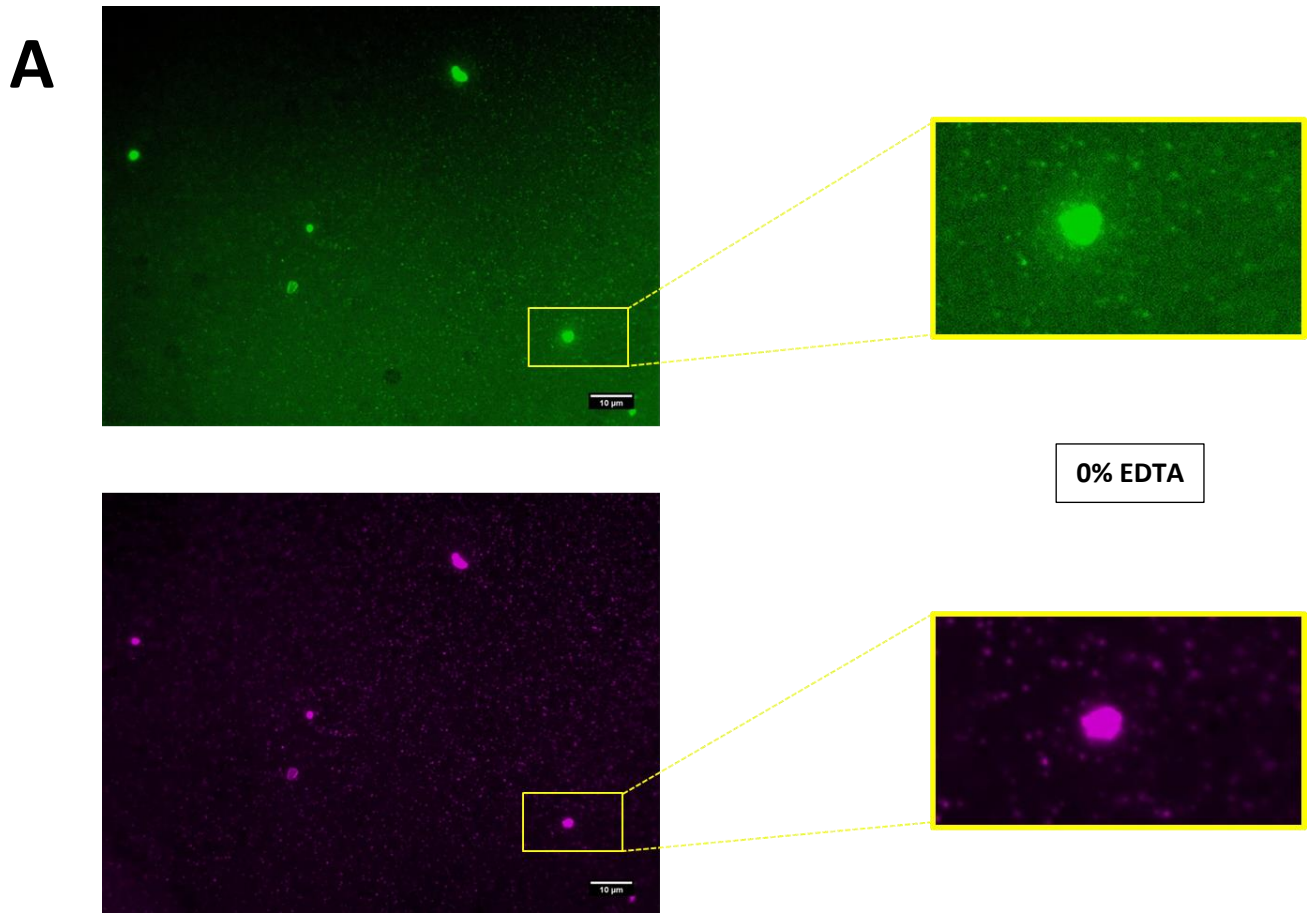


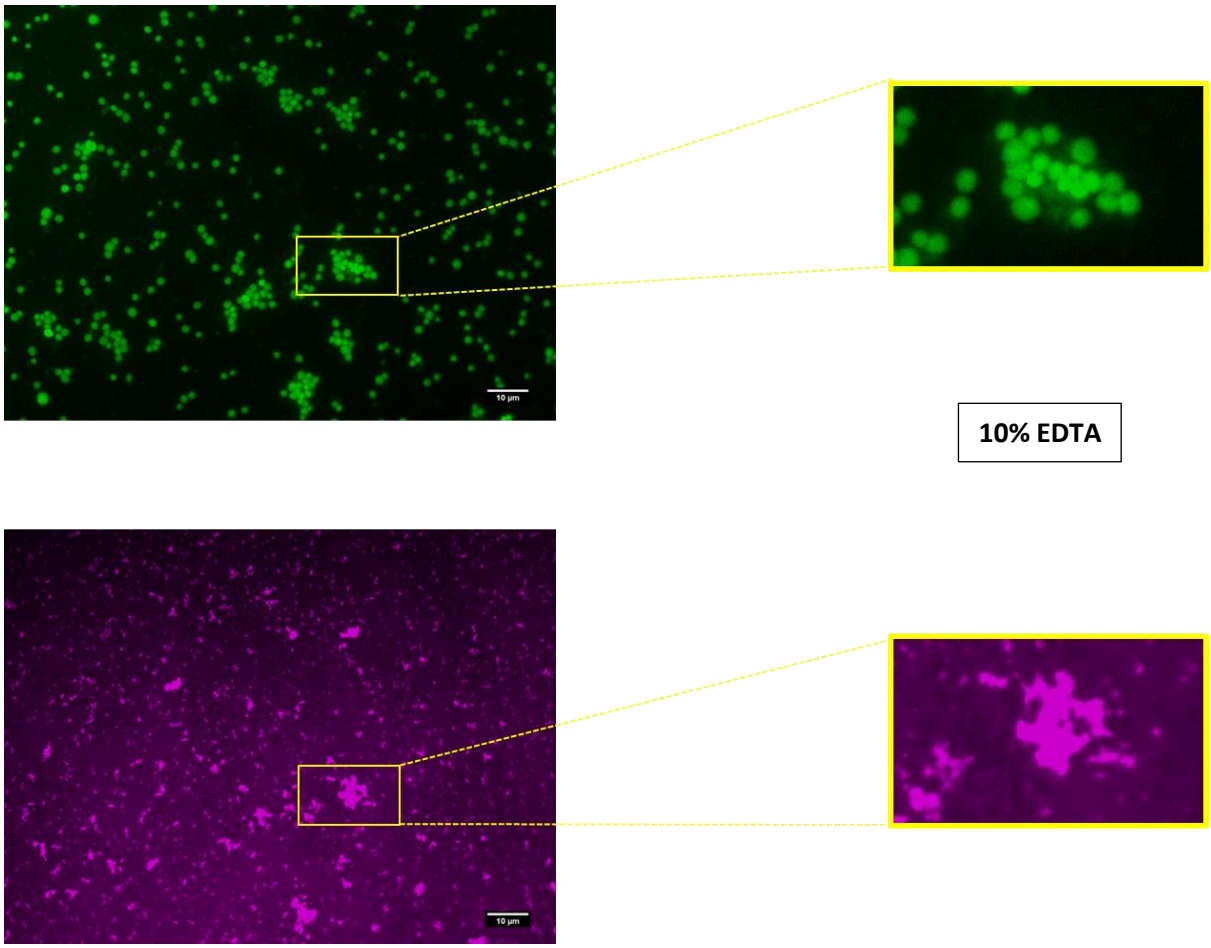
Figure 6.6 | ***Haloferax volcanii*** **ConA microscopy**. *H. volcanii* strain H1206 was spun down and stained with both Acradine Orange (A), and Convavalin A (B) to determine if ConA stains appropriately.

6.4.6 Specificity of S-Layer staining in *H. volcanii*

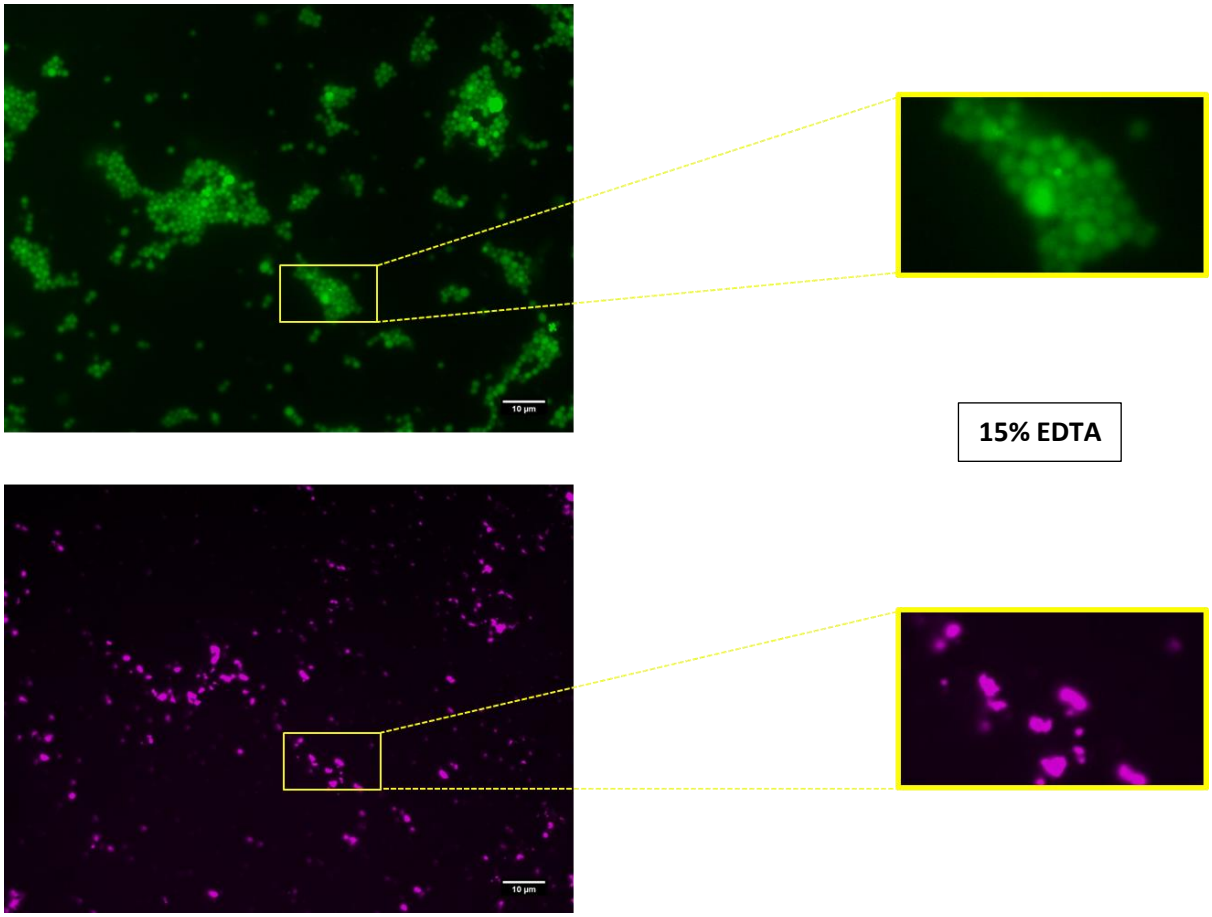
The specificity of ConA binding to glucose and mannose residues on the S-layer of *H. volcanii*, as opposed to other carbohydrates on and within the cell, can be determined via titration of EDTA to remove the S-layer, as it is reported to do so (Cline *et al.*, 1989). 5ml of relevant culture (H1206) was grown overnight in Hv-YPC broth and left to grow overnight at 45°C, rotating at 8rpm. The following day, the culture was diluted and left to grow to an A_{650} of 0.6. The culture was then pelleted and resuspended, as per the transformation protocol, in buffered spheroplasting solution, after which varying concentrations of EDTA were titrated in. After incubation for 5 minutes, 1ml of culture was spun at 3300xg for 8 minutes to pellet the S layer fragments, and resuspended in 18% SW, followed by staining with Acridine Orange and Concanavalin A, after which cells were imaged under a FITC or Cy5 channel, SLIM, University of Nottingham (Figure 6.7).



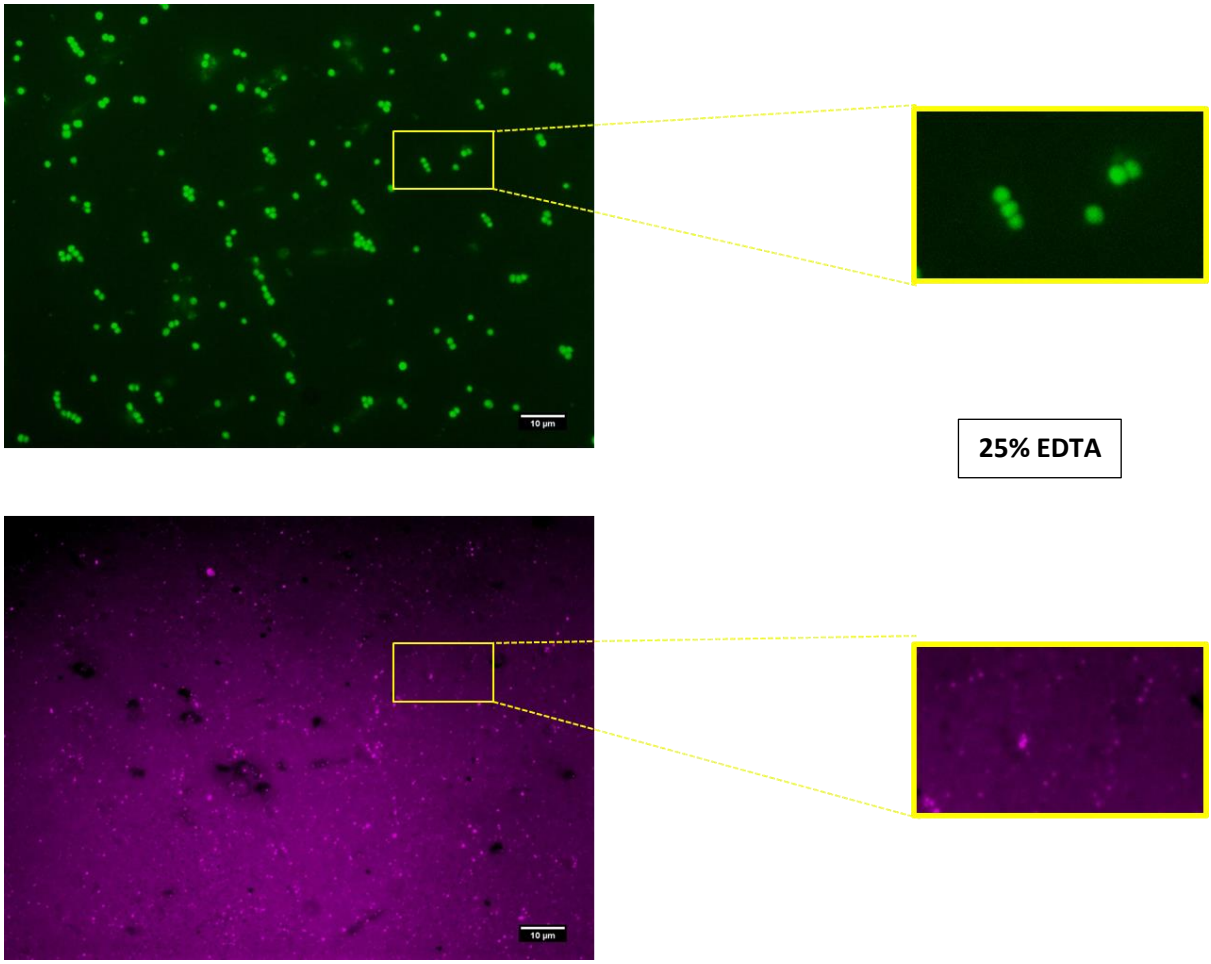
B



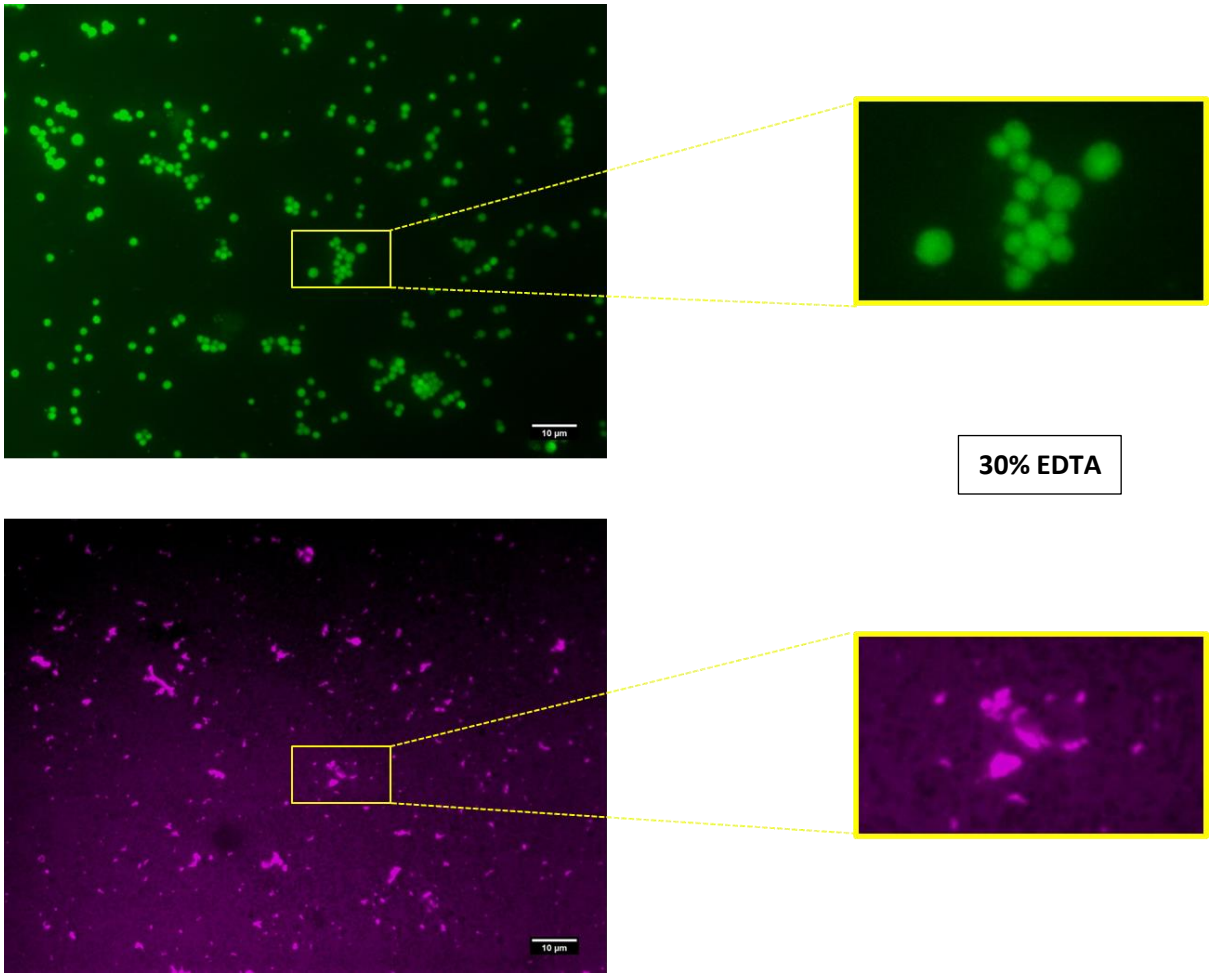
C



D



E



F

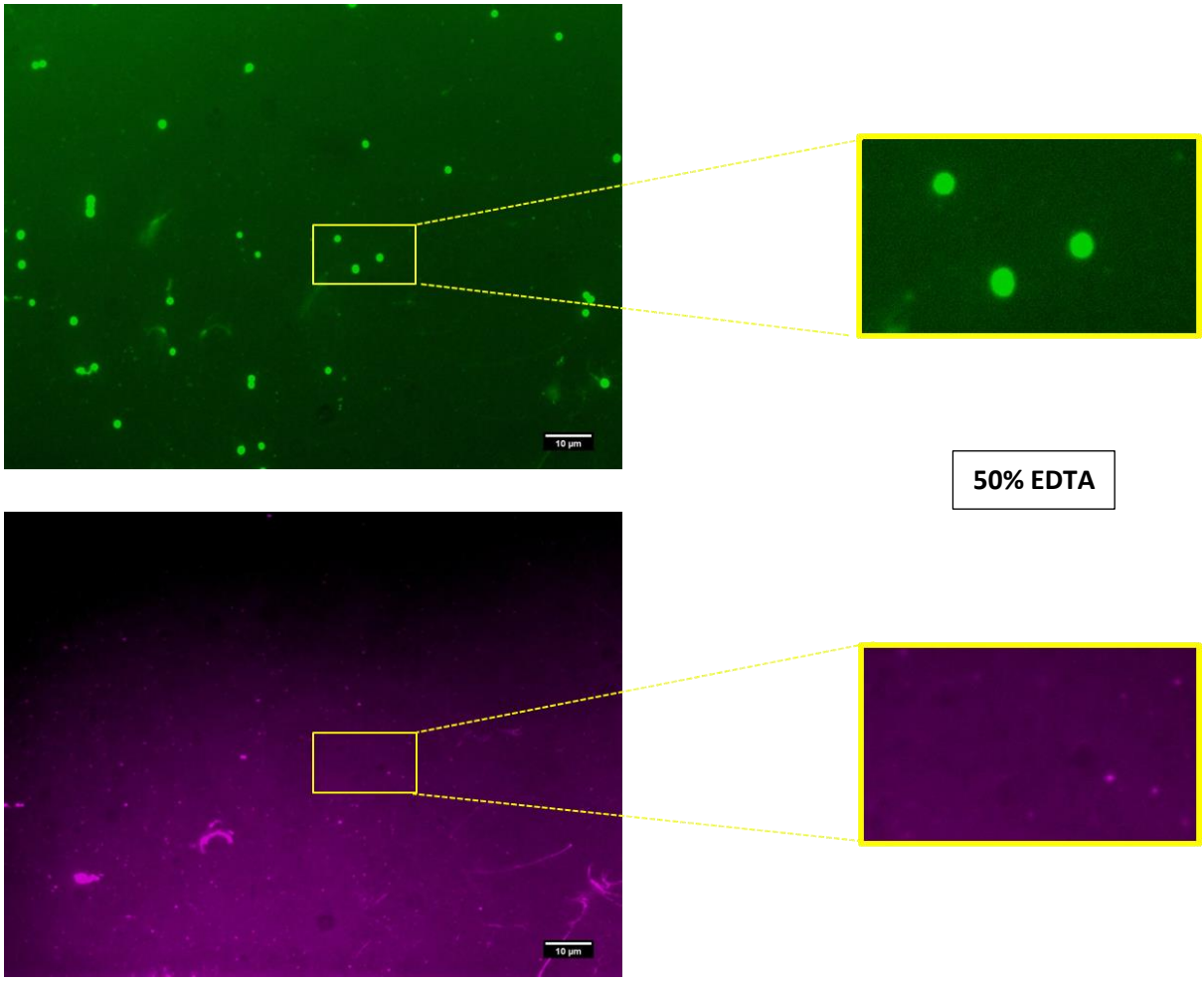


Figure 6.7 | ***Haloferax volcanii* EDTA Staining**. *H. volcanii* strain H1206 was spun down and stained with both Acridine Orange (top), and Concanavalin A (bottom) to determine if ConA stains appropriately. Selectivity of ConA for the S layer was determined by addition of titrated amounts of 0.5M EDTA w/v. 0% EDTA (A), 10% EDTA (B), 15% EDTA (C), 25% EDTA (D), 30% EDTA (E), 50% EDTA (F).

6.4.7 EDTA Viability Assay

The viability of EDTA treated cells in section 7.4.6 was tested alongside the staining assay. Prior to staining, a sample of cells was diluted and spotted onto Hv-YPC plates to quantitatively determine cell survival during the process (figure 6.8).

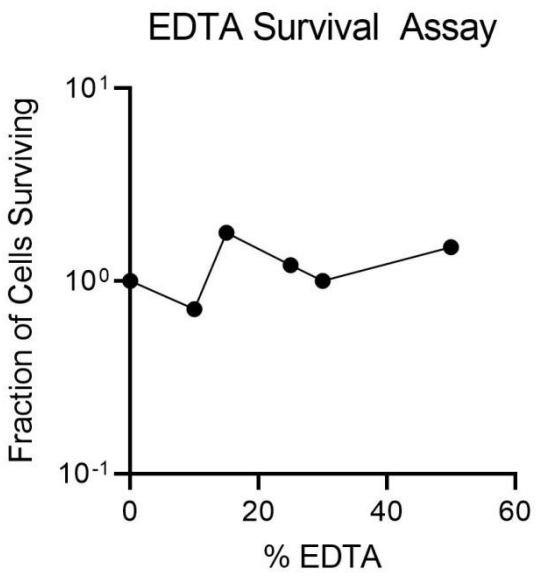


Figure 6.8 | **EDTA survival assay**. *H. volcanii* strains from the staining assay in 7.4.6 were also subjected to serial dilutions and spotting. After incubation for 5 days at 45°C, spots were counted to determine survival. Survival was calculated compared to number of cells in untreated sample (0% EDTA).

6.4.8 Quantification of S-Layer staining in *H. volcanii*

The staining procedure shown in section 7.4.6 allowed for fluorescence quantification, courtesy of the SLIM Zeim 200m microscope used for the assay. The intensity density and mean grey values for the images shown in 7.4.6 was then calculated and used to plot fluorescence of AO and ConA with increasing concentrations of EDTA (figure 6.9).

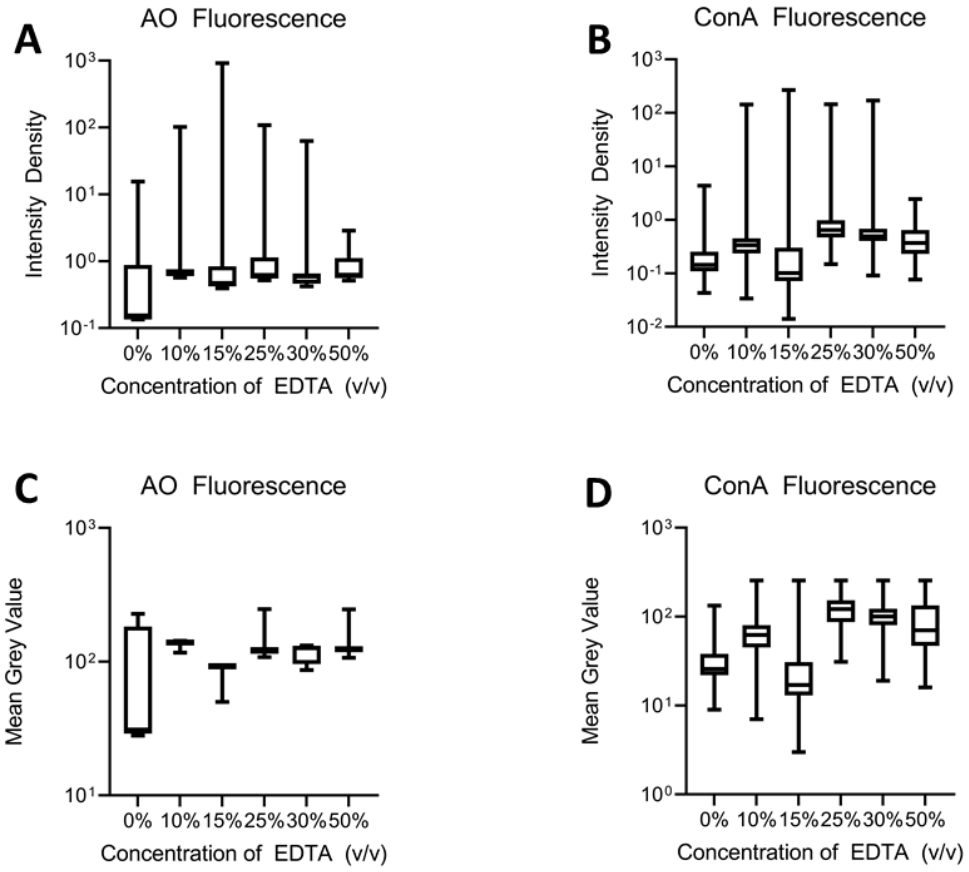


Figure 6.9 | **Quantification of EDTA S layer staining**. *H. volcanii* strains from the staining assay in 7.4.6 were also subjected to stain quantification using both Intensity Density and Mean Grey Values. Analysis was carried out using ImageJ (FIJI).

6.4.9 EDTA-mediated transformation assay

It was of use to determine if indeed the S-layer is being removed during sphere formation, if transformation efficiency was affected. It is known that EDTA removes the S-layer, allowing for PEG mediated introduction of exogenous DNA during the transformation procedure. The transformation process was followed as per the methods section, the only alteration was to remove all EDTA (replacing with water), or titrate increasing amounts of EDTA during spheroplast formation during the transformation protocol. The transformation efficiency was then calculated in each case, and the results are shown in figure 6.10. For the assay, pTA912 was used to transform the host H1206 strain to uracil prototrophy. Spheres of the same strain were also formed using standard protocol, prior to transformation.

