The discovery and characterization of NHJ-1, a novel regulator of canonical non-homologous end joining in the nematode *Caenorhabditis elegans*

Aleksandar Vujin

Department of Biology

McGill University, Montreal, Quebec

January 2019

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

Table of contents

List of Figures and Tables

Abstract

The physical and informational integrity of DNA is of paramount importance to the survival, development, and reproduction of living organisms. Of the various forms of damage which can arise from the numerous intrinsic and extrinsic forms of genotoxic stress, the most serious are double strand breaks (DSBs). These lesions are dangerous not only because of their mutagenic potential, but because they disrupt the physical continuity of DNA molecules and risk the loss of all genes distal to the break. Repair of DSBs is handled by several distinct pathways, of which the classical non-homologous end joining (cNHEJ) is the most prominent in somatic tissues. The mechanism of cNHEJ involves three steps: 1) the rapid recognition and binding of the free DNA end by the heterodimeric Ku ring protein; 2) the processing of the damaged DNA ends by a complement of nucleases, kinases, phosphatases, and polymerases to yield ligation-compatible ends; and 3) the terminal ligation by Ligase IV. Despite the wide conservation of cNHEJ proteins among eukaryotes, the nematode *Caenorhabditis elegans* is unique in that it does not possess homologs of known cNHEJ processing enzymes or structural scaffolds. This raises the question of whether a basic cNHEJ system consisting only of the Ku ring and the Ligase IV ortholog LIG-4 suffices in *C. elegans* or whether it possesses non-conserved functional analogs of other known eukaryotic cNHEJ factors.

In this thesis, I provide a definitive answer to this question by showing that *nhj-1*, a previously uncharacterized gene unique to the nematode order Rhabditida, is an indispensable cNHEJ factor in *C. elegans*. I show that some lines of the most commonly used *C. elegans* wild type strain, N2, exhibit unexpected stressor- and developmental stage-specific sensitivity to ionizing radiation exposure at the first larval stage (L1). This sensitivity is characterized by a high incidence of growth delay and abnormal morphological phenotypes, as well as a greatly reduced number of progeny. However, there is no evidence of germline DNA damage, which is evocative of known cNHEJ mutants in *C. elegans*. Using deep sequencing and CRISPR-Cas9 mutagenesis, I reveal the existence of a spontaneous *nhj-1* mutation in the sensitive N2 lines, and show that IR-sensitivity arises because of the resulting *nhj-1* loss of function. Furthermore, the IR-sensitivity of *nhj-1*

mutants is as severe as that of *lig-4* or Ku ring mutants, and double mutants do not show additive IR-sensitivity, demonstrating that *nhj-1* is part of the cNHEJ pathway. I also provide the first characterization of the subcellular localization of NHJ-1 and LIG-4 in the *C. elegans* L1 larva, the adult germline, and the adult intestine, showing a lack of complete overlap in the NHJ-1 and LIG-4 expression patterns in the L1 larva. Finally, I employ a germline cNHEJ-enabled genetic background to show that NHJ-1 likely acts downstream of Ku ring binding, at least in the context of the adult germline.

My work reveals that *C. elegans* has restructured the ancient cNHEJ pathway to include an entirely novel regulator. It raises further questions about the regulation and mechanism of cNHEJ activity in this important model organism, and opens several research avenues in which the answers to those questions may be sought.

Résumé

L'intégrité physique et informationnelle de l'ADN est d'une importance capitale pour la survie, le développement et la reproduction d'organismes vivants. Parmi les diverses formes de dommages pouvant résulter des nombreuses formes intrinsèques et extrinsèques de stress génotoxique, les plus graves sont les cassures à double brin (DSB). Ces lésions sont dangereuses non seulement en raison de leur potentiel mutagène, mais également parce qu'elles perturbent la continuité physique des molécules d'ADN et risquent la perte d'un grand nombre de gènes situés en aval de la rupture. La réparation des DSBs est gérée par plusieurs voies de réparation distinctes, parmi lesquelles la classique jonction d'extrémité non homologue (cNHEJ) est la plus importante dans les tissus somatiques. Le mécanisme de cNHEJ comprend trois étapes: 1) la reconnaissance et liaison rapides de l'extrémité de l'ADN libre par la protéine hétérodimère Ku en forme d'anneau; 2) le traitement des extrémités d'ADN endommagées par un complément de nucléases, kinases, phosphatases et polymérases afin d'obtenir des extrémités compatibles avec la ligature; et 3) la ligature terminale par la ligase IV. Malgré la large conservation des protéines cNHEJ chez les eucaryotes, le nématode *Caenorhabditis elegans* est unique en ce qu'il ne possède aucun homologue d'enzymes de traitement de la voie cNHEJ ni d'échafaudages structuraux. Cela soulève la question de savoir si un système de cNHEJ de base composé uniquement de l'anneau Ku et de l'orthologue LIG-4 de Ligase IV suffit chez *C. elegans* ou s'il possède des analogues fonctionnels non conservés d'autres facteurs cNHEJ eucaryotes bien connus.

Dans cette thèse, je réponds de manière définitive à cette question en montrant que *nhj-1*, un gène unique à l'ordre des nématodes Rhabditida et jusque-là non caractérisé, est un facteur cNHEJ indispensable chez *C. elegans*. Je montre que certaines lignées de la souche de type sauvage de *C. elegans* la plus couramment utilisée, N2, présentent une sensibilité inattendue à l'exposition aux rayonnements ionisants au premier stade larvaire (L1), qui est spécifique au type de stress et au stage de développement. Je montre que la lignée N2 sensible à l'IR présente une incidence élevée de retard de croissance et de phénotypes morphologiques anormaux, ainsi qu'un nombre considérablement réduit de descendants. Cependant, rien n'indique que l'ADN

germinal ait été endommagé, ce qui évoque des mutants cNHEJ connus chez *C. elegans*. En utilisant le séquençage en profondeur et la mutagenèse CRISPR-Cas9, je révèle l'existence d'une mutation spontanée de *nhj-1* dans les lignées sensibles de N2 et montre que cette sensibilité résulte de l'inactivation de ce gène. Je montre également que la sensibilité des mutants *nhj-1* est aussi sévère que celle des mutants des anneaux *lig-4* ou Ku, et que les doubles mutants ne montrent pas de sensibilité additive à l'IR, démontrant que *nhj-1* fait partie de la voie de cNHEJ. Je fournis également la première caractérisation de la localisation subcellulaire de NHJ-1 et LIG-4 dans les larves de *C. elegans* L1, la lignée germinale adulte et l'intestin adulte, montrant un manque de chevauchement complet des modèles d'expression de NHJ-1 et LIG-4 dans les larves L1. Enfin, j'utilise un contexte génétique dans lequel la voie cNHEJ est activée dans la lignée germinale pour montrer que NHJ-1 agit probablement en aval de la liaison de l'anneau Ku, au moins dans le contexte de la lignée germinale adulte.

Mon travail révèle que *C. elegans* a restructuré l'ancienne voie du cNHEJ pour y inclure un régulateur entièrement nouveau. Cela soulève d'autres questions sur la régulation et le mécanisme de l'activité du cNHEJ dans cet organisme modèle important et ouvre plusieurs pistes permettant de rechercher des réponses à ces questions.

List of abbreviations

53BP1 = p53-binding protein 1 (protein)

AID = activation induced cytidine deaminase (protein)

APE = AP endonuclease (protein)

APLF = aprataxin-like factor (protein)

AP site = apurinic/apyrimidinic site (DNA lesion)

APTX = aprataxin (protein)

ARD = adult reproductive diapause (developmental stage)

ATM = ataxia-telangiectasia mutated (protein)

BER = base excision repair (DNA repair pathway)

BIR = break-induced replication (DNA repair pathway)

BRCA1 = breast cancer 1 (protein)

cNHEJ = canonical non-homologous end joining (DNA repair pathway)

CPD = cyclobutane pyrimidine dimer (DNA lesion)

CSA = Cockayne syndrome protein A (protein)

CSB = Cockayne syndrome protein B (protein)

CSR = class switch recombination (process)

dHJ = double Holliday junction (transient DNA structure)

DNA = deoxyribonucleic acid (macromolecule)

DNA-PK = DNA-dependent protein kinase (protein complex)

DNA-PKcs = DNA-dependent protein kinase catalytic subunit (protein)

dNTP = deoxyribonucleotide triphosphate (molecule)

DSB = DNA double strand break (DNA lesion)

dsDNA = double stranded DNA (DNA type)

EJ = end joining (DNA repair pathway type)

FATC = FRAP-ATM-TRAPP-C-terminal (domain)

FHA = forkhead-associated (domain)

Gy = Gray (unit of ionizing radiation dose)

H2AX = histone 2AX (histone variant)

H3K9 = histone 3, lysine (K) 9 (histone residue)

H3K9Ac = acetylated H3K9 (histone mark)

H3K9me3 = tri(3)methylated H3K9 (histone mark)

Him = high incidence of males (phenotype)

HIT = histidine triad (protein family)

HR = homologous recombination (DNA repair pathway type)

ICL = interstrand crosslink (DNA lesion)

IDL = insertion-deletion loop (DNA lesion)

IR = ionizing radiation (genotoxic stressor)

MDC1 = mediator of DNA damage checkpoint 1 (protein)

MMR = mismatch repair (DNA repair pathway)

modENCODE = model organism Encyclopedia of DNA Elements (project)

mRNA = messenger RNA (RNA type)

NHEJ = see cNHEJ

PALF = PNKP and APTX-like FHA protein; see APLF

PARylation = poly-ADP ribosylation (process)

PAXX = paralog of XRCC4 and XLF (protein)

PIKK = phosphoinositide 3-kinase related kinases (protein class)

PNKP = polynucleotide kinase 3'-phosphatase (protein)

PRD = PIKK-regulatory domain (protein domain)

[R] = IR-resistant genetic background

RAD23B = Radiation sensitive 23 homolog B (protein)

RAG = recombinase activating gene (protein complex)

RdRP = RNA dependent RNA polymerase (protein class)

RNA = ribonucleic acid (macromolecule)

RNAi = RNA interference (process)

rNTP = ribonucleotide triphosphate (molecule)

[S] = IR-sensitive genetic background

SC = synaptonemal complex (proteinaceous structure)

SDSA = synthesis-dependent strand annealing (DNA repair pathway)

SHM = somatic hypermutation (process)

siRNA = small interfering RNA (RNA type)

SSA = single strand annealing (DNA repair pathway)

ssDNA = single stranded DNA (DNA type)

TDP1 = tyrosyl DNA phosphodiesterase 1 (protein)

TFIIH = transcription initiation factor II H (protein)

TLS = translesion synthesis (DNA damage tolerance mechanism)

UVR = ultraviolet radiation (genotoxic stressor)

XLF = XRCC4-like factor (protein)

XPC = xeroderma pigmentosum complementation group C (protein)

XRCC = X-ray repair cross complementing (protein class)

Acknowledgements

This thesis represents the culmination of extraordinary efforts that took place over the better part of the last decade, which would not have been possible without the patience and unwavering support of the many people who, whether in an official capacity or not, helped me achieve it. I apologize for slighting those persons who could not be individually thanked here because of space considerations. You are all remembered and cherished.

I would first like to thank my supervisor, Dr. Monique Zetka, for her guidance and supervision during these many years. Monique, I am immensely thankful that you allowed me the independence to pursue this most challenging project and uncover something that neither of us would have envisioned when I walked into your office to ask for a graduate student position. The perseverance of your confidence in me during the long periods when things looked bleak stands as a testament to exemplary mentorship.

I would also like to thank Dr. Richard Roy, whose lab is a close collaborator of ours, and who also served on my Supervisory Committee. Rik's insightful comments and critique, frankly and sincerely offered during the joint weekly group meetings and Supervisory Committee meetings, greatly benefited both the advancement of my project and my growth as a scientist.

I am also very grateful to Dr. Laura Nilson, who in her capacity as a member of my Supervisory Committee always provided astute and useful observations about my project, and made her expertise available whenever I asked for advice.

My thanks also go to my undergraduate and honours project supervisors, Dr. Thomas Bureau, Dr. François Fagotto, and Dr. Gregor Fussmann, whose invaluable training during the early steps of my research career gave me the skills and confidence to embark on the journey to a graduate degree.

I was particularly lucky to have worked with some exceptionally talented scientists and magnanimous people in the Zetka and Roy laboratories. To Dr. Sara Labella and Dr. Pratik Kadekar, who were there from the beginning to the end, thank you for all your scientific expertise

and advice, but above all your companionship on the long and difficult road we shared. To Ms. Jasmin Hanafi, thank you for being such a great friend from the early days all the way through the present. To Dr. Florence Couteau and Dr. Budget Saribek, thank you both for showing me what it means to be a dedicated researcher. To Ms. Hanh Nguyen, Dr. Annina Spilker, Dr. Sebastian Schmeißer, Mr. Nathan Navidzadeh, Ms. Anna Kazanets, Ms. Naomi Amberber, and Mr. Ryan Dawson, thank you for your friendship in the lab and beyond. To Dr. Meng Xie, Dr. Yu Lu, Dr. Mina Ahmadi, and other Roy and Zetka lab members too numerous to list, thank you for making the lab such a fun and stimulating environment to work in. To Mr. Shaolin Li, thank you very much for the technical prowess and troubleshooting skill you were always willing to share. I would be remiss if I didn't also thank the numerous work study students who have worked in the lab over the years, whose diligent work produced the seemingly inexhaustible stream of consumable reagents that my research required.

I would also like to thank my collaborators, Dr. Steven Jones and his student Mr. Harwood Kwan, for their hard work and bioinformatic expertise which were instrumental to the discovery of *nhj-1*. No less essential to the discovery was the work of my colleague Ms. Maja Kaplan from the lab of Dr. Thomas Bureau, and that of Ms. Anja Bošković, who not only a colleague but a great friend with whom I have spent countless fun-filled hours during and since our time in the lab.

During my tenure in Monique's lab, I also mentored several undergraduate students, including Mr. Kevin Dick and Mr. Julyan Baruch, who have since gone on to do great work in graduate school, and have become dear friends. In the last few years, I have also closely advised Mr. Muhammad Tabassum during his Masters project in the lab, who also became a treasured friend. To all of you, and to the students I have TAed over the years, thank you for teaching me what it truly means to be a teacher.

To my parents and my brother, thank you for providing me with unconditional emotional support and material abundance which made me want for nothing while I pursued the life of the mind, and for fighting so hard to secure for me a life in Canada. Najdraži moji, neizmerno vam hvala.

Finally, to my best friend, soul mate, and love, Camille, thank you, above all, for loving me and being there for me through thick and thin (mostly thick). Merci toi.

Contribution to original knowledge

This thesis represents a contribution to original knowledge by showing for the first time:

1. That a cryptic mutation in the previously uncharacterized gene *nhj-1* segregates in several *C. elegans* laboratory strains, including lines of the most commonly used wild type strain N2, sensitizing them to ionizing radiation during early development.

2. That NHJ-1 is a critical regulator of the canonical non-homologous end joining DSB repair pathway, and that it acts downstream of the Ku ring.

3. The subcellular localization of the canonical non-homologous end joining pathway components NHJ-1 and LIG-4 in the L1 larva, and the adult germline and intestinal cells.

Contribution of authors

The text of this thesis has been entirely composed and written by me, Aleksandar Vujin, and has been edited by me on the basis of the valuable advice and comments of professor Monique Zetka, Dr. Pratik Kadekar, Dr. Sara Labella, and Ms. Camille Delouche. Ms. Camille Delouche provided editorial comments and translated the abstract into French.

After the oral defense, additional edits were made based on the comments from the internal thesis examiner, Prof. Joseph Dent, and the external thesis examiner, Prof. Brent Derry, as well as the oral defense committee member Prof. Anna Naumova.

All experiments described herein are entirely my own work, with the supervision and investigative input of professor Monique Zetka, except as follows. Dr. Steven Jones (University of British Columbia) and his team (Mr. Harwood Kwan and Dr. Yaoqing Shen) have bioinformatically analyzed the deep sequencing data and provided the list of single nucleotide variants in the N2 [S] genome which have an impact on protein sequences (**Table 2.2**). Ms. Anja Bošković performed the genome alignment and Ms. Maia Kaplan generated the variant call format (VCF) files which enabled me to identify the IR-sensitizing mutation in *nhj-1*. Mr. Harwood Kwan uncovered the full sequence of the N2 [S] mutation in *nhj-1* (**Figure 3.1**).

I prepared all the Figures and Tables in this thesis, including the ones based on collaborative efforts, with the exception of **Figure 1.3** and **Figure 1.4**, which are adapted from Wormatlas under the non-profit scientific use license (http://www.wormatlas.org/copyrightanduse.htm).

A note on language

Throughout most of this thesis, the first person singular pronoun is used to describe experiments which I have performed alone. The use of the first person plural pronoun is reserved to describe experiments or procedures done with collaborators, which include Dr. Monique Zetka, and Dr. Steven Jones and his laboratory personnel (see Contribution of authors).

"Science in the service of appreciation,

and appreciation in the service of reverence,

which, in the face of wonders that are not of our making,

is our only proper response"

Colin Tudge, in "The Tree" (2006)

Chapter I: Introduction and literature review

Introduction

This thesis details the story of the discovery and initial characterization of a novel gene required for canonical non-homologous end joining (cNHEJ), a highly conserved DNA double-strand break (DSB) repair pathway in the nematode *Caenorhabditis elegans (C. elegans)*. **Chapter I** presents a review of the literature relevant for both understanding the experiments and theoretical frameworks which support the findings in this thesis, but also for situating them in a wider context. It is divided into three parts: Part I introduces the various types and sources of DNA damage, with emphasis on ionizing radiation (IR), the most efficient DSB-inducing genotoxic stressor. It also contains an overview of the best understood DNA repair pathways, placing cNHEJ in a larger context. This discussion focuses on mammalian systems, as our understanding of DNA repair pathways primarily comes from studies in mammals, but deviations from this pattern in *C. elegans* are noted in Part III. Part II discusses cNHEJ in more detail, introducing the hitherto identified molecular players which play a role in the process, the steps of the repair mechanism, its regulation in the context of competing alternative DSB repair pathways in different tissues or cell cycle stages, its evolutionary conservation and its relevance in mammalian development and human disease. Part III consists of the introduction to the roundworm *C. elegans* as a model organism, focusing on those aspects of its biology most relevant to understand the experimental procedures in the dissertation, as well as its conclusions. **Chapter II** describes the serendipitous discovery of phenotypic diversity in the IR response among lines of the wild type strain N2, and presents evidence that some N2 lines carry an unannotated recessive mutation which results in IR sensitivity similar to that of cNHEJ mutants. **Chapter III** describes the mapping of this mutation to *nhj-1*, a previously uncharacterized gene, and presents the evidence that NHJ-1 is a cNHEJ factor which acts downstream of the Ku ring. The cellular and tissue localization patterns of NHJ-1 and the terminal cNHEJ effector LIG-4 are also described. The significance of these findings and their implications are discussed in **Chapter IV**, as are the unanswered questions and avenues for further research. The experimental methodology is detailed in **Chapter V**.

Literature review, part I: DNA damage and repair

I.1.1 DNA and its structure as the basis of the genetic code

Life on planet Earth, and possibly in general, requires a set of instructions for growth, development, and reproduction, that can be inherited (transmitted to progeny). These instructions in turn require a physical basis, which in the case of Earth's biota is provided by the sturdy molecules called nucleic acids. In all non-viral forms of life discovered so far, the totality of heritable units (genes), also known as the genome, is based on deoxyribonucleic acid (DNA).

Nucleic acids are polymers composed of covalently linked nucleotide units, and non-viral genomic DNA is exclusively found in the form of two long strands bound to each other into a helical structure by non-covalent links like hydrogen bonding and nucleotide stacking (for a general review of DNA structure on which the summary presented below is based, see (Bowater and Waller 2014)). Nucleotides are composed of a pentameric sugar (deoxyribose in DNA, ribose in RNA) linked to a nitrogenous base via a glycosidic bond on its 1' carbon atom, and to a phosphate group on its 5' carbon. Nucleotides assemble into nucleic acid polymers through covalent phosphodiester bonding between the 5' phosphate group of one nucleotide and the 3' hydroxyl group of another nucleotide. A linear single stranded DNA molecule possesses polarity, with a phosphorylated 5' terminus and a hydroxyl 3' terminus. In double stranded DNA, the strands are held together by nitrogenous base interactions (see below) and are oriented in an antiparallel fashion, such that each end of double stranded DNA contains the 5' terminus of one strand and the 3' terminus of the other. Because the phosphodiester bonds that hold a single DNA strand together involve only deoxyribose and the phosphate groups, the term "sugar backbone" is often used to describe this part of the molecule, in contrast to the nitrogenous bases. There are five nitrogenous bases, also known as nucleobases, which are most commonly found in nucleotides. All are derived from nitrogen-containing heterocyclic aromatic carbon rings, with three nucleobases (uracil, thymine, and cytosine) based on pyrimidine, a six-atom ring, and two (adenine, guanine) based on purine, a nine-atom fusion of pyrimidine and imidazole. While not strictly correct, these nucleobases are often referred to as "purines" and "pyrimidines"

after the aromatic ring from which they derive, and such usage is retained here. Under physiological conditions, a purine from a nucleotide in one strand of DNA will interact by hydrogen bonding with a pyrimidine from the other strand to stabilize the double helix. Adenine bonds with thymine, while guanine binds to cytosine. Uracil is normally found only in RNA, where it replaces thymine in bonding with adenine. This specificity in base binding is what enables the mechanism of DNA replication, in which the hydrogen bonds between DNA strands are disrupted, and new DNA strands synthesized by the addition of nucleotides with bases complementary to the ones in the template strand.

The order of nucleobases in the DNA polymer represents the basis for the genetic code. Two macromolecules critical for life processes are synthesized based on the information contained in the DNA base sequence: ribonucleic acid (RNA) is synthesized much like DNA itself, using base complementarity with existing DNA to create new RNA chains, while proteins, chemically and structurally versatile polymers composed of amino acid monomers, are synthesized by the addition of one amino acid for every three nucleotides (a triplet, or codon) of a DNA sequence, via a messenger RNA (mRNA) intermediary.

An organism's genome is organized in the form of chromosomes, long DNA polymers, which may be circular or linear. On the chromosomes are located sequences which encode individual protein and RNA molecules - the coding sequences or genes (though the exact definition of a gene may be more inclusive than this – for an excellent review of the continually evolving definition of the term "gene", see (Portin and Wilkins 2017)). Chromosomes also contain sequences which regulate the temporal and spatial expression (activity) of genes, as well as non-coding sequences with important roles in chromosome maintenance, such as centromeres, which are required for proper chromosome segregation in cell division (Fukagawa and Earnshaw 2014), and telomeres, which cap linear chromosomes and prevent chromosomal fusion via inappropriate DNA repair as well as excessive chromosome shortening during DNA replication (O'Sullivan and Karlseder 2010). In addition to functional sequences, chromosomal DNA includes intergenic "filler" or "junk" sequences, and parasitic elements such as transposons which have the potential to replicate faster than the rest of the genome (Palazzo and Gregory 2014).

I.1.2 Genotoxic stressors

As a physical entity, the DNA polymer can be damaged in a variety of ways by genotoxic agents found in the environment (exogenous genotoxins) and within the cell or the organism itself (endogenous genotoxins). The exogenous/endogenous division is used here to broadly classify common DNA damaging agents, but it is not the only classification method. One could differentiate between chemical and physical agents, biogenic or non-biogenic, or categorize them based on the type of damage inflicted to the DNA (*eg*, base adduct, nick in the sugar backbone, double-strand break, *etc*), and such distinctions among genotoxic stressors described below are pointed out where salient.

I.1.3 Molecular, cellular, and physiological consequences of DNA damage

Before we move on to a discussion of specific genotoxic stressors, it is worthwhile to briefly consider the multitudinous effects DNA damage can have on a cell or an organism. On the lower end of the severity spectrum, a single nucleobase within a coding region may be chemically altered such that it no longer pairs with the appropriate partner, which can affect RNA transcription, and in turn the translation of mRNA into a protein. Because the genetic code is degenerate, *ie* an amino acid can be encoded by several codons (Hartman and Smith 2014), a single nucleotide change in the DNA will not necessarily result in an amino acid change in the protein sequence. Furthermore, a single nucleotide mutation need not impact protein function even if it does cause an amino acid substitution, or impact it only slightly. Greater in severity are changes which more acutely disrupt protein or RNA function, including both single nucleotide alterations which cause a substitution in a functionally critical amino acid in the protein or nucleotide residue in the RNA, as well as larger scale changes to the DNA, such as a loss of many nucleotides. In this latter case, many monomer residues may be lost from or inappropriately incorporate into the protein or RNA, with a greater chance of a negative functional outcome. In the case of protein-coding genes, a deletion or insertion of a number of nucleotides indivisible by three results in a frame-shift, *ie*, a change in codon identity downstream of the damaged site, leading to missense translation and loss of normal protein function. The most severe kinds of damage may result in a simultaneous loss of many genes, as in the case of large deletions caused

by unrepaired or improperly repaired DSBs. On a cellular level, the effects of such genomic injury may range from reduced function to senescence and death. If the damage to DNA is not repaired before DNA replication occurs in the affected cell, the changes to the nucleotide sequence may become fixed and transmitted to subsequent cellular generations as mutations. On the organismal level, cancer is one of the potential outcomes of such mutations (affecting cell proliferation, growth, and identity), while aging is hypothesized to be another (Bouwman and Jonkers 2012, Wolters and Schumacher 2013). It is therefore unsurprising that pathways to counteract DNA damage and its effects have evolved early in the history of life (Caplin and Willey 2018). Furthermore, cells have evolved to utilize and indeed produce DNA damage in a controlled setting to enable critical life processes. Examples include topoisomerase-mediated induction of transient breaks in the DNA to resolve the topological problems that arise from DNA replication and transcription of RNA (Wang 2002), generation of DSBs to promote genetic exchange between parental chromosomes in meiosis (Zickler and Kleckner 2015), the recombination which allows mate-type switching in *Saccharomyces cerevisiae (S. cerevisiae)* (Haber 2012) and the V(D)J recombination responsible for antigen receptor diversity in the mammalian immune system (Malu, Malshetty et al. 2012). The remainder of Part I will focus on pathological DNA damage induced by genotoxic agents.

I.1.4 Endogenous genotoxic stressors

As mentioned above, genotoxic stressors can be divided into two groups depending on whether their source is endogenous or exogenous. Endogenous genotoxic stress can be further subdivided into two groups: aberrant or erroneous enzymatic action by endogenous proteins, and chemical or mechanical damage caused by the interaction of the DNA chains with normal cellular metabolites and internal environments.

I.1.4.1 Replication errors

One source of endogenous enzymatic genotoxic stress comes from DNA polymerases during DNA synthesis, which may erroneously incorporate an inappropriate nucleotide into the nascent DNA strand (Loeb and Monnat 2008). The 14 known human DNA-dependent DNA polymerases vary greatly in error rate. High fidelity replicative polymerases Pol δ and Pol ε commit an incorporation

error at a rate of 10^{-6} or less. Low fidelity polymerases such as Pol η, which can synthesize DNA through lesions which would block the standard replicative complexes in a process called translesion synthesis (TLS), can have error rates as high as 10^{-2} (Loeb and Monnat 2008). High fidelity polymerases have a lower error rate not only because of an intrinsically higher selectivity, but because they often also contain a proofreading exonuclease activity which immediately removes misincorporated nucleotides (Kunkel 2009). Apart from the intrinsic error rate of the polymerase, a mismatched nucleotide is more likely to be incorporated when the ratio of deoxyribonucleotide triphosphates (dNTPs) is unbalanced (Buckland, Watt et al. 2014).

In addition to base mismatches, replication errors also include insertions or deletions caused by the slippage of the primer and template strands in relation to each other, which can occur in regions with repetitive sequences (Mittelman 2013). Moreover, ribonucleotide triphosphate (rNTP) units may be incorporated into the DNA strand instead of dNTPs, which can distort the shape of the DNA double helix and make the sugar backbone prone to breakage (Potenski and Klein 2014). Polymerases may also erroneously incorporate deoxyribouridine triphosphate (dUTP), which despite being chemically equivalent to deoxyribothymidine triphosphate (dTTP), presents a problem in the context of DNA as it may be misinterpreted by the DNA repair machinery as a product of cytosine deamination, which occurs frequently (Vertessy and Toth 2009). Most replication errors which are not immediately corrected by proofreading are repaired by mismatch repair (MMR, see section **I.1.6.3**).

I.1.4.2 Aberrant enzymatic activity

Pathological activity by topoisomerases and the RAG complex represent another source of endogenous enzymatic DNA damage. Topoisomerases are enzymes specialized to relieve the topological stress of supercoiling in the DNA which results from replication or transcription (Wang 2002). Two main classes of topoisomerases exist: class I topoisomerases cut DNA one strand at a time, creating transient single-strand breaks (SSBs); class II topoisomerases cut both DNA strands at the same time, creating a DSB (Wang 2002). After SSB or DSB induction, topoisomerases transiently bind to the DNA end via a covalent bond between a tyrosine residue in the enzyme and the phosphate of the nucleotide (Pommier, Barcelo et al. 2006). Known as

cleavage complexes, these structures are normally transient, but in certain cases may be stabilized, leaving an exposed DNA end (Adachi, Suzuki et al. 2003, Pommier, Barcelo et al. 2006). Class I topoisomerase-induced SSBs are repaired by the SSB repair pathway (section **I.1.6.4**), while the DSBs caused by class II topoisomerases are predominantly repaired by cNHEJ (see **Part II** of this chapter) or a homologous recombination-based pathway (sections **I.1.6.5.1a-d**).

