Foundations of experimentally-evolved ionizing radiation resistance in *Escherichia coli*

By

Steven T. Bruckbauer

A dissertation submitted in partial fulfillment of the requirements for the degree of

Degree of Philosophy

(Microbiology)

at the

University of Wisconsin – Madison

2018

Date of the final oral examination: May 7th, 2018

The dissertation is approved by the following members of the Final Oral Committee:

Michael M. Cox, Professor, Biochemistry

Jue D. Wang, Professor, Bacteriology

James L. Keck, Professor, Biochemistry

Richard L. Gourse, Professor, Bacteriology

David C. Schwartz, Professor, Genetics

Wesley Culberson, Assistant Professor, Medical Physics

dedicated to

my mentors

Acknowledgements

The following is a very long series of acknowledgments. This is for very good reason – I would not be where I am today without the support of very many people! If I have forgotten anyone, it isn't due to lack of recognizing your support, it is because I have been writing for a long time now and I am getting forgetful. Therefore, this first acknowledgement goes out to anyone that I may not have, but should have, mentioned!

There is no better place to begin than by thanking Mike. Not being a Biochemist by training, when I came to the University of Wisconsin I had no awareness of the Cox lab. I 'knew' that I wanted to work on pathogens for my thesis. However, as soon as Mike spoke to my cohort during orientation about the research in his lab I was immediately drawn in. I had a soft-spot for RecA since my bacterial genetics course at my undergraduate university, and the extremely ionizing radiation resistant bacterium (*Deinococcus radiodurans*) was new to me, but sounded amazing. Above it all, the project to evolve IR resistant *Escherichia coli* was too enticing to ignore. After working through my three lab rotations to determine where I would do my thesis work, I was conflicted. All were such great options! But the final straw came back to Mike. Running a research lab, teaching, and mentoring graduate and undergraduate students alike were my eventual goals; Mike looked to be the perfect mentor to get me to this position. To be sure of the choice I was considering, I contacted a former graduate student of Mike's who was also a Microbiology. She told me that throughout graduate school, my friends will likely often complain about their mentors – and that with Mike, I will have nothing to contribute. I can truly say that Mike has lived up to being the model of a mentor that I one day hope to become, and much more. I especially appreciate the independence that I have been given in my time here, and just how wide Mike has opened the door for me into the scientific community. I have been to nine conferences, worked in two other

labs, met scientists who have come to Mike's lab on sabbatical, and been given the opportunity to work on a collaborative project with the lab of Professor Michael Sussman here at UW. The amount of trust and support Mike has given me is something that I fully recognize and have worked to use well. Additionally, Mike has instilled in me a love of fine wine and scientific discussions (where some of the best ideas are made)!

I also especially thank each member of my thesis committee. I thank Professor Jade Wang, for consistently being a second mentor to me, and declaring me an honorary member of the Wang lab. Jade has always inspired me to think big, and to always think of a model! I thank Professor Jim Keck for being the model of exceptionally smart, attentive, supportive, and kind mentor. I am grateful for the group meetings that we have with the Keck lab so that I could be introduced to Jim, and have the pleasure of his consistently good feedback, advice, and company. I thank Professor Rick Gourse for excellent feedback on my thesis work, and being an model Professor, mentor, and scientist. I remember my introduction to Rick during my interview weekend at UW vividly, and that meeting was part of what drew me into the program (and for introducing me to the magic spot, ppGpp). I thank Professor Dave Schwartz for his enthusiasm, and inventive takes on my research and science in general. The class I took that Dave taught was one of the most enjoyable that I had at UW, and Dave's encouragement to think outside the box has been one of the defining lessons of my time here. Last but not least, I thank Assistant Professor Wes Culberson. Wes was a late-comer to my committee, but a much welcome addition. My research, and I think the field of radiation resistance in general, was much in need of the perspective of a Medical Physicist to set our views on radiation straight. Wes joined my committee with immediate interest and started providing great feedback right-away. I am very grateful for his enthusiasm, advice, and eagerness to help.

I have been lucky in that my mentors have extended beyond my immediate committee. Professor John Battista (of Louisiana State University) was gracious enough to welcome me into his lab, where my thesis work originated many years ago, for two weeks while I was a fairly new graduate student. His mentorship on radiobiology was greatly appreciated, and I wish that we would have had more discussions throughout my graduate career. I look forward to his feedback and comments on my thesis work. Professor Nick Dixon, Professor Antoine van Oijen, and Andrew Robinson were kind enough to welcome myself, Mike, and another graduate student Tyler Stanage into their labs at the University of Wollongong for two weeks. This trip was one of the most enlightening in my life. Their mentorship and unfailing willingness to help (extended to the rest of their labs, as well) was incredible. The scientific discussion with each of them and their labs members was always just as entertaining as enlightening. Working with all of them has shown me the great benefits of the international scientific community, and I hope to maintain these relationships and build new ones throughout my scientific career.

I thank Professor Mike Sussman here at the University of Wisconsin for his mentorship and willingness to involve me in a new collaborative effort with Mike Cox (part of which has become some of my thesis work presented here). Mike has an uncanny ability to motivate me and get my enthusiasm and creativity to overflow. This collaboration has been fruitful thus far, and has really opened my eyes into new methods and new ways of thinking. I look forward to where this collaboration has yet to go!

I thank the labs of all those mentioned above, with special thanks to Dr. Benjamin Minkoff (Sussman lab), Dr. Greg Sabat (UW Mass spectrometry facility), Tricia Windgassen (Keck lab), Kasia Dzubiel (Keck lab), Andy Voter (Keck lab), Brent Anderson (Wang lab), Sabari Thirupathy (formerly Wang lab), KQ Liu (formerly Wang lab), Jessica Tse (formerly Wang lab), Ponlkrit

Yessen (Wang lab), and Jeremy Schroeder (Wang lab). Without the advice, mentorship, and hard work of those listed here much of this work could not be possible.

I additionally thank all the other Professors, research scientists, postdoctoral researchers and graduate students that I have met at conferences. All of these meetings have expanded my view and perspective on the scientific community, and I have thoroughly enjoyed all the new connections that I have made.

Thank you to Dan Anderson of Medical Physics for spending two and a half years in the WIMR basement irradiating *E. coli* with me. Dan has always been a source of entertainment and physics knowledge. He (and the rest of Medical Physics) have been incredibly helpful in evolving new populations of *E. coli* using the Linac – without them, none of this would be possible. To Dan and whomever in Medical Physics that will be sitting through our weekly rounds of evolution – 'See you at the same Bat-place at the same Bat-time!' (quote courtesy Batman, via Dan Anderson).

Thank you to the Cox lab! Thank you Angie Gruber for being my first mentor and introduction to the Cox lab. Thank you to TJ Kim, Erin Ronayne, and Stefanie Chen for being the wise elders that I needed when first joining the lab. Stefanie especially was an excellent mentor, and was generous enough to help me during my (frequent) troubleshooting crises. Thank you to Zach Romero and Miguel Osorio, who have been fantastic additions to the Cox lab, and have taken the unusual path of mentoring the elder graduate student (myself) in Biochemistry. I appreciate it! Thank you to Kanika Jain, for being an excellent new bay-mate, and another great addition to the Cox lab! The future of the lab is in good hands.

A special thank you to Camille Henry, who is one of the kindest and most helpful people that I know. I was thrilled to here than she was coming back to the lab as a postdoctoral researcher,

and she has been an excellent source of advice and scientific discussion! The lab has been a better place with her here.

Thank you to Tyler Stanage, a former graduate student that joined the lab when I did. The infamous ‘Steven Tyler’ duo was a force to be reckoned with. I can’t express my appreciation more for all of his help in Biochemistry, and I hope my bacteriology advice has been just as helpful. I couldn’t have asked for a better lab mate to travel to Australia, Colorado, and the Library with. I wish him the best of luck in science, and can only hope that I will be lucky enough to remain in the same field as him so we may collaborate in the future!

I have been lucky to have two extraordinary undergraduate students to mentor in my time here: Mary Menhart and Joe Trimarco. I thank Mary for dealing with me as a brand new graduate student, as we had to learn alongside each other the ways of the lab. I appreciate the late-night help plating well over a hundred plates – it was much more enjoyable with her help! She was an incredible undergrad to mentor, and I wish her the best in her graduate studies on anthropology and evolution. Thank you to Joe Trimarco, who amazingly continued the streak of exceptional undergrad mentees. I think the accomplishment I am most proud of in grad school is convincing Joe to continue his research career in microbiology (although it may be focused on virology). He has been incredibly smart and reliable. I don’t know how I will carry on without a dependable extra set of hands! But I am very excited for Joe’s future career, and can tell that he is smart enough, talented enough, and driven enough to go far. I hope I have taught him the important ways of grad school (including the need to have fun!).

Thank you to Carol Pfeffer, who has made my time in the Cox lab flow much more smoothly than I could have managed on my own. Carol has to deal with very busy, very forgetful

people, and she somehow makes keeping all those balls in the air look easy. And, importantly, she is always such a pleasure to talk to!

Thank you to Sindhu Chitteni-Patu. I have only had the pleasure of working with her once on electron microscopy work, but from my own results and many others I know how extremely talented she is. Her talents are not the only reason that we are so lucky to have her in the lab, as she is one of the cheeriest, kindest people that I know, and it always brightens my day to talk to Sindhu.

Last but not least, thank you to Liz Wood. There are very few, if any other, labs that have someone like Liz. Although I had a cloning background in my undergraduate lab, I can say that there is no way I could be where I am today with Liz to help my project run smoothly. Amazingly, it is not just my work that she allows to run smoothly, she keeps the entire lab running smoothly (not just the projects!). She is an incredible resource of wisdom, and past Cox lab gossip. Liz is always fun to talk to, and has been an incredible lab member to share my entire time in the Cox lab with. To top it all of, her cats have amazing ‘Wood’ themed names!

Thank you to all the amazing friends that I have made here. Without all of you, none of this would be possible. The fun, the conversation, and the support of everyone is invaluable. It’s been my privilege to watch as everyone has progressed in their own scientific careers, has gotten engaged, been married, and has achieved so much more. I look forward to seeing what paths everyone takes, and I hope that our paths cross in the future!

I must especially thank Brent Anderson, who has been a consistent friend scientifically and beyond in my time here. We’ve shared many beers and many ideas, which has gotten me through many tough days of grad school. We also made it in and out of Minhas Distillery/Brewery alive,

which I'm still not convinced was a real place. I hope that we stay in similar fields, so that we can continue to collaborate and eventually tell embarrassing stories about each other to our students.

Thank you to the coordinator of the Microbiology Doctoral Training Program, Cathy Davis Gray. I can say without a doubt that Cathy is one of the reasons I came to UW. Cathy is incredibly helpful, kind, and funny, and has really instilled in me my love for the program and the university. She has consistently supported me and advocated for students in general. Without her enthusiasm, I never would have gotten involved in the Student Recruitment Team, which has been one of my greatest pleasures here at UW. And, thanks to her, we are now getting an MDTP student retreat started! Thank you Cathy!

Thank you to my previous mentors in the lab of Professor Herbert Schweizer at Colorado State University. Everyone there provided my first look into microbiology and research, and without such an incredible experience I never would have gotten to where I am today. In particular, thank you to Assistant Professor Brian Kvitko, who mentored me as a brand new undergraduate student that knew little to nothing when I started. Brian was an amazingly kind, fun, smart, and (when needed) tough mentor. His mentorship allowed me to become the scientist that I am, and I have taken the mentorship lessons learned from Brian and applied them in my own career. Brian's inventiveness and determination is something that I have carried with me in my approach to research.

Thank you to Associate Professor RoxAnn Karkhoff-Schweizer for being the one who first opened the door to the world of science for me. With nothing more than a high-school transcript, she let me into the lab and mentored me in the basics of microbiology and bacterial genetics. The investment that she put in me is something I have never forgotten, and never will forget. Her ability

to always stay in touch with the ever-growing list of students that she has trained is incredible. I can only hope to eventually be such a positive influence on future students!

I thank my friends Jason Lewis, John Bloemker, Amanda Nickels, and Jeno Saghi. They all have been a second family to me and have consistently supported me through my undergraduate and graduate careers in their own special way. I can only hope that we continue to have reunions throughout the years!

I thank my family for giving me the foundation that has allowed me to get where I am today. The hard work-ethic, critical thinking, and independence I developed from you all has given me the perfect set of tools to get me through graduate school, and to become a scientist. All the lessons that I have learned will serve me very well as my career continues! Thank you for your unconditional support and love, and for believing that I could make it to this point.

My final acknowledgment goes to Bayleigh Benner. Of all my ‘discoveries’, Bayleigh has been the greatest and most important. There are many reasons I am grateful to have come to Wisconsin for my graduate work, but above all is that if I had not come here I would never have met Bayleigh. Bayleigh is incredibly smart, kind, and funny, and has been my love of my life, perfect partner, and Chosen One for over two years now. She has listened to my many (many) complaints when my research goes wrong, been there for my triumphs and failures, and has consistently been a source of reliable advice, wisdom and ideas for life and for research (not to mention, she is the reason that this thesis is not riddled with errors!). Our adventures outside of the lab have made the tough days of graduate school, and the best days, so much better. Bayleigh has provided an example of the scientist, and person, that I strive to be. I could only wish that we could tackle a research project together, because I know we could accomplish anything we set out to do. Thank you, Bayleigh, for being by my side. It is no Gag that we were meant to evolve together.

Abstract

Ionizing radiation (IR) is a ubiquitous stress capable of directly ionizing and indirectly oxidizing all cellular macromolecules. All organisms encode in their genome the proteins necessary to repair DNA damage caused by IR and ameliorate reactive oxygen species (ROS) formed due to radiolysis of intracellular H_2O; however, the extent of damage caused by IR is such that most organisms are sensitive to irradiation. Despite the types of IR-induced damage being well-known, the determinants of whether a cell survives IR exposure are not clear. Studies of IR-resistant organisms have yielded insights into extraordinary methods of DNA repair and ROS amelioration, but there is no clear answer as to what the foundational requirements for IR resistance are. Utilizing directed evolution, we have generated populations of IR-resistant *E. coli* and harnessed deep-sequencing technology to determine the profile of mutations that occur in these populations. By isolating the mutations which increase IR-resistance, the evolving bacteria can then 'tell' us what alterations to existing cellular machinery can allow a cell to become IR-resistant. We are continuing to evolve four replicate populations using high energy electron beam IR in an ongoing effort to generate the most IR-resistant organisms ever produced in the laboratory. We are additionally utilizing state-of-the-art mass spectrometry technology to categorize the full profile of oxidative modifications to the proteome of *E. coli*. In sum, this thesis work aims to build and expand a catalogue of genetic changes that can incrementally enhance an organism's ability to withstand irradiation, and furthermore determine the nature of IR-induced protein damage in order to prevent this damage from occurring.

Table of Contents

Table of Figures

Table of Tables

Chapter I: Introduction

1.1 What is ionizing radiation?

Ionizing radiation (IR) is radiation capable of ionizing molecules. IR may be particulate (α particles [neutrons], β particles [electrons], protons, and more) or electromagnetic (γ rays or X rays [photons]). These particles have a differential ability to penetrate matter (in ascending order with from α, to β particles to γ rays and X rays) and therefore deposit energy along different track lengths (α particles deposit energy along a shorter track, while the same amount of energy is deposited along increasingly longer tracks with β particles to γ rays and X rays) [1](Figure 1.1). Both types of IR deposit energy through interactions with molecules; however, electromagnetic IR is capable of both directly and indirectly ionizing molecules through imparting energy to and ejecting electrons from molecules that are capable of further ionization events (known as the Compton effect) [1].The γ or X rays can continue to travel through matter after depositing some portion of energy to an electron, further potentiating damage. The dose deposited by electromagnetic and particulate IR is measured in a variety of ways, but for the purposes of this work the unit Gray (Gy) will be used, which is defined by the Joules deposited per kilogram of irradiated material.

The majority of background radiation is non-ionizing (ex: UV, radio waves, and microwaves); however, IR is produced from naturally decaying radioactive and cosmic sources. Though the value may vary by location, the annual background dose of IR does not exceed 400 mGy, with an average annual dose of approximately 2.4 mGy [1, 2]. This annual dose may increase; activities such as air travel and exposure to certain medical instruments will impart a small, but significant, dose of IR. However, without catastrophic events, a full-body dose of 5000 mGy, the lethal dose for a human, is highly unlikely.

Manmade irradiators can produce γ-rays (^{60}Co and ^{137}Cs irradiators), X-rays (through photon-emission of metal plates struck by electrically generated high-energy electrons), and high energy particles including β-particles (electron accelerators). Each type of irradiator has their own benefits and drawbacks **[3]**. Of interest to this thesis work is the clinical linear accelerator, or Linac for short. This instrument can generate both high-energy electron beams as well as photon beams, separately, and is commonly used for cancer radiotherapy. The University of Wisconsin Medical Radiation Research Center (UWMRRC), within the Department of Medical Physics, has a Linac appropriated solely for research use, and access to this instrument has made much of this thesis work possible. Importantly, the Linac does not utilize a decaying source to produce IR, and therefore can consistently produce the same known dose over time.

1.2 Damage caused by ionizing radiation

Damage to cells caused by IR is due to primary ionization events, secondary ionization events (from electromagnetic IR), and oxidative damage from reactive oxygen species (ROS) produced by radiolysis of H_2O (Figure 2.1). Radiolysis products of H_2O include hydrogen radicals and hydrated electrons, however the primary species generated are ionized water (H_2O^+) and hydroxyl radicals ($^{\bullet}OH$). Approximately two-thirds of damage caused by IR is due to the indirect effects of ROS, particularly hydroxyl radicals **[2, 4]**. Hydroxyl radicals damage cellular macromolecules indiscriminately and react within 2 nm; therefore, the damage caused by IR is primarily diffusion-limited [5]. Hydrogen peroxide (H_2O_2) and superoxide (O_2^-) are secondary ROS products generated by IR, which are capable of diffusing through cells, generating damage distant from the initial ionization event. H_2O_2 is a key component to the Fenton Reaction where H_2O_2 forces release of Fe^{2+} from protein-coordinated Fe-S clusters. The Fe^{2+} atom can react with

additional H_2O_2 to generate more $^{\bullet}OH$ (Fenton Reaction: $H_2O_2 + Fe^{2+} \Rightarrow Fe^{3+} + {}^{-}OH + {}^{\bullet}OH$) [6]. Within 1 ms post-irradiation, reactions with IR-generated ROS have completed [2]. In order to survive this oxidative burst, cells rely on both enzymatic amelioration of ROS and ROS scavenging by low molecular weight (LMW) antioxidants, such as glutathione. How quickly the LMW antioxidant supply becomes depleted likely determines an organisms' sensitivity to low levels of IR [2].

1.2.1 DNA damage caused by IR

High doses of IR can cause extreme damage to DNA. The abundance and reactivity of hydroxyl radicals generated by IR make them especially problematic; superoxide and hydrogen peroxide are not very reactive with DNA on their own [7, 8]. Oxidation of DNA from hydroxyl radicals may lead to base modification, formation of DNA cross-links, and DNA breaks. Hydroxyl radicals are three to four-fold more likely to induce base modifications than react with the deoxyribose sugar. Modified bases have a wide variety of effects on the helical structure of DNA (which may block DNA replication machinery) as well as base mispairing (which may lead to heritable mutations)[8, 9].

Common forms of base damage include formation of 8-hydroxyguanine (8-oxodG), and cytidine oxidation to cytidine glycol. Cytidine glycol is subject to deamination to uracil glycol and dehydration to 5-hydroxy-2'-deoxyuridine (5-OH-dU). Of the four bases, guanine is particularly susceptible to oxidative damage [10]. 8-oxodG is the most abundant oxidized purine product detected in DNA and in the intracellular nucleotide pool post-irradiation, and leads to GC > TA or AT > CG transversions, respectively (due to incorporated 8-oxodG being mispaired with A, versus

incorporated A being mispaired with free 8-oxodG). 5-OH-dU leads to GC > AT transitions with 83% frequency [8, 11]. Unsurprisingly, GC > AT transitions are the most common base substitution arising from oxidative DNA damage, followed by GC > TA transversions [8, 12, 13]. As such, *E. coli* encodes a variety of defenses against 8-oxoG and 5-OH-dU in its genome (see *1.3 Mechanisms of natural ionizing radiation resistance*).

Despite the mutagenic potential of 8-oxodG, the adenine analogue 8-oxodA is a neutral modification in prokaryotes (as it is preferentially paired with T), although it is mutagenic in eukaryotic cells. 8-oxodA causes only a minor conformational change in DNA, unlike 8-oxodG, and therefore is not a block to replication [8, 14]. 2-Hydroxy-2'-deoxyadenosine (2-OH-dA) is another, and more mutagenic, oxidized product of adenine. It has been observed that adenine in the nucleotide pool, as opposed to that in duplex DNA, is preferentially oxidized to 2-OH-dA, indicating that 2-OH-dA must be incorporated. 2-OH-dA can lead to differing mutations dependent in which strand it has been incorporated: -1 deletions if 2-OH-dA was incorporated in the '+' strand, or A > T or A > G substitutions if incorporated in the '-' strand [8].

A major oxidative product of thymidine is thymidine glycol or dTg. Thymidine glycol induces structural changes in the DNA helix, blocking DNA replication. When bypass of dTg does occur, introduction of a T > C transition mutation is the likely outcome [15].

ROS can additionally cause breaks in the DNA phosphate-backbone through oxidative attacks on deoxyribose. Radical formation on the C4 of deoxyribose due to hydroxyl radical damage may lead to abasic sites and DNA strand breaks. As single-strand breaks (SSBs) accumulate, the likelihood that breaks will occur in closer proximity on opposite strands increases which leads to double-strand breaks (DSBs) (when two SSBs occur on opposite strands within approximately 12 nucleotides). IR can also cause DSBs through direct ionization of DNA

molecules. A single unrepaired DSB is lethal and is a major cause of cell death from IR [1]. It is estimated that IR-induced DSBs accumulate at a rate of 0.003 – 0.006 DSBs/Gy/Mbp of DNA [16].

Importantly, DNA damage caused by IR-generated ROS is clustered as opposed to lone events caused by endogenous ROS. Clusters of DNA damage including base modifications, abasic sites, bulky adducts, along with SSBs and DSBs near these sites of damage can prove to be difficult to repair for typical cellular DNA repair mechanisms. [14, 17, 18].

1.2.2 Protein damage

Protein damage by IR has been argued to be an important factor, if not the main determinant, of cell survival [6, 16, 19-22]. Indeed, previous studies with eukaryotic cells have suggested that proteins are the primary target of IR-produced ROS [23]. However, it is important to note than in bacterial cells, unlike eukaryotes, DNA is the macromolecule which comprises most of the intracellular volume [1]. There are still many unknowns as to the importance of protein damage as a determinant of cell survival in bacterial cells, and part of this thesis work aims to begin to determine the role of IR-induced protein damage. However, studies of oxidative damage to proteins *in vitro* and in eukaryotic cells will likely provide a helpful guide to the study of protein damage in bacteria.

There are conflicting reports as to whether hydroxyl radicals more readily attack the protein amide backbone [2] or amino acid side chains [24]. However, studies that have irradiated lyophilized or frozen samples (to eliminate production of secondary ROS from radiolysis of H_2O) have shown that primary ionization events on proteins result in peptide cleavage [3]. Peptide

backbone cleavage most likely results from attack at the alpha-carbon in the protein backbone; however, cleavage of the peptide backbone has also been observed due to radical transfer from oxidized proline side chains, and the gamma and beta carbons of side chains [24]. As such, protein cleavage is a likely and severe outcome of irradiation.

Amino acid side chain modification is a major outcome of IR damage to proteins. Methionine, cysteine, tryptophan, tyrosine, phenylalanine and histidine are the most reactive amino acids with hydroxyl radicals, although there is disagreement as to the exact order of reactivity [2] [25] [24]. Oxidation of methionine to methionine sulfoxide (MetO) is a common event from ROS (although hydroxyl radicals preferentially attack the amino group of free methionine) [2] [25] [24]. Methionine sulfoxide can be irreversibly further oxidized to methionine sulfone. Cysteine residues can be oxidized first to sulfenic acid, then (irreversibly) to sulfinic acid, and finally sulfonic acid. Hydroxyl radicals and superoxide will first oxidize the thiol group of cysteine to a thiol radical, which can react again to form sulfenic acid or form disulfide bonds with a nearby cysteine or low molecular weight thiols, such as glutathione [26, 27]. Additionally, hydrogen radicals and aqueous electrons can reduce cysteine to alanine [2, 25].

Protein carbonylation is an irreversible modification due to ROS. In eukaryotic studies, it has been shown that protein carbonylation inactivates proteins and marks them for degradation [2]. Arginine, lysine, threonine, cysteine, histidine and proline side chains are susceptible to carbonylation, but solvent accessibility and proximity to another carbonylated residue further govern susceptibility [2, 28]**.** Protein carbonylation is commonly used as marker for protein oxidation due to IR [2, 22]. However, common detection methods of protein carbonylation using immunoblotting are likely to overestimate levels of protein carbonylation, as the reagent used to

derive carbonyls for antibody detection is cross-reactive with sulfenic acid, and nucleic acid contamination of samples can lead to false-positives [2].

While oxidation of cysteine and methionine, and carbonylation are major outcomes of irradiation, there are additional modifications. These species are numerous and have been extensively characterized by Xu and Chance [24]. However, carbonylation (+14 Da), oxidation (+16 Da) and peroxidation (+33 Da) are the most prominent species [23, 24].

The susceptibility of iron sulfur-clusters (commonly [4Fe-4S]), often coordinated by cysteine side chains, to ROS pose another problem for proteins exposed to IR. Iron-sulfur proteins are used to transfer electrons in central metabolic pathways and other reactions and are used to detect oxygen levels and oxidative stress. Destruction of iron-sulfur clusters is caused by electron transfer to one of the coordinated Fe atoms. This newly released Fe^{2+} can further potentiate oxidative stress through the Fenton Reaction (Fenton Reaction: $H_2O_2 + Fe^{2+} \Rightarrow Fe^{3+} + {}^{-}OH + {}^{\bullet}OH$ [6].

1.2.3 Damage to lipids

Oxidation of lipids from IR exposure in bacteria is apparently a fairly under-studied phenomenon. Poly-unsaturated fatty acids (PUFA) are a major target for ROS. Oxidation of PUFAs can lead to a decrease in membrane fluidity, and breakdown of oxidized PUFAs can generate highly reactive aldehyde, which can diffuse to and damage distant sites in the cell [29]. However, most bacteria do not produce PUFA [30], as such lipid peroxidation and ionization, to the extent that they occur in bacteria, likely forms through other pathways. Interestingly, it has been observed that plasma-induced oxidative stress in *E. coli* leads to loss of membrane potential and lipid peroxidation [31]. There is clearly more work to be done to determine the effects of IR on bacterial cell membranes.

1.3 Mechanisms of natural ionizing radiation resistance

1.3.1 DNA repair mechanisms

All organisms rely on a wide-array of DNA repair mechanisms to withstand the damage caused by IR. There are more pathways and proteins involved in DNA repair than are introduced here, however a list of the DNA repair proteins of *E. coli* is provided in Table 1.1.

Base excision repair (BER) through the MutM, MutY, Nth, and Nei glycosylases play an important role in the removal of oxidized DNA bases [8, 10, 15]. The MutM, MutY, and MutT proteins can efficiently remove incorporated 8-oxodG from the genome (through interplay between the purine-specific MutM and MutY glycosylases) and the nucleotide pool (MutT is capable of hydrolyzing 8-oxoGTP to 8-oxoGMP).

The glycosylases endonuclease III (Nth) and endonuclease VIII (Nei) remove oxidized pyrimidines (including 5-OH-dU) [15]. Surprisingly, *E. coli* strains lacking Nth are not IR sensitive, indicating that Nei, and perhaps other endonucleases, may be able complement the loss of Nth [15],[8]. BER through Nth and Nei are also capable of removing thymidine glycol [15]. Nucleotide excision repair (NER), consisting of the UvrABC proteins, has also been implicated in the removal of thymidine glycol [32, 33]. However, NER is not the major pathway for removal of oxidized bases in *E. coli*. This does not completely discount a role for NER in IR resistance, as removal of bulky nucleotide adducts or DNA-crosslinks requires NER [32].

To combat DNA DSBs, cells rely on non-homologous end joining (NHEJ) in which two DNA fragments are ligated together, and homologous recombination (HR) where a sister chromosome is used as a template to repair the broken DNA. NHEJ has potential for loss of genetic information or DNA translocation, as DNA fragments may be ligated together regardless of

sequence context. In contrast, HR is relatively error-free. HR is catalyzed by the RecA protein in bacteria, known as Rad51 in eukaryotes (Figure 3.1**)**. As this thesis work focuses on *E. coli*, the following discussion of homologous recombination will focus on the proteins known in this organism. At a DSB, the heterotrimer RecBCD helicase/exonuclease will load onto a blunt duplex DNA end and unwind DNA in a 3' to 5' direction while degrading DNA from 5' to 3'. This exonuclease activity is attenuated once the RecBCD enzyme reaches a chi sequence (5'-GCTGGTGG-3'), at which point RecA is loaded on the prepared single-strand DNA by RecBCD [34-36] [37, 38]. RecA-bound single-strand DNA can then invade a sister chromosome and search for homologous DNA and begin repair of the parent DNA. Additionally, RecA can repair single-strand gaps in DNA through a RecFOR-mediated pathway [39], and RecA bound to ssDNA can activate the global response to DNA damage, the SOS response [40, 41]. The functions of RecA in DNA repair, and the myriad of proteins involved, have been extensively reviewed [42-44].

1.3.2 ROS amelioration and repair of oxidized protein

Organisms encode a wide array of ROS ameliorating proteins such as catalases, superoxide dismutases, and oxidoreductases to reduce IR-induced ROS stress. These enzymatic mechanisms of reducing ROS come in addition to low molecular weight ROS scavengers, such as glutathione. Expression of genes involved in ROS stress are controlled, in part, by the SoxR and OxyR response regulators which respond to superoxide and hydrogen peroxide stress, respectively. Interestingly, these proteins use oxidation to their advantage; oxidation of the [2Fe-2S] cluster of SoxR and the CXXC motif of OxyR allow for the conformational changes necessary to activate transcription of the respective regulons [45]. The expression of many proteins listed below is controlled by OxyR or SoxR.

An important step to survive ROS stress from IR is to repair damage to highly susceptible methionine and cysteine resides. To combat methionine oxidation, *E. coli* utilizes the MsrA, MsrB, MsrC, and BisC proteins to reduce these oxidized residues [26, 46-50]. *E. coli* also encodes two Msr proteins to specifically repair methionine residues of membrane-bound proteins, MsrP and MsrQ [26, 51]. The effects of reducing agents produced by IR, hydrogen radicals and aqueous electrons, is much less studied than oxidizing agents. However, it has been observed that irradiation can generate methionine reduction to aminobutyric acid [2, 25].

Di-oxidation of cysteine residues results in the irreversible formation of sulfinic acid (cysteine thiols are oxidized to sulfenic acid, then again to sulfinic acid). To prevent sulfinic acid formation *E. coli* encodes thioredoxins (Trx1 and Trx2) as well as glutaredoxins (Grx1, Grx2, and Grx3). These proteins contain their own catalytically active cysteines (in a CXXC motif) which reduce sulfenic acid through oxidation of the CXXC motif to a disulfide bond. Oxidized Trx proteins can be reduced to active Trx through TrxR, and oxidized Grx proteins are reduced using the low-molecular weight antioxidant, glutathione [26]. To repair oxidized methionine residues in envelope proteins, *E. coli* encodes the Trx-like proteins DsbC and DsbG [26, 52, 53].

Destruction of Fe-S clusters and exacerbating ROS stress through participation of the released Fe in the Fenton Reaction (Fenton Reaction: $H_2O_2 + Fe^{2+} \Rightarrow Fe^{3+} + {}^{-}OH + {}^{\bullet}OH$ [6]) is countered through many pathways. In addition to encoding multiple catalase and superoxide dismutase enzymes to reduce H_2O_2 and superoxide, respectively, *E. coli* encodes the Dps protein, which can sequester free Fe^{2+} to prevent this Fenton chemistry. To repair and rebuild Fe-S clusters, the gene products of the *iscRSUA* and *sufABCDSE* operons are responsible for housekeeping maintenance and re-construction during oxidative stress of iron-sulfur cluster proteins,

respectively [54, 55]. Two additional proteins of unknown function, YggX and YtfE, are implicated in repair of Fe-S clusters [45].

1.3.3 IR resistance of *Deinococcus radiodurans*

The study of highly IR-resistant organisms is a subject of fascination as well as a potential source of insights into novel mechanisms of DNA repair and ROS amelioration. The Gram-positive bacterium *Deinococcus radiodurans* is a prominent subject of research into extreme IR resistance, as it is capable of surviving a dose of IR that exceeds the lethal dose for humans by 1000-fold or more [6, 19, 56].

D. radiodurans relies on remarkably efficient DNA repair mechanisms, some of which are unique to this bacterium, as well as an incredible capacity for ROS amelioration to achieve extreme IR resistance [1, 6, 16, 19-22, 57]. In addition to the usual complement of DNA repair pathways (including DNA replication restart, homologous recombination (HR), nucleotide excision repair (NER), base excision repair (BER) and mismatch repair (MMR)), *D. radiodurans* highly expresses several unique DNA repair-associated proteins post-irradiation and desiccation including DdrA [58], DdrB [59], DdrC [60, 61], DdrD [62], IrrE [63] and PprA [64]. See Table 2.1 for a summary of these proteins. PprA, DdrA, DdrB and DdrC are all implicated in single-strand DNA or DNA end binding and are the most abundant transcripts detected in *D. Radiodurans* post-irradiation and desiccation [56]; these observations are highly suggestive of the need to protect the ends of fragmented DNA after these stresses. IrrE appears to be a novel regulator of RecA post-IR exposure [63].

Another proposed mechanism of IR resistance of *D. radiodurans* is an extraordinary capacity to ameliorate ROS generated by IR to avoid protein oxidation [1, 6, 16, 19-22, 57]. In comparison to more IR-sensitive organisms, *D. radiodurans* has a remarkable resistance to protein oxidation caused by IR exposure. Despite this enhanced capacity to protect its proteome, *D. radiodurans* suffers DNA double-strand breaks (DSBs) at a rate similar to other organisms [6]. Taken together, these observations have led to the hypothesis that the integrity of the proteome, which can then repair the genome, is a key determinant for IR resistance of an organism. Indeed, it was discovered that protein-free *Deinococcus* cell lysate could protect protein, but not DNA, at doses of IR as high as 50 kGy. Complexes of Mn^{2+} ions, orthophosphates, and peptides provide the observed protection [16, 19, 22]. The hypothesis that accumulation of Mn^{2+} ions, in conjunction with cellular metabolites, is a strong determinant of IR resistance is supported by data indicating that intracellular manganese:iron ratio of organisms correlates with IR sensitivity (IR resistance increases as the ratio increases) [6, 16].

The radioresistant phenotype of *D. radiodurans* is seemingly not a product of natural selection for IR resistance, as an annual background doses of 2.4 mGy likely would not be a strong enough selection pressure for extreme IR resistance [1]. In the case of this bacterium, extreme IR resistance appears to be a byproduct of natural selection for desiccation tolerance [65]. Organisms with high levels of IR resistance can be found readily in arid environments [66, 67]. No known organism has likely evolved to withstand high doses of IR.

1.4 Experimental evolution

Studies of evolution are incredibly important to understanding the foundations of biology. However, evolution studies in natural contexts are typically limited to observation; it would be

impossible to ‘repeat’ evolution in an ecological context. This is where microbes become excellent models for studying evolution. Their short generation time, ability to be stored indefinitely and relative ease of genetic manipulation makes them ideal models for evolution. As sequencing technology has advanced to the point of whole-genome sequencing, evolution experiments in microbes have become incredibly powerful [68, 69]. Generations of change can be observed easily, allowing for evolution to be followed in real-time both phenotypically and genetically using next-generation sequencing technology. Since microbes can be readily stored, the original population, or the Founder, can be compared against any subsequent generation to directly compare fitness, phenotype, or genotype [68, 69].

There are two major types of evolution experiments: mutation accumulation (without selective pressure) and adaptive (directed) evolution (utilizing a selective pressure). Mutation accumulation experiments repeatedly propagate a single clone from an evolving population. This randomized selection allows for the study of mutation rates during evolution, though the rate will be slightly underestimated due to underrepresentation of lethal and deleterious mutations [69-72] Recent work from the laboratory of Patricia Foster has highlighted the benefits of mutation accumulation experiments through their exploration of the mutation topology accumulated across the *E. coli* genome, where mutation hotspots appear to be dependent on replication forks that are blocked or have experienced collisions [70] [73-75].

Directed evolution experiments can be utilized to select for a specific phenotype and allows the organism to ‘tell’ the researcher where fitness peaks for this selection pressure may be. These types of experiments can be fine-tuned to the goals of the experiment by applying either strong or weak pressures, though either has downsides. Too weak of selection runs the risk of producing populations that are more fit but will not reach the highest possible fitness peak

against that pressure; if the selection is too harsh, the organism could find the ‘easy way out’ and reach an easily attainable fitness peak that will constrain further evolvability [69]. Directed evolution studies can expose biological phenomena that need iterative cycles of selection to make the phenotype in question prominent enough to be observed. For example, studies on *E. coli* cultured to long-term stationary phase in iterative cycles has revealed the underlying selective forces at play in cultures typically thought to be ‘stationary’ [76-78].

As mutations accumulate in experimental evolution experiments, they lead to the formation of diverse fitness landscapes where organisms find their own path to fitness ‘peaks’ depending on the types and order of mutations gained. Adaptations to an environment can be classified as ‘optimization’ or ‘innovation’ depending on how quickly the accumulated mutations benefit fitness. A slow build-up of fitness is considered optimization, as the acquired phenotype is improved overtime. Conversely, innovation is characterized by a sudden major increase in fitness [69]. Optimization and innovation are not mutually exclusive; slow optimization may set the stage for innovation in an ‘all-or-none’ epistatic interaction where the innovative phenotype needs all the previous mutations to occur. Diminishing returns epistasis is also common, where early innovations rapidly increase fitness, but over time new beneficial mutations return decreasing improvements of relative fitness [79].

Sub-populations within evolving populations reach distinct peaks in fitness landscapes, increasing diversity. In asexual populations, the competition of sub-populations is known as clonal interference [80]. Clonal interference will prevent mutations which enhance fitness from fixing (reaching 100% frequency) in a population if other sub-populations with beneficial mutations exist. These differing lineages can co-exist in evolving populations until one is able to

accumulate the beneficial mutations necessary to sweep through the population and drive competitors extinct (Figure 4.1) [69].

One needs not look further than the *E. coli* Long-Term Evolution Experiment (LTEE) conducted by the laboratory of Richard Lenski over the past 30 years to understand the wealth of information packed into microbial evolution experiments [81, 82]. The LTEE was designed to track the adaption of *E. coli* to growth in glucose-limiting minimal medium in a chemostat environment. Now at over 60,000 generations, the 12 replicate populations in the LTEE have demonstrated a wide-variety of evolutionary trends, including heavy clonal interference [82], evolution of cross-feeding within a population (where one sub-population utilizes the glucose in the medium, and the other feeds off of the build-up of acetate waste from the former population) [77] [83], and even observation of a complex series of mutations beginning with gene duplication to allow for aerobic metabolism of citrate [84, 85]. The LTEE has proven to be a brilliant guide and comparator for any long-term evolution experiments.

1.5 Previous ventures in experimental evolution of ionizing radiation resistance

Rather than study the mechanisms of resistance in radioresistant organisms, we have taken a different approach: evolve extreme IR resistance in the laboratory. Directed evolution has been previously used to generate IR-resistance in previously IR-sensitive organisms ([86-89]). The extensively-studied model bacterium *E. coli*, has been observed to generate increasingly radioresistant progeny due to IR exposure [86] However, previous studies have been unable to determine the genetic basis of experimentally-evolved IR resistance. Due to current sequencing

technology, the mechanisms of radioresistance in experimentally-evolved *E. coli* can be investigated with relative ease.

We previously subjected four separate populations of *Escherichia coli* to 20 cycles of ^{60}Co irradiation (sufficient to kill up to 99.9% of the population) followed by outgrowth of the survivors. The 20 cycles resulted in large gains in IR resistance in all four populations, designated IR-1-20, IR-2-20, IR-3-20, and IR-4-20 [13, 90]. Isolates from these populations (designated CB1000, CB2000, CB3000, and CB4000) were highly IR resistant compared to the MG1655 parent strain, exhibiting up to a 1000-fold increase in percent survival at a dose of 3000 Gy (Figure 5.1) [13]. Sequencing of these and other isolates from each of the four populations opened the first-ever window onto evolved IR resistance in *E. coli*. Mutations in DNA repair, ROS amelioration and cell wall metabolism were prominent among these isolates. Of particular interest were the numerous mutations in RecA, the enzyme responsible for catalyzing DNA repair through homologous recombination and triggering the SOS response (which upregulates a wide array of DNA repair genes) in *E. coli* [42-44]. Indeed, the RecA A290S variant was shown to enhance IR-resistance when placed into the parent strain, indicating a role for modified DNA repair proteins in radioresistance [13]. Despite the implicated beneficial role of RecA variants, these mutations were only observed in three of four populations. In fact, only a single mutation, the excision of the cryptic prophage e14, occurred in each of the four populations. Even within population IR-1-20 there were no fixed mutations due to clonal interference.

An isolate from population IR-2-20, CB2000, was previously characterized [90]. Three variants affecting DNA metabolism were found to enhance IR resistance in this isolate: a variant of the DNA repair protein RecA (D276N), the replicative helicase DnaB (P80H), and the putative helicase YfjK (A151D) [90]. Reliance of CB2000 on three variant DNA metabolism proteins for

IR resistance further suggested modification of existing DNA repair mechanisms could enhance radioresistance. Further biochemical characterization of the RecA D276N variant revealed novel activities consistent with repair of genomes fragmented by IR exposure [91]

Unlike the other three populations, IR-3-20 contained no mutations in RecA, DnaB, and YfjK [13, 90]. This population had apparently taken a distinct evolutionary path to IR resistance, suggesting that *E. coli* may utilize mechanisms other than enhanced DNA repair to radioresistance. This thesis work in part focuses on characterizing the unique phenotypes of CB3000, an isolate from population IR-3-20.

As CB3000 work was underway, we came to the realization that CB3000, and each of these initial populations, were an evolutionary dead-end; decaying radioactive sources capable of delivering doses to evolve, and even assay, these populations were no longer available. As such, we sought out a new source to generate *E. coli* with IR resistance comparable to *D. radiodurans*. We have since been employing a clinical linear accelerator (Linac) to generate a high-energy electron beam with a dose rate approximately 4-fold higher than that used to evolve the previous populations (72 Gy/min versus 19 Gy/min). The process of directed evolution of these populations and characterization of these lineages to an unprecedented 100 rounds of selection is detailed in this thesis work.

The final chapter of this thesis focuses a new collaborative effort with the lab of Professor Michael Sussman (University of Wisconsin – Madison). In conjunction with their expertise in and access to state-of-the art mass spectrometry technology, we are undertaking the task of cataloguing all modifications to the *E. coli* proteome caused by IR. With the data generated from this work, we will be able to determine the molecular basis, if any, for cell death by IR due to protein damage. In all, this thesis work aims to determine the foundations of IR resistance in *E. coli*.

1.6 Tables

Protein	Subunit MW	Gene	AA	Molecular Form	Cellular Function
DnaA	52,551	*dnaA*	467	Extended complex	Replication initiation
DnaB	52,390	*dnaB*	471	Hexamer helicase	
DnaC	27,935	*dnaC*	245	Monomer	
PriA	81,655	*priA*	732	Monomer	Replication restart (varied complexes and subcomplexes with DnaB and DnaC)
PriB	11,442	*priB*	104	Dimer	
PriC	20,375	*priC*	175	Monomer	
DnaT	19,455	*dnaT*	179	Trimer	
DnaG	18,975		178	Tetramer	Primase
DiaA	21,106	*diaA*	196	Tetramer	Initiation regulation Magnesium binding
Dps	18,695	*dps*	167	Decamer	DNA protection Iron binding
Hda	26,633	*hda*	233	Monomer	Initiation regulation
HUα	9,535	*hupA*	90	Heterodimer w/HUβ	DNA binding
HUβ	9,226	*hupB*	90	Heterodimer w/HUα	DNA binding
PolA	103,100	*polA*	928	Monomer	DNA polymerase I Repair replication Magnesium binding
SSB	18,975	*ssb*	178	Tetramer	ssDNA binding
PolB	90,053	*polB*	782	Monomer	DNA polymerase II Translesion replication Magnesium binding
DnaE	129,905	*dnaE/ polC*	1,160	α subunit (2-3)	DNA polymerase III Chromosomal replication Magnesium binding
DnaQ	27,099	*dnaQ*	243	ε subunit (2-3)	
DnaX	71,138	*dnaX*	641	τ subunit (2-3)	
HolA	38,704	*holA*	343	δ subunit (1)	
HolB	36,937	*holB*	334	δ‘ subunit (1)	
HolC	16,633	*holC*	147	χ subunit (1)	
HolD	15,174	*holD*	137	ψ subunit (1)	
HolE	8,846	*holE*	76	θ subunit (2-3)	
DnaN	40,587	*dnaN*	366	β subunit (2-3 dimers)	
DinB	39,516	*dinB*	351	Monomer	DNA polymerase IV Translesion replication Magnesium binding
UmuC	47,680	*umuC*	422	UmuD′$_2$C+RecA	DNA polymerase V Tranlesion replication
UmuD′	12,290	*umuD*	115	UmuD′$_2$C+RecA	
UvrA	103,868	*uvrA*	940	Dimer; UvrA$_2$B	UvrABC Excinuclease
UvrB	76,226	*uvrB*	673	UvrA$_2$B or UvrBC	
UvrC	68,188	*uvrC*	610	UvrBC	
UvrD	81,990	*uvrD*	720	dimer	helicase
Rep	77,024	*rep*	673	Monomer/dimer	helicase
Ligase	73,606	*ligA*	671	monomer	DNA ligase Zinc binding
AlkA	31,393	*alkA*	282	dimer	Glycosylase (hypoxanthine)
UNG	25,693	*ung*	229	dimer	Glycosylase (uracil)
Fpg	30,290	*mutM*	269	monomer	Glycosylase (8-oxoG+) Zinc binding
EndoIII	223,562	*nth*	211	monomer	Glycosylase (varies)

Table 1.1A. The DNA repair proteome of *Escherichia coli*.

Protein	Subunit MW	Gene	AA	Molecular Form	Cellular Function
Fpg	30,290	*mutM*	269	monomer	Glycosylase (8-oxoG+) Zinc binding
EndoIII	223,562	*nth*	211	monomer	Glycosylase (varies) 4Fe-4S cluster
EndoVIII	29,845	*nei*	263	monomer	Glycosylase (varies)
MutY	39,149	*mutY*	350	monomer	Glycosylase (8-oxoG) 4Fe-4S cluster
EndoIV	31,480	*nfo*	285	monomer	AP endonuclease Zinc binding
ExoI	54,501	*xbcB* *xonA*	475	monomer	Exonuclease Magnesium binding
ExoIII	30,969	*xthA*	268	monomer	Glycosylase++ Magnesium binding
ExoVIIa	51,832	*xseA*	456	Heterooligomer 4/1	exonuclease
ExoVIIb	8,952	*xseB*	80	Heterooligomer 1/4	
MutS	92,247	*mutS*	853	dimer	Mismatch repair
MutL	67,924	*mutL*	615	dimer	Mismatch repair
MutH	25,527	*mutH*	229	monomer	Mismatch repair
VSR	18,016	*vsr*	156	monomer	Vsp mismatch repair Magnesium binding
Photolyase	53,667	*phrB*	472	monomer	Direct repair TTdimer
Ada	39,324	*ada*	354	monomer	Direct repair O6MeG Zinc binding
AlkB	24,076	*alkB*	216	monomer	Direct repair 3MeC+ Iron binding
RecA	37,842	*recA*	352	Filament on DNA	Recombinase
RecB	133,959	*recB*	1,180	RecBCD Magnesium binding	RecBCD helicase/nuclease
RecC	128,848	*recC*	1,122	RecBCD	
RecD	66,902	*recD*	608	RecBCD	
RecG	76,430	*recG*	693	monomer	Rec helicase
RecN	61,396	*recN*	553	Dimer SMC-like	DSBR
RecQ	68,364	*recQ*	609	monomer	Rec helicase
RecJ	63,389	*recJ*	577	Monomer/dimer	ssDNA exonuclease
RecF	40,514	*recF*	357	Dimer or $RecFR_4$	RecA loader
RecO	27,260	*recO*	242	Monomer/$RecOR_4$	RecA loader
RecR	21,965	*recR*	201	Tetramer/$RecOR_4$/ $RecFR_4$	RecA loader
RecE/ExoVIII	96,334	*recE*	866	Tetramer	Rec exonuclease
RecT	29,721	*recT*	269	Tetramer	DNA binding recomb
RadA	49,472	*radA*	460	monomer	Modulates RecA function
RadD	66,413	*yejH*	586	monomer	ATPase DSBR repair Zinc binding?
DinG	81,440	*dinG*	716	unknown	Helicase/DNA repair 4Fe-4S cluster
Uup	72,067	*uup*	635	unknown	ATPase DSBR repair
RecX	19,293	*recX*	165	monomer	RecA regulator
DinI	8,818	*dinI*	81	monomer	RecA regulator
PsiB	15,853	*psiB*	144	monomer	RecA regulator
RdgC	33,993	*rdgC*	303	dimer	RecA regulator
DinD	31,078	*dinD*	274	unknown	RecA regulator
DprA	40,955	*dprA*	374	monomer	RecA loader
Lon	87,438	*lon*	784	Dimer or tetramer	protease

Table 1.1B. The DNA repair proteome of *Escherichia coli*.