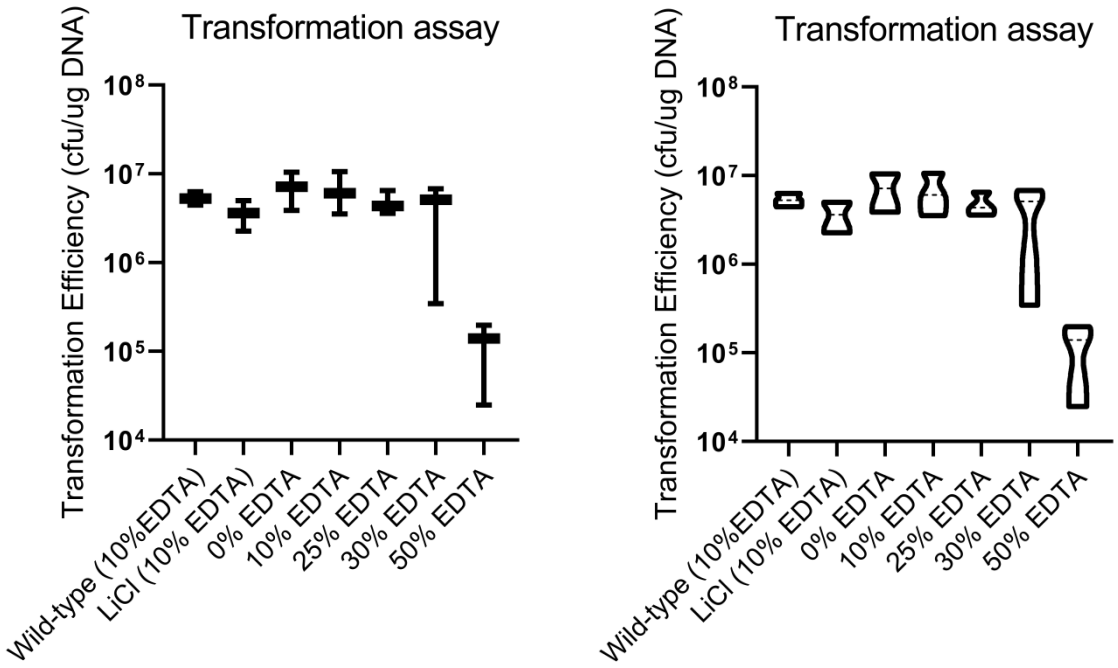


Figure 6.10 | **Transformation Assay with titration of EDTA**. *H. volcanii* strain H1206 (spheres and non-spheres) was transformed using standard procedure, except EDTA was titrated in varying concentrations. The transformation efficiency was plotted as a bar (left) and violin plot (right). N=3

6.5 Discussion

6.5.1 Sphere formation in *Haloferax volcanii*

Spheres were formed under LiCl buffer in order to lower the water activity of solution. *Haloferax mediterranei* was used as a positive control, owing to the previously reported ability of this species to form spheres under low water activity. Interestingly, although *H. mediterranei* cells did form spherical cells after a two hour incubation period in LiCl buffer, the most drastic morphological shift was observed with *Haloferax volcanii* cells, which exhibited a drastic change into spherical cells, from original wild-type cells that display a range of morphologies. It has therefore been shown that *Haloferax volcanii* should be added to the group of haloarchaeal species that can form spheres under low water activity. Of note, wild-type cells do show some degree of spherical morphology in some cells, although after incubation with LiCl buffer, every cell in the field of view has transformed into a much more rounded cell shape. Furthermore, spheres do appear to be mainly uniform in terms of morphology, however their sizes appear to be slightly less homogenous, perhaps indicating a metabolic shift, as well as a morphological change which would support evidence from (Fendrihan *et al.*, 2012). Either way, the ability of *H. volcanii* to form spheres shows this is a mechanism conserved between multiple species of Haloarchaea, potentially the Haloarchaea as a whole. Furthermore, this allows more in –depth analysis to be carried out on spheres in *H. volcanii* such as investigating other potential triggers, which would lend weight to spheres being 'real' biologically programmed events, and not a morphological response to low water activity, or indeed, Li salt.

6.5.2 Sphere reversion and viability

Spheres appear to be significantly disadvantaged when plated after forming under LiCl buffer, with and without a regeneration step. Wild-type cells are far more viable overall. This could be due to a range of factors that will need to be investigated further. The regeneration step appears to decrease the sphere viability significantly, perhaps indicating the spheres have had their S-layer removed and are thus prone to damage and therefore viability decreases. This may add weight to the S-layer theory, whereby, the spheres are a sole consequence of S-layer removal by the LiCl buffer, and not low water activity as has been suggested by (Fendrihan *et al.*, 2012). Secondly, and perhaps less likely, is that spheres have altered biochemical properties (perhaps supported by Fendrihan *et al*., 2012 who report ATP concentration alteration, amongst other things), meaning that most spheres can not recover immediately. This could also lend weight to the idea that they are dormant-like cells.

Recovery in YPC certainly appears to increase the yield of spheres that are recovered, perhaps pointing to an alteration from spheres, back to non-spherical, wild-type, cells, which cope better with the stresses of plating. This might be possible due to the fact that plating stress may, in part, be due to dehydration stress as cells transition from immersion to a much drier surface. Further research, especially into the nature of the S-layer whilst in spherical form would shed more light on this apparent lack of survivability of spheres once plated.

6.5.3 Spheres in *Haloferax volcanii* do not appear to be resistant to UV stress

Although the data here seem to indicate no resistance to UV stress over wild-type cells, an outside possibility here is that UV stress is itself a trigger for inducing sphere formation. This was considered as spheres may represent a strategy by which cells could deal with radiation found on other planets and in harsh environmental niches. This would be interesting to delve deeper into, and would explain the apparent similarity to wild-type cells, especially if the wild-type cells form spheres at low UV doses, therefore all cells in the control group are actually spheres.

6.5.4 Spheroplast formation resembles archaeal spheres

The microscopy of spheroplasts, formed via standard procedure using EDTA (0.5M, pH 8.0), shows that spheroplasts resemble the spherical cells formed under LiCl buffer (low water activity). Although this does not mean that the spheres formed under LiCl are just spheroplasts, it does allow further interrogation of S-layer structure whilst spheres are formed, and certainly does not rule out the possibility of low water activity not being the primary cause of sphere formation. Low water activity could be a by-product of the LiCl buffer, but in fact it could be the Li salt that removes either part or the entire S-layer from the archaeal cells, thus mediating a transition to a spheroplast.

Further study, perhaps staining the S-layer and forming spheres, might be able to shed light on the nature of the S-layer under such conditions and allow a clearer picture of if spheres are genuine morphological and biochemical cellular states, or simply spheroplasts. Another line of attack here would be to determine if the cells divide in order to form spheres. The paper by Fendrihan *et al*., 2012 indicates that one wild-type cell forms between three and four spheres, each of which can re-form a wild-type cell. This indicates a potential for asymmetric cell division, however the authors do not provide any evidence to support this claim.

6.5.5 S-layer removal does not pose as a negative control for ConA

As a result of the microscopy of cells stained for mannose residues on the S-layer by ConA fused to a fluorophore, it was determined that such a stain was not specific enough to be used and take this project futher. Without confirming if the S layer is indeed still present after sphere formation, a negative control must be established. The choice here was to remove the S-layer using different titrations of EDTA. Unfortunately, as can be seen in the results, although less ConA was detected, it still remained even after removal of the supernatant.

Therefore, it was never established as a good enough negative control, and therefore stain, to progress and stain the S-layer of wild-type cells vs spheres to say for certain if the S-layer was indeed still present, and therefore that sphere formation was a physiological response worthy of further study.

Other techniques to visualise the S-layer, such as CryoEM, would have been desirable. However, this was not possible due to time constraints and financial limitations. Alternatively, it might be possible to perform image-based analysis to examine the extent of ConA staining of archaeal cells – in EDTA treated samples there appears to be more patchy ConA coverage and colocalization of ConA staining and cell staining may be reduced. Super-resolution microscopy, such as structured light illumination may be usefully applied here in the future.

6.5.6 EDTA addition has no impact on cell viability

EDTA addition to cells during the staining assay in section 6.4.6 could have impacted on cellular viability, owing to chekatuib of divalent cations such as Mg^{2+}. Cells treated with varying amounts of EDTA were then subject to a survival assay. Although this was a pilot assay to look for any difference worth investigating, initial data shown in figure 6.8 shows that viability is not impacted.

6.5.7 Fluorescence analysis shows S-layer may not be removed by EDTA

The results seen from the attempted negative control study show that ConA is still present even with significant amounts of EDTA. Therefore is is surmised that whilst the literature reports EDTA 'stripping' off the S-layer to facilitate transformation, this study indicates that the S-layer is not stripped off in its entirety at all, and merely is disrupted by EDTA but the core constituents of the S-layer remain attached to the cell.

6.5.8 Transformation efficiency

The transformation efficiency was also looked into for cells treated with differing amounts of EDTA. EDTA is known to be important to the transformation process, mediated by PEG, allowing S-layer disruption to facilitate PEG-mediated DNA transfer. However, surprisingly results showed that when EDTA was omitted from the process, transformation efficiency appeared much higher than when the normal amount was added (10% v/v). This indicates that perhaps EDTA is not required, and that the S-layer may allow DNA into the cell via other means, perhaps through channels. More experiments would need to be carried out to confirm this.

Spheres were also generated, confirmed by microscopy, and subject to the same assay. Interestingly the transformation efficiency appeared to be less than that of wild-type cells treated with the same amount of EDTA. This perhaps indicates that spheres have less S-layer present.

50% EDTA (v/v) appears to drastically reduce the transformation efficiency of cells, perhaps due to cell death inflicted by such a high concentration of a divalent cation collator. This is therefore not surprising.

Overall, no significant differences were seen between spheres and non spheres and so the sphere project was abandoned half way through the PhD project, owing to this and the fact that there was no suitable negative control for S-layer staining. It had been hoped to perform a sphere-based assay with the genomic library, but the lack of suitable selection or quantification regime for sphere formation at scale prevented this.

Chapter 7: Discussion and future perspectives

XerCD recombinase increases resistance to UV radiation and Mitomycin C when overexpressed

A whole genome library screen for UV resistance genes indicated that a XerCD-like integrase present within the prophage region in *H. volcanii* contributes to increased cell survival under both UV and MMC induced genotoxic stress. When the specific coding sequence was placed under a strong constitutive promoter, p.syn, UV and MMC resistance was shown to be increased comparative to both wild-type and the already somewhat resistant *Δmre11-rad50* strain.

Multiple theories were considered as to why this could be the case, the most appropriate of which was that XerCD lies within a latent region of recent integration via phage activity and is therefore expressed at low levels. Previous RNA Seq analysis confirmed a three-fold increase in expression of XerCD upon treatment with MMC [TA lab, unpublished data]. Upon placement downstream of a strong promoter, XerCD appeared to act to increase levels of UV resistance.

It should be noted that this is not the 'day-job' of XerCD. In fact, the predominant function is to act to resolve chromosomal dimers and aberrant branched structures formed from colliding replication forks and unequal recombination events during DNA replication, thus facilitating replication termination and correct chromosomal disjunction. Overexpression could therefore be facilitating more efficient removal of such structures, therefore additional experiments such as overexpressing XerCD in a strain deleted for origins of replication would provide a useful and logical next step.

Of course, other possibilities could be true, such as XerCD being able to step in as a redundant, latent, activity. Data within this thesis indicates the possibility of XerCD assisting with UvrC, the canonical nuclease within the nucleotide excision repair pathway. Data on this was indictive, but not strong, so no further time was given to this line of enquiry.

ΔXerCD has no impact on growth rate or cell survival indicating redundant function with the remaining homologues

The deletion of the XerCD coding sequence identified in the screen resulted in no impact on growth rate, and no decrease in UV resistance compared to that of wild-type. It is therefore surmised that the remaining eleven homologues in *H. volcanii* can carry out similar functions. Furthermore, XerCD may be at transcriptionally low levels to begin with and therefore removing such activity may not have much of an impact on cell health. This also may be true given that the prophage has resulted from a recent horizontal gene transfer event, and so the genes within this region are unlikely to be essential to the host at this stage in its evolutionary history.

Δprophage increases levels of UV and MMC resistance beyond that of XerCD overexpression

Within the context of the DNA damaging assays in chapter four, it was found that deleting the entire prophage region from the main chromosome of *H. volcanii* increased levels of resistance to both MMC an UV induced DNA damage. The proposed explanation for this is phage induction after UV or MMC treatment, and thus subsequent entry into a lytic lifestyle leads to cellular death at a certain 'background' level. Therefore all DNA damaging assays are showing cellular death as a result of two factors:

(1) The initial genotoxic insult

(2) Phage induction and subsequent cell lysis

Removing the ability of the phage to lyse cells in the prophage deleted strain therefore brings cell survival fractions up, resulting in what appears to be UV or MMC resistance. Alternatively, phage genome excision, even without production of lytic phage may exacerbate host genome damage or related effects. Similarly, production of phage proteins mey have toxic effects without direct phage-induced lysis.

The presence of phage was confirmed by qPCR of both cell pellets and supernatant after incubation with MMC. Phage DNA was confirmed to be present in both samples and at greater levels in the supernatant compared to the endogenous control. The amount of phage DNA does decrease significantly in MMC treated cells (pellets, figure 5.14), even after 30 minutes. This is perhaps due to loss of phage either from lysis, or as an artefact of the experimental conditions leading to loss of phage during downstream detection. The assay did not demonstrate a large

increase in detectable phage DNA beyond that of an endogenous control. This assay was carried out in the final week or so of the PhD process, and so repeating this, perhaps with a different internal control, may provide a better set of results that can more clearly prove or disprove this theory.

Transcription of XerCD is the rate limiting step

Analysis via the use of a codon optomised XerCD under UV stress demonstrated that translation efficiency is less important than transcription for UV resistance. Therefore this explains why XerCD increases UV resistance under p.syn and lends further weight to the notion that the prophage is not particularly active under normal conditions, reflected by its recent genomic integration and lack of G/C content.

RNA sequencing data suggests RPA and RPA3AP are involved in the DNA damage response to MMC

Genes for RPA and its associated protein, RPA3AP are upregulated in wild-type cells when treated with MMC. This support the data shown by Stroud et al, 2015 and further confirms the key role RPA and its homologues play across the domains of life to maintain ssDNA integrity and stability.

Suggestions for further experiments

(1) ΔCas3 complementation

A recent paper has shown that a strain deleted for Cas3 demonstrates increased levels of UV resistance by acting with Mre11 and Rad50 to restrain homologous recombination under normal circumstances. Deletion of such proteins therefore increase levels of homologous recombination and thus the DNA damage response. It would therefore be of interest to overexpress XerCD in this Δ*cas3* background (Miezner *et al.*, 2023).

(2) Protein pulldowns with XerCD

As the tagged version of XerCD has been generated in this study, it would be useful to perform a pull down under both normal and UV/MMC conditions to look for potential interacting partners. A split GFP assay could then be used to confirm protein protein interactions (Winter, Born and Pfeifer, 2018, Bignon, Gruet and Longhi, 2022). This would also provide further insight into if XerCD is recruited to sites of UV damage by Uvr proteins. XerCD has not yet been confirmed at replication forks, so this may reveal a non-damage associated role also.

(3) Overexpress XerCD in an origin-less strain

Given the notion that XerCD may act to increase levels of UV resistance by dealing with aberrant branched structures and concatenated chromosomes during DNA replication, expressing XerCD at high levels in a Δ*ori* strain may prove a useful insight into how important XerCD is, and if this is the predominant route by which UV and MMC resistance is conferred. If XerCD increases homologous recombination then we would expect faster replication, if HR is restrained then we would expect lower growth rates.

(4) Complement a strain deleted for all XerCD homologues

Given that twelve homologues exist in *H. volcanii* it would be of interest to complement a strain deleted for each XerCD as well as investigate phenotypic properties of this strain such as resistance to DNA damaging agents and growth rates.

(5) Confirm phage presence outside cells using CryoEM before and after MMC treatment

The qPCR assay did not provide a definitive answer with regards to phage induction. It is therefore advised to use transmission electron microscopy on both a wild-type and MMC treated culture to look for increased phage particles outside cells.

Appendix I: RNASeq Data

RNA Seq analysis was carried out by DeepSeq, University of Nottingham. The following samples were used for differentially expressed gene analysis (DEG):

Sample A = H26 (Δ*pyrE2*)

Sample B = H26 (Δ*pyrE2*) + MMC

Sample C = H1192 (Δ*pyrE2,* Δ*phage*)

Sample D = H1192 (Δ*pyrE2,* Δ*phage*) +MMC

Table 1. Summary of raw read counts, percent passing quality control, percent mapped by 'subread', and percent uniquely mapped as paired reads.

Sample No.	ID	Raw Reads	Passing QC	Mapping Rate (% of QC reads)	Correctly paired, mapping (% of QC reads)
1	1	19,093,669	17,822,866	93.09	92.42
2	2	17,341,350	16,170,574	95.72	95.02
3	3	20,209,483	18,865,660	95.74	95.08
6	2 mmc	18,776,308	17,522,853	97.86	97.19
7	3 mmc	19,221,340	17,937,698	96.93	96.25
8	4 mmc	19,043,964	17,728,673	94.86	94.16
9	H1192 1	16,493,708	15,364,522	95.90	95.20
11	H1192 3	21,265,118	19,832,382	96.01	95.33
12	H1192 4	18,624,143	17,449,436	97.45	96.78
13	H1192 1 mmc	22,619,488	21,148,723	94.93	94.34
14	H1192 2 mmc	21,111,651	19,704,837	96.11	95.48
15	H1192 3 mmc	19,867,815	18,535,805	96.76	96.09
16	H1192 4 mmc	19,726,512	18,366,577	96.36	95.69

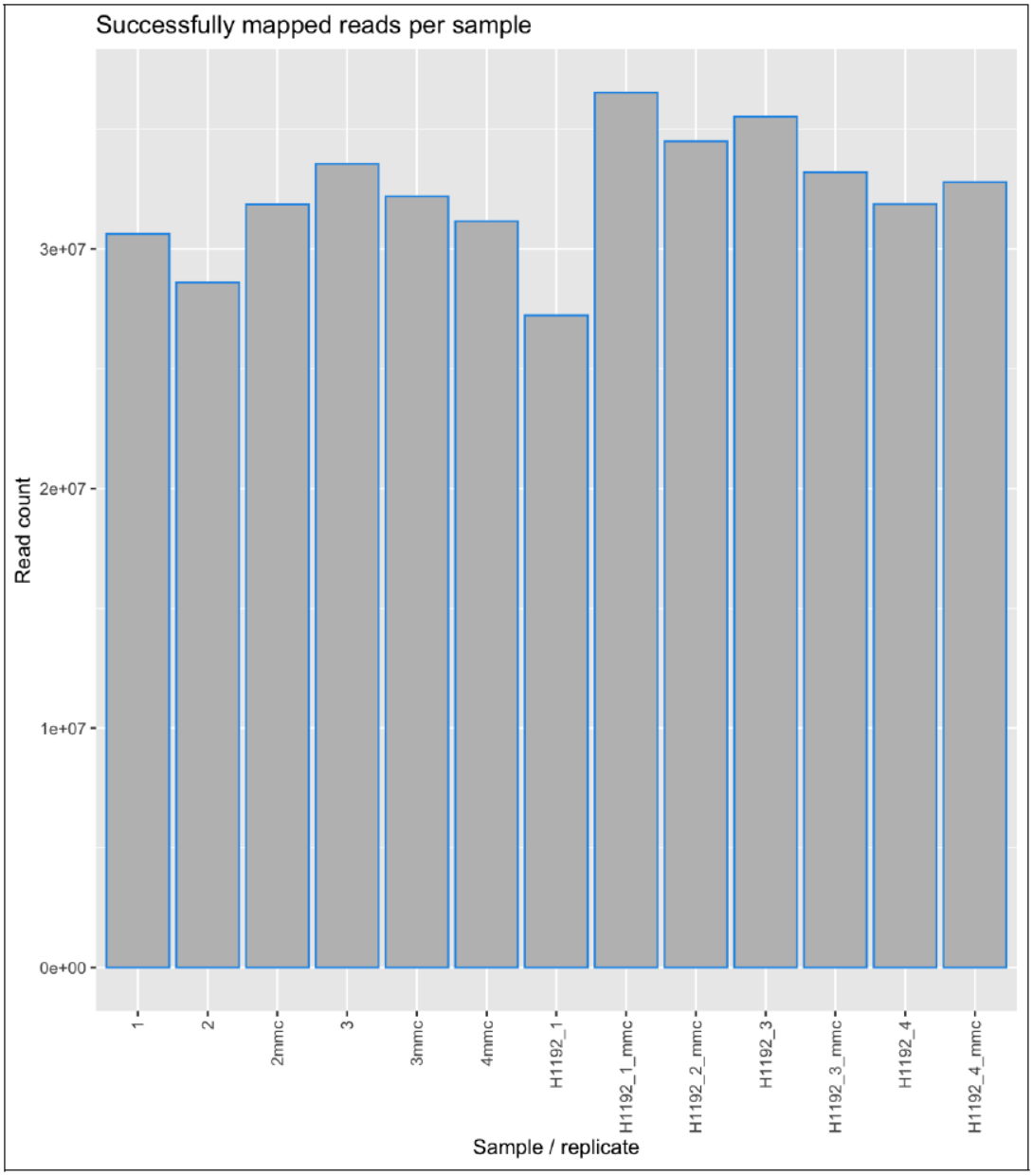


Figure 1. Bar chart showing the total numbers of quality trimmed reads successfully mapped in each sample/replicate. These are high alignment score mappings with both reads of each pair mapping successfully.

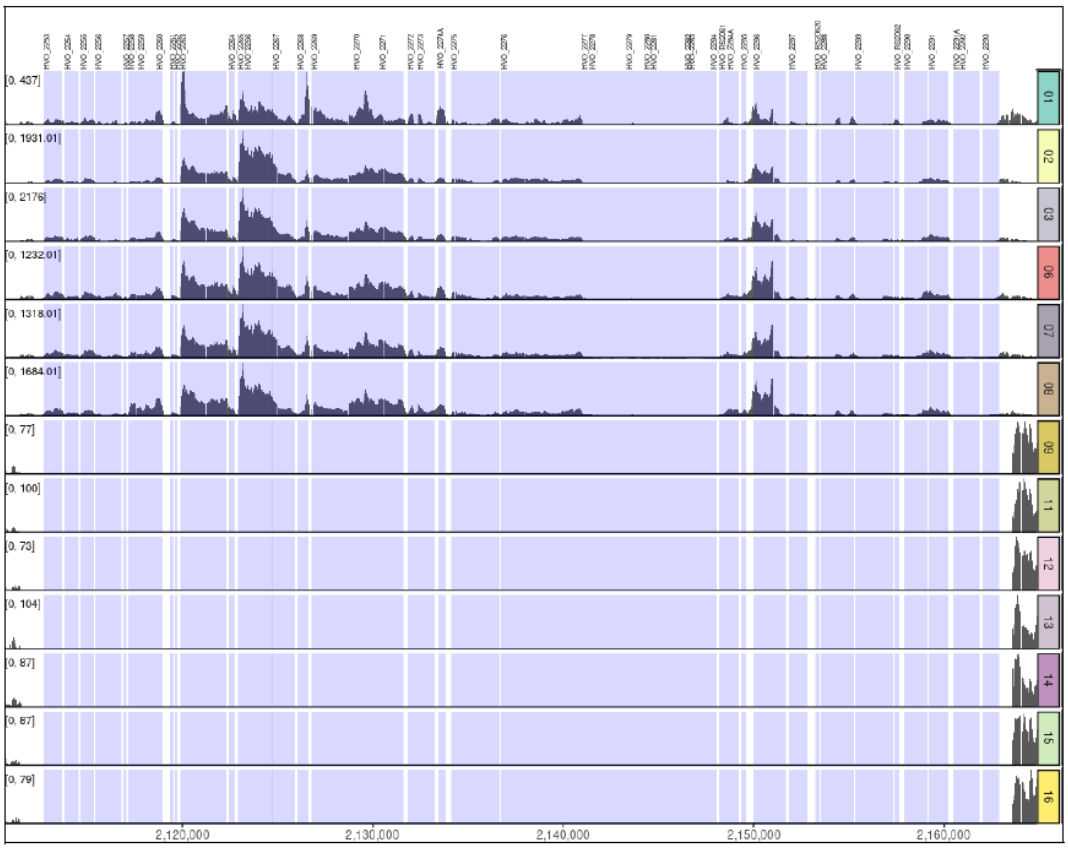


Figure 2. Coverage of mapped reads within the region of interest (HVO_2253 -> HVO_2293; NC_013967.1:2112703-2162873. Each sample is plotted as a row with genes highlighted in light blue columns. Groups A and B are very clearly separated from groups C and D in this region.

Table 2. Counts of significantly, differentially expressed genes by contrast.