Enzymes other than topoisomerases have been proposed to induce pathological DSBs. Activation induced cytidine deaminase (AID) normally functions to initiate an enzymatic cascade which produces DSBs essential for class switch recombination (CSR) and somatic hypermutation (SHM), mechanisms which are critical for antibody diversity (Robbiani, Bothmer et al. 2008). However, AID has been shown to be required for DSBs which lead to oncogenic translocations (Robbiani, Bothmer et al. 2008). AID also cooperates with the recombinase activating gene (RAG) endonuclease complex, which normally creates DSBs during the V(D)J recombination process to generate antigen receptor diversity in lymphocytes, to create off-target DSBs (Mahowald, Baron et al. 2008, Tsai, Lu et al. 2008).

I.1.4.3 Spontaneous hydrolysis

Spontaneous hydrolysis affects exocyclic amines in nucleobases and glycosidic bonds between nucleobases and deoxyribose, and is major non-enzymatic source of endogenous DNA damage. Base deamination produces uracil from cytosine, hypoxanthine from adenine, xanthine from guanine, and thymine from 5-methyl cytosine (Chatterjee and Walker 2017). Cytosine and 5-methyl cytosine are deaminated with much greater frequency than the other bases, and single stranded DNA, even in transient structures such as replication forks or transcription bubbles, is more prone to base deamination (Lindahl 1979, Lindahl 1993). Most of the resulting base mismatches are repaired by MMR (section **I.1.6.3**) before DNA replication. If left unrepaired, the conversion of cytosine to uracil or 5-methyl cytosine to thymine results in a G:C to A:T point mutation following DNA replication. These deamination-initiated base transitions play an important role in SHM, though in this context they are regulated by AID instead of occurring spontaneously (Chandra, Bortnick et al. 2015).

Spontaneous hydrolysis can also affect the glycosidic bond, cleaving off the nucleobase from the sugar backbone and leaving an abasic site, also known as an apurinic/apyrimidinic (AP) site (Chatterjee and Walker 2017). It has been estimated that between 2,000 and 10,000 abasic sites are created daily in a human cell, increasing in frequency with higher temperatures and pH extremes (Lindahl 1993). AP sites are readily repaired by base excision repair (BER, see section **I.1.6.1**) and less efficiently by nucleotide excision repair (NER, see section **I.1.6.2**) (Boiteux and Guillet 2004, Chatterjee and Walker 2017). If left unrepaired, AP sites present a problem for the replication machinery but can be bypassed by TLS polymerases, resulting in mutation (Boiteux and Guillet 2004). Furthermore, AP sites are unstable and can convert into SSBs (Chatterjee and Walker 2017).

I.1.4.4 Reactive oxygen species (ROS)

The final major source of endogenous genotoxic stress is reactive oxygen species (ROS). These include the superoxide anion (O_2^-), which is produced by mitochondrial aerobic metabolism; hydrogen peroxide (H_2O_2), which is generated by superoxide dismutases from O_2^-; and the highly reactive hydroxyl radicals ($^{\cdot}OH$), which are typically produced from H_2O_2 by ferrous or cuprous ions, although they can also be generated by the action of the exogenous genotoxic stressor ionizing radiation (IR, see section **I.1.5.1**) on water molecules (Cadet and Wagner 2013, Schieber and Chandel 2014). While aerobic metabolism is the main source of ROS in the cell, they can also be generated by catabolic oxidases, anabolic processes, and phagocytosis (Henle and Linn 1997). Despite playing a role in cellular signalling at low doses (Schieber and Chandel 2014), when produced in excess, ROS, and in particular $^{\cdot}OH$, readily react with any nearby macromolecules, including DNA. About 100 distinct lesions in the DNA are induced by oxidation of nucleobases or the sugar backbone, including nucleobase hydroperoxides, diastereomeric nucleosides, DNA intra- or inter-strand crosslinks, DNA-protein crosslinks, and SSBs (Cadet and Wagner 2013, Chatterjee and Walker 2017).

I.1.5 Exogenous sources of DNA damage

Exogenous DNA damaging agents originate outside of the cell or organism, and can be broadly divided into radiological, including ionizing radiation and ultraviolet radiation (UVR), and

chemical, including alkylating agents, aromatic amines, polycyclic aromatic hydrocarbons, and biological toxins. Beyond these two broad categories, genotoxic stress can also result from physical or chemical conditions in the environment, including high or low temperature and low oxygen availability (hypoxia).

I.1.5.1 Ionizing radiation (IR)

Ionizing radiation is composed of particles energetic enough to convert atoms in their path into ions. Examples include alpha particles (helium nuclei), beta particles (electrons or positrons), neutrons, and highly energetic photons (X-rays and gamma rays) (Donya, Radford et al. 2014). IR can originate from natural sources like radionuclides in the rocks or soil, radon gas, and cosmic rays from outer space, but also anthropogenic sources like radioactive waste, nuclear tests and accidents, and medical devices in the course of treatment (Desouky, Ding et al. 2015, Chatterjee and Walker 2017). At sea level, approximately 300 million IR particles (mostly gamma and X-rays) have been estimated to pass through a person per hour, about half of which come from terrestrial sources and the other half from space (Lieber 2010). Each type of IR may be further classified with respect to how it affects biological macromolecules (directly or indirectly), or by ionization density (or linear energy transfer, LET) (Chatterjee and Walker 2017), which is a measure of the energy lost by the ionizing particles as they traverse a medium (International Commission on Radiation and Measurements 2011). Alpha particles are considered high LET (Chatterjee and Walker 2017), and are thus more easily stopped by shielding than the more penetrative low LET particles like beta particles or photons. IR causes DNA damage either directly, striking the DNA molecule itself and disrupting its structure, or indirectly, by hitting the abundant water molecules or other organic molecules in the cell to generate hydroxyl radicals (see section **I.1.4.4**) or alkoxy radicals (·OR), both highly reactive species which can attack nucleobases or the sugar backbone (Desouky, Ding et al. 2015). Additionally, IR-induced radicals may in turn generate other DNA-reactive species (Wardman 2009). In total, indirect damage from hydroxyl radicals accounts for ~65% of IR-induced DNA lesions, and the spectrum of IR-induced DNA damage is therefore similar to that caused by ROS (Chatterjee and Walker 2017). However, IR also efficiently creates SSBs with the unique chemical signature of a 3' phosphate or phosphoglycolate, instead of the more usual hydroxyl group (Chatterjee and Walker 2017). Since it often results in clustered or complex

DNA damage (multiple lesions within one turn of the double helix) (Mladenov and Iliakis 2011), IR is also the most efficient genotoxic stressor in the formation of DSBs, which readily form from two SSBs in close proximity to another (but on opposite strands), or even two nucleobase lesions on opposite strands, if their repair mechanism involves the generation of SSBs (Hutchinson 1985, Iliakis 1991). If improperly repaired, DSBs can lead to chromosomal aberrations such as translocations. If left unrepaired, they present an insurmountable obstacle both to DNA replication and to faithful segregation of genetic material (Mehta and Haber 2014), and can lead to a loss of hundreds of genes located distally to the DSB on the chromosome. DSBs are repaired either by an end joining (EJ) pathway or a homologous recombination (HR)-based pathway, depending on the cellular context (see section **I.2.3**).

I.1.5.2 Ultraviolet (UV) radiation

UV radiation represents another form of a radiological genotoxic stressor. It is composed of photons with insufficient energy to cause ionization of matter, but enough to create or break covalent bonds. The overwhelmingly dominant source of UV radiation is the Sun (Kiefer 2007), though anthropogenic sources such as UV lamps are also widespread, ranging from germicidal lamps to tanning beds. The UV part of the electromagnetic spectrum is properly defined as being between 100 nm and 400 nm in wavelength, and is further subdivided into four categories (Kiefer 2007). The most energetic UV radiation is called cosmic UV or Schumann UV, and has a wavelength range of 100-200 nm. Because it is efficiently absorbed by the gases in the atmosphere (Kiefer 2007), cosmic UV may be of concern in interplanetary travel, but has little effect on the Earth's biota. The next most energetic part of the UV spectrum is UV-C (200-280 nm), which is strongly absorbed by the ozone and diatomic oxygen in the atmosphere and rarely reaches the surface except in ozone-depleted areas (Rastogi, Richa et al. 2010). However, UV-C is widely used in laboratory contexts because it is maximally absorbed by DNA (Chatterjee and Walker 2017). The less energetic UV-B (280-315 nm) and UV-A (315-380 nm) do reach the surface to affect the organisms living there, but represent only 5.1% and 0.3% of total sunlight, with visible spectrum light (62.7%) and infrared (31.9%) accounting for the rest (Kiefer 2007, Chatterjee and Walker 2017).

Like IR, UV radiation can cause DNA damage by direct excitation of biomolecules, leading to photochemical alteration, or indirectly, by energy transfer from other molecules, known as "photosensitizers", that have been directly excited by UV photons (Kiefer 2007, Rastogi, Richa et al. 2010). The two main types of DNA lesion caused by UV-C radiation are cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4PPs), both of which result from covalent bonding between neighboring nucleotides, with two bonds in CPDs and one bond in 6-4PPs linking the pyrimidine rings (Rastogi, Richa et al. 2010, Chatterjee and Walker 2017). Less frequently, UV-C exposure can generate other photoproducts, including thymine glycols, pyrimidine hydrates, and dipurine adducts (Chatterjee and Walker 2017). In contrast to UV-C, UV-B is much less efficient in forming pyrimidine dimers (Rastogi, Richa et al. 2010), while UV-A damages DNA by photooxidation, creating oxidized nucleobase adducts, or by excitation of endogenous photosensitizers (Douki, Perdiz et al. 1999, Chatterjee and Walker 2017). All forms of UV are also capable of attacking the sugar backbone and causing SSBs (Rastogi, Richa et al. 2010, Chatterjee and Walker 2017). CPDs, 6-4PPs, and oxidative lesions are bulky lesions which distort the shape of the double helix and interfere with transcription and replication (Rastogi, Richa et al. 2010). They are predominantly repaired by NER (section **I.1.6.2**), though interstrand crosslink (ICL) repair and HR may also contribute to repair in the case of more complex lesions (Chatterjee and Walker 2017). If left unrepaired, they can be bypassed in replication by the TLS polymerase Pol η (Goodman and Woodgate 2013).

I.1.5.3 Alkylating agents

As mentioned above, exogenous genotoxic stressors also include chemical agents which can react with DNA. One such group of DNA damaging molecules are alkylating agents, which are chemicals that can efficiently react with the highly nucleophilic nitrogen atoms in the heterocyclic rings of nucleobases to add an alkyl group and create an adduct (Fu, Calvo et al. 2012, Chatterjee and Walker 2017). The N7 position in guanine and the N3 position in adenine are particularly vulnerable to alkylation, but nitrogen and oxygen atoms in other nucleobases, as well as the phosphates in the sugar backbone, can also be alkylated (Chatterjee and Walker 2017). Alkylation of the major sites on adenine (N1 and N3) can result in a replication block or an A→T transversion (Fu, Calvo et al. 2012). In guanine, alkylation of N7 often leads to depurination (formation of an

AP site), which is itself unstable and potentially mutagenic, while alkylation of O6 can stall replication or produce a G→A transition (Fu, Calvo et al. 2012). Alkylating agents are commonly found in the environment, in various dietary sources, in tobacco smoke, employed as industrial processing agents and chemotherapeutic agents, and are also generated by biomass burning (Chatterjee and Walker 2017). A few alkylating agents find common laboratory use as mutagens or carcinogens. These include methyl methanesulfonate (MMS), ethyl methanesulfonate (EMS), methyl nitrosourea (MNU), and ethyl nitrosourea (ENU) (Acevedo-Arozena, Wells et al. 2008, Chatterjee and Walker 2017). Direct damage reversal (section **I.1.6**), BER (section **I.1.6.1**), and ICL repair are believed to be the primary pathways responsible for the repair of alkylation adducts (Wyatt and Pittman 2006).

I.1.5.4 Aromatic amines

Another class of genotoxic chemicals which can effect alkylation are aromatic amines. Examples include 2-acetylaminofluorene and 2-aminofluorene, which were used as insecticides before their carcinogenicity was established (Kriek 1992). Principally found in cigarette smoke, fuel, coal, industrial dyes, pesticides, and as by-products of high temperature cooking (Chatterjee and Walker 2017), aromatic amines are metabolized by the cytochrome P450 enzymes (Hammons, Milton et al. 1997) and converted into alkylating agents which target the C8 position of guanine, resulting in base substitutions and frameshift mutations (Heflich and Neft 1994, Hammons, Milton et al. 1997). Lesions produced by aromatic amines are repaired by NER (see section **I.1.6.2**) (Mu, Kropachev et al. 2012).

I.1.5.5 Polycyclic aromatic hydrocarbons (PAHs)

Polycyclic aromatic hydrocarbons (PAHs) represent another major class of chemical genotoxic stressors, with benzo(a)pyrene being the best-known example (Chatterjee and Walker 2017, Ewa and Danuta 2017). PAHs are composed of two or more aromatic carbon rings, and are commonly found in tobacco smoke, automobile exhaust, charred food, and incompletely combusted organic matter or fossil fuels (Chatterjee and Walker 2017). Like aromatic amines, PAHs are metabolized by the cytochrome P450 system to create DNA-reactive compounds which can intercalate into the double helix and react with the N2 atom in guanine and N6 in adenine to form adducts

(Cosman, de los Santos et al. 1992, Ewa and Danuta 2017). Like other DNA adducts, PAH lesions are both mutagenic and represent an obstacle to DNA replication (Ewa and Danuta 2017), though they can be bypassed by the TLS polymerase Pol κ (Jha, Bian et al. 2016). PAH-generated adducts are primarily repaired by NER (section **I.1.6.2**) or BER (section **I.1.6.1**) (Braithwaite, Wu et al. 1998, Ewa and Danuta 2017).

I.1.5.6 Toxins

DNA-reactive compounds can also be found among toxins of biological origin. Toxic molecules are often produced by bacteria, fungi, and plants for both offensive and defensive deployment (Ames, Profet et al. 1990, Dolan, Matulka et al. 2010). Among hundreds of examples, (Ames, Profet et al. 1990) some DNA-reactive toxins include furocoumarins in citrus fruit and the carrot/celery family *Umbelliferae*, which can generate DNA adducts and crosslinks; safrole, which is found in the aromatic oils of nutmeg, cinnamon, and camphor and can electrophilically attack DNA (Ewa and Danuta 2017); and aflatoxins, which are synthesized by two fungal species in the genus *Aspergillus*, and which attack the N7 atom in guanine, often resulting in depurination (Chatterjee and Walker 2017).

I.1.5.7 Physical stressors

The last type of exogenous genotoxic stressor to merit brief attention here is that of physical environmental causes. One such factor is temperature; Transient exposure to very low temperatures above freezing has been shown to lead to DNA fragmentation (Gregory and Milner 1994), and prolonged exposure to 25°C can result in ROS formation and DNA damage in the form of DSBs (Neutelings, Lambert et al. 2013). For a good overview of heat-stress induced DNA damage, see (Kantidze, Velichko et al. 2016). Low availability of oxygen (hypoxia) also promotes DNA damage by downregulating several DNA repair pathways, including cNHEJ, HR, and MMR (Luoto, Kumareswaran et al. 2013). Furthermore, a return to a fully oxygenated state may trigger a burst of ROS, resulting in concomitant DNA damage (Luoto, Kumareswaran et al. 2013).

I.1.6 DNA repair pathways

Given the diversity and ubiquity of both endogenous and exogenous genotoxic stressors, it is unsurprising that living organisms have evolved ways to cope with this unceasing assault on their genomes. In the following sections, a brief overview of major hitherto characterized DNA repair pathways will be presented. However, it is worth pointing out that DNA repair is but one of several mechanisms available to a cell to deal with DNA damage (Chatterjee and Walker 2017). First, cells may employ DNA damage checkpoints, which are signaling axes that interface with cell cycle machinery to temporarily halt cell cycle progression until the DNA damage is resolved (Houtgraaf, Versmissen et al. 2006, Shaltiel, Krenning et al. 2015). Second, certain less severe forms of DNA damage like nucleobase adducts may be tolerated instead of immediately repaired (Friedberg 2005). Mechanisms of DNA damage tolerance include the previously described translesion synthesis (TLS); post-replication repair, in which the replicative complex will "skip" over the (single strand) damaged lesion and assemble ~1000 nucleotides later to resume replication while the remaining gap is filled in by recombination between the daughter DNA molecules after the replication fork has passed; and replication fork regression, in which one of the two new DNA strands can use the other new DNA strand as a template instead of the damaged parental strand, forming a temporary four-stranded 'chicken foot' structure until the daughter strands can re-engage with the parental strands (Friedberg 2005). Finally, if the DNA damage can be neither tolerated nor repaired, the cell can opt to self-destruct by a mechanism of programmed cell death (PCD). The best known mechanism of PCD is apoptosis in the phylum Metazoa, but PCD has also been described in various unicellular organisms where it amounts to organismal suicide, a phenomenon which has been variously intepreted in terms of kin selection or group selection (Kaczanowski, Sajid et al. 2011, Nedelcu, Driscoll et al. 2011).

Our understanding of DNA repair pathways is continually expanding, as is their classification. Six main repair pathways, or groups of pathways, will be discussed in the following sections, including base excision repair, nucleotide excision repair, mismatch repair, single strand break repair, homologous recombination, and non-homologous end joining. What is common to all these types of repair is that the damaged lesion is removed, usually together with a surrounding segment on the DNA, and replaced with *de novo* synthesis. This contrasts with several ancient

repair pathways which are specialized to directly reverse the damage to nucleobases without removal of the lesion (Chatterjee and Walker 2017), which deserve brief mention. Examples include photoreactivation, a pathway conserved from bacteria to vertebrates (Lucas-Lledo and Lynch 2009), in which a photolyase restores original pyrimidine nucleobases from CPDs and 6-4PPs caused by UV radiation (Weber 2005); and the reversal of O-alkylation in nucleobases by O^6-alkylguanine alkyltransferase (AGT), and of N-alkylation by AlkB-related α-ketoglutarate dependent dioxygenases (Mishina, Duguid et al. 2006).

I.1.6.1 Base excision repair (BER)

Base excision repair is a pathway specialized to repair DNA lesions which do not greatly distort the shape of the DNA helix, such as nucleobase oxidation, deamination, or alkylation, as well as single abasic sites (Dianov and Hubscher 2013, Chatterjee and Walker 2017). BER begins with chromatin remodeling around the damaged lesion to allow the repair machinery to access the DNA (Odell, Wallace et al. 2013), which is a theme conserved across most DNA repair pathways. The damaged site is subsequently recognized by one of several DNA glycosylases which excise the damaged nucleobase from the double helix by severing its glycosidic bond to deoxyribose (Krokan and Bjoras 2013). The pathway forks into two possible directions at this step, depending on the nature of the glycosylase enzyme which performed the nucleobase excision. Monofunctional DNA glycosylases channel the lesion into the short-patch repair sub-pathway, while bifunctional DNA glycosylases, which in addition to the glycosylase activity, also possess a β-lyase activity, guide the damaged site into the long-patch repair sub-pathway (Dianov and Hubscher 2013). In short-patch repair, an AP endonuclease (APE1) cleaves the DNA backbone and creates an SSB, leaving a deoxyribose phosphate at the 5' end. DNA Pol β then removes this phosphate with its lyase activity, and simultaneously adds the missing nucleotide. In the final step, the nick in the sugar backbone is repaired by DNA ligase IIIα (Lig IIIα), in complex with X-ray repair cross complementing 1 (XRCC1) (Dianov and Hubscher 2013). The long-patch pathway is initiated either by bifunctional DNA glycosylases or in cases where SSB ends are too chemically complex to be immediately ligated (such as those created by IR). Here, after the initial nucleotide is added by DNA Pol β, replicative polymerases Pol δ and Pol ε continue synthesizing DNA, displacing one strand and creating a "flap" of 2-12 nucleotides. This flap is removed by flap

endonuclease I, and the resulting nick in the sugar backbone is sealed by DNA ligase I (Dianov and Hubscher 2013, Chatterjee and Walker 2017).

I.1.6.2 Nucleotide excision repair (NER)

Nucleotide excision repair is a pathway optimized to remove bulky single-strand lesions which distort the shape of the double helix, including the common UV photoproducts CPDs and 6-4PPs (section **I.1.5.2**) (Chatterjee and Walker 2017). Like BER and other repair pathways, NER requires chromatin remodeling to permit the repair proteins access to the damaged site (Scharer 2013). Two distinct subpathways of NER exist, distinguished on the basis of how the DNA damage is detected: global genome NER (GG-NER), and transcription-coupled NER (TC-NER). In GG-NER, a complex composed of the proteins Xeroderma pigmentosum complementation group C (XPC), UV excision repair protein Radiation sensitive 23 homolog B (RAD23B), and Centrin 2 (CENT2) scans the genome for the presence of single-stranded DNA (ssDNA) which arises from strand displacement caused by the damaged lesion (Chatterjee and Walker 2017). Upon binding to the damaged site, XPC recruits the ten-subunit transcription initiation factor II H (TFIIH) complex, which plays a role both in transcription of protein coding genes and in NER (Compe and Egly 2012). The final phase of GG-NER consists of three coordinated steps: the excision of a segment of the damaged DNA strand around the lesion, carried out by the structure-specific endonucleases XPF-ERCC1 and XPG (Fagbemi, Orelli et al. 2011); gap filling performed by DNA Pol δ, Pol ε, or Pol κ (Ogi, Limsirichaikul et al. 2010); and terminal ligation executed by either DNA ligase I or the DNA Ligase III-XRCC1 complex (Chatterjee and Walker 2017). By contrast, TC-NER initiates when RNA polymerase II encounters a bulky lesion and is blocked from proceeding. This stalling recruits CSA (Cockayne syndrome protein A) and CSB (Cockayne syndrome protein B), which backtrack the RNA Pol II holoenzyme to expose the lesion. CSA and CSB then assemble a set of core NER proteins (excluding XPC, which is specific to GG-NER) and several proteins specific to TC-NER (Fousteri, Vermeulen et al. 2006, Schwertman, Lagarou et al. 2012), after which TFIIH is recruited and the sequence of downstream events converges with GG-NER (Marteijn, Lans et al. 2014, Chatterjee and Walker 2017).

I.1.6.3 Mismatch repair (MMR)

Mismatch repair is a pathway that corrects errors arising in DNA replication, including both the removal of mismatched nucleotide which have escaped the proofreading activity of major replicative polymerases as well as the insertion-deletion loops (IDLs) which occur as a consequence of replicative strand slippage in repetitive sequence regions (Jiricny 2006, Kunkel 2011, Chatterjee and Walker 2017). As in other DNA repair pathways, initiation of repair is preceded by localized chromatin remodeling (Li 2014). The replication error is first recognized by a MutS heterodimer, either MutSα (composed of MSH2 and MSH6) in the case of base mismatches and short IDLs, or MutSβ (composed of MSH2 and MSH3) in the case of longer IDLs (Sachadyn 2010). MutL heterodimers (in humans, predominantly the MutLα dimer composed of MLH1 and PMS1) then bind to the MutS heterodimers, together forming a dynamic assembly anchored to the site of the mismatch (Qiu, Sakato et al. 2015, Chatterjee and Walker 2017). With its endonuclease activity MutLα generates a nick in the sugar backbone of the mismatched strand which is then further digested by the exonuclease EXO1, creating a gap (Jiricny 2013, Chatterjee and Walker 2017). In the final steps of repair, the gap is filled by DNA polymerase δ and accessory replication factors, followed by the terminal ligation executed by DNA ligase I (Jiricny 2013, Chatterjee and Walker 2017).

I.1.6.4 Single strand break repair (SSBR)

The three repair pathways discussed above deal with damage to nucleobases or errors in their incorporation into the polymer chain. The three pathways discussed in the following sections and in Part II are the primary repair mechanisms to deal with DNA damage which breaks the phosphodiester bonds in the sugar backbone. The less deleterious form of such DNA damage are single-strand breaks, which are discontinuities in the sugar backbone of only one of the strands of the double helix. SSBs may be accompanied by the loss of one or more nucleotides, and chemically altered (or blocked) 5’ or 3’ break termini (Caldecott 2008). The most common sources of SSBs include direct damage to deoxyribose by ROS (section **I.1.4.4**) or IR (section **I.1.5.1**), endonucleolytic cleavage that occurs as part of another repair process such as BER (section **I.1.6.1**) or NER (section **I.1.6.2**), and pathologic activity of class I topoisomerases (Caldecott 2014).

Left unrepaired, SSBs can lead to replication fork collapse, stall the progression of transcription, or even trigger necrosis or apoptosis by overactivation of the SSB sensor poly ADP-ribose polymerase-1 (PARP1) and depletion of NAD+ and ATP in the cell (Heeres and Hergenrother 2007, Caldecott 2014). Most SSBs are first detected by PARP1, which then covalently attaches branched chains of poly (ADP-ribose) to itself and other target proteins, including the neighboring chromatin, clearing the way for subsequent repair (Caldecott 2007, Caldecott 2014). The exception are SSBs caused by APE1 activity during BER, which in most cases do not require PARP1 activity. The next step in the repair process is the recruitment of processing enzymes, including APE1, polynucleotide kinase 3'-phosphatase (PNKP), aprataxin (APTX), aprataxin-like factor (APLF), which process the termini of general SSBs, and tyrosyl DNA phosphodiesterase 1 (TDP1), which is specialized for topoisomerase I-induced SSBs (Zhou, Lee et al. 2005, El-Khamisy, Hartsuiker et al. 2007, Caldecott 2014). An important role here is played by the molecular scaffold protein XRCC1, which directly interacts with PNKP, APTX, and APLF, and indirectly with TDP1, and also recruits DNA polymerase β and DNA ligase III, enzymes instrumental in the remaining two steps of SSB repair – gap filling and terminal ligation (Caldecott 2014). In these last steps, SSB repair converges with BER, with two distinct sub-pathways operating depending on the size of the gap in the damaged DNA strand. In the case of single nucleotide gap filling (the short-patch sub-pathway), the nucleotide is incorporated by Pol β and the nick is ligated by Lig III. As in BER, gap filling may be extended to between 2 and 12 nucleotides (the long-patch sub-pathway), in which the DNA synthesis is performed by Pol δ and Pol ε, and the terminal ligation executed by DNA ligase I (Caldecott 2014).

I.1.6.5 Double strand break repair

Because of the potential to affect chromosome segregation and effect large scale gene loss (Mehta and Haber 2014), DNA double-strand breaks represent the most severe form of DNA damage. Since at least 10 DSBs per day per cell are estimated to occur in human cells (Lieber 2010), with some estimates as high as 50 DSBs (Vilenchik and Knudson 2003), the need for DSB repair is acute. The diverse pathways which repair DSBs can be classified into two broad categories, depending on whether the repair mechanism requires the presence of sequences with extensive homology to the damaged region or not. Long homologous stretches are required

for HR-mediated DSB repair, which includes the double Holliday junction (dHJ) pathway, synthesis-dependent strand annealing (SDSA), and single-strand annealing (SSA). Much shorter or no sequence homology is required in EJ pathways, including canonical NHEJ and "alternative" end joining (Alt-EJ) or microhomology-mediated end joining (MMEJ). In the following sections, HR-based pathways are briefly discussed, and a more extensive discussion of cNHEJ as well as a concise overview of Alt-EJ are provided in Part II of this chapter.

I.1.6.5.1 Homologous recombination (HR)-based repair

All HR-based pathways start with the detection of the DSB by the MRN (MRE11-RAD50-NBS1) complex. Like other DNA repair pathways, HR-based DSB repair involves chromatin modification and signaling events known collectively as the DNA damage response (DDR). Major events in DDR include the MRN-mediated recruitment and activation of ataxia-telangiectasia mutated (ATM) kinase, phosphorylation of histone 2AX (H2AX), poly-ADP ribosylation (PARylation) of chromatin, and recruitment of mediator of DNA damage checkpoint 1 (MDC1), p53-binding protein 1 (53BP1), and breast cancer 1 (BRCA1), which collectively act to coordinate the proper cellular response to damage (Ciccia and Elledge 2010, Wei and Yu 2016, Chatterjee and Walker 2017).

I.1.6.5.1a Single-strand annealing (SSA)

The simplest of the HR-based pathways, SSA contrasts with the dHJ pathway and SDSA in that it occurs between direct repeats on the same chromosome by intra-molecular recombination. In this respect it is similar to Alt-EJ/MMEJ, which is also known as micro-SSA (Bhargava, Onyango et al. 2016), but it requires more extensive homology between the repeats, with between 20-30 nucleotides being the minimum extent of required homology (Mehta and Haber 2014, Chang, Pannunzio et al. 2017). Like other HR pathways, SSA begins with the resection of opposite strands on each side of the break, which is a two step-process. In the first, or "end clipping" phase, the MRN complex binds the DSB together with CtIP. This activates the endonuclease activity of MRE11, removing a small number of nucleotides (~20 in mammals, and up to 300 in yeast) (Ceccaldi, Rondinelli et al. 2016). At this stage, the DSB can be channeled into the Alt-EJ pathway (section **I.2.4**), or a second, more extensive resection phase can occur. This involves several helicases and exonucleases, including DNA2, BLM, WRN, CtIP, and EXO1, and commits the DSB

to HR-based repair (Ceccaldi, Rondinelli et al. 2016). After the second resection phase, the complementary strands of the identical or nearly identical repeats on either side of the DSB can anneal in a process mediated by RAD52, leaving unpaired ssDNA flaps of non-homologous sequences between the repeats (Mehta and Haber 2014, Morrical 2015). These flaps are removed by the XPF-ERCC1 endonuclease in complex with MutSβ (MSH2-MSH3), the remaining gaps presumably filled by one or more DNA polymerases (though it is not yet clear which polymerases act in SSA) (McVey, Khodaverdian et al. 2016), and finally sealed by DNA ligase I (Mehta and Haber 2014, Ceccaldi, Rondinelli et al. 2016). Because the sequences between the annealing repeats are lost, SSA is an intrinsically mutagenic repair pathway, although this appears to be an acceptable tradeoff for the restoration of chromosomal integrity.