Protein	Subunit MW	Gene	AA	Molecular Form	Cellular Function
LexA	22,358	*lexA*	202	dimer	SOS gene regulator
RNaseHI	17,597	*rnhA*	155	monomer	RNA removal from DNA Magnesium binding
RNaseHII	21,526	*rnhB*	198	monomer	RNA removal from DNA Magnesium binding
YfjK	83,062	*yfjK*	729	unknown	Putative helicase
MgsA or RarA	49,626	*mgsA/ rarA*	447	tetramer	Flap creator
Endo IX/ Endo Xni	28,166	*ygdG*	251	unknown	Flap endonuclease
SbcD	44,714	*sbcD*	400	SbcCD	Nuclease recombination
Lhr	169,381	*lhr*	1,538	unknown	Putative helicase
YoaA	70,378	*yoaA*	636	unknown	Putative helicase 4Fe-4S cluster
TopoI	97,350	*topA*	865	monomer	Type 1 topoisomerase Magnesium binding
TopoII or DNA gyrase	96,964 89,950	*gyrA* *gyrB*	875 804	heterotetramer	Type 2 topoisomerase Magnesium binding
TopoIII	73,217	*topB*	653	monomer	Type 1 topoisomerase Magnesium binding
TopoIV	83,831 70,244	*parC* *parE*	752 630	heterotetramer	Type 2 topoisomerase *parE* (Magnesium binding)

Table 1.1C. The DNA repair proteome of *Escherichia coli*. Table provided by Professor Michael Cox and was modified to include proteins which coordinate metal ions.

Protein	Function	Citation
DdrA	Binds to 3' ends of fragmented DNA and protects from degradation	Harris et al., 2004
DdrB	Single-strand DNA binding protein	Norais et al., 2009
DdrC	Single- and double-strand DNA binding protein; may protect DNA fragments from degradation; may participate in extended synthesis-dependent strand annealing (ESDSA) for repairing fragmented genomic DNA	Selvam et al., 2013 Ithurbide et al., 2015
DdrD	Implicated in UV and mitomycin C resistance along with DdrA and PprA	Selvam et al., 2013.
PprA	May stimulate non-homologous end joining (NHEJ) as a method of repairing DNA DSBs	Selvam et al., 2013 Earl et al., 2002
IrrE	Novel regulator of RecA expression post-IR exposure	Narumi et al., 2004

Table 2.1. Novel DNA repair proteins in *Deinococcus radiodurans*. Six of the most-studied novel DNA repair proteins of *D. radiodurans*. Of these six, the genes for DdrA, DdrB, DdrC, DdrD, and PprA are the most highly transcribed genes post-irradiation and desiccation in *D. radiodurans* [56, 58, 59, 61, 62, 64].

1.7 Figures

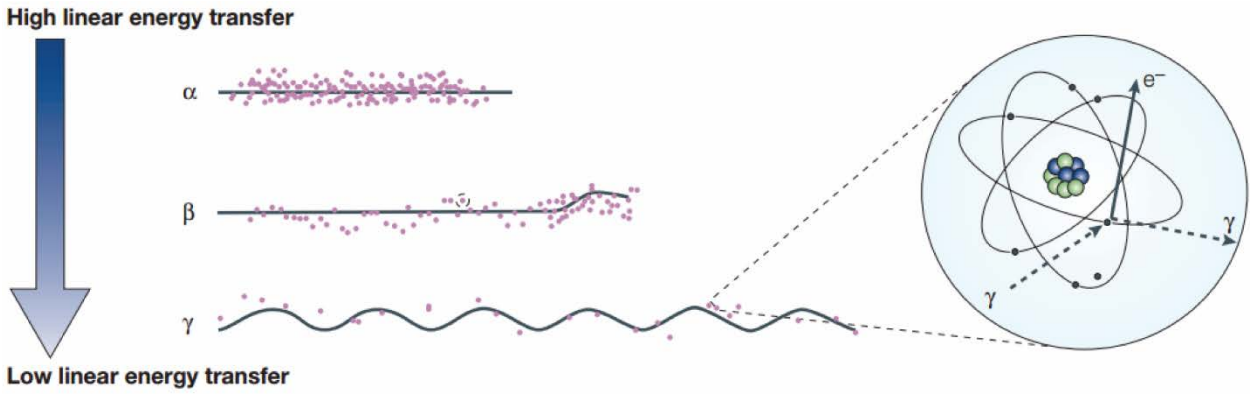


Figure 1.1. Types of ionizing radiation. Ionizing radiation (IR) consists of particulate (including α particles (neutrons) and β particles (electrons)) and electromagnetic (γ rays and X rays) radiation. α particles have the highest linear energy transfer; these particles transfer more energy over a shorter, linear track. Electromagnetic IR can penetrate the farthest into matter and has the lowest linear energy transfer. The Compton effect is also depicted, where γ or X rays can impart energy to an electron (thus generating a β particle) and continue along its track, ionizing additional atoms. Figured modified from Cox and Battista, 2005 [1].

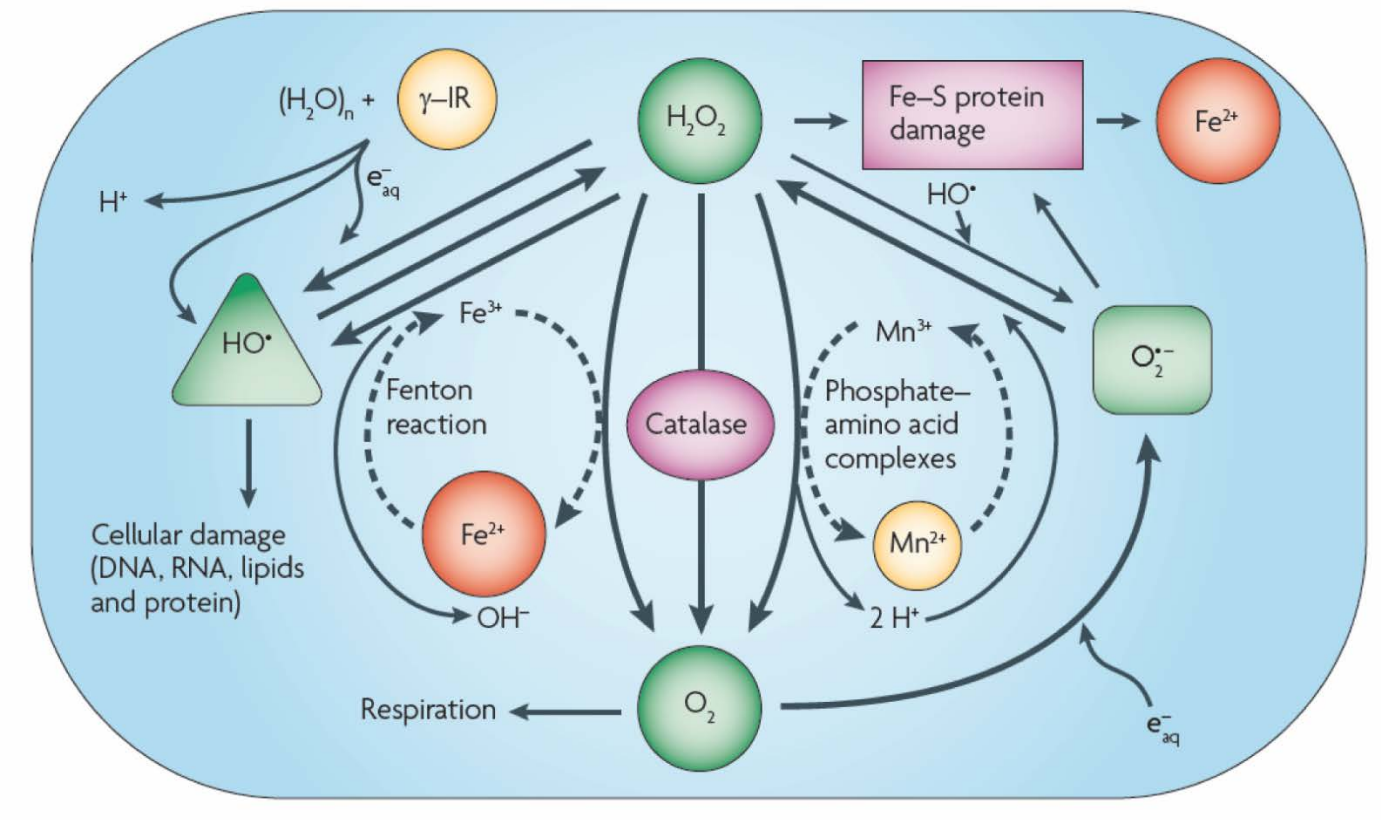


Figure 2.1. Reactive oxygen species produced by exposure to ionizing radiation. Ionizing radiation will hydrolyze intracellular water to ionized water (H_2O^+), hydrogen radicals ($H^\bullet$), hydrated electrons (e^-_{aq}), hydroxyl radicals ($^\bullet OH$), hydrogen peroxide (H_2O_2) and superoxide ($O2^-$). Reactive oxygen can oxidize all intracellular macromolecules with varying discrimination. Figure modified from Daly et al., 2009 [6].

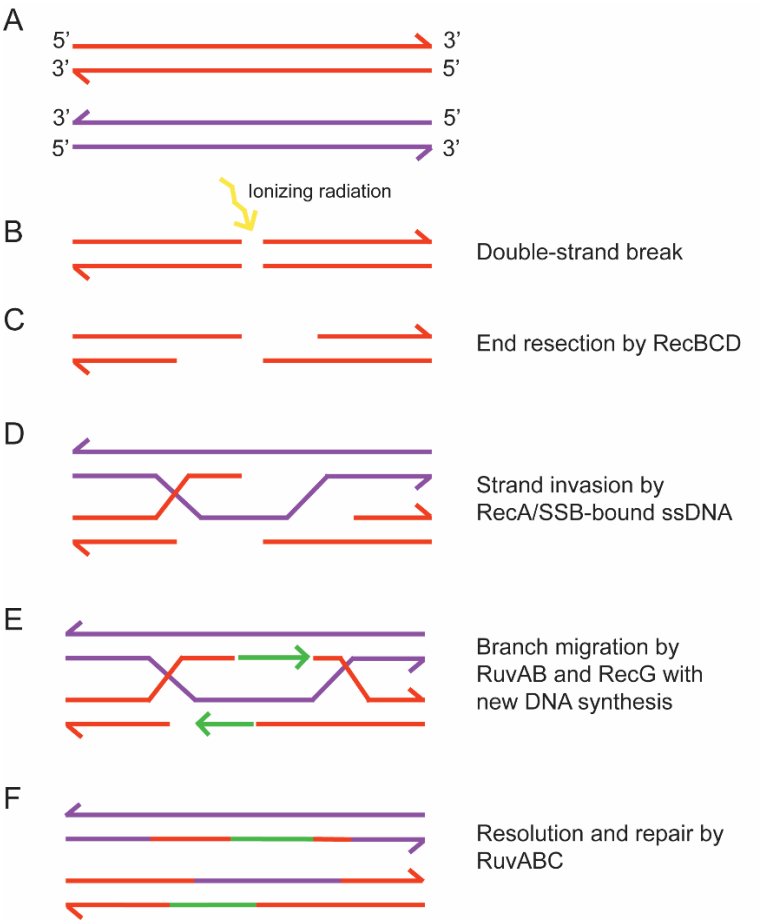


Figure 3.1. Cartoon depiction of RecA-mediated DNA double-strand break repair. A) Two sister chromosomes. B) Ionizing radiation causes a DNA double-strand break. C) The RecBCD heterotrimer binds to the fragmented blunt ends and unwinds DNA 3' to 5', while degrading the 5' strand. DNA degradation slows once RecBCD reaches a chi sequence. D) Single-strand binding protein (SSB) binds to single-strand DNA (ssDNA) and stimulates RecA filamentation. The RecA filament bound to ssDNA invades the sister chromosome and initiates the homology search. E) RuvAB and RecG stimulate branch migration of the Holiday junctions, and new DNA synthesis begins from the free 3' DNA ends. F) RuvABC process the Holiday junctions and the synthesized DNA is ligated to the existing chromosomes. Figure is modified from Masson and West, 2001. [92].

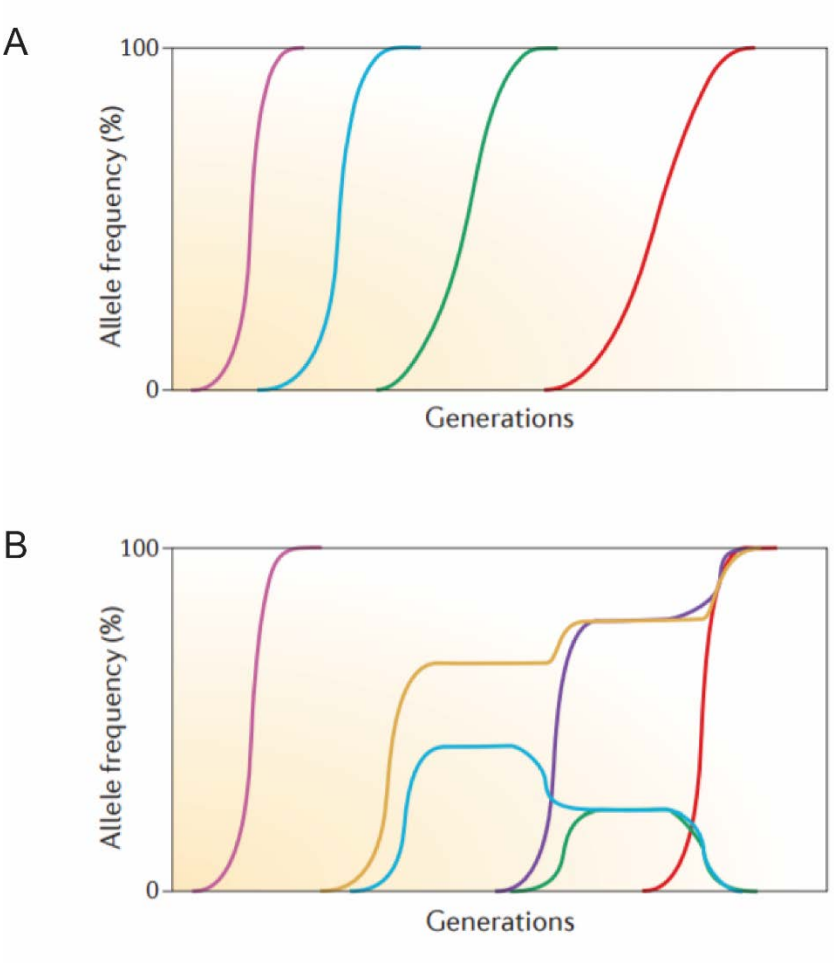


Figure 4.1. Cartoon depiction of allele frequencies in evolving populations. A) A single mutation within an evolving population is depicted by the differently-colored lines. Beneficial mutations increase in frequency in a population until they are fixed in the population (i.e. 100% frequency). Once a mutation fixes, each mutation afterwards appears within that new genetic background. B) Clonal interference impairs the ability of beneficial mutations to fix in a population. Sub-populations containing different beneficial mutations (yellow and blue) increase in frequency until they come into direct competition, at which point neither frequency continues to increase. The subpopulation with the yellow mutation gains new beneficial mutations, so it outcompetes the blue subpopulation and fixes. The blue subpopulation is driven extinct. Figure is modified from Barrick and Lenski, 2013 [69].

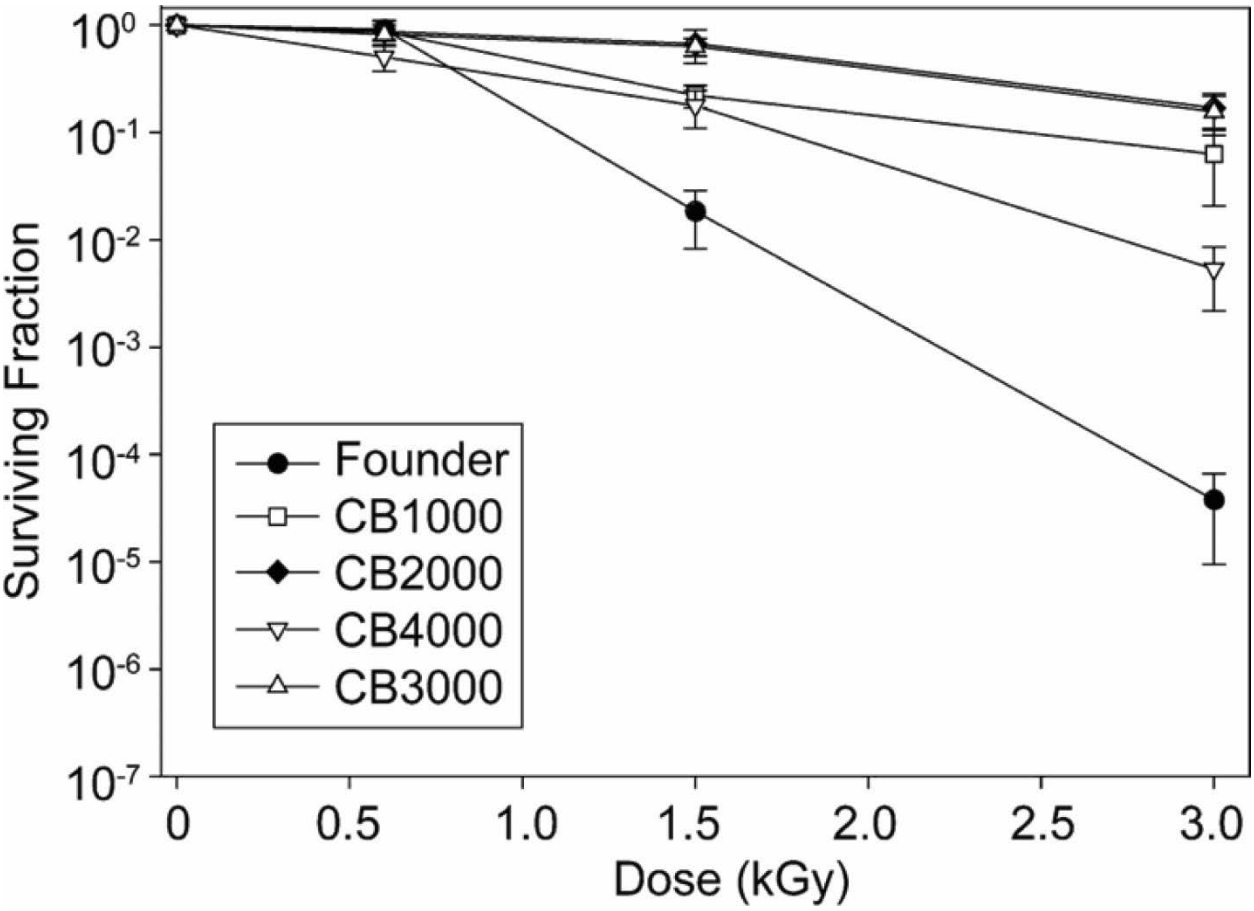


Figure 5.1. Survival curves of evolved *E. coli* isolates exposed to γ-ray irradiation. CB1000, CB2000, CB3000, and CB4000 are individual isolates from four separate populations; each population was started from the Founder strain. The four evolved populations were exposed to 20 iterative rounds of sufficient γ-ray IR to kill 99% of the population at each round of selection. These isolates have 100 to 1000-fold higher percent survival than the Founder strain at a dose of 3000 Gy. Figure modified from Harris, *et al*., 2009 [13].

Chapter II: Mechanisms of experimentally-evolved gamma ray ionizing radiation resistance

2.1 Abstract

Escherichia coli can acquire high levels of resistance to ionizing radiation via directed evolution. In previous work, the IR resistance of one of four evolved populations (IR-2-20) was explained by mutations in three different DNA repair functions: *recA*, *dnaB*, and *yfjK*. Another of the four populations (IR-3-20; characterized here) took a different path to the observed IR resistance. One mutation largely accounted for the phenotype, this time affecting the *fnr* gene (encoding the FNR F186I variant). Two additional mutations affecting the *clpX* and *priA* genes appear to make contributions to IR resistance. In population IR-3-20, regulatory changes affecting anaerobic metabolism and proteolysis is evidence that more pathways other than alterations in DNA repair processes exist as an IR resistance mechanism. The key mutations in IR-2-20 and IR-3-20 exhibited an additive effect on overall IR resistance levels.

2.2 Introduction

Organisms resistant to extreme doses of ionizing radiation (IR) have yielded insights into mechanisms that could modulate radiation effects in cells. The Gram-positive bacterium *Deinococcus radiodurans* has been a prominent subject in the study of radioresistance. Despite being able to survive doses of IR well over 10 kGy (over 1,000-fold higher than the lethal human dose), the IR-resistance phenotype of *D. radiodurans* appears to be a by-product of desiccation resistance [1, 56]. Indeed, IR-resistant species are readily isolated from desert environments [66, 67].

D. radiodurans relies on DNA repair mechanisms, some of which are unique to this bacterium, that can repair hundreds of DNA double-strand breaks (DSBs) without loss of genetic information **[1, 57]**. Additionally, *D. radiodurans* resists substantial protein oxidation due to IR exposure via an extraordinary ability to ameliorate reactive oxygen species (ROS) generated by radiolysis of intracellular H_2O [19, 22]. Although *D. radiodurans* has an extreme capacity for ROS amelioration, this bacterium accumulates DNA double-strand breaks (DSBs) induced by IR at a normal rate [6]. These observations have led to the hypothesis that protecting the proteome from damage, not the genome, is the primary determinant of IR resistance [6, 16, 20, 21]. Indeed, protein-free *Deinococcus* cell lysate protects protein, but not DNA, from doses of IR as high as 50 kGy [22].

Whereas *Deinococcus radiodurans* has provided some insight into mechanisms of extreme radiation resistance [1, 6, 57], we have pursued another strategy. Bacterial species that do not display unusual levels of resistance to ionizing radiation can acquire such resistance by directed evolution [86-89]. However, limits of sequencing technology prevented molecular characterization of evolved IR resistance in previous studies.

With the advent of modern sequencing techniques, we sought to generate IR resistance in the model bacterium *Escherichia coli* and deep-sequence these evolved populations. We previously subjected four separate populations of *E. coli* to 20 cycles of ^{60}Co irradiation (sufficient to kill up to 99.9% of the cells) followed in each cycle by survivor outgrowth. The 20 cycles resulted in large gains in IR resistance in all four populations, designated IR-1-20, IR-2-20, IR-3-20, and IR-4-20 [13, 90].

An isolate from population IR-2-20, CB2000, was previously characterized to identify the genetic alterations underlying the phenotype [90]. The effort focused on mutations that were fixed in the population and affected genes or pathways altered in additional populations. Though seven mutations made at least a minor contribution, three mutations affecting DNA metabolism accounted for the majority of the IR resistant phenotype of CB2000. These were variants of (a) the DNA repair protein RecA (D276N), (b) the replicative helicase DnaB (P80H), and (c) the putative helicase YfjK (A151D) [90]. Reliance of CB2000 on three variant DNA metabolism proteins for IR resistance implicated a role for enhanced DNA repair in IR resistance. Indeed, biochemical characterization of the RecA D276N variant revealed novel activities consistent with repair of genomes fragmented by IR exposure [91].

Sequencing of isolates from each population revealed a striking trend: RecA, DnaB, and YfjK variants appeared in IR-1-20, IR-2-20, and IR-4-20 [13, 90]. These sequencing results suggested that, in addition to IR-2-20, both IR-1-20 and IR-4-20 rely at least in part on enhanced DNA repair for IR resistance. However, of seven isolates sequenced from population IR-3-20, none contained mutations in *recA*, *dnaB*, or *yfjK*. Protein variants common to IR-3-20 which appear in the same protein or pathway in at least one other population include those involved in DNA replication restart (PriA V554I), amelioration of reactive oxygen species (ROS) (GsiB

L289P and RsxD V45A), cell wall metabolism (NanT F406S), and anaerobic metabolism (FNR F186I) [13, 90]. IR-3-20 also contains a unique intergenic SNP between the *clpP* and *clpX* genes (*clpP*/*clpX* int). Variants of the ClpP and ClpX proteins were also identified in population IR-1-20 [13, 90]. Of these pathways, only protein variants affecting ROS amelioration and cell metabolism have been previously associated with contributions to IR resistance [90].

We now present the characterization of an isolate from population IR-3-20, CB3000. IR resistance in CB3000 is a highly variable phenotype compared to that observed in CB2000. This variability likely stems from the different IR resistance mechanisms of CB3000. Primarily, IR resistance in CB3000 is founded in altered global regulation of anaerobic respiration (FNR F186I). Altered replication restart (PriA V554I) and proteolysis (*clpP*/*clpX* int) may also play a role. The precise reason for the variable IR resistant phenotype of CB3000 could not be determined, although the FNR F186I mutation contributes to IR resistance with regularity.

2.3 Results

2.3.1 Deep sequencing of the IR resistant populations reveals trends in molecular adaptation to high dose ionizing radiation.

The evolved IR-resistant populations IR-1-20, IR-2-20, IR-3-20 and IR-4-20 were previously characterized by sequencing individual isolates from each [13, 90]. In a new effort, all four populations were subjected to deep sequencing using Illumina MiSeq technology. This instrument was employed to achieve approximately 1000X coverage of the *E. coli* genome, allowing for reliable detection of genomic variants present in at least 2% of each population (Table 1.2). Patterns of mutations previously detected in each population were confirmed (Table 2.2) [13, 90].

As in the earlier study [90], we examined the genomic sequencing data for patterns in which the same gene or process was altered in multiple evolved populations. Mutations in DNA metabolism genes, including *recA*, *yfjK*, and genes encoding proteins involved in replication restart were prominent in three of the populations, including the poorly-characterized (until now) population IR-4-20. Much like IR-2-20, fixed (>95% frequency) mutations in the DNA metabolism genes encoding RecA, DnaB, and YfjK appeared independently in this population. IR-4-20 shares the same RecA variant (D276N) as IR-2-20. The DnaB and YfjK variants of IR-4-20 (P200Q and H652Y, respectively) are unique and remain to be tested for a contribution to IR resistance.

Population IR-1-20 has no fixed mutations due to clonal interference. We have previously sequenced seven isolates from IR-1-20 [90]**.** Therefore, we can identify three unique major subpopulations (termed population IR-1-20 A, B, and C) based on linked mutations in the

sequenced isolates. Each of these subpopulations prominently feature mutations in DNA repair pathways. IR-1-20 A (approximately 55% of the population) contains gene alterations encoding RecA (A2S) and DnaT (R145C) variants. Population IR-1-20 B (about 25% of the population) also contains mutations that produce a RecA variant (D276A), as well as a YfjK (K294E) variant. The third sub-population does not include a *recA* allele but does include a mutation in the gene encoding PriC (L162P). Additionally, low frequency mutations in the genes encoding DNA metabolism proteins DnaB (P80H and N398K), PriA (G279D), and YfjK (I172L) appear in IR-1-20 but cannot be linked to one of the three major subpopulations as they were not detected in the previously sequenced isolates [13, 90].

The mutations in RecA and variants of replication restart proteins (DnaT, DnaB, PriA, PriC; assuming the loading target protein DnaB is included in the category) appear often in each of the evolved populations (Table 2.2) [90]. Such prevalence presents a strong indication for a role for enhanced RecA activity [13, 90], and a previously undocumented role for modified replication restart in IR resistance.

There were a total of 46 mutations that appeared to be fixed (present in all, or nearly all, of the cells) in IR-3-20. As in the previous work, we focused on mutations that were not only prominent in the population (fixed or nearly so), but which occurred in genes or pathways that were also significant mutation targets in other populations. We followed these criteria to focus on mutations that are likely linked to radioresistance in multiple populations. Based on these criteria, six mutations attracted our interest. These were the *priA* mutation, a mutant *fnr* allele (encoding the alteration F186I), an intergenic SNP between the *clpP* and *clpX* genes (referred to hereafter as *clpP*/*clpX* int), a GsiB variant (L289P), a RsxD variant (V45A), and a NanT variant (F406S) proteins.

The mutations that are fixed in IR-3-20 (the primary subject of this report) are largely outside the pattern set by the remaining three populations. Aside from one fixed PriA allele (V554I), none of the 701 high-confidence variants detected in the population (>1% frequency and detected on both DNA strands) in population IR-3-20 are in the commonly mutated DNA metabolism genes or pathways seen in the other populations (such as *recA* and *dnaB*).

None of the other three evolved populations considered in Table 2.2 contain a mutation in the gene encoding FNR, a transcription factor that controls the transition between aerobic and anaerobic growth [93] [94] [95]. However, a pilot study in which a single isolate of IR-1-20 (CB1000) was taken through 30 more rounds of directed evolution using ^{60}Co radiation did produce a prominent FNR variant (a nonsense mutation at codon 156, or M156*) [13, 90], suggesting a potential role for FNR variants in IR resistance.

The ClpX and ClpP proteins form the multisubunit ClpXP protease, a AAA+ enzyme responsible for cleavage and elimination of an array of target proteins, including many involved in DNA metabolism [96]. The *clpP*/*clpX* int mutation is a guanine to adenine SNP thirteen nucleotides upstream of the *clpX* 5'-ATG-3' start codon, within the 125 bp intergenic region between *clpP* and *clpX*. While no other intergenic mutations between *clpP* and *clpX* were detected in the four evolved populations, mutations within the *clpP* and *clpX* genes are detectable in IR-1-20. The mutational pattern suggests that modification of ClpXP levels or activity may be beneficial for IR resistance.

The remaining three fixed protein variants are GsiB L289P, RsxD V45A, and NanT F406S. The *gsiB* gene encodes one of four types of subunits for an ABC transporter for glutathione [97], a prominent cellular antioxidant. RsxD is a membrane-spanning protein in the membrane-associated RsxABCDGE protein complex, which is involved in regulation of the SoxRS two-

component signaling pathway [98]. The *nanT* gene encodes a membrane-bound sialic acid/H+ symporter [99, 100]. Mutations in these genes/operons have been previously explored for contributions to IR resistance in IR-2-20 (GsiB N105S, RsxB K122E, and NanE A129T) [90].

2.3.2 The ionizing radiation resistant phenotype of CB3000 is highly variable

We have assayed CB3000 for IR resistance from 2014 to 2017. Over this time-period, the percent survival of CB3000 to 3000 Gy of IR has ranged from the level of Founder Δe14 (a derivative of the Founder ancestor strain which lacks the e14 prophage that was lost early in the directed evolution experiment [13]) to that of CB2000 (Figure 1.2). There is no apparent explanation for this highly variable phenotype. It is unlikely that CB3000 is no more IR resistant than Founder Δe14. Like CB2000, CB3000 is an isolate from a population that had undergone 20 cycles of selection with sufficient IR to kill 99% of the population at each round of selection. The isolates from these populations were highly IR resistant compared to the Founder strain [13]. It is likely, however, that the IR-resistant phenotype of CB3000 is highly dependent on irradiation conditions. Unfortunately, the available methods for assaying CB3000 are not well-controlled. The duration necessary to achieve a dose of 3000 Gy, the dose at which we assay mutations from contribution to IR resistance, has changed with the decay of the ^{137}Cs source (from ~7.75 hr and a dose rate of 6.6 Gy/min in 2014, to ~ 8.5 hr and a dose rate of 6.1 Gy/min in 2017). In addition to changing irradiation conditions, the irradiator is not temperature controlled, leading to changes in room temperature with the use of air conditioning and heating where the irradiator is housed, depending on the season. Finally, cells are suspended in an undefined rich medium (LB) for the duration of these irradiations, which likely allows for growth during the assay.

Despite these variable irradiation conditions, IR resistance of Founder Δe14 and CB2000 has not changed IR-resistance from 2014 to 2017. Therefore, the unstable phenotype of CB3000 is likely a biological problem, where the mutations which increase IR resistance of this isolate are affected by the variable irradiation conditions noted above. As such, despite being able to accurately determine the mutations which contribute to IR resistance in CB2000 (Byrne et al, 2014) [90], it is out of our ability to reproducibly determine which mutations contribute to IR resistance in CB3000. Additionally, new federal policies have incentivized the decommissioning of ^{137}Cs and ^{60}Co irradiators, which renders us unable to repeat previous experiments and further explore what factors affect the IR resistant phenotype of CB3000.

Though we have been unable to determine the extent of IR resistance of CB3000, one mutation from CB3000, FNR F186I, has reproducibly increased IR resistance of Founder Δe14 from 2014 to 2017 (Figure 1.2). In addition, over a two-month period in 2016 when temperature conditions where the ^{137}Cs irradiator is housed were likely stable, two mutations (the PriA V554I variant and in intergenic mutation between the *clpP* and *clpX* genes) appeared to epistatically contribute to IR resistance in CB3000. The data presented here on which mutations contribute to IR resistance in CB3000 will focus on this time-period in 2016. However, we recognize that any assertion of contribution by the PriA V554I and *clpP*/*clpX* int mutations is speculative, and that only the FNR F186I mutation from CB3000 can reliably enhance IR resistance of *E. coli*.

2.3.3 Altered regulatory pathways enhance IR-resistance in IR-3-20

Six fixed mutations from the IR-3-20 isolate CB3000 were tested for their contribution to IR resistance: *clpP*/*clpX* int, plus mutations in genes encoding: FNR F186I, PriA V554I, GsiB

L289P, RsxD V45A, and NanT F406A. As described previously [90], wild type genes were replaced by mutations inserted singly and in combination into the Founder Δe14 background. In addition, the mutated sequences of these loci were reverted, singly and in combination, to the wild-type sequence in CB3000, maintaining the otherwise evolved mutant background (Figure 2.2A, B).

As shown in Figure 2.2, the radioresistant phenotype of CB3000 was accounted for entirely by the mutations in the *fnr* and *priA* genes, along with the intergenic *clpP*/*clpX* int mutation (Figure 2.2B). Three variants tested had no effect on IR resistance as measured in these trials: GsiB L289P, RsxD V45A and NanT F406S (Supplemental Figure 1.2). All these mutants were testing over a two-month period in 2016, to eliminate variability of the CB3000 IR resistant phenotype that may be affected by ambient temperature at the location of the irradiator.

Surprisingly, the *clpP*/*clpX* int and PriA V554I mutations appear to work epistatically to enhance IR resistance. Alone, neither mutation contributes to IR resistance when placed into the otherwise wild-type Founder Δe14 background. Consistently, converting any of these single mutations to the wild-type sequence in CB3000 results in a decrease in IR resistance**.** FNR F186I has no epistatic relationship with other mutations and alone confers the greatest contribution to IR resistance. Taken together, these mutations indicate that IR resistance in CB3000 is largely explained by alterations in particular patterns of global regulation, a distinct path from the enzymatic changes previously noted in CB2000 [90]. Both altered proteomic regulation (*clpP*/*clpX* int) that acts with PriA-mediated replication restart (PriA V554I) and altered transcriptional regulation affecting anaerobic metabolism (FNR F186I) appear to be important.

2.3.4 The prominent mutations affecting IR-resistance in CB2000 and CB3000 act additively

As the isolates CB2000 and CB3000 appear to have developed IR resistance through two separate pathways, we investigated how these two pathways interact. We combined the prominent mutations contributing to IR-resistance from CB2000 (RecA D276N, DnaB P80H, and YfjK A151D) with those from CB3000 (*clpP*/*clpX* int, PriA V554I and FNR F186I) in the Founder Δe14 background. We found that these sets of mutations act additively to increase IR resistance (Figure 2.2C). This suggests that these two sets of mutations affect complementary pathways in the cell.

2.3.5 CB3000 is not resistant to non-IR DNA damaging agents

To gain a deeper understanding of how these evolved populations survive IR, we tested resistance of CB2000 and CB3000 to other types of DNA damaging agents related to the damage caused by IR. DNA double-strand breaks (DSBs) are a prominent, and lethal, form of damage caused by IR. The antibiotic ciprofloxacin was used as a proxy for the DSBs caused by IR to assay the DSB repair capability of our evolved populations. CB2000, whose IR resistance stems from DNA metabolism protein variants, was approximately two logs more resistant to ciprofloxacin at a concentration of 0.01 µg/ml compared to Founder Δe14 (Figure 3.2A). Interestingly, CB3000 was more sensitive to the antibiotic than Founder Δe14, further suggesting that the mechanism of IR resistance in CB3000 was distinct from that seen in CB2000.

CB3000 has been previously reported to exhibit enhanced UV resistance compared to the Founder strain [13]. However, UV resistance in this strain appears to be solely founded in the loss of the e14 prophage, a deletion, which appears in all the evolved populations. CB3000 only exhibits increased UV resistance compared to Founder, but not Founder Δe14 (Figure 3.2B). In

contrast, CB2000 is more UV resistant than both CB3000 and Founder Δe14, which is consistent with an enhanced DNA repair capability.

2.3.6 CB3000 has enhanced ROS amelioration capabilities.

Most damage caused by IR comes from the reactive oxygen species (ROS) formed from the interaction of IR with intracellular water molecules [4, 16]. Therefore, ROS amelioration could provide a path to surviving high levels of irradiation. Since we have seen that CB3000 is likely not radioresistant due to enhanced global DNA repair, we sought to determine the extent of ROS amelioration capability in CB3000. As a general representative of ROS, we challenged our evolved strains with H_2O_2. In direct contrast to the results obtained with ciprofloxacin, CB3000 was more resistant to 10 mM H_2O_2 compared to Founder Δe14. In contrast, CB2000 exhibited greater sensitivity to H_2O_2 (Figure 4.2).

We then sought to determine whether the H_2O_2 resistance of CB3000 stems from an enhanced capacity to ameliorate ROS levels or repair damage caused by intracellular ROS. The dye H_2DCFDA can be used to quantify intracellular ROS, as it must be modified by intracellular esterases and then oxidized to form a detectable fluorescent compound [101]. We therefore mixed H_2DCFDA with Founder Δe14, CB2000, and CB3000 after exposure to 10 mM H_2O_2 and detected resulting fluorescence to infer ROS levels. While fluorescence of the compound mixed with Founder Δe14 and CB2000 is similar, significantly less fluorescence is detected in CB3000, suggesting an enhanced capacity to ameliorate levels of intracellular ROS (Figure 5.2).

As measured by the H_2DCFDA assay, the enhanced ROS amelioration capacity of CB3000 does not reside in the three key mutations that account for the IR resistance phenotype. Founder

Δe14 carrying the three mutations from CB3000, which increase IR resistance (*clpP/clpX* int, PriA V554I, and FNR F186I), does not mimic the ROS amelioration capacity of CB3000 (Figure 5.2). The overall result indicates that CB3000 has enhanced ROS amelioration capacity, and that this capacity is not shared by CB2000. However, that capacity likely reflects other mutations in CB3000 that are not studied here, and it remains unclear as to what extent ROS amelioration contributes to the measured IR resistance phenotype in this strain.

2.3.7 FNR F186I is a hyperactive FNR variant

We sought to better understand IR resistance in CB3000 by further characterizing the mutations, which confer radioresistance in this isolate. FNR is a Fe-S binding transcription factor that regulates genes related to anaerobic metabolism [102, 103]. It has been previously reported that the F186 residue of FNR contacts the alpha C-terminal domain of RNA polymerase and that variants of residue F186 decrease activation of Class I and Class II FNR-dependent promoters [104, 105]. Therefore, we sought to determine the effects, if any, of the F186I variant on FNR-controlled transcription. We utilized a β-galactosidase fusion of the classical FNR-controlled *narG* promoter (P_{narG}) to *lacZ* [106] to characterize the FNR F186I variant activity. We observed no difference in promoter activity in Founder Δe14 with the FNR F186I variant compared to Founder Δe14 with the wild-type FNR F186 allele at early exponential (OD_{600} of 0.2) and stationary phase (overnight cultures grown for 15 to 18 hours) (Figure 6.2).

The conditions in which we irradiate strains (approximately an 8 hour duration for a dose of 3000 Gy, in sealed 1.5 mL tubes containing 1 mL of exponential phase culture) likely create a microaerophilic environment over time. A microaerophilic environment should increase the

amount of active FNR dimers, which are destroyed by O_2 [107] [108]. To determine the effect of these conditions on FNR activity, we performed a mock-irradiation, allowing culture to incubate at room temperature in microfuge tubes as they do during radiation assays. In mock-irradiation conditions, Founder Δe14 with the FNR F186I variant shows increased transcription from the P_{narG} – *lacZ* fusion compared to Founder Δe14 with the wild-type FNR (Figure 6.2). These results indicate that the FNR F186I variant allows for enhanced FNR activity under the conditions of our irradiation trials.

2.3.8 The *clpP/clpX* int mutation reduces levels of the ClpX protein

The *clpP/clpX* int mutation is a SNP thirteen nucleotides upstream of the *clpX* 'ATG' start codon, within the 125 bp intergenic region between the *clpP* and *clpX* genes. The *clpP* and *clpX* genes share a promoter, proximal to *clpP* [109, 110]. Therefore, it is unlikely that the *clpP*/*clpX* int mutation affects transcription of the *clpX* gene. Based on sequence similarity to the consensus *E. coli* ribosome binding site (RBS) (5'-AGGAGGT-3'), this SNP may be in the *clpX* mRNA RBS (Figure 7.2A). This SNP reduces the similarity of this RBS to the consensus sequence. Therefore, we hypothesized that the *clpP*/*clpX* int mutation reduces translation of the ClpX protein. To test this hypothesis, we determined the amount of ClpX in Founder Δe14 and CB3000 via Western blot (Figure 7.2B). We observed that CB3000 produces less ClpX protein, and this reduction of ClpX is abolished when the *clpP*/*clpX* int mutation is converted to the wild-type sequence in CB3000 (Figure 7.2C).

2.3.9 PriA V554I has reduced ATPase and helicase activity

PriA is a 3' to 5' DNA helicase which assists in replisome reassembly at the replication fork [111-114]. The V554 residue of PriA is in the helicase lobe II domain, C-terminal of the helicase motif V. In the published PriA structure[115] (*K. pneumoniae* PriA; 88 % identity with *E. coli* PriA), V554 appeared to pack with helicase motif VI (3.5 Å between V554 side chain carbon and a side chain carbon of V580 and 4.0 Å between the other V554 side chain carbon and the A584 side chain carbon within the motif VI QVSGRAGR) (Figure 8.2A). With the ATP pocket on the opposite side of motif VI, we hypothesized that this mutation would alter the ATP hydrolysis mechanism and the helicase activity of PriA. It has been previously reported that PriA variants, including a variant at residue L557, rescue the UV-sensitivity of Δ*recG* mutants [116] and that this rescue is due to decreased helicase activity [117]. We deleted *recG* from Founder Δe14 and CB3000 derived strains to determine the effect of PriA V554I on UV sensitivity in these backgrounds. PriA V554I partially rescues the UV sensitivity of Founder Δe14 Δ*recG* at a dose of 60 J/m^2. The effect is not as great as that seen with some other PriA variants [116], but it suggests that this variant may have decreased, but not abolished, helicase activity (Figure 9.2).

We purified PriA V554I to characterize the effect of this mutation *in vitro*. We first tested for retention of PriA DNA fork-binding using a fluorescence anisotropy assay. The V554I PriA K_D for a fluorescein labeled DNA fork (two partial complementary oligonucleotides) was equal to that of WT PriA (Figure 8.2B). We next analyzed the PriA DNA-dependent ATPase activity of this variant. PriA and ATP concentrations were held constant and poly dT was titrated into the reaction (Figure 8.2C). The k_{max} of V554I was 0.14 that of WT PriA (~7-fold reduction). This V554I ATPase activity was in a detectable range, ~12-fold above the ATPase-dead, Walker A mutant K230R. The K_{DNA} of V554I was near that of WT PriA. To analyze the effect of this ATPase

defect on PriA helicase activity, we next analyzed V554I in a variety of *in vitro* helicase assays. On the simple DNA fork, V554I helicase activity was significantly reduced from WT PriA, near background and the level of K230R PriA (Figure 8.2D). As a previously characterized PriA point mutation near V554, *recG* suppressor "srgA1" (L557P), was shown to be deficient in unwinding a fork of this type but to retain helicase activity when both nascent strands were present [117], we also tested V554I on a four-oligonucleotide fork. V554I helicase activity was similarly reduced on this fork type, near background levels with K230R PriA. To test for any detectable retention of PriA helicase activity, we tested PriA helicase activity under SSB and PriB-mediated stimulation of PriA helicase activity [118-120]. In the presence of SSB and PriB, V554I PriA helicase activity was detectable above background. In the presence of PriB, specifically, V554I helicase activity was near the WT PriA level, in conditions were WT PriA helicase activity appeared to be above a detectable range.

2.3.10 IR-resistant isolates CB2000 and CB3000 exhibit increased fitness in rich media

We have previously assessed fitness of evolved isolates in a direct competition assay [121] relative to Founder Δe14 during irradiation. The results of this assay confirmed that the evolved isolate CB2000 is more fit, i.e., it grows faster, than Founder Δe14 post-irradiation [90]. We now demonstrate that this trend continues without selection. In mixed cultures (Founder Δe14 versus CB2000; Founder Δe14 versus CB3000) grown for 72 hours, both CB2000 and CB3000 almost completely outcompete Founder Δe14 (Figure 10.2 A,B). To determine the basis of this phenotype, we competed Founder Δe14 with each of the IR resistance increasing mutations (FNR F186I, *clpP*/*clpX* int, and PriA V554I) against the wild-type Founder Δe14. We additionally calculated fitness statistics for each mutation to determine quantitative contributions, if any, to the growth

phenotype [121]. Despite an enhancement of growth from the F186I mutation, these three mutations did not completely recapitulate the growth phenotype of CB3000. We hypothesized that this growth phenotype was likely selected for in the outgrowth step of our directed evolution protocol. We therefore applied the criteria we used to determine candidate mutations, which increase IR resistance to find mutations that enhance growth. We assayed the GsiB L289P mutation from CB3000, which does not increase IR resistance (Figure 10.2C), for a contribution to enhanced growth. We found that the GsiB L289P variant enhances growth of Founder Δe14 almost to the level of CB3000. Adding the FNR F186I mutation, which enhances fitness of Founder Δe14 without selection, to Founder Δe14 with GsiB L289P completely recapitulates the enhanced growth phenotype of CB3000. Consistently, converting the *gsiB* and *fnr* genes to their wild-type sequence in CB3000 abolishes the enhanced growth phenotype (Figure 10.2D). In addition, the GsiB N104S mutation from CB2000 also enhances growth of Founder Δe14 (Figure 10.2E). Enhanced growth from two different GsiB variants indicates that despite differing mechanisms of IR resistance, CB2000 and CB3000 have evolved similar mechanisms of enhanced growth.

2.4 Discussion

Making use of four *E. coli* populations that have acquired enhanced resistance to ionizing radiation via directed evolution, we have further explored the molecular basis of the IR resistance phenotype. IR resistance of the CB3000 isolate derived from population IR-3-20 is at least partially due to the FNR F186I variant. Despite FNR F186I reproducibly increasing IR resistance of the Founder Δe14 background strain (Figure 1.2), the IR resistant phenotype of CB3000 is highly variable over time. However, over a two-month time-period the *clpP*/*clpX* int and PriA V554I mutations increased IR resistance in conjunction with FNR F186I, and these three mutations appeared to contribute to the entirety of the IR resistant phenotype of CB3000. Thus, IR resistance in CB3000 has come about via mechanisms distinct from those most prominent in CB2000. In addition to the distinct genetic alterations, CB3000 exhibits greater resistance to H_2O_2 than either Founder Δe14 cells or CB2000, whereas CB2000 exhibits greater resistance to the DNA damaging agents ciprofloxacin and UV irradiation. A similar pattern was seen in an earlier study in which CB2000 exhibited enhanced resistance to treatment with mitomycin C while CB3000 did not [13]. Second, the most prominent mutations conferring IR resistance in CB2000 and CB3000 are additive in their effects, again suggesting they affect complementary pathways. Third, some of the mutations arising in these populations, most notably mutations in the gene *gsiB*, enhance cellular growth rate and likely arose as a result of the outgrowth phase incorporated into our selection protocol.

We have only begun to peel back the layers of pathways which contribute to experimentally evolved resistance to extreme levels of IR. We have previously evolved four populations of highly IR-resistant *E. coli*, termed IR-1-20, IR-2-20, IR-3-20, and IR-4-20 [13, 90]. Three mutations in proteins involved in DNA metabolism were found to comprise most of the IR-resistant phenotype

of CB2000: variants of the RecA protein (D276N), DnaB (P80H), and YfjK (A151D) [13, 90]. The results provided evidence that enhancements in DNA metabolism can contribute to the IR resistance phenotype. The population IR-3-20 starkly contrasts with this trend. Other than the PriA variant, the three mutations which appear to increase IR resistance in CB3000 affect aspects of global regulation rather than DNA metabolism. The alteration in FNR lead to likely increases in transcription of genes in the FNR regulon under the microaerophilic conditions likely to be created during extended irradiation trials, as seen here with transcription from the *narG* gene promoter (Figure 6.2). It is not clear what parts of the regulon may be responsible for the observed IR resistance. The *clpP*/*clpX* int mutation decreases expression of the *clpX* protein, likely increasing steady state levels of all proteins degraded by ClpP/ClpX.

Next-generation sequencing of each population confirmed trends seen previously by whole-genome sequencing of isolates from each population. Populations IR-2-20, IR-3-20, and IR-4-20 are largely clonal, with 22, 31, and 21 mutations fixed in each population, respectively. IR-2-20 and IR-4-20 have a similar pattern of mutations, with fixed variants of RecA (D276N in both), DnaB (P80H and P199Q) and YfjK (A151D and H652Y). Given that the IR resistance phenotype of the IR-2-20 isolate CB2000 was previously shown to be largely dependent on variants of these three DNA metabolism proteins, it is likely that much of the IR resistance of IR-4-20 is founded in the variants of these same proteins.

IR-1-20 has at least three sub-populations. The three most prominent sub-populations, named IR-1-20 A, B, and C, make up approximately 60%, 15%, and 20% of IR-1-20, respectively. Considering each sub-population separately, IR-1-20 B appears similar to IR-2-20 and IR-4-20, with mutations in RecA (D276A), DnaB (L74S), and YfjK (K294E). Of these three DNA metabolism proteins, IR-1-20 A only contains a RecA variant (A290S) which was previously

shown to increase IR resistance in a wild-type background [13]. IR-1-20 A only contains two other mutations that are common to the other three main populations: one affecting the DNA replication restart protein DnaT (R145C) and the other a variant of the ClpX protein (Y384C), which is part of the ClpXP protease. Sub-population IR-1-20 C continues the trend of a replication restart variant (PriC L162P) alongside a variant affecting the ClpXP protease (ClpP Y76C). The final two mutations in IR-1-20 C which appear in commonly mutated genes or pathways are two synonymous mutations in the *dnaJ* and *wcaJ* genes.