Contrast	Up-regulated	Down-regulated
A vs B	4 (0.1%)	26 (0.7%)
A vs C	165 (4.1%)	107 (2.7%)
A vs D	414 (10.0%)	491 (12.0%)
B vs C	224 (5.6%)	149 (3.7%)
B vs D	355 (8.9%)	438 (11.0%)
C vs D	48 (1.2%)	155 (3.9%)
A vs D1	1190 (30.0%)	1094 (27%)
A vs D2	530 (13.0%)	446 (11.0%)
B vs D1	1254 (31.0%)	1150 (29.0%)
B vs D2	307 (7.7%)	258 (6.5%)
C vs D1	1182 (30.0%)	1122 (28.0%)
C vs D2	9 (0.22%)	36 (0.9%)

Contrast 1) A vs B

Table 3. A vs B*: Top four* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS06625	HVO_0399	7732.974	0.6125279	1.87E-05	hypothetical protein
gene-HVO_RS11415	HVO_1393	6319.939	0.606236	1.78E-08	hypothetical protein
gene-HVO_RS11420	HVO_1394	35188.137	0.601631	1.78E-08	Na+/H+ antiporter NhaC family protein
gene-HVO_RS12830	HVO_1691	4457.876	0.5539747	1.82E-02	PRC-barrel domain-containing protein

Table 4. A vs B*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS03060	HVO_A0274	4422.36078	-0.5287768	6.55E-03	inositol monophosphatase
gene-HVO_RS18350	HVO_2823	791.95565	-0.7079302	4.76E-03	hypothetical protein
gene-HVO_RS01270	HVO_B0261	1210.37521	-0.7649532	3.62E-03	aminotransferase class III-fold pyridoxal phosphate-dependent enzyme
gene-HVO_RS06565	HVO_0387	3790.10438	-0.904846	1.03E-04	hypothetical protein
gene-HVO_RS06105	HVO_0291	35807.11667	-0.9092942	2.62E-02	RPA family protein
gene-HVO_RS06110	HVO_0292	83191.06987	-0.9553441	3.17E-02	hypothetical protein
gene-HVO_RS07800	HVO_0644	21012.66936	-0.9858898	4.49E-05	2-isopropylmalate synthase
gene-HVO_RS01600	HVO_B0333	387.79305	-1.0198348	1.24E-04	Gfo/Idh/MocA family oxidoreductase
gene-HVO_RS07650	HVO_0613	116.76355	-1.0494212	3.05E-02	MBL fold metallo-hydrolase
gene-HVO_RS03210	HVO_A0303	230.25147	-1.0612525	1.54E-02	amidohydrolase family protein
gene-HVO_RS00620	HVO_B0128	337.03502	-1.0653907	8.95E-03	SDR family oxidoreductase
gene-HVO_RS08890	HVO_0874	2809.33182	-1.0809524	3.17E-02	beta-CASP ribonuclease aCPSF1
gene-HVO_RS04390	HVO_A0568	100.7093	-1.1750496	4.90E-02	hydroxymethylglutaryl-CoA lyase
gene-HVO_RS01670	HVO_B0346	77.14375	-1.201128	8.10E-04	mandelate racemase/muconate lactonizing enzyme family protein
gene-HVO_RS01655	HVO_B0343	129.16395	-1.2158533	4.49E-05	glycoside hydrolase family 4
gene-HVO_RS04380	HVO_A0566	153.20935	-1.2175359	2.08E-02	3-isopropylmalate dehydratase small subunit
gene-HVO_RS00605	HVO_B0125	42.52853	-1.2714974	2.62E-02	ABC transporter ATP-binding protein
gene-HVO_RS04385	HVO_A0567	199.66967	-1.2857231	1.24E-03	3-isopropylmalate dehydratase large subunit
gene-HVO_RS01490	HVO_B0310	123.55036	-1.3411941	2.62E-02	FAD binding domain-containing protein
gene-HVO_RS06560	HVO_0386	2977.70477	-1.362734	1.02E-06	hypothetical protein

Contrast 2) A vs C

Table 5. A vs C*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15590	HVO_2266	3883.9596	15.62632	2.82E-27	hypothetical protein
gene-HVO_RS15575	HVO_2263	3887.6862	15.46327	8.10E-29	winged helix-turn-helix transcriptional regulator
gene-HVO_RS21315	HVO_2286	2469.3796	14.64924	1.32E-24	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.45939	2.64E-26	type I restriction-modification system subunit M
gene-HVO_RS15640	HVO_2276	1777.7261	14.44317	2.60E-25	hypothetical protein
gene-HVO_RS15605	HVO_2269	1716.518	14.27566	2.64E-26	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS20265	HVO_2271	1548.562	14.0987	3.47E-22	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.97286	1.94E-22	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.67229	3.50E-22	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.22305	3.91E-22	N-6 DNA methylase
gene-HVO_RS15725	HVO_2291	566.281	12.69337	1.35E-19	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.61901	2.80E-22	TATA-box-binding protein
gene-HVO_RS20605	HVO_2253	488.4749	12.32601	6.58E-19	hypothetical protein
gene-HVO_RS15540	HVO_2255	391.8697	12.1255	1.43E-18	hypothetical protein
gene-HVO_RS15625	HVO_2273	399.0095	12.0058	4.47E-18	tyrosine-type recombinase/integrase
gene-HVO_RS15555	HVO_2259	459.3064	11.92631	6.60E-18	tyrosine-type recombinase/integrase
gene-HVO_RS15710	HVO_2288	340.4556	11.90331	1.43E-18	DUF2254 domain-containing protein
gene-HVO_RS15720	HVO_2290	328.8174	11.8155	2.61E-16	tyrosine-type recombinase/integrase
gene-HVO_RS15630	HVO_2274A	355.3856	11.80981	9.91E-20	hypothetical protein
gene-HVO_RS15560	HVO_2260	305.3267	11.59768	2.87E-17	hypothetical protein

Table 6. A vs C: *Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS11415	HVO_1393	6319.9394	-0.2999388	1.67E-02	hypothetical protein
gene-HVO_RS08910	HVO_0878	2089.7419	-0.3311125	2.29E-02	endonuclease III
gene-HVO_RS05560	HVO_0175	32030.5724	-0.3483447	3.19E-02	DNA polymerase sliding clamp
gene-HVO_RS10990	HVO_1306	3348.9428	-0.4547055	4.96E-02	chorismate synthase
gene-HVO_RS08195	HVO_0727	2254.4908	-0.4562275	3.33E-02	rhomboid family intramembrane serine protease
gene-HVO_RS08670	HVO_0827	2681.7809	-0.4630044	4.85E-02	MTH865 family protein
gene-HVO_RS11420	HVO_1394	35188.1374	-0.4672401	9.19E-06	Na+/H+ antiporter NhaC family protein
gene-HVO_RS03060	HVO_A0274	4422.3608	-0.4803934	4.47E-03	inositol monophosphatase
gene-HVO_RS12830	HVO_1691	4457.8756	-0.4845925	1.39E-02	PRC-barrel domain-containing protein
gene-HVO_RS18715	HVO_2897	1784.1103	-0.4846729	2.61E-04	hypothetical protein
gene-HVO_RS15780	HVO_2302	10340.2572	-0.4892794	3.82E-02	AarF/ABC1/UbiB kinase family protein
gene-HVO_RS09870	HVO_1077	1269.9664	-0.4896407	2.13E-02	diphthine--ammonia ligase
gene-HVO_RS15775	HVO_2301	1325.7711	-0.497989	3.06E-02	DUF5518 domain-containing protein
gene-HVO_RS18575	HVO_2870	4512.9726	-0.5023233	9.53E-04	AI-2E family transporter
gene-HVO_RS11905	HVO_1492	2621.4058	-0.5092804	2.54E-03	DUF5799 family protein
gene-HVO_RS19030	HVO_2963	749.9433	-0.5241551	1.01E-02	DUF2178 domain-containing protein
gene-HVO_RS06425	HVO_0357	5514.6091	-0.5306559	4.61E-02	hypothetical protein
gene-HVO_RS16295	HVO_2408	1690.5875	-0.5311936	1.31E-02	energy-coupling factor ABC transporter permease
gene-HVO_RS06330	HVO_0338	5212.8213	-0.5332689	9.69E-03	transcriptional regulator
gene-HVO_RS12300	HVO_1581	1643.8582	-0.536389	1.16E-02	cytidine deaminase

Table 7. A vs D: *Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15590	HVO_2266	3883.9596	15.60962	2.74E-35	hypothetical protein
gene-HVO_RS15575	HVO_2263	3887.6862	15.44657	9.55E-38	winged helix-turn-helix transcriptional regulator
gene-HVO_RS21315	HVO_2286	2469.3796	14.63254	8.42E-32	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.44269	1.39E-34	type I restriction-modification system subunit M
gene-HVO_RS15640	HVO_2276	1777.7261	14.42647	5.66E-33	hypothetical protein
gene-HVO_RS15605	HVO_2269	1716.518	14.25896	1.39E-34	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS20265	HVO_2271	1548.562	14.082	1.22E-28	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.95616	4.40E-29	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.65558	9.48E-29	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.20635	7.89E-29	N-6 DNA methylase
gene-HVO_RS15725	HVO_2291	566.281	12.67667	1.75E-25	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.60231	1.52E-29	TATA-box-binding protein
gene-HVO_RS20605	HVO_2253	488.4749	12.30931	1.08E-24	hypothetical protein
gene-HVO_RS15540	HVO_2255	391.8697	12.1088	2.59E-24	hypothetical protein
gene-HVO_RS15625	HVO_2273	399.0095	11.9891	1.25E-23	tyrosine-type recombinase/integrase
gene-HVO_RS15555	HVO_2259	459.3064	11.90961	2.07E-23	tyrosine-type recombinase/integrase
gene-HVO_RS15710	HVO_2288	340.4556	11.88661	2.22E-24	DUF2254 domain-containing protein
gene-HVO_RS15720	HVO_2290	328.8174	11.7988	4.65E-21	tyrosine-type recombinase/integrase
gene-HVO_RS15630	HVO_2274A	355.3856	11.79311	3.10E-26	hypothetical protein
gene-HVO_RS15560	HVO_2260	305.3267	11.58098	1.29E-22	hypothetical protein

Table 8. A vs D*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS08865	HVO_0868	7286.2338	-0.2164535	4.27E-02	NAD-dependent epimerase/dehydratase family protein
gene-HVO_RS17545	HVO_2659	1736.1653	-0.2868613	4.52E-02	tautomerase family protein
gene-HVO_RS17600	HVO_2671	9593.1202	-0.3040779	4.02E-02	alanine--glyoxylate aminotransferase family protein
gene-HVO_RS11275	HVO_1365	759.4844	-0.31293	4.05E-02	MBL fold metallo-hydrolase
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.31347	2.25E-03	2-dehydro-3-deoxygalactono kinase KdgK2
gene-HVO_RS17640	HVO_2679	1057.1735	-0.3234003	2.43E-02	hypothetical protein
gene-HVO_RS13880	HVO_1916	23094.007	-0.3322205	4.46E-02	cation:proton antiporter
gene-HVO_RS05625	HVO_0190	6332.1556	-0.3334189	1.57E-02	ATP-dependent sacrificial sulfur transferase LarE
gene-HVO_RS09110	HVO_0921	1051.2426	-0.3435349	2.90E-02	Lrp/AsnC family transcriptional regulator
gene-HVO_RS05485	HVO_0160	1452.1151	-0.3466044	4.26E-02	type II CAAX endopeptidase family protein
gene-HVO_RS08910	HVO_0878	2089.7419	-0.3507849	2.67E-03	endonuclease III
gene-HVO_RS18310	HVO_2814	1348.5724	-0.3665677	4.65E-02	DNA-3-methyladenine glycosylase
gene-HVO_RS06860	HVO_0449	3911.4596	-0.386314	4.64E-03	prephenate dehydratase
gene-HVO_RS19055	HVO_2968	16880.1031	-0.3894359	4.01E-02	phosphoglycerate dehydrogenase
gene-HVO_RS06145	HVO_0299	631.926	-0.3959748	2.79E-02	hypothetical protein
gene-HVO_RS18455	HVO_2843	1412.0349	-0.404496	2.01E-03	DASH family cryptochrome
gene-HVO_RS10345	HVO_1174	9919.8374	-0.4140035	4.28E-02	transcription factor
gene-HVO_RS15970	HVO_2341	1720.872	-0.41773	4.42E-02	TetR/AcrR family transcriptional regulator
gene-HVO_RS08485	HVO_0789	3201.6847	-0.4208549	4.26E-02	tryptophan synthase subunit alpha
gene-HVO_RS08860	HVO_0867	8498.7442	-0.4282836	3.73E-02	HD domain-containing protein

Table 9. B vs C*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15575	HVO_2263	3887.6862	15.57398	1.46E-29	winged helix-turn-helix transcriptional regulator
gene-HVO_RS15590	HVO_2266	3883.9596	15.40152	7.54E-27	hypothetical protein
gene-HVO_RS21315	HVO_2286	2469.3796	15.05253	4.80E-26	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.56472	7.54E-27	type I restriction-modification system subunit M
gene-HVO_RS15605	HVO_2269	1716.518	14.40212	7.54E-27	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS15640	HVO_2276	1777.7261	14.33648	4.57E-25	hypothetical protein
gene-HVO_RS20265	HVO_2271	1548.562	14.27907	9.85E-23	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.81106	4.32E-22	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.80241	1.40E-22	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.17444	5.29E-22	N-6 DNA methylase
gene-HVO_RS15555	HVO_2259	459.3064	12.81553	9.50E-21	tyrosine-type recombinase/integrase
gene-HVO_RS15725	HVO_2291	566.281	12.78598	4.86E-20	hypothetical protein
gene-HVO_RS20605	HVO_2253	488.4749	12.70373	3.17E-20	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.58856	2.90E-22	TATA-box-binding protein
gene-HVO_RS15625	HVO_2273	399.0095	12.4333	1.62E-19	tyrosine-type recombinase/integrase
gene-HVO_RS15540	HVO_2255	391.8697	12.28852	2.93E-19	hypothetical protein
gene-HVO_RS15630	HVO_2274A	355.3856	12.2876	1.66E-21	hypothetical protein
gene-HVO_RS15710	HVO_2288	340.4556	12.10263	2.39E-19	DUF2254 domain-containing protein
gene-HVO_RS15550	HVO_2258	257.5188	12.09416	2.05E-17	hypothetical protein
gene-HVO_RS15720	HVO_2290	328.8174	12.08436	3.59E-17	tyrosine-type recombinase/integrase

Table 10. B vs C: *Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS08910	HVO_0878	2089.742	-0.2951251	4.29E-02	endonuclease III
gene-HVO_RS14050	HVO_1949	3120.813	-0.3298014	2.57E-02	DUF5809 family protein
gene-HVO_RS18455	HVO_2843	1412.035	-0.3324829	4.29E-02	DASH family cryptochrome
gene-HVO_RS16855	HVO_2519	26050.842	-0.3362074	1.83E-02	hypothetical protein
gene-HVO_RS09775	HVO_1058	9788.399	-0.3464672	2.75E-02	Trk system potassium transporter TrkA
gene-HVO_RS12160	HVO_1553	1924.221	-0.3492534	2.34E-02	DUF4147 domain-containing protein
gene-HVO_RS03965	HVO_A0476	1419.967	-0.3558988	1.86E-02	hypothetical protein
gene-HVO_RS19000	HVO_2957	10551.424	-0.3595327	4.93E-02	lipoyl synthase
gene-HVO_RS15130	HVO_2168	10317.096	-0.3651527	1.84E-02	MoxR family ATPase
gene-HVO_RS15930	HVO_2333	4168.763	-0.3655593	4.52E-02	DUF6293 family protein
gene-HVO_RS16600	HVO_2468	1480.679	-0.3731881	3.29E-03	SDR family oxidoreductase
gene-HVO_RS05560	HVO_0175	32030.572	-0.3755109	1.36E-02	DNA polymerase sliding clamp
gene-HVO_RS08835	HVO_0863	9949.108	-0.3852913	3.05E-02	metal-dependent transcriptional regulator
gene-HVO_RS15165	HVO_2175	3985.757	-0.3855244	3.71E-02	AAA family ATPase
gene-HVO_RS05315	HVO_0128	3897.367	-0.3918083	7.08E-03	AAA family ATPase
gene-HVO_RS09945	HVO_1092	3800.341	-0.4242828	1.36E-02	Rpp14/Pop5 family protein
gene-HVO_RS11905	HVO_1492	2621.406	-0.4458781	9.70E-03	DUF5799 family protein
gene-HVO_RS12290	HVO_1579	9446.501	-0.4472294	1.67E-02	nucleoside phosphorylase
gene-HVO_RS08195	HVO_0727	2254.491	-0.4505839	2.99E-02	rhomboid family intramembrane serine protease
gene-HVO_RS17640	HVO_2679	1057.173	-0.4523072	4.10E-03	hypothetical protein

Table 11. B vs D*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15575	HVO_2263	3887.6862	15.55728	2.57E-38	winged helix-turn-helix transcriptional regulator
gene-HVO_RS15590	HVO_2266	3883.9596	15.38482	1.55E-34	hypothetical protein
gene-HVO_RS21315	HVO_2286	2469.3796	15.03582	1.46E-33	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.54802	5.11E-35	type I restriction-modification system subunit M
gene-HVO_RS15605	HVO_2269	1716.518	14.38542	5.11E-35	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS15640	HVO_2276	1777.7261	14.31978	1.50E-32	hypothetical protein
gene-HVO_RS20265	HVO_2271	1548.562	14.26237	2.84E-29	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.79436	1.58E-28	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.78571	2.84E-29	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.15774	1.17E-28	N-6 DNA methylase
gene-HVO_RS15555	HVO_2259	459.3064	12.79883	5.82E-27	tyrosine-type recombinase/integrase
gene-HVO_RS15725	HVO_2291	566.281	12.76928	6.22E-26	hypothetical protein
gene-HVO_RS20605	HVO_2253	488.4749	12.68702	2.97E-26	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.57186	2.12E-29	TATA-box-binding protein
gene-HVO_RS15625	HVO_2273	399.0095	12.4166	2.37E-25	tyrosine-type recombinase/integrase
gene-HVO_RS15540	HVO_2255	391.8697	12.27182	4.58E-25	hypothetical protein
gene-HVO_RS15630	HVO_2274A	355.3856	12.2709	1.74E-28	hypothetical protein
gene-HVO_RS15710	HVO_2288	340.4556	12.08593	2.76E-25	DUF2254 domain-containing protein
gene-HVO_RS15550	HVO_2258	257.5188	12.07745	1.88E-22	hypothetical protein
gene-HVO_RS15720	HVO_2290	328.8174	12.06766	4.63E-22	tyrosine-type recombinase/integrase

Table 12. B vs D*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.2205533	4.40E-02	2-dehydro-3-deoxygalactonokinase KdgK2
gene-HVO_RS11415	HVO_1393	6319.9394	-0.2847593	7.68E-03	hypothetical protein
gene-HVO_RS08865	HVO_0868	7286.2338	-0.2960955	4.74E-03	NAD-dependent epimerase/dehydratase family protein
gene-HVO_RS16170	HVO_2382	2392.0836	-0.3143829	4.43E-02	hypothetical protein
gene-HVO_RS08910	HVO_0878	2089.7419	-0.3147975	9.61E-03	endonuclease III
gene-HVO_RS12705	HVO_1665	1028.6613	-0.3191763	3.93E-02	hypothetical protein
gene-HVO_RS09945	HVO_1092	3800.3411	-0.3255112	3.52E-02	Rpp14/Pop5 family protein
gene-HVO_RS06860	HVO_0449	3911.4596	-0.3329367	1.96E-02	prephenate dehydratase
gene-HVO_RS18585	HVO_2872	2789.5539	-0.3362029	3.62E-02	hypothetical protein
gene-HVO_RS15930	HVO_2333	4168.7627	-0.342852	2.65E-02	DUF6293 family protein
gene-HVO_RS18590	HVO_2873	3558.1918	-0.3437996	1.85E-02	flap endonuclease-1
gene-HVO_RS05625	HVO_0190	6332.1556	-0.3470019	1.38E-02	ATP-dependent sacrificial sulfur transferase LarE
gene-HVO_RS13880	HVO_1916	23094.007	-0.3490052	3.78E-02	cation:proton antiporter
gene-HVO_RS08835	HVO_0863	9949.108	-0.351153	2.17E-02	metal-dependent transcriptional regulator
gene-HVO_RS06665	HVO_0408	1768.8822	-0.3559505	6.44E-03	PrsW family intramembrane metalloprotease
gene-HVO_RS03060	HVO_A0274	4422.3608	-0.3573098	1.86E-02	inositol monophosphatase
gene-HVO_RS12460	HVO_1614	756.3432	-0.3628191	4.59E-02	hypothetical protein
gene-HVO_RS09870	HVO_1077	1269.9664	-0.3659138	4.61E-02	diphthine--ammonia ligase
gene-HVO_RS16600	HVO_2468	1480.679	-0.3712487	9.59E-04	SDR family oxidoreductase
gene-HVO_RS18720	HVO_2898	2116.0099	-0.3763592	2.36E-02	hypothetical protein

Table 13. C vs D: *Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS04350	HVO_A0560	11079.5581	1.5227621	4.19E-02	imidazolonepropionase
gene-HVO_RS08430	HVO_0778	144993.8159	1.4957102	2.96E-02	thermosome subunit alpha
gene-HVO_RS14495	HVO_2037	10185.3616	1.4637219	3.69E-02	SHOCT domain-containing protein
gene-HVO_RS01720	HVO_B0356	602.60693	1.3137115	4.71E-02	hypothetical protein
gene-HVO_RS01425	HVO_B0294	85.30061	1.3066395	5.64E-03	carbohydrate ABC transporter permease
gene-HVO_RS01750	HVO_B0362	446.0108	1.2976726	2.98E-02	4Fe-4S ferredoxin N-terminal domain-containing protein
gene-HVO_RS06590	HVO_0392	3008.65636	1.2863389	2.96E-02	cell division protein SepF
gene-HVO_RS14985	HVO_2140	717.21341	1.1701959	4.68E-02	hypothetical protein
gene-HVO_RS01745	HVO_B0361	5575.90395	1.1307166	4.52E-02	helix-turn-helix domain-containing protein
gene-HVO_RS00670	HVO_B0138	4504.01336	1.1049682	4.14E-02	PQQ-binding-like beta-propeller repeat protein
gene-HVO_RS04345	HVO_A0559	18041.72397	1.0794072	3.93E-02	histidine ammonia-lyase
gene-HVO_RS00915	HVO_B0186	6472.92664	0.9797124	1.29E-02	ABC transporter permease
gene-HVO_RS14720	HVO_2084	3066.47714	0.976151	2.71E-02	ABC transporter permease
gene-HVO_RS05505	HVO_0164	6811.8887	0.9677412	1.15E-02	hypothetical protein
gene-HVO_RS16645	HVO_2477	1146.6647	0.9550505	4.94E-02	triphosphoribosyl-dephospho-CoA synthase
gene-HVO_RS08020	HVO_0690	3951.12147	0.9544371	3.32E-02	ScpA family protein
gene-HVO_RS01715	HVO_B0355	2354.66841	0.9173069	3.75E-02	hypothetical protein
gene-HVO_RS05760	HVO_0217	687.4953	0.9085923	4.94E-02	ABC transporter ATP-binding protein
gene-HVO_RS16970	HVO_2542	16290.82048	0.8522711	4.14E-02	uL15m family ribosomal protein
gene-HVO_RS14710	HVO_2082	6178.07655	0.8220296	3.34E-02	hypothetical protein

Table 14. C vs D*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS20870	HVO_0735	624.1493	-0.370722	2.14E-02	hypothetical protein
gene-HVO_RS13930	HVO_1926	892.0183	-0.4007092	6.40E-03	SIMPL domain-containing protein
gene-HVO_RS03060	HVO_A0274	4422.3608	-0.4056932	2.20E-02	inositol monophosphatase
gene-HVO_RS19265	HVO_3009	4311.2112	-0.434731	3.59E-02	universal stress protein
gene-HVO_RS06145	HVO_0299	631.926	-0.4390392	4.33E-02	hypothetical protein
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.4414819	3.62E-05	2-dehydro-3-deoxygalactono kinase KdgK2
gene-HVO_RS03175	HVO_A0296	353.6849	-0.4495673	4.07E-02	3-oxoacyl-ACP reductase family protein
gene-HVO_RS11130	HVO_1335	1427.6878	-0.4556121	2.96E-02	hypothetical protein
gene-HVO_RS10835	HVO_1275	2396.3075	-0.4884958	4.28E-02	secondary thiamine-phosphate synthase enzyme YjbQ
gene-HVO_RS03020	HVO_A0266	3114.7503	-0.4976969	4.26E-02	IclR family transcriptional regulator
gene-HVO_RS18180	HVO_2788	1187.4564	-0.5071865	4.07E-02	universal stress protein
gene-HVO_RS14960	HVO_2135	619.0261	-0.5133718	4.72E-02	hypothetical protein
gene-HVO_RS06765	HVO_0430	1212.8997	-0.5203442	1.63E-02	UPF0146 family protein
gene-HVO_RS08935	HVO_0883	1888.1298	-0.554243	3.32E-02	hypothetical protein
gene-HVO_RS06825	HVO_0442	3815.1098	-0.5626987	3.51E-02	DUF5786 family protein
gene-HVO_RS15340	HVO_2211	1216.862	-0.5676665	2.95E-02	TrkA family potassium uptake protein
gene-HVO_RS18350	HVO_2823	791.9557	-0.5704208	1.34E-02	hypothetical protein
gene-HVO_RS01940	HVO_A0018	588.0728	-0.5806408	4.93E-02	RNA-guided endonuclease TnpB family protein
gene-HVO_RS12235	HVO_1569	1406.3037	-0.5987965	9.21E-03	hypothetical protein
gene-HVO_RS08795	HVO_0855	2347.9533	-0.616775	1.63E-02	hypothetical protein

Table 15. A vs D1*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15590	HVO_2266	3883.9596	15.41581	3.16E-16	hypothetical protein
gene-HVO_RS15575	HVO_2263	3887.6862	15.25277	1.14E-18	winged helix-turn-helix transcriptional regulator
gene-HVO_RS21315	HVO_2286	2469.3796	14.43872	8.35E-15	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.24892	4.94E-19	type I restriction-modification system subunit M
gene-HVO_RS15640	HVO_2276	1777.7261	14.23267	4.08E-16	hypothetical protein
gene-HVO_RS15605	HVO_2269	1716.518	14.06519	7.53E-19	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS20265	HVO_2271	1548.562	13.88819	5.41E-13	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.76235	1.17E-13	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.46179	8.53E-14	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.01259	5.52E-16	N-6 DNA methylase
gene-HVO_RS15725	HVO_2291	566.281	12.48283	1.26E-12	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.40862	1.11E-15	TATA-box-binding protein
gene-HVO_RS20605	HVO_2253	488.4749	12.11557	7.37E-14	hypothetical protein
gene-HVO_RS15540	HVO_2255	391.8697	11.91506	1.49E-13	hypothetical protein
gene-HVO_RS15625	HVO_2273	399.0095	11.79536	3.13E-13	tyrosine-type recombinase/integrase
gene-HVO_RS15555	HVO_2259	459.3064	11.71588	4.24E-13	tyrosine-type recombinase/integrase
gene-HVO_RS15710	HVO_2288	340.4556	11.69288	2.08E-13	DUF2254 domain-containing protein
gene-HVO_RS15720	HVO_2290	328.8174	11.60495	3.49E-10	tyrosine-type recombinase/integrase
gene-HVO_RS15630	HVO_2274A	355.3856	11.59944	6.37E-14	hypothetical protein
gene-HVO_RS15560	HVO_2260	305.3267	11.38727	1.40E-12	hypothetical protein

Table 16. A vs D1*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS08865	HVO_0868	7286.234	-0.2448685	2.17E-02	NAD-dependent epimerase/dehydratase family protein
gene-HVO_RS08225	HVO_0734	7933.454	-0.2560859	4.69E-02	DNA repair protein NreA
gene-HVO_RS10165	HVO_1137	3272.417	-0.2655525	4.96E-02	NAD-binding protein
gene-HVO_RS18890	HVO_2933	1463.216	-0.2876969	4.17E-02	hypothetical protein
gene-HVO_RS07810	HVO_0646	1359.436	-0.2957235	7.44E-03	hypothetical protein
gene-HVO_RS06910	HVO_0458A	4991.983	-0.3017721	2.20E-02	hypothetical protein
gene-HVO_RS16760	HVO_2500	4366.071	-0.3072199	1.22E-02	amino acid permease
gene-HVO_RS10130	HVO_1130	16175.298	-0.3114894	4.23E-02	hypothetical protein
gene-HVO_RS08125	HVO_0712	2406.002	-0.3126028	3.16E-02	shikimate dehydrogenase
gene-HVO_RS14315	HVO_2002	2251.12	-0.3184464	4.57E-02	winged helix-turn-helix domain-containing protein
gene-HVO_RS03325	HVO_A0329	2482.77	-0.319785	7.02E-03	bifunctional 4-hydroxy-2-oxoglutarate aldolase/2-dehydro-3-deoxy-phosphogluconate aldolase
gene-HVO_RS10810	HVO_1271	38656.445	-0.3201059	1.29E-02	HalX domain-containing protein
gene-HVO_RS05620	HVO_0189	2559.046	-0.3218942	1.30E-02	histidine phosphatase family protein
gene-HVO_RS13425	HVO_1817	2909.871	-0.3244848	4.26E-02	hypothetical protein
gene-HVO_RS03320	HVO_A0328	1064.267	-0.325188	2.77E-03	2-dehydro-3-deoxygalactono kinase KdgK2
gene-HVO_RS07190	HVO_0514	1080.461	-0.3280555	4.63E-02	DUF309 domain-containing protein
gene-HVO_RS09275	HVO_0955	1305.313	-0.3294423	1.45E-02	hypothetical protein
gene-HVO_RS17600	HVO_2671	9593.12	-0.3321889	2.57E-02	alanine--glyoxylate aminotransferase family protein
gene-HVO_RS18895	HVO_2934	1543.421	-0.3344095	4.07E-02	hypothetical protein
gene-HVO_RS14285	HVO_1996	1004.145	-0.3361063	4.47E-02	NUDIX hydrolase

Table 17. A vs D2*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15590	HVO_2266	3883.9596	15.80334	8.50E-16	hypothetical protein
gene-HVO_RS15575	HVO_2263	3887.6862	15.6403	3.29E-18	winged helix-turn-helix transcriptional regulator
gene-HVO_RS21315	HVO_2286	2469.3796	14.82625	1.57E-14	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.63644	2.31E-18	type I restriction-modification system subunit M
gene-HVO_RS15640	HVO_2276	1777.7261	14.62019	8.65E-16	hypothetical protein
gene-HVO_RS15605	HVO_2269	1716.518	14.45272	2.31E-18	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS20265	HVO_2271	1548.562	14.27571	8.88E-13	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	14.14987	1.85E-13	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.84931	1.35E-13	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.40012	9.17E-16	N-6 DNA methylase
gene-HVO_RS15725	HVO_2291	566.281	12.87036	1.62E-12	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.79614	1.42E-15	TATA-box-binding protein
gene-HVO_RS20605	HVO_2253	488.4749	12.50309	1.03E-13	hypothetical protein
gene-HVO_RS15540	HVO_2255	391.8697	12.30259	1.85E-13	hypothetical protein
gene-HVO_RS15625	HVO_2273	399.0095	12.18288	3.77E-13	tyrosine-type recombinase/integrase
gene-HVO_RS15555	HVO_2259	459.3064	12.1034	5.07E-13	tyrosine-type recombinase/integrase
gene-HVO_RS15710	HVO_2288	340.4556	12.08041	2.50E-13	DUF2254 domain-containing protein
gene-HVO_RS15720	HVO_2290	328.8174	11.99247	6.09E-10	tyrosine-type recombinase/integrase
gene-HVO_RS15630	HVO_2274A	355.3856	11.98697	8.64E-14	hypothetical protein
gene-HVO_RS15560	HVO_2260	305.3267	11.77479	1.62E-12	hypothetical protein

Table 18. A vs D2*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS07225	HVO_0522	3916.7485	-0.2374146	4.65E-02	histone deacetylase
gene-HVO_RS19270	HVO_3010	52874.5837	-0.2746585	8.02E-03	DEAD/DEAH box helicase
gene-HVO_RS13850	HVO_1907	24534.5828	-0.283601	3.61E-02	ATP-binding protein
gene-HVO_RS19265	HVO_3009	4311.2112	-0.2917722	4.32E-02	universal stress protein
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.3020931	1.40E-02	2-dehydro-3-deoxygalactonokinase KdgK2
gene-HVO_RS06860	HVO_0449	3911.4596	-0.3055819	4.30E-02	prephenate dehydratase
gene-HVO_RS04885	HVO_0038	3934.0812	-0.3210806	2.03E-02	cytochrome P450
gene-HVO_RS15395	HVO_2225	5122.953	-0.3235983	6.94E-03	succinylglutamate desuccinylase/aspartoacylase family protein
gene-HVO_RS08910	HVO_0878	2089.7419	-0.3246281	1.95E-02	endonuclease III
gene-HVO_RS01370	HVO_B0283	1187.2593	-0.3463605	1.38E-03	helix-turn-helix domain-containing protein
gene-HVO_RS11275	HVO_1365	759.4844	-0.3515761	4.39E-02	MBL fold metallo-hydrolase
gene-HVO_RS05055	HVO_0073	2881.9596	-0.3535786	3.01E-02	DHH family phosphoesterase
gene-HVO_RS08330	HVO_0756	2981.2794	-0.3687477	1.62E-02	DHH family phosphoesterase
gene-HVO_RS17835	HVO_2718	5512.1905	-0.3710043	3.02E-02	multiprotein-bridging factor 1 family protein
gene-HVO_RS05675	HVO_0200	1441.0486	-0.3823419	2.66E-02	helix-turn-helix transcriptional regulator
gene-HVO_RS06665	HVO_0408	1768.8822	-0.3861746	6.87E-03	PrsW family intramembrane metalloprotease
gene-HVO_RS17545	HVO_2659	1736.1653	-0.3890446	2.92E-03	tautomerase family protein
gene-HVO_RS12550	HVO_1635	1945.7428	-0.3924427	4.76E-02	hypothetical protein
gene-HVO_RS18455	HVO_2843	1412.0349	-0.3932645	1.20E-02	DASH family cryptochrome
gene-HVO_RS19040	HVO_2965	2520.9009	-0.4004955	4.43E-02	phosphoserine phosphatase SerB