I.1.6.5.1b Synthesis-dependent strand annealing (SDSA)

In contrast to SSA, SDSA and the dHJ pathway usually require a homologous sequence on another DNA molecule, whether in the form of a sister chromatid, an allelic sequence on the homologous chromosome, or some other genomic extrachromosomal source, such as a plasmid. The extensive resection around the DSB takes place as in SSA, but the 3' ssDNA overhang created by the processing is then mobilized to perform a "strand-invasion" of the homologous repair template. The ssDNA is first bound by replication protein A (RPA), possibly in an effort to remove any secondary structures (*ie,* intra-strand hydrogen bonding) which may have formed, but is quickly replaced by RAD51 recombinase (Mehta and Haber 2014). The RAD51-ssDNA nucleoprotein filament then searches local cellular space for the presence of homologous double stranded DNA (dsDNA). Upon locating a homologous sequence, RAD51 mediates the base pairing between the ssDNA overhang from the damaged strand and the complementary strand of the homologous molecule (Mehta and Haber 2014). This displaces the other (non-complementary, or identical) strand of the template molecule, resulting in a so-called "D-loop" or "joint molecule" structure (Morrical 2015). DNA synthesis follows, with DNA Pol δ and/or Pol ε extending the 3' end of the invading strand (Mehta and Haber 2014).

In SDSA, this DNA synthesis is not semiconservative, because the newly synthesized strand eventually dissociates from the template, collapsing the D-loop, and pairs with the 3' resected

end on the other side of the DSB in the original DNA duplex (Mehta and Haber 2014, Morrical 2015). The mechanistic details of the dissociation from the template strand are not known, but in yeast it appears to require the helicases Mph1 and Srs2 (Mehta and Haber 2014). After the newly synthesized DNA has base-paired with the original molecule, the remaining gaps are filled by replicative polymerases and the ends ligated, completing the process (Morrical 2015). Inter-homolog repair by SDSA results in a non-crossover (NCO), and is therefore not favored in meiosis (Zickler and Kleckner 2015). However, SDSA appears to be the preferred mitotic form of HR repair in higher eukaryotes, presumably because of a lower probability of gene conversion and loss of heterozygosity when compared to dHJ (Morrical 2015).

I.1.6.5.1c Double Holliday junction (dHJ) pathway

In the dHJ pathway, the sequence of events initially proceeds in the same manner as in SDSA, including DSB resection, strand invasion and D-loop formation, and the initiation of DNA synthesis. The dHJ pathway diverges from SDSA in that the newly synthesized strand does not dissociate from the template to bind the ssDNA on the other side of the DSB in the original molecule. Instead, the displaced strand in the D-loop binds to the ssDNA on the other DSB end in a process called “second-end capture”, thus generating a second heteroduplex and providing a template along which the non-invading 3’ end of the DSB can be extended (Morrical 2015). After ligation of the extending strands to the 5’ ends of the DSB, the resulting structure of two entangled heteroduplexes is known as a double Holliday junction, named after Robin Holliday who first described it in 1964 (Holliday 2007). Double Holliday junctions may also be formed if both ends of the DSB are mobilized to invade the homologous template at the same time, obviating the need for second-capture (Mehta and Haber 2014). Further processing of the dHJ can result in three possible outcomes, which are controlled by enzymes known as resolvases and dissolvases: 1) it can be resolved to form a crossover (CO), *ie* a reciprocal translocation between homologous chromosomes; 2) it can be resolved to form an NCO, in which there is no translocation but small segments of the two original heteroduplexes switch places; and 3) it can be dissolved, with the template duplex remaining unchanged and the damaged duplex restored with just a small segment of the template duplex surrounding the location of the former DSB (Mehta and Haber 2014). Because of a requirement for both increased genetic diversity and

accurate segregation of homologous chromosomes, inter-homolog dHJs are heavily favored to resolve into COs during meiosis (Zickler and Kleckner 2015).

I.1.6.5.1d Break-induced replication (BIR)

In addition to the three major HR mechanisms, a final form of HR-repair, called break-induced replication (BIR), deserves brief mention. BIR occurs when a damaged DNA double helix presents only one end for repair, or if only one end shows extensive homology to the repair template. The strand invasion happens as in dHJ and SDSA, but instead of a limited DNA synthesis, a unidirectional replication fork is created which is capable of replicating hundreds of thousands of kilobases, down to the telomeric sequences (Mehta and Haber 2014). If the repair template is the homologous chromosome, BIR results in extensive loss of heterozygosity, which in combination with being much more mutagenic than normal replication because of reduced proofreading ability of Pol δ, has been suggested to contribute to genetic instability in cancer (Deem, Keszthelyi et al. 2011).

The other major group of DSB repair mechanisms, the pathways that fall under the umbrella of end joining, are discussed in the **Part II** below.

Literature review, part II: Non-homologous end joining

I.2.1 End joining pathways

End joining is a common descriptor of at least two distinct pathways that do not require a homologous template to execute DSB repair, and instead just process the DNA ends and ligate the chromosome back together. The first discovered of these pathways is non-homologous end joining (NHEJ), now more often known as canonical non-homologous end joining (cNHEJ) to distinguish it from the "non-canonical" counterpart termed alternative end joining (Alt-EJ). **Part II** of this chapter focuses on cNHEJ, discussing its mechanism of action, introducing the known cast of protein characters that play a role in the best described (mammalian) cNHEJ system,

comparing this system to those found in other studied taxa, and discussing its role in human health and disease. Brief attention is also given to Alt-EJ, and how it contrasts with cNHEJ.

I.2.2 The mechanism of cNHEJ

Although the precise sequence of events is not yet fully understood, the mechanism of cNHEJ can be divided into three conceptual phases (excellently reviewed in (Davis and Chen 2013), (Wang and Lees-Miller 2013), (Radhakrishnan, Jette et al. 2014), and (Chang, Pannunzio et al. 2017); see also **Figure 1.1**). During the first, "detection and tethering" phase, the DNA ends in the DSB are recognized by sensor molecules, stabilized, and committed into the cNHEJ repair pathway, *ie*, protected from processing by rival DSB repair pathways or non-specific nucleases (Davis and Chen 2013, Wang and Lees-Miller 2013). In the second, "processing" phase, the DNA ends are acted on by various exo- and endonucleases, helicases, DNA polymerases, kinases, and phosphatases. This action converts them into a form compatible for ligation, which occurs in the third, or "terminal ligation" phase (Wang and Lees-Miller 2013, Radhakrishnan, Jette et al. 2014).

I.2.2.1a Detection and tethering – the centrality of the Ku ring

At the core of all three phases of cNHEJ is Ku, a highly conserved protein which plays a critical role in the process in all three domains of life (Gu and Lieber 2008). Ku was first identified in the 1980s as the antigen targeted by autoantibodies in a population of Japanese patients with scleroderma-polymyositis overlap syndrome, an autoimmune disorder (Mimori, Akizuki et al. 1981). The antigen was named Ku after the first two letters of the name of the patient who provided the prototype serum (Mimori, Akizuki et al. 1981), and found to contain two polypeptides, Ku70 and Ku80 (70 kDa and 80 kDa in molecular weight, respectively) which precipitated together with DNA (Mimori, Hardin et al. 1986). It was soon demonstrated that Ku required free DNA ends to bind to dsDNA (Blier, Griffith et al. 1993), and that it forms a DNA-binding component of DNA-dependent protein kinase (DNA-PK), a nuclear kinase which is now understood to play important roles in cNHEJ (see section **I.2.2.1b**), forming a bridge between DNA and the catalytic subunit (DNA-PKcs) of the kinase (Gottlieb and Jackson 1993), which was in turn found to phosphorylate Ku (Leesmiller, Chen et al. 1990). Shortly after, Ku and DNA-PKcs were identified as the molecular identities of, respectively, *XRCC5* and *XRCC7*, two X-ray

sensitivity genes in human cells (Taccioli, Gottlieb et al. 1994, Kirchgessner, Patil et al. 1995, Leesmiller, Godbout et al. 1995, Weaver 1995). They were also demonstrated to be required for V(D)J recombination (see section **I.2.6**) and DSB repair by non-homologous end joining (Rathmell and Chu 1994, Smider, Rathmell et al. 1994, Taccioli, Gottlieb et al. 1994, Blunt, Finnie et al. 1995, Leesmiller, Godbout et al. 1995). As a phenomenon, NHEJ was first described in the early 1980s as "indiscriminate recombination" between DNA of diverse (eukaryotic and prokaryotic) origins co-transfected with simian virus 40 (SV40) into mammalian cells (Botchan, Stringer et al. 1980, Winocour and Keshet 1980). Shortly after, it was demonstrated that such "nonhomologous recombination" can occur with high efficiency in mammalian somatic cells by end-to-end joining of unrelated DNA segments (Wilson, Berget et al. 1982), and the specific term "non-homologous end joining" started being used to describe the process in the mid-1990s (Bowater and Doherty 2006). Further work on the Ku heterodimer provided the first crystal structure, revealing that Ku70 and Ku80 form a ring like structure which can encircle a DNA duplex, interacting with the sugar backbone while making no direct contact with the nucleobases (Walker, Corpina et al. 2001). This model explained the observation of "indiscriminate", or non-sequence-specific nature of Ku-DNA binding (Davis and Chen 2013). While the primary sequence similarity of Ku70 and Ku80 is not high overall, it is very high in several blocks of homology, indicating that they arose from a common ancestor (Downs and Jackson 2004). The two Ku proteins are structurally similar, with the major difference being the C-terminal helical domain unique to Ku80 and the C-terminal SAP domain unique to Ku70 (Gell and Jackson 1999, Rivera-Calzada, Spagnolo et al. 2007). The SAP domain of Ku70 is primarily responsible for DNA binding, forming a positively-charged surface which interacts with the phosphates in the sugar backbone (Aravind and Koonin 2000, Zhang, Zhu et al. 2001). The C-terminus of Ku80 contains six alpha helices with a topology similar to structural scaffold proteins (Zhang, Hu et al. 2004), and is primarily responsible for protein-protein interactions, acting in a manner likened to a "tool belt" which recruits other cNHEJ factors critical for DSB repair (Lieber 2008). The Ku ring is the initial responder to DSBs, binding broken DNA ends within seconds and reaching maximal intensity within 3 minutes of a laser-induced DSB event (Mari, Florea et al. 2006). It is thus essential in the "detection and tethering" phase of cNHEJ. In addition to its best-studied role in cNHEJ, the Ku proteins also play important roles in

other cellular processes, including inhibition of pro-apoptotic proteins, control of transcriptional machinery, telomere capping, localization, and length regulation, and retrotransposon/retrovirus mobility (Downs and Jackson 2004, Bailey and Murnane 2006, O'Sullivan and Karlseder 2010).

I.2.2.1b Detection and tethering – DNA-PKcs

The Ku-interacting DNA-PKcs belongs to phosphoinositide 3-kinase related kinases (PIKKs), an important group of enzymes with critical roles in cellular stress responses, which besides DNA-PKcs includes ATM, ATR, mTOR, and hSMG-1 (Abraham 2004, Bakkenist and Kastan 2004). The kinase domain of PIKK family members is located in the C-terminal region, and is flanked by the PIKK-regulatory domain (PRD) and the FRAP-ATM-TRAPP-C-terminal (FATC) domain which regulate kinase activity through conformational changes (Bakkenist and Kastan 2004, Lempiainen and Halazonetis 2009). Every member of the PIKK family possesses N-terminal helical HEAT repeats, which likely play a role as a protein-protein interaction site (Perry and Kleckner 2003). While the C-terminus of Ku80 has been shown to be important for the recruitment and activation of DNA-PKcs (Gell and Jackson 1999, Singleton, Torres-Arzayus et al. 1999, Falck, Coates et al. 2005), DNA-PKcs can be localized to laser-induced DSBs independently of Ku80 (Weterings, Verkaik et al. 2009), suggesting that the mode of DNA-PKcs recruitment may depend on the lesion type. This interpretation is strengthened by the existence of an N-terminal leucine-rich region in DNA-PKcs which may be responsible for its Ku-independent DNA binding (Gupta and Meek 2005). Crystallographic and cryo-electron microscopic studies have revealed that DNA-PKcs possesses an open-ring-like (or pincer-like) structure, with the HEAT repeats forming the “fingers” of the pincer, and the putative DNA-binding region located on the inside surface of the ring, through which a DNA duplex can easily pass (Williams, Lee et al. 2008, Sibanda, Chirgadze et al. 2010). DNA-PKcs likely makes contact with the Ku ring through several surfaces, but the N-terminus is indispensable for both the interaction with Ku and the Ku-DNA complex-dependent kinase activation (Spagnolo, Rivera-Calzada et al. 2006, Davis, Lee et al. 2013). DNA-PKcs phosphorylates multiple cNHEJ proteins in vitro and in vivo, including Ku70 and Ku80, Artemis, XRCC4, and XLF, and is also capable of autophosphorylation (Meek, Dang et al. 2008). However, apart from DNA-PKcs itself, phosphorylation of other cNHEJ factors does not play a functional role in the repair process (Meek, Dang et al. 2008). By contrast, DNA-PKcs autophosphorylation at several critical

sites (the so-called “ABCDE” sites) has been proposed to induce a conformational change which opens the pincer, allowing for the release of DNA-PKcs from DNA and the Ku ring, and the completion of repair (Dobbs, Tainer et al. 2010). Autophosphorylation also regulates DNA end processing (Cui, Yu et al. 2005), as well as DSB repair pathway choice, with phosphorylation of some sites inhibiting HR and promoting cNHEJ, and others promoting HR (Cui, Yu et al. 2005, Neal, Dang et al. 2011, Shibata, Conrad et al. 2011). Consistent with its role in regulating DNA end processing, DNA-PKcs is required only for the repair of chemically complex DSBs, such as those induced by a near infrared laser, and does not participate in the repair of simple DSBs such as those induced by the sparsely ionizing ultrasoft X-rays (Reynolds, Anderson et al. 2012). In addition to DNA-PKcs, ATM may also play a role in regulating the access of cNHEJ factors to the DSB and/or their activity, as it has been shown to phosphorylate DNA-PKcs, Artemis, XLF, APTX, and APLF (see below) (Radhakrishnan, Jette et al. 2014).

I.2.2.2 DNA end processing

As mentioned above, IR-induced DSBs are often chemically complex, *ie* associated with other types of base and sugar lesions, because of the oxidative damage caused by ionization of water and other molecules in the close proximity of the DNA chain (Cadet and Wagner 2013, Georgakilas, O'Neill et al. 2013). Before such complex DSBs can be repaired, extensive processing must take place to produce ligation compatibility at the DNA ends. In the context of cNHEJ, several important processing enzymes have been identified so far. The first cNHEJ processing factor is the Ku ring itself, which apart from the central role it plays in stabilizing DNA ends and recruiting processing enzymes also possesses an AP-lyase activity similar to that of Pol β in the short-patch BER and SSBR (sections **I.1.6.1** and **I.1.6.4**) (Reynolds, Anderson et al. 2012). Ku also recruits WRN, a helicase and exonuclease mutated in Werner’s Syndrome, a type of progeria, and stimulates its exonuclease activity (Orren, Machwe et al. 2001). DNA-PKcs recruits Artemis, a structure-specific endonuclease which is activated by DNA-PKcs mediated phosphorylation and plays an important role in the opening of RAG-induced hairpins in V(D)J recombination (see section **I.2.6**) (Ma, Pannicke et al. 2002, Ma, Schwarz et al. 2005). Artemis-deficient cells have increased radiosensitivity, which suggest that Artemis can act in IR-induced DSB repair in addition to V(D)J recombination (Moshous, Callebaut et al. 2001). In addition to the DNA-PKcs-activated

endonuclease activity, Artemis has also been reported to possess an intrinsic exonuclease activity, but it was subsequently demonstrated that this exonuclease activity is either a result of contamination in the assay or originates in an associated factor and not the Artemis polypeptide (Pawelczak and Turchi 2010). Another processing factor active in cNHEJ is PNKP. This enzyme possesses a 3' DNA phosphatase and 5' DNA kinase activity, allowing it to remove aberrant phosphates from 3' ends and add phosphates to the 5' ends which have lost them, restoring ligation compatibility (Weinfeld, Mani et al. 2011). PNKP is recruited to the DSB to participate in cNHEJ by CK2-phosphorylated XRCC4 (Koch, Agyei et al. 2004). The kinase CK2 can also phosphorylate XRCC1, which can recruit PNKP to participate in SSBR and Alt-EJ (see sections **I.1.6.4** and **I.2.4**) (Whitehouse, Taylor et al. 2001, Weinfeld, Mani et al. 2011). The binding of PNKP to phosphorylated XRCC1 and XRCC4 is mediated by its forkhead-associated (FHA) domain (Weinfeld, Mani et al. 2011). The same FHA-mediated interactions recruit aprataxin (APTX) to XRCC1 and XRCC4 (Clements, Breslin et al. 2004). This zinc-finger protein belongs to the histidine triad (HIT) family of nucleotide hydrolases and transferases, and acts to remove AMP from 5' DNA ends which is generated by abortive ligation attempts (Moreira, Barbot et al. 2001, Ahel, Rass et al. 2006). Although mutations in *APTX* impair SSBR (Moreira, Barbot et al. 2001), no defects in cNHEJ have so far been identified in *APTX* mutants (Radhakrishnan, Jette et al. 2014). Aprataxin and PNKP-like factor (APLF), also known as PNKP and APTX-like FHA protein (PALF), is another processing factor associated with cNHEJ (Iles, Rulten et al. 2007, Kanno, Kuzuoka et al. 2007). Like APTX and PNKP, APLF interacts with phosphorylated XRCC4 and XRCC1 by its FHA domain (Iles, Rulten et al. 2007, Kanno, Kuzuoka et al. 2007). APLF possesses both an endonuclease and 3' exonuclease activity (Li, Kanno et al. 2011), and also directly binds to Ku80 and promotes the retention of XRCC4-Lig4 and XLF (Grundy, Rulten et al. 2013, Shirodkar, Fenton et al. 2013) (see next section).

After the DNA ends have been cleaned up and associated lesions removed, several Pol X family polymerases act to fill in any resulting gaps, including (in reverse order of dependence on template complementarity) Pol β, Pol λ, Pol μ, and the template independent TdT, which acts only in the context of V(D)J recombination in lymphocytes (Ramsden and Asagoshi 2012).

I.2.2.3 Terminal ligation

Ligation represents the final step of cNHEJ, restoring the integrity of the double helix. In cNHEJ, it is mediated by the universally conserved DNA Ligase IV (Chang, Pannunzio et al. 2017). DNA Ligase IV exists in complex with XRCC4 (Sibanda, Critchlow et al. 2001, Wu, Frit et al. 2009), a scaffold protein which interacts with a structurally similar protein, XRCC4-like factor (XLF, also known as Cernunnos) (Ahnesorg, Smith et al. 2006, Buck, Malivert et al. 2006), to form long helical filaments which promote the critical pre-ligation steps of DNA end bridging, alignment, and protection (Hammel, Rey et al. 2011, Wu, Ochi et al. 2011, Mahaney, Hammel et al. 2013). XRCC4 stabilizes DNA Ligase IV and stimulates its activity (Grawunder, Wilm et al. 1997), while XLF greatly promotes the ligation of non-matching or noncohesive DNA ends in the absence of processing factors (Tsai, Kim et al. 2007). Recently, another paralog of XRCC4 and XLF, PAXX, has been identified in mammals (Craxton, Somers et al. 2015, Ochi, Blackford et al. 2015). The loss of PAXX results in radiosensitivity, consistent with a role in cNHEJ, and it appears to function in Ku-dependent ligation stimulation, much like XLF (Roy, de Melo et al. 2015, Tadi, Tellier-Lebegue et al. 2016). As in the other steps of cNHEJ repair, Ku also plays a central role in the terminal ligation step, primarily by organizing the recruitment of accessory factors. Ku70 interacts with and recruits XRCC4 to the DSB (Mari, Florea et al. 2006). The Ku ring can also directly recruit XLF to the DSB, although the presence of XRCC4 stabilizes XLF at the site of the DSB (Yano, Morotomi-Yano et al. 2008, Yano, Morotomi-Yano et al. 2011). PAXX can only be recruited to DSBs by interacting with Ku70 (Roy, de Melo et al. 2015, Tadi, Tellier-Lebegue et al. 2016). Finally, Ku also recruits APLF in its capacity as a structural ligation-promoting factor (Grundy, Rulten et al. 2013, Shirodkar, Fenton et al. 2013).

I.2.3 Regulation of DSB pathway choice

The regulation of cNHEJ and other DSB repair pathways is mediated by DDR kinase signaling, including DNA-PKcs, ATM, and ATR (Sirbu and Cortez 2013). This regulation has been described as occurring on three levels: 1) the modification of enzymes and structural proteins directly involved in DNA repair; 2) the regulation of chromatin accessibility prior to, during, and following repair; and 3) the alteration of the general cellular environment to create repair-conducive

conditions (reviewed in (Sirbu and Cortez 2013)). The first of these levels has been briefly touched on above (section **I.2.2.1b**) as a series of DNA-PKcs-mediated phosphorylation events in the context of cNHEJ (Meek, Dang et al. 2008), although the full story lies beyond the scope of this chapter. The second, chromatin level includes local chromatin relaxation and histone removal or exchange, which is necessary for efficient recruitment of DSB sensing proteins, as well as the creation of a platform for the recruitment of downstream DDR signaling factors (Sirbu and Cortez 2013). As mentioned in section **I.1.6.5.1**, the core histone modification for both purposes is the ATM-mediated phosphorylation of the conserved serine 139 on the histone variant H2AX, which produces γH2AX (Sirbu and Cortez 2013). This creates a chromatin platform for the recruitment of MDC1, which promotes more ATM activity and H2AX phosphorylation, but also recruits ubiquitin ligases such as RNF8 and RNF168 which ubiquitylate nucleosomes in proximity to the DSB (Sirbu and Cortez 2013). This ubiquitylation in turn serves to recruit (among others) the DSB repair proteins BRCA1 and 53BP1, which play an important role in the regulation of DSB repair pathway choice. Several excellent reviews have been published on the topic of DSB pathway choice regulation, and the reader is directed to them for a more detailed treatment (Chapman, Taylor et al. 2012, Grabarz, Barascu et al. 2012, Clouaire and Legube 2015, Ceccaldi, Rondinelli et al. 2016). In brief, the initial choice between DNA end resection (favoring HR, SSA, and Alt-EJ) and no DNA resection (favoring cNHEJ) is decided by the balance of activities of several factors. The key players favoring end resection are BRCA1, its ubiquitylation target CtIP, and the MRN complex. In contrast, 53BP1 promotes cNHEJ by opposing nucleolytic resection of DNA ends as well as acting as a synaptic bridge to promote ligation (Chapman, Taylor et al. 2012). Pre-existing chromatin structures and histone modification may play a role in pathway choice as well (Clouaire and Legube 2015). For example, HR has been proposed to be favored in actively transcribed regions because trimethylated H3K36, a histone mark associated with active transcription (Wagner and Carpenter 2012), recruits CtIP and therefore promotes end resection (Clouaire and Legube 2015). Similarly, acetylated H4K16, another histone mark associated with transcription (Zhang, Erler et al. 2017), antagonizes 53BP1 recruitment, which also promotes end resection (Tang, Cho et al. 2013, Clouaire and Legube 2015). In contrast, other histone marks, such as the highly abundant dimethylated H4K20 (present on more than 85% of all histones (Schotta,

Sengupta et al. 2008)) and the coding-region associated monomethylated H4K20 (Barski, Cuddapah et al. 2007), favor cNHEJ by recruiting 53BP1 and blocking end resection (Hartlerode, Guan et al. 2012, Clouaire and Legube 2015).

The third level of DSB repair regulation is represented by the wider effects of DDR on the cellular environment which promote successful repair (Sirbu and Cortez 2013). Foremost of these are cell cycle checkpoints, which block progression through the cell cycle by modulating the activities of cyclin-dependent kinases (CDKs), allowing time for repair (Shaltiel, Krenning et al. 2015). The cell cycle intersects with the repair pathway choice decision as well. As an example, CDK-dependent phosphorylation of CtIP and EXO1 greatly promotes end resection and therefore HR, ensuring that this DSB repair mode is favored in S phase and G2 phase, when an accurate repair template (*ie*, the sister chromatid) is available, in contrast to cNHEJ which is favored in the G1 phase (Ceccaldi, Rondinelli et al. 2016). Other alterations to the wider cellular environment which affect DNA repair include transcriptional upregulation, which increases the level of DNA repair proteins and histones; increased nucleotide synthesis, which provides the raw materials for the repair; and increased chromosome mobility, which promotes the search for a homologous template in pathways that require it (Sirbu and Cortez 2013).

I.2.4 Alternative end joining

Alternative end joining (a-EJ or Alt-EJ), also known as alternative non-homologous end joining (a-NHEJ or Alt-NHEJ), microhomology-mediated end joining (MMEJ), Pol θ-mediated end joining (TMEJ), and backup non-homologous end joining (bNHEJ) is the most recently discovered pathway (or a group of pathways) involved in DSB repair (Deriano and Roth 2013). This modality of DSB repair operates as a backup to both cNHEJ and HR, *ie*, is primarily active in conditions where these pathways are impaired (Iliakis, Murmann et al. 2015). While the nomenclature has not yet been stabilized, there appear to exist at least two distinct pathways of this backup repair, with MMEJ requiring pre-existing microhomologies for annealing and a-EJ either not requiring microhomologies at all or relying on polymerase activity to generate them (Decottignies 2013, Frit, Barboule et al. 2014). Because microhomologies are important in only a subset of these backup DSB repair pathways, the more general term a-EJ is used here, following prior practice

(Frit, Barboule et al. 2014). The a-EJ pathway can be divided into three steps parallel those of cNHEJ: end recognition/tethering, end processing, and ligation (Frit, Barboule et al. 2014). The only factor identified to play a role in DNA end recognition in a-EJ is PARP1, which also plays a role in BER (see section **I.1.6.1**) and SSBR (see section **I.1.6.4**) (Frit, Barboule et al. 2014, Sallmyr and Tomkinson 2018). PARP1 recruits the MRN complex and CtIP, which in addition to a possible role in end bridging, also initiate DNA end resection (Sallmyr and Tomkinson 2018). Besides MRN and CtIP, resection factors involved in the a-EJ pathway include the nucleases EXO1 and DNA2 (Sallmyr and Tomkinson 2018). The nucleases FEN1 and ERCC1-XPF, which otherwise play an important role in NER (see section **I.1.6.2**), have also been implicated (Frit, Barboule et al. 2014). After resection, which exposes stretches of ssDNA, the opposing ends are annealed using available microhomology, paralleling the annealing step of SSA (see section **I.1.6.5.1a**), although instead of RAD52, MMEJ likely employs PARP1 or the MRN complex in this step (Sallmyr and Tomkinson 2018). While cNHEJ joins ends with no homology or less than 5 nucleotides of homology, the MMEJ pathway of a-EJ requires 2-20 nucleotides of microhomology, and ends with more than 20 nucleotides of homology are joined by SSA (Chang, Pannunzio et al. 2017). End bridging and annealing is also promoted by the DNA polymerase Pol θ, whose critical role in a-EJ was first uncovered in *Drosophila* (Chan, Yu et al. 2010), and more recent studies have confirmed its involvement in mammals (Mateos-Gomez, Gong et al. 2015) and *C. elegans* (Koole, van Schendel et al. 2014). DNA Pol θ also plays a role in the gap-filling DNA synthesis that follows resection in MMEJ, but it is unclear if it plays a role in the other a-EJ pathway(s) (Sallmyr and Tomkinson 2018). Evidence from yeast and mammalian cells implicates the Pol X family polymerases Pol μ and Pol λ (and their yeast homolog Pol4) in the synthesis step of a-EJ (Frit, Barboule et al. 2014). Finally, PNKP may also have a role in the processing of DNA ends in a-EJ, as it is recruited to DNA ends containing PARP1 and the XRCC1-LIG3 ligation complex, which is responsible for the terminal ligation step of a-EJ (Audebert, Salles et al. 2006). While the involvement of XRCC1-LIG3 complex in mammalian MMEJ is well established, other a-EJ pathway(s) in mammalian cells may use LIG1 instead, because a-EJ can still occur in backgrounds lacking XRCC1, which is critical for the stability and activity of LIG3 (Sallmyr and Tomkinson 2018). In cNHEJ-competent backgrounds, a-EJ is relegated to a backup role, likely as a direct

consequence of the exceptional affinity of the Ku ring for free DNA ends, which initiates repair by cNHEJ (Frit, Barboule et al. 2014). In S and G2 phases of the cell cycle, CDK activity promotes the activation of BRCA1 and CtIP which favor more extensive end resection and initiation of HR, in contrast to the more limited resection required for a-EJ (Frit, Barboule et al. 2014). However, Pol θ has the ability to bind RAD51 and prevent the formation of the invasive nucleoprotein strand, thus antagonizing HR (Mateos-Gomez, Kent et al. 2017).