IR-3-20 continues a pattern observed in IR-1-20 subpopulations A and C, where a replication restart variant (PriA V554I) appears alongside a mutation affecting the ClpXP protease (in this case an intergenic mutation between the *clpP* and *clpX* genes). While mutations affecting cell wall metabolism and ROS amelioration appear in all four of the evolved populations, the FNR variant of IR-3-20 is unique. However, in a population started from the IR-1-20 isolate CB1000 which was exposed to an extra 30 iterative rounds of IR, a truncated FNR variant was observed (M157*) [13, 90].

Taken together, our results provide evidence for at least three distinct pathways that can contribute to IR resistance. One, previously explored [13, 90], involves enhancements in the cellular DNA repair systems. Here, we present two new pathways: altered anaerobic metabolism through the FNR transcription factor, and an interaction between replication restart machinery (PriA, PriC, and DnaT) and the ClpXP protease (or an altered ClpXP-dependent proteome).

We previously reported that CB2000 and CB3000 are both more resistant to UV compared to Founder [13]. We now present evidence that the enhanced resistance of CB3000 to UV is due entirely to the loss of the e14 prophage. CB3000 is more UV resistant than MG1655, but not Founder Δe14. However, CB2000 exhibits a higher UV resistance compared to Founder Δe14.

ROS formed during irradiation can damage cellular macromolecules indiscriminately [2, 4, 16]. Therefore, enhanced ROS amelioration could provide a pathway for IR resistance [6, 19, 22]. CB3000 was moderately more resistant to H_2O_2 exposure compared to Founder Δe14, although this enhanced resistance to H_2O_2 is not associated with the three mutations we focused on in the current study. Using the dye H_2DCFDA, which can be used to detect intracellular ROS, we found that after H_2O_2 exposure, CB3000 had reduced levels of intracellular ROS compared to Founder Δe14 and CB2000. Notably, CB2000 was nearly two logs more sensitive to 10mM H_2O_2 than Founder Δe14, in spite of its enhanced capacity to survive exposure to IR. The reason for this sensitivity remains unclear. Thus, significant enhancement of IR resistance can be achieved via directed evolution without an enhanced capacity to ameliorate ROS.

The FNR F186I mutation is the only mutation from CB3000 which consistently and reproducibly enhances IR resistance. The F186 residue of FNR is known to contact the alpha-subunit of RNA polymerase [104], making it likely that FNR F186I alters transcription of FNR-controlled genes. FNR is only active in a dimeric form created by coordination of an oxygen-labile Fe-S cluster between FNR monomers [94, 107, 108]. The microaerophilic environment of the mock irradiation assay may allow for increased levels of active, dimeric FNR protein. The downstream effects of the FNR F186I variant are unknown. Microarray analysis has implicated FNR as a regulator of 103 operons related to anaerobiosis, activating 68 and repressing 35 [93]. Of these FNR-controlled genes, *sodA* (which is repressed by FNR and encodes the Mn-binding superoxide dismutase) and *katG* (which is activated by FNR and encodes a bifunctional catalase/peroxidase) could (in principle) contribute to enhanced IR resistance through ROS amelioration activity. The FNR F186I variant may also reduce endogenously produced ROS. Transfer of electrons from NADH and $FADH_2$ to oxygen as a final electron acceptor in the electron

transport chain is a major source of ROS [2]. FNR activates transcription of nitrate reductase A (*narGHJI* operon) and fumarate reductase (*frdABCD*), which allow for use of nitrate and fumarate as alternative electron acceptors. Enhanced production of nitrate reductase A and fumarate reductase may thus allow CB3000 to reduce the levels of ROS produced by using O_2 as a terminal electron acceptor. Reducing endogenous ROS as means of IR resistance may parallel mechanisms used by *D. radiodurans*. It has previously been observed that in *D. radiodurans*, exposure to IR will alter gene expression to reduce metabolic ROS [122].

The protein products of *clpP* and *clpX* form the ClpXP protease, which is responsible for degrading approximately 96 proteins in *E. coli*, many of which are related to stress responses to DNA damage and ROS [96]. The SNP appears to be in the ribosome binding site of *clpX* which will affect the levels of ClpX, and therefore ClpXP. The *clpP*/*clpX* int mutation reduces the levels of ClpX protein approximately two-fold compared to wild-type. A number of DNA repair proteins are ClpXP substrates [96]. These include proteins involved in DNA double-strand break repair (RecN) and nucleotide excision repair (UvrA). However, no previously characterized ClpXP substrates are involved in DNA replication restart, leaving the relationship between the *clpP*/*clpX* int mutation and the PriA V554I variant unknown.

PriA is a 3′ to 5′ helicase which assists in DNA replication restart. The V554I variant partially rescues UV sensitivity of a Δ*recG* mutant compared to a Δ*recG* mutant with the wild-type PriA protein, suggesting at least a partial loss of helicase activity [116, 117]. Interestingly, this mutation is three residues before the previously characterized "*srgA1*" mutant, PriA L557P, that was found as a suppressor of the UV sensitivity of recG mutants [116]. This underlies the importance of this PriA region and altering the PriA ATPase/helicase activity in DNA-damage pathways. The PriA V554 residue is highly conserved (analysis of unique PriA sequences from

300 bacterial species) that packs against helicase motif VI in the *K. pneumoniae* PriA structure [115]. Our *in vitro* characterization of V554I PriA appeared to be consistent with the hypothesis that this mutation had reduced helicase activity due to reduced ATP hydrolysis rates. The PriA V554I variant helicase activity was only detectable above background upon SSB and PriB stimulation, with robust helicase activity in the presence of PriB. Interestingly, the PriA protein of *D. radiodurans* lacks helicase activity [123], though it is not clear if this helicase-deficiency is involved in the IR resistant phenotype of this bacterium.

Many mutations, such as those in *gsiB* or the *nan* operon, appear in multiple populations at high frequencies yet do not appear to contribute significantly to IR resistance. The directed evolution protocol utilized to generate these four populations included an outgrowth step post-irradiation. The intention was to allow unfit, slow-growing strains with deleterious genetic alterations to be outcompeted from each population. This step appears to have also selected for enhanced growth. Variants of GsiB in both CB2000 and CB3000 have substantial positive effects on fitness of growth in rich media without selection (Figure 10.2).

A problem with studies using irradiators employing radioactive elements is that those elements decay, and irradiation conditions change over time. This problem is likely the root cause of our inability to reliably characterize the IR resistant phenotype of CB3000. Here, we can report that alterations to global anaerobic respiration control (through FNR 186I) can enhance IR resistance in *E. coli*, with putative contributions from altered global proteolysis (through *clpP/clpX* int) and DNA replication restart (from PriA V554I). However, steps must be taken to ensure that the problems encountered with CB3000 are not built into future evolution studies. It is clear that IR resistance in *E. coli* can surpass what has been observed here. Utilization of more stable

irradiation methods may allow for generation of new populations of highly IR resistant *E. coli*, and deeper exploration of the genetic contributions to radiation resistance.

2.5 Tables

Tier	Variant Freq.	Detection in triplicates	Variant on both strands	IR-1-20	IR-2-20	IR-3-20	IR-4-20
Tier_1	>95%	All Three	Both Strands	13	37	44	44
Tier_2	5% - 95%	All Three	Both Strands	128	76	142	113
Tier_3	1-5%	All Three	Both Strands	312	87	123	131
Tier_4	1-10%	At least one	Both Strands	291	118	392	302
Tier_5	1-10%	At least one	One Strand	345	397	265	268
Tier_6	Any (0.3% to 10%)	0 - 3	1 or 2	868	910	1914	1209
Total number of high confidence variants (tiers 1 - 3)				453	200	309	288
Total number of variants (tiers 1 - 6)				1957	1625	2880	2067
Mean sequence coverage of variants				1947	1593	934	1391

Table 1.2. Tier definitions of SNPs detected by whole-population sequencing. Sequencing data was produced as described in the *Materials and Methods*. Data analysis performed by Xinmin Xhang (BioInfoRx, Inc; Madison, WI).

	Gene	Allele	IR-1-20			IR-2-20	IR-3-20	IR-4-20	IR-CB1000 +
			IR-1-20 A	IR-1-20 B	IR-1-20 C				
DNA Metabolism	RecA	D276A	-	15%	-	-	-	-	-
		D276N	-	-	-	100%	-	100%	-
		A289S	56%	-	-	-	-	-	-
	DnaG	R139H	-	-	-	84.56%	-	-	-
	DnaN	V310I	-	-	-	-	-	7.62%	-
	DnaB	P80H		(4.95%)		100%	-	-	-
		P199Q	-	-	-	-	-	100%	-
		L74S	-	3.93%	-	-	-	-	-
		N398K		(3.49%)		-	-	-	-
	DnaT	R145C	58.50%	-	-	-	-	-	-
	PriA	V554I	-	-	-	-	100%	-	-
		G279D		(2.23%)		-	-	-	-
	PriC	L162P	-	-	11%	-	-	-	+
	YfjK	A152D	-	-	-	100%	-	-	-
		K294E	-	15%	-	-	-	-	-
		H652Y	-	-	-	-	-	100%	-
		I172L		(2.98%)		-	-	-	-
		P252S	-	-	-	4.32%	-	-	-
ROS amelioration	GsiB	N105S	-	-	-	100%	-	-	-
		L289P	-	-	-	-	100%	-	-
		V481A	-	-	-	-	-	38.38%	-
	RsxB	K122E	-	-	-	100%	-	-	-
	RsxD	V44A	-	-	-	-	100%	-	-
	RsxC	I51I	-	-	-	-	-	57.63%	-
		P671Q		(1.84%)		-	-	-	-
		Q582E		(1.33%)		-	2.82%	-	-
		A441T	-	-	-	-	2.44%	-	-
		F344	-	-	-	-	-	-	+
Protein metabolism	ClpP	Y62C/Y76C	-	-	19%	-	-	-	+
		E51A/E65A		(3.74%)		-	-	-	-
		T133N/T147N		(1.41%)		-	-	-	-
	ClpX	Y384C	1.41%	-	-	-	-	-	-
	clpP/clpX	G -13 A	-	-	-	-	100%	-	-
	DnaJ	G229G	-	-	11.98%	-	-	-	+
Cell wall metabolism	NanA	V135M	-	-	-	-	-	-	+
	NanE	A129T	-	-	-	100%	-	-	-
	NanM	T237M	-	-	-	-	5.89%	-	-
	NanT	F406S	-	-	-	-	100%	-	-
	WcaA	K253I	-	-	-	-	4.44%	-	-
	WcaC	S314S	-	-	-	-	100%	-	-
	WcaG	- 1 frameshift at F173		(3.78%)		-	-	-	-
	WcaJ	D278D	-	-	9.70%	-	-	-	-
	WcaK	Y133C	-	-	-	86%	-	-	-
	WcaL	G236S	-	-	-	-	-	-	+
	WcaM	N157S	-	-	-	-	-	35.82%	-
	WcaN	-1 frameshift at H112	-	-	-	-	-	38.16%	-
G3P metabolism	GlpD	V256M	-	-	-	100%	-	-	-
	GlpQ	E171D/E146D	-	-	-	-	9.80%	-	-
Anaerobic metabolism	Fnr	F186I	-	-	-	-	100%	-	-
		M157*	-	-	-	-	-	-	+

Table 2.2. **Prominently mutated genes and pathways within IR resistant populations.** Mutations listed are high confidence mutations that have been detected in Illumina Mi-Seq sequencing of triplicate replicates of the population on both the forward and reverse strands of the genome (See *Materials and Methods*). The two far-left columns contain the pathway and genes in which the variant appears. Variants with two listings (ex: ClpP T133N/T147N) indicate the variant

in the mature peptide and full peptide sequence, respectively. The top row indicates the population in which the mutation occurs. The lack of a mutation in the given gene is indicated by a – symbol. The frequency of each variant within the mixed population is listed if the mutation is present in the given population. IR-1-20 has three prominent sub populations, designated IR-1-20 A, B and C. Mutations that have not been detected by sequencing of individual mutants from IR-1-20 [13] and therefore cannot be assigned to a particular sub-population but have been detected in whole-population sequencing are designated by parentheses.

Strain	Relevant Genotype	Source
Founder	MG1655 + RbsR L92R + CytR Q110 stop + RIP321 A 4296380 ACG + YbhJ L54I + MntP G25D + *fabI* /*ycjD* int (G 1351174* A) + *yifN* /*ppiC* int (T 3959934* C)	Harris 2009
EAW7704	Founder Δe14 prophage	Harris 2009
IR-1-20	Founder exposed to 20 iterative rounds of IR; mixed population	Harris 2009
IR-2-20	Founder exposed to 20 iterative rounds of IR; mixed population	Harris 2009
IR-3-20	Founder exposed to 20 iterative rounds of IR; mixed population	Harris 2009
IR-4-20	Founder exposed to 20 iterative rounds of IR; mixed population	Harris 2009
CB2000	Isolate from IR-resistant evolved population IR-2-20	Harris 2009
CB3000	Isolate from IR-resistant evolved population IR-3-20	Harris 2009
EAW156	Founder Δe14 + RsxD V45A	This study
EAW239	Founder Δe14 + RecA D276N + DnaB P80H + YfjK A151D	Byrne 2014
EAW348	CB3000 + *clpP* /*clpX* wt	This study
EAW349	CB3000 + GsiB L288 wt	This study
EAW350	Founder Δe14 + *clpP* /*clpX* int	This study
EAW352	Founder Δe14 + GsiB L288P	This study
EAW353	CB3000 + FNR F186 wt	This study
EAW354	Founder Δe14 + FNR F186I	This study
EAW355	CB3000 + PriA V554 wt	This study
EAW356	Founder Δe14 + PriA V554I	This study
EAW 365	Founder Δe14 + NanT F406S	This study
EAW493	Founder Δe14 + *clpP* /*clpX* int + FNR F186I	This study
EAW494	CB3000 + *clpP* /*clpX* wt+ FNR F186 wt	This study
EAW647	CB3000 + *clpP* /*clpX* wt+ FNR F186 wt + PriA V554 wt	This study
EAW648	Founder Δe14 + *clpP* /*clpX* int + FNR F186I + PriA V554I	This study
EAW655	Founder Δe14 + FNR F186I + PriA V554I	This study
EAW656	Founder Δe14 + *clpP* /*clpX* int+ PriA V554I	This study
EAW657	CB3000 + FNR F186 wt + PriA V554 wt	This study
EAW658	CB3000 + *clpP* /*clpX* wt + PriA V554 wt	This study
EAW662	Founder Δe14 + *clpP* /*clpX* int+ FNR F186I + PriA V554I + RecA D276N + DnaB P80H + YfjK A151D	This study
* nucleotide position compared to NCBI GenBank U00096.3 reference sequence **(citation)**		

Table 3.2. Strains Used in this study.

Name	Source	Sequence (5' - 3')	Structure
dT28	This study	TTTTTTTTTT TTTTTTTTTT TTTTTTTT	
1b-98	Heller & Marians, 2005	GCAAGCCTTC TACAGGTCGA CCGTCCATGG CGACTCGAGA CCGCAATACG GATAAGGGCT GAGCACGCCG ACGAACATTC ACCACGCCAG ACCACGTA	
3L-98 (+/- 5' fluorescin)	Heller & Marians, 2005	GACTATCTAC GTCCGAGGCT CGCGCCGCAG ACTCATTTAG CCCTTATCCG TATTGCGGTC TCGAGTCGCC ATGGACGGTC GACCTGTAGA AGGCTTGC	
11b-38	Heller & Marians, 2005	GAGCACGCCG ACGAACATTC ACCACGCCAG ACCACGTA	
B-33	Heller & Marians, 2005	AGTCTGCGGC GCGAGCCTCG GACGTAGATA GTC	

Table 4.2. Oligos used for *in vitro* assays in this study. [124, 125]

2.6 Figures

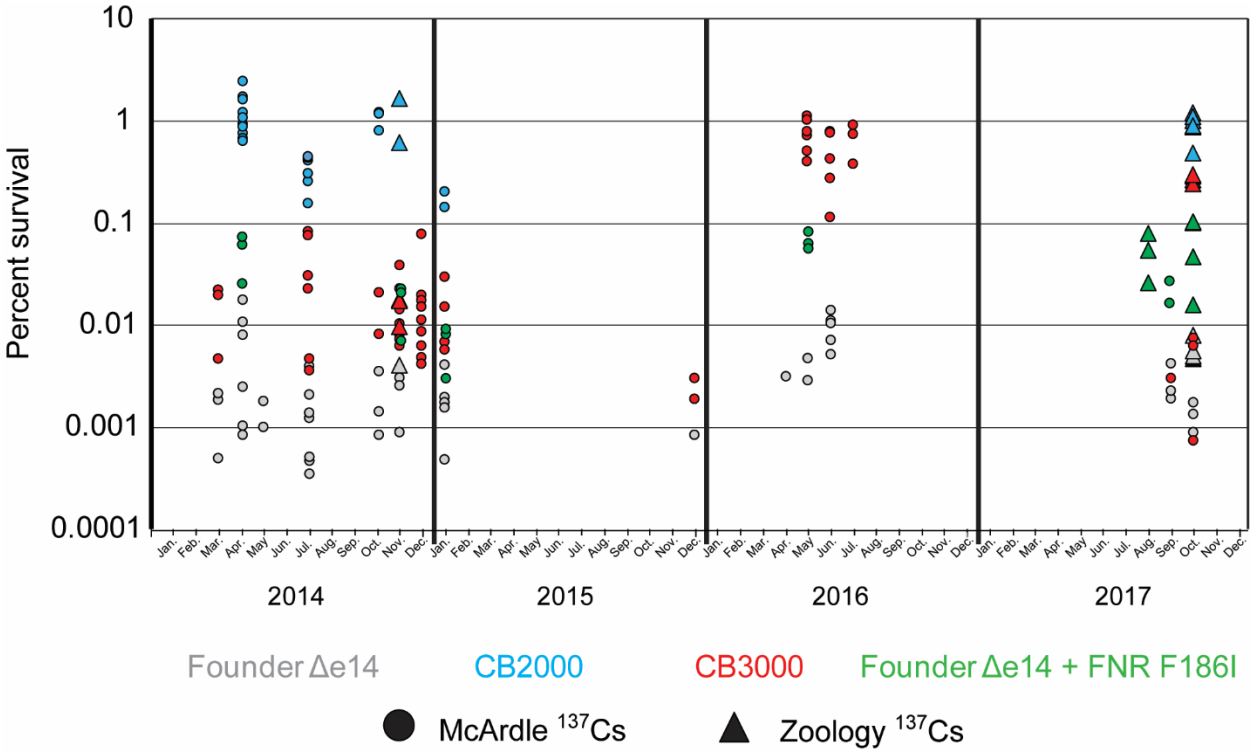


Figure 1.2. Variability of IR-resistance phenotype of evolved isolate CB3000. Each point represents the percent survival of a single replicate. Many replicates were tested on the same day or within the same month, and appear vertical relative to each other. Circles and triangles indicate that those samples were irradiated using two different irradiators, but are the same model (Mark I ^{137}Cs irradiator (J. L. Shepherd and Associates)). Data was obtained over the course of four consecutive years.

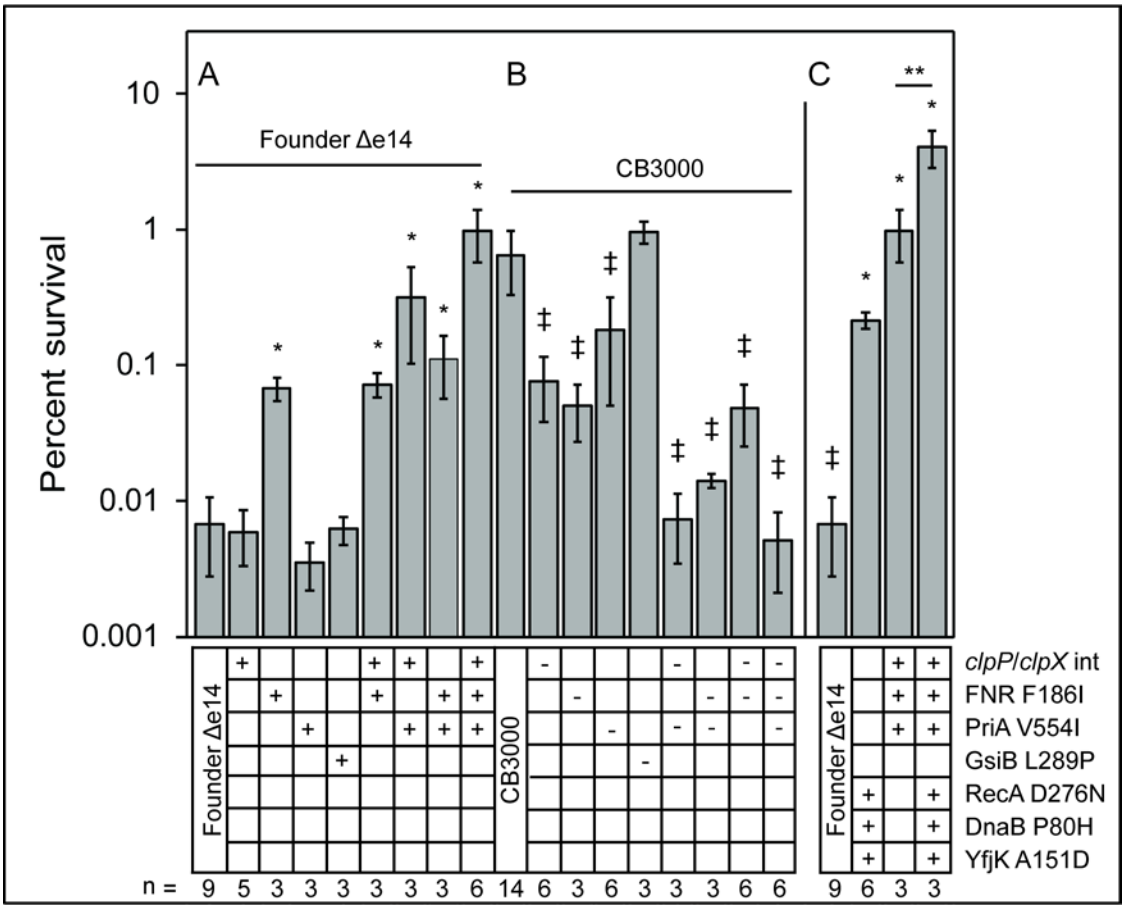


Figure 2.2. IR resistance of CB3000 and CB3000-derived mutants over a two month period in 2016. The Y axis represents percent survival. Mutants were assayed for survival of exposure to 3,000 Gy at exponential phase growth as described in the *Materials and Methods* section. (A) Mutations discovered in CB3000 were moved singly or in combination into the Founder Δe14 background. The mutation present is given by the + symbol. (B) Mutations discovered in CB3000 were reverted singly or in combination to the Founder sequence. Reversion of the mutation is given by the – symbol. (C) Mutations shown to increase IR resistance from IR-3-20 were compared and added to those discovered previously in IR-2-20 [90] in the Founder Δe14 background. The number of replicates for each strain, n, is presented below the appropriate strain. Error bars represent the standard deviation. A single * symbol indicates a statistically significant difference

(p-value < 0.01) between the percent survival of the indicated strain and Founder Δe14, as calculated with a two-tailed student's t-test. The ‡ symbol indicates a statistically significant difference (p-value < 0.01) between the percent survival of the indicated strain and CB3000, as calculated with a two-tailed student's t-test. The ** symbol indicates a statistically significant difference (p-value < 0.001) between indicated strains, as calculated with a two-tailed student's t-test.

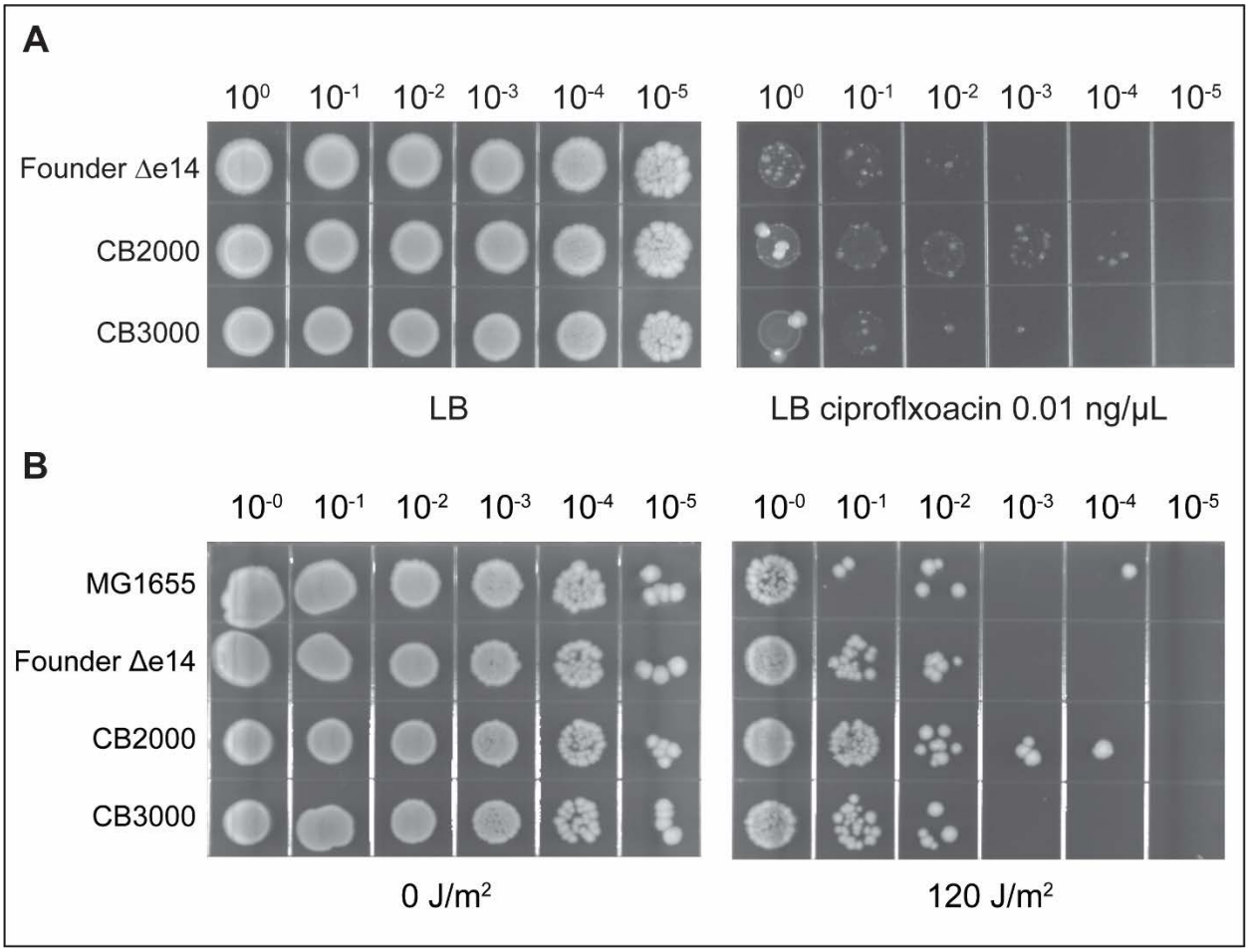


Figure 3.2. Resistance of evolved isolates CB2000 and CB3000 to DNA damaging agents. A) Comparison of the evolved isolates CB2000 and CB3000 to Founder Δe14 when spot plated onto LB plates containing 0.01 μg/ml ciprofloxacin as described in the *Materials and Methods* section. B) Comparison of the evolved isolates CB2000 and CB3000 to Founder Δe14 and MG1655 (+e14) when exposed to 120 J/m^2 of UV after spot plating on LB agar as described in the *Materials and Methods* section.

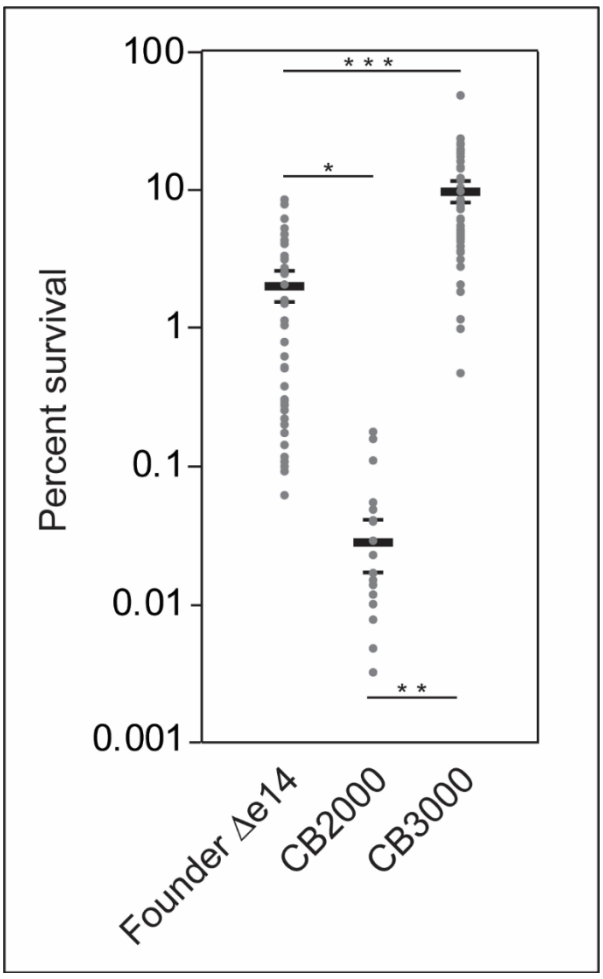


Figure 4.2. H_2O_2 resistance of evolved isolates and their derivatives. All strains were grown for 7 hr to stationary phase and exposed to H_2O_2 as described in the *Materials and Methods* section. The Y axis represents percent survival. Comparison of the evolved isolates CB2000 and CB3000 to Founder Δe14. Method modified from Gort *et al*. [126]. Data points represent percent survival across all independent experiments performed (n values: Founder Δe14: 36; CB2000: 14; and CB3000: 46). The center bar represents the average value. Error bars represent the standard deviation. The *, **, and *** symbols indicate a statistically significant difference (p-value < 0.01, <0.001, and $<10^{-5}$, respectively) between the indicated strains, as calculated with a two-tailed student's t-test.

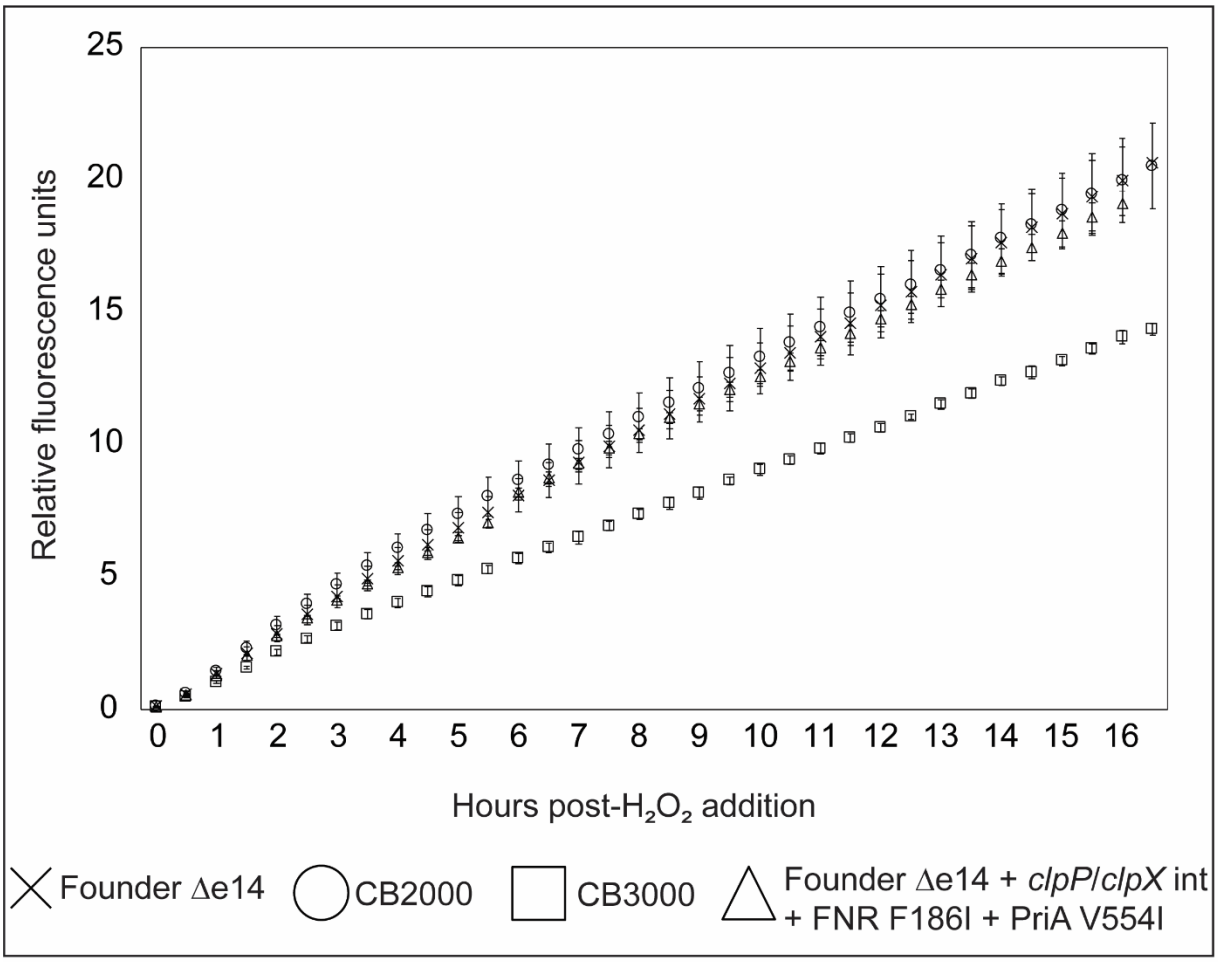


Figure 5.2. ROS amelioration capability of evolved isolates after exposure to 10 mM H_2O_2. These data represent the results of a representative single experiment. The Y axis represents raw fluorescence of oxidized H_2DCFDA normalized to OD_{600}. Error bars represent the standard deviation of biological triplicates. The ROS amelioration capacity assay was performed as described in the Materials in Methods section.

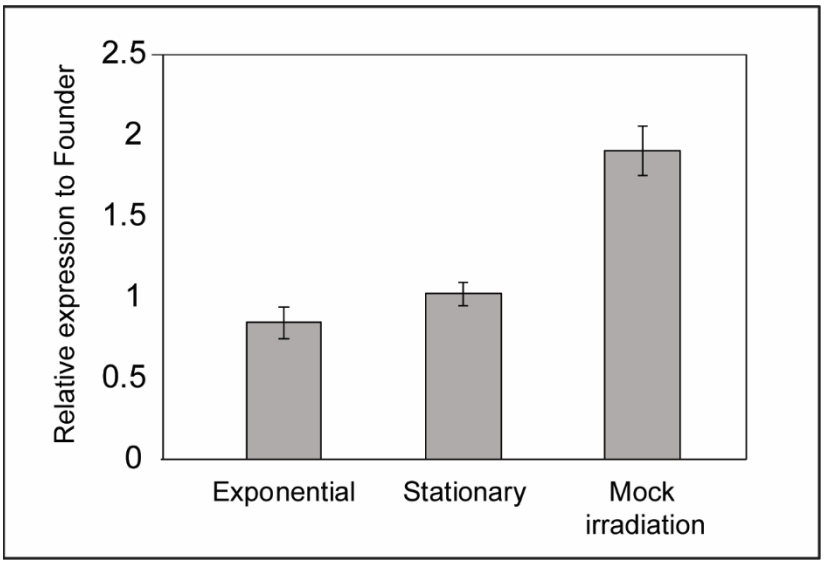


Figure 6.2. β-galactosidase activity of a P_{narG}-*lacZ* (fusion in Founder Δe14 + FNR F186I). β-galactosidase activity was assayed at exponential and stationary phase, and after a mock irradiation (exponential phase cells aliquoted in a 1.5 mL tube and incubated at room temperature for 8 hr). Beta gal activity of Founder Δe14 + FNR F186I P_{narG}-*lacZ* was normalized to the average Beta-gal activity of Founder Δe14 + P_{narG}-*lacZ* (containing the wild-type FNR F186 allele) to determine relative expression. These data represent the results of two experiments of biological triplicates. The β-galactosidase assay was performed as described in the Materials in Methods section. P_{narG}-*lacZ* was graciously given to us by the lab of Patricia Kiley (University of Wisconsin – Madison) [106].

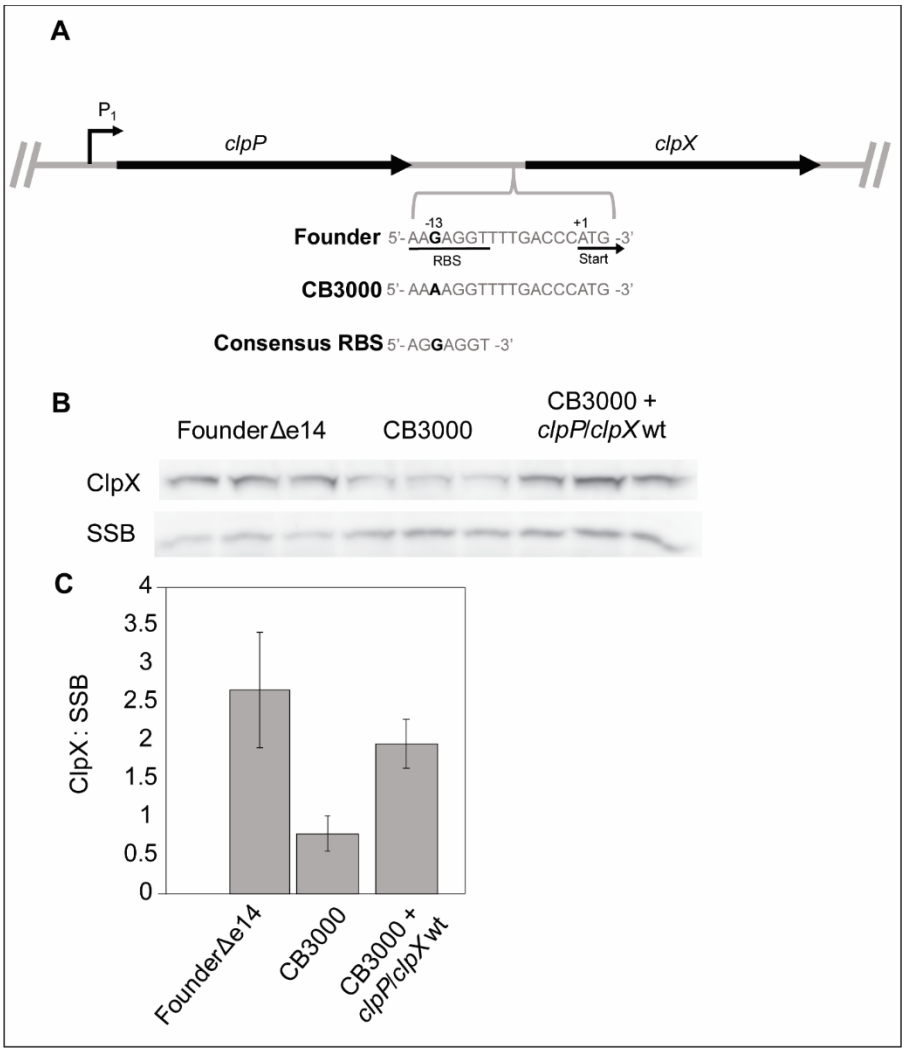


Figure 7.2. Effect of the *clpP*/*clpX* int mutation on ClpX expression levels in CB3000. A) Cartoon depiction of the *clpPX* operon. The *clpP*/*clpX* int mutation is located at the -13G of the *clpX* transcription start site. The *clpP*/*clpX* int may be located in the ribosome binding site for the *clpX* mRNA based on similarity to the *E. coli* consensus RBS sequence. B) Western blot of the ClpX protein in CB3000. α-ClpX antibodies were used to detect ClpX protein in Founder Δe14, CB3000, and CB3000 + *clpP*/*clpX* wt via Western blot using chemiluminescent detection. α-SSB antibody was used as an internal control. C) Levels of ClpX were normalized to levels of SSB to quantify the relative amount of ClpX between strains with and without the *clpP*/*clpX* int mutation. Western blot was performed as described in the *Materials and Methods.*

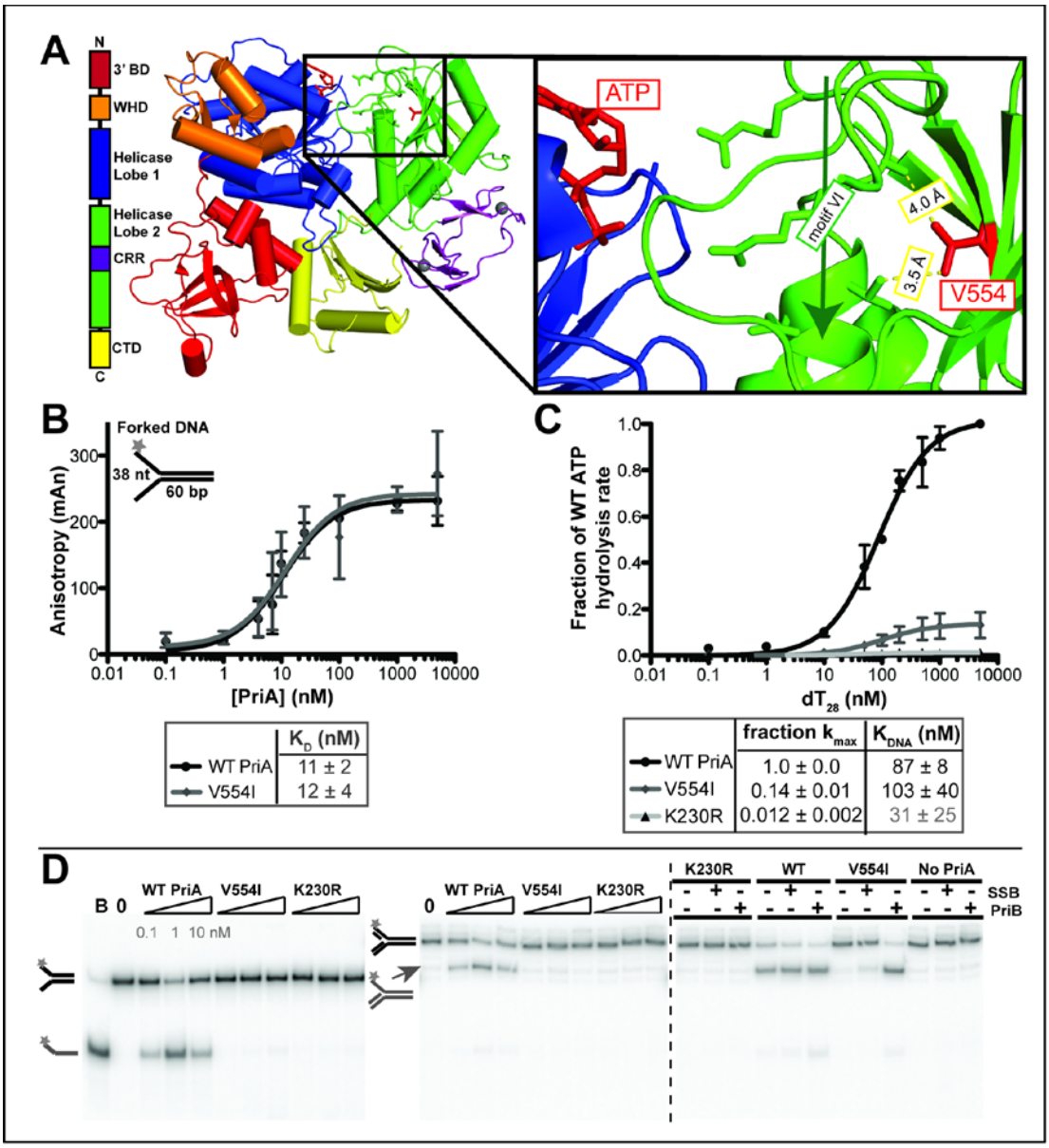


Figure 8.2. PriA V554I variant has significantly reduced helicase activity that can be stimulated by PriB. (A) Domain architecture (left) and the structure of the *K. pneumoniae* PriA (PDB: 4NL4) (middle) [115], with a zoom-in view of the ATP-pocket and helicase motif VI (right). ATP and the V554 side-chain are shown in red sticks. Helicase motif VI is shown in green sticks, with a green arrow indicating the beginning and end of the motif. **(B)** Fluorescence anisotropy of a DNA fork (1 nM) containing a 5'-fluorescein labeled template lagging strand with

increasing concentrations of PriA. Data from three replicates (± SD) were fit to a single site-specific model (solid line) and dissociation constant (± SD) from that fit is listed in the table (below). **(C)** PriA DNA-dependent ATP hydrolysis under increasing concentrations of dT_{28} oligonucleotide, compared to the max ATPase rate of WT PriA for each replicate. Helicase/ATPase-dead PriA (Walker A, K230R) was used as a negative control. Data for three replicates were fit to a Michaelis-Menten model (solid line) and fit parameters (± SD) listed in the table below. As PriA ATPase is DNA dependent and DNA was titrated (under excess ATP and 50 nM PriA), K_{DNA} was reported in place of the K_D parameter. **(D)** Polyacrylamide gel electrophoresis resolution of PriA helicase products upon incubation of 0, 0.1, 1, or 10 nM PriA variant with 1 nM DNA fork. "B" labels boiled substrate. Initial DNA forks were either the simple two oligonucleotide fork (left gel panel), using a 5'-^{32}P label in place of fluorescein in panel A, or a four oligonucleotide fork (middle and right gel panels), upon which PriA is known to preferentially unwind the nascent lagging strand. Initial forks are depicted in black to left of gel panel and helicase products are depicted in grey to left. Image is representative gel of three replicates. Figure and experiments were generously provided by Tricia Windgassen and colleagues (laboratory of Professor James Keck; University of Wisconsin – Madison).

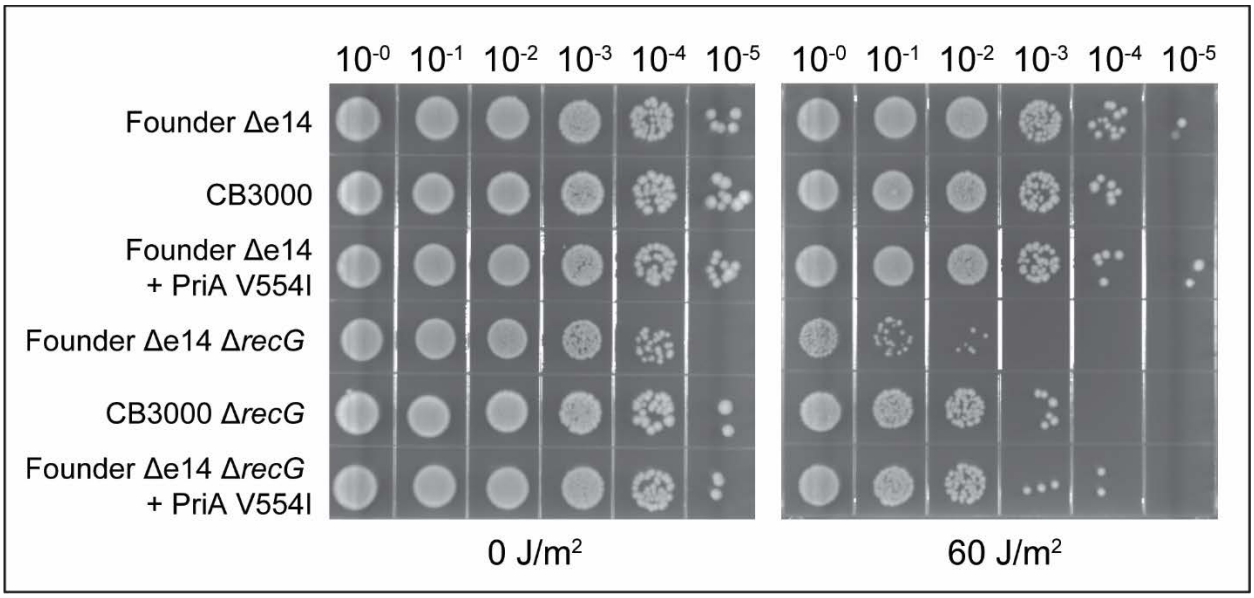


Figure 9.2. Rescue of UV sensitivity of Δ*recG* mutants by the PriA V554I variant. All strains were grown to mid-exponential phase before being serial diluted and spot plated onto LB plates exposed to 60 J/m^2 UV as described in the *Materials and Methods* section. Comparison of the evolved isolates CB2000 and CB3000 to Founder Δe14.