Table 19. B vs D1*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15575	HVO_2263	3887.6862	15.3635	4.69E-19	winged helix-turn-helix transcriptional regulator
gene-HVO_RS15590	HVO_2266	3883.9596	15.19103	7.13E-16	hypothetical protein
gene-HVO_RS21315	HVO_2286	2469.3796	14.84204	1.21E-15	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.35424	1.98E-19	type I restriction-modification system subunit M
gene-HVO_RS15605	HVO_2269	1716.518	14.19164	2.69E-19	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS15640	HVO_2276	1777.7261	14.126	5.60E-16	hypothetical protein
gene-HVO_RS20265	HVO_2271	1548.562	14.06859	2.10E-13	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.60058	1.82E-13	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.59192	3.92E-14	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	12.96396	5.86E-16	N-6 DNA methylase
gene-HVO_RS15555	HVO_2259	459.3064	12.60505	4.46E-15	tyrosine-type recombinase/integrase
gene-HVO_RS15725	HVO_2291	566.281	12.5755	7.03E-13	hypothetical protein
gene-HVO_RS20605	HVO_2253	488.4749	12.49325	9.38E-15	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.37808	1.07E-15	TATA-box-binding protein
gene-HVO_RS15625	HVO_2273	399.0095	12.22282	3.20E-14	tyrosine-type recombinase/integrase
gene-HVO_RS15540	HVO_2255	391.8697	12.07804	5.38E-14	hypothetical protein
gene-HVO_RS15630	HVO_2274A	355.3856	12.07712	4.37E-15	hypothetical protein
gene-HVO_RS15710	HVO_2288	340.4556	11.89215	6.26E-14	DUF2254 domain-containing protein
gene-HVO_RS15550	HVO_2258	257.5188	11.88369	5.17E-11	hypothetical protein
gene-HVO_RS15720	HVO_2290	328.8174	11.87389	1.14E-10	tyrosine-type recombinase/integrase

Table 20. B vs D1*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.2322728	3.40E-02	2-dehydro-3-deoxygalactonokinase KdgK2
gene-HVO_RS05130	HVO_0091	2613.3927	-0.2614819	1.41E-02	PHP domain-containing protein
gene-HVO_RS06180	HVO_0306	9897.1474	-0.2775295	4.70E-02	hypothetical protein
gene-HVO_RS11415	HVO_1393	6319.9394	-0.2836483	9.08E-03	hypothetical protein
gene-HVO_RS01765	HVO_B0365	589.5013	-0.2856043	2.85E-02	NrfD/PsrC family molybdoenzyme membrane anchor subunit
gene-HVO_RS12385	HVO_1598	7808.1357	-0.2858457	3.88E-02	DEAD/DEAH box helicase family protein
gene-HVO_RS00110	HVO_B0024	868.4152	-0.2989351	4.81E-02	ABC transporter ATP-binding protein
gene-HVO_RS09110	HVO_0921	1051.2426	-0.3055107	3.99E-02	Lrp/AsnC family transcriptional regulator
gene-HVO_RS14170	HVO_1974	3042.0278	-0.3062352	1.53E-02	hypothetical protein
gene-HVO_RS09945	HVO_1092	3800.3411	-0.3101217	4.31E-02	Rpp14/Pop5 family protein
gene-HVO_RS10440	HVO_1194	513.1632	-0.3110409	2.48E-02	branched-chain amino acid ABC transporter permease
gene-HVO_RS12785	HVO_1682	282.7099	-0.3127018	4.14E-02	mechanosensitive ion channel
gene-HVO_RS00020	HVO_B0004	1953.4585	-0.3170275	3.02E-02	serine hydroxymethyltransferase
gene-HVO_RS17545	HVO_2659	1736.1653	-0.3186961	6.94E-03	tautomerase family protein
gene-HVO_RS14330	HVO_2006	2891.979	-0.3201281	3.44E-02	PGF-CTERM sorting domain-containing protein
gene-HVO_RS16170	HVO_2382	2392.0836	-0.3202062	3.85E-02	hypothetical protein
gene-HVO_RS03060	HVO_A0274	4422.3608	-0.3208604	3.29E-02	inositol monophosphatase
gene-HVO_RS09275	HVO_0955	1305.3134	-0.3226579	1.58E-02	hypothetical protein
gene-HVO_RS08865	HVO_0868	7286.2338	-0.3245085	1.82E-03	NAD-dependent epimerase/dehydratase family protein
gene-HVO_RS10025	HVO_1108	1441.651	-0.3271989	2.53E-03	Gar1/Naf1 family protein

Table 21. B vs D2*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS15575	HVO_2263	3887.6862	15.75102	2.31E-18	winged helix-turn-helix transcriptional regulator
gene-HVO_RS15590	HVO_2266	3883.9596	15.57856	2.97E-15	hypothetical protein
gene-HVO_RS21315	HVO_2286	2469.3796	15.22957	4.47E-15	hypothetical protein
gene-HVO_RS20260	HVO_2270	1934.864	14.74176	1.55E-18	type I restriction-modification system subunit M
gene-HVO_RS15605	HVO_2269	1716.518	14.57917	1.55E-18	type I restriction-modification enzyme R subunit C-terminal domain-containing protein
gene-HVO_RS15640	HVO_2276	1777.7261	14.51353	2.97E-15	hypothetical protein
gene-HVO_RS20265	HVO_2271	1548.562	14.45611	5.24E-13	restriction endonuclease subunit S
gene-HVO_RS15585	HVO_2265	1260.0183	13.9881	4.39E-13	hypothetical protein
gene-HVO_RS15595	HVO_2267	1131.3428	13.97945	1.03E-13	hypothetical protein
gene-HVO_RS15635	HVO_2275	778.1596	13.35149	2.97E-15	N-6 DNA methylase
gene-HVO_RS15555	HVO_2259	459.3064	12.99258	1.16E-14	tyrosine-type recombinase/integrase
gene-HVO_RS15725	HVO_2291	566.281	12.96303	1.42E-12	hypothetical protein
gene-HVO_RS20605	HVO_2253	488.4749	12.88077	2.08E-14	hypothetical protein
gene-HVO_RS15600	HVO_2268	515.1045	12.76561	2.97E-15	TATA-box-binding protein
gene-HVO_RS15625	HVO_2273	399.0095	12.61035	7.54E-14	tyrosine-type recombinase/integrase
gene-HVO_RS15540	HVO_2255	391.8697	12.46557	1.11E-13	hypothetical protein
gene-HVO_RS15630	HVO_2274A	355.3856	12.46464	1.13E-14	hypothetical protein
gene-HVO_RS15710	HVO_2288	340.4556	12.27968	1.21E-13	DUF2254 domain-containing protein
gene-HVO_RS15550	HVO_2258	257.5188	12.27121	1.29E-10	hypothetical protein
gene-HVO_RS15720	HVO_2290	328.8174	12.26141	2.86E-10	tyrosine-type recombinase/integrase

Table 22. B vs D2*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS08865	HVO_0868	7286.234	-0.2672816	4.24E-02	NAD-dependent epimerase/dehydratase family protein
gene-HVO_RS11415	HVO_1393	6319.939	-0.2858606	3.47E-02	hypothetical protein
gene-HVO_RS15395	HVO_2225	5122.953	-0.3033651	1.93E-02	succinylglutamate desuccinylase/aspartoacylase family protein
gene-HVO_RS07225	HVO_0522	3916.748	-0.3358373	4.89E-03	histone deacetylase
gene-HVO_RS16095	HVO_2367	4044.539	-0.3407745	2.23E-02	ATP-binding protein
gene-HVO_RS16280	HVO_2405	1999.474	-0.3439608	8.10E-03	energy-coupling factor ABC transporter ATP-binding protein
gene-HVO_RS08225	HVO_0734	7933.454	-0.3473205	2.44E-02	DNA repair protein NreA
gene-HVO_RS05290	HVO_0123	23883.872	-0.3521249	5.12E-03	signal recognition particle protein Srp54
gene-HVO_RS10405	HVO_1187	13414.29	-0.3653845	3.65E-02	SRPBCC family protein
gene-HVO_RS11435	HVO_1396	17704.129	-0.3816101	3.78E-02	geranylgeranyl reductase family protein
gene-HVO_RS13850	HVO_1907	24534.583	-0.3835032	4.92E-03	ATP-binding protein
gene-HVO_RS17470	HVO_2645	1642.595	-0.3837542	1.34E-02	molybdopterin-binding protein
gene-HVO_RS12505	HVO_1625	7389.244	-0.3855253	4.36E-02	hypothetical protein
gene-HVO_RS08860	HVO_0867	8498.744	-0.3858965	2.24E-02	HD domain-containing protein
gene-HVO_RS03060	HVO_A0274	4422.361	-0.3927192	3.37E-02	inositol monophosphatase
gene-HVO_RS18715	HVO_2897	1784.11	-0.3933328	9.27E-03	hypothetical protein
gene-HVO_RS02865	HVO_A0229	3007.124	-0.3982051	4.97E-02	CBS domain-containing protein
gene-HVO_RS08330	HVO_0756	2981.279	-0.4024114	1.27E-02	DHH family phosphoesterase
gene-HVO_RS14040	HVO_1947	1749.466	-0.4071075	1.81E-02	rhodanese-like domain-containing protein
gene-HVO_RS16600	HVO_2468	1480.679	-0.4119965	1.48E-03	SDR family oxidoreductase

Table 23. C vs D1*: Top 20* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS00725	HVO_B0150	1187.35	5.179913	8.02E-21	ABC transporter substrate-binding protein
gene-HVO_RS20495	HVO_A0350	47026.7129	5.107392	2.32E-06	hypothetical protein
gene-HVO_RS20405	HVO_C0054	40058.2032	4.99361	1.43E-12	hypothetical protein
gene-HVO_RS02450	HVO_A0129	2076.8472	4.816274	7.91E-19	hypothetical protein
gene-HVO_RS19800	HVO_A0130	2249.5333	4.815721	1.39E-16	hypothetical protein
gene-HVO_RS19930	HVO_A0351	48151.7592	4.784673	7.42E-06	hypothetical protein
gene-HVO_RS10230	HVO_1150	22006.5636	4.774668	1.38E-19	IS200/IS605-like element ISHal1 family transposase
gene-HVO_RS17715	HVO_2695	51223.0837	4.75242	2.03E-10	substrate-binding domain-containing protein
gene-HVO_RS02255	HVO_A0084	1883.2822	4.744791	2.19E-13	amidohydrolase family protein
gene-HVO_RS20540	HVO_0379B	928.7626	4.694157	6.82E-21	hypothetical protein
gene-HVO_RS02345	HVO_A0101	360.9921	4.68859	3.15E-32	hypothetical protein
gene-HVO_RS19975	HVO_A0406	5319.8179	4.583627	1.52E-14	PH domain-containing protein
gene-HVO_RS08920	HVO_0880	157603.9167	4.540431	2.77E-17	coiled-coil protein
gene-HVO_RS20385	HVO_C0040	12306.9689	4.511192	1.14E-18	DNA cytosine methyltransferase
gene-HVO_RS02405	HVO_A0118	618.1256	4.491746	1.57E-17	PadR family transcriptional regulator
gene-HVO_RS03700	HVO_A0416	1010.0366	4.460221	4.64E-17	PadR family transcriptional regulator
gene-HVO_RS03880	HVO_A0459	679.2132	4.426388	1.02E-25	hypothetical protein
gene-HVO_RS02250	HVO_A0083	734.545	4.41255	2.80E-24	Rieske 2Fe-2S domain-containing protein
gene-HVO_RS19590	HVO_C0076	1290.0264	4.395833	2.86E-17	IclR family transcriptional regulator
gene-HVO_RS02480	HVO_A0140	2548.7647	4.389428	1.43E-33	HD family hydrolase

Table 24. C vs D1*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS08370	HVO_0765	2300.517	-0.2382795	4.60E-02	FAD-binding oxidoreductase
gene-HVO_RS05130	HVO_0091	2613.393	-0.2435531	2.39E-02	PHP domain-containing protein
gene-HVO_RS10915	HVO_1293	3718.466	-0.2443505	4.12E-02	gluconate 2-dehydrogenase subunit 3 family protein
gene-HVO_RS13950	HVO_1930A	4191.74	-0.2516234	1.23E-02	hypothetical protein
gene-HVO_RS19265	HVO_3009	4311.211	-0.2591863	4.08E-02	universal stress protein
gene-HVO_RS10025	HVO_1108	1441.651	-0.2618766	1.78E-02	Gar1/Naf1 family protein
gene-HVO_RS06665	HVO_0408	1768.882	-0.2699992	3.69E-02	PrsW family intramembrane metalloprotease
gene-HVO_RS08225	HVO_0734	7933.454	-0.2850703	2.55E-02	DNA repair protein NreA
gene-HVO_RS06860	HVO_0449	3911.46	-0.291763	2.59E-02	prephenate dehydratase
gene-HVO_RS19045	HVO_2966	2298.343	-0.2925986	3.12E-02	hypothetical protein
gene-HVO_RS05560	HVO_0175	32030.572	-0.29482	3.04E-02	DNA polymerase sliding clamp
gene-HVO_RS11035	HVO_1314	1911.694	-0.3099524	4.05E-02	5'/3'-nucleotidase SurE
gene-HVO_RS07855	HVO_0656	4117.447	-0.312861	5.21E-03	DUF2103 domain-containing protein
gene-HVO_RS13425	HVO_1817	2909.871	-0.31505	4.90E-02	hypothetical protein
gene-HVO_RS18575	HVO_2870	4512.973	-0.3193616	1.95E-02	AI-2E family transporter
gene-HVO_RS13580	HVO_1850	1425.722	-0.3199198	4.18E-02	hypothetical protein
gene-HVO_RS00080	HVO_B0018	1412.841	-0.3202654	2.19E-02	ParA family protein
gene-HVO_RS00020	HVO_B0004	1953.458	-0.3203139	2.97E-02	serine hydroxymethyltransferase
gene-HVO_RS08270	HVO_0744	2531.916	-0.3209594	1.19E-02	glutamate-cysteine ligase family protein
gene-HVO_RS09180	HVO_0936	1419.779	-0.327494	4.27E-02	selenium cofactor biosynthesis protein YqeC

Table 25. C vs D2*: Top nine* ***UP****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS19535	HVO_C0064	393.8631	0.9117181	1.56E-04	IS1595-like element ISHvo5 family transposase
gene-HVO_RS08245	HVO_0739	8675.9947	0.8785012	1.78E-02	hypothetical protein
gene-HVO_RS11415	HVO_1393	6319.9394	0.6203153	1.24E-07	hypothetical protein
gene-HVO_RS06625	HVO_0399	7732.9737	0.6025997	6.43E-06	hypothetical protein
gene-HVO_RS00110	HVO_B0024	868.4152	0.5117512	2.04E-02	ABC transporter ATP-binding protein
gene-HVO_RS11420	HVO_1394	35188.1374	0.5112158	3.15E-05	Na+/H+ antiporter NhaC family protein
gene-HVO_RS06915	HVO_0459	9555.4721	0.4533178	2.64E-02	Rid family detoxifying hydrolase
gene-HVO_RS16280	HVO_2405	1999.4742	0.3937117	8.89E-03	energy-coupling factor ABC transporter ATP-binding protein
gene-HVO_RS10025	HVO_1108	1441.651	0.3870088	1.36E-02	Gar1/Naf1 family protein

Table 26. C vs D2*: Top 20* ***DOWN****-regulated genes (significantly changed only, sorted by LFC).*

Gene ID	Old Locus Tag	baseMean	log2FoldChange	padj	Product
gene-HVO_RS03320	HVO_A0328	1064.2671	-0.4301068	1.93E-03	2-dehydro-3-deoxygalactonokinase KdgK2
gene-HVO_RS01945	HVO_A0019	359.8168	-0.5114361	1.68E-02	hypothetical protein
gene-HVO_RS19270	HVO_3010	52874.5837	-0.5817655	1.97E-09	DEAD/DEAH box helicase
gene-HVO_RS19310	HVO_C0005	3990.8852	-0.5846777	1.34E-03	hypothetical protein
gene-HVO_RS19265	HVO_3009	4311.2112	-0.5908032	3.61E-05	universal stress protein
gene-HVO_RS14795	HVO_2101	774.6513	-0.5982561	4.49E-02	HPr family phosphocarrier protein
gene-HVO_RS12250	HVO_1572	25145.5314	-0.6500052	7.24E-03	DNA topoisomerase (ATP-hydrolyzing) subunit B
gene-HVO_RS08790	HVO_0854	26899.9065	-0.6579891	1.67E-03	DNA double-strand break repair ATPase Rad50
gene-HVO_RS08785	HVO_0853	9401.2133	-0.738609	5.54E-09	DNA double-strand break repair protein Mre11
gene-HVO_RS00410	HVO_B0085	1127.5755	-0.7905662	5.99E-03	glycosyl hydrolase family 28 protein
gene-HVO_RS12235	HVO_1569	1406.3037	-0.7987373	8.15E-09	hypothetical protein
gene-HVO_RS03960	HVO_A0475	59658.7297	-0.8335661	1.56E-02	superoxide dismutase
gene-HVO_RS21370	HVO_C0014	1668.0092	-0.8924065	5.72E-03	hypothetical protein
gene-HVO_RS12240	HVO_1570	10385.9128	-0.9089378	1.56E-04	DNA topoisomerase IV subunit A
gene-HVO_RS12245	HVO_1571	24202.3135	-1.0010354	1.02E-06	DNA topoisomerase VI subunit B
gene-HVO_RS04380	HVO_A0566	153.2093	-1.0201077	4.18E-02	3-isopropylmalate dehydratase small subunit
gene-HVO_RS07155	HVO_0507	5534.2317	-1.0239877	7.35E-06	creatininase family protein
gene-HVO_RS03955	HVO_A0474	993.5913	-1.04796	5.88E-03	hypothetical protein
gene-HVO_RS14890	HVO_2120	280.5482	-1.176027	2.04E-02	sialidase family protein
gene-HVO_RS00620	HVO_B0128	337.035	-1.182405	1.92E-03	SDR family oxidoreductase

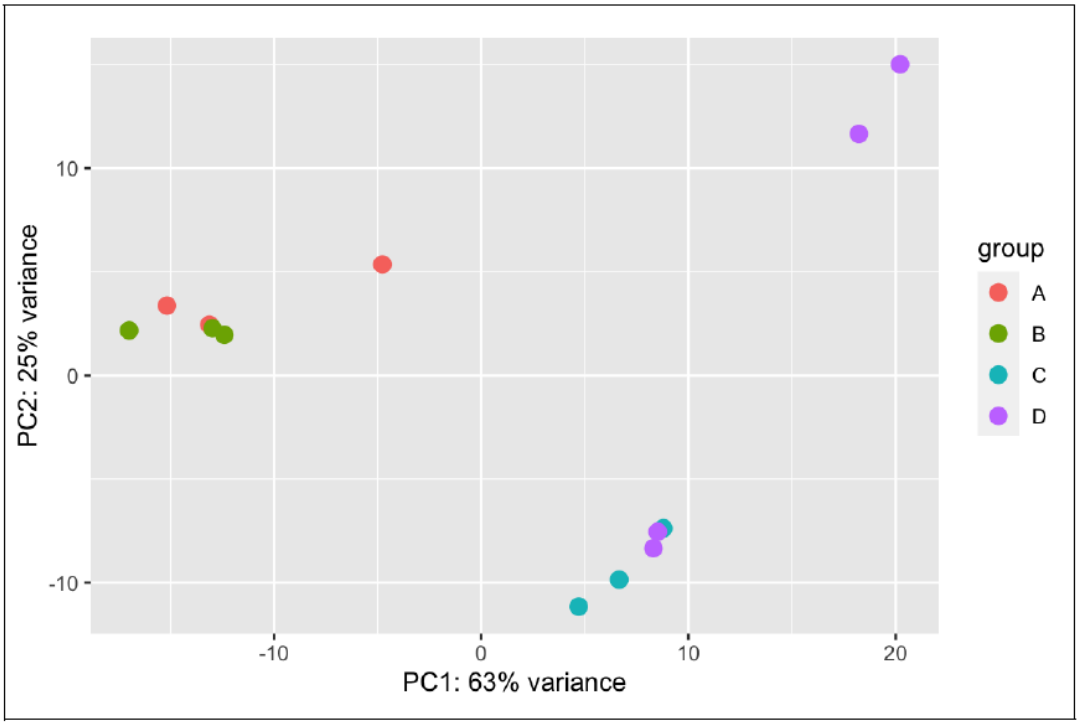


Figure 3. PCA plot showing extent of clustering of replicates and separation of treatments along the first two principal components. The two 'D' samples isolated from any other cluster were assigned as sub-group D1 and those 'D' samples clustering group C were assigned as sub-group D2 whilst those D groups.

Appendix II: PIP Placement Reflective Statement

Note to examiners:

This statement is included as an appendix to the thesis in order that the thesis accurately captures the PhD training experienced by the candidate as a BBSRC Doctoral Training Partnership student. The Professional Internship for PhD Students is a compulsory 3-month placement which must be undertaken by DTP students. It is usually centred on a specific project and must not be related to the PhD project. This reflective statement is designed to capture the skills development which has taken place during the student's placement and the impact it had on their career plans.

PIP Reflective Statement

During my PhD I've always enjoyed interacting with undergraduate students and helping deliver theory as part of my duties as a demonstrator. For my PIP placement I decided to explore teaching at secondary level. My placement was at Harrow School, UK. During my placement I co-taught year 9, year 10 and Lower-sixth classes whilst also assisting individuals in other year groups as required. In addition to this I ran a bespoke course on Bioinformatics as part of the school's 'Supercurriculum' which went down well with both students and staff. I helped coach various sports teams, including rugby and squash and partook in weekly parades and training packages as part of the Harrow Rifle Corps (CCF). I undertook academic and pastoral tutoring duties in one of the boarding houses, as well as acted as an academic mentor as part of the academic scholar's symposium orchestrated with Notting Hill and Ealing High School. During this time I developed my scientific communication skills, varying my explanations of scientific concepts to differing degrees of complexity depending on the academic capabilities of the audience. I enjoyed the challenge of monitoring my time-management effectively, often juggling many metaphorical plates at once, as well as gaining insight into the inner workings of one of the top schools in the country.

Appendix III: SARS-CoV-2 Impact Statement

Prior to my statement showing the impact COVID-19 had on my doctoral studies, I would like to take the opportunity to think of the profound impact this disease had on individuals, families and communities worldwide. As I reflect on the significance of this global crisis, my thoughts are with all those who have lost loved ones and been affected by the far-reaching consequences of the pandemic. Within the context of my PhD, the pandemic has introduced unprecedented challenges that have substantially affected the trajectory of my work. The Allers lab closed on the 10th March 2020, resulting in a sudden halt to all on-going laboratory work. The lab eventually reopened on the 29th July 2020. However, when the lab closed I had not even settled on a PhD project title, and was only a week in to my final 'lab rotation' as part of the doctoral training programme (DTP). During this time away from the lab, I decided on a project title and began writing my literature review, which has formed the basis of chapter 1 of this thesis. Although a useful way to spend my time, this accounted for between six and seven months of lost lab work time. Students above me in the DTP were allowed back into the lab in July 2020, as priority was given to these students – despite us all losing the same amount of 'lab hours'. Given I had no project at this stage, and certainly no data, I was not able to perform any data analysis at all. Once the lab reopened, this was on a split-shift pattern to allow for social distancing. I was assigned to the morning shift which ran from 0730hrs – 1230hrs. Given I lived an hour from the lab, this was even more tricky. My ability to get experiments done was compromised by being in the morning shift as there was no option to stay beyond the allocated time, and moving to the afternoon shift was not possible due to the 'bubbles' created by the initiative. Every other Friday was a full day for one shift, which helped a little, but social distancing meant I did not receive any training on laboratory techniques and was left to train myself on many techniques. Split shift patterns caused further issues as I was using the Cs-137 source housed in a different area of campus. The time required to irradiate meant that this could only ever be done every two weeks on the Friday when our shift was allocated to that day. This was also true for many other assays I was undertaking at the time, such as competition assays which run over multiple hours and microscopy. Supervision meetings via MS Teams also made this tricky. Resources were scarce and orders of materials had severe time delays which impacted my ability to perform various experiments. A large hurdle to overcome was the lack of radioactivity; radioactive work did not result until late January 2021. *Haloferax* genetics requires use of radiation to confirm strain generation due to their polyploid nature, so this was difficult to overcome.

However, with time I was able to adapt to this way of working before the normal working patterns were established a year later. Further complications arose as I was part of two lab groups. The lab I based myself in initially closed a number of times for PAT testing and HSE inspections, as we shared the lab space with a group working on category II pathogens. Eventually I was able to move into a different lab space.

In total, roughly 1 years' worth of lab time has been lost. Given the challenges I've faced, I'm rather proud to be submitting this thesis.

References

Abou-Zied, O. K., Jimenez, R. and Romesberg, F. E. (2001) 'Tautomerization Dynamics of a Model Base Pair in DNA', *Journal of the American Chemical Society*. American Chemical Society, 123(19), pp. 4613–4614. doi: 10.1021/ja003647s.

Acharya, S. *et al.* (2003) 'The Coordinated Functions of the E. coli MutS and MutL Proteins in Mismatch Repair', *Molecular Cell*. Elsevier, 12(1), pp. 233–246. doi: 10.1016/S1097-2765(03)00219-3.

Ahel, I. *et al.* (2006) 'The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates', *Nature*, 443(7112), pp. 713–716. doi: 10.1038/nature05164.

Al-Minawi, A. Z., Saleh-Gohari, N. and Helleday, T. (2008) 'The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells.', *Nucleic acids research*, 36(1), pp. 1–9. doi: 10.1093/nar/gkm888.

Alarcón-Schumacher, T. *et al.* (2022) 'Isolation of a virus causing a chronic infection in the archaeal model organism Haloferax volcanii reveals antiviral activities of a provirus', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 119(35), p. e2205037119. doi: 10.1073/pnas.2205037119.

Albers, S.-V. and Meyer, B. H. (2011) 'The archaeal cell envelope', *Nature Reviews Microbiology*, 9(6), pp. 414–426. doi: 10.1038/nrmicro2576.

Alexeev, A., Mazin, A. and Kowalczykowski, S. C. (2003) 'Rad54 protein possesses chromatin-remodeling activity stimulated by the Rad51–ssDNA nucleoprotein filament', *Nature Structural & Molecular Biology*, 10(3), pp. 182–186. doi: 10.1038/nsb901.

Allen, J. M. *et al.* (2011) 'Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication.', *Nucleic acids research*, 39(16), pp. 7020–7033. doi: 10.1093/nar/gkr157.

Allers, T., Ngo, H.-P., *et al.* (2004) 'Development of Additional Selectable Markers for the Halophilic Archaeon Haloferax volcanii Based on the leuB and trpA Genes', *Applied and Environmental Microbiology*, 70(2), p. 943 LP-953. doi: 10.1128/AEM.70.2.943-953.2004.

Allers, T., Ngo, H., *et al.* (2004) 'Development of Additional Selectable Markers for the Halophilic Archaeon Haloferax volcanii Based on the leuB and trpA Genes', 70(2). doi: 10.1128/AEM.70.2.943.

Allers, T. *et al.* (2010a) 'Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in Haloferax volcanii', *Applied and environmental microbiology*. 2010/01/22. American Society for Microbiology (ASM), 76(6), pp. 1759–1769. doi: 10.1128/AEM.02670-09.

Allers, T. *et al.* (2010b) 'Improved Strains and Plasmid Vectors for Conditional

Overexpression of His-Tagged Proteins in Haloferax volcanii ', 76(6), pp. 1759–1769. doi: 10.1128/AEM.02670-09.

Allers, T. (2010) 'Overexpression and purification of halophilic proteins in Haloferax volcanii', *Bioengineered bugs*. 2010/03/17. Landes Bioscience, 1(4), pp. 288–290. doi: 10.4161/bbug.1.4.11794.

Allers, T. and Lichten, M. (2001) 'Differential timing and control of noncrossover and crossover recombination during meiosis.', *Cell*. United States, 106(1), pp. 47–57. doi: 10.1016/s0092-8674(01)00416-0.