I.2.5 Evolution of cNHEJ

The known steps of the cNHEJ mechanism described above have to a large extent been elucidated in the mammalian system. However, cNHEJ is an ancient and virtually universally conserved mechanism of DSB repair (Gu and Lieber 2008), a fact that is perhaps less surprising in light of the estimates that IR levels from geologic sources and internal potassium isotopes were 5-fold higher during the period of life's emergence 3.5 to 4 billion years ago than they are today (Karam and Leslie 1999). In the following sections, a brief overview of what is known about cNHEJ is non-mammalian systems will be presented. The conservation of cNHEJ factors is shown in **Figure 1.2**.

I.2.5.1 cNHEJ in prokaryotes

Although it was initially presumed that cNHEJ is restricted to eukaryotes, distant homologs of Ku have been identified in several bacterial species (Aravind and Koonin 2001, Doherty, Jackson et al. 2001), and even bacteriophages (di Fagagna, Weller et al. 2003). Further studies revealed that bacterial Ku forms a homodimeric ring (Weller, Kysela et al. 2002). Many of the bacterial operons containing Ku sequences were also found to contain ATP-dependent ligases, some of which possessed nuclease and primase activities (Weller and Doherty 2001). It is now recognized that several ligase types may be active in bacterial cNHEJ. The primary bacterial ligase appears to be LigD, which possesses ligase, polymerase, and phosphoesterase activities, while LigC acts as a backup ligase (for a review of bacterial cNHEJ ligases and the process of bacterial cNHEJ in general, see (Shuman and Glickman 2007) and (Pitcher, Brissett et al. 2007)). Because bacteria generally possess multiple copies of the genome available to serve as HR repair templates (except spores which have only one genome copy), the role of NHEJ in non-spore contexts and non-sporulating bacteria is not fully understood (Pitcher, Brissett et al. 2007). However, at least one

report exists of cNHEJ being the dominant form of DSB repair during the stationary phase of growth and being required for DNA integrity during prolonged periods of desiccation in *Mycobacteria*, providing evidence for a physiological role in a non-spore context (Pitcher, Green et al. 2007).

I.2.5.2 cNHEJ in *Saccharomyces cerevisiae*

The model single-celled eukaryote *S. cerevisiae* possesses a simpler cNHEJ system in comparison to vertebrates, yet with many recognizable orthologs (for a review on cNHEJ in yeast, see (Daley, Palmbos et al. 2005) and (Emerson and Bertuch 2016)). Ku orthologs Yku70 and Yku80 are present in the yeast genome and are essential for cNHEJ (Barnes and Rio 1997), as is the Ligase IV ortholog Dnl4 (Wilson, Grawunder et al. 1997). Homologs of mammalian XRCC4 and XLF, called Lif1 and Lif2/Nej1, also participate in cNHEJ by recruiting Dnl4. Nej1 and Lif1-Dnl4 are able to independently recruit the processing polymerase Pol4 (Yang, Matsumoto et al. 2015), a DNA Pol β ortholog which is the only Pol X in yeast (Wilson and Lieber 1999), as well as Rad27, the ortholog of mammalian FEN-1 (Wu, Wilson et al. 1999, Yang, Matsumoto et al. 2015). No ortholog of DNA-PKcs exists in the budding yeast genome (Daley, Palmbos et al. 2005), and while an Artemis ortholog called Pso2/Snm1 is present, it seems to be active only in DNA hairpin processing and ICL repair, not general cNHEJ (Li and Moses 2003, Yu, Marshall et al. 2004). Similarly, the budding yeast ortholog of PNKP, Tpp1, does not possess a kinase domain and is dispensable for cNHEJ, but retains a role in BER (Daley and Wilson 2005). Another distinction of *S. cerevisiae* is the involvement of the Mre11-Rad50-Xrs2 (MRX) complex, the homolog of the vertebrate MRN which normally acts in HR-based repair, in NHEJ (Boulton and Jackson 1998). The role of MRX in the mechanism of NHEJ is not clear, but it has been postulated to assist in Ku binding and Dnl4 activity (Daley, Palmbos et al. 2005).

I.2.5.3 cNHEJ in *Drosophila melanogaster*

The fruit fly *Drosophila melanogaster* similarly has a reduced complement of NHEJ factors (for a recent review of DNA repair, including cNHEJ, in *D. melanogaster*, see (Sekelsky 2017)). The universally conserved Ku70 and Ku80 are present (Beall, Admon et al. 1994, Jacoby and Wensink 1996), but not DNA-PKcs, which has been lost in several insect orders, including the ultradiverse

Coleoptera (beetles) and partially in Diptera (lost in flies, retained in mosquitoes) (Sekelsky 2017). In terms of processing factors, Artemis has been lost in Schizophora, a section of the order Diptera to which *Drosophila* belongs, and like most insects *Drosophila* also lacks the polymerases Pol μ, Pol λ, and TdT (Sekelsky 2017). Despite this, DSB end processing occurs during cNHEJ in *Drosophila* (Sekelsky 2017). The terminal ligase LigIV is conserved in *Drosophila*, as is its role in cNHEJ (Gorski, Eeken et al. 2003). Finally, while the fly possesses an ortholog of XRCC4 (CG3448) and two orthologs of XLF (CG12728 and CG32756), no functional studies of these genes have yet been published (Sekelsky 2017).

I.2.5.4 cNHEJ in *Caenorhabditis elegans*

The present study was conducted in the nematode *C. elegans*, whose closest relative among the model organisms is *D. melanogaster* (see section **I.3.4c**). Unlike other animals or even eukaryotes as a whole, *C. elegans* conserves only the Ku dimer (CKU-70 and CKU-80) and the Ligase IV ortholog LIG-4 as components of its cNHEJ system (Clejan, Boerckel et al. 2006, Lemmens and Tijsterman 2011). Although an ortholog of Werner's helicase, WRN-1, is encoded by the *C. elegans* genome, it doesn't appear to possess a role in cNHEJ (Clejan, Boerckel et al. 2006). Furthermore, in contrast to the yeast or plant Ku ring, *C. elegans* Ku does not regulate telomere length (Lowden, Meier et al. 2008). Before the present study, the presence of only the Ku ring and LIG-4 was interpreted either as evidence of a "minimal system" of NHEJ in the worm, or evidence pointing to the existence of other, non-conserved worm cNHEJ factors (Lemmens and Tijsterman 2011). With the discovery of NHJ-1 as an indispensable cNHEJ factor in this study, this question has been answered in favor of the latter interpretation.

I.2.5.5 cNHEJ in the plant world

Finally, cNHEJ is also conserved in plants as an important mode of DSB repair. Because they are sessile, plants are especially vulnerable to environmental sources of genotoxic stress from which animals can escape, provided they can sense them (Gimenez and Manzano-Agugliaro 2017). DNA damage and repair in plants has primarily been studied in *Arabidopsis thaliana*, a small flowering plant that is used as a model organisms for plants in general, but some studies have also been performed in commercially important crop species (for a good overview of the DNA damage

response in plants, see (Manova and Gruszka 2015)). Unsurprisingly, Ku orthologs are present and required for cNHEJ in *Arabidopsis* (Tamura, Adachi et al. 2002, West, Waterworth et al. 2002), rice (Nishizawa-Yokoi, Nonaka et al. 2012), wheat (Gu, Wang et al. 2014), and barley (Stolarek, Gruszka et al. 2015). An ortholog of DNA-PKcs has not been identified in any plant genome studied to date (Manova and Gruszka 2015). Of the processing factors, three Artemis orthologs have been identified in *Arabidopsis*, but it is not known whether these proteins play a role in cNHEJ (Charbonnel, Gallego et al. 2010). The terminal DNA ligase IV is present and required for cNHEJ, as is XRCC4 (West, Waterworth et al. 2000, West, Waterworth et al. 2002). An ortholog of XLF has not yet been found (Manova and Gruszka 2015). Like most other multicellular eukaryotes (excepting vertebrates), plants deficient for Ku or Ligase IV do not show growth of viability defects (Bleuyard, Gallego et al. 2006).

I.2.6 cNHEJ and human health and disease

Despite the fact that several cNHEJ factors were first discovered in a clinical setting, cNHEJ mutations are rare in human disease, probably because the process is essential for survival in vertebrates (Chang, Pannunzio et al. 2017). cNHEJ plays a critical role in V(D)J recombination which generates antigen receptor diversity in lymphocytes by assembling the variable section of antigen receptor genes through recombination between variable (V), diversity (D), and joining (J) segments (reviewed in (Malu, Malshetty et al. 2012)). Briefly, the activity of the RAG complex recombinase, acting on recombination signal sequences (RSS) which flank the V(D)J segments, generates hairpin-containing “coding” or V(D)J segment DNA ends. The coding ends are then bound by Ku and DNA-PKcs, the hairpins opened by Artemis, several palindromic and random nucleotides added by Pol μ, Pol λ, and TdT, and the ends ligated by the Ligase IV-XRCC4-XLF complex (Malu, Malshetty et al. 2012). The DNA ends generated by the RAG complex and AID deamination cascade during CSR can be inappropriately joined by cNHEJ machinery with each other or random chromosomal breaks, leading to B- and T-cell malignancies such as lymphomas and leukemias (for a recent review of the role of cNHEJ in oncogenic translocations, see (Lieber 2016)). Because of its role in V(D)J recombination, mutations in Artemis are rare but when they do occur they can reduce antibody counts and, in serious cases, cause severe combined immunodeficiency (SCID) (de Villartay 2015, Volk, Pannicke et al. 2015). In the Athabascan-

speaking Navajo and Apache nations, a founder mutation in Artemis causes a high incidence of SCID, with 1 in 2000 live-born babies entirely lacking B- and T-lymphocytes (Li, Moshous et al. 2002). Unlike RAG mutants, cell lines from this and other populations with SCID which arise from Artemis mutation are also IR-sensitive, reflecting the importance of Artemis in general cNHEJ (de Villartay 2015, Volk, Pannicke et al. 2015). Mutations in DNA-PKcs have also been associated with SCID, as have hypomorphic mutations in LIG4 (Woodbine, Gennery et al. 2014). In addition to SCID, patients with mutations in LIG4, XLF, and at least one patient with a DNK-PKcs mutation have been reported to exhibit developmental delay, dwarfism, and neurological abnormalities such as microcephaly (van der Burg, Ijspeert et al. 2009, Woodbine, Gennery et al. 2014, Mathieu, Verronese et al. 2015). These defects have been postulated to result from a lower apoptotic threshold in certain neural stem cells, resulting in a substantial loss of neural progenitors during the rapid divisions of embryonic neurogenesis, during which too many DSBs may arise for the impaired cNHEJ to handle (Woodbine, Gennery et al. 2014). In this model, the developmental defects are not observed in patients with Artemis mutations because the rapid division-induced DSBs either do not require Artemis or are repaired to a sufficient level not to cause apoptosis (Woodbine, Gennery et al. 2014). The general growth defects are similarly hypothesized to arise from stem cell loss in other compartments/niches (Woodbine, Gennery et al. 2014). In contrast to the other cNHEJ factors, several mutations in XRCC4 have been identified that are immunoproficient, with either a small reduction or no reduction in lymphocyte numbers, while nevertheless resulting in severe neurological defects and dwarfism (de Villartay 2015, Saito, Kurosawa et al. 2016). One proposed explanation for this observation is that a reduced level of XRCC4 function, or basal LIG4 activity in the complete absence of XRCC4, may be sufficient to handle the repair of the small number of breaks created by the RAG complex in V(D)J recombination, but insufficient to manage the numerous endogenous DSBs which arise in embryogenesis (de Villartay 2015). In addition to its role in lymphocytic malignancies, polymorphisms in cNHEJ and reduced cNHEJ activity have also been reported in breast cancer patients, raising the possibility that impaired cNHEJ may be a susceptibility factor in breast cancer as well (Bau, Mau et al. 2007). No Ku mutations in humans have been associated with disease, with the possible exception of association with some cancer types (de Villartay 2015). In mice, Ku

mutations cause premature aging, with increased and earlier incidence of aging-associated phenotypes like alopecia and kyphosis (Li, Vogel et al. 2007). These aging phenotypes have been postulated to result from chromosomal abnormalities which are evident in Ku mutant mice, and which may result from improper DSB repair or telomere maintenance defects (Li, Vogel et al. 2007).

Literature review, part III: *C. elegans* as a model organism

I.3.1 *Caenorhabditis elegans* as a model system

C. elegans (or, "the worm") is a small, free-living (*ie*, non-parasitic) roundworm that is widely used as a model organism to study molecular, cellular, and developmental processes in the context of a moderately complex metazoan organism (for a good summary of *C. elegans* as a model organism and in the wild, see (Blaxter 2011)). This species was first described in the closing year of the 19th century by the French nematologist Emile Maupas. He isolated it from the "rich humus" (translation by prof Marie-Anne Félix) in the outskirts of Algiers in May and November 1897, and used it as one of the focus species in a treatise on reproductive modes in nematodes (Maupas 1900). It has been used throughout the 20th century to test fundamental questions in biology ranging from meiotic development, sex determination, ploidy-body size relationship, natural genetic and phenotypic variation, and acclimation to high temperature (for an excellent review of this first stage of *C. elegans* research, see (Nigon and Felix 2017)). The watershed moment in the history of *C. elegans* research came in the 1960s, when Sydney Brenner, believing that the classical questions of molecular biology have already been answered (Ankeny 2001), chose the worm to take molecular biological research into the domains of development and neurobiology (for a concise review of Brennerian era *C. elegans* research, see (Ankeny 2001)).

C. elegans has a number of advantages that make it an attractive model in genetics, cell biology, and developmental biology. It possesses a determinate and invariant number of cells, and the entire somatic cell lineage of the animal (*ie* the sequence of divisions from the zygote to

adulthood) has been mapped (Sulston and Horvitz 1977, Sulston, Schierenberg et al. 1983), making it amenable to the study of processes that control and coordinate development. Furthermore, its life cycle is rapid, maturing from fertilization to a reproductively capable adult in about three days at standard culture conditions (Brenner 1974). It is highly tractable genetically, being a diploid with 6 pairs of chromosomes (5 autosome pairs and one X chromosome pair) (Brenner 1974). *C. elegans* is predominantly found in the form of a protandric hermaphrodite, which produces both sperm and oocytes in that order (Pazdernik and Schedl 2013). However, a male form, which possess only one X chromosome (X0) compared to two (XX) in hermaphrodites, and produces only sperm, occurs at a low frequency in the population (Brenner 1974). The existence of males makes possible genetic crosses, which are a critical tool of the laboratory geneticist. The *C. elegans* genome is approximately 100 Mbp in size (Sulston and Brenner 1974), and it has been sequenced two decades ago – the first complete genomic sequence of a metazoan to be generated (Consortium 1998). Beyond the genomic sequence, a large-scale effort from the model organism Encyclopedia of DNA Elements (modENCODE) project has generated a genome-wide map of transcriptional activity and chromatin organization, from which better gene models, putative noncoding RNA loci, and transcription-regulatory networks have been inferred (Gerstein, Lu et al. 2010, Niu, Lu et al. 2011).

C. elegans can be efficiently transfected with exogenous DNA sequences by microinjection, with the foreign DNA either retained as extrachromosomal arrays (Stinchcomb, Shaw et al. 1985) or integrated into the genome (Fire 1986). RNA interference (RNAi), a mechanism of silencing gene expression through the action of small interfering RNA (siRNA)-guided complexes which degrade mRNAs in the cytoplasm or shut down transcription by chromatin modification in the nucleus (Castel and Martienssen 2013), was first discovered in *C. elegans* (Fire, Xu et al. 1998). In addition to microinjection, RNAi can also be efficiently delivered to *C. elegans* via its bacterial food source, and libraries of bacterial strains collectively covering more than 90% of genes in the genome have been created and made available to the research community (Kamath, Fraser et al. 2003, Rual, Ceron et al. 2004). Genomic editing by CRISPR-Cas9, an RNA-guided endonuclease complex which has recently been coopted from its original role in bacterial anti-viral defense to serve as a precise way to induce DSBs nearly anywhere in the genome and create desired mutations via an

exogenously provided DSB repair template (Hsu, Lander et al. 2014), has been quickly adapted for use in *C. elegans* (Paix, Folkmann et al. 2015, Dickinson and Goldstein 2016). Because of its amenability to genetic modification and its translucent cuticle, molecular and developmental processes can be tracked in *C. elegans* by fluorescent protein tagging and live (time-lapse) microscopy (Chalfie, Tu et al. 1994, Fire 1994, Schnabel, Hutter et al. 1997). In addition, immunocytological analysis of protein localization is easily achieved by either dissecting the desired tissue out of the animal or "freeze-cracking" – compressing the larvae or adults between a slide and a cover slip, freezing them in liquid nitrogen, and rapidly removing the coverslip and the animals' cuticles with it, to allow antibodies access to the tissues inside (Duerr 2006).

Finally, *C. elegans* research is greatly facilitated by the existence and continual maintenance of WormBase, an online repository of genomic, phenotypic, expression, experimental, and other data relating to *C. elegans* and a continually expanding cast of nematode species, both free-living and parasitic (Harris, Antoshechkin et al. 2010).

I.3.2 Development of Caenorhabditis elegans

In common with many animals of the protostome lineage, *C. elegans* exhibits determinate development, in which each blastomere possesses a pre-specified fate. Individual development begins with fertilization, which occurs internally with oocytes being fertilized as they pass through the spermatheca, a structure in the gonad which houses spermatozoa (Kipreos 2005). The zygote deposits a hard, chitinous shell within which the remainder of embryonic development occurs (Johnston and Dennis 2012), and then initiates a series of stereotypical cell divisions which ultimately give rise to the adult animal with an invariant number of somatic cells (Kipreos 2005). During embryonic development, two stages are distinguished – a proliferative stage in which cell division takes place, and the gastrulation and organogenesis stage in which the cells are rearranged into the larval tissues and organs (Sulston, Schierenberg et al. 1983). At 20°C, the temperature at which *C. elegans* is typically cultured in the laboratory, embryonic development lasts about 12 hours, of which 5-10 hours are spent outside the mother's body after the egg has been laid (Muschiol, Schroeder et al. 2009). The full developmental life cycle of the animal is depicted in **Figure 1.3**.

I.3.2.1 Somatic development

In total, 671 cells are generated during embryogenesis, of which 113 undergo programmed apoptosis, such that the animal hatches into the first larval stage, L1, with 556 somatic cells and 2 germline cells, called primordial germ cells (PGCs) (Sulston, Schierenberg et al. 1983, Kipreos 2005, Pazdernik and Schedl 2013). Under optimal environmental conditions, primarily an abundant food source, the animal quickly transitions through three additional larval stages (L2, L3, and L4) before molting into a reproductively capable adult approximately 3.5 days after fertilization (Sulston and Horvitz 1977). The post-embryonic development is continuous, with the larval stages separated by a period of low activity called lethargus, which allows for molting and restructuring of the cuticle (Rougvie and Moss 2013). Somatic development after hatching is largely driven by division of somatic blast cells, which account for 53 out of 556 cells in newly hatched L1s, but generate a further 403 nuclei in the hermaphrodite and 475 in the male for a final number of 959 somatic nuclei in the adult hermaphrodite versus 1031 somatic nuclei in the adult male (Sulston and Horvitz 1977, Kipreos 2005). While the animal continues to grow in size during the remainder of its lifespan (2-3 weeks), no further somatic divisions take place (Rougvie and Moss 2013).

I.3.2.2a Germline development

Post-embryonic germline development takes place in concert with somatic development. The PGCs, individually named Z2 and Z3, are positioned between the two somatic gonad precursor cells, Z1 and Z4 (for a review, see (Pazdernik and Schedl 2013)). Gonadogenesis begins in late L1, when the germline and somatic precursors start dividing mitotically, expanding the gonad throughout the L2 and L3 stages (Pazdernik and Schedl 2013). In late L3, some of the germline cells exit the mitotic cycle and enter meiotic development, producing spermatozoa throughout the L4 stage and storing them in the spermatheca, before switching to oogenesis after the adult molt (Pazdernik and Schedl 2013). The development of the germline is depicted in **Figure 1.4**.

A *C. elegans* hermaphrodite produces about 300 spermatozoa before it switches to oocyte production, limiting its brood size to that number of progeny (Hodgkin and Barnes 1991). However, oogenesis continues beyond the depletion of endogenous sperm, so that if exogenous

sperm is provided by mating with a male, additional progeny may be generated, up to or slightly exceeding 1000 per hermaphrodite (Hodgkin and Barnes 1991). The male form is relatively rare in the population, occurring at a rate of about 0.1% as a consequence of X chromosome non-disjunction events in the first meiotic division, but because 50% of the spermatozoa of males carry no X chromosome, 50% of cross-progeny is male (Pazdernik and Schedl 2013).

I.3.2.2b Meiotic development

The gonadal lobes of the L4 larva and the adult animal possess a spatio-temporally organized structure, which makes the *C. elegans* gonad an excellent model for the study of meiotic development (reviewed in (Lui and Colaiacovo 2013)). The distal gonad (defined by its position furthest from the vulva) houses the germline stem cells, which continuously undergo mitotic divisions to replenish cells lost by gametogenesis and physiological apoptosis (Lui and Colaiacovo 2013). Moving proximally toward the vulva, the germ cells enter meiosis, and the distinct temporal stages of meiotic prophase I are spatially separated. Leptotene and zygotene (known also as the "transition zone") are the first stages of meiotic prophase I, characterized by crescent-shaped DNA polarization; the pachytene stage follows, with its distinct "thick threads" of chromosomes from which it derives its name; the last two stages are diplotene, when the DNA begins condensing, and diakinesis, where six pairs of homologous chromosomes, also called bivalents, are visible as discrete condensed bodies (Lui and Colaiacovo 2013). The oocytes arrest in diakinesis, and continue meiotic development only following fertilization which occurs as they pass through the spermatheca to the uterus. The remaining stages of meiosis subsequently complete, the fusion of pronuclei occurs, embryonic development begins, the egg is laid, and the L1 larva hatches, beginning the developmental cycle anew (Marcello and Singson 2010).

I.3.2.3 Alternative life-cycle stages

In addition to the normal developmental cycle which occurs under optimal conditions, *C. elegans* is also able to slow down and alter development in response to adverse conditions. The best studied of the three known developmental stages is dauer (from the German verb *dauern*, meaning "to persist, last, endure"), which forms at the second molt, although the decision to commit to dauer is made at the end of L1 stage on the integrative basis of several environmental

signals, including overcrowding, lack of food, and high temperature (for an overview of the dauer stage biology, see (Hu 2007)). Dauer larvae are visibly slimmer than normal L3 larvae, possess a hardened cuticle, and do not eat (*ie*, their pharynx is constricted and they do not perform pharyngeal pumping) (Hu 2007). Furthermore, metabolism is altered to increase energy conservation, and behavior changes to favor dispersal, which are strategies to survive or escape the adverse conditions (Hu 2007). The dauer larva can survive for several months (Ewald, Castillo-Quan et al. 2018), and upon encountering favorable conditions, re-enters normal development.

Another alternative developmental stage is L1 arrest, also known as L1 diapause, which occurs when L1 larvae hatch into an environment lacking a food source (reviewed in (Baugh 2013)). Although the larval morphology is not changed, metabolism is altered and stress resistance increases, allowing the animal to survive for several weeks without food (Baugh 2013). In the context of laboratory research, L1 arrest has proven to be a useful tool in two respects: 1) the arrest allows for developmental synchronization of a heterogeneous population by hypochlorite treatment ("bleaching"), which destroys larvae and gravid adults to yield eggs protected by their eggshell, after which they can be washed and allowed to hatch in a foodless environment, initiating L1 arrest (Baugh 2013); and 2) the increased stress resistance makes arrested L1s the best adapted *C. elegans* life stage to survive freezing in liquid nitrogen, allowing strains to be indefinitely kept in cryopreservation (Lewis and Fleming 1995).

The last described form of developmental arrest in *C. elegans* is the adult reproductive diapause (ARD) (Angelo and Van Gilst 2009). Unlike L1 arrest and dauer, ARD affects only the germline, and is most effectively triggered by starvation during the mid-L4 stage (Angelo and Van Gilst 2009). Under these conditions, every germ cell except a small population of stem cells is destroyed by apoptosis, delaying the onset of reproduction and extending adult life span three-fold (Angelo and Van Gilst 2009). The starved adults are able to recover from ARD after finding a food source, and re-populate the gonad with new germ cells, although the reproductive potential (brood size) decreases with each day spent in starvation (Angelo and Van Gilst 2009).

I.3.3 *Caenorhabditis elegans* in the wild and the laboratory

Despite often being described as a soil-dwelling nematode (possibly because of the influence of Emile Maupas' 1900 publication), outside the laboratory *C. elegans* is usually found in rotting plant or fungal material (Blaxter 2011). The most commonly used wild type reference strain, N2 or Bristol, was isolated from a mushroom compost during a nematology course offered in Bristol in 1951 (Nigon and Felix 2017). The preferred habitat of *C. elegans* in the wild seems to be rotting fruit in the temperate regions of the planet (Barriere and Felix 2005, Cutter, Felix et al. 2006, Barriere and Felix 2007, Dolgin, Felix et al. 2008, Barriere and Felix 2014), with the dauer stage likely playing an important role in surviving the non-fruiting season (Blaxter 2011). As a species, *C. elegans* shows a cosmopolitan distribution (Kiontke and Sudhaus 2006, Dolgin, Felix et al. 2008), and appears to have been distributed globally at least in part by human activity, based on the observations that most wild type isolates have been isolated from anthropogenic sources like compost heaps and that genetic distance between wild type isolates does not correlate with geographic distance (Blaxter 2011). Despite relatively high local genetic diversity in *C. elegans* populations, which are indicative of low levels of outcrossing in the natural population, worldwide genetic diversity of *C. elegans* is relatively low (Barriere and Felix 2005). Six large haplotypes are shared by a set of almost 200 strains isolated from locations on all continents except Antarctica, excepting three strains isolated along the Pacific Rim; this has been interpreted as evidence of a selective sweep which occurred in the last few centuries, ostensibly as an adaptation to human-created habitats (Andersen, Gerke et al. 2012).

The most commonly used laboratory strain, N2, differs from the more recent wild isolates in several respects which may represent adaptation to life in captivity or neutral change following relaxed selection. Examples include the lack of a mating plug, a structure that is deposited by males during copulation to limit the mating efficiency of other males, because of a mutation in the key plug structural gene *plg-1* (Hodgkin and Doniach 1997, Palopoli, Rockman et al. 2008), and solitary feeding behavior (in contrast to more common social feeding), which arose as a result of an amino acid substitution in *npr-1*, a neuropeptide receptor (Rogers, Reale et al. 2003, Gloria-Soria and Azevedo 2008). However, it must be noted that several other wild isolates also lack a copulatory plug or engage in solitary feeding behavior, although such strains are in the minority.

I.3.4a *Caenorhabditis elegans* on the tree of life – related species

C. elegans shares its ecological niche with closely related species in the *Caenorhabditis* genus, including *C. briggsae* and 25 others identified so far (Kiontke and Fitch 2005, Kiontke and Sudhaus 2006, Felix and Braendle 2010, Blaxter 2011). These *Caenorhabditis* species as well as others from the related genera *Pristionchus* (Hong and Sommer 2006, Mayer, Herrmann et al. 2007) and *Oscheius* (Baille, Barriere et al. 2008) have been studied in comparison to *C. elegans* to model genomic evolution (Blaxter 2011). As part of the subclass Rhabditina, the genus *Caenorhabditis* belongs to a group that includes free-living, commensal, and parasitic species, which play critical roles in their ecological webs (Blaxter 2011). The parasitic species in particular have great economic importance, and some insect parasites such as those in the *Heterorhabditis* and *Steinernema* genera are used as biological pest control agents in agriculture (Blaxter 2011, Kenney and Eleftherianos 2016). Others, like the blood-feeding hookworms *Necator americanus* and *Ancylostoma duodenale* parasitize on hundreds of millions of people in the tropics and subtropics (Cantacessi, Mitreva et al. 2010), while plant-parasitic nematodes cause 8-15% of all crop failures (Kiontke and Fitch 2005). While parasitism has evolved independently in many nematode taxa (Blaxter and Koutsovoulos 2015), research conducted in *C. elegans*, particularly on the dauer stage which is similar to the infectious life stages of parasitic nematodes, provides a multitude of genomic and functional data which can be used as a baseline for inter-nematode comparison (Hu 2007, Blaxter 2011).