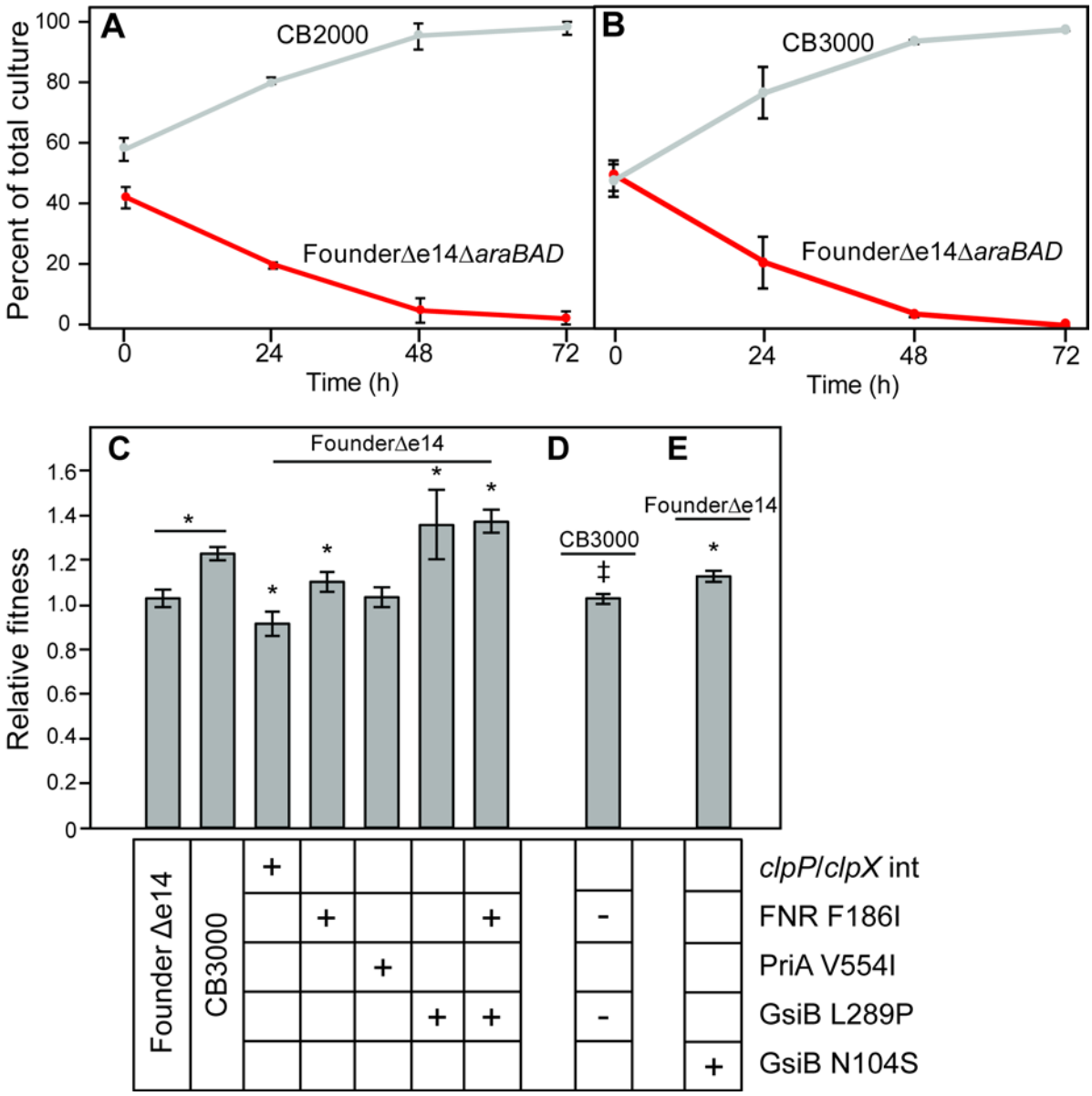
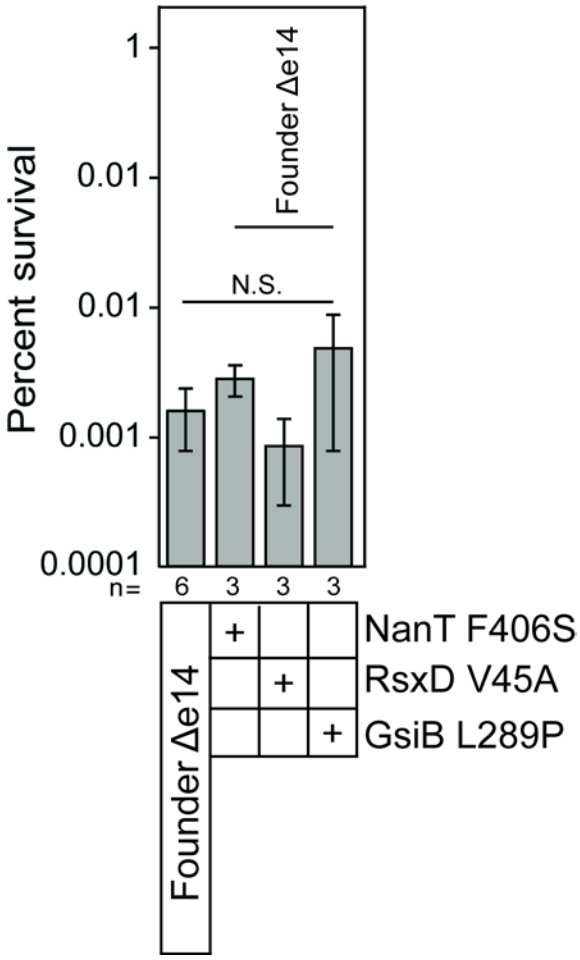


Figure 10.2. Evolved isolates outcompete Founder Δe14 in mixed culture without selection. A, B) CB2000 and CB3000 outcompete Founder Δe14. These data represent a single, representative competition against Founder Δe14 without the *araBAD* operon (a neutral mutation used to differentiate strains in the competition [121]. C) Mutations from CB3000 which contribute to the enhanced growth phenotype. To quantify contributions to the growth phenotype, we calculated a relative fitness statistic [121]. Mutations were placed singly or in combination into

Founder Δe14. All competitions were performed against Founder Δe14 Δ*araBAD*. D). CB3000 lacking mutations contributing to enhanced growth does not outcompete Founder Δe14. Mutations indicated were converted to the wild type sequence, and this mutant strain was competed against Founder Δe14 Δ*araBAD*. E) The GsiB N104S variant from population IR-2-20 contributes to the enhanced growth phenotype. Mutations were placed singly or in combination into Founder Δe14. The growth competition protocol was performed as previously described [121] [90] and as in the *Materials and Methods*. All data from C-E are the average of relative fitness statistics from at least six replicate competition assays performed against Founder Δe14 Δ*araBAD*. Error bars represent the standard deviation of these data. A single * symbol indicates a statistically significant difference (p-value < 0.01) between the relative fitness statistic of the indicated strain and Founder Δe14, as calculated with a two-tailed student's t-test. The ‡ symbol indicates a statistically significant difference (p-value < 0.01) between the relative fitness statistic of the indicated strain and CB3000, as calculated with a two-tailed student's t-test.



Supplemental Figure 1.2. IR resistance of Founder Δe14 with CB3000-derived mutations which do not affect IR resistance. The Y axis represents percent survival. Mutants were assayed for survival of exposure to 3,000 Gy at exponential phase growth as described in the *Materials and Methods* section. Mutations discovered in CB3000 were moved into the Founder Δe14 background. The mutation present is given by the + symbol. Strains were irradiated at approximately 6.6 Gy/min. Irradiations were performed using the same irradiator used to produce data in Figure 2.2, but approximately 3 years prior, thus producing the higher dose rate. Error bars represent the standard deviation. 'N.S.' indicates that there is no statistically significant difference between the three sets of data (p-value > 0.05), as calculated with a two-tailed student's t-test.

Chapter III: Directed evolution of extreme electron beam ionizing radiation resistance in *E. coli*: Selection cycles 1-100 and beyond

3.1 Abstract

We have undertaken an extended effort to generate *Escherichia coli* populations that are as resistant to ionizing radiation (IR) as *Deinococcus radiodurans*. We report here on results obtained after 100 cycles of selection, with special focus on isolates from the first 50 cycles of selection. Using high-energy electron beam IR, we have generated four replicate populations of *E. coli* that exhibit greater IR resistance than that seen in any comparable experiment to date. The appearance of genomic alterations has been monitored by regular deep-sequencing of each population, permitting the detection of evolutionary patterns. The resistance of one isolate, IR9-50-1, is derived largely from two mutations: a variant enhancing DNA double-strand break repair (RecD A90E) and mutations in RNA polymerase (RpoB S72N/RpoC K1172I). These adaptations reflect pathways of IR resistance developed in three of the four populations; additional mechanisms of resistance remain to be discovered.

3.2 Introduction

Ionizing radiation (IR), either direct ionization by high energy photons and electrons, or indirect ionization from reactive oxygen species (ROS) produced by interaction of IR with water molecules, can damage any cellular component. Fortunately, levels of IR sufficient to cause widespread oxidation of proteins or extensive DNA damage (including double strand breaks) are rarely encountered in the environment.

Substantial increases to bacterial IR resistance can be generated in the laboratory by directed evolution [86-89]. With the advent of advanced DNA sequencing technologies to reveal genomic changes, we have utilized directed evolution as a facile path to a better understanding of mechanisms underlying the extreme IR resistance phenotype. We previously exposed four replicate populations of *Escherichia coli* to 20 iterative rounds of gamma-ray IR and observed significant increases in IR resistance [13]. Focusing on isolates from each population, we determined that IR-resistance arose from enhanced DNA repair (reflected in mutant alleles of the recombinase RecA, replicative helicase DnaB, putative helicase YfjK, and replication restart machinery), alterations in global proteomic regulation (with alterations in expression of the ClpXP protease), and changes in the regulation of cellular responses to oxidative damage (through an allele of the anaerobic metabolism transcription factor, FNR) **[13, 90]**.

The overall IR resistance of these populations, while increased substantially, did not reach a level comparable to the incredibly radioresistant bacterium *Deinococcus radiodurans* [1]. Efforts to further enhance the IR resistance phenotype have been constrained both by source decay (altering irradiation dose rate parameters and increasing times required to impose a particular dose on a sample) and by new governmental policies mandating a general decommissioning of radioactive sources used for research. To generate *Deinococcus*-like levels of IR resistance in

evolved *E. coli*, we have embarked on a new and more ambitious effort with four new populations. This new effort employs a clinical linear accelerator (Linac) of a type employed for radiation treatment of cancer patients. Using this device, we can achieve a dose rate of 72 Gy/min using a high energy electron beam. This dose rate is nearly 4-fold higher than that used in the previous directed evolution study [13, 90]. Use of this instrument eliminates the problem of radioactive source decay and greatly reduces the amount of time required for exposure to kGy levels of IR. The higher dose rate also allows us to explore adaptation to a potentially much more challenging IR environment than that utilized in earlier trials and determine if dose rate can affect the results of an evolution trial.

In this report, we present data of the IR resistant phenotype of these four new *E. coli* lineages (IR9, IR10, IR11, and IR12) after over 100 cycles of selection, with deep-sequencing of populations at every other cycle of selection to round 60. The dose required to kill 99% of the cells in each population has increased from approximately 750 Gy to over 3000 Gy after over 100 rounds of selection. Additionally, we further characterize the four lineages after the first 50 rounds of selection (IR9-50, IR10-50, IR11-50, and IR12-50). These populations are highly resistant compared to the previously evolved isolates CB2000 and CB3000[13, 90], which turn out to be poorly adapted to the much higher dose rates applied with the electron beam IR. Deep genomic sequencing of populations after every other selection cycle allow us to explore the full breadth of mutations and provides a revealing window on the evolution of IR resistance. To unveil IR resistance mechanisms, we focus particular attention on the lineage IR-9 after 50 cycles of selection, termed IR9-50, as reflected in a representative isolate from this population IR9-50-1. The results both reinforce some earlier conclusions and offer new insights. The phenotype in IR9-50-1 is largely explained by two mutations. As seen previously, adaptations to the DNA double

strand break repair machinery (represented in this case by RecD A90E), can contribute to an increase in IR resistance. However, alterations in RNA polymerase (RpoB S72N/RpoC K1172I) can also contribute substantially to an IR resistant phenotype. The other three evolved populations exhibit patterns that both reinforce those seen in IR9-50 and also potentially reveal some additional novel approaches to IR resistance.

3.3 Results

3.3.1 High-energy electron beam IR kills E. coli at a similar rate to high-energy photon IR

We sought to determine if high-energy electron beam ionizing radiation (IR) kills *E. coli* similarly to high-energy photon IR. We note that the ^{60}Co source used in our previous directed evolution studies was a photon source [13], although the dose rate was substantially lower than the dose rates used here. The linear accelerator (Linac) used in this study has both an electron and photon mode and, by changing the distance of the *E. coli* cultures to the source of the IR beam, the Gray per minute (Gy/min) dose rate can be made equivalent between these modes. At a dose rate of 17 Gy/min, high energy electrons and photons kill nearly an identical percentage of *E. coli* MG1655 culture at 1000 Gy, with the electron beam being slightly more potent (Figure 1.3). Therefore, the two IR modes appear to have comparable effects at the level of bacterial survival.

3.3.2 Directed evolution of extreme IR resistance over 100 rounds of selection

Beginning with a single cultures of the *E. coli* strain MG1655 (four single colony isolates), we generated four highly IR resistant populations using a modified version of a previously established directed evolution protocol [13] as described in the *Materials and Methods* and depicted in Figure 2.3. Importantly, in this new set of evolution trials, we have utilized a linear accelerator (Linac) to deliver dose with a high energy electron beam. Our previous evolution trials to generate IR resistant *E. coli* utilized a ^{60}Co source with a dose rate of approximately 19 Gy/min [13]. Using the Linac in electron mode, we are utilizing a dose rate of 72 Gy/min. As there is no source decay with this device, the dose rate remains constant over the weeks and years needed for an extended evolution trial.

We exposed exponential phase cultures from four replicate *E. coli* populations to sufficient high energy electron beam IR to kill approximately 99% of each culture for 100 iterative cycles (with selection cycles occurring approximately once per week). The resulting populations are designated IR9-100, IR10-100, IR11-100, and IR12-100. A full explanation of the nomenclature of evolving lineages, populations, and isolates is described in Figure 3.3. Over these rounds of evolution, the dose required to kill 99% of each culture increased from approximately 750 to 3000 Gy (Figure 4.3).

3.3.3 Radioresistance of evolved populations at round 50 and round 100

At rounds 50 and 100, these populations are highly radioresistant compared to the parent MG1655 strains (Figure 5.3). Previously described IR resistant *E. coli* isolates [13, 90], which were evolved to withstand ^{60}Co IR at about 4-fold lower dose rates, are only moderately more resistant to the beam generated by the Linac compared to the parent strain and are greatly surpassed by the populations generated in this study. Populations IR9-100, IR10-100, IR11-100, and IR12-100 have not yet reached the level of IR resistance of the bacterium *Deinococcus radiodurans* but have made significant progress towards this goal, with a current shoulder of IR-resistance reaching approximately 2000 Gy. We note that *Deinococcus* is more sensitive to the higher dose rate in the Linac electron source than it is to ^{60}Co photon sources [19, 56].

3.3.4 Growth phenotypes of evolved population isolates

The IR resistance phenotype of these experimentally-evolved populations comes without an apparent growth defect in LB rich medium. An isolate from each population at round 50 of selection (IR9-50-1, IR10-50-1, IR11-50-1, and IR12-50-1) produced growth curves comparable

to the parent MG1655 strain in LB. (Figure 6.3A). However, growth defects of these isolates become apparent when inoculated in other growth media. In a different defined rich medium supplemented with glucose, EZ medium, isolate IR12-50-1 fails to grow and isolates IR9-50-1 and IR11-50-1 exhibit growth defects easily seen in a standard growth curve (Figure 6.3B). The isolate IR10-50-1 grows similarly to MG1655 in EZ medium and is the only of the four isolates to grow in M9 minimal medium supplemented with glucose (Figure 6.3C). Although IR9-50-1, IR11-50-1, and IR12-50-1 have potentially inactivating mutations in key amino acid biosynthetic pathways that could prevent growth in minimal medium (IR10-50-1 does not have a mutation which would clearly affect amino acid biosynthesis), supplementing minimal medium with these amino acids does not rescue the ability of these isolates to grow (Supplemental Figure 1.3).

Growth competition assays [121] in LB medium reveal growth phenotypes of these isolates that are not readily apparent when comparing standard growth curves (Figure 7.3). When competed against the parent wild type *E. coli* strain in LB, all four evolved isolates are rapidly outcompeted, even with significant dilutions to bias the starting culture ratios towards the evolved strains. These results indicate that 50 rounds of selection in this directed evolution protocol does not come without an underlying fitness cost without selection from IR.

3.3.5 Resistance of evolved isolates to other forms of DNA damage

Isolates from the four evolved populations after 50 and 100 cycles of selection reveal that resistance to IR does not correlate in any predictable way with resistance to other DNA damaging agents, even within a single population (Figure 8.3, 9.3). Multiple isolates from each population exhibit highly variable levels of resistance to UV radiation (Figure 8.3), mitomycin C (which causes DNA interstrand crosslinks), ciprofloxacin (which causes DNA double-strand breaks

through inhibition of DNA gyrase), bleomycin (which causes DNA double-strand breaks and apurinic/apyrimidinic sites through a ROS mediated mechanism [127], and hydroxyurea (which reduces dNTP production and may cause ROS-mediated DNA damage) [128-131] (Figure 9.3). While some isolates have increased resistance to these DNA-damaging agents, others exhibit no change or increased sensitivity. The results indicate that the directed evolution trials are generating specialists that are uniquely adapted to IR resistance.

3.3.6 Deep-sequencing of evolved populations reveals molecular trends in evolution

We utilized deep-sequencing technology to monitor genomic changes as IR resistance increased with continued selection cycles. At each even round of selection genomic DNA was prepared from each population and was submitted to the Joint Genome Institute (Walnut Creek, CA) for sequencing using an Illumina Hi-seq platform to determine all mutations present in each population. We currently have sequencing data for each population up to round 60 of evolution. Stitching together these snapshots reveals the underlying complexity of these evolving populations (Figure 10.3). The numbers of sub-populations and total number of mutations increases rapidly (Figure 11.3).

Despite starting from the same parent stain, the evolutionary path of each population is unique. The only event in common to each population is the excision of the e14 prophage, as indicated by four mutations in the *icd* gene in which the prophage is inserted (as depicted by the initial single black line appearing in each population in (Figure 10.3). When e14 is excised, the *icd* gene is restored to the wild-type sequence. Despite excision of e14 occurring quickly in each population, that excision becomes fixed in only three of the four populations. The sub-population

that lost the e14 prophage was displaced by another sub-population in the IR10 lineage, and e14 is retained in population IR10-60.

3.3.7 Numbers and types of mutations in evolved populations after 50 cycles of selection

At round 50 of selection, there are distinct difference between populations. Populations IR9-50 and IR10-50 are the most clonal of the four populations, with approximately 16 and 17%, respectively, of the total high confidence (> 2% allele frequency) mutations being fixed (detected at > 99% frequency). In contrast, only 4% and 11% of high confidence mutations in IR11-50 and IR12-50 are fixed, respectively (Table 1.3). Although IR9-50 and IR10-50 share a similar percentage of fixed mutations, IR9-50 and IR12-50 have the highest total number of mutations. Each population has a ratio of non-synonymous to synonymous mutations greater than two. This ratio, the dN/dS value, is considered indicative of positive selection on a population if greater than one [13].

Mutations in coding regions which likely cause loss of function of the protein product (single and double base insertions/deletions, gain or loss or stop codons, and loss of start codons) appear in each population. However, IR9-50 has the most predicted loss of function mutations (71) out of the other three populations (IR10-50: 47, IR11-50: 37, IR12-50: 50). Of these mutations in all populations, -1 frameshift mutations are the most common, followed by introduced stop codons.

SNPs in which guanine or cytidine are changed to adenine or thymine dominate, agreeing with mutational patterns detected in previously evolved IR-resistant *E. coli* populations [13, 90]. Surprisingly, base substitutions are approximately 50% transitions and 50% transversions, whereas the *E. coli* genomes exposed to ^{60}Co gamma-ray irradiation exhibit transversions at about four times the frequency of transitions [12, 13]. These results indicate that the selection utilized in the

current study produces a substantially different mutational spectrum relative to the populations evolved to be resistant to ^{60}Co.

3.3.8 Mutations which enhance IR resistance in isolate IR9-50-1

We have begun characterization of evolved IR resistance after 50 cycles of selection. To determine the molecular basis of IR resistance in one of the populations, we more carefully characterized one isolate from population IR9: IR9-50-1. This isolate is representative of several mutational patterns seen in most of the populations. IR9-50-1 contains over 300 mutations. To define mutations that made larger contributions to the phenotype, we focused on mutations meeting the following criteria: i) the mutations were non-synonymous; ii) mutations occurred in a gene that was mutated in at least two additional populations; iii) mutations in the gene achieved at least 10% abundance at some point during selection in each population in which they appeared; and iv) the mutation remained in the population at round 50 of selection. These criteria narrowed the search to mutations appearing in six genes: these included genes involved in recombinational DNA repair (*dinI*, *recD*, and *recN*), transcription (*rpoB*), regulation of anaerobic metabolism (*arcB*), and copper ion transport (*copA*). Two additional genes that did not meet the criteria in the current study were also included: *nth* and *recA*. Populations IR10-50 and IR12-20 have the same Nth variant (C203Y); this is the only example of the exact same mutation being present in multiple populations at round 50 aside from the loss of the e14 prophage. The *recA* gene was included as the RecA A290S variant, which is fixed in population IR10-50 and also appears in the IR11 lineage, is identical to a variant shown to contribute to IR resistance in the earlier study [13]. The relevant variants are listed in Table 2.3.

In IR9-50, mutations appear in each of these genes except for *arcB*, *nth*, and *recA*. However, a mutation in *arcB*, leading to the variant gene product ArcB N405D, did appear earlier in the IR9 lineage but was driven to extinction via clonal interference. As such, five protein variants present in IR9-50-1 were investigated for contributions to IR resistance: RecD A90E, RecN K429Q, DinI R28H, RpoB S72N, and CopA V270F. The ArcB N405D mutation was also examined as it represented a mutational pattern in other populations. A mutation in *rpoC* (which codes for the K1172I variant) was also included in the strains with RpoB S72N, as the genomic proximity of *rpoC* to *rpoB* did not allow for ready separation of these two mutations with the methods used here.

We moved these mutations singly and in combination into a derivative of the parent *E. coli* strain lacking the e14 prophage (as this prophage appears to excise from the chromosome after IR exposure and each subsequent mutation occurred in this genetic background [13, 90]. This strain, Founder Δe14 [13, 90], thus provided the genetic background to determine the contribution of each mutation, if any, to IR resistance. We determined that the RecD A90E and RpoB S72N variants each provided a significant increase in IR resistance. A combination of these two alleles increased IR resistance of the parent strain beyond either individual mutation, but not yet to the level of IR9-50-1, indicating that they likely account for most of IR resistance in population IR9-50 (Figure 12.3). The effects of these two alleles were confirmed by converting each of them singly and in combinations to the wild-type allele in the IR9-50 genetic background. When these three mutations are eliminated from a derivative of IR9-50-1 that retains the other 308 mutations, IR resistance was reduced to the levels seen in the Founder Δe14 strain.

The RecD A90E mutation appears to be at least a partial loss of function, as deleting the *recD* gene from Founder Δe14 increases IR resistance just as much as the RecD A90E variant.

These data agree with the fact that a truncated, and likely inactive, RecD variant appeared earlier in the IR9 lineage and in the IR10 lineage (Table 2.3). The DinI, RecN, CopA, and ArcB variants did not contribute significantly to the phenotype.

If we follow the appearance of mutations in genes that demonstrably contribute to IR resistance, a new pattern emerges (Figure 13.3). In the lineages IR9, IR11, and IR12 there is a conserved temporal order in the appearance of mutations. First, the e14 prophage is lost, followed by an *rpoB* mutation and then a *recD* mutation. In the IR12 lineage, the primary *rpoB* mutation is a synonymous coding mutation. However, this does not negate a possible effect of this mutation, potentially at the level of codon usage (RpoB D1203, GAC (WT) to GAT (mutant): 0.37 to 0.63 frequency). The trend is clear in linages IR9 and IR11, where all the mutations are non-synonymous. IR10 is the only lineage that does not conform to this trend. Although the e14 prophage did excise in lineage IR10, the sub-population that lost the prophage was outcompeted by another that maintained the prophage and contained a variant of both RecD (N124D) and RecA (A290S). After the appearance of the RecA A290S and RecD N124D variants, no RpoB variants appeared in the IR10 lineage until nearly round 40 of selection, and those variants (F15L and K1200E) have yet to reach 50% frequency in IR10. It appears that the evolutionary path observed in IR10 nearly occurred in population IR12 as well, where the sub-population that lost the e14 prophage was almost outcompeted by a different lineage (the competing sub-population again featuring RecD (T568A) and RecA variants (E19K)), before being carried to fixation with the apparent help of a different RecD variant (S92I). A sub-population with a RecD (W534R) and RecA variant (E286G) also reached high frequency in IR11 but was slowly outcompeted by a subpopulation with a different RecD variant (A550E) after it gained a second RpoB variant (P535L).

3.3.9 Prominent mutations can contribute to enhanced growth phenotypes

If variants in ArcB, CopA, DinI, and RecN do not contribute to IR resistance, do they make some other contribution to fitness that explains their prominence? We have previously observed that high frequency mutations in experimentally evolved populations may contribute to enhanced growth rates, rather than IR resistance (See *Chapter II* of this thesis). Consistently, the ArcB N405D mutation significantly enhances the growth phenotype of Founder Δe14 in a growth competition assay relative to the same strain with a wild-type ArcB (Figure 14.3). The RecN N405D and DinI R28H mutations make minor contributions to enhanced growth; the CopA V270F mutation appears to have no effect. Mutations that enhance the growth phenotype of these populations are likely important, as evidenced by their inability to grow well in media aside from LB (Figure 6.3) and their inability to compete with Founder Δe14 (Figure 7.3). In effect, certain mutations may arise that compensate for growth defects conferred by mutations that contribute to IR resistance. As a proof of concept, Founder Δe14 with the RpoB S72N variant is outcompeted quickly by Founder Δe14 with a wild-type RpoB (Figure 14.3). These results indicate that in order to develop high levels of IR resistance, mutations which buffer against deleterious effects must also rise to prominence in these populations.

3.4 Discussion

To our knowledge, we have generated the most IR resistant *E. coli* ever produced in a laboratory. After 100 rounds of selection with high-energy electron beam IR, the dose required to kill 99% of each of these four replicate populations has increased by over four-fold to over 3000 Gy (Figure 4.3). Additionally, the dose of electron beam IR that these populations can withstand far exceeds that of previously evolved IR-resistant isolates [13, 90], and is beginning to approach the level of the highly radioresistant bacterium *Deinococcus radiodurans* (Figure 5.3). These populations have variable growth phenotypes (Figures 6.3 and 7.3) and resistance to other forms of DNA damage (Figure 8.3, 9.3), indicating that they are specialized to withstand high doses of IR at the expense of other cellular functions. Utilizing deep-sequencing technology, we have an unprecedented window to the process of experimental evolution of IR resistance (Figure 10.3). It is apparent that each population has experienced unique evolutionary paths to reach the same goal. Despite varied levels of clonal interference in each population, these populations do share mutated genes in common which allow us to investigate contributors to extreme IR resistance. We have focused on which mutations found resistance in an isolate from population IR9-50, IR9-50-1. We have determined that mutations in *recD*, *rpoB*, and *rpoC* increase IR resistance of the Founder Δe14 parent-derivative to the level of IR9-50-1. Variants of these proteins have not been observed to increase IR resistance in previously evolved IR-resistant *E. coli* [13, 90].

Populations IR9-50, IR10-50, IR11-50, and IR12-50 provide us with the best platform yet for understanding the molecular fundamentals of extreme IR-resistance. These populations have been evolved using a source of IR that delivers dose at a much higher rate than that used previously to evolve IR resistance in *E. coli* (72 Gy/min versus 19 Gy/min) [13]. In addition, during each selection trial cells are irradiated in the absence of rich medium (which would provide cells with

nutrients and at least partial shielding from ROS generated by IR) and at 4° C. In this way, we prevent cells from responding to the damage caused by IR until they allowed to recover in rich medium post-irradiation (Figure 3.2). Following this selection protocol, modified from that used previously, these new populations have rapidly acquired resistance to electron beam IR. Isolates from the prior evolved populations, CB2000 and CB3000 [13, 90] are only moderately more resistant to this form of IR compared to the parent strain, indicating that the new populations are far more IR resistant than any evolved *E. coli* that we have ever generated.

At round 50 of selection, population IR9-50 appears to be radioresistant primarily through variants of RecD and RpoB/RpoC (Figure 13.3). The RecD A90E variant is likely a loss of function mutation, as a RecD deletion in the Founder Δe14 background increases IR resistance just as much as the A90E variant in this background. Previous studies have shown that a RecD inactivation can lead to a hyper-recombination phenotype [132, 133]. Increased homologous recombination could be of significant use to repair highly fragmented DNA post-IR (using an estimate for DNA DSBs produced by IR determined by Daly and colleagues [6], there would be approximately 15 DSBs at 1000 Gy). RpoB and RpoC are the beta and beta-prime subunits of RNA polymerase, respectively. Stringent mutations of RNAP, which mimic the effects of ppGpp binding and therefore the stringent response, are located primarily in RpoB and RpoC [134]. Some of these stringent mutants are capable of rescuing UV-sensitivity of *ruvABC* mutants of *E. coli*, potentially due to decreased stability of contacts of RNAP with DNA [135, 136]. Similar to these previous observations, the mutations in RpoB that are prevalent in our evolved populations do not locate to a single region of RpoB and therefore may affect DNA interaction throughout the DNA channel formed by RpoB/RpoC. Previously isolated stringent mutations affect residues near those affected in these evolved populations (such as the previously described L571Q and H1244Q mutants [135, 136]).

These mutations likely decrease stability of RNAP on DNA and may allow for easier removal of RNAP that has stalled at a DNA lesion. Removal of stalled RNAP may be crucial for efficient DNA repair due to RNAP occluding the lesion from repair machinery or providing a major obstacle to DNA replication [137]. Additionally, many stringent mutants of RNAP also confer resistance to the antibiotic Rifampicin. Mutated variants of RNAP at P535 [138] and S574 [139-141] have previously been isolated using selection for Rifampicin resistance. Recent work has implicated increasing backtracked RNAP in enhanced repair of DNA DSBs, and that the H1244Q variant of RpoB enhances DNA DSB repair [133].

Importantly, in this new undertaking we are documenting experimental evolution of IR resistance in an unprecedented manner through deep-sequencing technology. At every even round of selection, we have sequenced each mixed population to determine the full breadth of mutations down to 2% frequency in those populations. Put together, these data reveal the underlying competition of lineages within each population. While much of this study focuses on round 50 of selection, observing evolution over the remaining 49 cycles allows us to understand the genetic context in which new mutations arise and how the successful lineages were able to outcompete other fit, but ultimately unsuccessful, lineages.

Of particular interest is the apparent competition between lineages which contain alleles of *recD* and *rpoB*, and those with *recD* and *recA* alleles. The four replicate populations offer a view on four unique scenarios of competition between these two lineages. In IR9, no *recA* variant has yet appeared, and the population successfully lost the e14 prophage and then gained a *rpoB* followed by a *recD* allele over these 60 cycles of selection. In IR10, the RecA/RecD variant pair appears almost immediately, and outcompetes the lineage that lost the e14 prophage. After this event, the only of variant to fix within the population (of those which we are focusing on) is another

RecD variant. RpoB variants do not appear until at least round 40 of selection in IR10, and by round 50, these variants do not appear to be on a path to fixation. No *copA* allele has appeared in IR10. In the IR11 lineage, loss of the e14 prophage and subsequent gain of an *rpoB* allele fixed. In this genetic background, two separate sub-populations containing a RecD and RecA/RecD variants, respectively, reached prominence. The RecA/RecD sub-population appeared to be approaching fixation, but the RecD-only subpopulation gained a variant of RpoB, which was sufficient to drive this sub-population to fixation and the RecA/RecD to extinction. Finally, IR12 almost began on a similar evolutionary trajectory as IR10. A sub-population that lost the e14 prophage and gained an *rpoB* allele (which leads to a synonymous coding mutation) was nearly outcompeted by a RecA/RecD variant sub-population, before a RecD variant rescued the former sub-population. By round 40 of selection, the winning subpopulation reached fixation.

All together, these results suggest that over 50 cycles of selection there are two prominent paths to IR resistance (and likely others not discovered here). Time will tell if these two pathways will mix or continue to diverge over further rounds of selection. Considering that IR10-50 appears to be following a path to IR resistance that was outcompeted in the other three populations, yet is the second-most IR resistant population, it is likely that other mechanisms of IR resistance not yet documented exist in this population. The isolate IR10-50-1, isolated from IR10-50, is also the only isolate tested capable of growing in minimal medium (Figure 6.3C), indicating that this population may be 'healthier' than the others. Over the course of these 50 rounds of selection, lineage IR10 appears to be providing an opportunity to observe a completely different path to IR resistance than the other three populations.

There is no apparent explanation for the differences in the mutations which appear in these populations and those that were previously evolved [13, 90]. Surprisingly, the frequency of

transversion mutations detected in these populations (~50% of all SNPs) is far lower than previous studies reported in *E. coli* exposed to gamma-ray IR (~80% of all SNPs) [12, 13]. GC to TA transverions and GC to AT transitions dominate, which is consistent with previous sequencing results. However, AT to GC transitions are detected at high frequencies in these new populations. This transition, as well as -1 deletions (which are the most prominent source of frame-shift mutations in these populations), are consistent with incorporation of 2-Hydroxy-2'-deoxyadenosine (2-OH-dA). As this oxidation product of adenine is incorporated rather than generated from paired adenine [8], it may be that irradiation from the Linac more efficiently oxidizes the nucleotide pool than ^{60}Co irradiation. In addition, only a single mutation previously noted to enhance IR resistance, RecA A290S[13, 90], appears in these four populations. There may be a fundamental difference in how a high energy electron beam interacts with bacterial cells as opposed to high-energy photons, leading to a different spectrum of mutations. Alternatively, the higher dose rate used in this study may lead to differences in the damage caused, or the altered selection protocol changes how the cells repair IR-induced damage, leading to a different mutation spectrum. In any case, these new *E. coli* lineages are abundant with novel mechanisms of experimentally evolved IR-resistance.

We have not yet achieved experimentally evolved *E. coli* that reach the natural IR resistance of the bacterium *D. radiodurans.* However, these populations have taken great leaps towards this phenotype, and have given no indication that the full level of IR resistance in *E. coli* has been reached. As such, we are continuing cycles of selection to continue to push the limits of what these incredible organisms can endure.

3.5 Tables

	Parent bp	Mutant bp	IR9-50	IR10-50	IR11-50	IR12-50
Transitions	AT	GC	93	106	94	110
	GC	AT	180	158	130	223
Transversions	AT	CG	26	17	11	25
	GC	TA	154	116	84	143
	AT	TA	81	58	59	85
	GC	CG	41	25	31	51
	Transitions		273	264	224	333
	Transversions		302	216	185	304
	Total		575	480	409	637
Coding	Synonymous		127	124	98	171
	Non-synonymous		340	275	237	363
	dN/dS		2.67717	2.21774	2.41837	2.12281
	Stop gained		29	15	12	16
	Stop lost		1	2	0	1
	Start lost		0	0	1	2
	Insertions	+1	1	2	2	3
		+2	0	0	0	0
		+3	0	0	1	0
		+12	0	0	1	0
	Deletions	-1	40	26	21	26
		-2	0	1	1	2
		-3	1	0	0	0
Non-coding	SNPs		81	70	65	92
	Insertions	+1	0	0	0	0
		+2	0	0	0	1
		+3	0	0	0	0
	Deletions	-1	3	6	4	6
		-2	0	0	0	1
		-3	0	0	0	0
Allele frequencies	Fixed (>99%)		107	84	20	76
	>50%		242	109	130	150
	>10 %		309	199	205	320
	>2%		620	515	439	676

Table 1.3. Mutation totals in populations IR9-50, IR10-50, IR11-50, and IR12-50.

Pathway	Protein	IR9	IR10	IR11	IR12
Anaerobic respiration	ArcB	N405D	D166Y S280R L90 fs Y71C	R100H **S74W**	**S24P** K547 fs
Copper metabolism	CopA	V270F		T525A	A812V
DNA metabolism	DinI	**R28H** P14A	**W71***	I66N *82Q	S26Y
	Nth		C203Y E160K		C203Y K85N
	RecA	Y294C	**A290S**	A290S E286G	E19K
	RecD	**A90E** L188* L223* P99*	**N124D** **Q463*** L301M	A550E A39G C71 fs G307W W534R	**S92I** A271E C103* T568A
	RecJ	*578L R254C	G502D I332V R273S T15A	Q337R	F426L
	RecN	**K429Q** E271G R118L	F144L R415L	R102P E320K S382R	**A361T** R285C R368H S310L
RNA polymerase	RpoB	**S72N** S391P	F15L K1200E	P535L **S574F**	**T600I**

Table 2.3. Prevalent mutations in evolved populations at 60 rounds of selection. Mutations listed in black text are present above 2% frequency at round 60 of selection. Mutations listed in bold text are fixed in the population. Mutations listed in gray were in a sub-population that has gone extinct by round 60 of selection. A '*' symbol indicates a stop codon, and 'fs' denotes a frame-shift.

Strain	Relevant Genotype	Source
MG1655	YbhJ L54I + MntP G25D + RIP321 A 4296380* ACG + GlpR C 3560455* CG + GatC ACC 2173360* A	
EAW7704 (Founder Δe14)	MG1655 Δe14 + RbsR L92R + CytR Q110 stop + *fabI /ycjD* int (G 1351174* A) + *yifN /ppiC* int (T 3959934* C)	Harris 2009
IR9-50	MG1655 exposed to 50 iterative rounds of IR; mixed population	This study
IR10-50	MG1655 exposed to 50 iterative rounds of IR; mixed population	This study
IR11-50	MG1655 exposed to 50 iterative rounds of IR; mixed population	This study
IR12-50	MG1655 exposed to 50 iterative rounds of IR; mixed population	This study
IR9-100	MG1655 exposed to 100 iterative rounds of IR; mixed population	This study
IR10-100	MG1655 exposed to 100 iterative rounds of IR; mixed population	This study
IR11-100	MG1655 exposed to 100 iterative rounds of IR; mixed population	This study
IR12-100	MG1655 exposed to 100 iterative rounds of IR; mixed population	This study
IR9-50-1	Isolate from IR9-50	This study
IR9-50-2	Isolate from IR9-50	This study
IR9-50-3	Isolate from IR9-50	This study
IR9-50-4	Isolate from IR9-50	This study
IR9-50-5	Isolate from IR9-50	This study
IR10-50-1	Isolate from IR10-50	This study
IR10-50-2	Isolate from IR10-50	This study
IR10-50-3	Isolate from IR10-50	This study
IR10-50-4	Isolate from IR10-50	This study
IR10-50-5	Isolate from IR10-50	This study
IR11-50-1	Isolate from IR11-50	This study
IR11-50-2	Isolate from IR11-50	This study
IR11-50-3	Isolate from IR11-50	This study
IR11-50-4	Isolate from IR11-50	This study
IR11-50-5	Isolate from IR11-50	This study
IR12-50-1	Isolate from IR12-50	This study
IR12-50-2	Isolate from IR12-50	This study
IR12-50-3	Isolate from IR12-50	This study
IR12-50-4	Isolate from IR12-50	This study
IR12-50-5	Isolate from IR12-50	This study
CB2000	Isolate from IR-resistant evolved population IR-2-20	Harris 2009
CB3000	Isolate from IR-resistant evolved population IR-3-20	Harris 2009
EAW792	Founder Δe14 + ArcB N405D	This study
EAW971	Founder Δe14 + CopA V270F	This study
EAW766	Founder Δe14 + DinI R28H	This study
JDT56	Founder Δe14 + RecA A290S	This study
EAW726	Founder Δe14 + RecD A90E	This study
EAW848	Founder Δe14 + RecJ *578L	This study
EAW725	Founder Δe14 + RecN K429Q	This study
EAW756	Founder Δe14 + RpoB S72N/RpoC K1172I	This study
EAW748	Founder Δe14 + RecD A90E + RecN K428H	This study
EAW904	Founder Δe14 + RecD A90E + RpoB S72N/RpoC K1172I	This study
EAW781	Founder Δe14 + RecD A90E + RecN K428H + RpoB S72N/RpoC K1172I	This study
EAW980	Founder Δe14 + RecD A90E + RpoB S72N/RpoC K1172I + CopA V270F	This study
EAW972	IR9-50-1 + CopZ V270 wt	This study
STB75	Founder Δe14 Δ*araBAD*	This study
* nucleotide position compared to NCBI GenBank U00096.3 reference sequence		

Table 3.3. Strains used in this study.

3.6 Figures

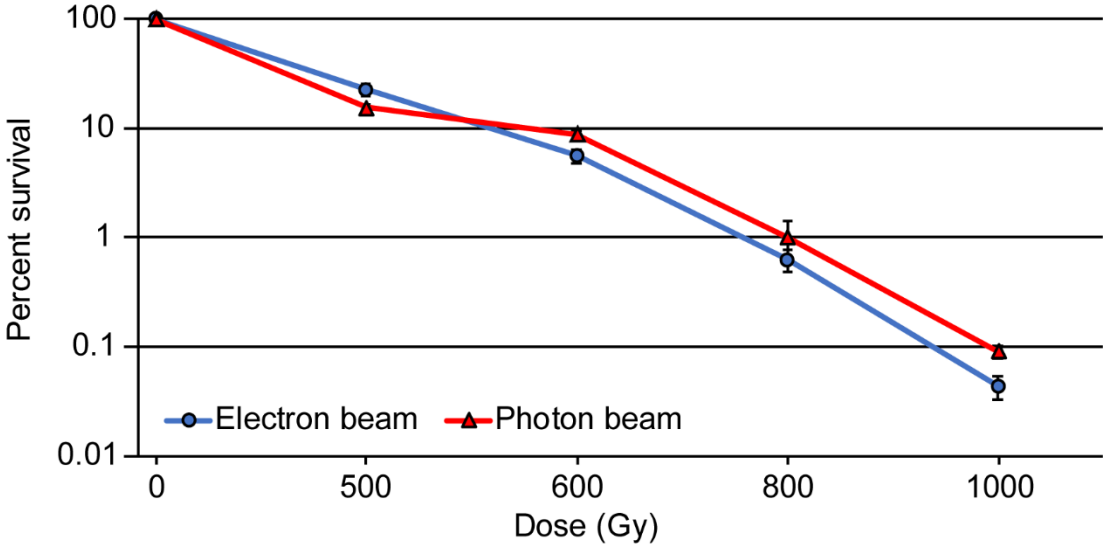


Figure 1.3. Cell killing of electron beam versus photon beam ionizing radiation. Early exponential phase cultures of MG1655 were exposed to electron beam or photon beam IR from the Linac as described in the *Materials and Methods* with a dose rate of 17 Gy/min. Percent survival was determined via CFU/mL counts pre- and post-irradiation.

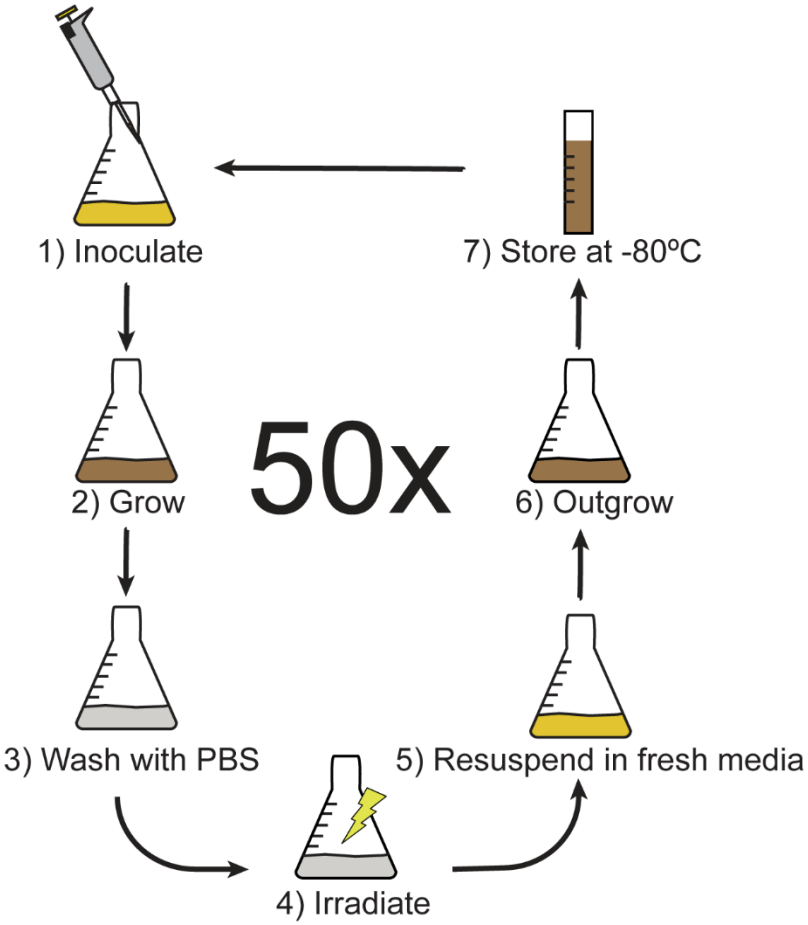


Figure 2.3. Directed evolution scheme. Briefly, overnight cultures were grown from a freezer stock of the parent strain/the evolved population from the previous round of selection. This overnight culture was used to inoculate fresh LB rich medium and cultures were grown to early exponential phase. These cultures were then separated into multiple aliquots and washed 3X in 1X PBS to remove LB medium. One aliquot of each population was irradiated with the same dose that killed 99% of the population in the previous round of selection, and the remaining aliquots were irradiated with higher doses. A portion of the irradiated aliquots were plated to determine percent survival, while the remaining culture was resuspended in fresh LB medium. These were then grown overnight, and the culture that was nearest to 1% survival (as determined by CFU counts from pre- and post-irradiation cultures) was stored at -80° C.

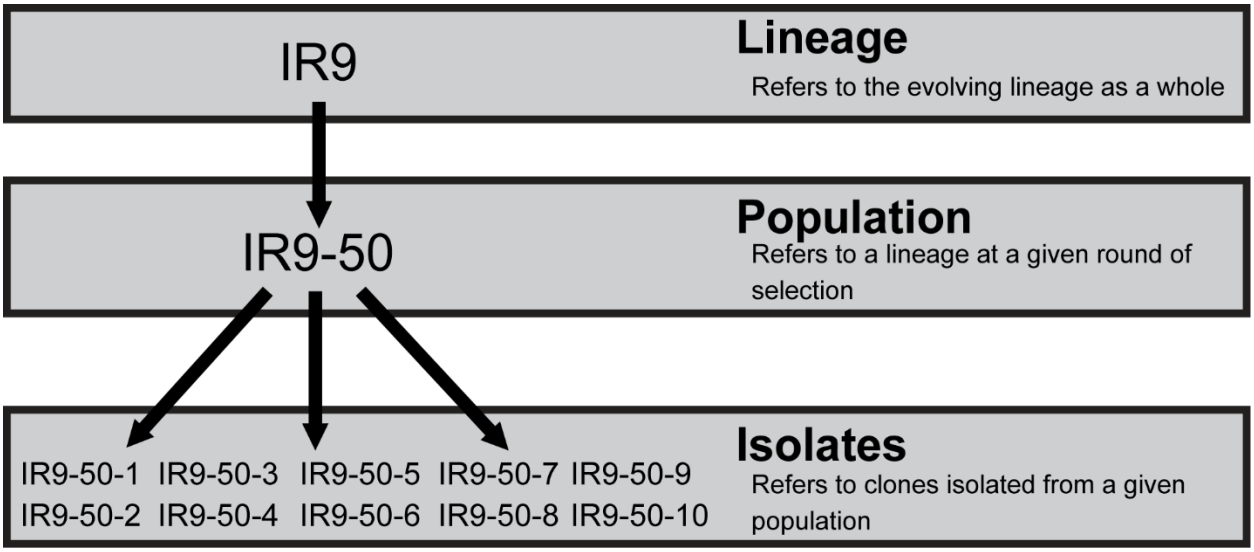


Figure 3.3. Nomenclature of evolved *E. coli* lineages. We have generated four lineages of highly IR-resistant *E. coli*, designated IR9, IR10, IR11, and IR12. A designated round of selection indicates a population at that round (Ex: IR9-50 is lineage IR9 after 50 rounds of selection). Each population has been stored at -80 °C as a 'fossil record' of evolution. Clonal isolates generated by streak plating a population then streak plating ten separate isolated colonies are designated by a numeral value added to the parent population nomenclature (Ex: The first isolate from IR9-50 is IR9-50-1).

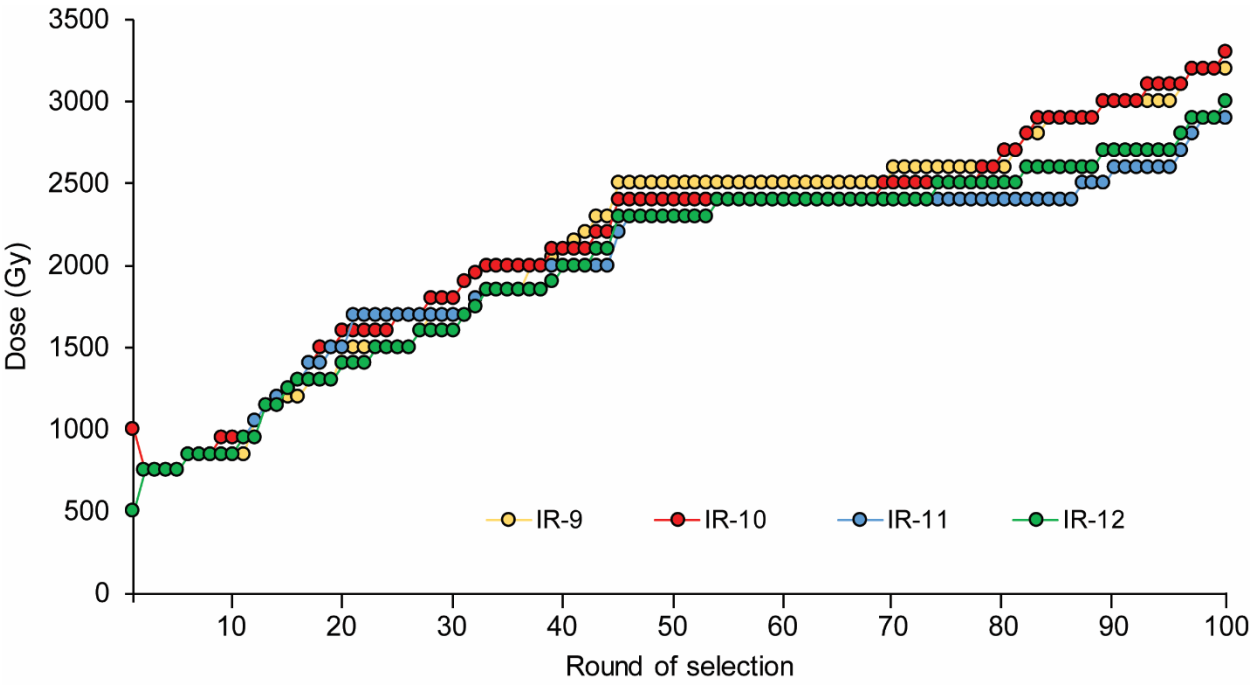


Figure 4.3. Dose required to kill 99% of each population has increased over 100 rounds of selection. Each data point indicates the dose of IR that each population was given prior to being outgrown overnight (Step 6 in Figure 2) and stored at -80° C (Step 7 in Figure 2.3).