Allers, T. and Mevarech, M. (2005) 'Archaeal genetics - the third way.', *Nature reviews. Genetics*. England, 6(1), pp. 58–73. doi: 10.1038/nrg1504.

Ambily Nath, I. V and Loka Bharathi, P. A. (2011) 'Diversity in transcripts and translational pattern of stress proteins in marine extremophiles.', *Extremophiles : life under extreme conditions*. Germany, 15(2), pp. 129–153. doi: 10.1007/s00792-010-0348-x.

Amundsen, S. K. *et al.* (2007) 'Intersubunit signaling in RecBCD enzyme, a complex protein machine regulated by Chi hot spots', *Genes & development*. Cold Spring Harbor Laboratory Press, 21(24), pp. 3296–3307. doi: 10.1101/gad.1605807.

Amundsen, S. K., Sharp, J. W. and Smith, G. R. (2016) 'RecBCD Enzyme "Chi Recognition" Mutants Recognize Chi Recombination Hotspots in the Right DNA Context', *Genetics*. 2016/07/08. Genetics Society of America, 204(1), pp. 139–152. doi: 10.1534/genetics.116.191056.

Anand, R. *et al.* (2016) 'Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection.', *Molecular cell*. United States, 64(5), pp. 940–950. doi: 10.1016/j.molcel.2016.10.017.

Anand, R. *et al.* (2019) 'NBS1 promotes the endonuclease activity of the MRE11-RAD50 complex by sensing CtIP phosphorylation.', *The EMBO journal*, 38(7). doi: 10.15252/embj.2018101005.

Anderson, D. G. and Kowalczykowski, S. C. (1997a) 'The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme.', *Genes & development*. United States, 11(5), pp. 571–581. doi: 10.1101/gad.11.5.571.

Anderson, D. G. and Kowalczykowski, S. C. (1997b) 'The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner.', *Cell*. United States, 90(1), pp. 77–86. doi: 10.1016/s0092-8674(00)80315-3.

Antón, J. (2011) 'Halophile', in Gargaud, M. et al. (eds) *Encyclopedia of Astrobiology*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 725–727. doi: 10.1007/978-3-642-11274-4_694.

Antony, E. and Hingorani, M. M. (2004) 'Asymmetric ATP binding and hydrolysis activity of the Thermus aquaticus MutS dimer is key to modulation of its interactions with mismatched DNA', *Biochemistry*, 43(41), pp. 13115–13128. doi: 10.1021/bi049010t.

Ashiuchi, M. and Misono, H. (2002) 'Biochemistry and molecular genetics of poly-gamma-

glutamate synthesis.', *Applied microbiology and biotechnology*. Germany, 59(1), pp. 9–14. doi: 10.1007/s00253-002-0984-x.

Azeroglu, B. *et al.* (2016) 'RecG Directs DNA Synthesis during Double-Strand Break Repair', *PLOS Genetics*. Public Library of Science, 12(2), p. e1005799. Available at: https://doi.org/10.1371/journal.pgen.1005799.

Barman, D., Jha, D. and Bhattacharjee, K. (2020) 'Metallotolerant Bacteria: Insights into Bacteria Thriving in Metal-Contaminated Areas', in, p. pp 135-164. doi: 10.1007/978-981-15-3028-9_9.

Barras, F. and Marinus, M. G. (1989) 'The great GATC: DNA methylation in E. coli', *Trends in Genetics*, 5, pp. 139–143. doi: https://doi.org/10.1016/0168-9525(89)90054-1.

Barry, E. R. and Bell, S. D. (2006) 'DNA replication in the archaea', *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology, 70(4), pp. 876–887. doi: 10.1128/MMBR.00029-06.

Barzel, A. and Kupiec, M. (2008) 'Finding a match: how do homologous sequences get together for recombination?', *Nature Reviews Genetics*, 9(1), pp. 27–37. doi: 10.1038/nrg2224.

Bell, J. C. *et al.* (2012) 'Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA', *Nature*. 2012/10/24, 491(7423), pp. 274–278. doi: 10.1038/nature11598.

Bell, S. D. (2022) 'Form and function of archaeal genomes.', *Biochemical Society transactions*. England, 50(6), pp. 1931–1939. doi: 10.1042/BST20221396.

Benison, K. C. and Karmanocky III, F. J. (2014) 'Could microorganisms be preserved in Mars gypsum? Insights from terrestrial examples', *Geology*, 42(7), pp. 615–618. doi: 10.1130/G35542.1.

Berenguer, J. (2011) 'Thermophile', in Gargaud, M. et al. (eds) *Encyclopedia of Astrobiology*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1666–1667. doi: 10.1007/978-3-642-11274-4_1583.

Bergink, S. *et al.* (2012) 'Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex.', *The Journal of cell biology*, 196(6), pp. 681–688. doi: 10.1083/jcb.201107050.

Bielski, C. M. *et al.* (2018) 'Genome doubling shapes the evolution and prognosis of advanced cancers.', *Nature genetics*. United States, 50(8), pp. 1189–1195. doi: 10.1038/s41588-018-0165-1.

Bignon, C., Gruet, A. and Longhi, S. (2022) 'Split-GFP Reassembly Assay: Strengths and Caveats from a Multiparametric Analysis.', *International journal of molecular sciences*. Switzerland, 23(21). doi: 10.3390/ijms232113167.

Bishop, A. J. R. and Schiestl, R. H. (2000) 'Homologous recombination as a mechanism for genome rearrangements: environmental and genetic effects', *Human Molecular Genetics*,

9(16), pp. 2334–2427. doi: 10.1093/hmg/9.16.2427.

Bitan-Banin, G., Ortenberg, R. and Mevarech, M. (2003) 'Development of a gene knockout system for the halophilic archaeon Haloferax volcanii by use of the pyrE gene.', *Journal of bacteriology*, 185(3), pp. 772–778. doi: 10.1128/jb.185.3.772-778.2003.

Boiteux, S., O'Connor, T. R. and Laval, J. (1987) 'Formamidopyrimidine-DNA glycosylase of Escherichia coli: cloning and sequencing of the fpg structural gene and overproduction of the protein', *The EMBO journal*, 6(10), pp. 3177–3183. Available at: https://pubmed.ncbi.nlm.nih.gov/3319582.

Bolhuis, H. *et al.* (2006) 'The genome of the square archaeon Haloquadratum walsbyi : life at the limits of water activity', *BMC genomics*. BioMed Central, 7, p. 169. doi: 10.1186/1471-2164-7-169.

Bolt, E. L., Lloyd, R. G. and Sharples, G. J. (2001) 'Genetic analysis of an archaeal Holliday junction resolvase in Escherichia coli.', *Journal of molecular biology*. England, 310(3), pp. 577–589. doi: 10.1006/jmbi.2001.4791.

Bonura, T., Schultz, R. and Friedberg, E. C. (1982) 'An enzyme activity from Escherichia coli that attacks single-stranded deoxyribopolymers and single-stranded deoxyribonucleic acid containing apyrimidinic sites.', *Biochemistry*. United States, 21(10), pp. 2548–2556. doi: 10.1021/bi00539a039.

Borrego-Soto, G., Ortiz-López, R. and Rojas-Martínez, A. (2015) 'Ionizing radiation-induced DNA injury and damage detection in patients with breast cancer', *Genetics and molecular biology*. 2015/11/24. Sociedade Brasileira de Genética, 38(4), pp. 420–432. doi: 10.1590/S1415-475738420150019.

Bose, S. N. and Davies, R. J. H. (1984) 'The photoreactivity of T-A sequences in oligodeoxyribonucieotides and DNA', *Nucleic Acids Research*, 12(20), pp. 7903–7914. doi: 10.1093/nar/12.20.7903.

Boyd, J. B. *et al.* (1981) 'Third-chromosome mutagen-sensitive mutants of Drosophila melanogaster', *Genetics*, 97(3–4), pp. 607–623. doi: 10.1093/genetics/97.3-4.607.

Bramhill, D. and Kornberg, A. (1988) 'Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the E. coli chromosome.', *Cell*. United States, 52(5), pp. 743–755. doi: 10.1016/0092-8674(88)90412-6.

Braña, M. F. *et al.* (2001) 'Intercalators as anticancer drugs.', *Current pharmaceutical design*. United Arab Emirates, 7(17), pp. 1745–1780. doi: 10.2174/1381612013397113.

Brasser, R. and Mojzsis, S. J. (2017) 'A colossal impact enriched Mars' mantle with noble metals', *Geophysical Research Letters*. John Wiley & Sons, Ltd, 44(12), pp. 5978–5985. doi: 10.1002/2017GL074002.

Breimer, L. H. (1990) 'Molecular mechanisms of oxygen radical carcinogenesis and mutagenesis: the role of DNA base damage.', *Molecular carcinogenesis*. United States, 3(4), pp. 188–197. doi: 10.1002/mc.2940030405.

Breuert, S. *et al.* (2006) 'Regulated Polyploidy in Halophilic Archaea', *PLOS ONE*. Public Library of Science, 1(1), p. e92. Available at: https://doi.org/10.1371/journal.pone.0000092.

Brill, S. J. (2013) 'Linking the Enzymes that Unlink DNA', *Molecular Cell*. Elsevier, 52(2), pp. 159–160. doi: 10.1016/j.molcel.2013.10.007.

Bruner, S. D., Norman, D. P. and Verdine, G. L. (2000) 'Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA.', *Nature*. England, 403(6772), pp. 859–866. doi: 10.1038/35002510.

Burdett, V. *et al.* (2001) 'In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair.', *Proceedings of the National Academy of Sciences of the United States of America*, 98(12), pp. 6765–6770. doi: 10.1073/pnas.121183298.

Burton, A. S. *et al.* (2012) 'Understanding prebiotic chemistry through the analysis of extraterrestrial amino acids and nucleobases in meteorites', *Chemical Society Reviews*. The Royal Society of Chemistry, 41(16), pp. 5459–5472. doi: 10.1039/C2CS35109A.

Cadet, J., Sage, E. and Douki, T. (2005) 'Ultraviolet radiation-mediated damage to cellular DNA.', *Mutation research*. Netherlands, 571(1–2), pp. 3–17. doi: 10.1016/j.mrfmmm.2004.09.012.

Le Caer, S. (2011) 'Water Radiolysis: Influence of Oxide Surfaces on H2 Production under Ionizing Radiation', *Water*, 3. doi: 10.3390/w3010235.

Cannavo, E., Reginato, G. and Cejka, P. (2019) 'Stepwise 5' DNA end-specific resection of DNA breaks by the Mre11-Rad50-Xrs2 and Sae2 nuclease ensemble', *Proceedings of the National Academy of Sciences*, 116(12), p. 5505 LP-5513. doi: 10.1073/pnas.1820157116.

Carbone, V. *et al.* (2015) 'Structure and Evolution of the Archaeal Lipid Synthesis Enzyme sn-Glycerol-1-phosphate Dehydrogenase', *The Journal of biological chemistry*. 2015/07/14. American Society for Biochemistry and Molecular Biology, 290(35), pp. 21690–21704. doi: 10.1074/jbc.M115.647461.

Caron, P. R., Kushner, S. R. and Grossman, L. (1985) 'Involvement of helicase II (uvrD gene product) and DNA polymerase I in excision mediated by the uvrABC protein complex.', *Proceedings of the National Academy of Sciences of the United States of America*, 82(15), pp. 4925–4929. doi: 10.1073/pnas.82.15.4925.

Carré, L. *et al.* (2023) 'Effects of chaotropic salts on global proteome stability in halophilic archaea: Implications for life signatures on Mars.', *Environmental microbiology*. England. doi: 10.1111/1462-2920.16451.

Castillo, F., Benmohamed, A. and Szatmari, G. (2017) 'Xer Site Specific Recombination: Double and Single Recombinase Systems.', *Frontiers in microbiology*. Switzerland, 8, p. 453. doi: 10.3389/fmicb.2017.00453.

Chaban, B., Ng, S. Y. M. and Jarrell, K. F. (2006) 'Archaeal habitats--from the extreme to the ordinary.', *Canadian journal of microbiology*. Canada, 52(2), pp. 73–116. doi: 10.1139/w05-147.

Chakravarti, D. *et al.* (1991) 'Cloning and expression in Escherichia coli of a human cDNA encoding the DNA repair protein N-methylpurine-DNA glycosylase.', *The Journal of biological chemistry*. United States, 266(24), pp. 15710–15715.

Chen, Z., Yang, H. and Pavletich, N. P. (2008) 'Mechanism of homologous recombination from the RecA–ssDNA/dsDNA structures', *Nature*, 453(7194), pp. 489–494. doi: 10.1038/nature06971.

Chien, A., Edgar, D. B. and Trela, J. M. (1976) 'Deoxyribonucleic acid polymerase from the extreme thermophile Thermus aquaticus', *Journal of bacteriology*, 127(3), pp. 1550–1557. Available at: https://pubmed.ncbi.nlm.nih.gov/8432.

Chodavarapu, S. and Kaguni, J. M. (2016) 'Replication Initiation in Bacteria', *The Enzymes*. 2016/04/20, 39, pp. 1–30. doi: 10.1016/bs.enz.2016.03.001.

Chow, K.-H. and Courcelle, J. (2007) 'RecBCD and RecJ/RecQ initiate DNA degradation on distinct substrates in UV-irradiated Escherichia coli.', *Radiation research*. United States, 168(4), pp. 499–506. doi: 10.1667/RR1033.1.

Chu, W. K. and Hickson, I. D. (2009) 'RecQ helicases: multifunctional genome caretakers.', *Nature reviews. Cancer*. England, 9(9), pp. 644–654. doi: 10.1038/nrc2682.

Clark, A. B. *et al.* (2000) 'Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes.', *The Journal of biological chemistry*. United States, 275(47), pp. 36498–36501. doi: 10.1074/jbc.C000513200.

CLARK, B. C. *et al.* (1976) 'Inorganic Analyses of Martian Surface Samples at the Viking Landing Sites', *Science*, 194(4271), p. 1283 LP-1288. doi: 10.1126/science.194.4271.1283.

Clarke, L. and Carbon, J. (1976) 'A colony bank containing synthetic Col El hybrid plasmids representative of the entire E. coli genome.', *Cell*. United States, 9(1), pp. 91–99. doi: 10.1016/0092-8674(76)90055-6.

Cline, S. W. *et al.* (1989) 'Transformation methods for halophilic archaebacteria.', *Canadian journal of microbiology*. Canada, 35(1), pp. 148–152. doi: 10.1139/m89-022.

Cockell, C. *et al.* (2005) 'Effects of a Simulated Martian UV Flux on the Cyanobacterium, Chroococcidiopsis sp. 029', *Astrobiology*, 5, pp. 127–140. doi: 10.1089/ast.2005.5.127.

Constantin, N. *et al.* (2005) 'Human mismatch repair: reconstitution of a nick-directed bidirectional reaction.', *The Journal of biological chemistry*, 280(48), pp. 39752–39761. doi: 10.1074/jbc.M509701200.

Constantinescu-Aruxandei, D. *et al.* (2016) 'Mechanism of DNA loading by the DNA repair helicase XPD.', *Nucleic acids research*, 44(6), pp. 2806–2815. doi: 10.1093/nar/gkw102.

Cooper, D. L., Lahue, R. S. and Modrich, P. (1993) 'Methyl-directed mismatch repair is bidirectional.', *The Journal of biological chemistry*. United States, 268(16), pp. 11823–11829.

Copeland, W. C. and Longley, M. J. (2008) 'DNA2 resolves expanding flap in mitochondrial base excision repair', *Molecular cell*, 32(4), pp. 457–458. doi: 10.1016/j.molcel.2008.11.007.

Corn, J. E. and Berger, J. M. (2006) 'Regulation of bacterial priming and daughter strand synthesis through helicase-primase interactions', *Nucleic acids research*. 2006/08/25. Oxford University Press, 34(15), pp. 4082–4088. doi: 10.1093/nar/gkl363.

Crameri, A. *et al.* (1996) 'Improved green fluorescent protein by molecular evolution using DNA shuffling.', *Nature biotechnology*. United States, 14(3), pp. 315–319. doi: 10.1038/nbt0396-315.

Crickard, J. B. *et al.* (2020) 'Rad54 Drives ATP Hydrolysis-Dependent DNA Sequence Alignment during Homologous Recombination', *Cell*. Elsevier. doi: 10.1016/j.cell.2020.04.056.

Crowley, D. J. *et al.* (2006) 'The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1.', *Saline systems*, 2, p. 11. doi: 10.1186/1746-1448-2-11.

Cubeddu, L. and White, M. F. (2005) 'DNA damage detection by an archaeal single-stranded DNA-binding protein.', *Journal of molecular biology*. England, 353(3), pp. 507–516. doi: 10.1016/j.jmb.2005.08.050.

D'Amours, D. and Jackson, S. P. (2002) 'The MRE11 complex: at the crossroads of DNA repair and checkpoint signalling', *Nature Reviews Molecular Cell Biology*, 3(5), pp. 317–327. doi: 10.1038/nrm805.

Daly, M. J. *et al.* (1994) 'In vivo damage and recA-dependent repair of plasmid and chromosomal DNA in the radiation-resistant bacterium Deinococcus radiodurans.', *Journal of bacteriology*. United States, 176(12), pp. 3508–3517. doi: 10.1128/jb.176.12.3508-3517.1994.

Daly, M. J. and Minton, K. W. (1996) 'An alternative pathway of recombination of chromosomal fragments precedes recA-dependent recombination in the radioresistant bacterium Deinococcus radiodurans.', *Journal of bacteriology*. United States, 178(15), pp. 4461–4471. doi: 10.1128/jb.178.15.4461-4471.1996.

Dang, B. *et al.* (2014) 'Simulated microgravity increases heavy ion radiation-induced apoptosis in human B lymphoblasts.', *Life sciences*. Netherlands, 97(2), pp. 123–128. doi: 10.1016/j.lfs.2013.12.008.

Dao, V. and Modrich, P. (1998) 'Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex.', *The Journal of biological chemistry*. United States, 273(15), pp. 9202–9207. doi: 10.1074/jbc.273.15.9202.

Davey, M. J. *et al.* (2002) 'The DnaC helicase loader is a dual ATP/ADP switch protein', *The EMBO journal*. Oxford University Press, 21(12), pp. 3148–3159. doi: 10.1093/emboj/cdf308.

David, S. S. (2005) 'Structural biology: DNA search and rescue.', *Nature*. England, pp. 569–570. doi: 10.1038/434569a.

Dawson, K. S., Freeman, K. H. and Macalady, J. L. (2012) 'Molecular characterization of core lipids from halophilic archaea grown under different salinity conditions', *Organic Geochemistry*, 48, pp. 1–8. doi: https://doi.org/10.1016/j.orggeochem.2012.04.003.

Deguchi, S. *et al.* (2011) 'Microbial growth at hyperaccelerations up to 403,627 x g.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 108(19), pp. 7997–8002. doi: 10.1073/pnas.1018027108.

Delmas, S. *et al.* (2009) 'Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon Haloferax volcanii by restraining homologous recombination.', *PLoS genetics*, 5(7), p. e1000552. doi: 10.1371/journal.pgen.1000552.

Demple, B. *et al.* (1999) 'Roles of AP Endonucleases in Repair and Genetic Stability', in, pp. 59–66. doi: 10.1007/978-1-4615-4865-2_6.

Demple, B. and Linn, S. (1982) '5,6-Saturated thymine lesions in DNA: production by ultraviolet light or hydrogen peroxide', *Nucleic acids research*, 10(12), pp. 3781–3789. doi: 10.1093/nar/10.12.3781.

Deshpande, R. A. *et al.* (2016) 'Nbs1 Converts the Human Mre11/Rad50 Nuclease Complex into an Endo/Exonuclease Machine Specific for Protein-DNA Adducts.', *Molecular cell*. United States, 64(3), pp. 593–606. doi: 10.1016/j.molcel.2016.10.010.

Desmarais, D. *et al.* (1997) '2-Sulfotrehalose, a novel osmolyte in haloalkaliphilic archaea', *Journal of bacteriology*, 179(10), pp. 3146–3153. doi: 10.1128/jb.179.10.3146-3153.1997.

DeVeaux, L. C. *et al.* (2007) 'Extremely radiation-resistant mutants of a halophilic archaeon with increased single-stranded DNA-binding protein (RPA) gene expression.', *Radiation research*. United States, 168(4), pp. 507–514. doi: 10.1667/RR0935.1.

Dewar, J. M. and Walter, J. C. (2017) 'Mechanisms of DNA replication termination', *Nature reviews. Molecular cell biology*. 2017/05/24, 18(8), pp. 507–516. doi: 10.1038/nrm.2017.42.

Dixon, D. A. and Kowalczykowski, S. C. (1993) 'The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the E. coli RecBCD enzyme.', *Cell*. United States, 73(1), pp. 87–96. doi: 10.1016/0092-8674(93)90162-j.

Dizdaroglu, M. *et al.* (2002) 'Free radical-induced damage to DNA: mechanisms and measurement.', *Free radical biology & medicine*. United States, 32(11), pp. 1102–1115. doi: 10.1016/s0891-5849(02)00826-2.

Dolganov, G. M. *et al.* (1996) 'Human Rad50 is physically associated with human Mre11: identification of a conserved multiprotein complex implicated in recombinational DNA repair.', *Molecular and cellular biology*, 16(9), pp. 4832–4841. doi: 10.1128/mcb.16.9.4832.

Donnianni, R. A. and Symington, L. S. (2013) 'Break-induced replication occurs by conservative DNA synthesis.', *Proceedings of the National Academy of Sciences of the United States of America*, 110(33), pp. 13475–13480. doi: 10.1073/pnas.1309800110.

Doolittle, W. F. (2020) 'Evolution: Two Domains of Life or Three?', *Current biology : CB*. England, 30(4), pp. R177–R179. doi: 10.1016/j.cub.2020.01.010.

Dorazi, R. *et al.* (2007) 'Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus.', *Molecular microbiology*. England, 63(2), pp. 521–529. doi: 10.1111/j.1365-2958.2006.05516.x.

Douki, T. *et al.* (2000) 'Formation of the main UV-induced thymine dimeric lesions within isolated and cellular DNA as measured by high performance liquid chromatography-tandem mass spectrometry.', *The Journal of biological chemistry*. United States, 275(16), pp. 11678–11685. doi: 10.1074/jbc.275.16.11678.

Douki, T. and Sage, E. (2016) 'Dewar valence isomers, the third type of environmentally relevant DNA photoproducts induced by solar radiation.', *Photochemical & photobiological sciences : Official journal of the European Photochemistry Association and the European Society for Photobiology*. England, 15(1), pp. 24–30. doi: 10.1039/c5pp00382b.

Duggin, I. G. *et al.* (2008) 'The replication fork trap and termination of chromosome replication', *Molecular Microbiology*. John Wiley & Sons, Ltd, 70(6), pp. 1323–1333. doi: 10.1111/j.1365-2958.2008.06500.x.

Durbeej, B. and Eriksson, L. A. (2003) 'On the Formation of Cyclobutane Pyrimidine Dimers in UV-irradiated DNA: Why are Thymines More Reactive?¶', *Photochemistry and Photobiology*. John Wiley & Sons, Ltd, 78(2), pp. 159–167. doi: 10.1562/0031-8655(2003)0780159OTFOCP2.0.CO2.

Dusseau, C. *et al.* (2001) 'Analysis of uracil DNA glycosylase in human colorectal cancer.', *International journal of oncology*. Greece, 18(2), pp. 393–399. doi: 10.3892/ijo.18.2.393.

Egelman, E. H. and Stasiak, A. (1986) 'Structure of helical RecA-DNA complexes. Complexes formed in the presence of ATP-gamma-S or ATP.', *Journal of molecular biology*. England, 191(4), pp. 677–697. doi: 10.1016/0022-2836(86)90453-5.

Elsila, J. E., Glavin, D. P. and Dworkin, J. P. (2009) 'Cometary glycine detected in samples returned by Stardust', *Meteoritics & Planetary Science*. John Wiley & Sons, Ltd, 44(9), pp. 1323–1330. doi: 10.1111/j.1945-5100.2009.tb01224.x.

Eme, L. *et al.* (2023) 'Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes', *Nature*, 618(7967), pp. 992–999. doi: 10.1038/s41586-023-06186-2.

Engelward, B. P. *et al.* (1998) 'A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome.', *The Journal of biological chemistry*. United States, 273(9), pp. 5412–5418. doi: 10.1074/jbc.273.9.5412.

Errington, J. (2003) 'Regulation of endospore formation in Bacillus subtilis.', *Nature reviews. Microbiology*. England, 1(2), pp. 117–126. doi: 10.1038/nrmicro750.

Essen, L. O. and Klar, T. (2006) 'Light-driven DNA repair by photolyases.', *Cellular and molecular life sciences : CMLS*. Switzerland, 63(11), pp. 1266–1277. doi: 10.1007/s00018-005-5447-y.

Evans, E. *et al.* (1997) 'Mechanism of open complex and dual incision formation by human nucleotide excision repair factors.', *The EMBO journal*, 16(21), pp. 6559–6573. doi: 10.1093/emboj/16.21.6559.

Fang, W. H. and Modrich, P. (1993) 'Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction.', *The Journal of biological*

chemistry. United States, 268(16), pp. 11838–11844.

Farrokhi, A., Liu, H. and Szatmari, G. (2019) 'Characterization of the Chromosome Dimer Resolution Site in Caulobacter crescentus.', *Journal of bacteriology*. United States, 201(24). doi: 10.1128/JB.00391-19.

Fendrihan, S. *et al.* (2012) 'Spherical particles of halophilic archaea correlate with exposure to low water activity--implications for microbial survival in fluid inclusions of ancient halite', *Geobiology*. 2012/07/15. Blackwell Publishing Ltd, 10(5), pp. 424–433. doi: 10.1111/j.1472-4669.2012.00337.x.

Ferreira, A. C. *et al.* (1997) 'Deinococcus geothermalis sp. nov. and Deinococcus murrayi sp. nov., two extremely radiation-resistant and slightly thermophilic species from hot springs.', *International journal of systematic bacteriology*. England, 47(4), pp. 939–947. doi: 10.1099/00207713-47-4-939.

Fiala, G. and Stetter, K. O. (1986) 'Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C', *Archives of Microbiology*, 145(1), pp. 56–61. doi: 10.1007/BF00413027.

Firth, A. E. and Patrick, W. M. (2005) 'Statistics of protein library construction.', *Bioinformatics (Oxford, England)*. England, 21(15), pp. 3314–3315. doi: 10.1093/bioinformatics/bti516.

Fischer, C. J., Maluf, N. K. and Lohman, T. M. (2004) 'Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA.', *Journal of molecular biology*. England, 344(5), pp. 1287–1309. doi: 10.1016/j.jmb.2004.10.005.

Fitch, M. E. *et al.* (2003) 'In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product.', *The Journal of biological chemistry*. United States, 278(47), pp. 46906–46910. doi: 10.1074/jbc.M307254200.

Fogg, J. M. and Lilley, D. M. (2000) 'Ensuring productive resolution by the junction-resolving enzyme RuvC: large enhancement of the second-strand cleavage rate.', *Biochemistry*. United States, 39(51), pp. 16125–16134. doi: 10.1021/bi001886m.

Foley, C. *et al.* (2008) 'Martian surface chemistry: APXS results from the Pathfinder landing site', *The Martian Surface - Composition, Mineralogy, and Physical Properties*.

Forget, A. L. and Kowalczykowski, S. C. (2012) 'Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search', *Nature*, 482(7385), pp. 423–427. doi: 10.1038/nature10782.

Frick, D. N. and Richardson, C. C. (2001) 'DNA primases.', *Annual review of biochemistry*. United States, 70, pp. 39–80. doi: 10.1146/annurev.biochem.70.1.39.

Friedmann, E. I. (1982) 'Endolithic Microorganisms in the Antarctic Cold Desert', *Science*, 215(4536), p. 1045 LP-1053. doi: 10.1126/science.215.4536.1045.

Frigola, J. *et al.* (2017) 'Cdt1 stabilizes an open MCM ring for helicase loading', *Nature Communications*, 8(1), p. 15720. doi: 10.1038/ncomms15720.

Fu, D., Calvo, J. A. and Samson, L. D. (2012) 'Balancing repair and tolerance of DNA damage caused by alkylating agents', *Nature Reviews Cancer*, 12(2), pp. 104–120. doi: 10.1038/nrc3185.

Fujikane, R. *et al.* (2010) 'Genetic analysis of DNA repair in the hyperthermophilic archaeon, Thermococcus kodakaraensis', *Genes & genetic systems*, 85, pp. 243–257. doi: 10.1266/ggs.85.243.

Fujikane, R., Shinagawa, H. and Ishino, Y. (2006) 'The archaeal Hjm helicase has recQ-like functions, and may be involved in repair of stalled replication fork', *Genes to Cells*. John Wiley & Sons, Ltd, 11(2), pp. 99–110. doi: 10.1111/j.1365-2443.2006.00925.x.

Ganesan, S. and Smith, G. R. (1993) 'Strand-specific binding to duplex DNA ends by the subunits of the Escherichia coli RecBCD enzyme.', *Journal of molecular biology*. England, 229(1), pp. 67–78. doi: 10.1006/jmbi.1993.1008.