I.3.4b *Caenorhabditis elegans* on the tree of life – the phylum Nematoda

As a nematode, *C. elegans* belongs to a phylum of immense diversity, with approximately 30,000 described species and possibly more than a million total species (Lambshead and Boucher 2003). Most of the described species are benthic (*ie*, seafloor-dwelling) marine organisms which heavily outnumber other animals in this environment, but nematodes also dominate in other sampled environments, such as the rainforest canopy and top soil layers (Creer, Fonseca et al. 2010). Nathan Augustus Cobb, the founder of American nematology, famously illustrated the abundance of nematodes on the planet by stating that, were all other matter on Earth to disappear, the outline of major geographical features would still be recognizable because of the

nematode biomass, and the distribution of other species inferable from the distribution of their nematode parasites (Cobb 1915, Huettel and Golden 1991). In addition to global distribution and abundance, nematodes are characterized by hardiness, perhaps best exemplified by sensationalized but accurate accounts of *C. elegans* surviving the atmospheric disintegration of the Space Shuttle Columbia and the resulting 42-32 kilometer fall to Earth (Szewczyk, Mancinelli et al. 2005), or of nematodes from the orders Rhabditida and Plectida being cryopreserved in permafrost during the Pleistocene and recovered after 40,000 years of dormancy (Shatilovich, Tchesunov et al. 2018).

I.3.4c *Caenorhabditis elegans* on the tree of life – the wider context

In the tree of life, the phylum Nematoda is positioned inside a larger group of invertebrates called Ecdysozoa, or molting animals (Aguinaldo, Turbeville et al. 1997, Philippe, Lartillot et al. 2005). Ecdysozoa most prominently also includes arthropods, a highly successful group to which the model organism *D. melanogaster* belongs. Together with Lophotrochozoa, the other major branch of invertebrates which includes molluscs and segmented worms, Ecdysozoa belong to a larger group of animals called Protostomia (Halanych 1995, Philippe, Lartillot et al. 2005). The protostomes belong to Bilateria, or bilaterally-symmetrical animals, within which they are the sister group to Deuterostomia, a clade which includes chordates and, ultimately, mammals (Nielsen 2017). Animals (the kingdom Animalia) are closely related to the kingdom Fungi, and belong to a larger grouping of eukaryotes called Opisthokonta, which is more distantly related to plants and other photosynthesizing eukaryotes (Steenkamp, Wright et al. 2006). Ultimately, the domain Eukaryota is a close relative of the domain Archea, and together with the domain Bacteria they comprise all known cellular life (Walsh and Doolittle 2005).

I.3.5 DNA repair in *C. elegans*

C. elegans conserves all of the DNA repair pathways introduced in **Part I** of this chapter, although as has already been mentioned for the case of cNHEJ (section **I.2.5.4**), not every vertebrate DNA repair gene has a *C. elegans* ortholog (for a brief overview, see (O'Neil and Rose 2006)). Starting in 1976, IR has been used on *C. elegans* as a means of generating genetic balancers, but proper radiological studies in *C. elegans* began in the 1980s with the isolation of the first nine *rad*

mutants using IR and UV radiation, and have greatly expanded in the following decades (O'Neil and Rose 2006, Sakashita, Takanami et al. 2010, Lans and Vermeulen 2011). Today, *C. elegans* has become an important model for the study of DNA damage, especially that of radiological causes, owing to the utility of studying DDR *in vivo* both in the entire animal and in specific tissue contexts (Sakashita, Takanami et al. 2010, Lans and Vermeulen 2015). Some limitations of *C. elegans* as a general model of DDR exist, however. In general, the repair pathways appear less complex (*ie*, involve fewer players), a fact particularly true of the "minimal" cNHEJ system. Additionally, *C. elegans* appears to lack several key components of the classic DDR signaling cascade (see sections **I.1.6.5.1** and **I.2.3**), including the histone variant H2AX, the DDR mediator MDC1, and the ubiquitin E3 ligase RNF8 (Lans and Vermeulen 2015).

I.3.5.1 The tissue context of DNA repair

The ability to study the effects of DNA damage in a small number of well-defined tissue contexts is a particular strength of the *C. elegans* model (for a brief review on this topic, see (Lans and Vermeulen 2015)). At one level, somatic tissues can be contrasted with the germline in both the objective and the strategy of DNA repair. In the soma, which dies with every brief generation, the aim of DNA repair is to maintain function and enable survival and reproduction; in the germline, which represents the immortal proliferative potential of the species, the goal of DNA repair is the faithful replication of genetic information. Beyond the soma/germline division, DNA repair in *C. elegans* can be studied in several tissue and developmental stage subcontexts, including (Lans and Vermeulen 2015): 1) the adult germline stem cells (the mitotic zone of the adult gonad), where several repair pathways, including HR, NER, BER, and TLS, as well as strong checkpoints ensure the high fidelity of repair; 2) pachytene cells, in which DSBs are initially preferentially repaired by inter-homolog HR, which ensures proper segregation in anaphase I and promotes genetic diversity, but are later in pachytene shifted to inter-sister HR for high fidelity repair; 3a) the early embryo, in which the demands for rapid cell division and synchronous development result in tolerance of ssDNA lesions via TLS, and where DSB repair preferentially occurs by HR; 3b) the late embryo, in which the preferred mode of DSB repair in most blastomeres (excepting the PGCs) is cNHEJ because they are G1-arrested while gastrulation and organogenesis takes place; 4) somatic cells of developing larvae, which are served by several repair pathways,

including cNHEJ, HR, NER, and Alt-EJ, but are remarkably resistant to genotoxic stress; and 5) systemic, or non-cell autonomous responses, in which signals from one tissue may influence the DDR of another (*eg* neurons to germline (Sendoel, Kohler et al. 2010), or germline to the entire soma (Ermolaeva, Segref et al. 2013)) (Lans and Vermeulen 2015). A brief overview of some of the conserved DNA repair pathways in *C. elegans*, as well as the physiological consequences of their dysfunction, is given below.

I.3.5.2 Single-strand lesion repair

C. elegans conserves at least two glycosylases (UNG-1 and NTH-1) and two AP endonucleases (EXO-1 and APN-1) that are active in BER. However, no strong phenotypes have been associated with the loss of these enzymes (Shatilla, Leduc et al. 2005, Nakamura, Morinaga et al. 2008, Hunter, Gustafson et al. 2012). By contrast, many of the *C. elegans* proteins orthologous to mammalian NER factors, including XPC-1, XPF-1, XPG-1, and ERCC-1, show increased UV sensitivity (for a review of NER in *C. elegans*, see (Lans and Vermeulen 2011)). The physiological effects of UV exposure depend on the developmental stage, with UV irradiation at the L1 stage causing a terminal developmental arrest, and exposure in adulthood resulting in a decreased lifespan and an increased embryonic lethality in the progeny (Lans and Vermeulen 2011). Several orthologs of MMR proteins, including MSH-2, MSH-6, and MLH-1, have been described. The loss of these factors results in increased mutation rates and DNA repeat instability (Tijsterman, Pothof et al. 2002). Finally, several PARP orthologs exist in *C. elegans*, and the loss of these is synthetic lethal with a depletion of the cohesin HIM-1, consistent with a role of PARPs in SSB repair and prevention of replication fork collapse (McLellan, O'Neil et al. 2012, Rose 2014).

I.3.5.3 DSB repair

C. elegans possesses both of the two primary modes of DSB repair, cNHEJ and HR (for a review of DSB repair in the worm, see (Lemmens and Tijsterman 2011)). All of the major factors involved in the HR pathways are conserved, including the MRN complex, the strand invasion protein RAD-51, the resection enzymes DNA-2 and EXO-1, BRC-1, and others (Lemmens and Tijsterman 2011). Single strand annealing in the worm requires XPF-1 (Pontier and Tijsterman 2009). In contrast to HR, and as mentioned in section **I.2.5.4**, the Ku ring components CKU-70 and CKU-80, and the

DNA Ligase IV ortholog LIG-4, are the only cNHEJ factors identified before the present study (Clejan, Boerckel et al. 2006). The existence of Alt-EJ pathways in *C. elegans* has been demonstrated by the ability of the worm to perform end joining in the absence of *lig-4*, *brc-1*, and *xpf-1* (Pontier and Tijsterman 2009). More recently, an Alt-EJ pathway dependent on the polymerase θ ortholog POLQ-1, has been described (van Schendel, van Heteren et al. 2016, Macaisne, Kessler et al. 2018). As in other organisms, DSBs represent the most deleterious form of genotoxic stress in *C. elegans*. Unrepaired DSBs effect distinct outcomes in the somatic and germline tissues. In the soma, the chromosome mis-segregation that results from DSBs causes developmental defects, which are most noticeable as morphological abnormalities in the vulva, the egg-laying and mating organ of the hermaphrodite animal (Clejan, Boerckel et al. 2006, Weidhaas, Eisenmann et al. 2006). Chromosome mis-segregation in the adult germline results in embryonic lethality in the case of autosomes, or a high incidence of males (Him phenotype) in the case of the X chromosome, while mutagenic repair can result in an increased embryonic lethality (Emb phenotype) (Lemmens and Tijsterman 2011).

Rationale for the present study

The initial rationale for the present study was to investigate the potential role of the germline protein HTP-3, which has well described roles in meiotic development ranging from coordination of key prophase I processes to the regulation of the DNA damage response at the chromosome axes (Goodyer, Kaitna et al. 2008, Severson, Ling et al. 2009, Couteau and Zetka 2011, Lui and Colaiacovo 2013), in the early germline response to IR-induced genotoxic stress. The project began as an RNAi screen to identify chromatin factors which may collaborate with HTP-3 in effecting DDR in the germlines of L1 larvae after IR treatment. However, anomalies in the wild type response to IR channeled the project onto a different path, which led to the discovery and initial characterization of NHJ-1, a novel regulator of cNHEJ in *C. elegans*. That story is described in the following two chapters (**Chapter II** and **Chapter III**) of this thesis.

Figures

Figure 1.1

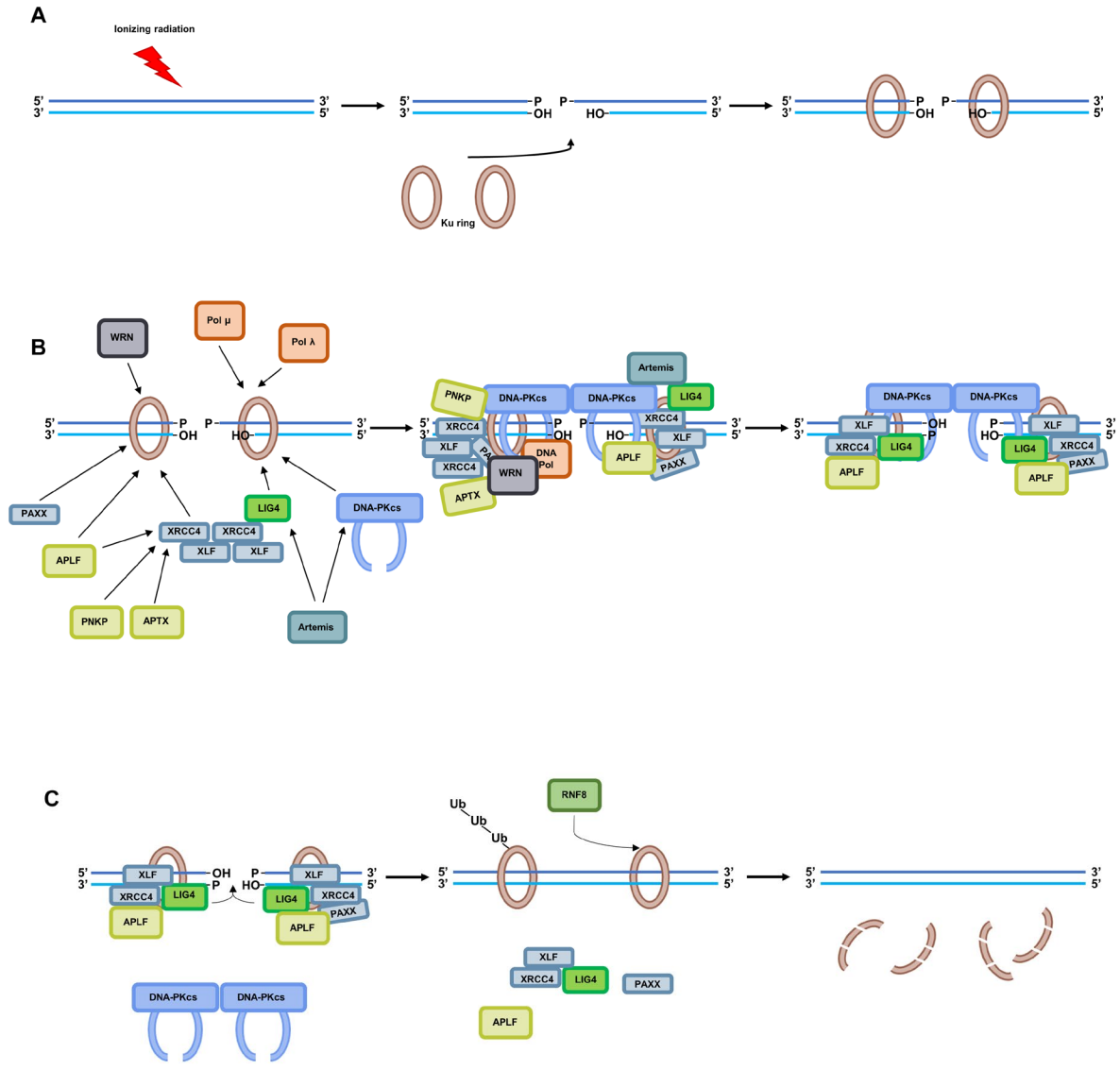


Figure 1.1 - Model of cNHEJ in mammalian cells

This figure represents the current understanding of the repair of chemically complex DSBs by the mammalian cNHEJ pathway. It is based on similar graphical summaries of cNHEJ presented in (Davis and Chen 2013, Wang and Lees-Miller 2013, Radhakrishnan, Jette et al. 2014). The exact stoichiometry of proteins involved is not known.

(A) Ionizing radiation generates a chemically complex DSB, characterized by loss of nucleotides (ssDNA overhangs), loss of 5' phosphate groups, and abberrant phosphate groups on the 3' termini. The Ku ring quickly and efficiently binds the free DNA ends, stabilizing them and protecting them from non-specific nucleases.

(B) The Ku ring acts as a "toolbelt" to recruit the kinase DNA-PKcs, the structural scaffold proteins XRCC4, XLF, and PAXX, DNA end-processing factors such the WRN helicase, the nucleases APLF, APTX (via XRCC4), and Artemis (via DNA-PKcs and LIG4), the kinase/phosphatase PNKP, the DNA polymerases Pol μ and Pol λ, and the ligase LIG4 (in complex with XRCC4 and XLF). The collective action of these enzymes results in ssDNA overhang removal, gap-filling DNA synthesis, phosphorylation of 5' ends, and dephosphorylation of 3' ends, making the processed ends compatible for ligation. Not shown are the phosphorylation events that mediate recruitment and coordinate activity of the proteins involved.

(C) Terminal ligation is preceded by the autophosphorylation of DNA-PKcs (not shown), which mediates its dissociation from the cNHEJ complex. The ligase activity of LIG4 is promoted by APLF. Once the DNA ends have been ligated, the Ku ring is trapped on the repaired chromosome, and is removed by ubiquitin-mediated degradation. The E3 ligase RNF8 is currently understood to mediate the ubiquitylation of Ku.

See section **I.2.2** for more detail.

Figure 1.2

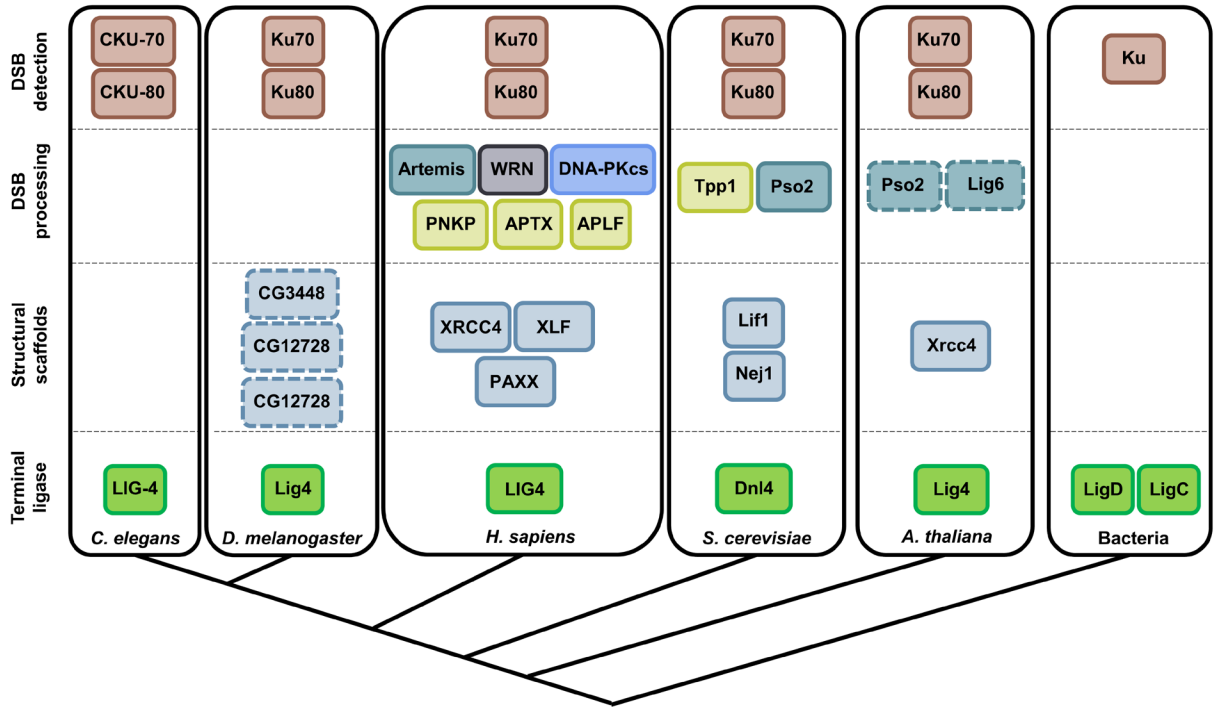


Figure 2.2 - Conservation of cNHEJ factors

This figure shows the evolutionary conservation of cNHEJ factors in several well-studied taxa, organized by the distinct steps of the cNHEJ process. Although located within the eukaryotic lineage whose other members share several DSB processing factors and structural scaffold proteins, *C. elegans* appears to possess only the 3 universally conserved cNHEJ pathway components.

Proteins with sequence homology are color-coded. Proteins inferred but not confirmed to be active in cNHEJ are indicated by a dotted outline. The DNA polymerases known to be involved in mammalian and yeast cNHEJ are not shown. Branch lengths are not to scale and do not indicate a quantitative degree of taxon relatedness. See section **I.2.5** for more detail.

Figure 1.3

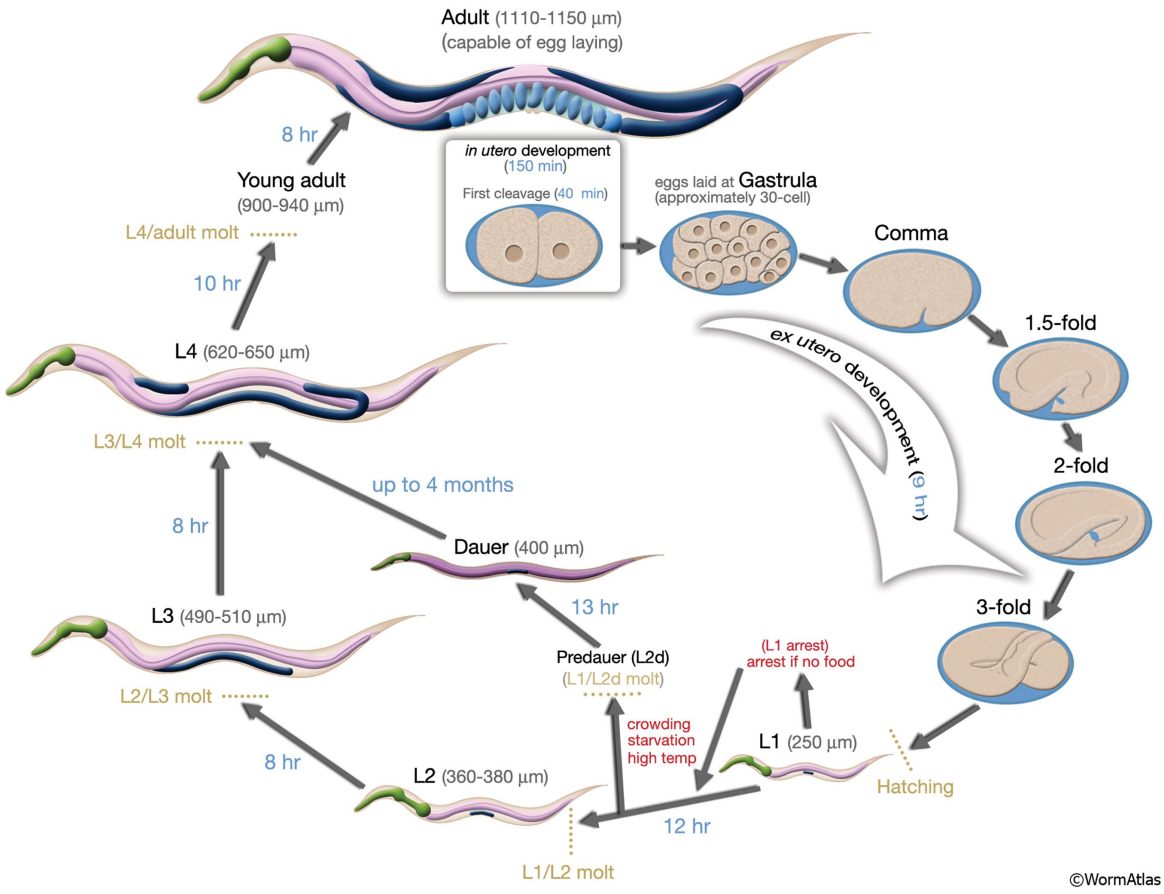


Figure 3.3 - Life cycle of *Caenorhabditis elegans*

This figure presents the development of *C. elegans* at 22°C, showing time spent in and the sizes of the different larval stages. Note that experiments in the present study were performed at 20°C, at which temperature the egg hatches 10-12 hours post-fertilization, and the animals spend 14 hours in L1, 8.5 hours in L2, 9 hours in L3, 12.5 hours in L4, and start laying eggs approximately 9 hours after molting into adults (Byerly, Cassada et al. 1976). See sections **I.3.2**, **I.3.2.1** and **I.3.2.3** for more information. This figure was made by Wormatlas, which holds the copyright thereto. It is reproduced here for an approved non-profit educational and scientific purpose.

Figure 1.4

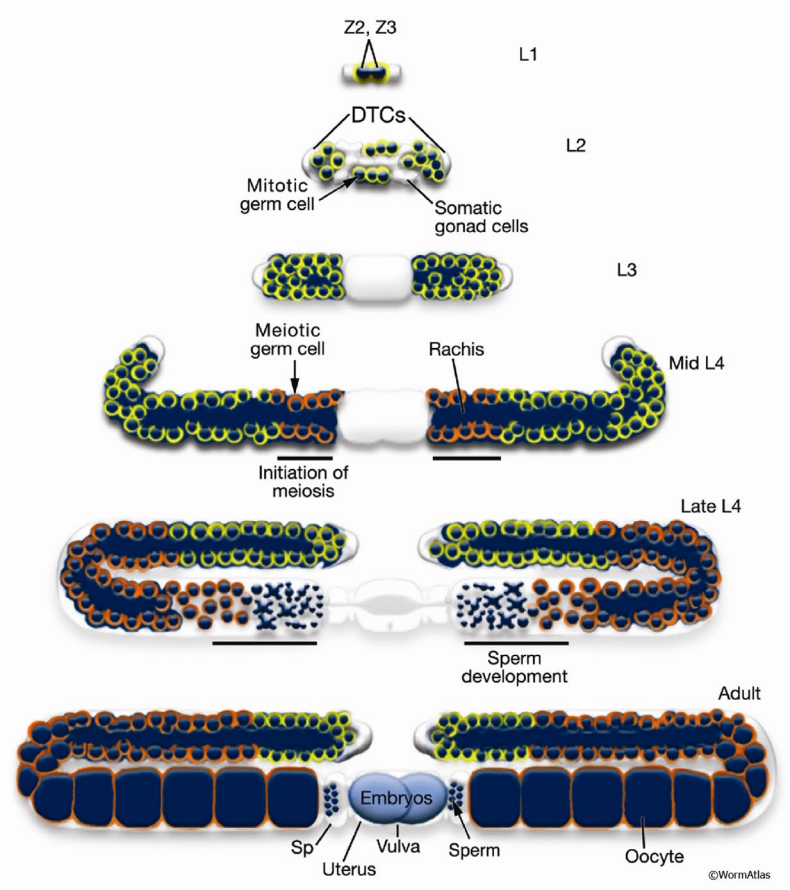


Figure 4.4 - Germline development in *Caenorhabditis elegans*

This figure shows the development of the *C. elegans* germline, starting with the primordial germ cells (PGCs) at the L1 stage and ending with the oocyte-producing adult germline. See sections **I.3.2.2a-b** for more detail. Cells outlined in yellow are mitotically proliferating, while cells outlined in orange are undergoing meiosis. This figure was made by Wormatlas, which holds the copyright thereto. It is reproduced here for an approved non-profit educational and scientific purpose.

Chapter II: Discovery of a novel regulator of ionizing radiation response in early *C. elegans* larvae

II.1 Introduction and screening for chromatin regulators of the DNA damage response in the L1 germline

The project that led to the discovery of NHJ-1 as a critical regulator of cNHEJ began as an investigation into the role of the germline protein HTP-3 in the IR response of early larvae. HTP-3 is an essential component of the synaptonemal complex (SC), a conserved proteinaceous structure which plays an essential role in stabilizing the interactions between homologous chromosomes during the first meiotic prophase (reviewed in (Zickler and Kleckner 2015) and (Cahoon and Hawley 2016)). The SC is composed of two main parts: 1) the meiotic chromosome axis elements, also known as axial elements or lateral elements, called such because they assemble along the length of the chromosomes in prophase I; and 2) the transverse elements, also called central elements, which assemble between the axial elements of two homologous chromosomes to stabilize the inter-homolog interaction and effect synapsis (Cahoon and Hawley 2016). HTP-3 is the most upstream of the lateral SC elements, interacting with cohesins on the chromosomes and forming a platform required for the localization of other lateral SC elements including HIM-3, HTP-1, and HTP-2 (Goodyer, Kaitna et al. 2008, Severson, Ling et al. 2009, Lui and Colaiacovo 2013). Therefore, HTP-3 is essential for homolog alignment, synapsis, and crossing over. Furthermore, it interacts with the MRN complex and is required for the formation of programmed DSBs in prophase I (Goodyer, Kaitna et al. 2008). Beyond being a protein of central importance to several meiotic processes, two observations about HTP-3 prompted us to study its potential extra-meiotic roles: 1) It is expressed in germline nuclei prior to the onset of meiosis, notably in the PGCs during the L1 stage, suggesting a possible function during the early larval stages; and 2) It is required for the acquisition of wild type levels of acetylated histone 2A lysine 5 (H2AK5Ac), a DNA-damage responsive chromatin mark, suggesting that HTP-3 may play a role in establishing the appropriate chromatin architecture in meiotic DDR (Couteau and Zetka 2011).

Joining the observation of HTP-3 expression in the PGCs with its role in meiotic DDR, we hypothesized that it may play a role in DDR in the early germline. Because HTP-3 is essential for the critical meiotic processes which ensure proper chromosome segregation in the gametes, *htp-3* null mutants are generally inviable, and show ~90% embryonic lethality (Monique Zetka, unpublished data). However, *htp-3(vc75)*, a missense allele generated in a collaborative reverse genetic screen in which the histidine at position 97 in the protein chain has been substituted by a tyrosine (Gilchrist, O'Neil et al. 2006), produces viable progeny (Monique Zetka, unpublished data). In *htp-3(vc75)* animals, levels of H2AK5Ac on the meiotic axes are reduced, and the axes exhibit localized separation or desynapsis, which are phenotypes also observed in wild type animals following IR treatment (Couteau and Zetka 2011). I therefore aimed to identify possible genetic interactors of the *htp-3(vc75)* mutation whose loss may sensitize the animals to IR at the L1 stage, particularly in terms of brood size outcomes, *ie*, in the number of progeny generated in adulthood after IR treatment during the L1 stage. Because in the context of meiotic chromosomes HTP-3 directs histone modification, I decided to limit my search to chromatin modifying enzymes and associated proteins. As mentioned in **Chapter I** (section **I.3.1**), *C. elegans* is an excellent model for RNAi-based screens because of the ease of knockdown by feeding as well as the availability of RNAi libraries which cover most protein-coding genes in the genome. For my screen, I made use of a published list (sub-library) of genes encoding chromatin modifiers and related factors, which includes a total of 464 RNAi clones (Tursun, Patel et al. 2011), although only 429 of these were in our possession. I performed the screen simultaneously in three genotypes: the Bristol wild type strain N2, *htp-3(vc75)*, and *rrf-1(pk1417)*. The primary aim was to identify genes whose knockdown, in conjunction with IR treatment, affects *htp-3(vc75)* fertility more than wild type fertility (see **Figure 5.1** and section **V.7** for more details on methodology). The RNA-dependent RNA polymerase (RdRP) *rrf-1(pk1417)* mutant was included to identify genes which may have a germline-specific effect on post-IR fertility, as *rrf-1* mutants were believed at the time to be refractory to somatic RNAi, although this was soon after demonstrated not to be true in all somatic tissues (Kumsta and Hansen 2012).