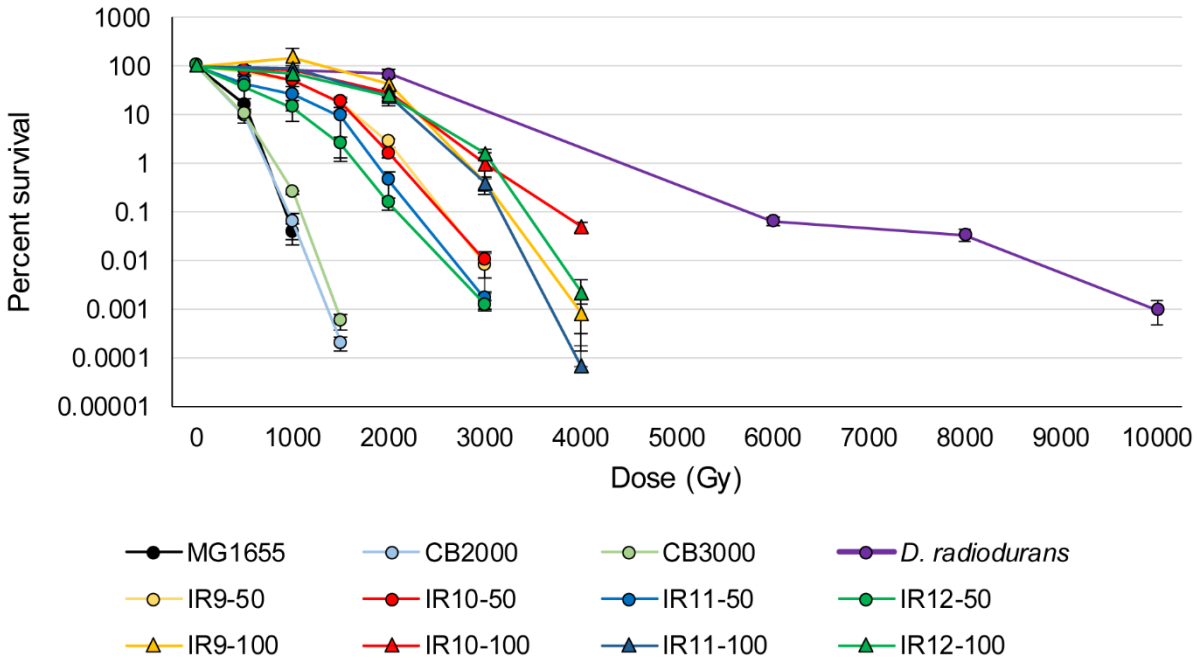


Figure 5.3. Survival curves of evolved populations compared to previously evolved *Escherichia coli* isolates and *Deinococcus radiodurans*. Early exponential phase cultures of the indicated strains were exposed to electron beam IR as described in the *Materials and Methods*. CB2000 and CB3000 are isolates from *E. coli* populations evolved to withstand γ-ray IR [13, 90]. Error bars represent the standard deviation of CFU/mL calculations of biological triplicate.

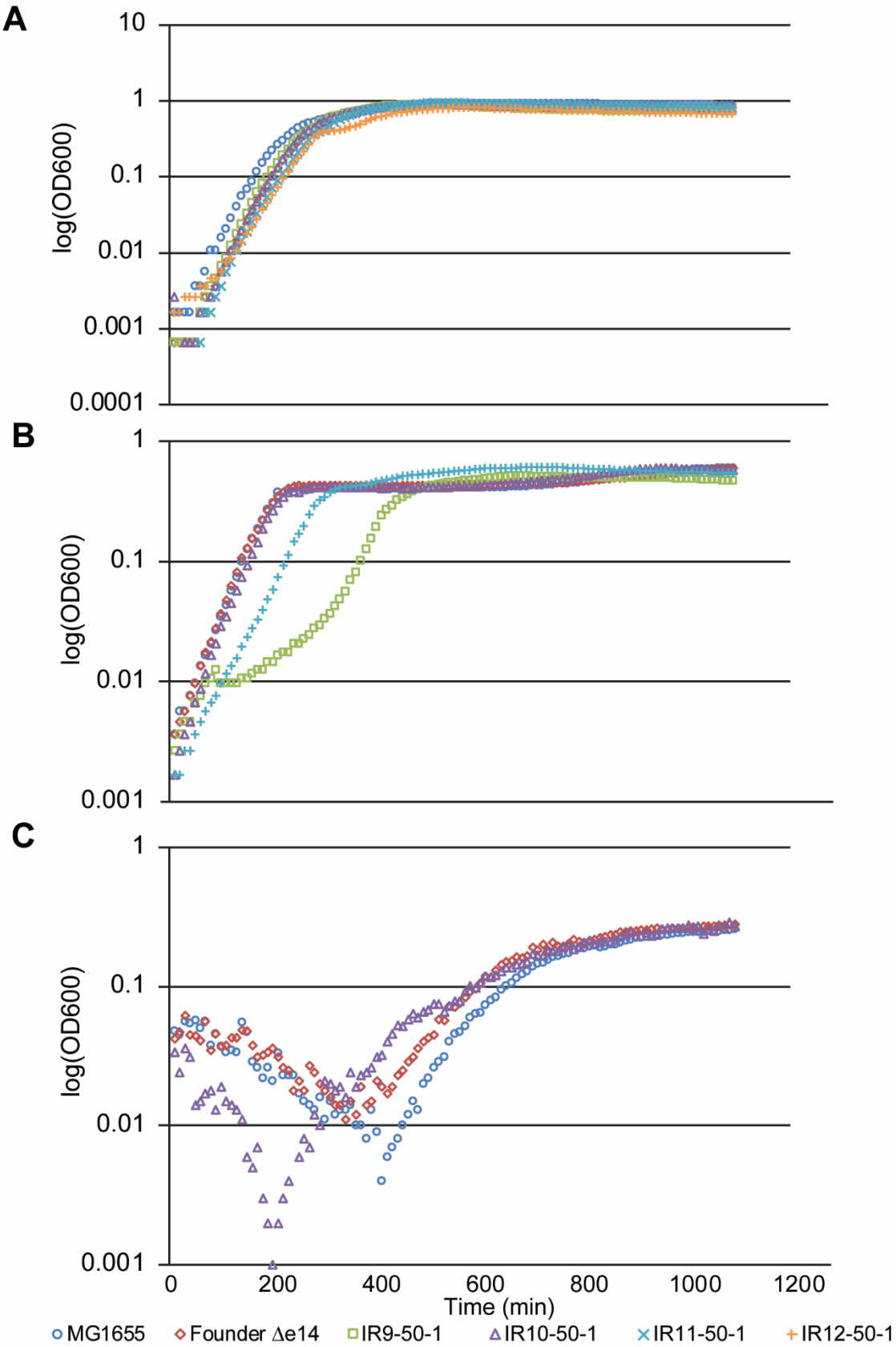

A
B
C
log(OD600)
10
1
0.1
0.01
0.001
0.0001
0
200
400
600
800
1000
1200
Time (min)
MG1655
Founder Δe14
IR9-50-1
IR10-50-1
IR11-50-1
IR12-50-1

Figure 6.3. Growth curves of evolved isolates from round 50 of selection in rich and minimal medium. A. Growth curves of isolates in LB rich medium. B. Growth curves of isolates in EZ defined rich medium supplemented with 0.2% glucose. C. Growth curves of isolates in M9 defined minimal medium supplemented with 0.2% glucose. Each growth curve is a representative replicate from an experiment performed in biological triplicate. Cultures of indicated strains were grown in the appropriate medium overnight and then to early exponential phase as described in the *Materials and Methods*. Early exponential phase cultures were diluted 1:100 in the appropriate medium, and then incubated overnight in a Biotek Synergy 2 plate reader, with OD_{600} measurements taken automatically every 5 minutes.

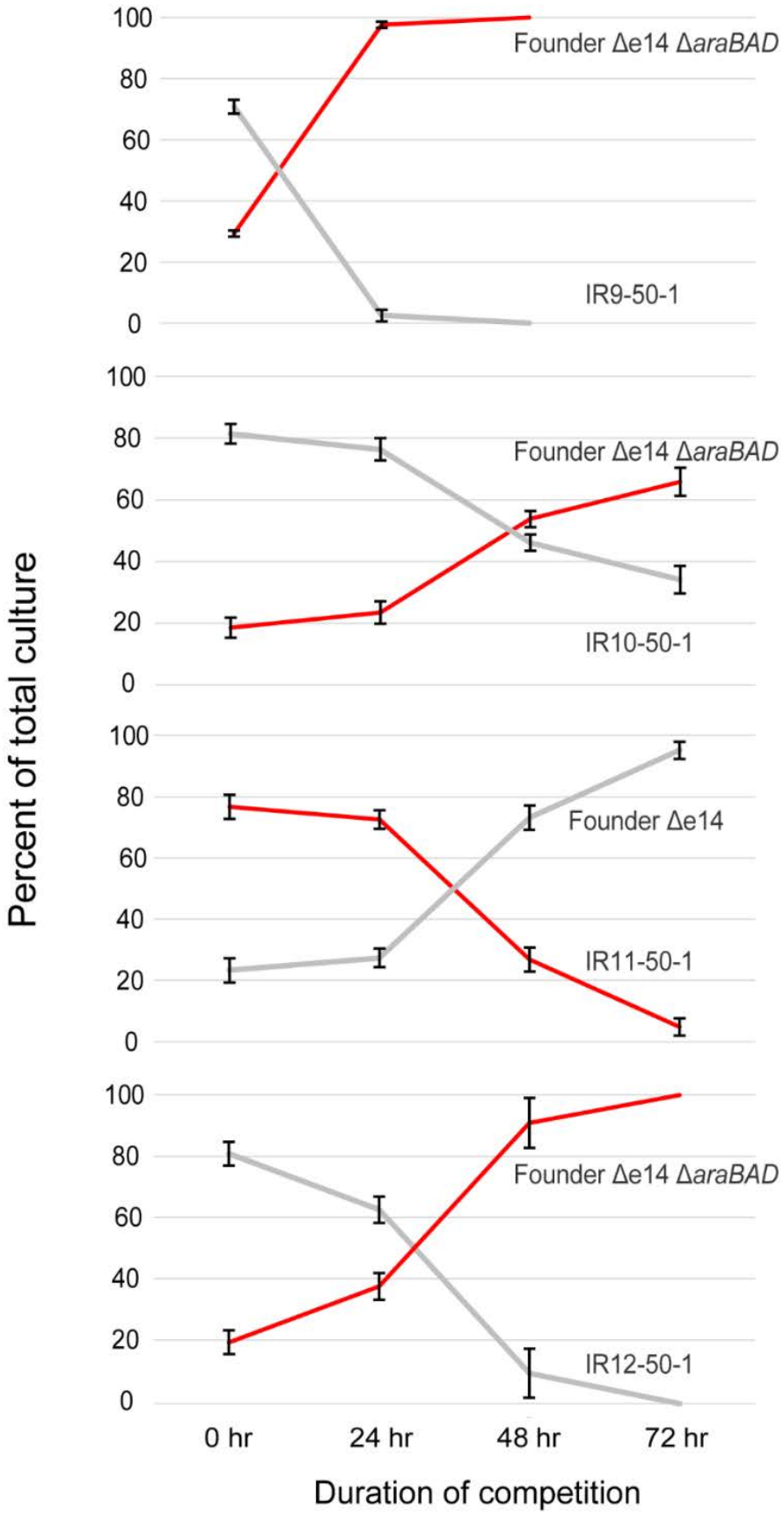

Percent of total culture
100
80
60
40
20
0
Founder Δe14 ΔaraBAD
IR9-50-1
Founder Δe14 ΔaraBAD
IR10-50-1
Founder Δe14
IR11-50-1
Founder Δe14 ΔaraBAD
IR12-50-1
0 hr
24 hr
48 hr
72 hr
Duration of competition

Figure 7.3. Growth competitions of evolved isolates from round 50 of selection. Growth competitions were performed as described in the *Materials and Methods*. Briefly, competitions were started with an excess of the evolved isolate so that competitions could be carried out to 72 hr. Even with an excess of IR9-50-1, this isolate was outcompeted by 48 hr. Deleting the *araBAD* operon causes a red colony phenotype with no fitness cost that allows for differentiation of the two strains in competition on TA medium. IR11-50-1 is red on TA medium without alteration of the *araBAD* operon, so this competition was performed against Founder Δe14 *araBAD*$^{+}$. Error bars represent the standard deviation of CFU/mL calculations of biological triplicate.

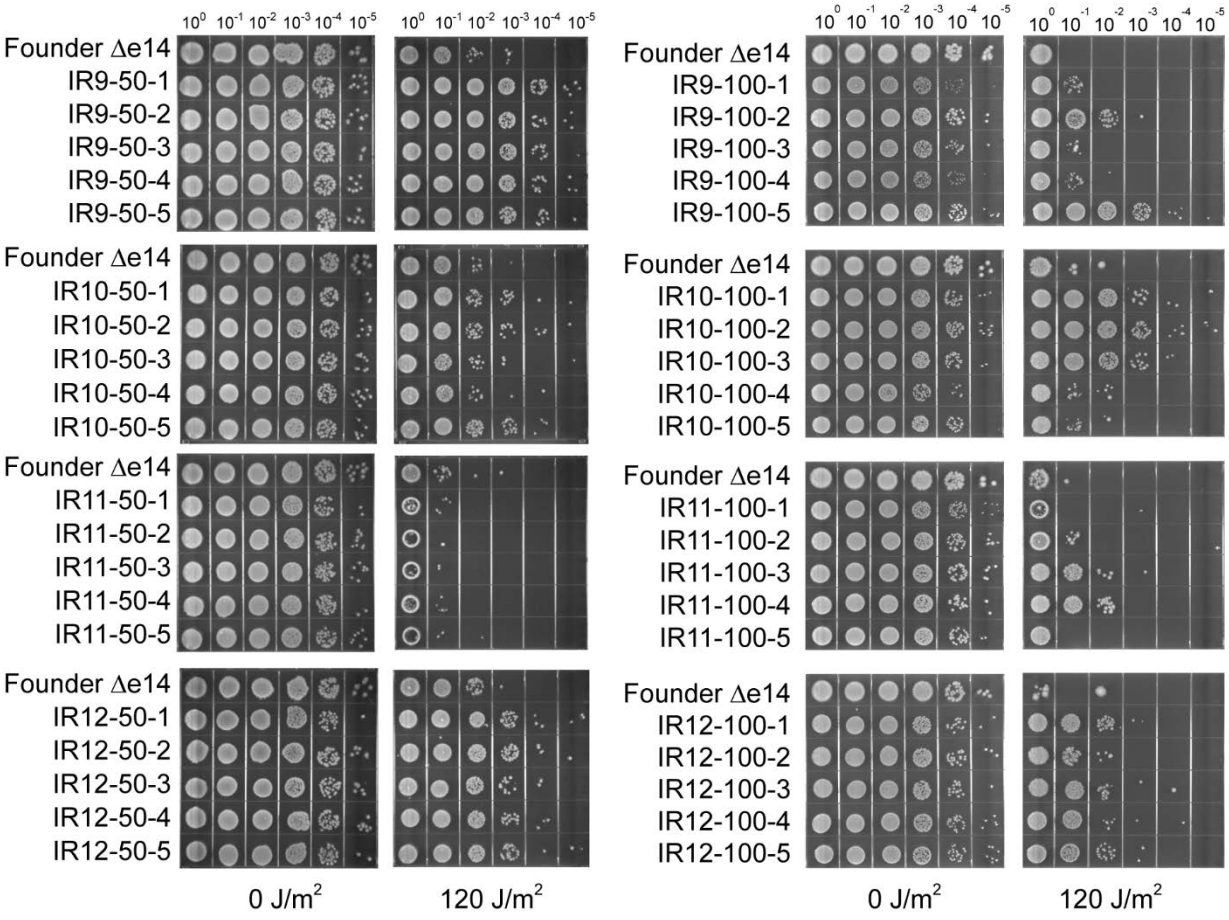


Figure 8.3. Evolved isolates exhibit variable survival to UV irradiation. Five isolates from each population were grown in LB medium overnight and then to early exponential phase as described in the *Materials and Methods*. Cultures were serial diluted in 1X PBS, and then 10 µL of each dilution was spot plated onto LB agar. Once spots dried, plates were exposed to 120 J/m^2 of UV irradiation and were then incubated overnight before imaging.

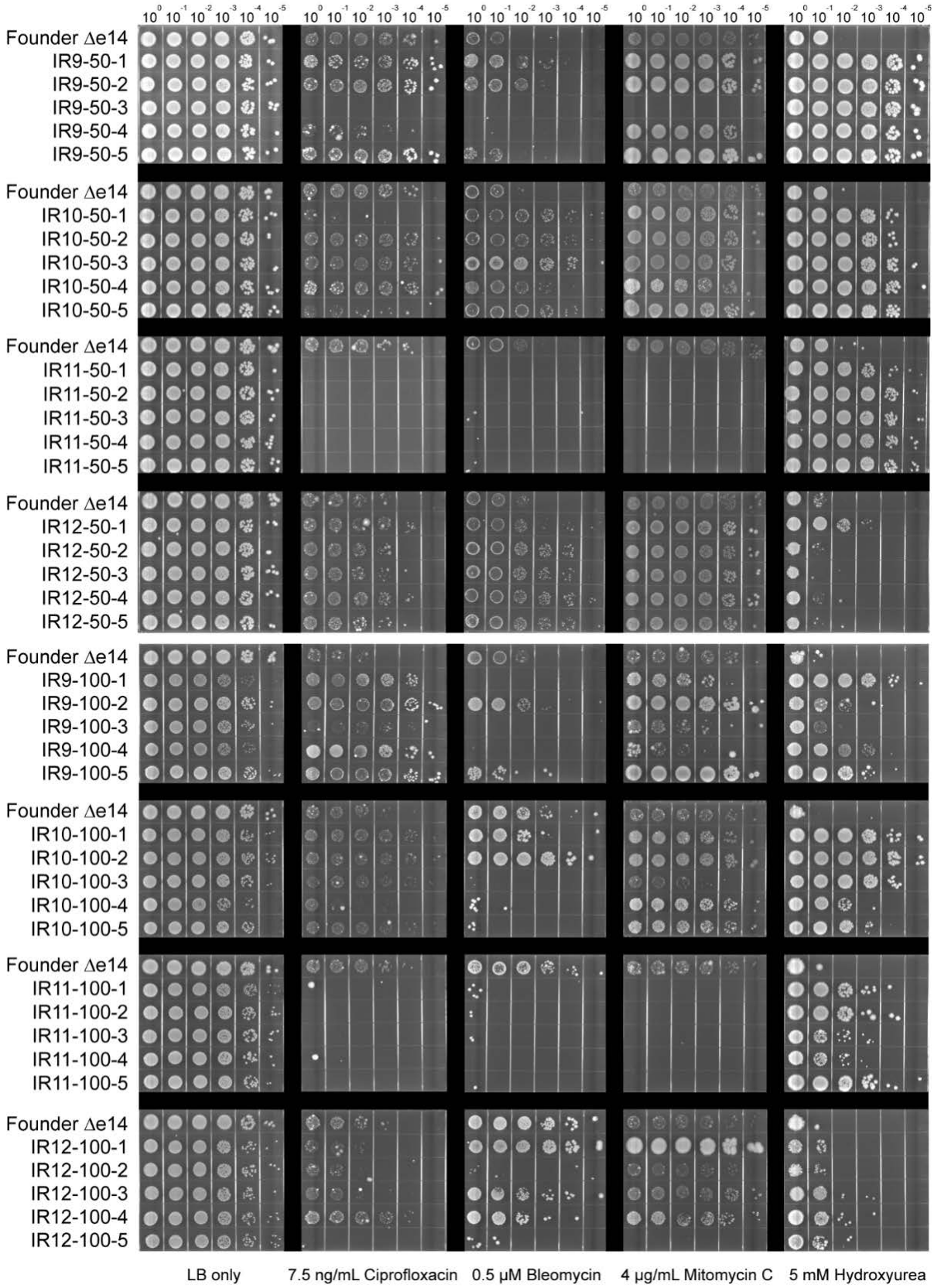

10^{0} 10^{-1} 10^{-2} 10^{-3} 10^{-4} 10^{-5}
Founder Δe14
IR9-50-1
IR9-50-2
IR9-50-3
IR9-50-4
IR9-50-5
Founder Δe14
IR10-50-1
IR10-50-2
IR10-50-3
IR10-50-4
IR10-50-5
Founder Δe14
IR11-50-1
IR11-50-2
IR11-50-3
IR11-50-4
IR11-50-5
Founder Δe14
IR12-50-1
IR12-50-2
IR12-50-3
IR12-50-4
IR12-50-5
Founder Δe14
IR9-100-1
IR9-100-2
IR9-100-3
IR9-100-4
IR9-100-5
Founder Δe14
IR10-100-1
IR10-100-2
IR10-100-3
IR10-100-4
IR10-100-5
Founder Δe14
IR11-100-1
IR11-100-2
IR11-100-3
IR11-100-4
IR11-100-5
Founder Δe14
IR12-100-1
IR12-100-2
IR12-100-3
IR12-100-4
IR12-100-5
LB only
7.5 ng/mL Ciprofloxacin
0.5 μM Bleomycin
4 μg/mL Mitomycin C
5 mM Hydroxyurea

Figure 9.3. Evolved isolates exhibit variable survival of various DNA damaging agents. Five isolates from each population were grown in LB medium overnight and then to early exponential phase as described in the *Materials and Methods*. Cultures were serial diluted in 1X PBS, and then 10 µL of each dilution was spot plated onto LB agar with the indicated DNA-damaging agent. Once spots dried the plates were then incubated overnight before imaging.

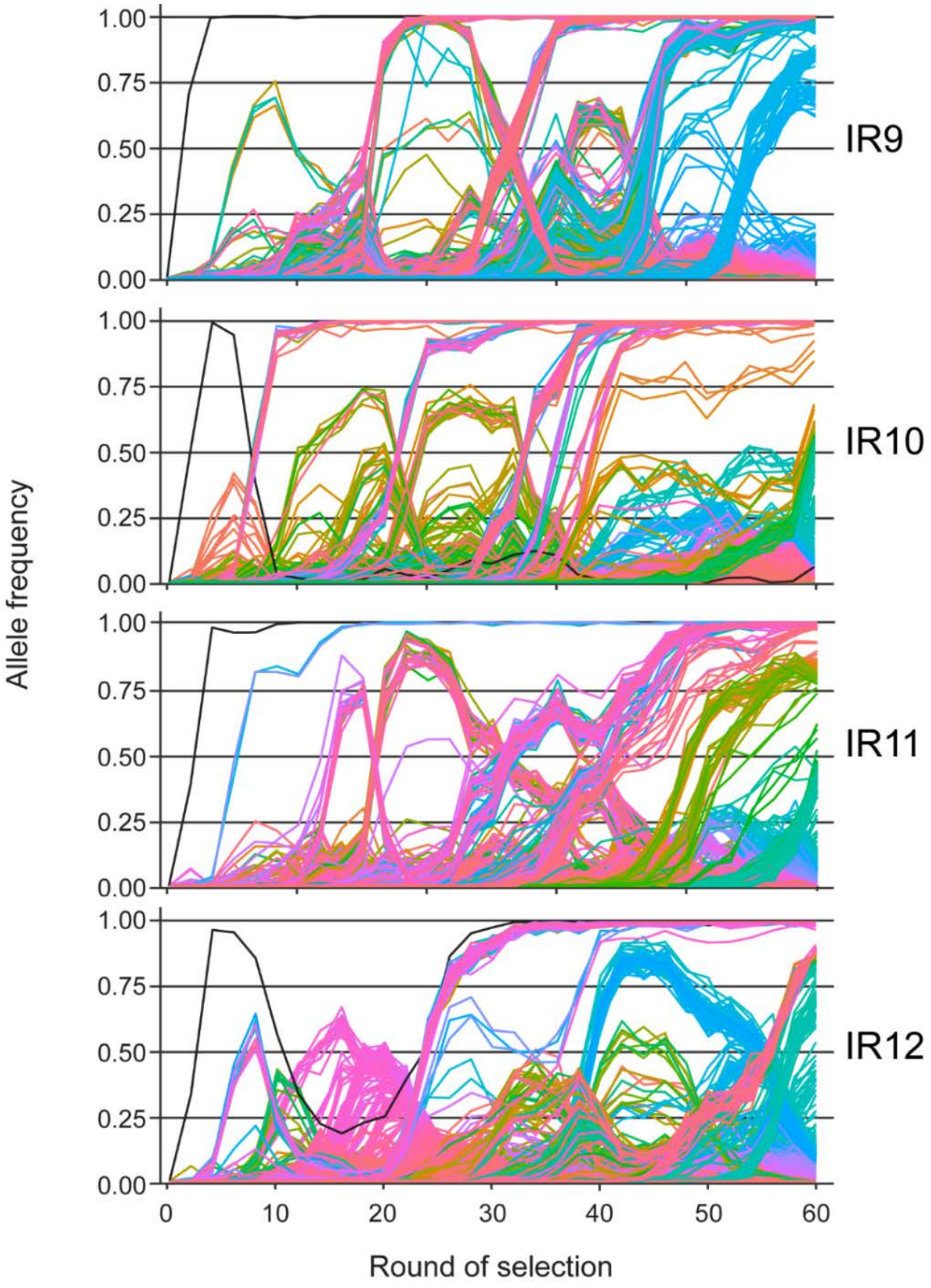

Allele frequency
1.00
0.75
0.50
0.25
0.00
IR9
IR10
IR11
IR12
0
10
20
30
40
50
60
Round of selection

Figure 10.3. Mutation frequencies over 60 rounds of selection in evolving populations. Whole populations at each even round of selection were deep-sequenced by the Joint Genome Institute (JGI) as described in the *Materials and Methods*. The frequency throughout these 60 rounds of selection of each mutation that reached at least 2% frequency in a single population are depicted as single lines. The sole black line in each graph represents the loss of the e14 prophage. Mutations that reach a frequency of "1.00" are fixed, and each subsequent mutation occurs within that genetic background. The graphs were generated using the R library "ggplot2" and mutations are not colored in any particular order.

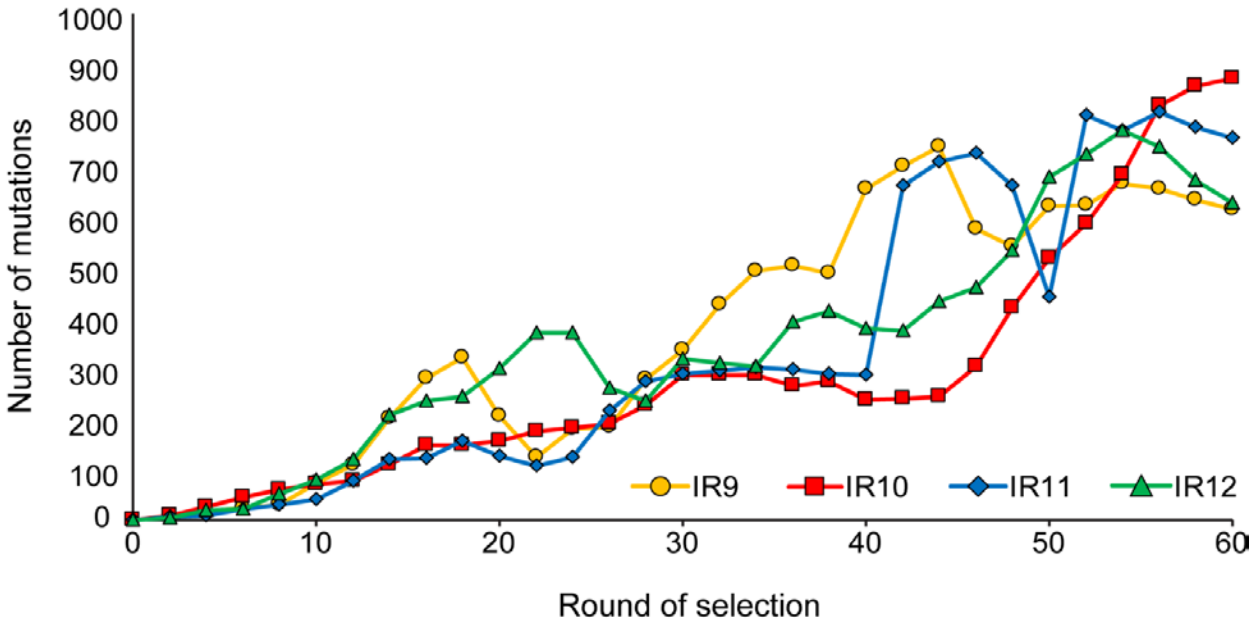


Figure 11.3. Number of mutations over 60 rounds of selection in evolving populations. Each data point represents the total number of detected mutations at or above 2% frequency in each sequenced whole population. The number of mutations generally increases but sometimes dips as major sub-populations are driven to extinction.

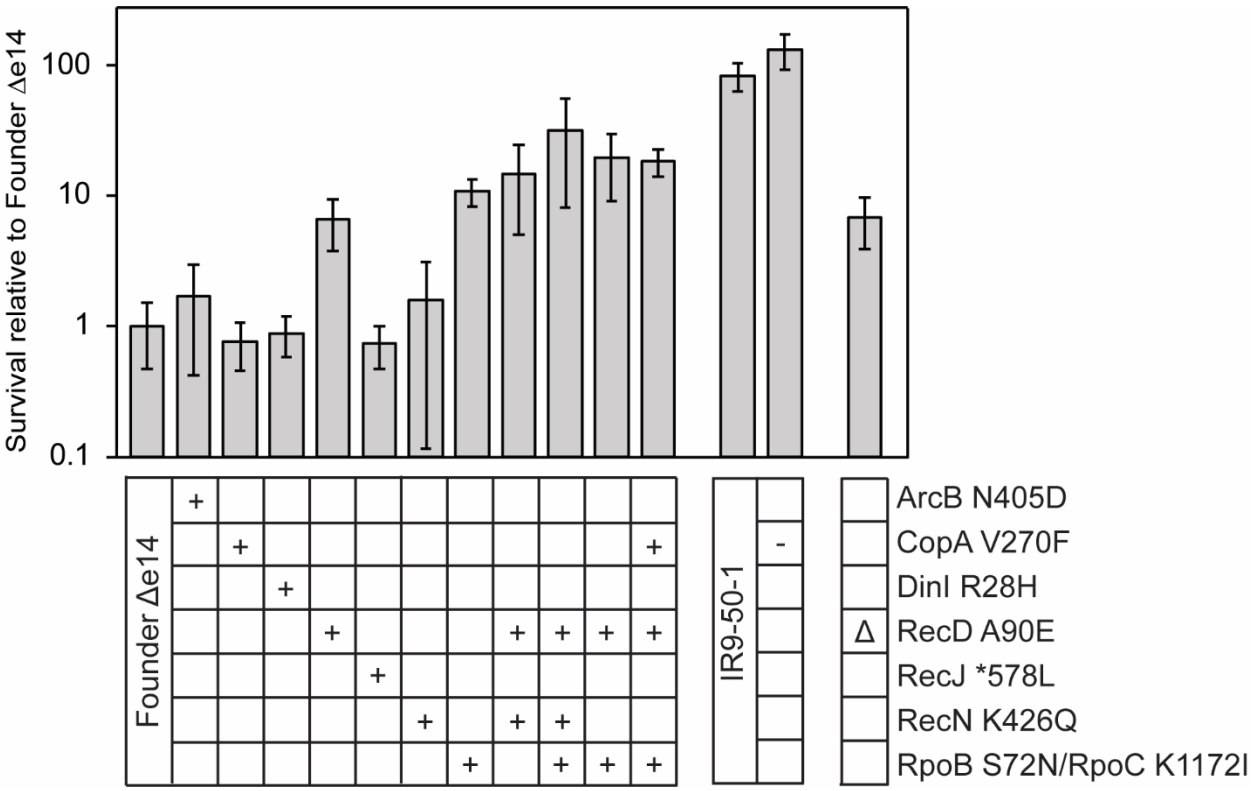


Figure 12.3. Mutations tested for contribution to IR resistance in IR9-50-1. All strains were dosed with 1000 Gy of electron beam IR and percent survival was determined via calculating CFU/mL before and after irradiation. Mutations were moved from IR9-50-1 to the Founder Δe14 genetic background singly and in combination as described in the *Materials and Methods*. Mutations were transferred from IR9-50-1 to Founder Δe14 as described in the *Materials and Methods*. The 'Δ' symbol represents a full deletion of the indicated gene. Error bars represent the standard deviation of CFU/mL calculations of at least two experiments performed in biological triplicate.

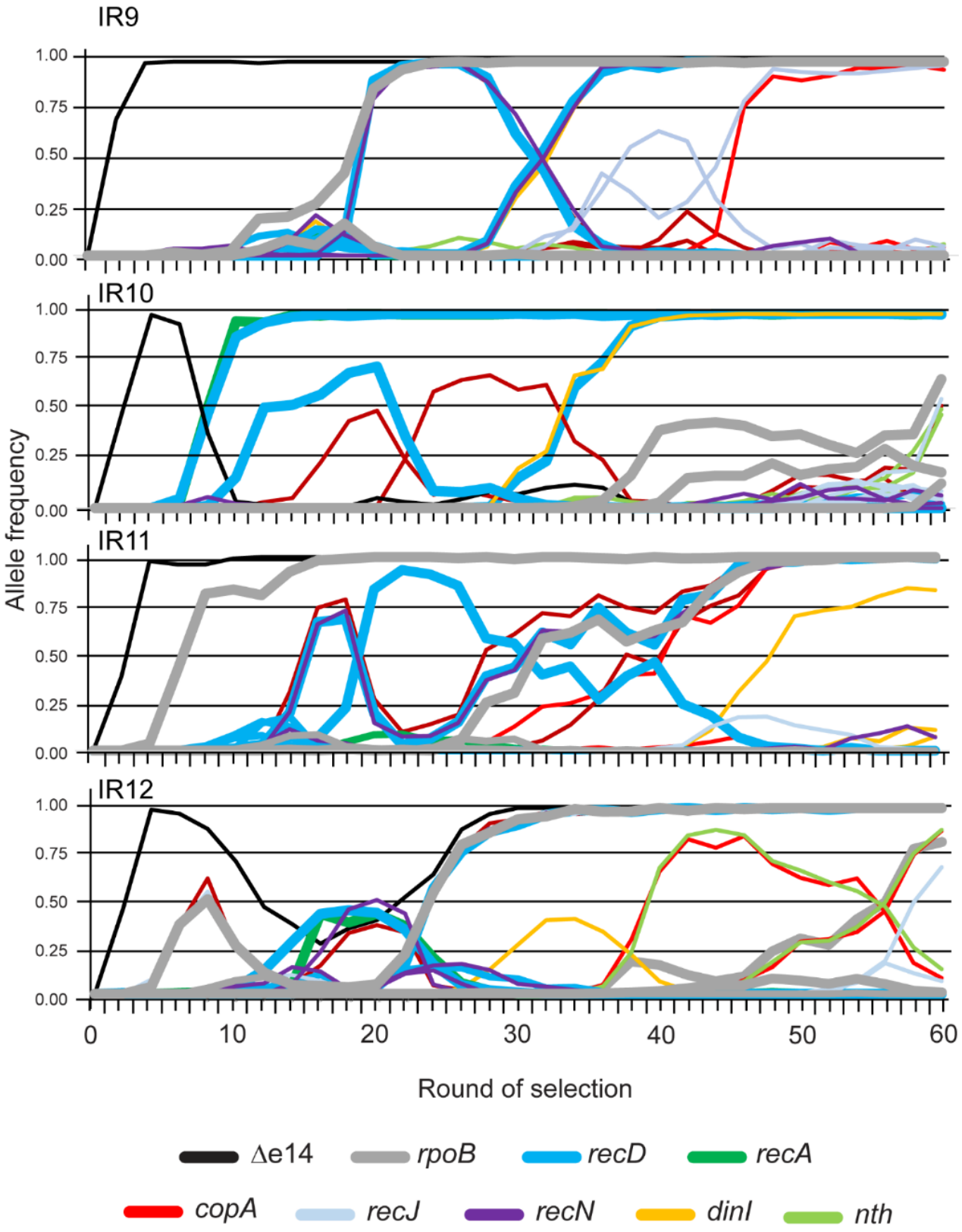

IR9
IR10
IR11
IR12
Allele frequency
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
60
Round of selection
Δe14
rpoB
recD
recA
copA
recJ
recN
dinI
nth

Figure 13.3. Frequencies of mutations implicated in IR resistance over rounds of selection. Frequencies of mutations in genes implicated in evolved IR resistance (*rpoB, recD, copA* [This study]; *recA* [13] and the loss of the e14 prophage are depicted. Both synonymous and non-synonymous mutations are included. Thick lines indicate the allele frequency of mutations in genes in which another mutation has been verified to enhance IR resistance. Thin lines indicate allele frequencies of mutant genes where a mutation in that gene appears in at least three populations but does not appear to enhance IR resistance. These data seemingly indicate two separate paths to acquiring IR resistance: loss of the e14 prophage, gain of an *rpoB* then *recD* in that order (IR9, IR11, IR12), or gain of a *recA* and *recD* mutation in concert (IR10). In lineages IR11 and IR12, these two pathways to IR resistance appear in direct competition, with the *rpoB*, and *recD* path being the apparent ‘winner’.

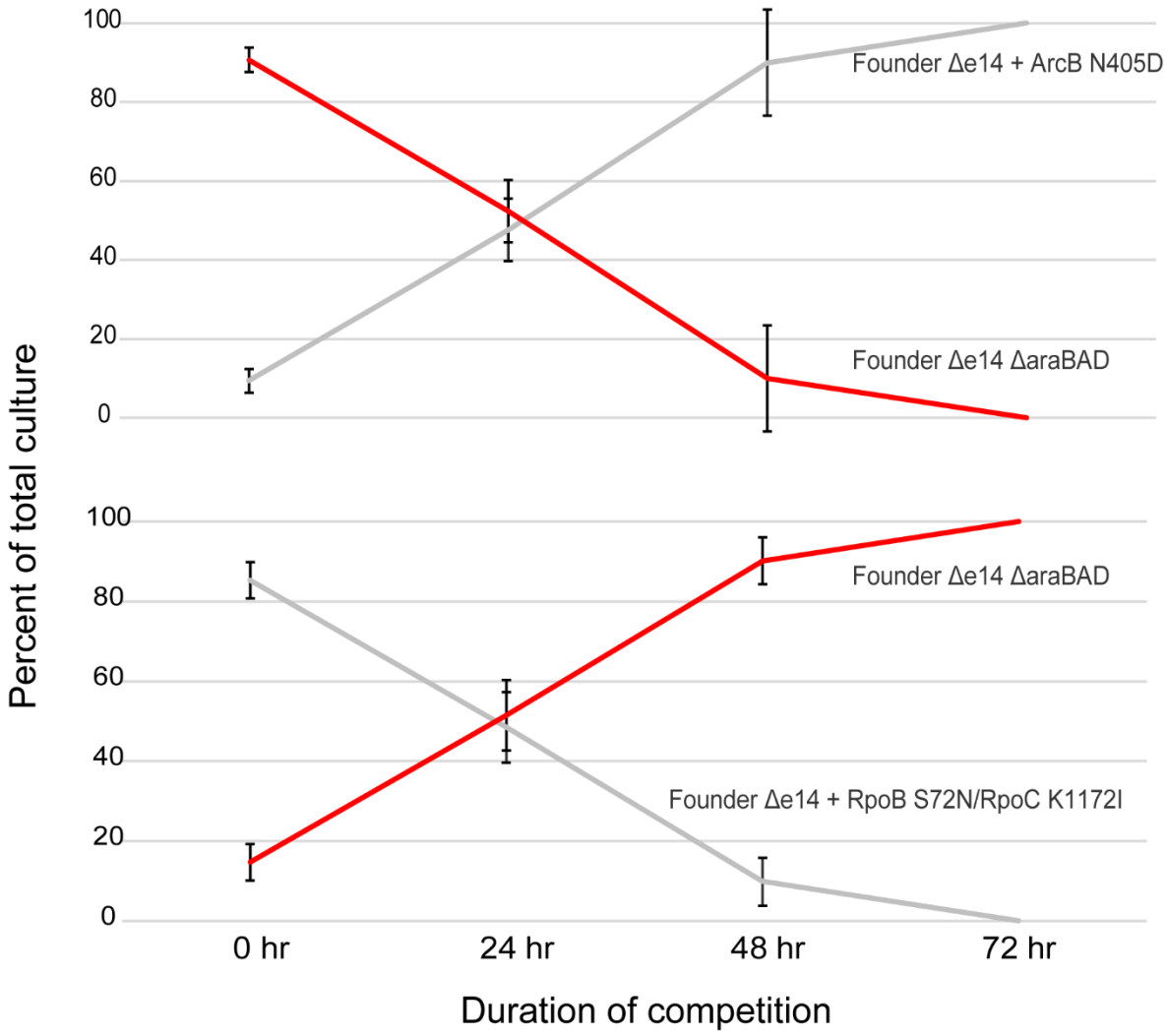
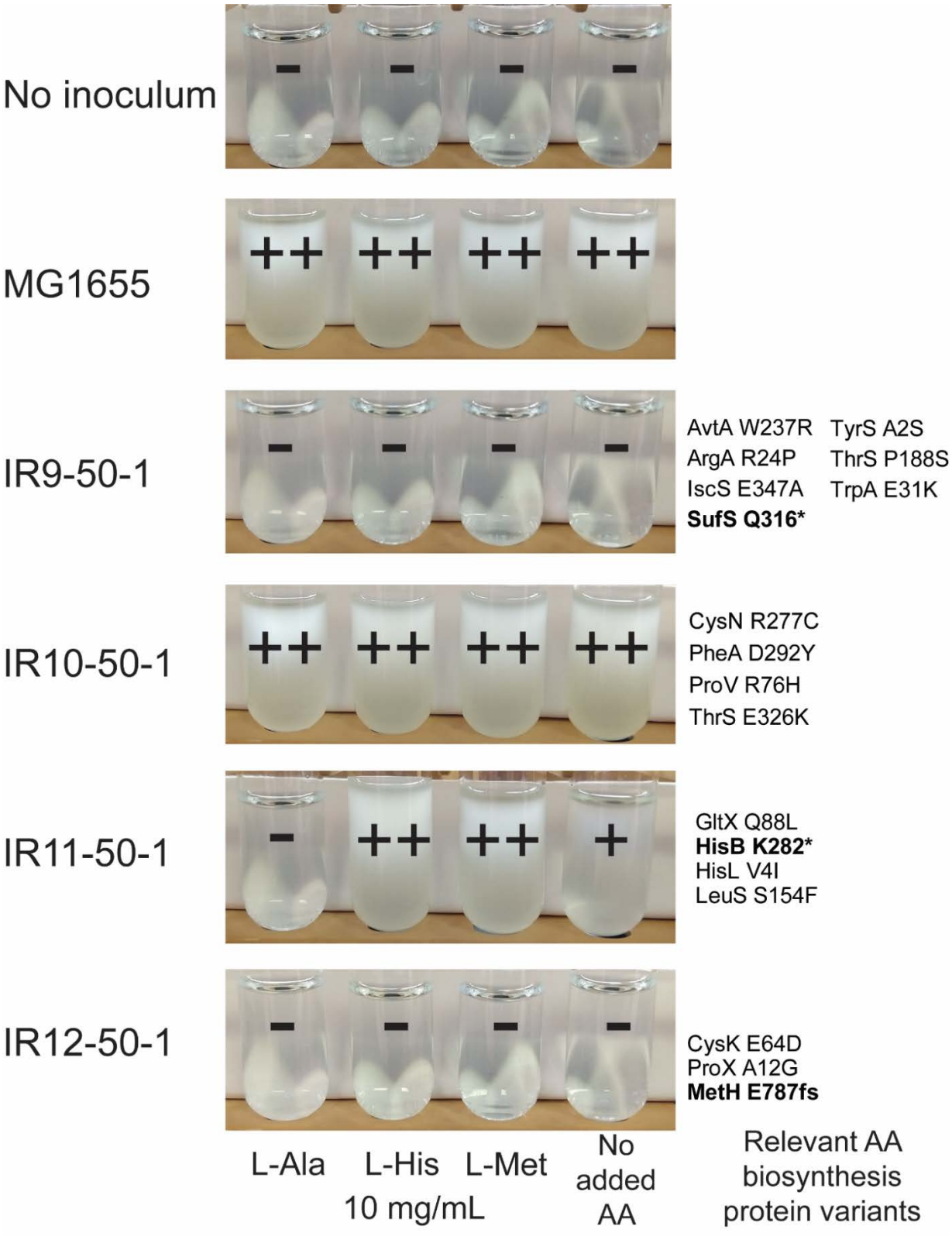
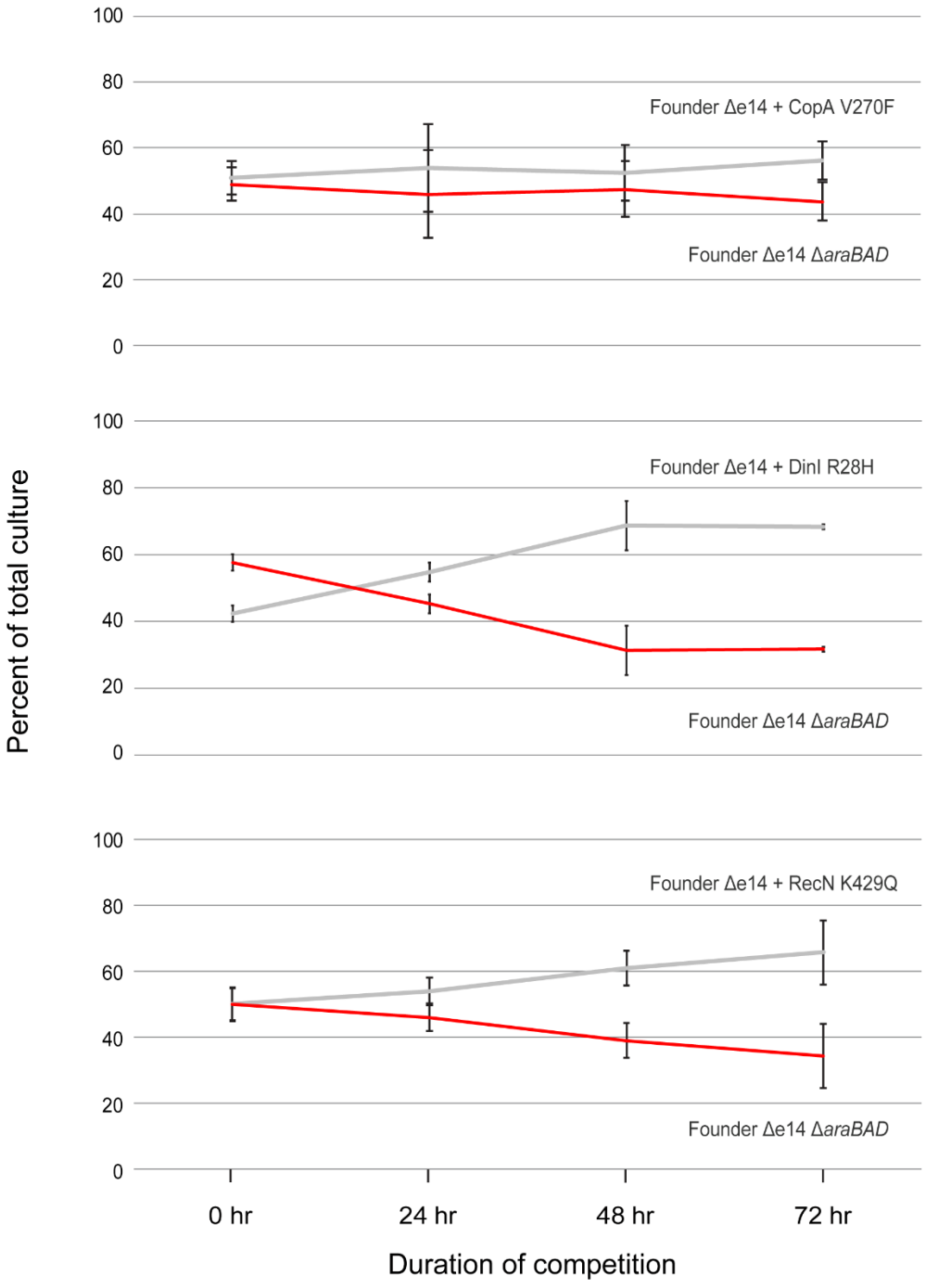


Figure 14.3. Effects of prevalent mutations on growth. Two mutations from lineage IR9 have opposite effects on growth; ArcB N405D greatly enhances growth of Founder e14 whereas the RpoB S72N/RpoC K1172I mutations greatly hinder growth. These results suggest that mutations not seen to affect IR resistance (ArcB N405D) may be selected for in these populations to counteract the deleterious growth effects of mutations which enhance IR resistance. Growth competition assays were performed as described in the *Materials and Methods*. Initial ratios of each strain were purposefully mixed at 9 (loser):1 (winner) so that growth competition assays could be carried out for longer than 24 hr.