Gari, K. *et al.* (2008) 'The Fanconi anemia protein FANCM can promote branch migration of Holliday junctions and replication forks.', *Molecular cell*. United States, 29(1), pp. 141–148. doi: 10.1016/j.molcel.2007.11.032.

Ghodgaonkar, M. M. *et al.* (2013) 'Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair.', *Molecular cell*, 50(3), pp. 323–332. doi: 10.1016/j.molcel.2013.03.019.

Gilboa-Garber, N., Mymon, H. and Oren, A. (1998) 'Typing of halophilic Archaea and characterization of their cell surface carbohydrates by use of lectins', *FEMS Microbiology Letters*. John Wiley & Sons, Ltd, 163(1), pp. 91–97. doi: https://doi.org/10.1111/j.1574-6968.1998.tb13031.x.

Gillet, L. C. J. and Schärer, O. D. (2006) 'Molecular mechanisms of mammalian global genome nucleotide excision repair.', *Chemical reviews*. United States, 106(2), pp. 253–276. doi: 10.1021/cr040483f.

Gobbini, E. *et al.* (2016) 'Functions and regulation of the MRX complex at DNA double-strand breaks', *Microbial cell (Graz, Austria)*. Shared Science Publishers OG, 3(8), pp. 329–337. doi: 10.15698/mic2016.08.517.

Goldmark, P. J. and Linn, S. (1972) 'Purification and properties of the recBC DNase of Escherichia coli K-12.', *The Journal of biological chemistry*. United States, 247(6), pp. 1849–1860.

Goldstein, R. A. (2007) 'Amino-acid interactions in psychrophiles, mesophiles, thermophiles, and hyperthermophiles: insights from the quasi-chemical approximation', *Protein science : a publication of the Protein Society*. Wiley-Blackwell, 16(9), pp. 1887–1895. doi: 10.1110/ps.072947007.

Gómez, F. (2011) 'Acidophile BT - Encyclopedia of Astrobiology', in Gargaud, M. et al. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 10–12. doi: 10.1007/978-3-642-11274-4_22.

Goodman, M. F. and Woodgate, R. (2013) 'Translesion DNA polymerases', *Cold Spring*

Harbor perspectives in biology. Cold Spring Harbor Laboratory Press, 5(10), pp. a010363–a010363. doi: 10.1101/cshperspect.a010363.

Górecka, K. M. *et al.* (2019) 'RuvC uses dynamic probing of the Holliday junction to achieve sequence specificity and efficient resolution', *Nature Communications*, 10(1), p. 4102. doi: 10.1038/s41467-019-11900-8.

Gradia, S. *et al.* (1999) 'hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA.', *Molecular cell*. United States, 3(2), pp. 255–261. doi: 10.1016/s1097-2765(00)80316-0.

Grafstrom, R. H. (1986) 'The repair of pyrimidine dimers via a DNA-glycosylase mechanism.', *Basic life sciences*. United States, 38, pp. 281–286. doi: 10.1007/978-1-4615-9462-8_29.

Grant, W. D. (2004) 'Life at low water activity.', *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 359(1448), pp. 1247–1249. doi: 10.1098/rstb.2004.1502.

Gray, S. and Cohen, P. E. (2016) 'Control of Meiotic Crossovers: From Double-Strand Break Formation to Designation', *Annual review of genetics*. 2016/09/14, 50, pp. 175–210. doi: 10.1146/annurev-genet-120215-035111.

Greenberg, M. M. (2012) 'The formamidopyrimidines: purine lesions formed in competition with 8-oxopurines from oxidative stress', *Accounts of chemical research*. 2011/11/11, 45(4), pp. 588–597. doi: 10.1021/ar2002182.

Greenough, L., Kelman, Z. and Gardner, A. F. (2015) 'The roles of family B and D DNA polymerases in Thermococcus species 9°N Okazaki fragment maturation.', *The Journal of biological chemistry*. United States, 290(20), pp. 12514–12522. doi: 10.1074/jbc.M115.638130.

Grogan, D. W. (2004) 'Stability and repair of DNA in hyperthermophilic Archaea.', *Current issues in molecular biology*. England, 6(2), pp. 137–144.

Grogan, D. W. (2015) 'Understanding DNA Repair in Hyperthermophilic Archaea: Persistent Gaps and Other Reasons to Focus on the Fork', *Archaea (Vancouver, B.C.)*. Hindawi Publishing Corporation, 2015, p. 942605. doi: 10.1155/2015/942605.

Guldan, H., Sterner, R. and Babinger, P. (2008) 'Identification and Characterization of a Bacterial Glycerol-1-phosphate Dehydrogenase: Ni2+-Dependent AraM from Bacillus subtilis', *Biochemistry*. American Chemical Society, 47(28), pp. 7376–7384. doi: 10.1021/bi8005779.

Gunz, D., Hess, M. T. and Naegeli, H. (1996) 'Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism.', *The Journal of biological chemistry*. United States, 271(41), pp. 25089–25098. doi: 10.1074/jbc.271.41.25089.

Gupta, S., Yeeles, J. T. P. and Marians, K. J. (2014) 'Regression of replication forks stalled by leading-strand template damage: I. Both RecG and RuvAB catalyze regression, but RuvC cleaves the holliday junctions formed by RecG preferentially', *The Journal of biological*

chemistry. 2014/08/19. American Society for Biochemistry and Molecular Biology, 289(41), pp. 28376–28387. doi: 10.1074/jbc.M114.587881.

Gutman, P. D. *et al.* (1993) 'Identification, sequencing, and targeted mutagenesis of a DNA polymerase gene required for the extreme radioresistance of Deinococcus radiodurans.', *Journal of bacteriology*. United States, 175(11), pp. 3581–3590. doi: 10.1128/jb.175.11.3581-3590.1993.

Gutman, P. D. *et al.* (1994) 'Sequencing, targeted mutagenesis and expression of a recA gene required for the extreme radioresistance of Deinococcus radiodurans.', *Gene*. Netherlands, 141(1), pp. 31–37. doi: 10.1016/0378-1119(94)90124-4.

Gutman, P. D., Fuchs, P. and Minton, K. W. (1994) 'Restoration of the DNA damage resistance of Deinococcus radiodurans DNA polymerase mutants by Escherichia coli DNA polymerase I and Klenow fragment.', *Mutation research*. Netherlands, 314(1), pp. 87–97. doi: 10.1016/0921-8777(94)90064-7.

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

Hall, M. C., Jordan, J. R. and Matson, S. W. (1998) 'Evidence for a physical interaction between the Escherichia coli methyl-directed mismatch repair proteins MutL and UvrD.', *The EMBO journal*, 17(5), pp. 1535–1541. doi: 10.1093/emboj/17.5.1535.

Hall, M. C. and Matson, S. W. (1999) 'The Escherichia coli MutL protein physically interacts with MutH and stimulates the MutH-associated endonuclease activity.', *The Journal of biological chemistry*. United States, 274(3), pp. 1306–1312. doi: 10.1074/jbc.274.3.1306.

Hallsworth, J. E. (2019) 'Wooden owl that redefines Earth's biosphere may yet catapult a fungus into space', *Environmental Microbiology*. John Wiley & Sons, Ltd, 21(7), pp. 2202–2211. doi: https://doi.org/10.1111/1462-2920.14510.

Hamilton, C. M. *et al.* (2002) 'Multichannel plating unit for high-throughput plating of cell cultures', *BioTechniques*, 33(2), pp. 420–423. doi: 10.2144/02332ht02.

Hanawalt, P. C. and Spivak, G. (2008) 'Transcription-coupled DNA repair: two decades of progress and surprises.', *Nature reviews. Molecular cell biology*. England, 9(12), pp. 958–970. doi: 10.1038/nrm2549.

Hanford, M. J. and Peeples, T. L. (2002) 'Archaeal tetraether lipids', *Applied Biochemistry and Biotechnology*, 97(1), pp. 45–62. doi: 10.1385/ABAB:97:1:45.

Hansen, M. T. (1978) 'Multiplicity of genome equivalents in the radiation-resistant bacterium Micrococcus radiodurans', *Journal of bacteriology*, 134(1), pp. 71–75. Available at: https://pubmed.ncbi.nlm.nih.gov/649572.

Haque, R. U., Paradisi, F. and Allers, T. (2019) 'Haloferax volcanii as immobilised whole cell biocatalyst: new applications for halophilic systems.', *Applied microbiology and biotechnology*. Germany, 103(9), pp. 3807–3817. doi: 10.1007/s00253-019-09725-y.

Hartman, A. L. *et al.* (2010) 'The complete genome sequence of Haloferax volcanii DS2, a model archaeon.', *PloS one*, 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*. 2013/11/03, 503(7477), pp. 544–547. doi: 10.1038/nature12650.

Heimroth, R. D., Casadei, E. and Salinas, I. (2018) 'Effects of Experimental Terrestrialization on the Skin Mucus Proteome of African Lungfish (Protopterus dolloi)', *Frontiers in immunology*. Frontiers Media S.A., 9, p. 1259. doi: 10.3389/fimmu.2018.01259.

Heller, R. C. and Marians, K. J. (2006) 'Replication fork reactivation downstream of a blocked nascent leading strand.', *Nature*. England, 439(7076), pp. 557–562. doi: 10.1038/nature04329.

Henneke, G. (2012) 'In vitro reconstitution of RNA primer removal in Archaea reveals the existence of two pathways', *Biochemical Journal*, 447(2), pp. 271–280. doi: 10.1042/BJ20120959.

Henner, W. D. *et al.* (1983) 'gamma Ray induced deoxyribonucleic acid strand breaks. 3' Glycolate termini.', *The Journal of biological chemistry*. United States, 258(2), pp. 711–713.

Hilario, J. *et al.* (2009) 'Direct imaging of human Rad51 nucleoprotein dynamics on individual DNA molecules', *Proceedings of the National Academy of Sciences*, 106(2), p. 361 LP-368. doi: 10.1073/pnas.0811965106.

Hoang, M. L. *et al.* (2010) 'Competitive Repair by Naturally Dispersed Repetitive DNA during Non-Allelic Homologous Recombination', *PLOS Genetics*. Public Library of Science, 6(12), p. e1001228. Available at: https://doi.org/10.1371/journal.pgen.1001228.

Hollingsworth, M. A. and Swanson, B. J. (2004) 'Mucins in cancer: protection and control of the cell surface.', *Nature reviews. Cancer*. England, 4(1), pp. 45–60. doi: 10.1038/nrc1251.

Holloman, W. K. (2011) 'Unraveling the mechanism of BRCA2 in homologous recombination', *Nature structural & molecular biology*, 18(7), pp. 748–754. doi: 10.1038/nsmb.2096.

Holmes, M. L. and Dyall-Smith, M. L. (1990) 'A plasmid vector with a selectable marker for halophilic archaebacteria', *Journal of bacteriology*, 172(2), pp. 756–761. doi: 10.1128/jb.172.2.756-761.1990.

Holmes, M. L. and Dyall-Smith, M. L. (2000) 'Sequence and expression of a halobacterial beta-galactosidase gene.', *Molecular microbiology*. England, 36(1), pp. 114–122. doi: 10.1046/j.1365-2958.2000.01832.x.

Holmes, M. L., Nuttall, S. D. and Dyall-Smith, M. L. (1991) 'Construction and use of halobacterial shuttle vectors and further studies on Haloferax DNA gyrase', *Journal of bacteriology*, 173(12), pp. 3807–3813. doi: 10.1128/jb.173.12.3807-3813.1991.

Hopkins, B. B. and Paull, T. T. (2008) 'The P. furiosus mre11/rad50 complex promotes 5' strand resection at a DNA double-strand break.', *Cell*, 135(2), pp. 250–260. doi: 10.1016/j.cell.2008.09.054.

Horikoshi, K. (1999) ‘Alkaliphiles: some applications of their products for biotechnology.’, *Microbiology and molecular biology reviews : MMBR*. United States, 63(4), p. 735–50, table of contents.

Horneck, G., Klaus, D. M. and Mancinelli, R. L. (2010) ‘Space microbiology’, *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology (ASM), 74(1), pp. 121–156. doi: 10.1128/MMBR.00016-09.

Hurst, L. D. and Merchant, A. R. (2001) ‘High guanine-cytosine content is not an adaptation to high temperature: a comparative analysis amongst prokaryotes’, *Proceedings. Biological sciences*, 268(1466), pp. 493–497. doi: 10.1098/rspb.2000.1397.

Ibarra, A., Schwob, E. and Méndez, J. (2008) ‘Excess MCM proteins protect human cells from replicative stress by licensing backup origins of replication.’, *Proceedings of the National Academy of Sciences of the United States of America*, 105(26), pp. 8956–8961. doi: 10.1073/pnas.0803978105.

Ilves, I. *et al.* (2010) ‘Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins.’, *Molecular cell*, 37(2), pp. 247–258. doi: 10.1016/j.molcel.2009.12.030.

Imshenetsky, A. A., Kouzyurina, L. A. and Jakshina, V. M. (1973) ‘On the multiplication of xerophilic micro-organisms under simulated Martian conditions.’, *Life sciences and space research*. Netherlands, 11, pp. 63–66.

Ingleston, S. M., Sharples, G. J. and Lloyd, R. G. (2000) ‘The acidic pin of RuvA modulates Holliday junction binding and processing by the RuvABC resolvasome’, *The EMBO journal*. Oxford University Press, 19(22), pp. 6266–6274. doi: 10.1093/emboj/19.22.6266.

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.

Ivančić-Baće, I. *et al.* (2003) ‘RecFOR Function Is Required for DNA Repair and Recombination in a RecA Loading-Deficient recB Mutant of Escherichia coli’;, *Genetics*, 163(2), p. 485 LP-494. Available at: http://www.genetics.org/content/163/2/485.abstract.

Ivancic-Bace, I., Salaj-Smic, E. and Brcic-Kostic, K. (2005) ‘Effects of recJ, recQ, and recFOR mutations on recombination in nuclease-deficient recB recD double mutants of Escherichia coli.’, *Journal of bacteriology*, 187(4), pp. 1350–1356. doi: 10.1128/JB.187.4.1350-1356.2005.

Iwabe, N. *et al.* (1989) ‘Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes.’, *Proceedings of the National Academy of Sciences of the United States of America*, 86(23), pp. 9355–9359. doi: 10.1073/pnas.86.23.9355.

Iwasaki, H. *et al.* (1992) ‘Escherichia coli RuvA and RuvB proteins specifically interact with Holliday junctions and promote branch migration.’, *Genes & development*. United States, 6(11), pp. 2214–2220. doi: 10.1101/gad.6.11.2214.

Iyer, R. R. *et al.* (2008) 'The MutSalpha-proliferating cell nuclear antigen interaction in human DNA mismatch repair.', *The Journal of biological chemistry*, 283(19), pp. 13310–13319. doi: 10.1074/jbc.M800606200.

Jaakkola, S. T. *et al.* (2016) 'Buried Alive: Microbes from Ancient Halite', *Trends in Microbiology*. Elsevier, 24(2), pp. 148–160. doi: 10.1016/j.tim.2015.12.002.

Jaciuk, M. *et al.* (2011) 'Structure of UvrA nucleotide excision repair protein in complex with modified DNA', *Nature structural & molecular biology*. 2011/01/16, 18(2), pp. 191–197. doi: 10.1038/nsmb.1973.

Jeong, C. *et al.* (2011) 'MutS switches between two fundamentally distinct clamps during mismatch repair', *Nature Structural & Molecular Biology*, 18(3), pp. 379–385. doi: 10.1038/nsmb.2009.

Jiricny, J. (2006) 'The multifaceted mismatch-repair system.', *Nature reviews. Molecular cell biology*. England, 7(5), pp. 335–346. doi: 10.1038/nrm1907.

John, S. (1998) 'A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 95(4), pp. 1460–1465. doi: 10.1073/pnas.95.4.1460.

Jolivet, E. *et al.* (2003) 'Thermococcus gammatolerans sp. nov., a hyperthermophilic archaeon from a deep-sea hydrothermal vent that resists ionizing radiation.', *International journal of systematic and evolutionary microbiology*. England, 53(Pt 3), pp. 847–851. doi: 10.1099/ijs.0.02503-0.

Jolivet, E. *et al.* (2004) 'Thermococcus marinus sp. nov. and Thermococcus radiotolerans sp. nov., two hyperthermophilic archaea from deep-sea hydrothermal vents that resist ionizing radiation.', *Extremophiles : life under extreme conditions*. Germany, 8(3), pp. 219–227. doi: 10.1007/s00792-004-0380-9.

Jones, J. A. *et al.* (2007) 'Ionizing radiation-induced bioeffects in space and strategies to reduce cellular injury and carcinogenesis.', *Aviation, space, and environmental medicine*. United States, 78(4 Suppl), pp. A67-78.

Jones, P. G. and Inouye, M. (1994) 'The cold-shock response--a hot topic.', *Molecular microbiology*. England, 11(5), pp. 811–818. doi: 10.1111/j.1365-2958.1994.tb00359.x.

Jozwiakowski, S. K., Borazjani Gholami, F. and Doherty, A. J. (2015) 'Archaeal replicative primases can perform translesion DNA synthesis', *Proceedings of the National Academy of Sciences*, 112(7), p. E633 LP-E638. doi: 10.1073/pnas.1412982112.

Kadyrov, F. A. *et al.* (2006) 'Endonucleolytic function of MutLalpha in human mismatch repair.', *Cell*. United States, 126(2), pp. 297–308. doi: 10.1016/j.cell.2006.05.039.

Kaguni, J. M. (2011) 'Replication initiation at the Escherichia coli chromosomal origin', *Current opinion in chemical biology*. 2011/08/18, 15(5), pp. 606–613. doi: 10.1016/j.cbpa.2011.07.016.

Kasai, H. and Nishimura, S. (1984) 'Hydroxylation of deoxyguanosine at the C-8 position by

ascorbic acid and other reducing agents', *Nucleic acids research*, 12(4), pp. 2137–2145. doi: 10.1093/nar/12.4.2137.

Kawakami, H. *et al.* (2015) 'Specific binding of eukaryotic ORC to DNA replication origins depends on highly conserved basic residues', *Scientific Reports*, 5(1), p. 14929. doi: 10.1038/srep14929.

Keeney, S., Giroux, C. N. and Kleckner, N. (1997) 'Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family.', *Cell*. United States, 88(3), pp. 375–384. doi: 10.1016/s0092-8674(00)81876-0.

Kellermann, M. Y. *et al.* (2016) 'Important roles for membrane lipids in haloarchaeal bioenergetics', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1858(11), pp. 2940–2956. doi: https://doi.org/10.1016/j.bbamem.2016.08.010.

Kelman, Z. and O'Donnell, M. (1995) 'DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine.', *Annual review of biochemistry*. United States, 64, pp. 171–200. doi: 10.1146/annurev.bi.64.070195.001131.

Kelner, A. (1949) 'PHOTOREACTIVATION OF ULTRAVIOLET-IRRADIATED ESCHERICHIA COLI, WITH SPECIAL REFERENCE TO THE DOSE-REDUCTION PRINCIPLE AND TO ULTRAVIOLET-INDUCED MUTATION', *Journal of bacteriology*, 58(4), pp. 511–522. Available at: https://pubmed.ncbi.nlm.nih.gov/16561813.

Kerr, I. D. *et al.* (2003) 'Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein.', *The EMBO journal*, 22(11), pp. 2561–2570. doi: 10.1093/emboj/cdg272.

Kikuchi, K. *et al.* (2013) 'Structure-specific endonucleases xpf and mus81 play overlapping but essential roles in DNA repair by homologous recombination.', *Cancer research*, 73(14), pp. 4362–4371. doi: 10.1158/0008-5472.CAN-12-3154.

Kil, Y. V *et al.* (2000) 'Efficient strand transfer by the RadA recombinase from the hyperthermophilic archaeon Desulfurococcus amylolyticus', *Journal of bacteriology*. American Society for Microbiology, 182(1), pp. 130–134. doi: 10.1128/jb.182.1.130-134.2000.

Kim, Y.-J. and Wilson 3rd, D. M. (2012) 'Overview of base excision repair biochemistry', *Current molecular pharmacology*, 5(1), pp. 3–13. doi: 10.2174/1874467211205010003.

Kimura, J. and Kitadai, N. (2015) 'Polymerization of Building Blocks of Life on Europa and Other Icy Moons', *Astrobiology*. Mary Ann Liebert, Inc., 15(6), pp. 430–441. doi: 10.1089/ast.2015.1306.

Kleczkowska, H. E. *et al.* (2001) 'hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci.', *Genes & development*, 15(6), pp. 724–736. doi: 10.1101/gad.191201.

Koga, Y. (2012) 'Thermal adaptation of the archaeal and bacterial lipid membranes', *Archaea (Vancouver, B.C.)*. 2012/08/15. Hindawi Publishing Corporation, 2012, p. 789652. doi: 10.1155/2012/789652.

Koonin, E. V and Wolf, Y. I. (2008) 'Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world', *Nucleic acids research*. 2008/10/23. Oxford University Press, 36(21), pp. 6688–6719. doi: 10.1093/nar/gkn668.

Koonin, E. V and Yutin, N. (2014) 'The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes.', *Cold Spring Harbor perspectives in biology*, 6(4), p. a016188. doi: 10.1101/cshperspect.a016188.

Kottemann, M. *et al.* (2005) 'Physiological responses of the halophilic archaeon Halobacterium sp. strain NRC1 to desiccation and gamma irradiation', *Extremophiles*, 9(3), pp. 219–227. doi: 10.1007/s00792-005-0437-4.

Kow, Y. W. (2002) 'Repair of deaminated bases in DNA12 1Guest Editor: Miral Dizdaroglu 2This article is part of a series of reviews on "Oxidative DNA Damage and Repair." The full list of papers may be found on the homepage of the journal.', *Free Radical Biology and Medicine*, 33(7), pp. 886–893. doi: https://doi.org/10.1016/S0891-5849(02)00902-4.

Krupovic, M., Dolja, V. V and Koonin, E. V (2023) 'The virome of the last eukaryotic common ancestor and eukaryogenesis.', *Nature microbiology*. England, 8(6), pp. 1008–1017. doi: 10.1038/s41564-023-01378-y.

Kucukyildirim, S., Ozdemirel, H. O. and Lynch, M. (2023) 'Similar mutation rates but different mutation spectra in moderate and extremely halophilic archaea.', *G3 (Bethesda, Md.)*. England, 13(3). doi: 10.1093/g3journal/jkac303.

Kumar, R. and De Massy, B. (2010) 'Initiation of meiotic recombination in mammals', *Genes*. MDPI, 1(3), pp. 521–549. doi: 10.3390/genes1030521.

Kumari, R., Singh, K. P. and Dumond, J. W. J. (2009) 'Simulated microgravity decreases DNA repair capacity and induces DNA damage in human lymphocytes.', *Journal of cellular biochemistry*. United States, 107(4), pp. 723–731. doi: 10.1002/jcb.22171.

Kurosawa, N. *et al.* (2010) 'Archaeal and bacterial community structures in the anoxic sediment of Antarctic meromictic lake Nurume-Ike', *Polar Science*, 4(2), pp. 421–429. doi: https://doi.org/10.1016/j.polar.2010.04.002.

Kushner, S. R., Nagaishi, H. and Clark, A. J. (1974) 'Isolation of exonuclease VIII: the enzyme associated with sbcA indirect suppressor', *Proceedings of the National Academy of Sciences of the United States of America*, 71(9), pp. 3593–3597. doi: 10.1073/pnas.71.9.3593.

Kvaratskhelia, M., Wardleworth, B. N. and White, M. F. (2001) 'Multiple Holliday junction resolving enzyme activities in the Crenarchaeota and Euryarchaeota.', *FEBS letters*. England, 491(3), pp. 243–246. doi: 10.1016/s0014-5793(01)02200-1.

Lahue, R. S., Au, K. G. and Modrich, P. (1989) 'DNA mismatch correction in a defined system.', *Science (New York, N.Y.)*. United States, 245(4914), pp. 160–164. doi: 10.1126/science.2665076.

Lamprecht-Grandío, M. *et al.* (2020) 'Novel Genes Involved in Resistance to Both Ultraviolet Radiation and Perchlorate From the Metagenomes of Hypersaline Environments.', *Frontiers in microbiology*, 11, p. 453. doi: 10.3389/fmicb.2020.00453.

Lange, C. *et al.* (2011) 'Gene conversion results in the equalization of genome copies in the polyploid haloarchaeon Haloferax volcanii', *Molecular Microbiology*. John Wiley & Sons, Ltd, 80(3), pp. 666–677. doi: 10.1111/j.1365-2958.2011.07600.x.

Lee, S.-J. *et al.* (2007) 'The ABC of binding-protein-dependent transport in Archaea.', *Trends in microbiology*. England, 15(9), pp. 389–397. doi: 10.1016/j.tim.2007.08.002.

Lehman, I. R. (1974) 'DNA ligase: structure, mechanism, and function.', *Science (New York, N.Y.)*. United States, 186(4166), pp. 790–797. doi: 10.1126/science.186.4166.790.

Lestini, R., Duan, Z. and Allers, T. (2010) 'The archaeal Xpf/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in Haloferax volcanii.', *DNA repair*. Netherlands, 9(9), pp. 994–1002. doi: 10.1016/j.dnarep.2010.06.012.

Lewis Jr, C. A. *et al.* (2016) 'Cytosine deamination and the precipitous decline of spontaneous mutation during Earth's history', *Proceedings of the National Academy of Sciences of the United States of America*. 2016/07/05. National Academy of Sciences, 113(29), pp. 8194–8199. doi: 10.1073/pnas.1607580113.

Li, L. *et al.* (1994) 'Specific association between the human DNA repair proteins XPA and ERCC1.', *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), pp. 5012–5016. doi: 10.1073/pnas.91.11.5012.

Li, M. and Wilson 3rd, D. M. (2014) 'Human apurinic/apyrimidinic endonuclease 1', *Antioxidants & redox signaling*. 2013/08/20. Mary Ann Liebert, Inc., 20(4), pp. 678–707. doi: 10.1089/ars.2013.5492.

Lim, J., Thomas, T. and Cavicchioli, R. (2000) 'Low temperature regulated DEAD-box RNA helicase from the Antarctic archaeon, Methanococcoides burtonii.', *Journal of molecular biology*. England, 297(3), pp. 553–567. doi: 10.1006/jmbi.2000.3585.

Lindahl, T. (1980) 'Uracil-DNA glycosylase from Escherichia coli.', *Methods in enzymology*. United States, 65(1), pp. 284–290. doi: 10.1016/s0076-6879(80)65038-1.

Lindahl, T. *et al.* (1988) 'REGULATION AND EXPRESSION OF THE ADAPTIVE RESPONSE TO ALKYLATING AGENTS', *Annual Review of Biochemistry*. Annual Reviews, 57(1), pp. 133–157. doi: 10.1146/annurev.bi.57.070188.001025.

Lindahl, T. and Karlström, O. (1973) 'Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution.', *Biochemistry*. United States, 12(25), pp. 5151–5154. doi: 10.1021/bi00749a020.

Lindahl, T. and Nyberg, B. (1972) 'Rate of depurination of native deoxyribonucleic acid.', *Biochemistry*. United States, 11(19), pp. 3610–3618. doi: 10.1021/bi00769a018.

Lindås, A.-C. and Bernander, R. (2013) 'The cell cycle of archaea', *Nature Reviews Microbiology*, 11(9), pp. 627–638. doi: 10.1038/nrmicro3077.

Liu, B. *et al.* (2017) 'Direct Visualization of RNA-DNA Primer Removal from Okazaki Fragments Provides Support for Flap Cleavage and Exonucleolytic Pathways in Eukaryotic

Cells', *The Journal of biological chemistry*. 2017/02/03. American Society for Biochemistry and Molecular Biology, 292(12), pp. 4777–4788. doi: 10.1074/jbc.M116.758599.

Liu, J. *et al.* (2017) 'Srs2 promotes synthesis-dependent strand annealing by disrupting DNA polymerase δ-extending D-loops.', *eLife*, 6. doi: 10.7554/eLife.22195.

Liu, J., Wu, T. C. and Lichten, M. (1995) 'The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate.', *The EMBO journal*, 14(18), pp. 4599–4608.

Liu, Y. *et al.* (2021) 'Expanded diversity of Asgard archaea and their relationships with eukaryotes', *Nature*, 593(7860), pp. 553–557. doi: 10.1038/s41586-021-03494-3.

Liu, Z. *et al.* (2011) 'Dynamics and mechanism of cyclobutane pyrimidine dimer repair by DNA photolyase', *Proceedings of the National Academy of Sciences of the United States of America*. 2011/07/29. National Academy of Sciences, 108(36), pp. 14831–14836. doi: 10.1073/pnas.1110927108.

Lloyd, R. G. and Sharples, G. J. (1993) 'Dissociation of synthetic Holliday junctions by E. coli RecG protein.', *The EMBO journal*, 12(1), pp. 17–22.

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), p. 6271 LP-6275. doi: 10.1073/pnas.81.20.6271.

Lovett, S. T. (2012) 'A glimpse of molecular competition', *Nature*, 491(7423), pp. 198–200. doi: 10.1038/nature11639.

Lovett, S. T. (2017) 'Template-switching during replication fork repair in bacteria', *DNA repair*. 2017/06/13, 56, pp. 118–128. doi: 10.1016/j.dnarep.2017.06.014.