II.2 Endogenous RNAi factors and HTP-3 appear to sensitize L1 larvae to ionizing radiation

Before I commenced screening in earnest, I noticed an anomaly in the brood size response of N2 compared to *htp-3(vc75)* and *rrf-1(pk1417)* mutants. Irradiated at the L1 stage with 75 Gy of IR, N2 animals exhibited a markedly reduced fertility compared to unirradiated controls, with only 28% of the animals producing more than 50 progeny (**Figure 2.1A**). Considering that IR is a potent genotoxic stressor, this observation is not surprising on its own. However, both *htp-3(vc75)* and *rrf-1(pk1417)* showed a significantly more modest brood size reduction, with 88% and 93% of irradiated animals, respectively, producing more than 50 progeny ($p<0.001$ vs N2 for both genotypes) (**Figure 2.1A**). It thus appeared that the *htp-3(vc75)* missense mutation and the *rrf-1(pk1417)* deletion conferred IR-resistance, even in absence of any RNAi knockdown. This finding did not impact the decision to perform the screen, as IR-resistance could be modulated by chromatin factors in addition to newly implicated RRF-1 and HTP-3. Furthermore, the resistant phenotype of *rrf-1(pk1417)* mutant raised the possibility that the endogenous small interfering RNA (endo siRNA) pathway may play a role in sensitizing the animals to IR, since RRF-1 plays a critical role in this process. Small interfering RNAs are small (20-30 nucleotide long) RNA molecules which can interact with proteins called Argonautes to form mRNA silencing complexes (Boisvert and Simard 2008). Endogenous siRNAs are encoded by the genome and are involved in the regulation of endogenous gene expression. Exogenous siRNAs are created from extra-genomic sources and play a role in the anti-viral immune response in *C. elegans*, though this mechanism has been co-opted for the silencing of endogenous genes in the laboratory (Boisvert and Simard 2008). Endogenous and exogenous siRNA biogenesis is effected by distinct but partially overlapping pathways (Gent, Lamm et al. 2010, Vasale, Gu et al. 2010). Endogenous siRNAs are generated by a two-phase process. In the first phase, the relatively rare 26 nucleotide long primary endo siRNAs are generated from genomic transcripts by the activity of the RNAse III-like enzyme Dicer (DCR-1), the RdRP RRF-3, and other accessory factors, and are then bound and stabilized by the primary Argonaute ERGO-1 (Gent, Lamm et al. 2010). In the second phase, the more abundant 22 nucleotide long secondary endo siRNAs are produced by the action of the RdRP RRF-1 and possibly other RNA polymerases, and are then bound by one of the more than 20 secondary Argonautes in *C. elegans* (Boisvert and Simard 2008, Gent, Lamm et al. 2010).

Broadly similar to endo siRNA biogenesis, the exo siRNA synthesis pathway involves the processing of exogenous RNA sources by DCR-1 and accessory proteins, but does not require RRF-3 and possesses its own primary Argonaute, RDE-1, rather than ERGO-1. The downstream amplification requires RRF-1 in the soma and EGO-1 in the germline, though some EGO-1 activity may occur in certain somatic tissues, notably the intestine (Boisvert and Simard 2008, Gent, Lamm et al. 2010, Kumsta and Hansen 2012).

Because there was no trigger for exogenous RNAi in the initial L1 irradiation experiment (**Figure 2.1A**), I interpreted the *rrf-1(pk1417)* phenotype as a consequence of endo siRNA pathway disruption. This interpretation seemed even more plausible in light of the discoveries that both endo and exo RNAi pathways are capable of gene silencing via histone 3 lysine 9 trimethyl (H3K9me3) deposition and heterochromatinization (Burton, Burkhart et al. 2011, Gu, Pak et al. 2012), which I considered a possible link to HTP-3 or other chromatin factors. Therefore, I wanted to assay whether other endo siRNA biogenesis mutants also exhibit IR-resistance. I tested this at a higher dose of 100 Gy, at which the difference between N2 and *rrf-1* or *htp-3(vc75)* is even more pronounced (**Figure 2.1B**). At this dose, *rrf-3* and *ergo-1* null mutants showed a significantly higher brood size, with 38% and 43% of the animals producing more than 50 progeny, compared to less than 1% in N2 ($p<0.001$) (**Figure 2.1B**). Similar levels of IR-resistance were also observed in the four secondary Argonaute null mutants tested, *ppw-1, ppw-2, sago-1,* and *sago-2* (**Figure 2.1B**). Interestingly, the *dcr-1(mg375)* mutant was not significantly different than N2 in its post-IR fertility ($p>0.05$) (**Figure 2.1B**). I used this allele because null mutants of *dcr-1* are not viable. It is a missense mutation converting glycine 492 in Dicer's helicase domain to an arginine, which greatly reduces the protein's helicase activity (Welker, Pavelec et al. 2010). While this does not affect exogenous siRNA biogenesis, it greatly reduces the production of endogenous siRNAs from a subset of genomic sources (Welker, Pavelec et al. 2010). I therefore initially concluded that IR-sensitivity must be mediated by Dicer helicase-independent endo siRNAs. However, the fact that an inactivating mutation in any of the tested secondary Argonautes, which are known to be redundant for exogenous RNAi (Yigit, Batista et al. 2006), produces an IR-resistant phenotype not different than *rrf-3* or *rrf-1* (**Figure 2.1B**), raised the possibility that the IR-resistance of these strains may be caused not by the endo siRNA pathway mutations, but may instead be the default

IR response. The IR-sensitivity in N2 might therefore be caused by a cryptic, unannotated mutation not present in the other tested strains, and not represent the true wild type response. Adding credence to this hypothesis, mutants of one of the first two hits from the RNAi screen, *nurf-1* (a component of the Nucleosome Remodeling Factor chromatin remodeler complex) and the methyltransferase *set-22* both proved to be resistant to IR (**Figure 2.1B**), even though *nurf-1* was identified as a negative regulator of post-IR brood size in the RNAi screen, and *set-22* as a positive regulator (**Figure 2.2A**). I additionally tested a completely unrelated strain for the L1 IR response, an *unc-119(ed9)* mutant with an integrated construct (*orIs20*) carrying the *unc-119(+)* rescuing sequence and a moesin gene tagged with GFP. At this stage, I also decided to abandon the brood size binning approach and quantify the actual brood size of each animal. As expected based on previous experiments, the brood size of N2 animals irradiated with 75 Gy was drastically reduced to a median of 77 progeny (inter-quartile range (IQR) 74.75) compared to 333.5 progeny (IQR 44.5) in unirradiated controls (**Figure 2.1C**). By contrast, the median number of progeny in *orIs20; unc-119(ed9)* animals after IR treatment was 160 (IQR 215.5), significantly higher than in irradiated N2 animals ($p<0.001$) (**Figure 2.1C**). I therefore began strongly favoring the idea that a cryptic mutation in the N2 strain kept in the Zetka lab was causing the animals to become sensitive to IR.

The RNAi screen yielded 33 candidates with an effect on post-IR brood size (**Figure 2.2**). However, no candidates were found in which the genetic background (whether N2, *htp-3(vc75)*, or *rrf-1(pk1417)*) modulated the IR response in combination with the RNAi knockdown. Because the initial rationale of the screen was to find genetic interactors of *htp-3* in the IR response, following up on the candidates from the screen appeared less promising than furthering the investigation of the severe IR-sensitivity discovered in the N2 line. I thus decided to focus exclusively on this latter phenomenon, and more deeply examine its nature and genetic basis.

II.3 The Bristol wild type strain N2 exhibits diversity in the L1 radiation response

As a first step in this investigation, I obtained the N2 strain from the Caenorhabditis Genetics Center (CGC), a depository of strains at the University of Minnesota Twin Cities, which stores *C. elegans* and other nematode strains and distributes them to the research community. I also

obtained the N2 line from the neighbouring laboratory of Dr. Richard Roy at McGill University. While the Roy lab N2 exhibited a resistant phenotype with a median of 172.5 (IQR 51) progeny after L1 irradiation, both the Zetka lab N2 and the CGC N2 displayed significantly lower post-IR brood sizes of 21 (IQR 46.5) and 30 (IQR 51.75) progeny, respectively ($p<0.001$ for both genotypes vs Roy lab N2) (**Figure 2.3**). While the resistant phenotype of Roy lab N2 confirmed that there exists unannotated genetic variation among N2 lines, the sensitive phenotype of the CGC N2 suggested that the sensitivity of the Zetka lab N2 was not due to an isolated mutation, and that it may be more widespread. To further investigate the distribution of IR-resistance and -sensitivity, I tested N2 lines from four additional sources – Hekimi lab from McGill University, Andersen lab from Northwestern University (USA), Zhen lab from Lunenfeld-Tanenbaum Research Institute (Canada), and the National Bioresource Project (NBRP) of Japan. I also tested 18 non-N2 wild type isolates from diverse geographic locations covering all major landmasses on the planet except for Antarctica (**Figure 2.4**), as well as two strains of *C. briggsae*. Every strain tested, except the Zetka lab N2 and the CGC N2, showed a resistant phenotype following IR treatment at the L1 stage (**Table 2.1**). The most parsimonious explanation of this distribution pattern is that IR-sensitivity arose relatively recently in the N2, during its existence as a commonly used laboratory strain, and proliferated among some of the laboratory lines since.

II.4.1 The sensitive N2 [S] strain shows dose-dependent IR-sensitivity at L1

I next turned my attention to a better characterization of the L1 IR response. Using a single isolated animal, I derived a resistant N2 strain, henceforth called N2 [R], from the Andersen lab N2, and a sensitive N2 strain, henceforth called N2 [S], from the CGC N2. I had heretofore worked primarily with very high IR doses of 75 Gy and 100 Gy, which had a striking impact on the sensitive animals' post-IR fertility. To test whether this response is a consequence of an IR dose threshold or whether it scales with the dose, I looked at the response of N2 [R] and N2 [S] at two lower doses of 25 Gy and 50 Gy, in addition to 75 Gy. While a significant difference between the two N2 lines is detectable even at 25 Gy ($p<0.05$), the magnitude of the difference is far greater at 50 Gy and 75 Gy ($p<0.001$ for both doses) (**Figure 2.5A**). Furthermore, irradiated N2 [S] animals show a significantly lower brood size compared to unirradiated controls already at 25 Gy ($p<0.01$), while irradiated N2 [R] animals only produce a significantly lower brood size than unirradiated

controls at 75 Gy ($p<0.001$) (**Figure 2.5A**). Thus, a dose-response curve could be derived for both N2 [R] and N2 [S], with the response of N2 [S] being consistently more severe at any dose tested.

II.4.2 N2 [S] is not IR-sensitive at the L4 stage

I then wanted to test whether the IR-sensitivity was specific to L1 larvae, or whether it is present later in development as well. The initial rationale for using L1 animals was to screen for *htp-3* interactors which may modulate the IR response in the PGCs, which are the only germ cells present in the L1 larva at hatching. However, the serendipitously discovered divergence in the IR response of N2 [R] and N2 [S] may be general instead of life-stage delimited. To test this, I compared the post-IR brood size of L4 larvae, in which somatic development is largely completed with regard to cell division, and in which the germline comprises hundreds of cells undergoing meiosis and spermatogenesis (Sulston and Horvitz 1977, Kipreos 2005, Pazdernik and Schedl 2013). While IR treatment with 75 Gy reduced the number of progeny in N2 [R] from an unirradiated median of 321 (IQR 58) to 145 (IQR 39), this was not significantly different than N2 [S] ($p>0.05$), in which IR reduced the brood size from a median of 315.5 (IQR 45.25) to a median of 141.5 (IQR 47.75) (**Figure 2.5B**). The IR-sensitivity of N2 [S] is therefore developmentally modulated, and is no longer apparent by the L4 stage.

II.4.3 N2 [S] sensitivity is specific to ionizing radiation

I next wanted to test whether the sensitivity of N2 [S] is specific to IR as a genotoxic stressor, or whether it is reflective of a general sensitivity to genotoxic stress. To test this, I chose UV radiation, which predominantly causes inter-base crosslinks in neighbouring nucleotides (see section **I.1.5.2**), and ENU, an alkylating agent which can cause base transitions or transversions (see section **I.1.5.3**), as additional genotoxic stressors. Irradiation of L1 larvae with 50 J/m^2 significantly reduces the brood size of both N2 [R] and N2 [S] ($p<0.001$ for both genotypes), but there is no significant difference between the two genotypes at this dose ($p>0.05$) (**Figure 2.5D**). At 100 J/m^2, the brood size analysis was precluded by terminal arrest of irradiated larvae in both N2 backgrounds (**Figure 2.5D**). Thus, no differential sensitivity to UV radiation exists between N2 [R] and N2 [S]. Similarly, both N2 [R] and N2 [S] animals show a significant brood size reduction in adulthood after exposure to 5 mM or 10 mM ENU at the L1 stage ($p<0.01$ for 5mM ENU N2 [R]

against untreated controls; $p<0.001$ for 10mM ENU N2 [R] and N2 [S] at both ENU doses against untreated controls) (**Figure 2.5C**). However, no significant difference is apparent between the two N2 backgrounds after ENU treatment ($p>0.05$ at both ENU doses) (**Figure 2.5C**), showing that N2 [S] is not hypersensitive to DNA alkylation. In conclusion, the differential sensitivity of the N2 [S] background to IR seemed to be specific to that genotoxic stressor, raising the possibility of a defect in DSB repair, since DSBs are most efficiently induced by IR.

II.5 Reduced fertility in N2 [S] does not result from apparent germline defects

I then wanted to better characterize the physiological characteristics of the L1 IR response. In particular, why do the animals show a reduced brood size following irradiation? Is this a consequence of germline or somatic IR-induced defects? In *C. elegans,* a key indicator of DNA damage in the germline is increased embryonic lethality (the Emb phenotype) which results from chromosome segregation defects arising from unrepaired DSBs leading to aneuploidy, or mutagenic repair of survival-critical genes (Lemmens and Tijsterman 2011). Indeed, in L4 animals of both N2 [R] and N2 [S] lines, the post-IR (75 Gy) brood size decrease can be explained primarily by high embryonic lethality, 54% in N2 [R] and 51% in N2 [S], which is not significantly different ($p>0.05$) between the two lines (**Table 2.2**). By contrast, the embryonic lethality in post-L1-IR (75 Gy) N2 [S] animals is 0.67%, and is not significantly different than 0.85% Emb observed in unirradiated controls (**Table 2.2**). However, the post-L1-IR embryonic lethality in N2 [R] is 9.45%, significantly higher ($p<0.001$) than either the unirradiated N2 [R] (0.63%) or irradiated N2 [S] (0.67%) (**Table 2.2**). These observations on embryonic lethality suggest that a small part of the post-L1-IR brood size reduction in N2 [R] can be explained by unrepaired germline DNA damage or mutagenic repair, while no such damage is present in irradiated N2 [S] animals. Furthermore, neither N2 [R] nor N2 [S] animals exhibit gross cytological defects during the first meiotic prophase in the adult gonad following irradiation at the L1 stage (**Figure 2.6A** and **Figure 2.6B**). Marked by the germline chromatin marker HTP-3 and the DNA stain DAPI, the mitotic stem cells of irradiated animals appear the same as in unirradiated controls, DNA polarization in the transition zone occurs normally, no defects in the synapsed chromosome tracks are apparent, and diakinesis nuclei contain six DAPI-stained bodies, indicative of properly formed bivalents (**Figure 2.6A** and **Figure 2.6B**). Despite the lack of evidence for germline defects, the limiting

factor for post-IR fertility in both N2 [R] and N2 [S] appears to be oocyte production, as mating with unirradiated males, which increases the brood size of non-irradiated hermaphrodites of both N2 backgrounds by providing exogenous sperm, is not able to rescue the brood size of either irradiated N2 [R] or irradiated N2 [S] (**Figure 2.7**). Collectively, these results argue that DNA repair occurs normally in the N2 [S] germline following IR exposure. In N2 [R], the increased embryonic lethality supports either the possibility that DNA repair in the germline is impaired, that mutagenic repair occurs leading to an increased incidence of lethal mutations, or that the effects are germline non-autonomous. Because irradiated N2 [R] animals display no meiotic phenotypes which are a hallmark of unrepaired DSBs, and because the somatic development in N2 [R] is much less severely affected, I favor the mutagenic repair interpretation.

II.6 N2 [S] displays developmental delay and vulval defects

In addition to the brood size reduction, I noticed that several somatic defects occur much more prominently in N2 [S] than in N2 [R] following IR treatment at the L1 stage. The first of these is a developmental growth delay (the Gro phenotype). Three days after the L1 stage, all unirradiated N2 [R] and N2 [S] animals develop into adults, as do the vast majority of irradiated N2 [R] animals (94%) (**Figure 2.8A**). However, more than half (61%) of N2 [S] animals were still at the L4 or an earlier larval stage at this time point (**Figure 2.8A**), indicating a developmental delay. A small proportion of animals (2%) was found as thin, opaque, L3-like larvae (**Figure 2.8A**), similar to the L3-like larvae described previously in cNHEJ-defective mutants post-IR (Clejan, Boerckel et al. 2006). By contrast, only 6% of N2 [R] animals were still in a larval stage at the same time point (**Figure 2.8A**). Four days after the L1 stage, the great majority (89.5%) of the irradiated N2 [S] animals have developed into adults, but a large proportion (76%) exhibited vulval defects, including protruding vulva (the Pvl phenotype) and ruptured through vulva (the Rup phenotype) (**Figure 2.8B** and **Figure 2.8C**). Among irradiated N2 [R] animals, however, only 5% showed vulval defects (**Figure 2.8B**). Vulval dysgenesis is a hallmark of unrepaired DSBs or mutagenic repair in the soma (Lemmens and Tijsterman 2011). In addition to Pvl and Rup, another vulval phenotype that can occur as a consequence of IR treatment is vulvaless (the Vul phenotype), *ie*, a complete absence of the vulva (Weidhaas, Eisenmann et al. 2006). While I haven't directly scored for Vul, irradiated N2 [S] animals showed a much higher incidence than N2 [R] animals of the bag of

worms phenotype, which is the pathological occurrence of hatched larvae inside the body of the mother that results from an inability or impaired ability to lay eggs (Riddle, Blumenthal et al. 1997). One day after L4 stage, only 2% of IR-treated N2 [R] animals were bags of worms, compared to 25% of N2 [S] animals ($p<0.001$) (**Figure 2.8D**). This did not significantly change two days after the L4 stage ($p>0.05$ against the same genotype/treatment one day post-L4), with 3% and 31% bag of worm animals in N2 [R] and N2 [S], respectively (**Figure 2.8D**). Although efficient induction of vulval phenotypes in wild type animals requires very large doses of IR (100-400 Gy) (Weidhaas, Eisenmann et al. 2006), Gro and vulval phenotypes are apparent at much lower doses (60-120 Gy) in cNHEJ mutants, including *cku-70*, *cku-80*, and *lig-4* (Clejan, Boerckel et al. 2006). This raised the possibility that the N2 [S] background carried a cryptic mutation in one of these genes or possibly, in a novel factor involved in *C. elegans* cNHEJ (see **Chapter III**).

II.7 Ionizing radiation sensitivity is recessive

Identifying the mutation(s) causative of sensitivity in N2 [S] and N2 [R] required the full genomic sequence of both N2 genetic backgrounds, to reveal both the candidates for the causative locus (loci), as well as other mutations, such as restriction fragment length polymorphisms, which could be used in mapping. However, before I sent the genomic DNA for deep sequencing, I wanted to test whether IR-sensitivity is dominant or recessive to IR-resistance, so that the sequencing results could be better interpreted. I therefore compared the IR response of heterozygous N2 [R/S] animals derived from a cross between an N2 [R] male and an N2 [S] hermaphrodite to that of homozygous N2 [S/S] animals derived from a cross between two N2 [S] animals. When exposed to 50 Gy, N2 [R/S] heterozygotes had a significantly higher brood size (median 276 progeny, IQR 67.5) compared to N2 [S/S] animals (median 24 progeny, IQR 60.5; $p<0.001$) (**Figure 2.9A**), and exhibited a significantly lower penetrance of vulval phenotypes (60% in N2 [S/S] and 5% in N2 [R/S]; $p<0.001$) (**Figure 2.9B**), demonstrating that IR-resistance is dominant over IR-sensitivity, which is consistent with a genetic loss of function in the N2 [S] line.

II.8 Abrogation of Dicer helicase activity does not result in IR-sensitivity

Before proceeding with whole genome sequencing, I wanted to test the one candidate mutation associated with L1 IR-sensitivity that I already possessed: *dcr-1(mg375).* This *dcr-1* mutant was

the only genotype among those tested in the early stages of this investigation which didn't have a significantly different response than the IR-sensitive Zetka lab N2 (**Figure 2.1B**). To test whether the *dcr-1(mg375)* mutation can cause IR-sensitivity, I compared the post-IR brood size outcome of N2 [R], N2 [S], and *dcr-1(mg375)* with *dcr-1(vv121)*, a CRISPR-generated allele carrying the same G492R substitution as *dcr-1(mg375)* but in the N2 [S] background, and with *dcr-1(vv122)*, another CRISPR-generated carrying the G492R substitution in the N2 [R] background. While *dcr-1(mg375)* animals did not differ in post-IR brood size when compared to N2 [S] or *dcr-1(vv121)* [S], all three showed a markedly reduced brood size compared to *dcr-1(vv122)* [R] animals (**Figure 2.10**). This result demonstrated that the Dicer helicase G492R mutation does not impact IR-sensitivity, and that another mutation in the [S] background is responsible. Further progress now necessitated looking for other candidates, and I sent samples of N2 [R] and N2 [S] genomic DNA for deep sequencing at the Michael Smith Genome Sciences Centre in British Columbia.

II.9 Deep sequencing reveals IR-sensitizing candidate mutations in the N2 [S] genome

Deep sequencing results revealed 9,792 variants in the N2 [R] genome and 10,500 variants in the N2 [S] genome which were different than the reference genome (**Figure 2.11A**). Most of the variants in both genomes are not fixed in the population, but 32% of N2 [R] variants and 30% of N2 [S] variants are (**Figure 2.11A**), which reduced the list of possible causative variants. Furthermore, of the fixed variants, a large majority is shared between the two genomes, with only 16% being unique to N2 [R] and 17% unique to N2 [S] (**Figure 2.11B**), which suggests that N2 [R] and N2 [S] have diverged from each other only after their shared lineage diverged from the reference genome. The most important information that came out of the deep sequencing was the list of fixed variants specific to N2 [S], among which the IR-sensitivity causing variant was to be found. However, since this still represented 551 possible candidates, we partnered with the bioinformatics laboratory of Dr. Steven Jones at the British Columbia Cancer Agency to identify variants predicted to affect protein sequences, as we reasoned that a protein loss-of-function would be the most likely cause of IR-sensitivity. Dr. Jones' group identified 15 single nucleotide mutations which directly affect protein sequence, including 12 which result in amino acid substitutions, 2 which create a premature stop codon, and 1 which results in the abolition of a splice site (**Table 2.3**). As a first line strategy to map the causative mutation, I undertook the

sequencing of these 15 mutations in phenotypically sensitive N2 hybrid lines derived from singled F2 animals from an N2 [R] X N [S] cross (**Figure 5.2A**). I created 21 such sensitive lines from a total of 80 F2 animals, a 3:1 ratio consistent with a single causative locus, or multiple linked loci. Only two mutations, the amino acid substitutions in the inverted formin *inft-2* and in the putative hexosyl transferase *F10D2.12* (**Table 2.3**), were present in every sequencing reaction of every hybrid strain (**Figure 2.12**). These two mutations are both located on chromosome V, and are strongly linked (~4.2 cM apart). Thus, two possibilities presented themselves: 1) either the *inft-2* mutation, or the *F10D2.12* mutation, or both were truly causative of the IR-sensitivity in N2 [S]; or 2) the causative mutation is located somewhere on chromosome V, closely linked to *inft-2* and *F10D2.12*, but is not in a protein coding gene, affecting either an RNA gene or a DNA regulatory element. To test the first of these two possibilities, I used CRISPR mutagenesis to create the N2 [S]-like mutations, *inft-2(vv135[S239Y])* and *F10D2.12 (vv136[P325L])* in N2 [R], hypothesizing that this would result in IR-sensitivity if either of these two loci, or both, are causative. However, the *inft-2(vv135) F10D2.12(vv136)* double mutant proved to be IR-resistant, with a median post-IR brood size of 213 progeny (IQR 160.75), compared with a median of 161 progeny (IQR 184) in N2 [R] and 0 median progeny (IQR 9.25) in N2 [S] (**Figure 2.13**). I therefore concluded that the causative mutation is merely linked to *inft-2* and *F10D2.12* loci, and may not be in a protein coding gene.

II.10 Loss of cNHEJ sensitizes N2 [R] to ionizing radiation

Despite this conclusion, the similarity of post-IR N2 [S] phenotypes to the published post-IR phenotypes of cNHEJ pathway mutants (Clejan, Boerckel et al. 2006) prompted me to investigate the possibility that IR-sensitivity in N2 [S] is caused by a loss of cNHEJ activity. As an initial test of this possibility, I examined the post-IR phenotypes of *lig-4(ok716)*, a published deletion mutant of the terminal cNHEJ ligase *lig-4* (Clejan, Boerckel et al. 2006), as well as *lig-4(vv134[R18STOP])*, an allele converting codon 18 from an arginine to a stop codon, which I generated using CRISPR in the N2 [R] background. Like N2 [S], whose post-IR brood size is reduced to a median of 3 progeny (IQR 19), the post-IR brood size of both *lig-4* mutants is reduced, with a median of 0 progeny (IQR 0) in *lig-4(ok716)* and a median of 0 progeny (IQR 7.5) in *lig-4(vv134)* (**Figure 2.14A**). All three are significantly lower than the median post-IR brood size of N2 [R] of 126 progeny (IQR

220.25; p<0.001 for all three comparisons) (**Figure 2.14A**). Additionally, a significantly higher incidence of vulval (Pvl and Rup) phenotypes is apparent in N2 [S] (63%), *lig-4(ok716)* (66%), and *lig-4(vv134)* (75%), compared to N2 [R] (4%; p<0.001 for all three comparisons) (**Figure 2.14B**). Thus, the post-IR phenotypes of N2 [S] are not only present in a previously described null mutant of *lig-4*, but the normally resistant N2 [R] genetic background can be made IR-sensitive by inactivating *lig-4*. While this finding did not represent proof that the sensitivity of the N2 [S] strain is caused by a loss of cNHEJ activity (for direct evidence of that, see **Chapter III**), it did strengthen this hypothesis. Based on the deep sequencing data, the coding sequences of the three known *C. elegans* cNHEJ genes, *cku-70, cku-80,* and *lig-4*, did not contain any variants which might affect the protein function. This raised the possibility that a non-coding mutation somewhere on chromosome V may abolish the expression of one or more of the cNHEJ genes, which are all located on chromosome III. To test this possibility, I performed RT-PCR for each of the three known cNHEJ genes on cDNA converted from total RNA extracted from N2 [R] and N2 [S] animals. RT-PCR products for all of the cNHEJ gene transcripts were present in both N2 [R] and N2 [S] cDNAs (**Figure 2.15**), and I thus concluded that the N2 [S] mutation does not result in a loss of *cku-70*, *cku-80,* or *lig-4* transcription.

This left three possibilities open: 1) that the L1 IR-sensitivity is due to a disruption in the translational regulation of the known cNHEJ factors, or a loss of IR-dependent expression control; 2) that is not the result of a cNHEJ deficiency and is rather caused by a failure of another DSB repair pathway with a similar spectrum of post-IR phenotypes; and 3) that there is a novel cNHEJ factor somewhere on chromosome V. The first step toward testing these possibilities required the identification and characterization of the causative mutation in the N2 [S] genome, which is described in the following chapter.