No inoculum
- - - -
MG1655
++ ++ ++ ++
IR9-50-1
- - - -
AvtA W237R
ArgA R24P
IscS E347A
SufS Q316*
TyrS A2S
ThrS P188S
TrpA E31K
IR10-50-1
++ ++ ++ ++
CysN R277C
PheA D292Y
ProV R76H
ThrS E326K
IR11-50-1
- ++ ++ +
GltX Q88L
HisB K282*
HisL V4I
LeuS S154F
IR12-50-1
- - - -
CysK E64D
ProX A12G
MetH E787fs
L-Ala
L-His
L-Met
10 mg/mL
No added AA
Relevant AA biosynthesis protein variants

Supplemental Figure 1.3. Supplement with amino acids does not rescue growth of evolved isolates at round 50 of selection in minimal medium. IR9-50-1, IR11-50-1, and IR12-50-1 do not grow in M9 minimal medium after 24 hr of incubation at 37 °C (Figure 6.3). Isolates from all four evolved populations have mutations in amino acid biosynthesis proteins, but only the three that could not grow in minimal medium have predicted loss-of-function mutations (either an introduced stop codon '*', or a frame-shift, 'fs'. IR9-50-1 is predicted to be an alanine auxotroph (SufS Q316*), IR11-50-1 is predicted to be a histidine auxotroph (HisB K282*), and IR12-50-1 is predicted to be a methionine auxotroph (E787fs). However, supplementing minimal medium with each of these amino acids does not rescue the growth phenotypes of these isolates. IR11-50-1 appears able to grow in minimal medium with 48 hr of incubation, with additional rescue by supplementing with His or Met (surprisingly, supplementing with Ala inhibits growth). IR9-50-1 and IR12-50-1 cannot grow in minimal medium with amino acid supplementation or extended incubation time. A '-' symbol indicates no growth, a '+' symbol indicates some growth, and '++' indicates heavy growth.



Supplemental Figure 2.3. Minor effects of prevalent mutations on growth. Growth competition assays were performed as described in the *Materials and Methods*.

Chapter IV: Effects of ionizing radiation on the *E. coli* proteome

4.1 Abstract

Ionizing radiation (IR) causes multi-faceted and wide-spread damage to all cellular macromolecules. Repair of damaged DNA, particularly DNA double-strand breaks (DSBs) to promote faithful replication of the genome is necessary for cell survival. However, cell recovery from irradiation assumes that DNA repair, translation, transcription, and metabolic proteins remain intact. Indeed, the highly radioresistant bacterium *Deinococcus radiodurans* is IR-resistant in part due to an extraordinary ability to ameliorate reactive oxygen species (ROS) produced by IR in order to protect the integrity of the proteome. However, the full extent of oxidative damage to intracellular proteins and how this damage determines cell survival is not well understood. Here we present data from a new effort to characterize IR-induced damage to the *E. coli* proteome. Despite receiving a dose of irradiation that kills 99.9% of cells within an *E. coli* culture, intracellular protein appears to be protected from extensive degradation. Utilizing mass spectrometry to search for carbonylation, oxidation, and peroxidation events has revealed that protein oxidation is the primary form of IR-induced protein damage. Oxidation appears to preferentially occur on methionine and cysteine resides.

4.2 Introduction

Ionizing radiation (IR) can cause lethal damage to cells through ionization of DNA, RNA, protein, and lipids. The majority of damage from IR is due to reactive oxygen species (ROS) produced by interaction of IR with cellular H_2O. Hydroxyl radicals ($^{\bullet}OH$) are a primary product of ionization of intracellular water, which oxidize cellular macromolecules indiscriminately on nanosecond time-scales [4, 16]. Repair of IR-induced damage on DNA is particularly important for cell survival of IR exposure. Oxidation events on opposing DNA strands can lead to DNA double-strand breaks (DSBs), which are lethal if unrepaired [1].

Despite the importance of repairing DNA DSBs generated by IR, the cellular proteome is often considered to be the primary target of damage caused by IR [16, 20, 23] . In particular, the ability of the extremely IR resistant bacterium *Deinococcus radiodurans* to ameliorate ROS and survive 5,000 Gy of IR without loss in viability, despite sustaining similar numbers of DNA DSBs as IR sensitive organisms, has led to the hypothesis that ROS amelioration to protect the *proteome* is the primary determinant of IR resistance [6, 16, 19-22]. Oxidative protein damage in cells includes proteolysis, carbonylation, oxidation and peroxidation, and further, less prominent, species [2, 3, 23, 25, 26, 28, 142]. Carbonylation in particular is often used as a marker for oxidative protein damage [22].With the wide array of potentially inactivating forms of protein damage, significant irradiation may lead to cell death due to proteome damage [16].

In collaboration with the lab of Professor Michael Sussman (University of Wisconsin – Madison), we have begun to explore the effects of high levels of IR on the *Escherichia coli* proteome. Using mass spectrometry, we aim determine the effects of IR on the *E. coli* proteome composition and what oxidative modifications occur on irradiating protein *in vivo* and *in vitro*.

Taken together, we hope to achieve a complete profile of how IR modifies *E. coli* proteins, and how these effects determine the outcome of a cell exposed to IR.

Here we present evidence that the *E. coli* proteome is largely protected from degradation due to a dose of IR by components of the cell cytoplasm. Oxidative modifications detected on amino acids irradiated *in vitro* have informed out search for modifications *in vivo*; despite being protected from degradation, oxidative modifications are prevalent. Additionally, we have observed significant increases in peptides with oxidative modifications due to irradiation, and these increased modifications occur primarily on high-abundance proteins.

4.3 Results

4.3.1 *Escherichia coli* cell lysate is radioprotective

We sought to determine how IR affects the integrity of the *E. coli* proteome at extreme doses. Therefore, we exposed cells to a dose of 1000 Gy, which kills 99.9% of *E. coli* MG1655 (Figure 1.4) and observed the proteome via SDS-PAGE gel electrophoresis (Figures 2.4). Surprisingly, the proteome appeared largely intact. In addition, using an antibody to detect protein carbonylation via Western blot only yielded a single (obvious) carbonylated protein that was only detected in 1000 Gy samples compared to the controls (Figure 3.4). To determine the root of this apparent discrepancy (why is the proteome intact if the majority of cells cannot recover?), we irradiated whole cells, cell lysates, and the purified *E. coli* proteome and again observed the proteome integrity. While the proteome in whole cells appears to be well-protected from degradation, the purified proteome appears nearly completely degraded. The proteins from cell lysates appear protected, but to a lesser extent than those in whole cells (Figure 4.4). Mass spectrometry analysis of these samples confirms the observations from gels. Total numbers of detected proteins are lower in the lysate and pure protein samples, indicating protein degradation. Additionally, oxidized peptides increase in the pure protein samples compared the whole cell and cell lysate samples (Table 1.4).

4.3.2 Reactivity of L-amino acids with high-energy electron beam IR

In order to determine the reactivity of proteins and amino acid residues with IR in *E. coli* cells, we first set out to profile the reactivity of L-amino acids with IR and the modifications that occur *in vitro*. We irradiated 1 mL of 50 µM solutions in dH_2O with 100, 200, 500, and 1000 Gy

using a high-energy electron beam from a linear accelerator (Linac). A summary of the reactivity of amino acids to these doses is presented in Figure 5.4. Our results generally agree with previous studies (Figure 6.4); however, histidine appears to be more reactive than previously reported (in our hands, His is the second-most reactive amino acid behind cystine [in solution, we only observe di-sulfide bonded cysteine: cystine]). Twelve of the amino acids (Cys, Met, Trp, Arg, Leu, Val, His, Tyr, Lys, Phe, Ile, and Thr) are highly reactive with IR (here defined as less than 50% of amino acids in solution are unmodified at a dose of 200 Gy) and the remaining eight are only moderately reactive (>50% are unmodified at 200 Gy). In general, for each amino acid a dose of 1000 Gy will ensure modification of all amino acids in solution. The four least reactive amino acids (in order of decreasing reactivity, Ala, Glu, Asn, and Gly) have an unmodified fraction at 1000 Gy.

Generally, the most prominent modifications detected on amino acid side chains are oxidation (+16 Da) and peroxidation (+32 Da); carbonylation (+14 Da) is also apparent. The aromatic amino acids (Trp, Phe, and Tyr) are all subject to extreme oxidation (3 to 9 oxidation events [+16X Da]) (Table 8.4). The only other amino acid detected with tri-oxidation is cystine.

4.3.3 Targets of oxidation in the *Escherichia coli* proteome

With our profile of electron beam IR-induced damage to amino acids, we sought to determine what modifications occur on proteins *in vivo*. Although the proteome appears, at least superficially, protected from IR-induced degradation, mass spectrometry utilizing TMT-labeling revealed that the *E. coli* proteome is subject to oxidative modifications. We subjected 5 replicate early exponential phase cultures of *E. coli* MG1655 to 1000 Gy of high energy electron beam IR,

and prepared protein from each sample (with and without 1000 Gy) for mass spectrometry analysis.

We detected a total of 1938 unique proteins with 13563 peptide reads. Of these peptides, 13358 were found in all 10 samples (the 5 replicates that received 1000 Gy and the 5 that received 0 Gy). A total of 1440 oxidative modifications were detected, with the majority being oxidation (+16 Da; 1371), followed by carbonylation (+14 Da; 49 detected) and peroxidation (+32; 20 detected). Two-hundred thirty-nine oxidized peptides had a statistically significant ($p<0.01$) fold increase or decrease at 1000 Gy (abundance of peptide in 1000 Gy samples divided by the control samples). One-hundred sixty-four peptides have a fold change greater than 1.2 due to a dose of 1000 Gy, and only 18 had a fold change less than 0.8 (see Table 2.4 for a summary of the numbers of detected peptides and oxidative modifications).

Of the total number of peptides which were modified, we could determine the exact residue modified for approximately 60%. Oxidation (+16 Da) showed a clear preference for methionine residues (approximately 60% of all oxidized residues are methionine, followed by tryptophan which accounts for approximately 10%). Peroxidation (+32 Da) did not show any clear residue bias. Carbonylation events (+14 Da) were searched for on EIKLPQRV residues due to previously reported bias for carbonylation on these amino acids [24, 28]. Of these, carbonylation appears to occur preferentially on leucine residues (approximately one-third). See Tables 3.4 and 4.4 for a summary of modified amino acids.

Table 5.4 details the proteins detected with significant ($p < 0.01$) fold changes for their respective peptide reads. This list represents candidate proteins whose abundance are affected by IR, but are also detectable in samples without irradiation. Of these proteins, those associated with transcription (RpoB [RNA polymerase subunit β] and RpoC [RNA polymerase subunit β’]) and

translation (ribosomal large and small subunit proteins, and particularly elongation factor Tu [TufA]) are prominent, implicating an effect of IR on transcription and translation machinery in the cell.

Of the peptides with significant fold changes induced by IR exposure, 90 exhibit a major increase (fold-change greater than two) (Table 6.4), and 38 a major decrease (fold-change less than 0.6) (Table 7.4). Expectedly, oxidized peptides are more prevalent among peptides with increased fold-change due to irradiation (only a single oxidized peptide appears in the 'major decrease' subset). Peptides mapped to ribosomal proteins are those among the significantly increased peptides; surprisingly, the *E. coli* ribosome recycling factor (Frr) is among the most significantly decreased peptides. The peptide with the greatest increase after 1000 Gy is a peptide of GapA (glyceraldehyde-3-phosphate dehydrogenase A). This peptide has three oxidation events (S148, T150, and T151) which are three of the four resides implicated in the G3P binding site of GapA [143]. The peptide with the greatest fold decrease, YaaA, (approximately 0.3) is involved in the peroxide stress response of *E. coli* [144]. In general, these results underline the significance of IR-induced oxidation in the cell.

4.4 Discussion

IR induced damage to cells is predominantly caused by ROS produced by ionization of intracellular H_2O [4]. As such, it is no surprise that cellular proteins sustain oxidative damage due to IR exposure. However, our work represents, to our knowledge, the only undertaking to determine the full breadth of oxidative damage to the *E. coli* proteome. We have unique access to a linear accelerator (Linac) which generates a high-energy electron beam capable of irradiating samples at extremely high dose rates (72 Gy/min) in controlled conditions. The Linac therefore provides us the opportunity to determine the effects of IR on cells in a manner that is highly repeatable (the Linac does not use a decaying source to produce IR) and that is not available to most laboratories. At the outset of this work, we initially observed a surprising result: at a dose of 1000 Gy, which kills 99.9% of *E. coli* MG1655 cells (Figure 1.4), the proteome appears to be intact and free from massive oxidative damage (Figure 2.4 and 3.4). We subsequently determined that components of the *E. coli* cell lysate are radioprotective, as protein purified from *E. coli* is degraded (Figure 4.4). In order to dig deeper into what oxidative modifications are occurring on the *E. coli* proteome, we first profiled the effects of high energy electron beam IR on L-amino acids. Using the reactivity of amino acids and the side-chain modifications observed as a guide, we performed TMT-labeled mass spectrometry on *E. coli* cultures with and without exposure to 1000 Gy of IR. Our results demonstrated that oxidized peptides are significantly increased in abundance due to irradiation, and that oxidation (+ 16 Da) comprises the majority of IR-induced protein damage compared to carbonylation (+14 Da) and peroxidation (+32 Da). Surprisingly, methionine is the most prevalent oxidized amino acid, despite being the third-most reactive amino acid *in vitro*. Translation and transcription machinery appear to be prevalent targets of oxidation in the cell, and oxidation of these proteins increases with irradiation.

We have observed that at a dose which kills 99.9% of *E. coli* (1000 Gy), the proteome appears intact (Figure 2.4). Additionally, protein carbonylation (an irreversible, likely inactivating oxidative modification of proteins) is prevalent without irradiation and only modestly increases after a dose of 1000 Gy (Figure 3.4). These results contrast without previous studies showing protein carbonylation is not observed without irradiation [22]. Our results suggest that indirect assays for determining protein carbonylation may not be an effective means for quantifying proteome oxidation, and that at least protein degradation due to IR exposure may be less prominent than originally thought. We subsequently assayed proteome integrity of *E. coli* cells lysed prior to irradiation and of purified and irradiated *E. coli* protein. There was clear protection of the proteome in irradiated whole cell and cell lysate samples; however, purified protein experienced significant degradation due to IR exposure (Figure 4.4). Clearly, there are components of *E. coli* cytoplasm that are radioprotective. Low molecular-weight antioxidants are present in all cells and are implicated as a major determinant of IR resistance in the highly radioresistant bacterium *D. radiodurans* [19, 22]. In particular, glutathione is a known antioxidant in *E. coli* used to reduce oxidized cysteine residues [26, 27].

Although the *E. coli* proteome appears intact despite significant lethality, this does not rule out a role for other oxidative modifications in inactivation of *E. coli* proteins. We set out to determine the effects of IR on the *E. coli* proteome. We first began by cataloguing the IR-induced modifications on L-amino acids using high-energy electron beam IR. Although previous work has been published on the reactivity of amino acids with IR in a variety of conditions, we set out to confirm our conditions affect amino acids similarly [24, 145, 146]. While our data largely agree with previously published results, we observe His to be more reactive than in previous studies (Figure 6.4). Oxidation (+16 Da) and peroxidation (+32 Da) are the most common oxidative

modifications observed; carbonylation (+14 Da) is also prevalent. Aromatic amino acids (Thr, Trp, and Phe) are particularly susceptible to multiple oxidation events, and are the only amino acids with more than 3 oxidation events on a single side chain (Table 8.4).

Using the reactivity of amino acids *in vitro* as a guide, we set out to determine the full breadth of oxidative damage to cellular proteins due to irradiation. We utilized quintuplicate replicate cultures of *E. coli* that were either unirradiated or exposed to 1000 Gy along with TMT-labelled mass spectrometry to determine the abundance of oxidized peptides in the cell. The 13708 peptide reads (that were detected in all ten samples) mapped to 1938 proteins, and of these peptides approximately 12% were oxidatively modified (Table 2.4). Of the total number of detected peptides, 2207 had a significant ($p < 0.01$) fold change due to irradiation. While approximately 37% of detected peptides had a major fold change increase (defined as a fold change greater than 1.2), nearly 60% of oxidized peptides fall into this category indicating that oxidized peptides are more prevalent due to irradiation (Table 2.4).

Although cysteine appears to be the most reactive amino acid *in vitro*, the amino acid residue most susceptible to IR-induced oxidation is methionine *in vivo* (Table 3.4). Tryptophan is the second-most prominent modified residue. Why Met and Trp are clearly the favored residues for oxidation behind Cys and His (the top two most reactive amino acids *in vitro*) is not clear. Although there are abundant mechanisms in the cell for repair of oxidized Cys residues, repair of oxidized cellular Met is also well-known [26]. Additionally, cultures in these experiments are incubated at 4 °C, are suspended in 1X PBS (which contains no metabolites capable of supporting growth) and irradiations occur over short (15 minute) time-scales. Cells are flash-frozen and stored at -80 °C within 15 min after irradiation. Taken together, the cultured cells likely are unable to mount a response to damage induced by IR. Therefore, the apparent preference for oxidation on

Met and Trp is likely a result of reactivity of these amino acids in the context of a peptide bond as opposed to amino acids free in solution.

Protein abundance appears to be a major factor in susceptibility to oxidation. Peptides mapped to the most abundant *E. coli* protein, TufA (elongation factor Tu) [147], peptides account for eight of 90 peptides with a fold change increase greater than two, and each of these TufA peptides are oxidized (Table 6.4). Other prominent peptides with fold change increases greater than 2 map to large and small ribosomal proteins, the outer-membrane protein Pal, and glyceraldehyde-3-phosphate dehydrogenase A (GapA). Unlike the peptides of TufA, not all peptides mapping to these proteins are oxidized. Interestingly, the peptide with the largest fold change (approximately 19) is a peptide of GapA with two oxidation and one peroxidation event. These three modifications lie in the SCTT motif of GapA, which is implicated in glyceraldehyde-3-phosphate (G3P) binding [143]. Three other peptides of GapA are detected with a fold change greater than 2, but none are oxidized. It is unclear why the G3P-binding portion of GapA appears to be particularly prone to oxidation compared to other, more abundant proteins.

Again, it is important to note that these fold-change values likely do not reflect a transcriptional or translational cellular 'response' to IR, but rather susceptibility to oxidation. As noted in the *Materials and Methods* section, cells should be metabolically inactive due to cold temperatures (4 °C) over short time-scales and being suspended in PBS, which contains no nutrients capable of supporting growth. Decreased fold-change of peptides could reflect further oxidation events on those peptides due to IR, therefore the original peptide is detected less with IR due to a change in mass. The increase in fold change of oxidized peptides is logical for the same reasons; peptides are being increasingly oxidized due to IR, so peptides susceptible to oxidation appear more often. However, it is not clear why unoxidized peptides would increase in abundance

due to IR. Oxidative modifications to other areas on the protein may modify the protein structure in such a way that proteases utilized for mass spectrometry may more easily access the residues necessary to produce said peptides.

Our endeavor to catalogue the effect of IR on the *E. coli* proteome is still in its early stages. We have only just begun to characterize what modifications occur, where on a protein they will occur, and what proteins are more or less susceptible to these IR-induced modifications. An important step in this work will be to repeat our mass spectrometry experiments with *E. coli* evolved to be IR resistant, and compare how the proteome of these experimentally-evolved extremophiles differs from their parent strain. What changes there may be, or lack thereof, will be illuminating. We hope to use *E. coli* as a predictive model; given a dose of IR, we aim to predict which proteins will be modified, on which residues, and how this may impact overall physiology of the cell.

4.5 Tables

	Whole Cells	Lysate	Pure protein
Total peptides	3127	2432	1875
Oxidized peptides	94	98	115

Table 1.4. Numbers of detected peptides in samples of whole cells, cell lysate, and purified protein exposed to 1000 Gy of IR. Samples of *E. coli* whole cells, cell lysates, and pure proteins were prepared and irradiated as described in the *Materials and Methods* section. Detected peptide values indicate increased protein degradation in purified protein samples compared to whole cell and cell lysate samples. Consistently, oxidized peptides are increased in purified protein samples. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Total proteins detected			1938	
Total peptides detected			13708	**% of total**
Carbonylations (+14 Da)			56	0.41
Oxidations (+16 Da)			1564	11.41
Peroxidations (+32 Da)			25	0.18
	Significant fold change (1000 Gy abundance/0 Gy abundance)		2207	**% of total**
	Fold change > 1.2		825	37.38
	Fold change < 0.8		593	26.87
	Carbonylations (+14 Da)		11	0.50
	Oxidations (+16 Da)		277	12.55
	Peroxidations (+32 Da)		2	0.09
		With Oxidation	232	**% of total**
		Fold change > 1.2	136	58.62
		Fold change < 0.8	18	7.76

Table 2.4. Summary of total peptides detected, numbers of modifications detected, and numbers of peptides with significant fold changes. Peptides reported were detected in all ten *E. coli* samples; 5 replicate cultures exposed to 0 and 1000 Gy. Peptides were searched for carbonylations (+14 Da), oxidations (+16 Da), and peroxidations (+32 Da); peptides with these modifications make up approximately 12% of all detected peptides. Significant fold changes are defined as the fold difference between abundances of peptides in 1000 Gy samples and the same peptides in 0 Gy samples with a p-value less than 0.01. It should be noted that with a total of 13000 detected peptides, this means that approximately 130 of the fold changes will appear significant purely due to chance. Numbers of peptides with fold changes greater than 1.2 and less than 0.8 are singled out as peptides with ‘major’ increases or decreases. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Modifications per Amino Acid on all detected peptides			
Amino Acid	Oxidation (+16)	Peroxidation (+32)	Carbonylation (+14)
A	36	0	0
C	2	0	0
D	12	1	0
E	9	0	3
F	8	0	0
G	0	2	0
H	0	1	0
I	14	1	2
K	0	0	0
L	33	3	10
M	615	3	0
N	12	0	0
P	12	0	5
Q	13	0	2
R	4	0	0
S	14	0	0
T	11	0	0
V	35	0	5
W	112	1	0
Total modifications with known AA target	942	12	27
Overall total modifications	1564	25	56

Table 3.4. Numbers of modifications on each amino acid on all detected peptides isolated from *Escherichia coli* irradiated with 1000 Gy. Oxidation (+16 Da) appears to be heavily biased towards Met residues, followed by Trp. There is no obvious bias for residues for peroxidation. Carbonylation appears to occur primarily on Leu residues. Out of 1640 total modifications, we could determine with confidence which residue was modified for 981. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Modifications per Amino Acid on peptides with statisitcally significant fold changes (1000 Gy/ 0 Gy abundance)			
Amino Acid	Oxidation (+16)	Peroxidation (+32)	Carbonylation (+14)
A	18	0	0
C	0	0	0
D	9	0	0
E	1	0	1
F	4	0	0
G	0	0	0
H	0	1	0
I	2	0	2
K	0	0	0
L	6	0	3
M	94	0	0
N	6	0	0
P	1	0	3
Q	2	0	0
R	1	0	0
S	6	0	0
T	9	0	0
V	4	0	2
W	24	1	0
Total modifications with known AA target	187	2	11
Overall total modifications	277	2	11

Table 4.4. Numbers of modifications on each amino acid on peptides with significant ($p < 0.01$) fold changes isolated from *Escherichia coli* irradiated with 1000 Gy. As with the total peptides, oxidation (+16 Da) appears to be heavily biased towards Met residues, followed by Trp. There is no obvious bias for residues for peroxidation or carbonylation. Out of 290 total modifications, we could determine with confidence which residue was modified for 200. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Protein	Protein description	ID	Number of unique reads
AccC	Biotin carboxylase	P24182	1
AceE	Pyruvate dehydrogenase E1 component	P0AFG8	1
AckA	Acetate kinase	P0A6A3	1
AcnA	Aconitate hydratase A	P25516	1
AcnB	Aconitate hydratase B	P36683	2
AcrB	Multidrug efflux pump subunit AcrB	P31224	1
Adk	Adenylate kinase	P69441	2
AhpF	Alkyl hydroperoxide reductase subunit F	P35340	1
AlaA	Glutamate-pyruvate aminotransferase AlaA	P0A959	1
AlaS	Alanine--tRNA ligase	P00957	2
ArgG	Argininosuccinate synthase	P0A6E4	1
AroK	Shikimate kinase 1	P0A6D7	1
AroP	Aromatic amino acid transport protein AroP	P15993	1
AspA	Aspartate ammonia-lyase	P0AC38	1
AspC	Aspartate aminotransferase	P00509	1
AtpD	ATP synthase subunit beta	P0ABB4	2
AtpG	ATP synthase gamma chain	P0ABA6	1
BamD	Outer membrane protein assembly factor BamD	P0AC02	1
Cann	Carbonic anhydrase 2	P61517	1
ClpB	Chaperone protein ClpB	P63284	1
Crr	Glucose-specific phosphotransferase enzyme IIA component	P69783	2
CspE	Cold shock-like protein CspE	P0A972	1
CyoA	Cytochrome bo(3) ubiquinol oxidase subunit 2	P0ABJ1	2
CyoB	Cytochrome bo(3) ubiquinol oxidase subunit 1	P0ABI8	1
CysJ	Sulfite reductase [NADPH] flavoprotein alpha-component	P38038	1
DacA	D-alanyl-D-alanine carboxypeptidase DacA	P0AEB2	1
DacC	D-alanyl-D-alanine carboxypeptidase DacC	P08506	1
DapA	4-hydroxy-tetrahydrodipicolinate synthase	P0A6L2	1
DcrB	Phage protein DcrB	P0AEE1	1
DeaD	ATP-dependent RNA helicase DeaD	P0A9P6	2
Derr	GTPase Der	P0A6P5	1
DnaK	Chaperone protein DnaK	P0A6Y8	1
Eno	Enolase	P0A6P9	1
FabG	3-oxoacyl-[acyl-carrier-protein] reductase FabG	P0AEK2	3
FadL	Long-chain fatty acid transport protein	P10384	1
FbaA	Fructose-bisphosphate aldolase class 2	P0AB71	1
FtsH	ATP-dependent zinc metalloprotease FtsH	P0AAI3	2
FtsZ	Cell division protein FtsZ	P0A9A6	2
FumA	Fumarate hydratase class I, aerobic	P0AC33	1
FusA	Elongation factor G	P0A6M8	3
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	P0A9B2	3
GlmU	Bifunctional protein GlmU	P0ACC7	1
GlnS	Glutamine--tRNA ligase	P00962	1
GlpD	Aerobic glycerol-3-phosphate dehydrogenase	P13035	1
GlpK	Glycerol kinase	P0A6F3	1
GlpQ	Glycerophosphoryl diester phosphodiesterase	P09394	1
GltA	Citrate synthase	P0ABH7	2
GltB	Glutamate synthase [NADPH] large chain	P09831	1
GlyS	Glycine--tRNA ligase beta subunit	P00961	2
Gnd	6-phosphogluconate dehydrogenase, decarboxylating	P00350	1

Table 5.4A. Number of peptide reads for proteins with significant fold changes and oxidative modifications.

Protein	Protein description	ID	Number of unique reads
GrxD	Glutaredoxin-4	P0AC69	1
GuaA	GMP synthase [glutamine-hydrolyzing]	P04079	3
HemL	Glutamate-1-semialdehyde 2,1-aminomutase	P23893	2
HtpG	Chaperone protein HtpG	P0A6Z3	1
Icd	Isocitrate dehydrogenase [NADP]	P08200	2
IleS	Isoleucine--tRNA ligase	P00956	1
InfB	Translation initiation factor IF-2	P0A705	2
InfC	Translation initiation factor IF-3	P0A707	1
IspG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	P62620	1
LeuS	Leucine--tRNA ligase	P07813	2
LipA	Lipoyl synthase	P60716	1
Lon	Lon protease	P0A9M0	1
LpdA	Dihydrolipoyl dehydrogenase	P0A9P0	1
LpoB	Penicillin-binding protein activator LpoB	P0AB38	1
Lpp	Major outer membrane lipoprotein Lpp	P69776	1
LptB	Lipopolysaccharide export system ATP-binding protein LptB	P0A9V1	1
LptD	LPS-assembly protein LptD	P31554	1
LptE	LPS-assembly lipoprotein LptE	P0ADC1	1
MdaB	Modulator of drug activity B	P0AEY5	1
MetE	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	P25665	1
MreB	Rod shape-determining protein MreB	P0A9X4	1
MurA	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	P0A749	1
Ndh	NADH dehydrogenase	P00393	1
NuoC	NADH-quinone oxidoreductase subunit C/D	P33599	3
NuoF	NADH-quinone oxidoreductase subunit F	P31979	2
NusG	Transcription termination/antitermination protein NusG	P0AFG0	1
OmpA	Outer membrane protein A	P0A910	2
OmpR	Transcriptional regulatory protein	P0AA16	1
OppD	Oligopeptide transport ATP-binding protein	P76027	1
OppF	Oligopeptide transport ATP-binding protein	P77737	2
Pall	Peptidoglycan-associated lipoprotein	P0A912	2
ParE	DNA topoisomerase 4 subunit B	P20083	1
PflB	Formate acetyltransferase 1	P09373	2
Pgi	Glucose-6-phosphate isomerase	P0A6T1	3
PitA	Low-affinity inorganic phosphate transporter 1	P0AFJ7	1
PlsB	Glycerol-3-phosphate acyltransferase	P0A7A7	1
PntB	NAD(P) transhydrogenase subunit beta	P0AB67	1
PpsA	Phosphoenolpyruvate synthase	P23538	2
Prc	Tail-specific protease	P23865	1
PrfC	Peptide chain release factor 3	P0A7I4	1
PtsI	Phosphoenolpyruvate-protein phosphotransferase	P08839	1
PurA	Adenylosuccinate synthetase	P0A7D4	1
PurB	Adenylosuccinate lyase	P0AB89	2
PyrG	CTP synthase	P0A7E5	1
PyrH	Uridylate kinase	P0A7E9	1
RbsB	Ribose import binding protein RbsB	P02925	1
RecA	DNA strand exchange and recombination protein with protease and nuclease activity	P0A7G6	2

Table 5.4B. Number of peptide reads for proteins with significant fold changes and oxidative modifications.

Protein	Protein description	ID	Number of unique reads
RfbA	Glucose-1-phosphate thymidylyltransferase 1	P37744	1
Rho	Transcription termination factor Rho	P0AG30	1
RibE	6,7-dimethyl-8-ribityllumazine synthase	P61714	1
RimM	Ribosome maturation factor RimM	P0A7X6	1
RlmH	Ribosomal RNA large subunit methyltransferase H	P0A8I8	1
RplB	50S ribosomal protein L2	P60422	3
RplE	50S ribosomal protein L5	P62399	2
RplF	50S ribosomal protein L6	P0AG55	1
RplK	50S ribosomal protein L11	P0A7J7	1
RplN	50S ribosomal protein L14	P0ADY3	1
RplO	50S ribosomal protein L15	P02413	1
RplP	50S ribosomal protein L16	P0ADY7	2
RplQ	50S ribosomal protein L17	P0AG44	1
RplS	50S ribosomal protein L19	P0A7K6	1
RplU	50S ribosomal protein L21	P0AG48	1
RplY	50S ribosomal protein L25	P68919	1
RpmB	50S ribosomal protein L28	P0A7M2	1
RpmD	50S ribosomal protein L30	P0AG51	1
RpmI	50S ribosomal protein L35	P0A7Q1	1
RpoA	DNA-directed RNA polymerase subunit alpha	P0A7Z4	1
RpoB	DNA-directed RNA polymerase subunit beta	P0A8V2	6
RpoC	DNA-directed RNA polymerase subunit beta'	P0A8T7	7
RpsB	30S ribosomal protein S2	P0A7V0	2
RpsE	30S ribosomal protein S5	P0A7W1	1
RpsG	30S ribosomal protein S7	P02359	1
RpsI	30S ribosomal protein S9	P0A7X3	2
RpsK	30S ribosomal protein S11	P0A7R9	2
RpsN	30S ribosomal protein S14	P0AG59	1
RsmH	Ribosomal RNA small subunit methyltransferase H	P60390	1
SdaB	L-serine dehydratase 2	P30744	1
SecA	Protein translocase subunit SecA	P10408	1
SecF	Protein translocase subunit SecF	P0AG93	1
SlyB	Outer membrane lipoprotein SlyB	P0A905	2
SpeA	Biosynthetic arginine decarboxylase	P21170	1
SspA	Stringent starvation protein A	P0ACA3	1
SucA	2-oxoglutarate dehydrogenase E1 component	P0AFG3	1
SucC	Succinyl-CoA ligase [ADP-forming] subunit beta	P0A836	1
TalB	Transaldolase B	P0A870	1
Tdh	L-threonine 3-dehydrogenase	P07913	1
Tig	Trigger factor	P0A850	3
TopA	DNA topoisomerase 1	P06612	1
TrpS	Tryptophan--tRNA ligase	P00954	1
TrxB	Thioredoxin reductase	P0A9P4	1
TrxC	Thioredoxin-2	P0AGG4	1
Tsff	Elongation factor Ts	P0A6P1	2
TufA	Elongation factor Tu 1	P0CE47	10
ValS	Valine--tRNA ligase	P07118	2
WecB	UDP-N-acetylglucosamine 2-epimerase	P27828	1
YcaO	Ribosomal protein S12 methylthiotransferase accessory factor Yca	P75838	1
YebC	Probable transcriptional regulatory protein YebC	P0A8A0	1
YgfZ	tRNA-modifying protein YgfZ	P0ADE8	2
YgiB	UPF0441 protein YgiB	P0ADT2	1

Table 5.4C. Number of peptide reads for proteins with significant fold changes and oxidative modifications.

Protein	Protein description	ID	Number of unique reads
YhcB	Inner membrane protein YhcB	P0ADW3	1
YheS	Uncharacterized ABC transporter ATP-binding protein YheS	P63389	1
YhhX	Uncharacterized oxidoreductase YhhX	P46853	1
YibN	Uncharacterized protein YibN	P0AG27	1
YihD	Uncharacterized protein YihD	P0ADP9	1
YjbJ	Uncharacterized protein YjbJ	P68206	1
YncE	Uncharacterized protein YncE	P76116	1
Zwf	Glucose-6-phosphate 1-dehydrogenase	P0AC53	1

Table 5.4D. Number of peptide reads for proteins with significant fold changes and oxidative modifications. Proteins were detected from mapped peptide reads detected from TMT-labelled mass spectrometry of protein purified from quintuplicate *Escherichia coli* samples irradiated with and without 1000 Gy or IR. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Protein	Description	ID	Peptide Sequence	Fold Change	p-value	Modification?
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	P0A9B2	YAGQDIVSNA**S**C**TT**NCLAPLAK	19.90	9.50E-06	3 Oxidations
TufA	Elongation factor Tu 1	P0CE47	**AFD**QI**DNA**PEEK	6.22	8.72E-06	1 Oxidation
Pal	Peptidoglycan-associated lipoprotein	P0A912	GVSADQISIVSYGK	5.91	1.66E-03	None
Pal	Peptidoglycan-associated lipoprotein	P0A912	LQMQQLQQNNIVYFDLDK	5.50	8.23E-03	None
TufA	Elongation factor Tu 1	P0CE47	VGEEVEIVGIK	5.43	2.82E-05	1 Oxidation
RplB	50S ribosomal protein L2	P60422	DGAYVTLR	5.14	8.86E-03	None
TufA	Elongation factor Tu 1	P0CE47	STCTGVE**MF**R	4.76	1.37E-06	2 Oxidations
RplB	50S ribosomal protein L2	P60422	AGDQIQSGVDAAIKPGNTLPMR	4.53	7.69E-03	None
Pal	Peptidoglycan-associated lipoprotein	P0A912	GTPEYNISLGER	4.53	7.72E-03	None
RpsK	30S ribosomal protein S11	P0A7R9	NLEV**M**VK	4.40	6.43E-03	1 Oxidation
RpsI	30S ribosomal protein S9	P0A7X3	QPLELVD**M**VEK	4.19	9.67E-03	1 Oxidation
RplB	50S ribosomal protein L2	P60422	GKPFAPLLEK	4.15	2.59E-04	None
TufA	Elongation factor Tu 1	P0CE47	AGENVGV**L**LR	4.15	4.30E-06	1 Oxidation
RplB	50S ribosomal protein L2	P60422	SANIALVLYK	4.03	3.66E-03	None
Pal	Peptidoglycan-associated lipoprotein	P0A912	**LQMQQLQQNN**IVYFD**L**DK	3.85	7.17E-03	1 Oxidation
RpsK	30S ribosomal protein S11	P0A7R9	QGN**A**LG**WATA**GGSGFR	3.83	7.63E-03	1 Oxidation
RpsL	30S ribosomal protein S12	P0A7S3	SNVPALEACPQK	3.82	8.92E-03	None
RplB	50S ribosomal protein L2	P60422	SAGTYVQIVAR	3.78	5.15E-03	None
RpsI	30S ribosomal protein S9	P0A7X3	GGGISGQAGAIR	3.71	9.17E-03	None
RpsK	30S ribosomal protein S11	P0A7R9	QGN**A**LG**WATA**GG**S**GFR	3.67	5.10E-03	2 Oxidations
RpsK	30S ribosomal protein S11	P0A7R9	STPFAAQVAAER	3.62	7.81E-03	None
RplB	50S ribosomal protein L2	P60422	VVNPELHK	3.62	2.67E-03	None
RpsI	30S ribosomal protein S9	P0A7X3	LDLYITVK	3.60	8.22E-03	None
RplB	50S ribosomal protein L2	P60422	HPVTPWGVQTK	3.60	7.82E-04	None
RpmH	50S ribosomal protein L34	P0A7P5	TFQPSVLK	3.42	5.65E-03	None
Lpp	Major outer membrane lipoprotein Lpp	P69776	VDQLSNDVNA**M**R	3.40	9.06E-03	1 Oxidation
RpsK	30S ribosomal protein S11	P0A7R9	QGNALGWATAGGSGFR	3.37	6.92E-03	None
Pal	Peptidoglycan-associated lipoprotein	P0A912	EKPAVLGHDEAAYSK	3.36	1.27E-04	None
RpsK	30S ribosomal protein S11	P0A7R9	ALNAAGFR	3.35	7.00E-03	None
RplB	50S ribosomal protein L2	P60422	DGIPAVVER	3.34	8.63E-03	None
FusA	Elongation factor G	P0A6M8	IATDP**F**VGNLTFFR	3.31	7.26E-06	1 Oxidation
TufA	Elongation factor Tu 1	P0CE47	**L**LDEGR	3.31	1.73E-05	1 Oxidation
RpsL	30S ribosomal protein S12	P0A7S3	GALDCSGVK	3.26	8.01E-03	None
RpsL	30S ribosomal protein S12	P0A7S3	SNVPALEACPQKR	3.24	7.97E-03	None
PalI	Peptidoglycan-associated lipoprotein	P0A912	GTPEYNISLGERR	3.19	2.82E-03	None
FusA	Elongation factor G	P0A6M8	ILFYTGVN**H**K	3.18	9.90E-06	1 Peroxidation
TalB	Transaldolase B	P0A870	LTIAPA**L**LK	3.14	2.60E-05	1 Oxidation
RplO	50S ribosomal protein L15	P02413	GFEGGQ**M**PLYR	3.11	7.65E-03	1 Oxidation
FtnA	Bacterial non-heme ferritin	P0A998	SGEGLYFIDK	3.01	7.81E-03	None
LptE	LPS-assembly lipoprotein LptE	P0ADC1	SFFDNPQ**M**ALAK	2.95	1.67E-06	1 Oxidation
RpsL	30S ribosomal protein S12	P0A7S3	VYTTTPK	2.91	7.13E-03	None
RplB	50S ribosomal protein L2	P60422	LEYDPNR	2.89	4.59E-03	None
RpsI	30S ribosomal protein S9	P0A7X3	SLEQ**Y**FGR	2.85	9.52E-03	1 Oxidation
RplQ	50S ribosomal protein L17	P0AG44	LFNELGPR	2.80	9.96E-03	None
TufA	Elongation factor Tu 1	P0CE47	QVGVPYIIV**FL**NK	2.76	1.13E-04	1 Oxidation
RplB	50S ribosomal protein L2	P60422	H**PVTPW**GVQTK	2.75	3.31E-04	1 Oxidation
RpsK	30S ribosomal protein S11	P0A7R9	ITNITDVTPIPHNGCRPPK	2.73	2.74E-03	None
RpmA	50S ribosomal protein L27	P0A7L8	FISIEAE	2.73	7.20E-03	None
RplB	50S ribosomal protein L2	P60422	RYILAPK	2.64	3.75E-04	None
FtnA	Bacterial non-heme ferritin	P0A998	LFDYLTDTGNLPR	2.59	7.59E-03	None
Dcp	Peptidyl-dipeptidase dcp	P24171	VLNTEAATLTSQFNQR	2.59	9.66E-07	None
Pal	Peptidoglycan-associated lipoprotein	P0A912	SDFAQMLDAHANFLR	2.55	1.47E-03	None
RluB	Ribosomal large subunit pseudouridine synthase B	P37765	EIESIIEAGR	2.52	7.23E-03	None
RpmA	50S ribosomal protein L27	P0A7L8	FGGESVLAGSIIVR	2.51	8.48E-03	None
RpmI	50S ribosomal protein L35	P0A7Q1	GDLGLVIACLPYA	2.45	4.73E-03	None
RplQ	50S ribosomal protein L17	P0AG44	NMAGSLVR	2.44	7.35E-03	None

Table 6.4A. Detected peptides with a fold change value greater than two.

Protein	Description	ID	Peptide Sequence	Fold Change	p-value	Modification?
RpsU	30S ribosomal protein S21	P68679	ENEPFDVALRR	2.43	4.37E-04	None
RplP	50S ribosomal protein L16	P0ADY7	VLYE**MD**GVPEELAR	2.41	8.36E-03	1 Oxidation
RpsN	30S ribosomal protein S14	P0AG59	AIISDVNASDED**R**W**N**AVLK	2.38	1.96E-03	2 Oxidations
Pal	Peptidoglycan-associated lipoprotein	P0A912	SDFAQ**M**LDAHANFLR	2.38	2.53E-03	1 Oxidation
RpsS	30S ribosomal protein S19	P0A7U3	GPFIDLHLLK	2.38	8.29E-03	None
RluB	Ribosomal large subunit pseudouridine synthase B	P37765	YGDIPLPK	2.35	7.55E-03	None
RpsM	30S ribosomal protein S13	P0A7S9	AILAAAGIAEDVK	2.34	6.49E-03	None
HinT	Purine nucleoside phosphoramidase	P0ACE7	IAEQEGIAEDGYR	2.31	1.12E-03	None
RplQ	50S ribosomal protein L17	P0AG44	**A**GDN**APMA**YIELVDR	2.31	5.86E-03	1 Oxidation
RplP	50S ribosomal protein L16	P0ADY7	GLAQGTDVSFGSFGLK	2.27	4.83E-03	None
PanD	Aspartate 1-decarboxylase	P0A790	FSTYAIAAER	2.25	1.04E-06	None
RpsN	30S ribosomal protein S14	P0AG59	AIISDVNASDEDRWNAVLK	2.23	3.25E-03	None
Pall	Peptidoglycan-associated lipoprotein	P0A912	VTVEGHADER	2.23	7.25E-04	None
HsdS	Type-1 restriction enzyme EcoKI specificity protein	P05719	VNNLTQSILAK	2.22	8.21E-03	None
RplB	50S ribosomal protein L2	P60422	ATLGEVGNAEHMLR	2.20	9.52E-04	None
RsxC	Electron transport complex subunit RsxC	P77611	AAVEAAIAR	2.17	7.83E-03	None
RluB	Ribosomal large subunit pseudouridine synthase B	P37765	GGWTELDLAQTNYLR	2.15	8.52E-03	None
RluB	Ribosomal large subunit pseudouridine synthase B	P37765	WIAVGR	2.15	4.75E-03	None
RplO	50S ribosomal protein L15	P02413	AANIIGIQIEFAK	2.12	3.52E-03	None
FabR	HTH-type transcriptional repressor FabR	P0ACU5	EAGIAPTSFYR	2.12	7.56E-03	None
RpsU	30S ribosomal protein S21	P68679	EFYEKPTTER	2.10	3.54E-03	None
YihD	Protein YihD	P0ADP9	EPDLNLLQFLQK	2.09	1.63E-04	None
RpsM	30S ribosomal protein S13	P0A7S9	ISELSEGQIDTLRDEVAK	2.09	1.97E-03	None
TufA	Elongation factor Tu 1	P0CE47	TVGAGVV**A**K	2.09	3.76E-04	1 Oxidation
HsdS	Type-1 restriction enzyme EcoKI specificity protein	P05719	SQVVLLPPVK	2.08	6.11E-03	None
HtpG	Chaperone protein HtpG	P0A6Z3	LTDTPAIVSTDADE**M**STQ**M**AK	2.08	4.48E-05	2 Oxidations
FtnA	Bacterial non-heme ferritin	P0A998	ELSTLDTQN	2.07	9.51E-03	None
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	P0A9B2	VGINGFGR	2.05	1.99E-05	None
TufA	Elongation factor Tu 1	P0CE47	**TTLTAAITTVLA**K	2.04	1.87E-04	1 Oxidation
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	P0A9B2	AGIALNDNFVK	2.04	1.48E-05	None
RplO	50S ribosomal protein L15	P02413	GFEGGQMPLYR	2.03	7.16E-03	None
RpmB	50S ribosomal protein L28	P0A7M2	**FW**VESEK	2.03	9.63E-03	1 Oxidation
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	P0A9B2	AGIALNDNFVK	2.01	7.66E-06	None

Table 6.4B. Detected peptides with a fold change value greater than two. All peptides with fold changes greater than two due to irradiation are listed. The protein they map to and the peptide sequence are also indicated. Where an oxidative modification occurs, the amino acid residue is bolded. If the residue is also underlined, that indicates that the modification has been identified on that particular residue; if the residue is only bolded then that residue is only a candidate for where the modification occurs. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Protein	Description	ID	Peptide Sequence	Fold Change	p-value	Modification?
GlnS	Glutamine--tRNA ligase	P00962	AGGFEEGK	0.60	3.40E-03	None
IspG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	P62620	RIDVQQVEK	0.60	9.75E-03	None
IspG	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	P62620	SGLYEDGVRK	0.60	6.89E-03	None
Efpp	Elongation factor P	P0A6N4	SGEYVSR	0.60	1.23E-03	None
FtsZ	Cell division protein FtsZ	P0A9A6	VVNDNAPQTAK	0.60	5.66E-04	None
GshA	Glutamate--cysteine ligase	P0A6W9	TPSEEYAK	0.59	3.59E-04	None
PurA	Adenylosuccinate synthetase	P0A7D4	AEAVDYQK	0.59	8.08E-04	None
UxaC	Uronate isomerase	P0A8G3	LCTGDASDREK	0.59	2.76E-03	None
GlyA	Serine hydroxymethyltransferase	P0A825	ANITVNK	0.59	4.24E-03	None
Kbl	2-amino-3-ketobutyrate coenzyme A ligase	P0AB77	FICGTQDSHK	0.59	9.02E-03	None
GltX	Glutamate--tRNA ligase	P04805	SASAFNTDK	0.59	1.98E-03	None
YeiP	Elongation factor P-like protein	P0A6N8	GAATLYK	0.59	9.58E-04	None
FbaA	Fructose-bisphosphate aldolase class 2	P0AB71	DSQEYVSK	0.58	1.97E-03	None
AsnA	Aspartate--ammonia ligase	P00963	ALRPDEDR	0.58	8.68E-03	None
TktA	Transketolase 1	P27302	ESAWNEK	0.56	6.21E-04	None
GlpK	Glycerol kinase	P0A6F3	TAEICEHLK	0.56	9.70E-03	None
YjbJ	Putative stress response protein	P68206	DEAGGN**W**K	0.56	7.62E-04	1 Oxidation
PmbA	Metalloprotease PmbA	P0AFK0	AEQAALQADKR	0.56	7.62E-03	None
SdhA	Succinate dehydrogenase flavoprotein subunit	P0AC41	TGHALLHTLYQQNLK	0.56	3.61E-03	None
GpmA	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	P62707	YGDEQVK	0.55	9.33E-04	None
Icd	Isocitrate dehydrogenase	P08200	KITLQNGK	0.55	8.52E-03	None
AldA	Lactaldehyde dehydrogenase	P25553	IPDGQAEDARK	0.55	7.46E-03	None
ArcA	Aerobic respiration control protein ArcA	P0A9Q1	SVESYK	0.55	4.52E-03	None
FabD	Malonyl CoA-acyl carrier protein transacylase	P0AAI9	VWQQQGGK	0.55	4.38E-03	None
FabB	3-oxoacyl-[acyl-carrier-protein] synthase 1	P0A953	YNDTPEK	0.55	3.81E-03	None
AspC	Aspartate aminotransferase	P00509	GSALINDKR	0.54	7.79E-03	None
Agp	Glucose-1-phosphatase	P19926	SDQQWK	0.53	6.39E-03	None
Frr	Ribosome-recycling factor	P0A805	KIEAALADK	0.53	8.76E-03	None
RimO	Ribosomal protein S12 methylthiotransferase Rim	P0AEI4	RLVDAGVK	0.52	9.32E-03	None
PflA	Pyruvate formate-lyase 1-activating enzyme	P0A9N4	DTWDTHGGK	0.52	1.14E-03	None
YjbJ	Putative stress response protein	P68206	DEAGGNWK	0.51	4.00E-04	None
SecA	Protein translocase subunit SecA	P10408	HNVLNAK	0.50	6.58E-03	None
RpoC	DNA-directed RNA polymerase subunit beta'	P0A8T7	DTTVGR	0.49	1.34E-03	None
YaaA	Peroxide stress response protein.	P0A8I3	TLDYQSPLTTTR	0.32	2.65E-10	None

Table 7.4. Detected peptides with a fold change value less than 0.6. All peptides with fold changes less than 0.6 due to irradiation are listed. The protein they map to and the peptide sequence are also indicated. Where an oxidative modification occurs, the amino acid residue is bolded. If the residue is also underlined, that indicates that the modification has been identified on that particular residue; if the residue is only bolded then that residue is only a candidate for where the modification occurs. Figure was adapted from work generously provided by Dr. Benjamin Minkoff (Michael Sussman Lab, University of Wisconsin – Madison).