Lu, A. L. (1987) 'Influence of GATC sequences on Escherichia coli DNA mismatch repair in vitro.', *Journal of bacteriology*, 169(3), pp. 1254–1259. doi: 10.1128/jb.169.3.1254-1259.1987.

Ludt, K. and Soppa, J. (2018) 'Influence of Origin Recognition Complex Proteins on the Copy Numbers of Three Chromosomes in Haloferax volcanii. *J Bacteriol. 2018 Aug 10;200(17):e00161-18. doi: 10.1128/JB.00161-18. PMID: 29941422; PMCID: PMC6088171.*

Luncsford, P. J. *et al.* (2013) 'Coordination of MYH DNA glycosylase and APE1 endonuclease activities via physical interactions.', *DNA repair*, 12(12), pp. 1043–1052. doi: 10.1016/j.dnarep.2013.09.007.

Lutsenko, E. and Bhagwat, A. S. (1999) 'Principal causes of hot spots for cytosine to thymine mutations at sites of cytosine methylation in growing cells. A model, its experimental support and implications.', *Mutation research*. Netherlands, 437(1), pp. 11–20. doi: 10.1016/s1383-5742(99)00065-4.

Lyndaker, A. M. and Alani, E. (2009) 'A tale of tails: insights into the coordination of 3' end processing during homologous recombination', *BioEssays : news and reviews in molecular, cellular and developmental biology*, 31(3), pp. 315–321. doi: 10.1002/bies.200800195.

Lynn, A., Soucek, R. and Börner, G. V. (2007) 'ZMM proteins during meiosis: crossover artists at work.', *Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*. Netherlands, 15(5), pp. 591–605. doi: 10.1007/s10577-007-1150-1.

Mackwan, R. R. *et al.* (2007) 'An unusual pattern of spontaneous mutations recovered in the halophilic archaeon Haloferax volcanii.', *Genetics*, 176(1), pp. 697–702. doi: 10.1534/genetics.106.069666.

Madru, C. *et al.* (2023) 'DNA-binding mechanism and evolution of replication protein A', *Nature Communications*, 14(1), p. 2326. doi: 10.1038/s41467-023-38048-w.

Majka, J. *et al.* (2012) 'ATP hydrolysis by RAD50 protein switches MRE11 enzyme from endonuclease to exonuclease', *The Journal of biological chemistry*. 2011/11/18. American Society for Biochemistry and Molecular Biology, 287(4), pp. 2328–2341. doi: 10.1074/jbc.M111.307041.

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*. BioMed Central, 7, p. 7. doi: 10.1186/1745-6150-7-7.

Malkova, A. and Ira, G. (2013) 'Break-induced replication: functions and molecular mechanism', *Current opinion in genetics & development*. 2013/06/18, 23(3), pp. 271–279. doi: 10.1016/j.gde.2013.05.007.

Mancinelli, R. *et al.* (2007) 'Hydration of Sodium, Potassium, and Chloride Ions in Solution and the Concept of Structure Maker/Breaker', *The Journal of Physical Chemistry B*. American Chemical Society, 111(48), pp. 13570–13577. doi: 10.1021/jp075913v.

Manelyte, L. *et al.* (2009) 'The unstructured C-terminal extension of UvrD interacts with UvrB, but is dispensable for nucleotide excision repair', *DNA repair*. 2009/09/16. Elsevier, 8(11), pp. 1300–1310. doi: 10.1016/j.dnarep.2009.08.005.

Mansky, J. *et al.* (2023) 'Role of XerCD in release of over-replicated DNA through Outer Membrane Vesicles in <em>Escherichia coli</em&gt';, *bioRxiv*, p. 2023.05.30.542921. doi: 10.1101/2023.05.30.542921.

Margesin, R. and Miteva, V. (2011) 'Diversity and ecology of psychrophilic microorganisms.', *Research in microbiology*. France, 162(3), pp. 346–361. doi: 10.1016/j.resmic.2010.12.004.

Marion, G. M. *et al.* (2003) 'The Search for Life on Europa: Limiting Environmental Factors, Potential Habitats, and Earth Analogues', *Astrobiology*. Mary Ann Liebert, Inc., publishers, 3(4), pp. 785–811. doi: 10.1089/153110703322736105.

Mathieu, N. *et al.* (2013) 'DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH.', *Current biology : CB*. England, 23(3), pp. 204–212. doi: 10.1016/j.cub.2012.12.032.

Matos, J. *et al.* (2011) 'Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis.', *Cell*, 147(1), pp. 158–172. doi: 10.1016/j.cell.2011.08.032.

Matsumiya, S., Ishino, Y. and Morikawa, K. (2001) 'Crystal structure of an archaeal DNA sliding clamp: proliferating cell nuclear antigen from Pyrococcus furiosus', *Protein science : a publication of the Protein Society*. Cold Spring Harbor Laboratory Press, 10(1), pp. 17–23. doi: 10.1110/ps.36401.

Matsumoto, Y. and Kim, K. (1995) 'Matsumoto, Y. & Kim, K. Excision of deoxyribose phosphate residues by DNA polymerase betaduring DNA repair. Science 269, 699-702', *Science (New York, N.Y.)*, 269, pp. 699–702. doi: 10.1126/science.7624801.

Mattimore, V. and Battista, J. R. (1996) 'Radioresistance of Deinococcus radiodurans: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation', *Journal of bacteriology*, 178(3), pp. 633–637. doi: 10.1128/jb.178.3.633-637.1996.

Maul, R. W. and Gearhart, P. J. (2010) 'AID and somatic hypermutation', *Advances in immunology*, 105, pp. 159–191. doi: 10.1016/S0065-2776(10)05006-6.

Mazin, A. V *et al.* (2010) 'Rad54, the motor of homologous recombination', *DNA repair*. 2010/01/20, 9(3), pp. 286–302. doi: 10.1016/j.dnarep.2009.12.006.

McIlwraith, M. J. and West, S. C. (2008) 'DNA repair synthesis facilitates RAD52-mediated second-end capture during DSB repair.', *Molecular cell*. United States, 29(4), pp. 510–516. doi: 10.1016/j.molcel.2007.11.037.

McKay, C. P. (2010) 'An origin of life on Mars', *Cold Spring Harbor perspectives in biology*. 2010/03/03. Cold Spring Harbor Laboratory Press, 2(4), pp. a003509–a003509. doi: 10.1101/cshperspect.a003509.

McKay, C. P., Toon, O. B. and Kasting, J. F. (1991) 'Making Mars habitable', *Nature*, 352(6335), pp. 489–496. doi: 10.1038/352489a0.

Mees, A. *et al.* (2004) 'Crystal Structure of a Photolyase Bound to a CPD-Like DNA Lesion After in Situ Repair', *Science*, 306(5702), p. 1789 LP-1793. doi: 10.1126/science.1101598.

Meng, X. and Zhao, X. (2017) 'Replication fork regression and its regulation', *FEMS yeast research*. Oxford University Press, 17(1), p. fow110. doi: 10.1093/femsyr/fow110.

Messer, W. (2002) 'The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication', *FEMS Microbiology Reviews*, 26(4), pp. 355–374. doi: 10.1111/j.1574-6976.2002.tb00620.x.

Meyer, B. H. and Albers, S.-V. (2014) 'Archaeal Cell Walls', *eLS*. (Major Reference Works). doi: doi:10.1002/9780470015902.a0000384.pub2.

Midonet, C. *et al.* (2019) 'The TLCΦ satellite phage harbors a Xer recombination activation factor', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 116(37), pp. 18391–18396. doi: 10.1073/pnas.1902905116.

Miezner, G. *et al.* (2023) 'An archaeal Cas3 protein facilitates rapid recovery from DNA damage.', *microLife*. England, 4, p. uqad007. doi: 10.1093/femsml/uqad007.

Miné-Hattab, J. and Rothstein, R. (2012) 'Increased chromosome mobility facilitates

homology search during recombination', *Nature Cell Biology*, 14(5), pp. 510–517. doi: 10.1038/ncb2472.

Minton, K. W. (1994) 'DNA repair in the extremely radioresistant bacterium Deinococcus radiodurans', *Molecular Microbiology*. John Wiley & Sons, Ltd, 13(1), pp. 9–15. doi: 10.1111/j.1365-2958.1994.tb00397.x.

Miroshnichenko, M. L. *et al.* (2001) 'Isolation and characterization of Thermococcus sibiricus sp. nov. from a Western Siberia high-temperature oil reservoir.', *Extremophiles : life under extreme conditions*. Germany, 5(2), pp. 85–91. doi: 10.1007/s007920100175.

Miyabe, I., Kunkel, T. A. and Carr, A. M. (2011) 'The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved', *PLoS genetics*. 2011/12/01. Public Library of Science, 7(12), pp. e1002407–e1002407. doi: 10.1371/journal.pgen.1002407.

Moldovan, G.-L., Pfander, B. and Jentsch, S. (2007) 'PCNA, the maestro of the replication fork.', *Cell*. United States, 129(4), pp. 665–679. doi: 10.1016/j.cell.2007.05.003.

Moolenaar, G. F. *et al.* (2002) 'Cho, a second endonuclease involved in Escherichia coli nucleotide excision repair.', *Proceedings of the National Academy of Sciences of the United States of America*, 99(3), pp. 1467–1472. doi: 10.1073/pnas.032584099.

Morimatsu, K. and Kowalczykowski, S. C. (2014) 'RecQ helicase and RecJ nuclease provide complementary functions to resect DNA for homologous recombination', *Proceedings of the National Academy of Sciences*, 111(48), p. E5133 LP-E5142. doi: 10.1073/pnas.1420009111.

Mormile, M. R. *et al.* (2003) 'Isolation of Halobacterium salinarum retrieved directly from halite brine inclusions.', *Environmental microbiology*. England, 5(11), pp. 1094–1102. doi: 10.1046/j.1462-2920.2003.00509.x.

Mosbaugh, D. W. and Linn, S. (1980) 'Further characterization of human fibroblast apurinic/apyrimidinic DNA endonucleases. The definition of two mechanistic classes of enzyme.', *Journal of Biological Chemistry* , 255(24), pp. 11743–11752. Available at: http://www.jbc.org/content/255/24/11743.short.

Moyer, S. E., Lewis, P. W. and Botchan, M. R. (2006) 'Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase.', *Proceedings of the National Academy of Sciences of the United States of America*, 103(27), pp. 10236–10241. doi: 10.1073/pnas.0602400103.

Mullakhanbhai, M. F. and Larsen, H. (1975) 'Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement.', *Archives of microbiology*. Germany, 104(3), pp. 207–214. doi: 10.1007/bf00447326.

Mykytczuk, N. C. S. *et al.* (2013) 'Bacterial growth at -15 degrees C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1.', *The ISME journal*. England, 7(6), pp. 1211–1226. doi: 10.1038/ismej.2013.8.

Nagata, M. *et al.* (2017) 'The Cdc45/RecJ-like protein forms a complex with GINS and MCM, and is important for DNA replication in Thermococcus kodakarensis.', *Nucleic acids*

research, 45(18), pp. 10693–10705. doi: 10.1093/nar/gkx740.

Nakabeppu, Y. (2014) 'Cellular levels of 8-oxoguanine in either DNA or the nucleotide pool play pivotal roles in carcinogenesis and survival of cancer cells', *International journal of molecular sciences*. MDPI, 15(7), pp. 12543–12557. doi: 10.3390/ijms150712543.

Nakabeppu, Y., Kondo, H. and Sekiguchi, M. (1984) 'Cloning and characterization of the alkA gene of Escherichia coli that encodes 3-methyladenine DNA glycosylase II.', *The Journal of biological chemistry*. United States, 259(22), pp. 13723–13729.

Nakae, S. *et al.* (2016) 'Structure of the EndoMS-DNA Complex as Mismatch Restriction Endonuclease.', *Structure (London, England : 1993)*. United States, 24(11), pp. 1960–1971. doi: 10.1016/j.str.2016.09.005.

Nasir, A., Mughal, F. and Caetano-Anollés, G. (2021) 'The tree of life describes a tripartite cellular world', *BioEssays*. John Wiley & Sons, Ltd, 43(6), p. 2000343. doi: https://doi.org/10.1002/bies.202000343.

Natrajan, G. *et al.* (2003) 'Structures of Escherichia coli DNA mismatch repair enzyme MutS in complex with different mismatches: a common recognition mode for diverse substrates', *Nucleic acids research*. Oxford University Press, 31(16), pp. 4814–4821. doi: 10.1093/nar/gkg677.

Neale, M. J. and Keeney, S. (2006) 'Clarifying the mechanics of DNA strand exchange in meiotic recombination', *Nature*, 442(7099), pp. 153–158. doi: 10.1038/nature04885.

Neddermann, P. and Jiricny, J. (1993) 'The purification of a mismatch-specific thymine-DNA glycosylase from HeLa cells.', *The Journal of biological chemistry*. United States, 268(28), pp. 21218–21224.

Newman, T. J. *et al.* (2013) 'Replisome stall events have shaped the distribution of replication origins in the genomes of yeasts', *Nucleic acids research*. 2013/08/19. Oxford University Press, 41(21), pp. 9705–9718. doi: 10.1093/nar/gkt728.

Neylon, C. *et al.* (2005) 'Replication termination in Escherichia coli: structure and antihelicase activity of the Tus-Ter complex.', *Microbiology and molecular biology reviews : MMBR*, 69(3), pp. 501–526. doi: 10.1128/MMBR.69.3.501-526.2005.

Ng, J. M. Y. *et al.* (2003) 'A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein', *Genes & development*. 2003/06/18. Cold Spring Harbor Laboratory Press, 17(13), pp. 1630–1645. doi: 10.1101/gad.260003.

Nicholson, W. L. *et al.* (2000) 'Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments', *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology, 64(3), pp. 548–572. doi: 10.1128/MMBR.64.3.548-572.2000.

Nilsen, H., Lindahl, T. and Verreault, A. (2002) 'DNA base excision repair of uracil residues in reconstituted nucleosome core particles.', *The EMBO journal*, 21(21), pp. 5943–5952. doi: 10.1093/emboj/cdf581.

Nishi, R. *et al.* (2005) 'Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein.', *Molecular and cellular biology*, 25(13), pp. 5664–5674. doi: 10.1128/MCB.25.13.5664-5674.2005.

Nissenbaum, A. (1975) 'The microbiology and biogeochemistry of the Dead Sea.', *Microbial ecology*. United States, 2(2), pp. 139–161. doi: 10.1007/BF02010435.

Norais, C. *et al.* (2007) 'Genetic and physical mapping of DNA replication origins in Haloferax volcanii.', *PLoS genetics*, 3(5), p. e77. doi: 10.1371/journal.pgen.0030077.

NORTON, C. F. and GRANT, W. D. (1988) 'Survival of Halobacteria Within Fluid Inclusions in Salt Crystals', *Microbiology*, pp. 1365–1373. doi: 10.1099/00221287-134-5-1365.

Nußbaum, P. *et al.* (2021) 'The archaeal protein SepF is essential for cell division in Haloferax volcanii', *Nature Communications*, 12(1), p. 3469. doi: 10.1038/s41467-021-23686-9.

Nuttall, S. D. *et al.* (2000) 'The ShBle resistance determinant from Streptoalloteichus hindustanus is expressed in Haloferax volcanii and confers resistance to bleomycin.', *The Biochemical journal*, 346 Pt 2(Pt 2), pp. 251–254.

O'Brien, P. J. and Ellenberger, T. (2004) 'The Escherichia coli 3-Methyladenine DNA Glycosylase AlkA Has a Remarkably Versatile Active Site', *Journal of Biological Chemistry* , 279(26), pp. 26876–26884. doi: 10.1074/jbc.M403860200.

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*, 5(7). doi: 10.1101/cshperspect.a010108.

O'Donovan, A. *et al.* (1994) 'XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair.', *Nature*. England, 371(6496), pp. 432–435. doi: 10.1038/371432a0.

Oger, P. M. and Cario, A. (2013) 'Adaptation of the membrane in Archaea', *Biophysical Chemistry*, 183, pp. 42–56. doi: https://doi.org/10.1016/j.bpc.2013.06.020.

Olsson, M. and Lindahl, T. (1980) 'Repair of alkylated DNA in Escherichia coli. Methyl group transfer from O6-methylguanine to a protein cysteine residue.', *The Journal of biological chemistry*. United States, 255(22), pp. 10569–10571.

Onyenwoke, R. U. *et al.* (2004) 'Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes).', *Archives of microbiology*. Germany, 182(2–3), pp. 182–192. doi: 10.1007/s00203-004-0696-y.

Oren, A. (1999) 'Bioenergetic aspects of halophilism', *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology, 63(2), pp. 334–348. Available at: https://pubmed.ncbi.nlm.nih.gov/10357854.

Oren, A. (2002) 'Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications.', *Journal of industrial microbiology & biotechnology*. Germany, 28(1), pp. 56–63. doi: 10.1038/sj/jim/7000176.

Oren, A. (2008) 'Microbial life at high salt concentrations: phylogenetic and metabolic

diversity', *Saline systems*. BioMed Central, 4, p. 2. doi: 10.1186/1746-1448-4-2.

Orosei, R. *et al.* (2018) 'Radar evidence of subglacial liquid water on Mars', *Science*, 361(6401), p. 490 LP-493. doi: 10.1126/science.aar7268.

Orren, D. K. *et al.* (1992) 'Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I.', *The Journal of biological chemistry*. United States, 267(2), pp. 780–788.

Ortenberg, R., Rozenblatt-Rosen, O. and Mevarech, M. (2000) 'The extremely halophilic archaeon Haloferax volcanii has two very different dihydrofolate reductases.', *Molecular microbiology*. England, 35(6), pp. 1493–1505. doi: 10.1046/j.1365-2958.2000.01815.x.

Owczarzy, R. *et al.* (2004) 'Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures.', *Biochemistry*. United States, 43(12), pp. 3537–3554. doi: 10.1021/bi034621r.

Ozaki, S. and Katayama, T. (2009) 'DnaA structure, function, and dynamics in the initiation at the chromosomal origin.', *Plasmid*. United States, 62(2), pp. 71–82. doi: 10.1016/j.plasmid.2009.06.003.

Pappalardo, R. T. *et al.* (1999) 'Does Europa have a subsurface ocean? Evaluation of the geological evidence', *Journal of Geophysical Research: Planets*. John Wiley & Sons, Ltd, 104(E10), pp. 24015–24055. doi: 10.1029/1998JE000628.

Parsons, C. A. *et al.* (1995) 'Structure of a multisubunit complex that promotes DNA branch migration.', *Nature*. England, 374(6520), pp. 375–378. doi: 10.1038/374375a0.

Passot, F. M. *et al.* (2015) 'Nucleoid organization in the radioresistant bacterium Deinococcus radiodurans', *Molecular Microbiology*. John Wiley & Sons, Ltd, 97(4), pp. 759–774. doi: 10.1111/mmi.13064.

Patel, B. K. C., Skerratt, J. H. and Nichols, P. D. (1991) 'The Phospholipid Ester-linked Fatty Acid Composition of Thermophilic Bacteria', *Systematic and Applied Microbiology*, 14(4), pp. 311–316. doi: https://doi.org/10.1016/S0723-2020(11)80304-8.

Patrick, W. M., Firth, A. E. and Blackburn, J. M. (2003) 'User-friendly algorithms for estimating completeness and diversity in randomized protein-encoding libraries.', *Protein engineering*. England, 16(6), pp. 451–457. doi: 10.1093/protein/gzg057.

Paull, T. T. and Gellert, M. (1998) 'The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks.', *Molecular cell*. United States, 1(7), pp. 969–979. doi: 10.1016/s1097-2765(00)80097-0.

Pérez-Arnaiz, P. *et al.* (2020) 'Haloferax volcanii-a model archaeon for studying DNA replication and repair.', *Open biology*. England, 10(12), p. 200293. doi: 10.1098/rsob.200293.

Petrini, J. H. J. and Stracker, T. H. (2003) 'The cellular response to DNA double-strand breaks: defining the sensors and mediators.', *Trends in cell biology*. England, 13(9), pp. 458–462. doi: 10.1016/s0962-8924(03)00170-3.

Pierazzo, E. and Chyba, C. F. (2002) 'Cometary Delivery of Biogenic Elements to Europa', *Icarus*, 157(1), pp. 120–127. doi: https://doi.org/10.1006/icar.2001.6812.

Piggot, P. J. and Hilbert, D. W. (2004) 'Sporulation of Bacillus subtilis.', *Current opinion in microbiology*. England, 7(6), pp. 579–586. doi: 10.1016/j.mib.2004.10.001.

Polavarapu, A. *et al.* (2012) 'The Mechanism of Guanine Alkylation by Nitrogen Mustards: A Computational Study', *The Journal of Organic Chemistry*. American Chemical Society, 77(14), pp. 5914–5921. doi: 10.1021/jo300351g.

Prieur, D. (2014) 'Piezophile BT - Encyclopedia of Astrobiology', in Amils, R. et al. (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–2. doi: 10.1007/978-3-642-27833-4_151-3.

Qiu, R. *et al.* (2015) 'MutL traps MutS at a DNA mismatch', *Proceedings of the National Academy of Sciences of the United States of America*. 2015/08/17. National Academy of Sciences, 112(35), pp. 10914–10919. doi: 10.1073/pnas.1505655112.

Rafferty, J. B. *et al.* (1996) 'Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction.', *Science (New York, N.Y.)*. United States, 274(5286), pp. 415–421. doi: 10.1126/science.274.5286.415.

Rafferty, J. B. *et al.* (1998) 'Structural similarities between Escherichia coli RuvA protein and other DNA-binding proteins and a mutational analysis of its binding to the holliday junction.', *Journal of molecular biology*. England, 278(1), pp. 105–116. doi: 10.1006/jmbi.1998.1697.

Ragunathan, K., Liu, C. and Ha, T. (2012) 'RecA filament sliding on DNA facilitates homology search.', *eLife*, 1, p. e00067. doi: 10.7554/eLife.00067.

Rampelotto, P. H. (2013) 'Extremophiles and extreme environments', *Life (Basel, Switzerland)*. MDPI, 3(3), pp. 482–485. doi: 10.3390/life3030482.

Reginato, G. and Cejka, P. (2020) 'The MRE11 complex: A versatile toolkit for the repair of broken DNA', *DNA Repair*, 91–92, p. 102869. doi: https://doi.org/10.1016/j.dnarep.2020.102869.

Ren, B. *et al.* (2009) 'Structure and function of a novel endonuclease acting on branched DNA substrates.', *The EMBO journal*, 28(16), pp. 2479–2489. doi: 10.1038/emboj.2009.192.

Rescifina, A. *et al.* (2014) 'Recent advances in small organic molecules as DNA intercalating agents: Synthesis, activity, and modeling', *European Journal of Medicinal Chemistry*, 74, pp. 95–115. doi: https://doi.org/10.1016/j.ejmech.2013.11.029.

Reuter, C. J. and Maupin-Furlow, J. A. (2004) 'Analysis of proteasome-dependent proteolysis in Haloferax volcanii cells, using short-lived green fluorescent proteins', *Applied and environmental microbiology*. American Society for Microbiology, 70(12), pp. 7530–7538. doi: 10.1128/AEM.70.12.7530-7538.2004.

Rink, S. M. *et al.* (1993) 'Covalent structure of a nitrogen mustard-induced DNA interstrand cross-link: an N7-to-N7 linkage of deoxyguanosine residues at the duplex sequence 5'-

d(GNC)', *Journal of the American Chemical Society*. American Chemical Society, 115(7), pp. 2551–2557. doi: 10.1021/ja00060a001.

Robertson, A. B. *et al.* (2009) 'DNA repair in mammalian cells: Base excision repair: the long and short of it.', *Cellular and molecular life sciences : CMLS*. Switzerland, 66(6), pp. 981–993. doi: 10.1007/s00018-009-8736-z.

Rogers, C. M. *et al.* (2020) 'Fanconi anemia-independent DNA inter-strand crosslink repair in eukaryotes.', *Progress in biophysics and molecular biology*. England, 158, pp. 33–46. doi: 10.1016/j.pbiomolbio.2020.08.005.

Romanovskaia, V. A., Rokitko, P. V and Malashenko, I. R. (2000) '[Unique properties of highly radioresistant bacteria].', *Mikrobiolohichnyi zhurnal (Kiev, Ukraine : 1993)*. Ukraine, 62(1), pp. 40–63.

De Rosa, M. and Gambacorta, A. (1988) 'The lipids of archaebacteria.', *Progress in lipid research*. England, 27(3), pp. 153–175. doi: 10.1016/0163-7827(88)90011-2.

Rothschild, L. J. (1990) 'Earth analogs for Martian life. Microbes in evaporites, a new model system for life on Mars', *Icarus*, 88(1), pp. 246–260. doi: https://doi.org/10.1016/0019-1035(90)90188-F.

Rouillon, C. and White, M. F. (2010) 'The XBP-Bax1 helicase-nuclease complex unwinds and cleaves DNA: implications for eukaryal and archaeal nucleotide excision repair.', *The Journal of biological chemistry*, 285(14), pp. 11013–11022. doi: 10.1074/jbc.M109.094763.

Rowen, L. and Kornberg, A. (1978) 'Primase, the dnaG protein of Escherichia coli. An enzyme which starts DNA chains.', *The Journal of biological chemistry*. United States, 253(3), pp. 758–764.

Russell, N. J. (1989) 'Adaptive modifications in membranes of halotolerant and halophilic microorganisms', *Journal of Bioenergetics and Biomembranes*, 21(1), pp. 93–113. doi: 10.1007/BF00762214.

Sakofsky, C. J. and Malkova, A. (2017) 'Break induced replication in eukaryotes: mechanisms, functions, and consequences', *Critical reviews in biochemistry and molecular biology*. 2017/04/21, 52(4), pp. 395–413. doi: 10.1080/10409238.2017.1314444.

Sakumi, K. *et al.* (1986) 'Purification and structure of 3-methyladenine-DNA glycosylase I of Escherichia coli.', *The Journal of biological chemistry*. United States, 261(33), pp. 15761–15766.

Samson, L. *et al.* (1991) 'Cloning and characterization of a 3-methyladenine DNA glycosylase cDNA from human cells whose gene maps to chromosome 16', *Proceedings of the National Academy of Sciences of the United States of America*, 88(20), pp. 9127–9131. doi: 10.1073/pnas.88.20.9127.

Sancar, A. (1994) 'Structure and function of DNA photolyase.', *Biochemistry*. United States, 33(1), pp. 2–9. doi: 10.1021/bi00167a001.

Sancar, A. and Rupp, W. D. (1983) 'A novel repair enzyme: UVRABC excision nuclease of

Escherichia coli cuts a DNA strand on both sides of the damaged region', *Cell*. Elsevier, 33(1), pp. 249–260. doi: 10.1016/0092-8674(83)90354-9.

Sandler, S. J. and Marians, K. J. (2000) 'Role of PriA in replication fork reactivation in Escherichia coli', *Journal of bacteriology*. American Society for Microbiology, 182(1), pp. 9–13. doi: 10.1128/jb.182.1.9-13.2000.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors', *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), pp. 5463–5467. doi: 10.1073/pnas.74.12.5463.

Sartori, A. A. *et al.* (2007) 'Human CtIP promotes DNA end resection.', *Nature*, 450(7169), pp. 509–514. doi: 10.1038/nature06337.

Sawyer, S. *et al.* (2012) 'Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA', *PloS one*. 2012/03/30. Public Library of Science, 7(3), pp. e34131–e34131. doi: 10.1371/journal.pone.0034131.

Schorghofer, N. (2020) 'Mars: Quantitative Evaluation of Crocus Melting behind Boulders', *The Astrophysical Journal*, 890, p. 49. doi: 10.3847/1538-4357/ab612f.

Schormann, N., Ricciardi, R. and Chattopadhyay, D. (2014) 'Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes', *Protein science : a publication of the Protein Society*. 2014/10/25. BlackWell Publishing Ltd, 23(12), pp. 1667–1685. doi: 10.1002/pro.2554.

Schubert, B. A., Lowenstein, T. K. and Timofeeff, M. N. (2009) 'Microscopic identification of prokaryotes in modern and ancient halite, Saline Valley and Death Valley, California.', *Astrobiology*. United States, 9(5), pp. 467–482. doi: 10.1089/ast.2008.0282.

Schwartz, E. K. and Heyer, W.-D. (2011) 'Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes.', *Chromosoma*, 120(2), pp. 109–127. doi: 10.1007/s00412-010-0304-7.

Sehorn, M. G. *et al.* (2004) 'Human meiotic recombinase Dmc1 promotes ATP-dependent homologous DNA strand exchange.', *Nature*. England, 429(6990), pp. 433–437. doi: 10.1038/nature02563.

Sekimizu, K., Bramhill, D. and Kornberg, A. (1987) 'ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome.', *Cell*. United States, 50(2), pp. 259–265. doi: 10.1016/0092-8674(87)90221-2.

Setlow, P. and Kornberg, A. (1970) 'Biochemical studies of bacterial sporulation and germination. XXII. Energy metabolism in early stages of germination of Bacillus megaterium spores.', *The Journal of biological chemistry*. United States, 245(14), pp. 3637–3644.

Sharma, M. *et al.* (2014) 'Differential mismatch recognition specificities of eukaryotic MutS homologs, MutSα and MutSβ.', *Biophysical journal*, 106(11), pp. 2483–2492. doi: 10.1016/j.bpj.2014.04.026.