Figures and Tables

Figure 2.1

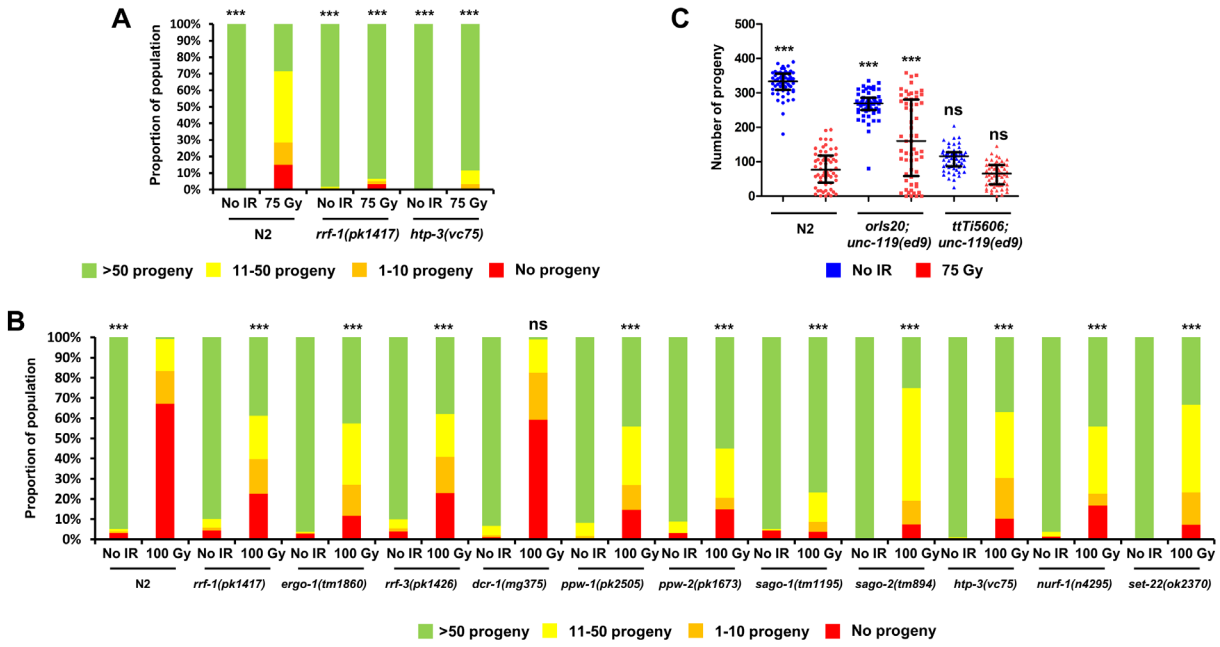


Figure 2.1 - The Bristol wild type strain N2 is sensitive to ionizing radiation

(A) Quantification of post-IR fertility by brood size category in N2, *rrf-1(pk1417),* and *htp-3(vc75).* When irradiated with 75 Gy of ionizing radiation at the L1 stage, wild type N2 animals show a significantly reduced brood size compared to unirradiated controls ($p<0.001$), with a large proportion of the irradiated population producing less than 50 progeny. Irradiated *rrf-1(pk1417)* and *htp-3(vc75)* mutants, however, exhibited a much more modest brood size reduction compared with unirradiated controls of the same genotype, which was not significant. All statistical comparisons shown in the figure are to N2 75 Gy (Kruskal-Wallis test and Dunn's post-hoc tests). Sample size (n) is 59 for unirradiated N2 and 60 for irradiated N2, both *rrf-1(pk1417)* groups, and both *htp-3(vc75)* groups.

Figure 2.1 The Bristol wild type strain N2 is sensitive to ionizing radiation (*continued)*

(B) Quantification of post-IR fertility by brood size category in N2, endo siRNA biogenesis mutants, and three chromatin-associated factors. After 100 Gy of IR at the L1 stage, the brood size of N2 is significantly reduced ($p<0.001$ in all cases) compared to the post-IR brood size of endo siRNA biogenesis pathway mutants *rrf-3(pk1426), ergo-1(tm1860), rrf-1(pk1417)*, the secondary Argonaute mutants *ppw-1(pk2505), ppw-2(pk1673), sago-1(tm1195),* and *sago-2(tm894)*, as well as several mutants of chromatin associated proteins, including *htp-3(vc75), nurf-1(n4295),* and *set-22(ok2370).* However, there was no significant difference in the post-IR brood size of N2 and *dcr-1(mg375),* a helicase dead mutant of the small RNA biogenesis protein Dicer. All statistical comparisons shown in the figure are to N2 100 Gy (Kruskal-Wallis test and Dunn's post-hoc tests). Sample size (n) is 1083 for unirradiated N2, 1067 for irradiated N2, 209 for both *rrf-1(pk1417)* groups, 209 and 204 respectively for unirradiated and irradiated *ergo-1(tm1860)*, 202 and 201 respectively for unirradiated and irradiated *rrf-3(pk14136)*, 198 and 194 respectively for unirradiated and irradiated *dcr-1(mg375)*, 197 and 193 respectively for unirradiated and irradiated *ppw-1(pk2505)*, 194 and 189 respectively for unirradiated and irradiated *ppw-2(pk1673)*, 209 and 210 respectively for unirradiated and irradiated *sago-1(tm1198)*, 70 and 68 respectively for unirradiated and irradiated *sago-2(ok894)*, 207 and 205 respectively for unirradiated and irradiated *htp-3(vc75)*, 209 for both *nurf-1(n4295)* groups, and 68 and 69 respectively for unirradiated and irradiated *set-22(ok2370).*

(C) Quantification of total brood size in N2, an otherwise wild-type transgenic strain, and an *unc-9* mutant. The post-IR brood size of N2 is significantly lower ($p<0.001$) than the post-IR brood size of *orIs20* [*pie-1p::GFP::MOE unc-119(+)*]; *unc-119(ed9)*, a MosSCI transgenic strain where the only mutation is rescued by an integrated construct. Although the N2 post-IR brood size is not significantly different ($p>0.05$) than the post-IR brood size of the *ttTi5606; unc-119(ed9)* strain, which is rescued for Unc by the *orIs20* transgene, it is also not different ($p>0.05$) from the unirradiated brood size of *ttTi5606; unc-119(ed9).* All statistical comparisons shown in the figure are to N2 75 Gy (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 60 for all groups.

IR = ionizing radiation
Gy = Gray (unit)

orIs20 = [*pie-1p::GFP::MOE unc-119(+)*] insertion
ttTi5606 = MosSCI insertion site
ns = not significant ($p>0.05$); *** = $p<0.001$

Figure 2.2

A

Gene	Class/Function	Effect on brood size
lin-40	Chromatin remodeling (NuRD)	-
nurf-1	Chromatin remodeling (NURF)	-
pbrm-1	Chromatin remodeling (PB/Rsc)	+
set-5	Chromatin remodeling (Polycomb/SET)	-
set-15	Chromatin remodeling (Polycomb/SET)	-
set-20	Chromatin remodeling (Polycomb/SET)	+
set-22	Chromatin remodeling (Polycomb/SET)	+
swsn-3	Chromatin remodeling (SWI/SNF2)	+
swsn-4	Chromatin remodeling (SWI/SNF2)	+
swsn-7	Chromatin remodeling (SWI/SNF2)	-
dpy-30	H3K4 methyltransferase	-
mys-1	Histone acetyltransferase; SynMuv C	-
trr-1	TRAAP kinase in a HAT complex; SynMuv C	-
gei-8	NCoR/SMRT corepressor orthologue	-
mlh-1	Mismatch repair	variable
brc-1	DSB repair; Ubiquitylation (E3)	+
ubc-1	Ubiquitylation (E2)	-
T13F2.2	Transcriptional co-activator; PC4 superfamily	+
Y113G7B.14	RNA Pol II transcription factor	+
mab-10	NAB family transcription factor	+
cin-4	Topoisomerase II - like	-
top-2	Topoisomerase II - like	-
cpar-1	CENP-A homologue	-
hcp-3	CENP-A homologue	-
rskn-2	Ribosomal S6 protein kinase	+
tpa-1	Protein kinase C	+
cdc-37	Hsp90 co-chaperone	-
lst-3	DNA binding; Hemidesmosome	variable
sop-2	Polycomb-related; miRNA antagonist	+
zfp-1	Zinc finger protein required for RNAi	+
nlp-17	Neuropeptide	variable
Y18D10A.1	Uncharacterized	-
CD4.8	Uncharacterized; SPK/SET domain	-

B

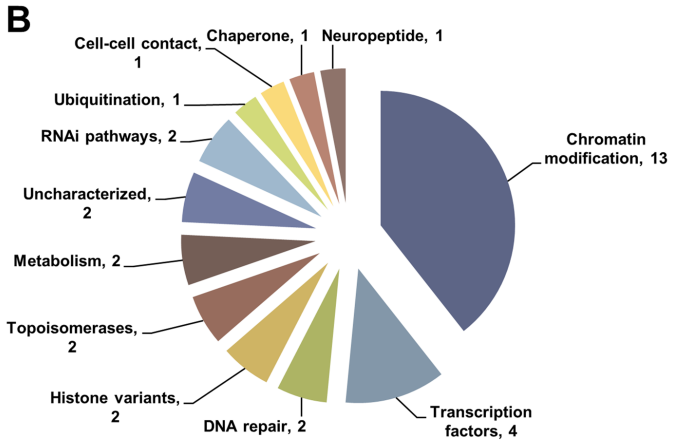


Figure 2.2 - Validated hits from the RNAi screen

(A) List of the 33 thrice-validated positive hits from the screen for chromatin-associated factors which modulate IR-sensitivity in L1 animals. Of the hits, 17 had a negative impact on post-IR brood size and 13 had a positive impact on post-IR brood size. Three candidates, *mlh-1*, *lst-3*, and *nlp-17*, always influenced the brood size, but the effect was variable. The Class/Function column is shaded with the respective functional class color as in **(B)**.

(B) Pie chart showing the categorization of the 33 positive hits by known or inferred cellular function of the confirmed candidate.

+ = increased brood size

- = decreased brood size

Figure 2.3

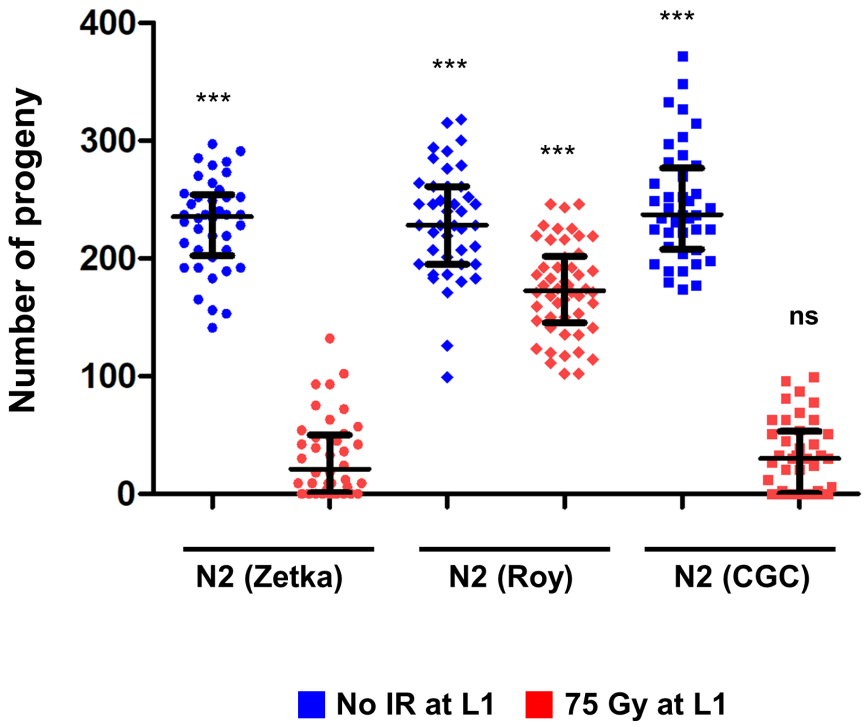


Figure 2.3 - CGC N2 is IR-sensitive, while the Roy lab N2 is IR-resistant

Quantification of brood size in N2 lines from the Zetka lab, the Roy lab, and the CGC. Following IR treatment at the L1 stage, the N2 line from the Caenorhabditis Genetics Center exhibits the same reduction in brood size as the N2 line from the Zetka lab ($p>0.05$). However, the N2 from the neighbouring laboratory of Dr. Richard Roy at McGill University is much more resistant, exhibiting a significantly higher brood size than either the Zetka lab N2 or the CGC N2 ($p<0.001$ for both comparisons), although still lower than in unirradiated controls ($p<0.01$).

Figure 2.3 - CGC N2 is IR-sensitive, while the Roy lab N2 is IR-resistant (*continued*)

All statistical comparisons shown in the figure are to N2 (Zetka) 75 Gy (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 40 for both N2 (Zetka) groups, 42 and 50 respectively for unirradiated and irradiated N2 (Roy), and 40 for both N2 (CGC) groups.

IR = ionizing radiation
Gy = Gray (unit)
CGC = Caenorhabditis Genetics Center
N2 (Zetka) = IR-sensitive N2 strain originating from the laboratory of Monique Zetka (McGill University)
N2 (Roy) = IR-resistant N2 strain originating from the laboratory of Richard Roy (McGill University)

ns = not significant ($p>0.05$); *** = $p<0.001$

Table 2.1

N2 lines		
Strain name	**Origin**	**L1 IR response**
N2	Zetka laboratory	Sensitive
N2	CGC	Sensitive
N2	Roy laboratory	Resistant
N2	Hekimi laboratory	Resistant
N2	Andersen laboratory	Resistant
N2	Zhen laboratory	Resistant
N2	NBRP	Resistant
Other *C. elegans* wild type isolates		
Strain name	**Origin**	**L1 IR response**
QX1211	San Francisco, California	Resistant
JU775	Lisbon, Portugal	Resistant
CB4856	Oahu, Hawaii	Resistant
DL238	Hawaii, Hawaii	Resistant
MY16	Mecklenbeck, Germany	Resistant
MY23	Roxel, Germany	Resistant
EG4724	Amares, Portugal	Resistant
JU258	Ribeiro Frio, Madeira	Resistant
JU1491	Le Blanc, France	Resistant
LKC34	Madagascar	Resistant
JU1088	Kakegawa, Japan	Resistant
ED3073	Limuru, Kenya	Resistant
DL226	Corvallis, Oregon	Resistant
AB4	Adelaide, Australia	Resistant
JU1171	Concepcion, Chile	Resistant
AB1	Adelaide, Australia	Resistant
JU1652	Montevideo, Uruguay	Resistant
JU1896	Athens, Greece	Resistant
***C. briggsae* strains**		
Strain name	**Origin**	**L1 IR response**
VT847	Hawaii, Hawaii	Resistant
PB800	Dayton, Ohio	Resistant

Table 2.1 - List of wild type *Caenorhabditis* strains tested for L1 IR-sensitivity

Seven N2 lines from different sources, 18 non-N2 *C. elegans* wild type isolates from all continents and diverse geographic latitudes, and two *C. briggsae* wild type isolates have been tested for the L1 response to IR. Only the N2s from the Zetka lab and the CGC are sensitive.

Zetka laboratory = laboratory of Dr. Monique Zetka, McGill University, Montreal, Canada
CGC = Caenorhabditis Genetics Center, University of Minnesota Twin Cities, United States of America
Roy laboratory = laboratory of Dr. Richard Roy, McGill University, Montreal, Canada
Hekimi laboratory = laboratory of Dr. Siegfried Hekimi, McGill University, Montreal, Canada
Andersen laboratory = laboratory of Dr. Eric Andersen, Northwestern University, Chicago, United States of America
Zhen laboratory = laboratory of Dr. Mei Zhen, Lunenfeld-Tanenbaum Research Institute, Toronto, Canada
NBRP = National Bioresource Project, Shinjuku-ku, Tokyo, Japan

IR = ionizing radiation

Figure 2.4


A
Finland
Iceland
Sweden
Norway
Russia
Canada
Poland
Germany
Ukraine
Kazakhstan
Mongolia
France
Spain
Italy
Turkey
China
South Korea
United States
North Atlantic Ocean
Iraq
Iran
Afghanistan
Pakistan
Algeria
Libya
Egypt
Saudi Arabia
India
Thailand
Mexico
Mali
Niger
Sudan
Chad
Nigeria
Venezuela
Colombia
Kenya
DR Congo
Tanzania
Indonesia
Papua New Guinea
Brazil
Peru
Angola
Bolivia
Namibia
Botswana
Madagascar
Indian Ocean
South Pacific Ocean
South Atlantic Ocean
South Africa
Argentina
IR resistant
IR sensitive

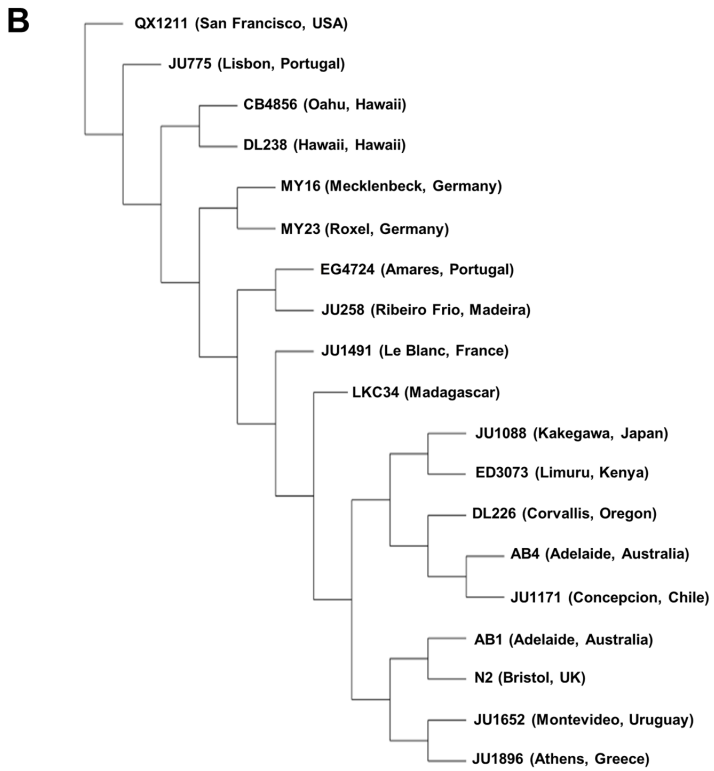

B
QX1211 (San Francisco, USA)
JU775 (Lisbon, Portugal)
CB4856 (Oahu, Hawaii)
DL238 (Hawaii, Hawaii)
MY16 (Mecklenbeck, Germany)
MY23 (Roxel, Germany)
EG4724 (Amares, Portugal)
JU258 (Ribeiro Frio, Madeira)
JU1491 (Le Blanc, France)
LKC34 (Madagascar)
JU1088 (Kakegawa, Japan)
ED3073 (Limuru, Kenya)
DL226 (Corvallis, Oregon)
AB4 (Adelaide, Australia)
JU1171 (Concepcion, Chile)
AB1 (Adelaide, Australia)
N2 (Bristol, UK)
JU1652 (Montevideo, Uruguay)
JU1896 (Athens, Greece)

Figure 2.4 - Locations of *C. elegans* wild type isolates tested for IR-sensitivity

(A) World map showing the geographic positions of the original locations where N2 and the 18 non-N2 *C. elegans* wild type strains tested for the L1 IR response (**Table 2.1**) have been isolated. The strains have been selected to cover all the major landmasses on Earth, though the Northern hemisphere and Europe in particular are overrepresented because of sampling bias.

(B) A phylogenetic tree of the 19 *C. elegans* isotypes tested for the L1 IR response, based on the more extensive neighbour-joining tree constructed in (Andersen, Gerke et al. 2012) on the basis of 40,857 single nucleotide polymorphisms. Distances do not accurately represent degrees of divergence.

IR = ionizing radiation

Figure 2.5

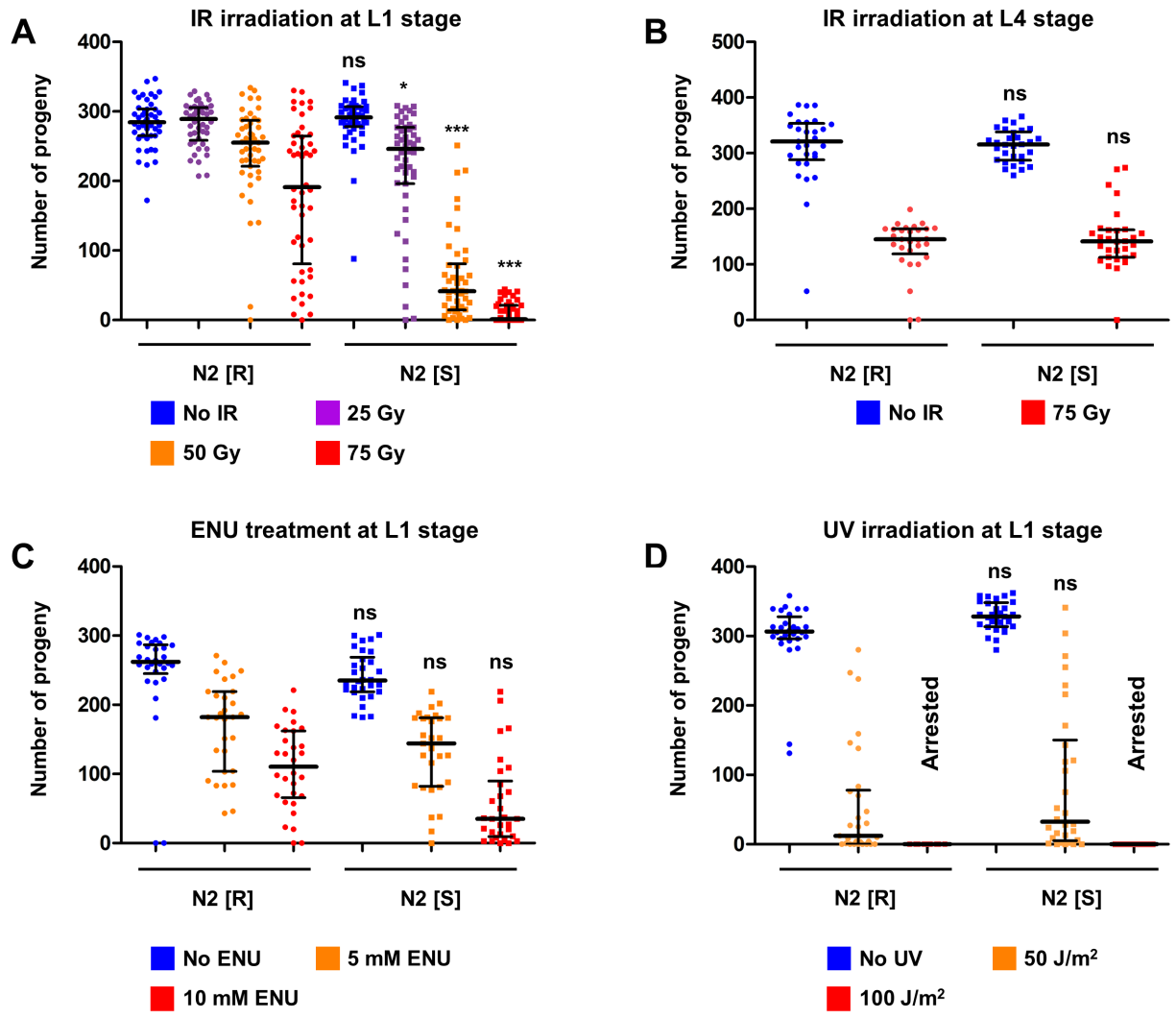


Figure 2.5 - N2 [R] and N2 [S] show a dose-dependent, L1-specific difference in brood size following treatment with ionizing radiation, but not ethyl-nitrosourea or ultraviolet radiation

(A) Multi-dose brood size quantification of N2 [R] and N2 [S] post-L1 IR treatment. The IR-sensitive N2 [S] line shows a significantly reduced brood size compared to unirradiated animals at 25 Gy ($p<0.01$), 50 Gy ($p<0.001$), and 75 Gy ($p<0.001$). Irradiated N2 [R] animals show a significantly reduced brood size compared to unirradiated controls only at 75 Gy ($p<0.001$). At every tested IR dose, N2 [S] is significantly more severely affected than N2 [R]. All statistical comparisons shown in the figure are to N2 [R] at the equivalent IR dose (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and

Figure 2.5 - N2 [R] and N2 [S] show a dose-dependent, L1-specific difference in brood size following treatment with ionizing radiation, but not ethyl-nitrosourea or ultraviolet radiation (*continued)*

(A, *continued*) interquartile range. Sample size (n) is 46 for unirradiated N2 [R], 45 for N2 [R] 25 Gy, 47 for N2 [R] 50 Gy, 48 for N2 [R] 75 Gy, 44 for unirradiated N2 [S], 48 for N2 [S] 25 Gy, 46 for N2 [S] 50 Gy, and 49 for N2 [S] 75 Gy.

(B) Brood size quantification of N2 [R] and N2 [S] post-L4 IR treatment. Both N2 [S] and N2 [R] animals show a reduced brood size in adulthood ($p<0.001$ vs unirradiated controls), but there is no significant difference between the two backgrounds, suggesting that the IR-sensitivity of N2 [S] is specific to the L1 stage. All statistical comparisons shown in the figure are to N2 [R] at the equivalent IR dose (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 29 for both N2 [R] groups, and 30 for both N2 [S] groups.

(C) Brood size quantification of N2 [R] and N2 [S] post-L1 ENU treatment. Both N2 [S] and N2 [R] animals show a dose-dependent reduction in brood size ($p<0.01$ for 5mM ENU and $p<0.001$ for 10mM ENU for N2 [R] versus untreated control; $p<0.001$ for both 5mM and 10mM ENU for N2 [S] versus untreated control), but there is no significant difference between the two backgrounds at the doses tested. All statistical comparisons shown in the figure are to N2 [R] at the equivalent ENU dose (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 29 for N2 [S] 5mM ENU, and 30 for all other groups.

(D) Brood size quantification of N2 [R] and N2 [S] post-L1 UV treatment. Both N2 [S] and N2 [R] animals show a reduction in adult brood size ($p<0.001$ versus unirradiated controls for both genotypes), but there is no significant difference between the two backgrounds. At the higher of the two tested doses, 100 J/m^2, animals of both backgrounds exhibit terminal larval arrest and do not produce progeny. All statistical comparisons shown in the figure are to N2 [R] at the equivalent ENU dose (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 28 for unirradiated N2 [R], 29 for unirradiated N2 [S], and 30 for all other groups.

IR = ionizing radiation; UV = ultraviolet radiation; ENU = N-ethyl-N-nitrosourea
Gy = Gray (unit); J = Joule (unit); m^2 = meter squared; mM = millimolar
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

ns = not significant ($p>0.05$); * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$

Figure 2.6

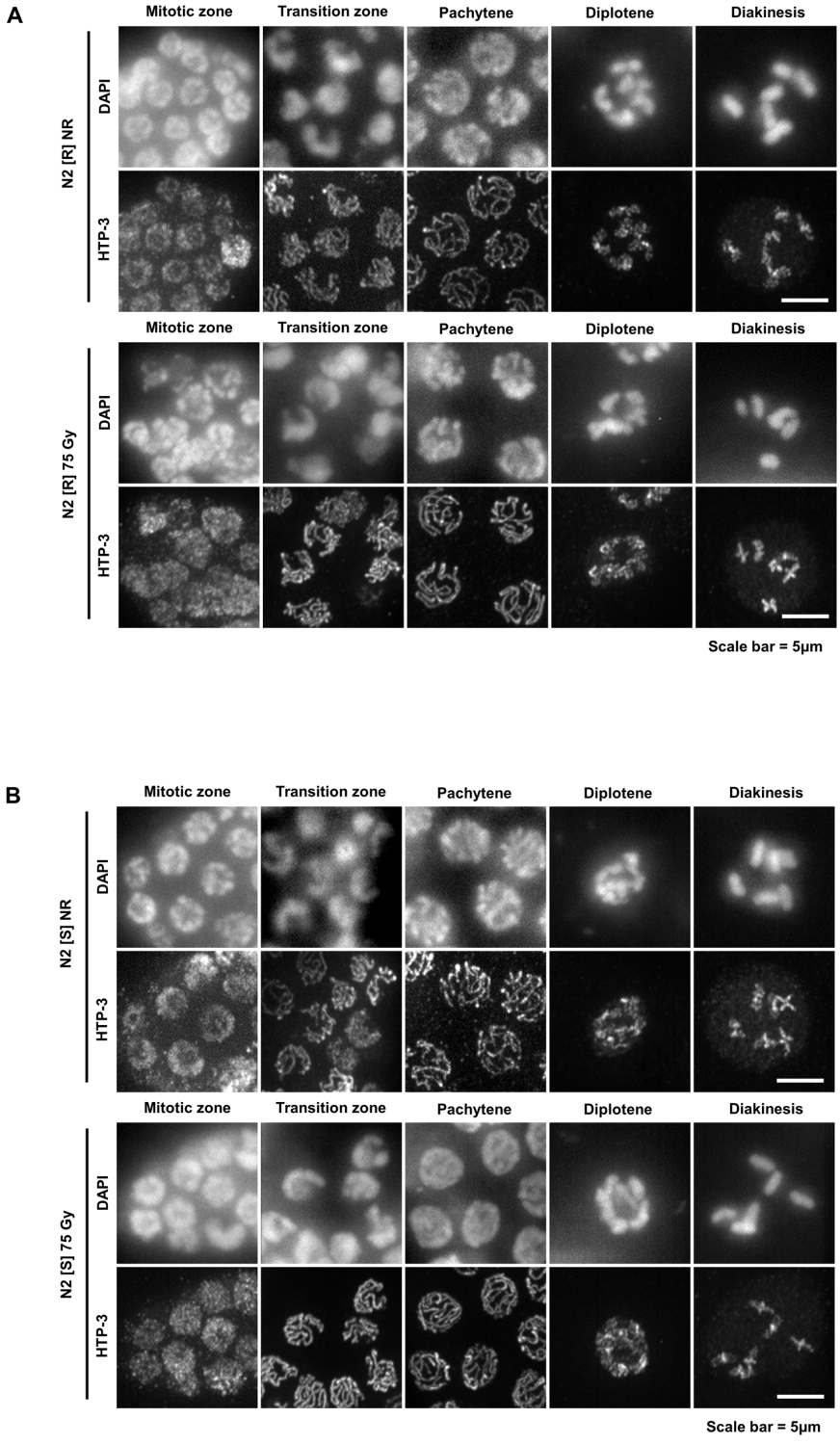

A
Mitotic zone
Transition zone
Pachytene
Diplotene
Diakinesis
N2 [R] NR
DAPI
HTP-3
N2 [R] 75 Gy
Scale bar = 5μm
B
N2 [S] NR
N2 [S] 75 Gy

Figure 2.6 - No gross cytological defects are apparent in either N2 [R] or N2 [S] adult germlines following IR treatment at the L1 stage

(A) Representative images of DAPI and anti-HTP-3 antibody staining of germline stem cells (mitotic zone) and meiotic prophase cells in leptotene/zygotene (transition zone), pachytene, diplotene, and diakinesis in unirradiated adult N2 [R] animals and adult N2 [R] animals irradiated with 75 Gy of IR at the L1 stage. No gross defects are visible in the mitotic cell morphology, chromosome axis formation and DNA polarization in the transition zone, axis morphology in pachytene, DNA condensation in diplotene, or the number of DAPI-staining bodies in diakinesis (6 bivalents).