Amino acid	Observed primary modifications	Observed secondary modifications	Previously reported primary modifications	Previously reported secondary modifications
cystine	[+48] sulfonic acid	[+32] sulfinic acid	[+48]	[+32]
* His	[+32] / [+5] ring-open	[-28] / [-22] / [-10] / [+8] / [+16] / [+35] / [+48]	[+16] / [-22] / [-10] / [+5]	[-23] plus intermediates[a]
Met	[+16] sulfoxide-	none	[+16]	[+32] / [-14] / [-16] / [-30] / [-32]
* Trp	[+32] dihydroxy-	[+16n, n=1-9] / [+18n, n=1-9] / [+20n, n=4-9]	[+16] / [+32]	[+4] / [+16n, n=1-5]
* Tyr	[+16] hydroxy-	[+16n, n=1-7] / [+18n, n=1-7] / [+20n, n=4-7]	[+16]	[+32]
* Phe	[+16] hydroxy-	[+16n, n=1-7] / [+18n, n=3-7] / [+20n, n=4-7]	[+16]	[+32]
* Lys	[+32] dihydroxy-	[+16] / [+14] / [+48]	[+14]	[+16]
* Arg	[+32] dihydroxy-	[+16] / [-61] / [-43] / [+14]	[-43] deguanidination	[+16] / [+14]
* Leu	[+32] dihydroxy-	[+16] / [+14]	[+16]	[+14] / [+32]
* Ile	[+14] carbonyl-	[+16] / [+32]	[+16]	[+14] / [+32]
* Val	[+32] dihydroxy-	[+16] / [+14]	[+16]	[+14]
* Ser	[-28] (?)	[+16] / [-2]	[+16] / [-2] trivial amounts[a]	none
* Thr	[-46] (?)	[+16] / [+32]	[+16] / [-2] trivial amounts[a]	none
* Pro	[+32] dihydroxy-	[+16] / [+14]	[+16] / [+14]	[+32]
* Glu	[-14] (?)	[-30] / [+32] / [+16] / [+14]	[-30]	[+16] / [+14]
* Gln	[+32] dihydroxy-	[+16] / [+14]	[+16]	[+14]
* Asp	[+137] (dimer-like?)	[-30] decarboxylation	[-30]	[+16] / [+14]
Asn	[+16] hydroxy-	[+32]	[+16]	none
* Ala	[+32] dihydroxy-	[+16]	[+16]	none

Table 8.4. Detected modifications on L-amino acids. Amino acids with oxidative modifications detected in our work but previously unreported are indicated by an '*'. Modifications are reported as mass differences in Daltons. Major products include +14 (carbonylation), +16 (oxidation) and +32 (peroxidation). Previously reported modifications are in a large part drawn from a) Xu and Chance, 2007 [24]. Table was generously provided by Dr. Grzegorz 'Greg' Sabat (Mass Spectrometry Facility, Biotechnology Center, University of Wisconsin – Madison).

4.6 Figures

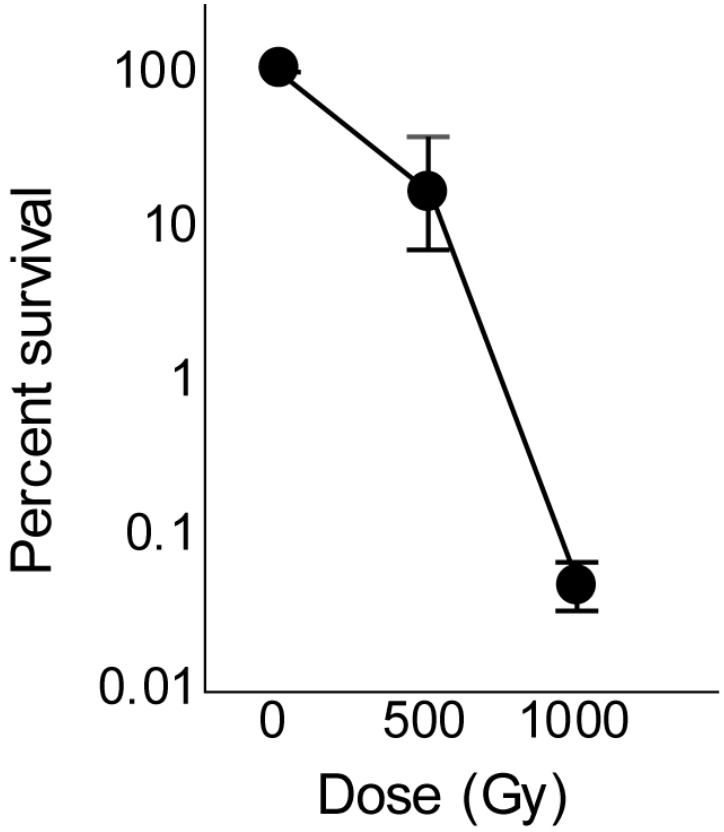


Figure 1.4. Survival curve of *E. coli* MG1655 exposed to high-energy electron beam IR. *E. coli* cultures were subjected to 500 and 1000 Gy of irradiation from a 21EX Varian clinical linear accelerator (Linac) as described in the *Materials and Methods* section. Error bars represent standard deviation of percent survival values determined from an experiment performed in biological triplicate.

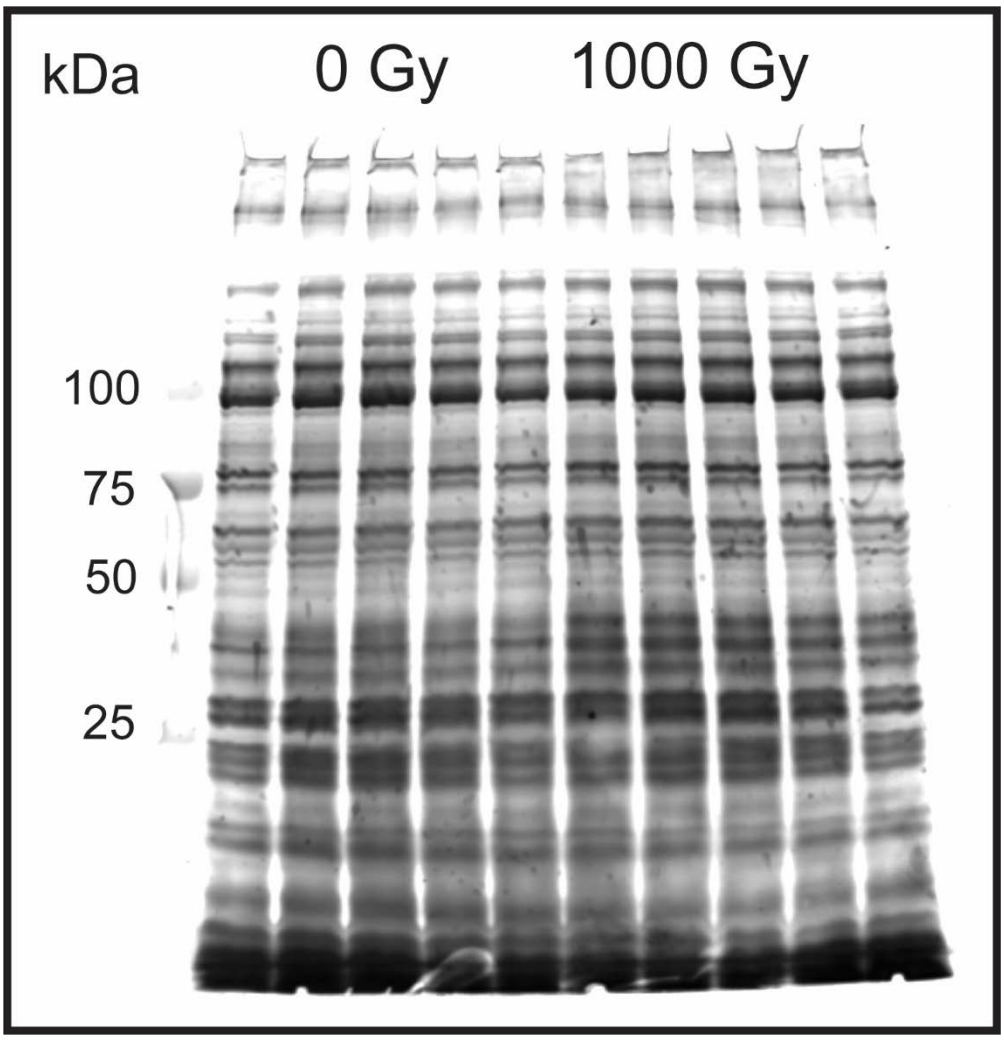


Figure 2.4. Effect of 1000 Gy of IR on the *E. coli* proteome. Quintuplicate replicates of *E. coli* were exposed to 1000 Gy of IR before cells were lysed and run on a SDS-PAGE gel as described in the *Materials and Methods* section. The majority of protein banding appears to be unaffected by irradiation.

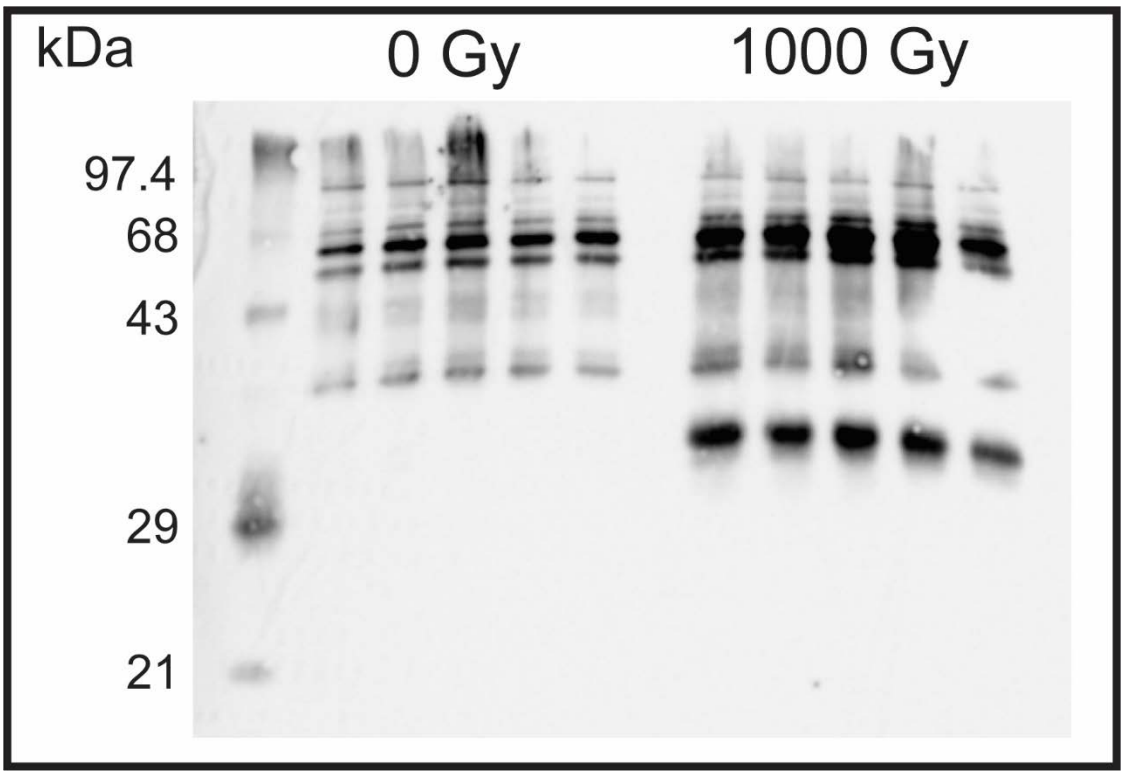


Figure 3.4. Detection of protein carbonylation due to IR exposure. Quintuplicate replicates of *E. coli* were exposed to 1000 Gy of IR before cells were lysed processed for Western blot detection of protein carbonylations as described in the *Materials and Methods* section. Protein carbonylation is an irreversible oxidative modification caused by exposure to IR. Only a single unique band, indicating a carbonylated protein, is observed in the irradiated samples.

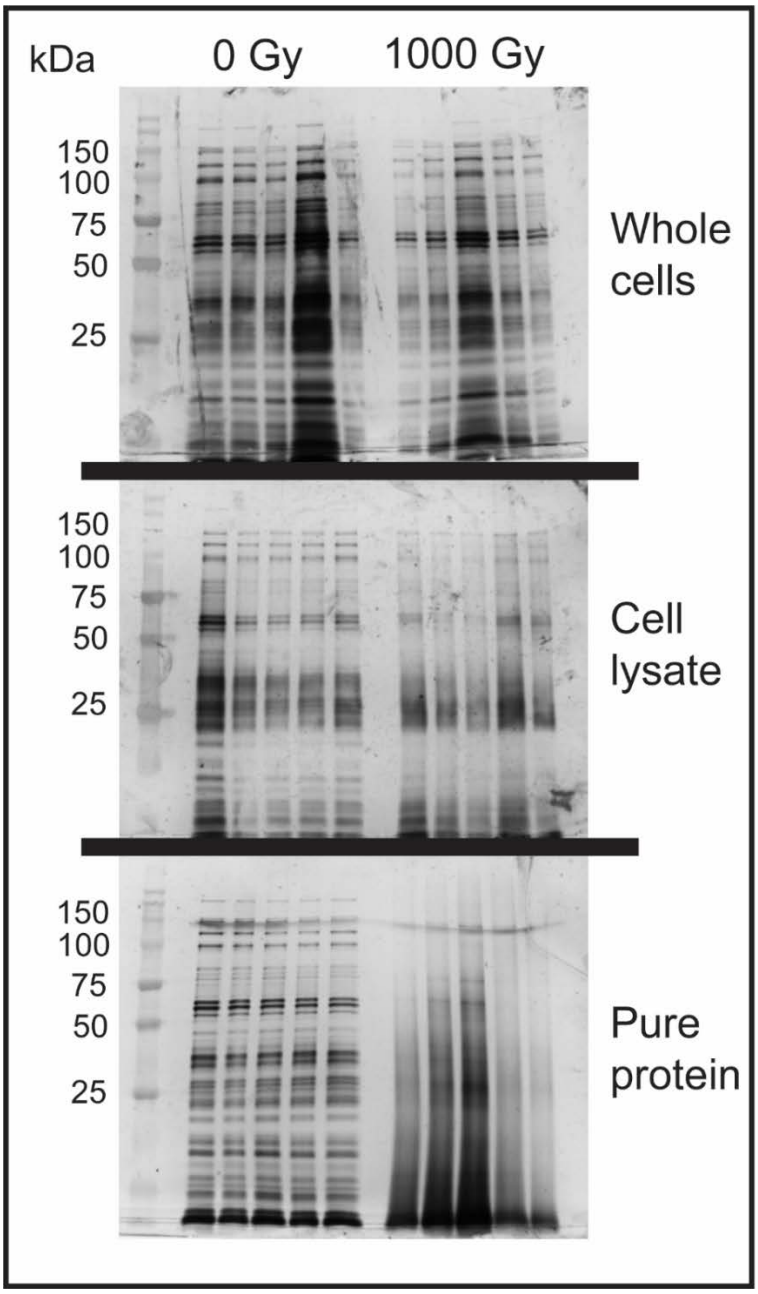


Figure 4.4. Effect of 1000 Gy of IR on the *E. coli* proteome in whole cells, cell lysate, and purified *E. coli* protein. Samples of *E. coli* whole cells, cell lysates, and pure proteins were prepared and irradiated as described in the *Materials and Methods* section. Protection of the proteome from IR induced degradation is evident in whole cell, and to a lesser extent cell lysate, samples. Purified protein samples experience heavy degradation due to irradiation.

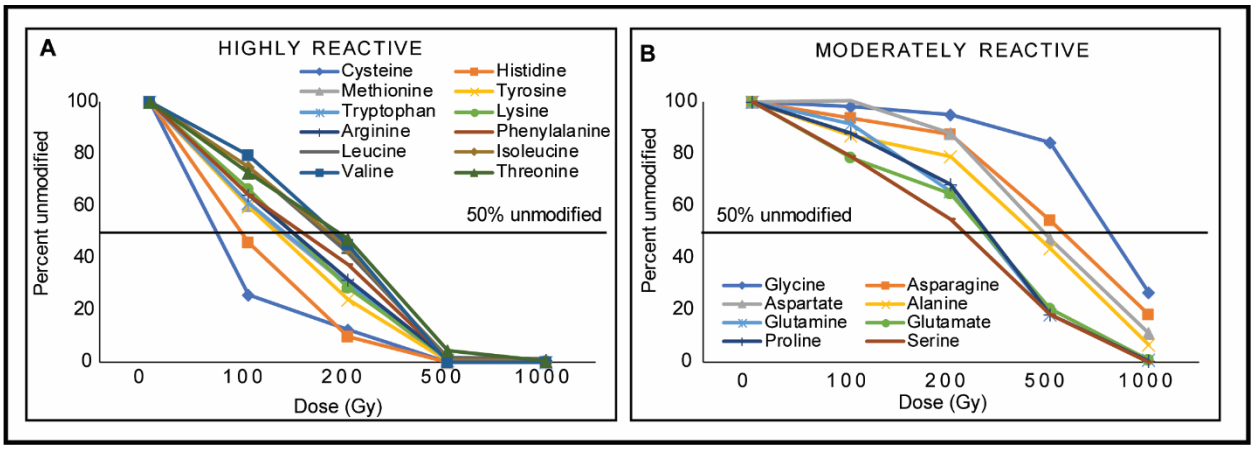


Figure 5.4. Reactivity of L-amino acids with IR. A). Highly reactive amino acids as defined by less than 50% of free amino acids are unmodified at a dose of 200 Gy. B). Moderately reactive amino acids as defined by greater than 50% of free amino acids are unmodified at a dose of 200 Gy. All amino acids were irradiated in a volume of 1 mL at a concentration of 50 µM as described in the *Materials and Methods* section. Figure was generously provided by Dr. Grzegorz 'Greg' Sabat (Mass Spectrometry Facility, Biotechnology Center, University of Wisconsin – Madison).

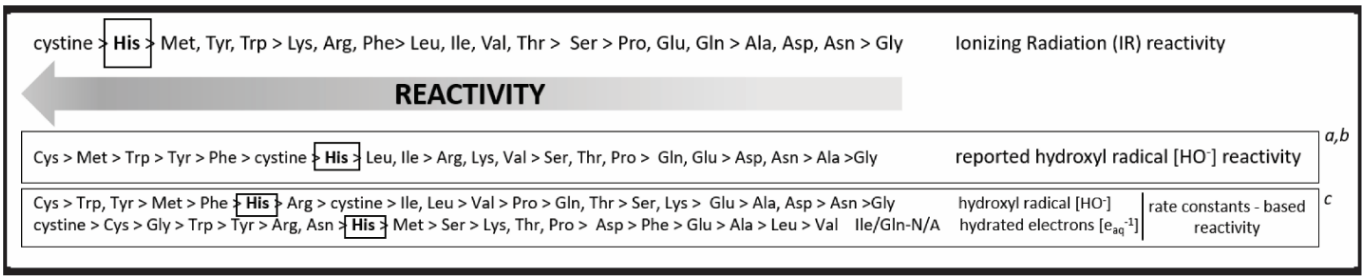


Figure 6.4. Summary of reactivity of amino acids with IR. The top row of amino acids represents data generated in this study, and the bottom rows are derived from the indicated sources a)[145], b)[24], c)[146]. The position of histidine is indicated in each row, to highlight that the reactivity of His varies across studies. Figure was generously provided by Dr. Grzegorz 'Greg' Sabat (Mass Spectrometry Facility, Biotechnology Center, University of Wisconsin – Madison).

Chapter V: Conclusion

The ability of a cell to survive exposure to ionizing radiation (IR) is incredible. IR can modify and fragment all cellular macromolecules through ionization and oxidation, rendering cells with much to repair and with a hampered ability to do so [2, 8, 18, 26]. Despite this, all organisms encode in their genome the machinery capable of dealing with damage caused by IR, though the extent of the damage that an organism can withstand varies wildly. Amazingly, resistance to high doses of IR does occur in nature and the study of these organisms has expanded views of the extreme conditions that life can endure [1, 6, 66, 67].

IR itself is complex. Electromagnetic and particulate IR differ in the ability to penetrate matter, how energy is deposited, and the sources capable of producing each [1]. As such, a cell's ability to survive IR exposure may depend on the conditions of the exposure as much as its complement of DNA and protein repair machinery.

The biological damage caused by IR affects DNA, RNA, proteins, and lipids [2]. The importance of damage to each of these molecules differs by the organism and is a subject of debate in radiation biology. It is likely that damage to each of these components plays a role in whether a cell will live or die due to IR exposure. However, it is true for all life that a single DNA double-strand break (DSB) prevents an organisms' ability to faithfully perform DNA replication and is therefore lethal if left unrepaired [1]. DNA DSBs accumulate in cells with prolonged exposure to IR, and as such DSB repair remains a crucial step to surviving irradiation.

Though DNA repair is crucial for surviving IR exposure, research into how cells survive IR must be as multi-faceted as the damage caused by IR. Each approach to studying IR resistance offers new insights into mechanisms of repair and avoidance of IR-induced damage, and therefore

increases the toolbox for ameliorating IR-induced trauma. As such, this thesis work has focused on furthering a novel approach to understanding IR-resistance: generating radioresistant organisms via experimental evolution.

The model bacterium *Escherichia coli* has unremarkable resistance to IR [6]. Much of our understanding of DNA repair and reactive oxygen species (ROS) amelioration stems from research on this organism; however, despite an impressive array of repair pathways (that are still yielding surprises and keeping secrets), a dose of 500 Gy will kill 90% of an *E. coli* MG1655 culture [6]. This starkly contrasts with the radioresistant organism *D. radiodurans*, which can survive a 10-fold higher dose without any loss in viability [1, 6, 56]. This comparison should not diminish what *E. coli* is capable of; it is important to remember predictions of DSB accumulation rate indicates that at a dose of 500 Gy, the surviving *E. coli* are repairing at least 7 DSBs in order to survive. This ignores base modifications, blocks to DNA replication, and damage to protein and the cell membrane that must additionally be dealt with [18].

A latent ability for *E. coli* to withstand extreme doses of IR has become clear throughout directed evolution of *E. coli* for IR resistance [13, 90]. We have evolved two sets of four replicate populations through iterative cycles of irradiation to kill 99% of each population with subsequent outgrowth of the survivors. Taken together, we thus far have a catalogue of six point-mutations known to enhance IR resistance of *E. coli*, and many more mutations which appear in the same proteins but have yet to be validated. These mutations span DNA repair proteins (RecA and RecD), DNA helicases (DnaB and YfjK), a transcription factor involved in anaerobic respiration (FNR) and a subunit of RNA polymerase (RpoB). Though the downstream effects of these mutations are unclear, it is telling that three of these mutations are involved in DNA repair pathways (RecA, RecD, and DnaB), and all of these proteins are either DNA-binding or are components of DNA-

binding complexes. Thus far, experimentally evolved IR-resistance appears to revolve around genome maintenance.

It has been argued that protection of the proteome from IR-induced ROS is the primary determinant of survival of irradiation [6, 16, 20]. In particular, if DNA repair proteins are damaged by IR, then the cell will not be able to then repair DNA damage. Our collaborative work with Professor Michael Sussman (University of Wisconsin – Madison) has been our first step determining the impact of IR-induced protein damage on cell survival. Our initial experiments have suggested that, at doses which kill 99 % of cells, components of the cell cytoplasm are capable of protecting the proteome from significant degradation and may ameliorate oxidative modification of proteins. However, mass spectrometry experiments have confirmed *in vivo* that oxidative protein modification does increase with irradiation, and that these modifications occur primarily on high abundance proteins. Although protein carbonylation is considered to be irreversible and inactivating [2, 28], most of the modifications detected are oxidation events on methionine residues, for which the cell does have the ability to repair. However, irradiation may overwhelm a cell's ability to reduce oxidized proteins. Additionally, oxidative modification of abundant, but critical, proteins involved in translation, transcription, and central glucose metabolism will likely inhibit recovery of a cell post-irradiation.

An outstanding question remains: what is the determinant of cell death due to irradiation? Our evolving *E. coli* populations suggest that enhanced DNA repair significantly increases IR resistance beyond the point where viable cells of the parent strain can be isolated. This observation suggests that at doses relevant to this thesis work (up to 4000 Gy), DNA damage, not protein damage, is the primary determinant of IR-induced cell death. However, there must be a point at

which the damage to translational, transcriptional, and metabolic machinery that we have observed does not allow for cell recovery.

To end this thesis document, I would like to propose a model for IR-induced cell death, that may serve as a testable hypothesis for our evolving populations. For each organism, there must be an amount of DNA and protein damage that can be tolerated. Since inactivating mutations within essential genes (which make up a minimum of 5% of the genome of *E. coli*, depending on growth condition [148]) and unrepaired DSBs are lethal, compounded by the fact that the genome exists in fewer copies than essential proteins, potentially lethal DNA damage is likely less tolerated than protein damage. Therefore, initial cell death due to irradiation is may result from DNA damage surpassing the amount of damage a cell can repair. Any shoulder of resistance, where little to no cell death occurs with increasing dose, reflects the DNA repair capability of that organism. Once DNA repair capacity is overwhelmed, percent survival of a population of cells will decrease until all cells are rendered non-viable; unless the organism is equipped with a unique ability to maintain proteome integrity. At extreme doses (a dose at which all cells within a culture would be rendered non-viable if survival relied solely on DNA repair capability), the proteome of a cell may provide a second chance for cell survival. If key components of the proteome, such as transcription and translation machinery, can remain active, cells may retain an ability to slowly recover from irradiation despite extensive DNA damage. This ability to recover would likely depend on the integrity of the genome at this dose point; cells with extensive genome degradation would be unable to recover despite an intact proteome. Therefore, proteome integrity comprises a secondary, but important, factor in governing cell survival from irradiation. The ability to protect the proteome from extensive damage makes a second shoulder in the survival curve, and once protein damage

is overwhelming, the last remaining option is cell death. See Figure 1.5 for a depiction of this model.

D. radiodurans appears to fit this model well. This bacterium has been previously reported to have an impressive initial shoulder of IR resistance out to nearly 10 kGy [6, 56]. Deletion of key DNA repair proteins significantly shortens this shoulder (such that the dose at which 10 percent of cells survive is 1000 to 6000 Gy), but a percentage of cells still survives at doses of 10 kGy [56]. Conversely, growing cells in medium depleted of Mn^{2+} (Mn^{2+} is implicated in forming ROS-ameliorating complexes capable of protecting the *D. radiodurans* proteome from oxidative damage) affects survival of cells at extreme dose but not the survival of cells at lower doses (the dose at which 10 percent of cells survive is reduced from 16 kGy to 10 kGy without Mn^{2+} present) [19]. Interestingly, when irradiated with a high-energy electron beam, *D. radiodurans* is far more radiosensitive than previously reported (99.9% of cells are killed by a dose of 6000 Gy), and a survival curve with two shoulders of resistance is apparent (the increased killing by the electron beam may exhaust DNA repair and ROS amelioration capabilities at low enough doses to fully observe the survival curve of *D. radiodurans* without experimentally perturbing one or both mechanisms of IR resistance as done previously).

Thus far, our evolved *E. coli* populations appear to be primarily evolving mechanisms of enhanced or altered DNA repair. If DNA repair is the initial determinant of survival of IR exposure as the model I present here suggests, beneficial mutations which enhance this process would provide (relatively) easy routes to increased fitness and would therefore be selected for often. Fixation of few mutations which confer large increases in fitness (as opposed to many mutations which contribute little) would enhanced by a strong selection pressure (such as killing 99% of a population at each round of selection). As mutations affected DNA repair accumulate, the initial

shoulder of resistance of these populations would lengthen, as they have. This model predicts that there will come a point in our evolving populations where diminishing returns from DNA repair mutations leads to fitness gains lower than those conferred by mutations that would enhance ROS amelioration/proteome protection and a second shoulder of resistance will appear, allowing survival of extremely high doses of irradiation. It may be only a matter of time for our evolving populations to prove or disprove this model of IR-induced cell death.

5.1 Figures

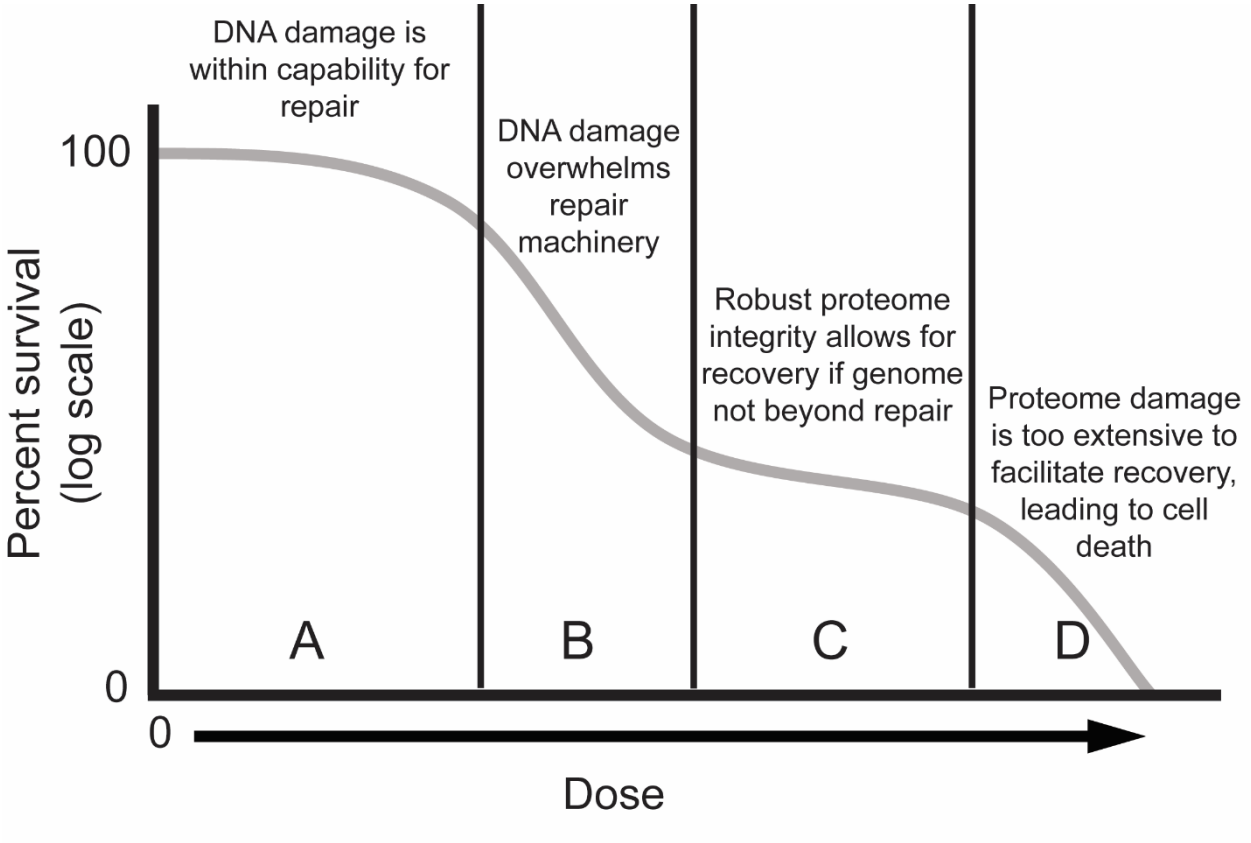


Figure 1.5. Proposed model for IR-induced cell death. A) IR induced DNA damage will kill little to no cells of an irradiated population if the average damage induced is within the DNA repair capacity of the organism. B) Once the average amount of DNA damage caused by IR exceeds the cells ability to repair that damage, cell death rapidly increases with IR dose. C) In a percentage of cells within the population, DNA damage exceeds normal DNA repair capacity but is not irreparable. In these cells, protection of key components of the proteome, such as transcriptional and translation machinery, may allow slow recovery from doses received in this range. D) At extremely high doses, relative to a given organism, DNA and proteins will be damaged too extensively to facilitate recovery and cell death with result.

Chapter VI: *Materials and Methods*

6.1 Growth conditions and bacterial strains used in this study

Unless otherwise stated, *E. coli* cultures were grown in Luria-Bertani (LB) broth [149] at 37°C with aeration. *E. coli* were plated on 1.5% LB agar medium [149]and incubated at 37°C. Overnight cultures were grown in a volume of 3 mL for 16 to 18 hr. Exponential phase cultures were routinely diluted 1:100 in 10 mL of LB medium in a 50 mL Erlenmeyer flask and were grown at 37°C with shaking at 200 rpm and were harvested at an OD_{600} of 0.2, unless otherwise noted. After growth to an OD_{600} of 0.2, cultures were placed on ice for 10 min to stop growth before being used for assays.

Cultures were plated on tetrazolium agar (TA) for growth competition assays when noted [121]. The defined rich medium EZ was mixed per manufacturer's specifications and was supplemented with 0.2% glucose or 0.2% glycerol as indicated (Teknova; Hollister, CA). M9 minimal medium (supplemented with 0.2% glucose) was used when indicated [149].

All strains used for *in vivo* assays in this study are mutants of *E. coli* K-12 derivative MG1655. Genetic manipulations to transfer mutations or delete genes were performed as previously described [150, 151].

6.2 Serial dilutions and CFU/mL determination

All serial dilutions were performed in 1X phosphate-buffered saline (PBS) (for 1 L: 8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, KH_2PO_4 0.24 g with 800 mL dH_2O, adjust pH with HCl to 7.4, then add remaining 200 mL dH_2O**).** Unless otherwise stated, serial dilutions were performed with serial 1:10 dilutions of 100 µL of culture or previous dilution into 900 µL 1X PBS. Before transfer

to the next dilution tube, samples were vortexed for 2 seconds and mixed by pipetting to ensure mixing. One-hundred µL of appropriate dilutions were aliquoted onto agar plates of the appropriate medium and were spread-plated utilizing an ethanol-sterilized, bent glass rod. For spot plating, 10 µL of each dilution was aliquoted onto agar plates of the appropriate medium and spots were allowed to dry before plates were incubated as in *Growth conditions*.

CFU/mL was calculated using the highest CFU count for each strain assayed that remained between 30 and 300 CFU (ex: 250 CFU on a 10^{-4} dilution plate would be used for calculation over 40 CFU on a 10^{-5} dilution plate).

6.3 Lambda transductions

Lambda transductions to generate P_{narG} – *lacZ* strains from parent strain RZ7350 [106] were carried out as follows. Incubated 10 mL of culture of the parent strain and the strains to be lysogenized overnight for approximately 18 hr at 37°C. To prepare the phage, added 300 µL to the turbid culture of the parent strain and shook at 200 rpm for 10 min at 37°C. Centrifuging for 10 min at 4700 xg pelleted the lysed cells, and the phage-containing supernatant was transferred to a new glass test tube. Five-hundred µL of the recipient strains were mixed with 6 mL of melted, 45 °C maximum, top agar and poured onto LB agar supplemented with 10 mM $MgSO_4$ and 40 ug/mL X-gal. Once the top agar solidified, a 10 uL drop of prepared lambda phage was placed onto each plate. After the spot had dried, the plates were incubated overnight at 37°C. Blue plaques were streaked to isolation onto LB supplemented with 40 ug/mL X-gal until uniform, blue colonies were isolated. Lambda lysogens were confirmed as previously described [152].

6.4 Deep sequencing of individual populations IR-1-20, IR-2-20, IR-3-20, and IR-4-20

For each population, three replicate cultures were grown overnight, and genomic DNA was prepared from each using a Wizard Genomic DNA Purification Kit (Promega; Madison, WI). Samples were sequenced with an Illumina MiSeq instrument (Quintara Biosciences; Richmond, CA**)**. IR-3-20 samples were sequenced with one flow cell. For IR-1-20, IR-2-20 and IR-4-20 samples, data from two flowcells were pooled together for each population because the second read files from the first run had low quality reads.

BioInfoRx, Inc. (Madison, WI), analyzed the raw sequencing reads. Briefly, Illumina MiSeq reads (2X250 bp) were trimmed to only keep high quality (>Q20) reads using trimmomatic [153], then mapped to *E. coli* MG1655 genome (Genbank accession number NC_000913.2) using Burrows-Wheeler Aligner [154]. Duplicated reads were removed using Picard Tools (http://broadinstitute.github.io/picard/index.html). Variants were called using samtools [155] and VarScan2 [156]. and the minimum variant frequency was set at 1%. The variants were annotated with SnpEff [157]. The variant files were further classified into tiers using custom Perl scripts. Briefly, tiers were separated by frequency of a variant within a population, the number of samples for each population the variant was detected in, and if the variant was detected on both DNA strands. Variants detected at a frequency greater than 1%, on both DNA strands, and in all the triplicate samples for each population were considered as reliable and not subjected to further verification.

6.5 Deep sequencing of evolving lineages IR9, IR10, IR11 and IR12, and isolates

Genomic DNA was prepped from overnight cultures prepared from frozen stocks of populations from every even round of selection using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI). DNA samples were submitted to the Department of Energy Joint Genome Institute (Walnut Creek, CA) for sequencing and analysis. DNA was randomly sheared into ~500bp fragments and the resulting fragments were used to create an Illumina library. This library was sequenced on Illumina HiSeq generating 100bp paired end reads. Reads were aligned to the reference genome using BWA [154] and putative SNPs and small indels were called using callvariants.sh from BBMap (sourceforge.net/projects/bbmap).

6.6 Ionizing radiation resistance assay using γ-ray sources

Cells from a single colony of each strain were cultured overnight and then grown to an OD_{600} of ~ 0.2 as in *Growth conditions (6.1).* One mL aliquots in 1.8 mL Eppendorf tubes were mixed by vortexing for 2 s and a 100 µL aliquot was removed and added to 900 µL PBS on ice as an initial 1:10 dilution for the non-irradiated control. Undiluted samples were then irradiated in a Mark I ^{137}Cs irradiator (J. L. Shepherd and Associates) for a time corresponding to 3 kGy (~ 6.5 Gy/min). Irradiated samples as well as the non-irradiated samples were serial diluted and plated using typical methods. CFU/mL was determined for irradiated and un-irradiated samples, and CFU/mL was determined. Initial cell densities ranged from 1 - 6 x 10^7 CFU/ml. For each experiment, each individual strain was tested in biological triplicate.

6.7 Generalized Linac irradiation protocol

Samples were maintained at 4 °C and taken to a Varian 21EX clinical linear accelerator (Linac) for irradiation (approximately 15 min transit time to and from the Linac). The Linac was set to 6 MeV HDTSe^{-} mode (a special high dose rate mode with 888 MU [monitor units]/minute) and tubes containing the samples were submerged in ice-water at a source-to-surface distance (SSD) of 61.7 cm and at a depth of 1.3 cm (measured to the center of the tube's volume). A 30x30 cm Linac field size setting gave an effective field size at this SSD of 18.5x18.5 cm. Samples were irradiated at 8.128 cGy/MU (approximately 72 Gy/min; verified using the TG-51 protocol with a field size of 30x30 cm [no electron cone] at SSD = 100 cm) until the desired dose was achieved.

6.8 Linac photon mode

In order to determine if there are any effective differences between electron and photon modes during irradiations, the SSD for photons was found that gave the same dose to the samples as in electron mode. In photon mode, the electron beam first strikes a flattening filter (not present in electron mode) to produce bremsstrahlung photons and also flatten the beam intensity profile.

6.9 Ionizing radiation resistance assay using Linac

Strains were grown in biological triplicate overnight and to an OD_{600} of approximately 0.2 in LB as routinely performed. One mL sample for each dose tested (including 0 Gy) were removed and aliquoted into sterile 1.5 mL microfuge tubes. Samples were pelleted by centrifugation at 13 xg for 1 min, and the supernatant was poured off. Samples were resuspended in 1 mL ice-cold 1X phosphate-buffered saline (PBS), and pelleting was repeated. This process was repeated three more to wash cells. A 100µL aliquot of each culture was removed, serial diluted 1:10 in 900µL of PBS

to a final 10,000-fold dilution and 100µL was plated on LB agar to determine the colony forming units (CFU)/mL before irradiation. Samples were maintained at 4 °C and irradiated with the appropriate doses as described. A 100µL aliquot of each culture was removed and plated to determine CFU/mL and percent survival as described.

6.10 Directed evolution protocol using Linac

For each round of directed evolution, separate aliquots of 2 mL of LB medium was inoculated with frozen stock of each population from the previous round of selection. These were incubated overnight with aeration at 37 °C and were grown with usual practices in LB medium to an OD_{600} of 0.2 the next day. Each culture was incubated on ice for 10 minutes to stop growth. Three 1 mL samples were removed and aliquoted into sterile 1.5 mL microfuge tubes. Samples were washed three times with 1 mL ice-cold 1X phosphate-buffered saline (PBS) and resuspended in a final volume 1 mL of 1X PBS. A 100µL aliquot of each culture was removed, serial diluted 1:10 in 900µL of PBS to a final 10,000-fold dilution and 100µL was plated on LB agar to determine the colony forming units (CFU)/mL before irradiation. Samples were maintained at 4 °C and taken to a Varian 21EX clinical linear accelerator (Linac) for irradiation.

After irradiation, an aliquot of each culture was removed, serial diluted 1:10 in 900 µL of PBS to a final 1,000-fold dilution, and 100 µL of each serial dilution was plated on LB agar for each dose to determine the CFU/mL after irradiation. LB agar plates were incubated overnight at 37 °C. Remaining irradiated cultures pelleted by centrifugation at 13 xg and supernatant was discarded. These pellets were resuspended in 1 mL of fresh LB medium, and this was added to 1 mL LB medium in a 5 mL glass culture tube. These resuspensions were incubated overnight with aeration at 37 °C. The following day, the percent survival for each dose was calculated using

CFU/mL calculations before and after irradiation at each dose. The overnight culture of each population replicate showing closest to 1% survival was stored at -80 °C and used for the next cycle of selection.

The initiating round of selection was done as described above, except the original culture used was an overnight culture of MG1655 prepared from an isolated colony.

Protocol was adapted from a previously used protocol [13].

6.11 Irradiation of amino acids

Stocks of amino acids were made at a concentration of 500 µM in 10 mL dH_2O. Stocks were kept at -20 °C and thawed when necessary. The 500 µM stocks were diluted 1:10 in 1 mL total dH_20 in 1.8 mL Eppendorf tubes for each sample to be irradiated. Samples were maintained at 4 °C and taken to a Varian 21EX clinical linear accelerator (Linac) for irradiation (~15 min duration each way). After irradiation, samples were frozen at -20 °C until needed. The following amino acids were used: L-Isoleucine (Product #: I2752; Sigma-Aldrich), L-Proline (Product #: P0380; Sigma-Aldrich), and the remaining amino acids came from a L-amino acids kit (Product #: LAA-21; Sigma-Aldrich).

6.12 UV resistance assay

Cells from a single colony of each strain were cultured overnight and then grown to an OD_{600} of ~ 0.2 as in *Growth conditions (6.1).* Samples were diluted and spotted onto 25 mL 1.5% LB agar described in the *Serial dilutions (6.2).* Spots were dried before the plate lid was removed and spots were exposed to the appropriate dose of UV irradiation using a Spectrolinker XL-1000

UV Crosslinker (Spectronics Corporation, Westbury,NY). Plates were imaged after incubation for 24 hr.

6.13 Resistance to chemical DNA-damaging agents

Cells were grown overnight and to an OD_{600} of approximately 0.2 in LB as routinely performed. Samples were mixed by vortexing for 5 s and were serial diluted 1:10 in 900 µl phosphate-buffered saline (PBS) to a final 100,0000-fold dilution. Ten µl was removed from each dilution and spotted onto 30mL 1.5% LB agar medium supplemented with 10 or 7.5 ng/ml ciprofloxacin hydrochloride as specified**,** 0.5 µM bleomycin, 4 µg/mL mitomycin C, or 5 mM hydroxyurea. Spots were dried before being incubated overnight at 37 °C. Plates were imaged after 48 hr for ciprofloxacin and bleomycin plates.

6.14 H_2O_2 resistance assay

This assay was adapted from a previously described protocol [126]. Cells from a single colony of each strain were cultured overnight at 37 °C with aeration. After growth overnight, 10 mL of LB broth in a 50 mL Erlenmeyer flask was inoculated with 70 µL of culture and grown at 37 °C with shaking for 7 hr in order to test stationary phase cells. Each culture was incubated on ice for at least 5 min. Samples were mixed by vortexing for 5 s and were serial diluted 1:10 in 900 µL phosphate-buffered saline (PBS) to a final dilution of 1:1000. These were then diluted 1:10 into two separate 12 mL conical tubes containing 3.6 mL PBS. One conical was designated as the experimental sample and 10 mM H_2O_2 (Fisher Scientific, Fair Lawn, NJ Cat#: H325) was added to this sample. Experimental and control samples were then incubated at 37 °C with shaking for 30 min. After incubation, 1000 U/mL catalase (MP Biomedicals, Solon, OH Cat#: 100429) was

then added to the experimental samples, which were mixed by vortexing for 5 s. To compensate for extra volume from H_2O_2 and catalase, an equal volume of PBS was then added to control samples. Samples were then serial diluted in 900 µL PBS as necessary. Dilutions were then plated onto 15% LB agar medium to determine the total number of CFUs. Percent survival was calculated by dividing the titer of the surviving H_2O_2-treated sample by the titer of the control sample. For each experiment, each individual strain was tested in biological triplicate.

6.15 ROS amelioration assay

Cells from a single colony of each strain were cultured overnight and then grown to an OD_{600} of ~ 0.2 as in *Growth conditions (6.1).* Each culture was incubated on ice for at least 5 min to stop growth. One mL of each culture was pelleted via centrifugation at 17000 xg for 1 min, supernatant was removed, and cells were resuspended in 1 mL M9 medium supplemented with 2 mg/mL L-Arginine (Sigma-Aldrich; St. Louis, MO; Cat#: A6969) (CB3000 is Arg-, therefore all media was supplemented with this amino acid). Ten µL of a 10 mM stock of H_2O_2 (Fisher Scientific; Fair Lawn, NJ; Cat#: H325) was added to appropriate wells of a black, clear-bottom 96-well plate. Ten µL of PBS were added to separate wells to be assayed as a no H_2O_2 control. Ninety µL of each strain was aliquoted into appropriate wells, then mixed with 100 uL of M9 + 10 mM 2',7'-dichlorodihydrofluorescein diacetate (H_2DCFDA) (Fisher Scientific; Fair Lawn, NJ; Cat#: D399). These were then placed in a Biotek Synergy 2 plate reader warmed to 37 °C. The plate was shaken continuously and a reading (OD_{600}, and fluorescence with excitation at 485/20 and emission at 510/20) was taken automatically every 5 min for 20 to 24 hr by the instrument. Blank readings (200 µL M9 medium) of fluorescence were divided by the blank OD_{600} reading.

Readings of experimental wells were treated in the same manner, and then the blank reading was subtracted.

6.16 β-galactosidase assay

Cultures to be assayed were prepared by incubating cells from a single colony of each strain overnight at 37 °C with aeration. This resulting overnight culture (grown 15-18 hr) was treated as the stationary phase culture to be assayed. Exponential phase cultures were grown by inoculating 10 mL of LB broth in a 50 mL Erlenmeyer flask with 70 µL of overnight culture and grown at 37 °C with shaking to an OD_{600} 0.2. These cultures were placed on ice for at least 5 minutes to stop growth before use. To perform a mock irradiation, two separate aliquots of 900 µL of exponential phase cultures in 1.5 mL Eppendorf tubes were inverted and incubated in the dark at room temperature (~24 °C) for 8 hr. Two aliquots for each replicate ensures that there is sufficient volume for 1 mL of culture to determine β-galactosidase activity and 100 µL to determine OD_{600}.

The β-galactosidase assay was performed as follows. An OD_{600} reading was taken of each culture, and an appropriate amount (1 mL for exponential phase and mock irradiation cultures, and 50 µL for stationary phase) was aliquoted into 2 mL microcentrifuge tubes. Cells were pelleted via centrifugation at 6900 xg for 3 min, and supernatant was removed. Cells were resuspended in 1 mL Z buffer (0.06 M Na_2HPO_4, 0.04 NaH_2PO4, 0.01 KCl, 0.001M $MgSO_4$, to volume with purified dH_2O), and 1 mL Z buffer was aliquoted in a 2 mL microcentrifuge tube for a blank sample. One-hundred µL chloroform and 50 uL 0.1% SDS were added to each tube. Each sample was then vortexed for 10 s and incubated at 4 °C for at least 10 min. Three samples at a time were removed from 4 °C and placed in a 28 °C water bath for 5 min. Two-hundred uL 4 mg/mL O-

Nitrophenyl β-D-Galactopyranoside (ONPG) (Sigma-Aldrich, St. Louis, MO Cat#: N1127) dissolved in Z buffer was added to each sample. The development of yellow coloration for each sample was timed. Once the sample had become yellow (or after 30 minutes) the reaction was stopped by adding 500 uL 1M Na_2CO_3 and samples were placed on ice. All samples were spun for 10 min at 17000 xg in a microcentrifuge at 4 °C. One mL was removed from each sample, and the OD_{420} and OD_{550} was read. To determine β-galactosidase activity, the following equation was used where *t* is time of the reaction and *v* is the volume of culture used:

$$\boldsymbol{Activity} = \frac{\mathbf{1000} * \left(\left(\boldsymbol{OD_{420}} - (\mathbf{1.75} * \boldsymbol{OD_{550}})\right)\right)}{\boldsymbol{t} * \boldsymbol{v} * \boldsymbol{OD_{600}}}$$

All β-galactosidase assays were performed using biological triplicate. To determine relative β-galactosidase activity compared to the parent strain, the activity of each mutant strain replicate was divided by the average activity of the wild-type triplicate in the given experiment.