Sharples, G. J. *et al.* (1990) 'Molecular and functional analysis of the ruv region of

Escherichia coli K-12 reveals three genes involved in DNA repair and recombination.', *Molecular & general genetics : MGG*. Germany, 221(2), pp. 219–226. doi: 10.1007/BF00261724.

Sharples, G. J. *et al.* (1994) 'Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions.', *The EMBO journal*, 13(24), pp. 6133–6142.

Shereda, R. D. *et al.* (2008) 'SSB as an organizer/mobilizer of genome maintenance complexes.', *Critical reviews in biochemistry and molecular biology*, 43(5), pp. 289–318. doi: 10.1080/10409230802341296.

Shimokawa, K. *et al.* (2013) 'FtsK-dependent XerCD-dif recombination unlinks replication catenanes in a stepwise manner', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 110(52), pp. 20906–20911. doi: 10.1073/pnas.1308450110.

Shivji, M. K. *et al.* (1995) 'Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA.', *Biochemistry*. United States, 34(15), pp. 5011–5017. doi: 10.1021/bi00015a012.

Shrivastav, N., Li, D. and Essigmann, J. M. (2010) 'Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation', *Carcinogenesis*. 2009/10/29. Oxford University Press, 31(1), pp. 59–70. doi: 10.1093/carcin/bgp262.

Siglioccolo, A. *et al.* (2011) 'Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface.', *BMC structural biology*, 11, p. 50. doi: 10.1186/1472-6807-11-50.

Simic, M. G. (1994) 'DNA Markers of Oxidative Processes in Vivo: Relevance to Carcinogenesis and Anticarcinogenesis', *Cancer Research*, 54(7 Supplement), p. 1918s LP-1923s. Available at: http://cancerres.aacrjournals.org/content/54/7_Supplement/1918s.abstract.

Singh, V., Fedeles, B. I. and Essigmann, J. M. (2015) 'Role of tautomerism in RNA biochemistry', *RNA (New York, N.Y.)*. Cold Spring Harbor Laboratory Press, 21(1), pp. 1–13. doi: 10.1261/rna.048371.114.

Singleton, M. R. *et al.* (2004) 'Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks', *Nature*, 432(7014), pp. 187–193. doi: 10.1038/nature02988.

Slupphaug, G. *et al.* (1996) 'A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA.', *Nature*. England, 384(6604), pp. 87–92. doi: 10.1038/384087a0.

Smith, G. R. *et al.* (1981) 'Structure of chi hotspots of generalized recombination.', *Cell*. United States, 24(2), pp. 429–436. doi: 10.1016/0092-8674(81)90333-0.

Smith, G. R. (2012) 'How RecBCD enzyme and Chi promote DNA break repair and recombination: a molecular biologist's view', *Microbiology and molecular biology reviews : MMBR*. American Society for Microbiology, 76(2), pp. 217–228. doi: 10.1128/MMBR.05026-

11.

Snowden, T. *et al.* (2004) 'hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes.', *Molecular cell*. United States, 15(3), pp. 437–451. doi: 10.1016/j.molcel.2004.06.040.

Soppa, J. (2011) 'Ploidy and gene conversion in Archaea', *Biochemical Society Transactions*, 39(1), pp. 150–154. doi: 10.1042/BST0390150.

Soulas-Sprauel, P. *et al.* (2007) 'V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining', *Oncogene*, 26(56), pp. 7780–7791. doi: 10.1038/sj.onc.1210875.

Stasiak, A. and Di Capua, E. (1982) 'The helicity of DNA in complexes with recA protein.', *Nature*. England, 299(5879), pp. 185–186. doi: 10.1038/299185a0.

Stiff, T. *et al.* (2005) 'Nbs1 is required for ATR-dependent phosphorylation events.', *The EMBO journal*, 24(1), pp. 199–208. doi: 10.1038/sj.emboj.7600504.

Strillinger, E. *et al.* (2016) 'Production of halophilic proteins using Haloferax volcanii H1895 in a stirred-tank bioreactor.', *Applied microbiology and biotechnology*. Germany, 100(3), pp. 1183–1195. doi: 10.1007/s00253-015-7007-1.

Stroud, A., Liddell, S. and Allers, T. (2012) 'Genetic and Biochemical Identification of a Novel Single-Stranded DNA-Binding Complex in Haloferax volcanii', *Frontiers in microbiology*. Frontiers Research Foundation, 3, p. 224. doi: 10.3389/fmicb.2012.00224.

Su, S. S. *et al.* (1988) 'Mispair specificity of methyl-directed DNA mismatch correction in vitro.', *The Journal of biological chemistry*. United States, 263(14), pp. 6829–6835.

Sugasawa, K. *et al.* (1996) 'HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro.', *Molecular and cellular biology*, 16(9), pp. 4852–4861. doi: 10.1128/mcb.16.9.4852.

Sugasawa, K. *et al.* (2001) 'A multistep damage recognition mechanism for global genomic nucleotide excision repair.', *Genes & development*, 15(5), pp. 507–521. doi: 10.1101/gad.866301.

Sugasawa, K. *et al.* (2005) 'UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex.', *Cell*. United States, 121(3), pp. 387–400. doi: 10.1016/j.cell.2005.02.035.

Sun, J. and Kong, D. (2010) 'DNA replication origins, ORC/DNA interaction, and assembly of pre-replication complex in eukaryotes', *Acta Biochimica et Biophysica Sinica*, 42(7), pp. 433–439. doi: 10.1093/abbs/gmq048.

Sung, P. and Klein, H. (2006) 'Mechanism of homologous recombination: mediators and helicases take on regulatory functions.', *Nature reviews. Molecular cell biology*. England, 7(10), pp. 739–750. doi: 10.1038/nrm2008.

Svendsen, J. M. *et al.* (2009) 'Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair', *Cell*, 138(1), pp. 63–77. doi:

10.1016/j.cell.2009.06.030.

Sweet, D. M. and Moseley, B. E. (1976) 'The resistance of Micrococcus radiodurans to killing and mutation by agents which damage DNA.', *Mutation research*. Netherlands, 34(2), pp. 175–186. doi: 10.1016/0027-5107(76)90122-6.

SZYBALSKI, W. and BRYSON, V. (1952) 'Genetic studies on microbial cross resistance to toxic agents. I. Cross resistance of Escherichia coli to fifteen antibiotics.', *Journal of bacteriology*, 64(4), pp. 489–499. doi: 10.1128/jb.64.4.489-499.1952.

Tadeo, X. *et al.* (2009) 'Structural basis for the aminoacid composition of proteins from halophilic archea', *PLoS biology*. 2009/12/15. Public Library of Science, 7(12), pp. e1000257–e1000257. doi: 10.1371/journal.pbio.1000257.

Takahashi, K. *et al.* (1988) 'Activation of Ada protein as a transcriptional regulator by direct alkylation with methylating agents.', *The Journal of biological chemistry*. United States, 263(27), pp. 13490–13492.

Takai, K. *et al.* (2000) 'Palaeococcus ferrophilus gen. nov., sp. nov., a barophilic, hyperthermophilic archaeon from a deep-sea hydrothermal vent chimney.', *International journal of systematic and evolutionary microbiology*. England, 50 Pt 2, pp. 489–500. doi: 10.1099/00207713-50-2-489.

Tapias, A., Leplat, C. and Confalonieri, F. (2009) 'Recovery of ionizing-radiation damage after high doses of gamma ray in the hyperthermophilic archaeon Thermococcus gammatolerans.', *Extremophiles : life under extreme conditions*. Germany, 13(2), pp. 333–343. doi: 10.1007/s00792-008-0221-3.

Taylor, A. F. *et al.* (1985) 'RecBC enzyme nicking at Chi sites during DNA unwinding: location and orientation-dependence of the cutting.', *Cell*. United States, 41(1), pp. 153–163. doi: 10.1016/0092-8674(85)90070-4.

Taylor, A. F. and Smith, G. R. (2003) 'RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity.', *Nature*. England, 423(6942), pp. 889–893. doi: 10.1038/nature01674.

Taylor, P. W. (2015) 'Impact of space flight on bacterial virulence and antibiotic susceptibility', *Infection and drug resistance*. Dove Medical Press, 8, pp. 249–262. doi: 10.2147/IDR.S67275.

Thangavel, S. *et al.* (2015) 'DNA2 drives processing and restart of reversed replication forks in human cells', *The Journal of cell biology*. The Rockefeller University Press, 208(5), pp. 545–562. doi: 10.1083/jcb.201406100.

Thorslund, T. and West, S. C. (2007) 'BRCA2: a universal recombinase regulator.', *Oncogene*. England, 26(56), pp. 7720–7730. doi: 10.1038/sj.onc.1210870.

Tokunaga, H., Arakawa, T. and Tokunaga, M. (2008) 'Engineering of halophilic enzymes: two acidic amino acid residues at the carboxy-terminal region confer halophilic characteristics to Halomonas and Pseudomonas nucleoside diphosphate kinases', *Protein science : a publication of the Protein Society*. 2008/06/23. Cold Spring Harbor Laboratory Press, 17(9),

pp. 1603–1610. doi: 10.1110/ps.035725.108.

Trujillo, K. M. *et al.* (1998) 'Nuclease activities in a complex of human recombination and DNA repair factors Rad50, Mre11, and p95.', *The Journal of biological chemistry*. United States, 273(34), pp. 21447–21450. doi: 10.1074/jbc.273.34.21447.

Tyrrell, R. M. (1995) 'Ultraviolet radiation and free radical damage to skin.', *Biochemical Society symposium*. England, 61, pp. 47–53. doi: 10.1042/bss0610047.

Udupa, K. S. *et al.* (1994) 'Novel ionizing radiation-sensitive mutants of Deinococcus radiodurans.', *Journal of bacteriology*. United States, 176(24), pp. 7439–7446. doi: 10.1128/jb.176.24.7439-7446.1994.

Vaithiyalingam, S. *et al.* (2014) 'Insights into eukaryotic primer synthesis from structures of the p48 subunit of human DNA primase', *Journal of molecular biology*. 2013/11/13, 426(3), pp. 558–569. doi: 10.1016/j.jmb.2013.11.007.

Venkitaraman, A. R. (2002) 'Cancer susceptibility and the functions of BRCA1 and BRCA2.', *Cell*. United States, 108(2), pp. 171–182. doi: 10.1016/s0092-8674(02)00615-3.

Verhoeven, E. E. *et al.* (2000) 'Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC.', *The Journal of biological chemistry*. United States, 275(7), pp. 5120–5123. doi: 10.1074/jbc.275.7.5120.

Verweij, J. and Pinedo, H. M. (1990) 'Mitomycin C: mechanism of action, usefulness and limitations.', *Anti-cancer drugs*. England, 1(1), pp. 5–13.

Vrtis, K. B. *et al.* (2021) 'Single-strand DNA breaks cause replisome disassembly.', *Molecular cell*, 81(6), p. 1309–1318.e6. doi: 10.1016/j.molcel.2020.12.039.

Wadsworth, R. I. and White, M. F. (2001) 'Identification and properties of the crenarchaeal single-stranded DNA binding protein from Sulfolobus solfataricus.', *Nucleic acids research*, 29(4), pp. 914–920. doi: 10.1093/nar/29.4.914.

Wakasugi, M. *et al.* (2001) 'Damaged DNA-binding protein DDB stimulates the excision of cyclobutane pyrimidine dimers in vitro in concert with XPA and replication protein A.', *The Journal of biological chemistry*. United States, 276(18), pp. 15434–15440. doi: 10.1074/jbc.M011177200.

Wakasugi, M. *et al.* (2002) 'DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair.', *The Journal of biological chemistry*. United States, 277(3), pp. 1637–1640. doi: 10.1074/jbc.C100610200.

Walker, J. J. and Pace, N. R. (2007) 'Endolithic microbial ecosystems.', *Annual review of microbiology*. United States, 61, pp. 331–347. doi: 10.1146/annurev.micro.61.080706.093302.

Wang, F. *et al.* (2010) 'Crystal structures of RMI1 and RMI2, two OB-fold regulatory subunits of the BLM complex', *Structure (London, England : 1993)*, 18(9), pp. 1159–1170. doi: 10.1016/j.str.2010.06.008.

Wang, G. *et al.* (2004) 'Arsenic resistance in Halobacterium sp. strain NRC-1 examined by

using an improved gene knockout system', *Journal of bacteriology*. American Society for Microbiology, 186(10), pp. 3187–3194. doi: 10.1128/jb.186.10.3187-3194.2004.

Wang, H. and Xu, X. (2017) 'Microhomology-mediated end joining: new players join the team', *Cell & Bioscience*, 7(1), p. 6. doi: 10.1186/s13578-017-0136-8.

Wang, M.-J. *et al.* (2017) 'Hepatocyte polyploidization and its association with pathophysiological processes', *Cell death & disease*. Nature Publishing Group, 8(5), pp. e2805–e2805. doi: 10.1038/cddis.2017.167.

Ward, J. F. (1988) 'DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability.', *Progress in nucleic acid research and molecular biology*. United States, 35, pp. 95–125. doi: 10.1016/s0079-6603(08)60611-x.

Wardell, K. *et al.* (2017) 'RadB acts in homologous recombination in the archaeon Haloferax volcanii, consistent with a role as recombination mediator.', *DNA repair*, 55, pp. 7–16. doi: 10.1016/j.dnarep.2017.04.005.

Was, H. *et al.* (2022) 'Polyploidy formation in cancer cells: How a Trojan horse is born', *Seminars in Cancer Biology*, 81, pp. 24–36. doi: https://doi.org/10.1016/j.semcancer.2021.03.003.

Wendoloski, D., Ferrer, C. and Dyall-Smith, M. L. (2001) 'A new simvastatin (mevinolin)-resistance marker from Haloarcula hispanica and a new Haloferax volcanii strain cured of plasmid pHV2.', *Microbiology (Reading, England)*. England, 147(Pt 4), pp. 959–964. doi: 10.1099/00221287-147-4-959.

Weng, M. *et al.* (2010) 'Repair of mitomycin C mono- and interstrand cross-linked DNA adducts by UvrABC: a new model.', *Nucleic acids research*. England, 38(20), pp. 6976–6984. doi: 10.1093/nar/gkq576.

West, J. B. (2000) 'Physiology in microgravity.', *Journal of applied physiology (Bethesda, Md. : 1985)*. United States, 89(1), pp. 379–384. doi: 10.1152/jappl.2000.89.1.379.

Whitby, M. C. *et al.* (1996) 'Interactions between RuvA and RuvC at Holliday junctions: inhibition of junction cleavage and formation of a RuvA-RuvC-DNA complex.', *Journal of molecular biology*. England, 264(5), pp. 878–890. doi: 10.1006/jmbi.1996.0684.

White, M. F. and Allers, T. (2018) 'DNA repair in the archaea—an emerging picture', *FEMS Microbiology Reviews*, 42(4), pp. 514–526. doi: 10.1093/femsre/fuy020.

White, R. H. (1984) 'Hydrolytic stability of biomolecules at high temperatures and its implication for life at 250 degrees C.', *Nature*. England, 310(5976), pp. 430–432. doi: 10.1038/310430a0.

Williams, T. A. *et al.* (2020) 'Phylogenomics provides robust support for a two-domains tree of life.', *Nature ecology & evolution*. England, 4(1), pp. 138–147. doi: 10.1038/s41559-019-1040-x.

Wilson, J. W. *et al.* (2007) 'Space flight alters bacterial gene expression and virulence and reveals a role for global regulator Hfq', *Proceedings of the National Academy of Sciences of*

the United States of America. 2007/09/27. National Academy of Sciences, 104(41), pp. 16299–16304. doi: 10.1073/pnas.0707155104.

Winker, S. and Woese, C. R. (1991) 'A definition of the domains Archaea, Bacteria and Eucarya in terms of small subunit ribosomal RNA characteristics.', *Systematic and applied microbiology*. Germany, 14(4), pp. 305–310. doi: 10.1016/S0723-2020(11)80303-6.

Winter, K., Born, J. and Pfeifer, F. (2018) 'Interaction of Haloarchaeal Gas Vesicle Proteins Determined by Split-GFP.', *Frontiers in microbiology*. Switzerland, 9, p. 1897. doi: 10.3389/fmicb.2018.01897.

Winters, T. A. *et al.* (1994) 'Removal of 3'-phosphoglycolate from DNA strand-break damage in an oligonucleotide substrate by recombinant human apurinic/apyrimidinic endonuclease 1', *Nucleic Acids Research*, 22(10), pp. 1866–1873. doi: 10.1093/nar/22.10.1866.

Woese, C. R. and Fox, G. E. (1977) 'Phylogenetic structure of the prokaryotic domain: The primary kingdoms', *Proceedings of the National Academy of Sciences*, 74(11), p. 5088 LP-5090. doi: 10.1073/pnas.74.11.5088.

Woese, C. R., Kandler, O. and Wheelis, M. L. (1990) 'Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.', *Proceedings of the National Academy of Sciences of the United States of America*. United States, 87(12), pp. 4576–4579. doi: 10.1073/pnas.87.12.4576.

van Wolferen, M. *et al.* (2013) 'Molecular analysis of the UV-inducible pili operon from Sulfolobus acidocaldarius', *MicrobiologyOpen*. 2013/09/19. Blackwell Publishing Ltd, 2(6), pp. 928–937. doi: 10.1002/mbo3.128.

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*, 293(27), pp. 10524–10535. doi: 10.1074/jbc.TM118.000372.

Wu, H.-Y., Lu, C.-H. and Li, H.-W. (2017) 'RecA-SSB Interaction Modulates RecA Nucleoprotein Filament Formation on SSB-Wrapped DNA', *Scientific Reports*, 7(1), p. 11876. doi: 10.1038/s41598-017-12213-w.

Wu, J.-H. *et al.* (2012) 'Induced polyploidy dramatically increases the size and alters the shape of fruit in Actinidia chinensis.', *Annals of botany*, 109(1), pp. 169–179. doi: 10.1093/aob/mcr256.

Wyatt, H. D. M. *et al.* (2013) 'Coordinated actions of SLX1-SLX4 and MUS81-EME1 for Holliday junction resolution in human cells.', *Molecular cell*. United States, 52(2), pp. 234–247. doi: 10.1016/j.molcel.2013.08.035.

Wyatt, H. D. M. *et al.* (2017) 'The SMX DNA Repair Tri-nuclease.', *Molecular cell*. United States, 65(5), p. 848–860.e11. doi: 10.1016/j.molcel.2017.01.031.

Xu, Z.-Q. and Dixon, N. E. (2018) 'Bacterial replisomes', *Current Opinion in Structural Biology*, 53, pp. 159–168. doi: https://doi.org/10.1016/j.sbi.2018.09.006.

Yakimov, M. M. *et al.* (2015) 'Microbial community of the deep-sea brine Lake Kryos

seawater-brine interface is active below the chaotropicity limit of life as revealed by recovery of mRNA.', *Environmental microbiology*. England, 17(2), pp. 364–382. doi: 10.1111/1462-2920.12587.

Yamaguchi, M., Dao, V. and Modrich, P. (1998) 'MutS and MutL activate DNA helicase II in a mismatch-dependent manner.', *The Journal of biological chemistry*. United States, 273(15), pp. 9197–9201. doi: 10.1074/jbc.273.15.9197.

Yamamoto, A. *et al.* (2000) 'Requirement for Phe36 for DNA binding and mismatch repair by Escherichia coli MutS protein.', *Nucleic acids research*, 28(18), pp. 3564–3569. doi: 10.1093/nar/28.18.3564.

Yang, X.-W. *et al.* (2022) 'MutS functions as a clamp loader by positioning MutL on the DNA during mismatch repair', *Nature Communications*, 13(1), p. 5808. doi: 10.1038/s41467-022-33479-3.

Yang, Y. and Davis, T. M. (2017) 'A New Perspective on Polyploid Fragaria (Strawberry) Genome Composition Based on Large-Scale, Multi-Locus Phylogenetic Analysis', *Genome biology and evolution*. Oxford University Press, 9(12), pp. 3433–3448. doi: 10.1093/gbe/evx214.

Yokoi, M. *et al.* (2000) 'The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA.', *The Journal of biological chemistry*. United States, 275(13), pp. 9870–9875. doi: 10.1074/jbc.275.13.9870.

Yokoyama, H. and Mizutani, R. (2014) 'Structural biology of DNA (6-4) photoproducts formed by ultraviolet radiation and interactions with their binding proteins', *International journal of molecular sciences*. MDPI, 15(11), pp. 20321–20338. doi: 10.3390/ijms151120321.

Zack, T. I. *et al.* (2013) 'Pan-cancer patterns of somatic copy number alteration.', *Nature genetics*. United States, 45(10), pp. 1134–1140. doi: 10.1038/ng.2760.

Zakharyevich, K. *et al.* (2012) 'Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase.', *Cell*, 149(2), pp. 334–347. doi: 10.1016/j.cell.2012.03.023.

Zeitlin, C. *et al.* (2004) 'Overview of the Martian radiation environment experiment.', *Advances in space research : the official journal of the Committee on Space Research (COSPAR)*. England, 33(12), pp. 2204–2210. doi: 10.1016/s0273-1177(03)00514-3.

Zenke, R. *et al.* (2015) 'Fluorescence microscopy visualization of halomucin, a secreted 927 kDa protein surrounding Haloquadratum walsbyi cells', *Frontiers in microbiology*. Frontiers Media S.A., 6, p. 249. doi: 10.3389/fmicb.2015.00249.

Zerulla, K. *et al.* (2014) 'DNA as a phosphate storage polymer and the alternative advantages of polyploidy for growth or survival', *PloS one*. Public Library of Science, 9(4), pp. e94819–e94819. doi: 10.1371/journal.pone.0094819.

Zhang, J. and Walter, J. C. (2014) 'Mechanism and regulation of incisions during DNA interstrand cross-link repair.', *DNA repair*. Netherlands, 19, pp. 135–142. doi:

10.1016/j.dnarep.2014.03.018.

Zhang, Q. M. *et al.* (1998) 'Escherichia coli MutY protein has a guanine-DNA glycosylase that acts on 7,8-dihydro-8-oxoguanine:guanine mispair to prevent spontaneous G:C-->C:G transversions.', *Nucleic acids research*, 26(20), pp. 4669–4675. doi: 10.1093/nar/26.20.4669.

Zhao, X. *et al.* (2017) 'Cell cycle-dependent control of homologous recombination.', *Acta biochimica et biophysica Sinica*. China, 49(8), pp. 655–668. doi: 10.1093/abbs/gmx055.

Zheng, L. and Shen, B. (2011) 'Okazaki fragment maturation: nucleases take centre stage', *Journal of molecular cell biology*. Oxford University Press, 3(1), pp. 23–30. doi: 10.1093/jmcb/mjq048.

Zhou, Z. *et al.* (2018) 'Two or three domains: a new view of tree of life in the genomics era', *Applied Microbiology and Biotechnology*, 102. doi: 10.1007/s00253-018-8831-x.

Zhu, Z. *et al.* (2008) 'Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends.', *Cell*, 134(6), pp. 981–994. doi: 10.1016/j.cell.2008.08.037.

Zilsel, J., Ma, P. H. and Beatty, J. T. (1992) 'Derivation of a mathematical expression useful for the construction of complete genomic libraries.', *Gene*. Netherlands, 120(1), pp. 89–92. doi: 10.1016/0378-1119(92)90013-f.

Zivanovic, Y. *et al.* (2009) 'Genome analysis and genome-wide proteomics of Thermococcus gammatolerans, the most radioresistant organism known amongst the Archaea', *Genome biology*. 2009/06/26. BioMed Central, 10(6), pp. R70–R70. doi: 10.1186/gb-2009-10-6-r70.

van Wolferen M, Ma X, Albers SV. DNA processing proteins involved in the UV- induced stress response of Sulfolobales. *J Bacteriol* 2015; 197:2941-51.

Stroud A, Liddell S, Allers T. Genetic and Biochemical Identification of a Novel Single-Stranded DNA-Binding Complex in Haloferax volcanii. Front Microbiol. 2012 Jun 18;3:224. doi: 10.3389/fmicb.2012.00224. PMID: 22719738; PMCID: PMC3376784.

Jang SW, Kim JM. The RPA inhibitor HAMNO sensitizes Fanconi anemia pathway-deficient cells. Cell Cycle. 2022 Jul;21(14):1468-1478. doi:10.1080/15384101.2022.2074200. Epub 2022 May 11: PMID: 35506981; PMCID: PMC9278452

Abdullah UB, McGouran JF, Brolih S, Ptchelkine D, El-Sagheer AH, Brown T, McHugh PJ. RPA activates the XPF-ERCC1 endonuclease to initiate processing of DNA interstrand crosslinks. EMBO J. 2017 Jul 14;36(14):2047-2060. doi: 10.15252/embj.201796664. Epub 2017 Jun 12. PMID: 28607004; PMCID: PMC5510000.

Nowosielska A, Smith SA, Engelward BP, Marinus MG. Homologous recombination prevents methylation-induced toxicity in Escherichia coli. Nucleic Acids Res. 2006 May 2;34(8):2258-68. doi:10.1093/nar/gkl222. PMID: 16670432; PMCID: PMC1456334,

Pecoraro V, Zerulla K, Lange C, Soppa J. Quantification of ploidy in proteobacteria revealed the existance of monoploid, (mero-)oligoploid and polyploid species. PLoS ONE. 2011 Jan 31;6(1):e16392. doi:10.1371/journal.pone.0016392. PMID:21305010; PMCID: PMC3031548.

Vilenchik MM, Knudson AG 2003. Endogenous DNA double-strand breaks: Production,

fidelity of repair and induction of cancer. *Proc Natl Acad Sci* 100: 12871-12876

Eker AP, Quayle C, Chaves I, van der Horst GT. DNA repair in mammalian cells: Direct DNA damage reversal: elegant solutions for nasty problems. Cell Mol Life Sci. 2009 Mar;66(6):968-80. doi: 10.1007/s00018-009-8735-0. PMID: 19153659.

Li X, Heyer WD. Homologous recombination in DNA repair and DNA damage tolerance. Cell Res. 2008 Jan;18(1):99-113. doi: 10.1038/cr.2008.1. PMID: 18166982; PMCID: PMC3087377,

Ohnishi K, Ohnishi T. The biological effects of space radiation during long stays in space. Biol Sci Space. 2004 Dec;18(4):201-5. doi: 10.2187/bss.18.201. PMID: 15858386

May PF, Zawadzki P, Sherratt DJ, Kapanidis AN, Arciszewska LK. Assembly, translocation, and activation of XerCD-dif recombination by FtsK translocase analyzed in real-time by FRET and two-color tethered fluorophore motion. Proc Natl Acad Sci U S A. 2015 Sep 15;112(37):E5133-41. doi: 10.1073/pnas.1510814112. Epub 2015 Aug 31. PMID: 26324908; PMCID: PMC4577173.

Graham JE, Sivanathan V, Sherratt DJ, Arciszewska LK. FtsK translocation on DNA stops at XerCD-dif. Nucleic Acids Res. 2010 Jan;38(1):72-81. doi: 10.1093/nar/gkp843. Epub 2009 Oct 23. PMID: 19854947; PMCID: PMC2800217.

Sherratt DJ, Arciszewska LK, Crozat E, Graham JE, Grainge I. The Escherichia coli DNA translocase FtsK. Biochem Soc Trans. 2010 Apr;38(2):395-8. doi: 10.1042/BST0380395. PMID: 20298190.

Grainge I, Lesterlin C, Sherratt DJ. Activation of XerCD-dif recombination by the FtsK DNA translocase. Nucleic Acids Res. 2011 Jul;39(12):5140-8. doi: 10.1093/nar/gkr078. Epub 2011 Mar 2. PMID: 21371996; PMCID: PMC3130261.

Shimokawa K, Ishihara K, Grainge I, Sherratt DJ, Vazquez M. FtsK-dependent XerCD-dif recombination unlinks replication catenanes in a stepwise manner. Proc Natl Acad Sci U S A. 2013 Dec 24;110(52):20906-11. doi: 10.1073/pnas.1308450110. Epub 2013 Nov 11. PMID: 24218579; PMCID: PMC3876235.

Ferreira H, Butler-Cole B, Burgin A, Baker R, Sherratt DJ, Arciszewska LK. Functional analysis of the C-terminal domains of the site-specific recombinases XerC and XerD. J Mol Biol. 2003 Jun 27;330(1):15-27. doi: 10.1016/s0022-2836(03)00558-8. PMID: 12818199.

Stolz R, Yoshida M, Brasher R, Flanner M, Ishihara K, Sherratt DJ, Shimokawa K, Vazquez M. Pathways of DNA unlinking: A story of stepwise simplification. Sci Rep. 2017 Sep 29;7(1):12420. doi: 10.1038/s41598-017-12172-2. PMID: 28963549; PMCID: PMC5622096.

Schmidt TG, Batz L, Bonet L, Carl U, Holzapfel G, Kiem K, Maatulewicz K, Niermeier D, Schuchardt I, Stanar K. Development of the Twin-Strep-tag and its application for purification

of recombinant proteins from cell culture supernatants. Protein Expr Purif. 2013 Nov;92(1):54-61. Doi: 10.1016/j.pep.2013.08.021. Epub 2013 Sep 6. PMID: 24012791.