6.17 Growth competition assay

This assay was adapted from a previously published protocol [121]. To differentiate strains within the competition, a fitness-neutral deletion of the *araBAD* operon was introduced into one of the two strains. This deletion results in red colonies on tetrazolium arabinose (TA) agar plates [121]. Briefly, overnight culture of each strain to be competed were mixed 1:1, or 1:10 (for growth competitions with CB3000 and GsiB mutations, in favor of the losing competitor), or 1:9 (for growth competitions using Linac isolates or RpoB or ArcB mutations, in favor of the losing competitor) in a 1.5 mL microcentrifuge tube. Samples were mixed by vortexing for 5 s and were serial diluted 1:10 in 900 µl phosphate-buffered saline (PBS) to a final dilution of 1:100,000. One

hundred uL of the final dilution was spread plated onto TA agar plates to assay for CFU. Seventy µl of the remaining cell mixture was used to inoculate 5 mL of fresh LB media for growth overnight. This overnight culture was used to inoculate fresh media the following day, and 100 µl was serially diluted and plated as noted above. This procedure was repeated twice more over a period of two days. The number of white versus red CFU was noted after each day of the competition and the total percentage of the culture for each competitor was determined.

The following equation was used to calculate a relative fitness statistic for each strain, where *w* is the fitness statistic, *A* and *B* are the CFU/mL of the competing strains *A* and *B*, *f* denotes the final CFU/mL, and *i* the initial CFU/mL [158]:

$$\boldsymbol{w = \left(\frac{ln\left(\frac{A_f}{A_i}\right)}{ln\left(\frac{B_f}{B_i}\right)}\right)}$$

Fitness statistics presented were calculated from at least two separate experiments. For strains which had a fitness statistic of 1.2 or greater relative to wild-type, rather than mixing strains 1:1 at the start of the competition, they were mixed 10 (parent strain): 1 (more-fit competitor strain) in order to obtain countable colonies of the wild-type strain after 72 hr of competition.

6.18 Growth curves

Strains were cultured described as in *Growth conditions (6.1)* overnight and to an OD_{600} of 0.2 in LB medium, EZ medium supplemented with 0.2% glucose (Teknova; Hollister, CA**)** and M9 minimal medium [149] supplemented with 0.2% glucose. Cultures were then diluted 1:100 in the appropriate medium in a clear, flat bottom 96-well plate (Fisher product #: 07-200-656**)** and

incubated overnight in a Biotek Synergy 2 plate reader **(**Biotek; Winooski, VT**)** at 37 °C with shaking, with OD_{600} readings taken by the plate reader every 10 minutes.

6.19 SDS-PAGE Protein Gel Electrophoresis

Samples were mixed 1:1 2X Laemilli sample buffer (250 mM Tris-HCl pH 6.8, 10% SDS, 20% glycerol, 10% β-Mercaptoethanol, sufficient Bromophenol Blue for dark blue coloration), and boiled for 5 min. Fifteen µL of each sample was loaded onto a Bio-Rad pre-cast 12% SDS-PAGE protein gel (Hercules, CA; 10 well Product #: 4561043; 15 well Product #: 4561046). Gels were run at 100 V until dye from Laemilli sample buffer had run off the gel.

6.20 Western blots for ClpX detection

Cells from a single colony of each strain were cultured overnight and then grown to an OD_{600} of ~ 0.2 as in *Growth conditions (6.1).* These cultures were placed on ice for at least 5 minutes to stop growth before use. One mL of each culture was pelleted via centrifugation at 17000 xg for 1 min, supernatant was removed, and cells were resuspended in 20 µl dH_2O and mixed with Laemilli sample buffer and run on a 12% SDS-PAGE gel as described. The gel was then placed in Western Transfer Buffer (per 1 L: 14 g glycine, 3 g Tris base, 200 mL methanol, 800 mL dH_20) for at least 15 min. A transfer sandwich was then prepared using Amersham Protean 0.2 µm Nitrocellulose Blotting Membrane (GE Healthcare Life Science, Cat# 10600001). The proteins were transferred from the SDS-PAGE gel to the transfer membrane as previously described [159] at 300 mA at 4 °C for 30 min. Blocked the transfer membrane in 15 mL 5% milk blocking solution (2.5 g powdered milk in 50 mL PBS-T (for 1 L, 50 mL 20X Phosphate buffered saline, 5 mL Tween 20,

and 945 mL purified dH_2O)) with rocking for 1 hr at room temperature (~24 °C). Removed blocking solution, then added 10 mL of 1:10000 Rabbit α-ClpX (Tania Baker lab, Massachusetts Institute of Technology) Rabbit α-SSB (James Keck lab;University of Wisconsin, Madison) in blocking solution. Samples were rocked for 1 hr at room temperature (~24 °C). Primary antibody was poured off, the transfer membrane washed by rinsing 3X in 25 mL 1X PBS-T. Washed transfer membranes were placed in 25 mL 1X PBS-T and rocked for 15 min. The membrane was removed, and two additional rinses performed. The PBS-T was replaced, but the membranes were only rocked for 5 min. After washing, 10 mL of 1:10000 Chicken α-Rabbit HRP-conjugated secondary antibody (Sigma-Aldrich, St. Louis, MO Cat#: A9046) in blocking solution was added. Samples were rocked for 1 hr at room temperature (~24 °C). Secondary antibody was poured off and washed as described previously. Imaged antibody-bound protein was detected using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo-Scientific; Rockford, IL; Prod#: 34577).

6.21 Silverstain of irradiated *E. coli* samples

Grew quintuplicate overnight cultures of *E. coli* MG1655 as described in *Growth conditions (6.1)*. Twenty-five mL LB medium was inoculated with 175 uL of overnight culture and grown to an OD_{600} of 0.2 as described. Two 1 mL aliquots were washed using PBS, and one sample for each replicate culture was irradiated with 1000 Gy of high energy electron beam IR as described in *6.9 Ionizing radiation resistance assay using Linac*. Post irradiation, pelleted samples via centrifugation at 13 xg for 1 min. Supernatant was removed, and samples were resuspended in via centrifugation at 17000 xg for 1 min, supernatant was removed, and cells were resuspended in 12.5 µl dH_2O and mixed with 12.5 µl Laemilli sample buffer. Fourteen µl

of each sample was loaded and run on a 12% SDS-PAGE gel as described. Gel was then stained per manufacturer's protocol using the Silver Stain Plus kit (Bio-Rad; Hercules, CA; Cat #: 1610449).

6.22 Western blots for protein carbonylation detection

Used sample prepped in *Silverstain of irradiated* E. coli *samples (6.21).* On the day of treatment (sample quality rapidly decreases), samples were then treated with DNPH (or control solution) as described by manufacturer's protocol using Oxyblot Protein Oxidation Detection kit (EMD Milipore; Burlington, MA; Cat #: S7150). Fourteen µl of each sample was loaded and run on a 12% SDS-PAGE gel as described in *6.19 SDS-PAGE Protein Gel Electrophoresis*. The gel was then placed in Western Transfer Buffer (per 1 L: 7.2 g glycine, 1.45 g Tris base, 200 mL methanol, 800 mL dH_20) for at least 15 min. Western Blot was then performed) as described by manufacturer's protocol using Oxyblot Protein Oxidation Detection kit (EMD Milipore; Burlington, MA; Cat #: S7150) using 1% BSA blocking solution and Amersham Protean 0.2 µm Nitrocellulose Blotting Membrane (GE Healthcare Life Science, Cat# 10600001). Imaged antibody-bound protein was detected using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo-Scientific, Rockford, IL, Prod#: 34577).

6.23 Protein Purification

Purified proteins were generously provided by Tricia Windgassen from the lab of Professor James Keck (University of Wisconsin - Madison). PriA was purified as described previously (Bhattacharyya 2014 PNAS). V554I was incorporated into the PriA overexpression plasmid,

pET15-*E. coli* PriA, using site-directed mutagenesis. PriA was over expressed using 1 mM IPTG for 3 hours in Rosetta 2 *E. coli* (Millipore Sigma, Burlington, MA, Prod#: 71405-3) containing pET15-*E. coli* PriA after reaching OD_{600} 0.5-0.6 (in LB supplemented with 100 ng/µl ampicillin and 20 ng/µl chloramphenicol). Cells were lysed using sonication and N-terminal 6xHis tagged PriA was purified using nickel affinity, ion exchange (SPFF), and size exclusion (S300) chromatography. Single-stranded DNA binding protein (SSB) was purified as described previously (George et al 2012). PriB was purified as described previously (Lopper et al, 2004).

6.24 Fluorescence anisotropy assay

Fluorescence anisotropy was performed by Tricia Windgassen from the lab of James Keck (University of Wisconsin – Madison). PriA DNA fork-binding fluorescence anisotropy assay was performed, analyzed, and corrected for intensity change as described previously [160]. Briefly, PriA (0.1 to 5000 nM) was incubated with 1 nM fluorescein-labeled DNA fork (two annealed oligonucleotides from [124, 125]) containing 60 bp complementary and 38 nt non-complementary sequence, with the template lagging strand 5’ fluorescein labeled) in 20 mM HEPES-HCl pH 7.0, 5 % glycerol, 75 mM potassium glutamate, and 10 mM dithiothreitol for 30 minutes at room temperature before fluorescence measurements on a Beacon 2000 fluorescence polarization instrument. Data from at least three replicates were fit to a single site-specific model using GraphPad Prism.

6.25 ATPase assay

ATPase assays were performed by Tricia Windgassen from the lab of James Keck (University of Wisconsin – Madison) PriA DNA-dependent ATPase was measured as previously described [160]. Briefly, 50 nM PriA was incubated with increasing concentrations of ssDNA (dT^{28} 0.1-5000 nM) in 20 mM HEPES-HCl, pH 8.0, 50 mM sodium chloride, 1 mM β-mercaptoethanol, 5 mM magnesium chloride, 0.10 mg/mL BSA, 2 mM phosphoenolpyruvate, 0.2 mM nicotinamide adenine dinucleotide, 3 U/ml pyruvate kinase, and 4.5 U/ml lactate dehydrogenase for 10 minutes at room temperature. Absorbance at 340 nm was monitored for 1 hour at room temperature upon addition of 1 mM ATP. Data from three replicates were analyzed as described previously [161].

6.26 Helicase assay

Helicase assays were performed by Tricia Windgassen from the lab of James Keck (University of Wisconsin – Madison). PriA helicase activity was analyzed on two synthetic DNA fork types: one composed of two oligonucleotides (as used in the fluorescence anisotropy assay) and the other composed of four oligonucleotides (adding "nascent" leading and lagging strands to the first fork type, yielding fully double-stranded DNA except for a 5 nt gap on the lagging strand before the fork junction). The synthetic template lagging strand was 5'-^{32}P labeled, annealed into either fork type, purified, and tested in helicase assays as described previously [160]. Briefly, PriA (0.1, 1, or 10 nM) was incubated with 1 nM radiolabeled fork in 50 mM HEPES-HCl, pH 8.0, 40 mg/L BSA, 2 mM dithiothreitol, 2 mM ATP, and 4 mM magnesium acetate for 30 minutes at 37 °C. For samples analyzed for SSB and PriB helicase stimulation, 250 nM SSB or 10 nM PriB were

preincubated with DNA, before addition of PriA (1 nM). Samples were stopped by incubation with 20 mM EDTA, 0.5 % SDS, 0.2 mg/mL proteinase K, and 2.5 ng/μL unlabeled "trap" oligonucleotide for 30 minutes at 37 °C. DNA-loading buffer was added, and samples were resolved through 10 % polyacrylamide gel electrophoresis in 1X TBE. Gels were fixed, dried, exposed to PhosphoImager screen (GE Healthcare; Chicago, IL), and imaged using a Typhoon FLA 9000.

6.27 Preparation and irradiation of whole cell, cell lysate, and pure protein samples

Grew quintuplicate overnight cultures of *E. coli* MG1655 as described in *Growth conditions (6.1).* Fifty mL LB medium was inoculated with 500 uL of overnight culture and grown to an OD_{600} of 0.2 as described. Forty mL of cultures were aliquoted into separate 50 mL conical tubes and were pelleted by centrifugation at 5000 rpm, 4 °C, for 10 min. Supernatant was removed and cells were resuspended in 10 mL ice cold 1X PBS. Pelleting and resuspension was repeated three more times, however final resuspension was done in 800 μL PBS. Made six 100 μL aliquots of each culture into separate 1.5 mL tubes. Placed each on ice.

For lysate and pure protein prep at 0 and 1000 Gy, froze four 100 μL aliquots for each sample at -80 °C. This is the first of two freeze/thaw cycles to lyse cells.

Kept two of the whole cell samples on ice (two for each replicate cultures) and took to the Linac irradiator in the Wisconsin Institute for Medical Research (WIMR) building. One aliquot for each replicate was irradiated as described with a dose of 1000 Gy; the other replicate received no dose. Samples that received no dose and 1000 Gy were returned to lab and frozen at -80 °C.

After 2 hr, removed all samples from -80 °C and thawed at room temperature for 10 min. Froze all samples overnight at -80 °C.

The following day, removed and thawed the lysate samples at room temperature for 10 min. These samples were irradiated as described for the whole cell samples. Upon returning to lab, removed and thawed the whole cell samples at room temperature for 10 min. Pelleted all whole cell and cell lysate samples vis microcentrifugation at 13 xg for 2 min. One-hundred µL aliquots of supernatant for each sample and added to 3500 MWCO Slide-a-lyzer Mini-Dialysis Units (ThermoFisher; Madison, WI; Cat #69550). Mini-Dialysis Units were floated on 500 mL 1X PBS with very slow stirring with stir bar. Let dialyze for 2 hr at room temperature. PBS was then replaced with fresh 500 mL 1X PBS, and dialysis was repeated (2 hr at room temperature). PBS was replaced a final time, and dialysis was performed overnight (~ 16 hr) at room temperature.

The next day, to set-aside an aliquot of purified protein for each sample to run on an SDS-PAGE gel, 25 µL of each sample was removed and added to 25 uL of 2X Laemilli sample buffer, boiled 5 min, and placed in 4 °C.

Placed remaining whole cell and cell lysate samples from dialysis into fresh 1.5 mL tubes, flash-froze with liquid N_2, and stored at -80 °C.

Removed and thawed the pure protein samples at room temperature for 10 min. Repeated steps for pelleting and dialyzing samples as described above for the whole cell and lysate samples. Since dialyzed protein ends up with greater than 100 µL, split volumes into two tubes as close to 100 uL as possible. Samples were irradiated as described above.

After irradiation, the appropriate samples were re-combined in a single 1.5 mL tube. As described above, 25 µL aliquots were removed and stored at 4 °C to run on an SDS-PAGE gel.

Placed remaining pure protein samples from dialysis into fresh 1.5 mL tubes and flash-froze with liquid N_2, and stored at -80 °C.

Frozen samples were delivered to Sussman lab on dry ice for further processing and mass spectrometry analysis. A final volume of 2% SDS was added to each sample (5 0Gy and 5 1000Gy for whole cell, cell lysate, and pure protein samples) and then immediately subjected to protein extraction and concentration using a standard methanol:chloroform protocol. TMT mass spectrometry was performed as described below, without pH pre-fractionation.

6.28 Preparation of wild-type *E. coli* samples for 'deep' mass spectrometry

Samples were grown in biological quintuplicate overnight and to an OD600 of 0.2 in EZ + 0.2% glycerol as routinely performed; however, cultures were grown to an OD600 of 0.2 in a total volume of 50 mL for each replicate. Once early exponential phase cultures were obtained, 40 mL of each culture was pelleted by centrifugation at 3500 rpm for 10 min at 4 °C. Supernatants were poured off, and then samples were resuspended in 40 mL of ice-cold 1X PBS. Cells were pelleted and supernatant was poured off again as described, but cells were resuspended in 20 mL ice-cold PBS. This process was repeated three more times, but the final resuspension was done with 500 µL ice-cold 1X PBS. Four 100 µL aliquots were made for each culture into 1.5 mL microfuge tubes. Two for 0 and 1000 Gy for mass spectrometry analysis, and two for plating to determine lethality as described in *Serial dilutions (6.2).*

6.29 TMT-labelled mass spectrometry

Samples were lysed by addition SDS to a final volume of 2.0%, then immediately subjected to protein extraction and concentration using a standard methanol:chloroform protocol. Purified protein pellets were solubilized into 8M Urea in 50mM triethylammonium bicarbonate (TEAB) and subjected to a standard BCA assay to determine protein concentration.

For each of the 10 samples, 10ug (varying volumes) of each was diluted to 4Murea with 50mM TEAB, and treated with 2mM dithiothreitol for 30 min at 50 °C, 5mM iodoacetamide for 30 min at room temperature in darkness, and then 2mM dithiothreitol for 5 min at room temperature. Samples were diluted further to 1Murea with TEAB, and 0.5ug of Trypsin and Lys-C proteases were each added. Samples were incubated overnight at 37 °C, for 15hr total.

Digestions were stopped with addition of 1.0% formic acid, subjected to sold phase cleanup using Agilent C18 OMIX tips, according to manufacturer's protocol, and dried down to completion using a vacuum centrifuge. Dried samples were resolubilized into 50mM TEAB. To each tube, 40ug of each of the 10 tandem mass tag chemical adducts were added as follows (4:1 tag:protein): 126 - 0gy1; 127N - 0gy2; 127C - 0gy3; 128N - 0gy4; 128C - 0gy5; 129N - 1Kgy1; 129C - 1Kgy2; 130N - 1Kgy3; 130C - 1Kgy4; 131 - 1Kgy5.

Samples were incubated for 1 hour at room temperature, after which 5% hydroxylamine was added to quench reaction. Samples were immediately combined and purified/concentrated using a Waters 1cc C18 Sep-pak solid phase chromatographic column, according to manufacturer's protocol, then dried to completion with a vacuum centrifuge.

Samples were resolubilized into 300uL 0.1% formic acid for high-pH tip-based fractionation. Fractionation was performed using a Pierce high pH reversed phase peptide

fractionation spin column kit, according to manufacturer's protocol. This resulted in nine fractions, which were dried down to completion using a vacuum centrifuge.

Samples were resolubilized into 10uL 0.1%formic acid for injection onto a ThermoFisher Orbitrap Lumos mass spectrometer instrument. For liquid chromatographic conditions, stationary phase was C18 and flow rate was 275nL/min. Three µL of each fraction was injected for analysis. Mobile phases A and B were 0.2% formic acid in water and 0.2% formic acid in 70% acetonitrile, respectively. For separation and elution, a 60 min gradient to 55% buffer B was used. HCD fragmentation was used to obtain high resolution MS2 spectra for TMT mass tag quantification.

Data was analyzed using the Sequest algorithm within Proteome Discoverer (ThermoFisher Scientific; Madison, WI). The Uniprot *E. coli* proteome (PID: UP000000625) was searched with the specified parameters: trypsin with 2 possible missed cleavages, precursor and fragment mass tolerance 10 ppm and 0.6 Da, respectively, and a max amount of 4 dynamic modifications per peptide. Dynamic modifications were specified as carbamidomethyl (on C), deamidation (on NQ), oxidation (on ACDEFHILMNPQRSTVWY), carbonylation (on EILPQRV), and peroxidation (on ACDEFGHILMNPQRSTVWY). Fixed modifications of tandem mass tag were specified on N-termini and Lys residues. A false discovery rate (FDR) for peptide spectral matches (PSMs), peptides, and proteins of 1.0 % was used. For TMT quantification, tag abundances were normalized to the total tag amount per channel.

Post searching, peptide identifications with an XCorr > 2.0 were considered (high confidence sequence match), and fold change values were obtained by averaging normalized reporter ion abundances pep identified peptide in the 5 treated channels vs. the 5 control channels. P-values were calculated using 2-tailed, unpaired Student's t-testing, assuming equal variance. Q values were calculated using Benjamini-Hochberg FDR testing.

6.30 Mass spectrometry of irradiated L-amino acids

LC-HR-ESI-TOF analysis of IR-dosed amino acids in water was done on Agilent LC/MSD TOF system (model 6210) equipped with 1200 series HPLC liquid handling system (Agilent Technologies; Santa Clara, CA). Measurements of starting compounds and their sister products generated from abiotic exposure was conducted under positive or negative ion polarity, where 2µl of control and IR-dosed sample was injected from an autosampler vial sitting at 6 °C into an isocratic 0.05 mL/min flow of 50:50 [Water / 0.1% Formic acid] : [Acetonitrile / 0.1% Formic acid]. The following instrumental parameters were used to generate the most optimum protonated [M+H+] or deprotonated [M-H-] ions under their respective acquisition polarity: Capillary voltage [3000V]; Drying gas [7 l/min]; Gas temperature [300 °C]; Nebulizer [15psig]; Oct DC1 [35V] for positive and [-34V] for negative ionization; Fragmentor [140V]; Oct RF [200V]; Skimmer [60V]. Internal calibration was achieved with assisted spray of two reference masses, 112.9856 m/z and 1033.9881 m/z. Data was acquired in profile mode scanning from 50-1,600 AMU at 0.89 cycles per second and 10,000 transients per scan.

Data processing was executed using Analysis QS 1.1 (build:9865) software (Agilent Technologies; Santa Clara, CA) to extract unique masses observed in the spectrum. This semi-manual analysis strategy was further complemented and verified with fully-automatic processing using Agilent's MassHunter Workstation Qualitative Analysis software (version B.01.03, build: 1.3.157.0), where centroided data (filtered based on absolute peak height of 200 counts per second, minimum to differentiate peaks from the baseline on our instrument) was used to assign compounds from their molecular features and ultimately generate extracted ion abundance chromatograms (using original profile data) to aid in quantification and distribution over the whole IR dosage spectrum.

Chapter VII: References

1. Cox, M.M. and J.R. Battista, *Deinococcus radiodurans - The consummate survivor*. Nature Reviews Microbiology, 2005. **3**(11): p. 882-892.
2. Reisz, J.A., et al., *Effects of ionizing radiation on biological molecules--mechanisms of damage and emerging methods of detection.* Antioxid Redox Signal, 2014. **21**(2): p. 260-92.
3. Kempner, E.S., *Effects of high-energy electrons and gamma rays directly on protein molecules.* J Pharm Sci, 2001. **90**(10): p. 1637-46.
4. Desouky, O., N. Ding, and G. Zhou, *Targeted and non-targeted effects of ionizing radiation.* Journal of Radiation Research and Applied Sciences, 2015. **8**(2): p. 247-254.
5. Cadet, J. and J.R. Wagner, *DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation.* Cold Spring Harbor perspectives in biology, 2013. **5**(2): p. a012559.
6. Daly, M.J., *A new perspective on radiation resistance based on Deinococcus radiodurans.* Nat Rev Microbiol, 2009. **7**(3): p. 237-45.
7. Aruoma, O.I., B. Halliwell, and M. Dizdaroglu, *Iron ion-dependent modification of bases in DNA by the superoxide radical-generating system hypoxanthine/xanthine oxidase.* J Biol Chem, 1989. **264**(22): p. 13024-8.
8. Wang, D., D.A. Kreutzer, and J.M. Essigmann, *Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions.* Mutat Res, 1998. **400**(1-2): p. 99-115.
9. Imlay, J.A., *The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium.* Nature Reviews Microbiology, 2013. **11**(7): p. 443.

10. David, S.S., V.L. O'Shea, and S. Kundu, *Base-excision repair of oxidative DNA damage.* Nature, 2007. **447**(7147): p. 941-50.

11. Tremblay, S. and J.R. Wagner, *Dehydration, deamination and enzymatic repair of cytosine glycols from oxidized poly(dG-dC) and poly(dI-dC).* Nucleic Acids Res, 2008. **36**(1): p. 284-93.

12. Xie, C.-X., et al., *Comparison of base substitutions in response to nitrogen ion implantation and 60Co-gamma ray irradiation in Escherichia coli.* Genetics and Molecular Biology, 2004. **27**(2): p. 284-290.

13. Harris, D.R., et al., *Directed evolution of radiation resistance in Escherichia coli.* Journal of Bacteriology, 2009. **191**: p. 5240-5252.

14. Talhaoui, I., et al., *7,8-Dihydro-8-oxoadenine, a highly mutagenic adduct, is repaired by Escherichia coli and human mismatch-specific uracil/thymine-DNA glycosylases.* Nucleic Acids Res, 2013. **41**(2): p. 912-23.

15. Krwawicz, J., et al., *Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease.* Acta Biochim Pol, 2007. **54**(3): p. 413-34.

16. Daly, M.J., *Death by protein damage in irradiated cells.* DNA Repair, 2012. **11**(1): p. 12-21.

17. Cunniffe, S., et al., *Reduced repair capacity of a DNA clustered damage site comprised of 8-oxo-7,8-dihydro-2'-deoxyguanosine and 2-deoxyribonolactone results in an increased mutagenic potential of these lesions.* Mutat Res, 2014. **762**: p. 32-9.

18. Eccles, L.J., P. O'Neill, and M.E. Lomax, *Delayed repair of radiation induced clustered DNA damage: friend or foe?* Mutat Res, 2011. **711**(1-2): p. 134-41.

19. Daly, M.J., et al., *Accumulation of Mn(II) in, Deinococcus radiodurans facilitates gamma-radiation resistance.* Science, 2004. **306**(5698): p. 1025-1028.

20. Daly, M.J., et al., *Protein Oxidation Implicated as the Primary Determinant of Bacterial Radioresistance.* PLoS Biol, 2007. **5**(4): p. e92.

21. Fredrickson, J.K., et al., *Protein oxidation: key to bacterial desiccation resistance?* Isme Journal, 2008. **2**: p. 393-403.

22. Daly, M.J., et al., *Small-molecule antioxidant proteome-shields in Deinococcus radiodurans.* PLoS One, 2010. **5**(9): p. e12570.

23. Du, J. and J.M. Gebicki, *Proteins are major initial cell targets of hydroxyl free radicals.* Int J Biochem Cell Biol, 2004. **36**(11): p. 2334-43.

24. Xu, G. and M.R. Chance, *Hydroxyl radical-mediated modification of proteins as probes for structural proteomics.* Chem Rev, 2007. **107**(8): p. 3514-43.

25. Girod, M., et al., *Structural basis of protein oxidation resistance: a lysozyme study.* PLoS One, 2014. **9**(7): p. e101642.

26. Ezraty, B., et al., *Oxidative stress, protein damage and repair in bacteria.* Nat Rev Microbiol, 2017. **15**(7): p. 385-396.

27. Fahey, R.C., et al., *Occurrence of glutathione in bacteria.* J Bacteriol, 1978. **133**(3): p. 1126-9.

28. Maisonneuve, E., et al., *Rules governing selective protein carbonylation.* PLoS One, 2009. **4**(10): p. e7269.

29. Cabiscol Català, E., J. Tamarit Sumalla, and J. Ros Salvador, *Oxidative stress in bacteria and protein damage by reactive oxygen species.* International Microbiology, 2000, vol. 3, núm. 1, p. 3-8, 2000.

30. Imlay, J.A., *Diagnosing oxidative stress in bacteria: not as easy as you might think.* Curr Opin Microbiol, 2015. **24**: p. 124-31.

31. Joshi, S.G., et al., *Nonthermal dielectric-barrier discharge plasma-induced inactivation involves oxidative DNA damage and membrane lipid peroxidation in Escherichia coli.* Antimicrob Agents Chemother, 2011. **55**(3): p. 1053-62.

32. Truglio, J.J., et al., *Prokaryotic nucleotide excision repair: the UvrABC system.* Chem Rev, 2006. **106**(2): p. 233-52.

33. Kow, Y.W., S.S. Wallace, and B. Van Houten, *UvrABC nuclease complex repairs thymine glycol, an oxidative DNA base damage.* Mutat Res, 1990. **235**(2): p. 147-56.

34. Anderson, D.G. and S.C. Kowalczykowski, *The recombination hot spot chi is a regulatory element that switches the polarity of DNA degradation by the RecBCD enzyme.* Genes Dev, 1997. **11**(5): p. 571-81.

35. Dixon, D.A. and S.C. Kowalczykowski, *The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the E. coli RecBCD enzyme.* Cell, 1993. **73**(1): p. 87-96.

36. Bianco, P.R. and S.C. Kowalczykowski, *The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single strand of DNA containing the sequence 5'-GCTGGTGG-3'.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6706-11.

37. Wigley, D.B., *Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB.* Nat Rev Microbiol, 2013. **11**(1): p. 9-13.

38. Churchill, J.J., D.G. Anderson, and S.C. Kowalczykowski, *The RecBC enzyme loads RecA protein onto ssDNA asymmetrically and independently of chi, resulting in constitutive recombination activation.* Genes Dev, 1999. **13**(7): p. 901-11.

39. Bell, J.C., et al., *Direct imaging of RecA nucleation and growth on single molecules of SSB-coated ssDNA.* Nature, 2012. **491**(7423): p. 274-8.

40. Simmons, L.A., et al., *The SOS Regulatory Network.* EcoSal Plus, 2008. **2008**.

41. Radman, M., *SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis.* Basic Life Sci, 1975. **5A**: p. 355-67.

42. Cox, M.M., *Historical overview: searching for replication help in all of the rec places.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8173-80.

43. Kuzminov, A., *Recombinational repair of DNA damage in Escherichia coli and bacteriophage lambda.* Microbiol Mol Biol Rev, 1999. **63**(4): p. 751-813, table of contents.

44. Kowalczykowski, S.C., et al., *Biochemistry of homologous recombination in Escherichia coli.* Microbiol Rev, 1994. **58**(3): p. 401-65.

45. Imlay, J.A., *Cellular defenses against superoxide and hydrogen peroxide.* Annu Rev Biochem, 2008. **77**: p. 755-76.

46. Rahman, M.A., et al., *Cloning, sequencing, and expression of the Escherichia coli peptide methionine sulfoxide reductase gene.* J Biol Chem, 1992. **267**(22): p. 15549-51.

47. Brot, N., et al., *Enzymatic reduction of protein-bound methionine sulfoxide.* Proc Natl Acad Sci U S A, 1981. **78**(4): p. 2155-8.

48. Lin, Z., et al., *Free methionine-(R)-sulfoxide reductase from Escherichia coli reveals a new GAF domain function.* Proc Natl Acad Sci U S A, 2007. **104**(23): p. 9597-602.

49. Ezraty, B., et al., *Methionine sulfoxide reduction and assimilation in Escherichia coli: new role for the biotin sulfoxide reductase BisC.* J Bacteriol, 2005. **187**(1): p. 231-7.

50. Grimaud, R., et al., *Repair of oxidized proteins. Identification of a new methionine sulfoxide reductase.* J Biol Chem, 2001. **276**(52): p. 48915-20.

51. Gennaris, A., et al., *Repairing oxidized proteins in the bacterial envelope using respiratory chain electrons.* Nature, 2015. **528**(7582): p. 409-412.

52. Depuydt, M., et al., *A periplasmic reducing system protects single cysteine residues from oxidation.* Science, 2009. **326**(5956): p. 1109-11.

53. Denoncin, K., et al., *A new role for Escherichia coli DsbC protein in protection against oxidative stress.* J Biol Chem, 2014. **289**(18): p. 12356-64.

54. Outten, F.W., O. Djaman, and G. Storz, *A suf operon requirement for Fe-S cluster assembly during iron starvation in Escherichia coli.* Mol Microbiol, 2004. **52**(3): p. 861-72.

55. Imlay, J.A., *Iron-sulphur clusters and the problem with oxygen.* Mol Microbiol, 2006. **59**(4): p. 1073-82.

56. Tanaka, M., et al., *Analysis of Deinococcus radiodurans's transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance.* Genetics, 2004. **168**(1): p. 21-33.

57. Slade, D. and M. Radman, *Oxidative stress resistance in Deinococcus radiodurans.* Microbiol Mol Biol Rev, 2011. **75**(1): p. 133-91.

58. Harris, D.R., et al., *Preserving Genome Integrity: The DdrA Protein of Deinococcus radiodurans R1.* PLoS Biology, 2004. **2**(10): p. e304.

59. Norais, C.A., et al., *DdrB protein, an alternative Deinococcus radiodurans SSB induced by ionizing radiation.* J Biol Chem, 2009. **284**(32): p. 21402-11.

60. Zahradka, K., et al., *Reassembly of shattered chromosomes in Deinococcus radiodurans.* Nature, 2006. **443**(7111): p. 569-73.

61. Ithurbide, S., et al., *Single Strand Annealing Plays a Major Role in RecA-Independent Recombination between Repeated Sequences in the Radioresistant Deinococcus radiodurans Bacterium.* PLoS Genet, 2015. **11**(10): p. e1005636.

62. Selvam, K., et al., *DdrA, DdrD, and PprA: components of UV and mitomycin C resistance in Deinococcus radiodurans R1*. PLoS One, 2013. **8**(7): p. e69007.

63. Earl, A.M., et al., *The IrrE protein of Deinococcus radiodurans R1 is a novel regulator of recA expression.* J Bacteriol, 2002. **184**(22): p. 6216-24.

64. Narumi, I., et al., *PprA: a novel protein from Deinococcus radiodurans that stimulates DNA ligation.* Mol Microbiol, 2004. **54**(1): p. 278-85.

65. Mattimore, V. and J.R. Battista, *Radioresistance of Deinococcus radiodurans: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation.* Journal of Bacteriology, 1996. **178**: p. 633-637.

66. Musilova, M., et al., *Isolation of Radiation-Resistant Bacteria from Mars Analog Antarctic Dry Valleys by Preselection, and the Correlation between Radiation and Desiccation Resistance.* Astrobiology, 2015. **15**(12): p. 1076-90.

67. Rainey, F.A., et al., *Extensive diversity of ionizing-radiation-resistant bacteria recovered from Sonoran Desert soil and description of nine new species of the genus Deinococcus obtained from a single soil sample.* Appl Environ Microbiol, 2005. **71**(9): p. 5225-35.

68. Elena, S.F. and R.E. Lenski, *Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation.* Nat Rev Genet, 2003. **4**(6): p. 457-69.

69. Barrick, J.E. and R.E. Lenski, *Genome dynamics during experimental evolution.* Nat Rev Genet, 2013. **14**(12): p. 827-39.

70. Lee, H., et al., *Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing.* Proc Natl Acad Sci U S A, 2012. **109**(41): p. E2774-83.

71. Lind, P.A. and D.I. Andersson, *Whole-genome mutational biases in bacteria.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17878-83.

72. Sung, W., et al., *Drift-barrier hypothesis and mutation-rate evolution.* Proc Natl Acad Sci U S A, 2012. **109**(45): p. 18488-92.

73. Foster, P.L., et al., *Determinants of spontaneous mutation in the bacterium Escherichia coli as revealed by whole-genome sequencing.* Proc Natl Acad Sci U S A, 2015. **112**(44): p. E5990-9.

74. Sung, W., et al., *Asymmetric Context-Dependent Mutation Patterns Revealed through Mutation-Accumulation Experiments.* Mol Biol Evol, 2015. **32**(7): p. 1672-83.

75. Foster, P.L., et al., *On the mutational topology of the bacterial genome.* G3 (Bethesda), 2013. **3**(3): p. 399-407.

76. Kram, K.E., et al., *Adaptation of Escherichia coli to Long-Term Serial Passage in Complex Medium: Evidence of Parallel Evolution.* mSystems, 2017. **2**(2).

77. Rozen, D.E. and R.E. Lenski, *Long-Term Experimental Evolution in Escherichia coli. VIII. Dynamics of a Balanced Polymorphism.* Am Nat, 2000. **155**(1): p. 24-35.

78. Finkel, S.E., *Long-term survival during stationary phase: evolution and the GASP phenotype.* Nat Rev Microbiol, 2006. **4**(2): p. 113-20.

79. Chou, H.H., et al., *Diminishing returns epistasis among beneficial mutations decelerates adaptation.* Science, 2011. **332**(6034): p. 1190-2.

80. Fogle, C.A., J.L. Nagle, and M.M. Desai, *Clonal interference, multiple mutations and adaptation in large asexual populations.* Genetics, 2008. **180**(4): p. 2163-73.

81. Lenski, R.E., *Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations.* ISME J, 2017. **11**(10): p. 2181-2194.

82. Good, B.H., et al., *The dynamics of molecular evolution over 60,000 generations.* Nature, 2017. **551**(7678): p. 45-50.

83. Grosskopf, T., et al., *Metabolic modelling in a dynamic evolutionary framework predicts adaptive diversification of bacteria in a long-term evolution experiment.* BMC Evol Biol, 2016. **16**(1): p. 163.

84. Blount, Z.D., C.Z. Borland, and R.E. Lenski, *Historical contingency and the evolution of a key innovation in an experimental population of Escherichia coli.* Proc Natl Acad Sci U S A, 2008. **105**(23): p. 7899-906.

85. Blount, Z.D., et al., *Genomic analysis of a key innovation in an experimental Escherichia coli population.* Nature, 2012. **489**(7417): p. 513-8.

86. Witkin, E.M., *Inherited Differences in Sensitivity to Radiation in Escherichia coli.* Proc Natl Acad Sci U S A, 1946. **32**(3): p. 59-68.

87. Parisi, A. and A. Antoine, *Increased radiation resistance of vegetative Bacillus pumilus.* Applied microbiology, 1974. **28**(1): p. 41-46.

88. Davies, R. and A.J. Sinskey, *Radiation-resistant mutants of Salmonella typhimurium LT2: development and characterization.* Journal of bacteriology, 1973. **113**(1): p. 133-144.

89. Erdman, I., F. Thatcher, and K. MacQueen, *STUDIES ON THE IRRADIATION OF MICROORGANISMS IN RELATION TO FOOD PRESERVATION: II. IRRADIATION RESISTANT MUTANTS.* Canadian journal of microbiology, 1961. **7**(2): p. 207-215.

90. Byrne, R.T., et al., *Evolution of extreme resistance to ionizing radiation via genetic adaptation of DNA Repair* eLife, 2014. **3**: p. e01322.

91. Piechura, J.R., et al., *Biochemical characterization of RecA variants that contribute to extreme resistance to ionizing radiation.* DNA Repair (Amst), 2015. **26**: p. 30-43.

92. Masson, J.Y. and S.C. West, *The Rad51 and Dmc1 recombinases: a non-identical twin relationship.* Trends Biochem Sci, 2001. **26**(2): p. 131-6.

93. Myers, K.S., et al., *Genome-scale analysis of escherichia coli FNR reveals complex features of transcription factor binding.* PLoS Genet, 2013. **9**(6): p. e1003565.

94. Kiley, P.J. and H. Beinert, *Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster.* FEMS microbiology reviews, 1998. **22**(5): p. 341-352.

95. LINN, S. and J.A. IMLAY, *Toxicity, mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide.* J Cell Sci, 1987. **1987**(Supplement 6): p. 289-301.

96. Neher, S.B., et al., *Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon.* Molecular cell, 2006. **22**(2): p. 193-204.

97. Suzuki, H., et al., *The yliA,-B,-C, and-D genes of Escherichia coli K-12 encode a novel glutathione importer with an ATP-binding cassette.* Journal of bacteriology, 2005. **187**(17): p. 5861-5867.

98. Koo, M.S., et al., *A reducing system of the superoxide sensor SoxR in Escherichia coli.* The EMBO journal, 2003. **22**(11): p. 2614-2622.

99. Martinez, J., S. Steenbergen, and E. Vimr, *Derived structure of the putative sialic acid transporter from Escherichia coli predicts a novel sugar permease domain.* Journal of bacteriology, 1995. **177**(20): p. 6005-6010.

100. Rodríguez-Aparicio, L.B., A. Reglero, and J.M. Luengo, *Uptake of N-acetylneuraminic acid by Escherichia coli K-235. Biochemical characterization of the transport system.* Biochemical Journal, 1987. **246**(2): p. 287.

101. Appukuttan, D., et al., *Engineering synthetic multistress tolerance in Escherichia coli by using a deinococcal response regulator, DR1558.* Applied and environmental microbiology, 2016. **82**(4): p. 1154-1166.

102. Lambden, P.R. and J.R. Guest, *Mutants of Escherichia coli K12 unable to use fumarate as an anaerobic electron acceptor.* J Gen Microbiol, 1976. **97**(2): p. 145-60.

103. Salmon, K., et al., *Global gene expression profiling in Escherichia coli K12. The effects of oxygen availability and FNR.* J Biol Chem, 2003. **278**(32): p. 29837-55.

104. Williams, S.M., et al., *Transcription activation at class I FNR-dependent promoters: identification of the activating surface of FNR and the corresponding contact site in the C-terminal domain of the RNA polymerase alpha subunit.* Nucleic Acids Res, 1997. **25**(20): p. 4028-34.

105. Weber, K.D., O.D. Vincent, and P.J. Kiley, *Additional determinants within Escherichia coli FNR activating region 1 and RNA polymerase alpha subunit required for transcription activation.* J Bacteriol, 2005. **187**(5): p. 1724-31.

106. Kiley, P.J. and W.S. Reznikoff, *Fnr mutants that activate gene expression in the presence of oxygen.* J Bacteriol, 1991. **173**(1): p. 16-22.

107. Khoroshilova, N., et al., *Iron-sulfur cluster disassembly in the FNR protein of Escherichia coli by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity.* Proc Natl Acad Sci U S A, 1997. **94**(12): p. 6087-92.

108. Sutton, V.R., et al., *Kinetic analysis of the oxidative conversion of the [4Fe-4S]2+ cluster of FNR to a [2Fe-2S]2+ Cluster.* J Bacteriol, 2004. **186**(23): p. 8018-25.

109. Yoo, S.J., et al., *clpX encoding an alternative ATP-binding subunit of protease Ti (Clp) can be expressed independently from clpP in Escherichia coli.* Biochem Biophys Res Commun, 1994. **203**(2): p. 798-804.

110. Gottesman, S., et al., *ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities.* J Biol Chem, 1993. **268**(30): p. 22618-26.

111. Nurse, P., et al., *Molecular cloning and DNA sequence analysis of Escherichia coli priA, the gene encoding the primosomal protein replication factor Y.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4615-9.

112. Lee, E.H., et al., *The priA gene encoding the primosomal replicative n' protein of Escherichia coli.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4620-4.

113. Ng, J.Y. and K.J. Marians, *The ordered assembly of the phiX174-type primosome. I. Isolation and identification of intermediate protein-DNA complexes.* J Biol Chem, 1996. **271**(26): p. 15642-8.

114. Rangarajan, S., R. Woodgate, and M.F. Goodman, *Replication restart in UV-irradiated Escherichia coli involving pols II, III, V, PriA, RecA and RecFOR proteins.* Mol Microbiol, 2002. **43**(3): p. 617-28.

115. Bhattacharyya, B., et al., *Structural mechanisms of PriA-mediated DNA replication restart.* Proc Natl Acad Sci U S A, 2014. **111**(4): p. 1373-8.

116. Al-Deib, A.A., A.A. Mahdi, and R.G. Lloyd, *Modulation of recombination and DNA repair by the RecG and PriA helicases of Escherichia coli K-12.* J Bacteriol, 1996. **178**(23): p. 6782-9.

117. Gregg, A.V., et al., *Direct rescue of stalled DNA replication forks via the combined action of PriA and RecG helicase activities.* Mol Cell, 2002. **9**(2): p. 241-51.

118. Cadman, C.J. and P. McGlynn, *PriA helicase and SSB interact physically and functionally.* Nucleic Acids Res, 2004. **32**(21): p. 6378-87.

119. Cadman, C.J., et al., *PriB stimulates PriA helicase via an interaction with single-stranded DNA.* J Biol Chem, 2005. **280**(48): p. 39693-700.

120. Chen, H.W., S.H. North, and H. Nakai, *Properties of the PriA helicase domain and its role in binding PriA to specific DNA structures.* J Biol Chem, 2004. **279**(37): p. 38503-12.

121. Lenski, R.E., et al., *Long-term experimental evolution in Escherichia-coli.1. Adaptation and divergence during 2,000 generations.* American Naturalist, 1991. **138**(6): p. 1315-1341.

122. Liu, Y., et al., *Transcriptome dynamics of Deinococcus radiodurans recovering from ionizing radiation.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 4191-6.

123. Lopper, M.E., J. Boone, and C. Morrow, *Deinococcus radiodurans PriA is a Pseudohelicase.* PLoS One, 2015. **10**(7): p. e0133419.

124. Heller, R.C. and K.J. Marians, *The disposition of nascent strands at stalled replication forks dictates the pathway of replisome loading during restart.* Molecular Cell, 2005. **17**(5): p. 733-743.

125. Heller, R.C. and K.J. Marians, *Unwinding of the nascent lagging strand by rep and PriA enables the direct restart of stalled replication forks.* Journal of Biological Chemistry, 2005. **280**(40): p. 34143-34151.

126. Gort, A.S., D.M. Ferber, and J.A. Imlay, *The regulation and role of the periplasmic copper, zinc superoxide dismutase of Escherichia coli.* Mol Microbiol, 1999. **32**(1): p. 179-91.

127. Povirk, L.F., *Bleomycin-induced mutagenesis in repackaged lambda phage: base substitution hotspots at the sequence C-G-C-C.* Mutat Res, 1987. **180**(1): p. 1-9.

128. Singh, A. and Y.J. Xu, *The Cell Killing Mechanisms of Hydroxyurea.* Genes (Basel), 2016. **7**(11).

129. Hendricks, S.P. and C.K. Mathews, *Differential effects of hydroxyurea upon deoxyribonucleoside triphosphate pools, analyzed with vaccinia virus ribonucleotide reductase.* J Biol Chem, 1998. **273**(45): p. 29519-23.

130. Sakano, K., et al., *Hydroxyurea induces site-specific DNA damage via formation of hydrogen peroxide and nitric oxide.* Jpn J Cancer Res, 2001. **92**(11): p. 1166-74.

131. Weinberger, M., et al., *Induction by adozelesin and hydroxyurea of origin recognition complex-dependent DNA damage and DNA replication checkpoints in Saccharomyces cerevisiae.* J Biol Chem, 1999. **274**(50): p. 35975-84.

132. Amundsen, S.K., A.F. Taylor, and G.R. Smith, *The RecD subunit of the Escherichia coli RecBCD enzyme inhibits RecA loading, homologous recombination, and DNA repair.* Proceedings of the National Academy of Sciences, 2000. **97**(13): p. 7399-7404.

133. Sivaramakrishnan, P., et al., *The transcription fidelity factor GreA impedes DNA break repair.* Nature, 2017. **550**(7675): p. 214-218.

134. Cashel, M., D. R. Gentry, V. J. Hernandez, and D. Vinella. , *The Stringent Response.* Escherichia coli and Salmonella: cellular and molecular biology, 1996. **1**(ASM Press): p. 1458-95.

135. Trautinger, B.W. and R.G. Lloyd, *Modulation of DNA repair by mutations flanking the DNA channel through RNA polymerase.* EMBO J, 2002. **21**(24): p. 6944-53.

136. McGlynn, P. and R.G. Lloyd, *Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression.* Cell, 2000. **101**(1): p. 35-45.

137. Epshtein, V., et al., *UvrD facilitates DNA repair by pulling RNA polymerase backwards.* Nature, 2014. **505**(7483): p. 372-7.

138. Landick, R., J. Stewart, and D.N. Lee, *Amino acid changes in conserved regions of the beta-subunit of Escherichia coli RNA polymerase alter transcription pausing and termination.* Genes Dev, 1990. **4**(9): p. 1623-36.

139. Barrick, J.E., et al., *Escherichia coli rpoB mutants have increased evolvability in proportion to their fitness defects.* Mol Biol Evol, 2010. **27**(6): p. 1338-47.

140. Severinov, K., et al., *Rifampicin region revisited. New rifampicin-resistant and streptolydigin-resistant mutants in the beta subunit of Escherichia coli RNA polymerase.* J Biol Chem, 1993. **268**(20): p. 14820-5.

141. Zhou, Y.N., et al., *Isolation and characterization of RNA polymerase rpoB mutations that alter transcription slippage during elongation in Escherichia coli.* J Biol Chem, 2013. **288**(4): p. 2700-10.

142. Davies, M.J., *Protein oxidation and peroxidation.* Biochem J, 2016. **473**(7): p. 805-25.

143. Yun, M., et al., *Structural analysis of glyceraldehyde 3-phosphate dehydrogenase from Escherichia coli: direct evidence of substrate binding and cofactor-induced conformational changes.* Biochemistry, 2000. **39**(35): p. 10702-10710.

144. Liu, Y., S.C. Bauer, and J.A. Imlay, *The YaaA protein of the Escherichia coli OxyR regulon lessens hydrogen peroxide toxicity by diminishing the amount of intracellular unincorporated iron.* Journal of bacteriology, 2011. **193**(9): p. 2186-2196.

145. Kaur, P., J.G. Kiselar, and M.R. Chance, *Integrated algorithms for high-throughput examination of covalently labeled biomolecules by structural mass spectrometry.* Anal Chem, 2009. **81**(19): p. 8141-9.

146. Davies, M.J.D., R. T. , *Radical-mediated protein oxidation: from chemistry to medicine*. 1997: Oxford University Press. 44-5.

147. Weijland, A., et al., *Elongation factor Tu: a molecular switch in protein biosynthesis.* Molecular microbiology, 1992. **6**(6): p. 683-688.

148. Goodall, E.C., et al., *The essential genome of Escherichia coli K-12.* mBio, 2018. **9**(1): p. e02096-17.

149. Miller, J.H., *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. 1992, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

150. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.

151. Warming, S., et al., *Simple and highly efficient BAC recombineering using galK selection.* Nucleic Acids Res, 2005. **33**(4): p. e36.

152. Powell, B.S., et al., *Rapid confirmation of single copy lambda prophage integration by PCR.* Nucleic Acids Research, 1994. **22**(25): p. 5765.

153. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-2120.

154. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.

155. Li, H., et al., *The sequence alignment/map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-2079.

156. Koboldt, D.C., et al., *VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing.* Genome research, 2012. **22**(3): p. 568-576.

157. Cingolani, P., et al., *A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3.* Fly, 2012. **6**(2): p. 80-92.

158. Wiser, M.J. and R.E. Lenski, *A Comparison of Methods to Measure Fitness in Escherichia coli.* PLoS One, 2015. **10**(5): p. e0126210.

159. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4350-4.

160. Windgassen, T.A. and J.L. Keck, *An aromatic-rich loop couples DNA binding and ATP hydrolysis in the PriA DNA helicase.* Nucleic acids research, 2016. **44**(20): p. 9745-9757.

161. Manthei, K.A., et al., *Structural mechanisms of DNA binding and unwinding in bacterial RecQ helicases.* Proceedings of the National Academy of Sciences, 2015. **112**(14): p. 4292-4297.