(B) Representative images of the same adult germline regions as in (A), but for irradiated and control N2 [S] animals. As in N2 [R], no cytological defects are apparent.

IR = ionizing radiation
Gy = Gray (unit)
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

Table 2.2

	Control L1 unhatched/total (%)	Irradiated L1 (75 Gy) unhatched/total (%)
N2 [R]	7/1111 (0.63%)^{ns}	55/582 (9.45%)***
N2 [S]	11/1293 (0.85%)	4/593 (0.67%)
	Control L4 unhatched/total (%)	**Irradiated L4 (75 Gy)** unhatched/total (%)
N2 [R]	7/922 (0.76%)^{ns}	617/1139 (54.17%)^{ns}
N2 [S]	6/948 (0.63%)	665/1297 (51.27%)

Table 2.2 - Embryonic lethality is increased in N2 [R] after IR treatment at the L1 stage, and in both N2 [R] and N2 [S] after IR treatment at the L4 stage

The progeny of N2 [S] mothers irradiated with 75 Gy at the L1 stage do not exhibit an increased embryonic lethality compared to the progeny of unirradiated mothers, strongly suggesting that the reduced brood size of IR-treated N2 [S] animals is not caused by impaired or mutagenic DNA repair in the germline. The progeny of N2 [R] mothers irradiated with 75 Gy show a significantly higher embryonic lethality than either the progeny of unirradiated N2 [R] mothers or the progeny of N2 [S] mothers irradiated at the same dose ($p<0.001$ for both comparisons). Irradiation with 75 Gy at the L4 stage induces much higher embryonic lethality in the progeny in both N2 genetic backgrounds, in comparison to unirradiated controls ($p<0.001$ for both N2 [R] and N2 [S]). However, there is no significant difference in the level of embryonic lethality among the progeny of irradiated N2 [R] and N2 [S]. All statistical comparisons shown in the figure are to N2 [S] at the equivalent IR dose (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.01$).

IR = ionizing radiation
Gy = Gray (unit)
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

ns = not significant ($p>0.01$); *** = $p<0.001$

Figure 2.7

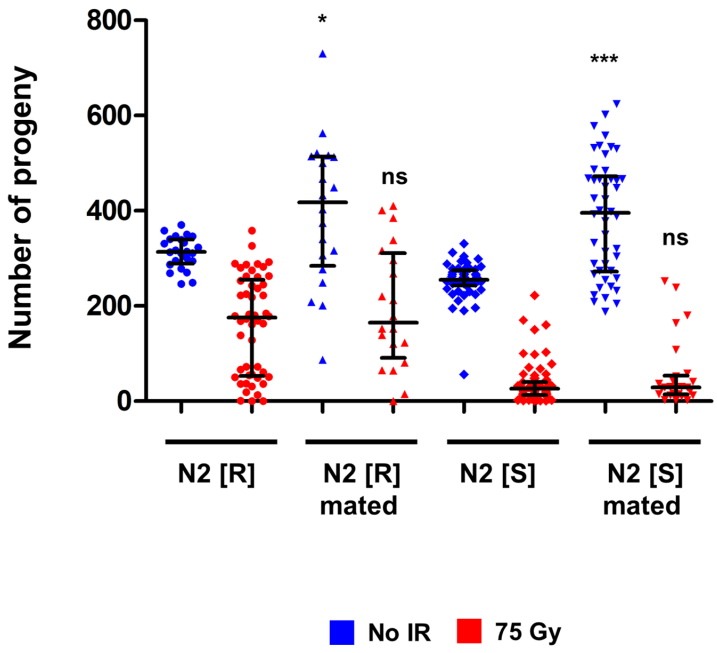

Number of progeny
800
600
400
200
0
*
ns
***
ns
N2 [R]
N2 [R] mated
N2 [S]
N2 [S] mated
No IR
75 Gy

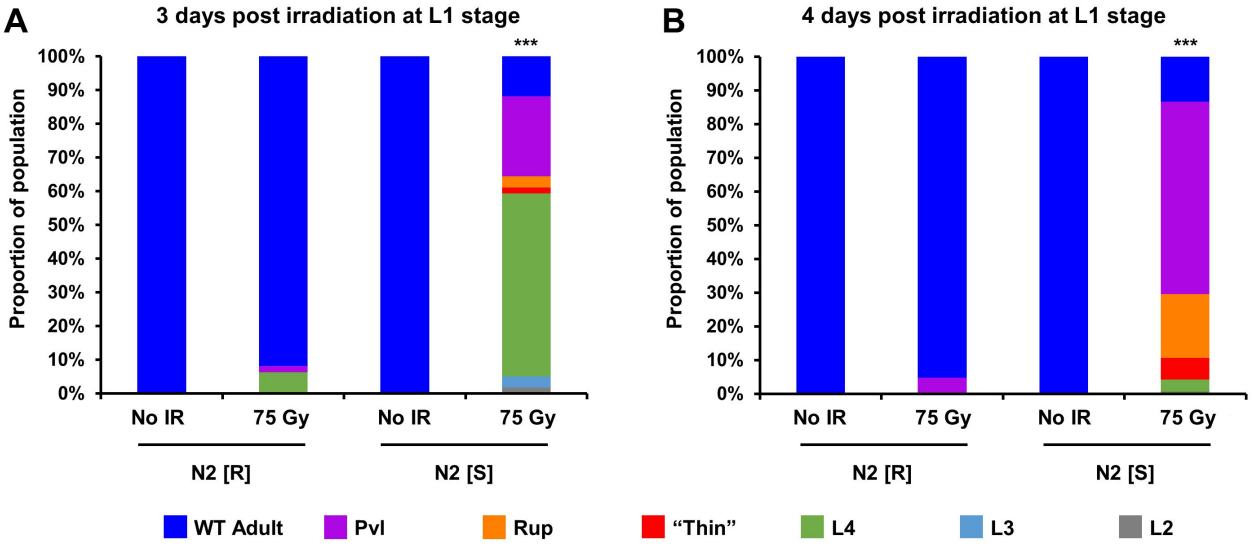
Figure 2.7 - Mating does not rescue post-IR brood size of N2 [R] or N2 [S]

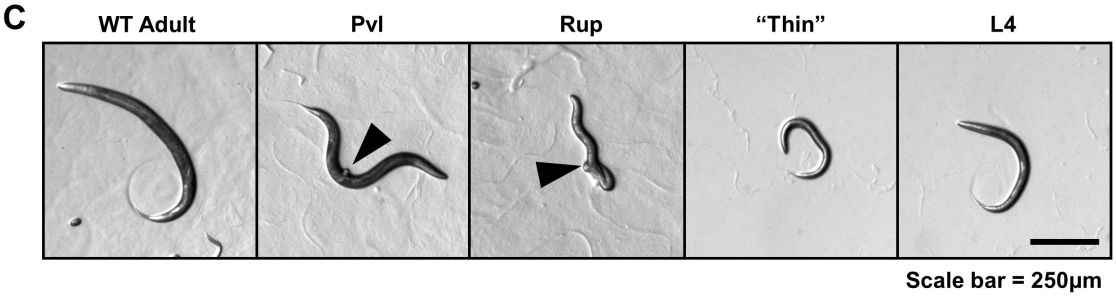
Quantification of total brood size in mated and unmated, irradiated and control N2 [R] and N2 [S] animals. The introduction of exogenous sperm via mating significantly increases the brood size of unirradiated N2 [R] hermaphrodites ($p<0.05$) and N2 [S] hermaphrodites ($p<0.001$). Mating of unirradiated males with irradiated N2 [R] hermaphrodites or irradiated N2 [S] hermaphrodites does not increase the number of progeny in either group, demonstrating that post-IR brood size is not sperm limited. All statistical comparisons shown in the figure are to the non-mated group of the corresponding genotype and treatment (Kruskal-Wallis test and Dunn's post-hoc tests). Error bars represent the median and interquartile range. Sample size (n) is 24 for non-mated unirradiated N2 [R], 50 for non-mated irradiated N2 [R], 20 for mated unirradiated and mated irradiated N2 [R], 52 for non-mated unirradiated N2 [S], 59 for non-mated irradiated N2 [S], 46 for mated unirradiated N2 [S], and 26 for mated irradiated N2 [S].

IR = ionizing radiation
Gy = Gray (unit)
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

ns = not significant ($p>0.01$); * = $p<0.05$, *** = $p<0.001$

Figure 2.8





D

	No IR at L1 bag of worms/total (%)	75 Gy at L1 bag of worms/total (%)
N2 [R] 1 days post-L4	0/100 (0%)	2/100 (2%)
N2 [S] 1 days post-L4	0/100 (0%)	25/100 (25%)***
N2 [R] 2 days post-L4	0/100 (0%)	3/100 (3%)
N2 [S] 2 days post-L4	0/100 (0%)	31/100 (31%)***

Figure 2.8 - N2 [S] displays several distinct somatic phenotypes post-IR

(A) Quantification of growth delay and vulval phenotypes three days after IR treatment at the L1 stage. When irradiated at 75 Gy of IR, N2 [S] animals show a developmental delay, as a large proportion is still in the L4 stage when all unirradiated controls and almost all irradiated N2 [R] animals have developed into adults. A small proportion of irradiated N2 [S] animals develops into thin, whitish larvae approximately the size of L3 larvae. The majority of irradiated N2 [S] animals that does develop into adults by this stage exhibits vulval phenotypes, most prominently protruding vulva (Pvl) and ruptured through vulva (Rup). The statistical comparison shown in the figure is to irradiated N2 [R] (Chi-squared test). Sample size (n) is 148 for unirradiated N2 [R], 111 for irradiated N2 [R], 141 for unirradiated N2 [S], and 118 for irradiated N2 [S].

(B) Quantification of the same phenotypes as in **(A)**, four days after treatment at L1 stage. Four days after irradiation with 75 Gy of IR, almost all N2 [S] animals develop into adults, but exhibit a high incidence of Pvl and Rup phenotypes, as well as occasional thin, whitish larvae. The statistical comparison shown in the figure is to irradiated N2 [R] (Chi-squared test). Sample size (n) is 109 for unirradiated N2 [R], 148 for irradiated N2 [R], 134 for unirradiated N2 [S], and 142 for irradiated N2 [S].

(C) Representative images of phenotypes quantified in (A) and (B). Black arrows point to the protruding vulva in the "Pvl" panel and the burst vulva and partial extrusion of internal organs in the "Rup" panel.

(D) Quantification of the bag of worms phenotype in irradiated and control animals. While a small proportion of N2 [R] animals develops a bag of worms phenotype, it is significantly more common in N2 [S] animals. The statistical comparisons shown in the figure are to irradiated N2 [R] at the equivalent time points (Chi-squared test).

IR = ionizing radiation
Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L4/L3/L2 = larva of the L4/L3/L2 stage
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

ns = not significant ($p>0.01$); * = $p<0.05$, *** = $p<0.001$

Figure 2.9

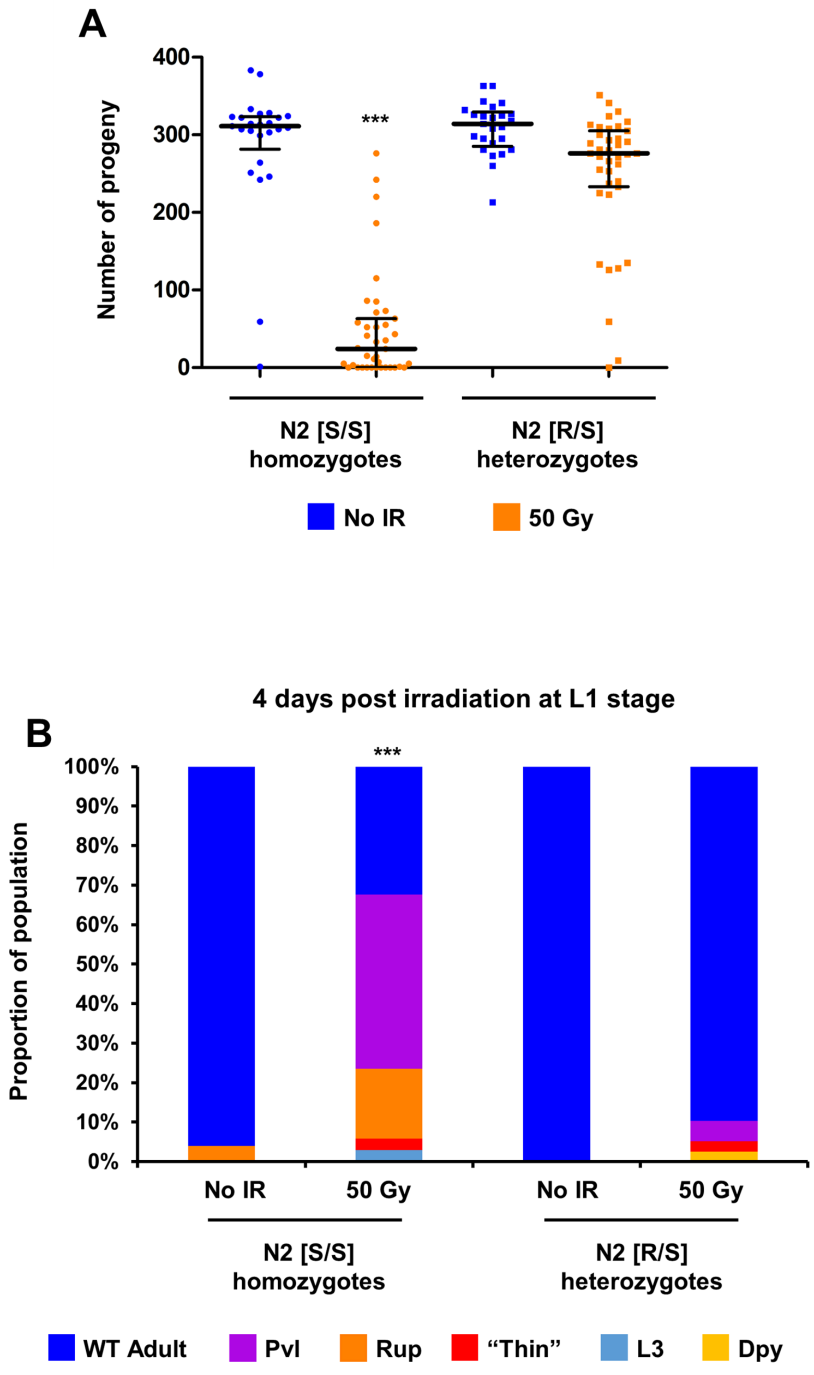

A
Number of progeny
400
300
200
100
0
***
N2 [S/S]
homozygotes
N2 [R/S]
heterozygotes
No IR
50 Gy
4 days post irradiation at L1 stage
B
Proportion of population
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
***
No IR
50 Gy
No IR
50 Gy
N2 [S/S]
homozygotes
N2 [R/S]
heterozygotes
WT Adult
Pvl
Rup
"Thin"
L3
Dpy

Figure 2.9 - IR-resistance is dominant to IR-sensitivity

(A) Total brood size quantification in N2 [S/S] homozygotes and N2 [R/S] heterozygotes after 50 Gy of IR at the L1 stage. N2 [S/S] show a significantly lower post-IR brood size than N2 [R/S] heterozygotes, demonstrating that IR-sensitivity is a recessive trait. The statistical comparison shown in the figure is to irradiated N2 [R/S] heterozygotes (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 25 for both unirradiated groups and 39 for both irradiated groups.

(B) Quantification of somatic phenotypes four days after IR treatment at the L1 stage. Irradiated N2 [S/S] homozygotes show a significantly higher incidence of Pvl and Rup phenotypes than N2 [R/S] heterozygotes, corroborating the conclusion that IR-sensitivity is recessive to IR-resistance. The statistical comparison shown in the figure is to irradiated N2 [R] (Chi-squared test). Sample size (n) is 25 for both unirradiated groups, 34 for irradiated N2 [S/S] homozygotes, and 39 for irradiated N2 [R/S] heterozygotes.

IR = ionizing radiation
Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L3 = larva of the L3 stage
Dpy = dumpy phenotype
N2 [S/S] = homozygous animals obtained from an N2 [S] male and N2 [S] hermaphrodite
N2 [R/S] = heterozygous animals obtained from an N2 [R] male and N2 [S] hermaphrodite

*** = $p<0.001$

Figure 2.10

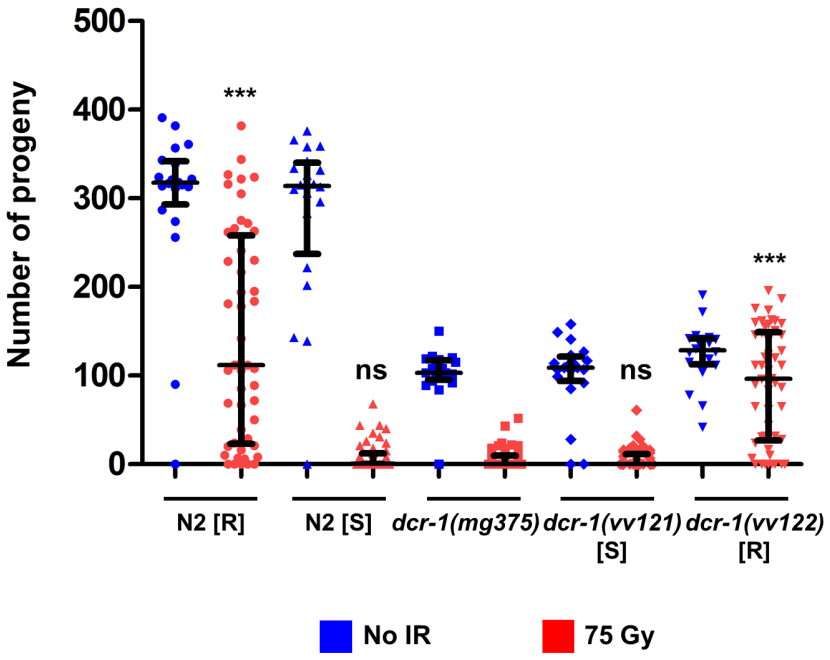

Number of progeny
500
400
300
200
100
0
***
ns
ns
***
N2 [R]
N2 [S]
dcr-1(mg375)
dcr-1(vv121) [S]
dcr-1(vv122) [R]
No IR
75 Gy

Figure 2.10 - The helicase domain of DCR-1 does not confer IR-resistance

Total brood size quantification in N2 [R], N2 [S], and *dcr-1* helicase mutants following IR treatment at the L1 stage. Irradiated *dcr-1(mg375)* animals (genetic background unknown) do not exhibit a significantly different brood size than either irradiated N2 [S] animals or irradiated *dcr-1(vv121)* animals, which harbour the same mutation as *dcr-1(mg375)* but in the N2 [S] background. *dcr-1(vv122)* animals, which harbour the same *dcr-1* helicase-inactivating mutation in the N2 [R] background, show a significantly higher brood size than either of N2 [S], *dcr-1(mg375)*, or *dcr-1(vv121)*, showing that post-IR brood size is determined by genetic background and not the *dcr-1* helicase mutation. All statistical comparisons shown in the figure are to irradiated *dcr-1(mg375)* (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 20 for all unirradiated groups, and 50 for all irradiated groups.

IR = ionizing radiation
Gy = Gray (unit)
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
dcr-1(mg375) = *dcr-1* helicase dead mutant [G492R] in unknown genetic background
dcr-1(vv121) = *dcr-1* helicase dead mutant [G492R] in N2 [S] background
dcr-1(vv122) = *dcr-1* helicase dead mutant [G492R] in N2 [R] background

ns = not significant ($p>0.05$), *** = $p<0.001$

Figure 2.11

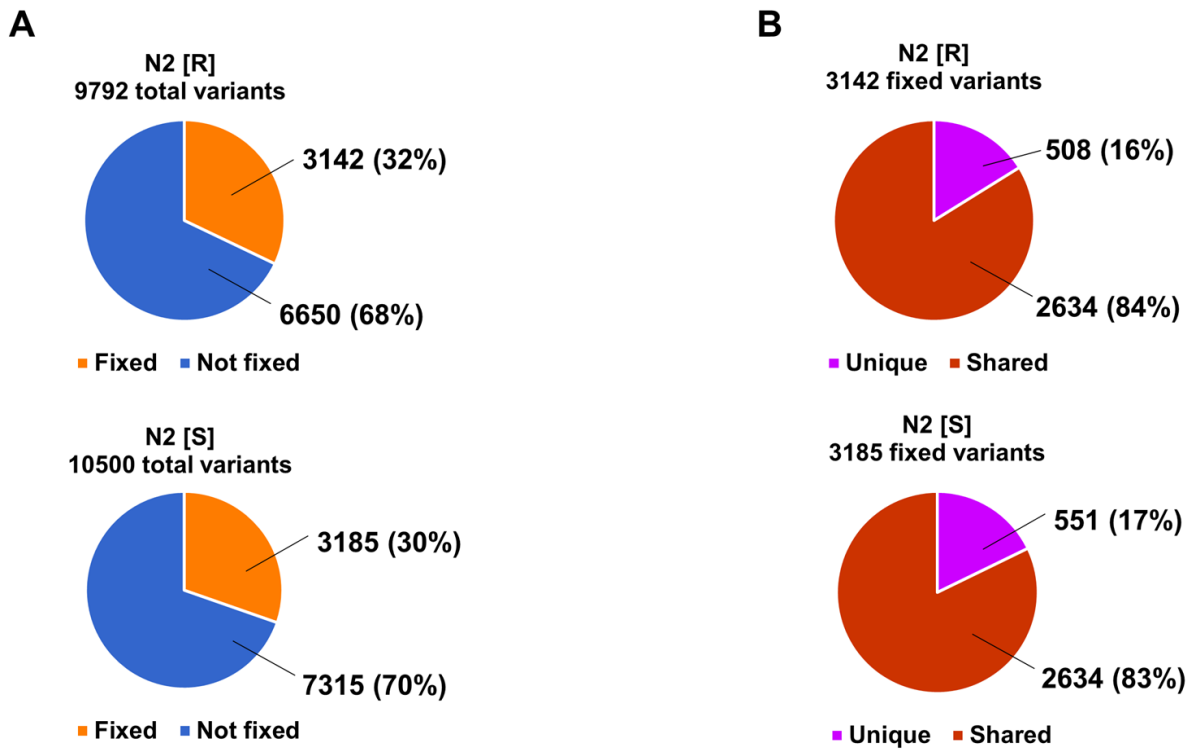


Figure 2.11 - Variants in the N2 [R] and N2 [S] genomes

(A) Pie charts showing the proportion of variants in the N2 [R] and N2 [S] genomes that are fixed (invariant in the population) and not fixed (polymorphic). Although most of the variants are not fixed, just under a third of variants is fixed in both genomes.

(B) Pie charts showing the proportion of fixed variants that is shared between the two sequenced N2 backgrounds and the proportion that is unique to each. Most of the fixed variants are shared by N2 [R] and N2 [S], suggesting that these lineages split from each other after their ancestor diverged from the reference genome.

N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

Table 2.3

Gene name	Different from reference in	Mutation	Chromosomal location	Mutation effect	Molecular identity	Potential function
trpp-8	N2 [S]	C->A	Chr I, 5616196	AA substitution	Transport protein	Involved in intracellular transport and meiosis
ZC239.15	N2 [S]	C->G	Chr II, 3223743	AA substitution	BTB2 domain containing KCTD10 ortholog	Potassium chanel (inferred)
emb-27	N2 [S]	G->C	Chr II, 8160254	AA substitution	APC subunit, cdc16 ortholog	Mitotic Exit
Y22D7AL.7	N2 [S]	G->A	Chr III, 1578431	AA substitution	Unknown protein, some homology to Rouxiella DNA binding response regulator	UNKNOWN
C50C3.2	N2 [S]	G->T	Chr III, 8188649	AA substitution	SPTAN1 ortholog (Spectrin)	Cell cortex structural
kin-24	N2 [S]	G->T	Chr IV, 9840056	AA substitution	Kinase, human FES ortholog	Kinase, SH2 domain containing, PKC superfamily
F07C6.4	N2 [S]	C->T	Chr IV, 12809030	AA substitution	FRMD4B ortholog	Putative actin binding, FERM domain
F10D2.12	N2 [S]	G->A	Chr V, 7149488	AA substitution	Hexosyl transferase (putative)	Glycosylation (inferred)
inft-2	N2 [S]	C->A	Chr V, 13003815	AA substitution	Inverted formin	FH2 superfamily domain, cytoskeletal organization
sru-48	N2 [S]	C->T	Chr X, 5180366	AA substitution	Serpentine GPCR receptor	Cell signaling
F48F7.3	N2 [S]	G->T	Chr X, 13964441	AA substitution	Galactosyl transferase (putative)	Galactosylation (inferred)
C06G1.6	N2 [S]	G->A	Chr X, 16652703	AA substitution	PPP2R3B ortholog (putative phosphatase)	Dephosphorylation (inferred)
fbxb-69	N2 [S]	T->A	Chr I, 14283531	STOP gained	F-box protein	UNKNOWN
C33H5.1	N2 [S]	G->T	Chr IV, 7810427	STOP gained	Nematode-specific putative methyltransferase	UNKNOWN
bre-5	N2 [S]	G->T	Chr IV, 12050973	Splice site loss	Galactosyl transferase (putative)	Galactosylation (inferred)

Table 2.3 - Unique variants which affect protein coding genes in N2 [S]

This table presents the list of 15 fixed (homozygous) single nucleotide variants unique to N2 [S] which affect the protein sequence of the genes in which they are found. Thirteen variants result in amino acid substitutions, two introduce a premature stop codon, and one results in the abolition of a splice site. None of the variants affects a gene with a known role in DNA repair or DDR.

Figure 2.12

		inft-2	*C33H5.1*	*bre-5*	*trpp-8*	*kin-24*	*fbxb-69*	*emb-27*	*Y22D7AL.7*	*ZC239.15*	*F10D2.12*	*F07C6.4*	*sru-48*	*C06G1.6*	*F48F7.3*	*C50C3.2*
SR 1	Hybrid line 1	[S]	[S]	[S]	[S]	[S]	[S]				[S]	[S]	[R]	[R]	Het	[R]
	Hybrid line 2	[S]	[R]	[R]	Het	[R]	Het	[S]		Het	[S]	[R]				
	Hybrid line 3	[S]		Het	[S]	Het	[S]	Het			[S]	Het				
	Hybrid line 4	[S]	[S]	[S]	[S]	[S]	Het	[R]			[S]	[S]				
SR 2	Hybrid line 5	[S]		Het	[S]	Het	[R]	[R]		[S]	[S]	[R]				
	Hybrid line 6	[S]		Het	[S]	Het	Het	Het			[S]	Het				
	Hybrid line 7	[S]	[S]	[S]	Het	[S]	[S]	[S]		[R]	[S]	[S]				
	Hybrid line 8	[S]	[S]	[S]	[R]	Het	[R]	[R]		Het	[S]	Het				
	Hybrid line 9	[S]	[R]	[R]	Het	[R]		Het	[R]	[S]	[S]	[R]				
	Hybrid line 10	[S]									[S]		Het			
	Hybrid line 11	[S]											Het			
	Hybrid line 12	[S]									[S]		[S]			
	Hybrid line 13	[S]									[S]		Het			
	Hybrid line 14	[S]									[S]		[R]			
SR 3	Hybrid line 15	[S]									[S]		[S]			
	Hybrid line 16	[S]									[S]		Het			
	Hybrid line 17	[S]									[S]		[S]			
	Hybrid line 18	[S]									[S]		[R]			
	Hybrid line 19	[S]									[S]		[S]			
	Hybrid line 20	[S]									[S]		[R]			
	Hybrid line 21	[S]									[S]		[S]			

[S]	[S/S] Homozygote
[R]	[R/R] Homozygote
Het	[S/R] Heterozygote
	Not sequenced/sequencing error

Figure 2.12 - First line mapping of the IR-sensitivity-causative mutation shows linkage to *inft-2* and *F10D2.12*

This figure presents the results of the sequencing of the 15 protein-sequence affecting homozygous variants unique to N2 [S] in 21 phenotypically sensitive hybrid lines chosen among 80 hybrid lines created from individual F2 animals obtained from a cross between an N2 [S] male and an N2 [R] hermaphrodite (**Figure 5.2**). A heterozygous or [R/R] homozygous sequencing outcome for a particular locus in even one of the hybrid lines excludes that locus as causative. The lines were sequenced in three rounds: only hybrid line 1 in the first round, which eliminated *sru-48, C06G1.6, F48F7.3,* and *C50C3.2*; hybrid lines 2-9 in the second round, which eliminated all loci except *inft-2* and *F10D2.12*; and hybrid lines 10-21 in the third round, which sequenced only *inft-2* and *F10D2.12* to test the extent of linkage of these two loci with the IR-sensitive phenotype, and *sru-48* as a control for DNA source.

SR = sequencing round

Figure 2.13

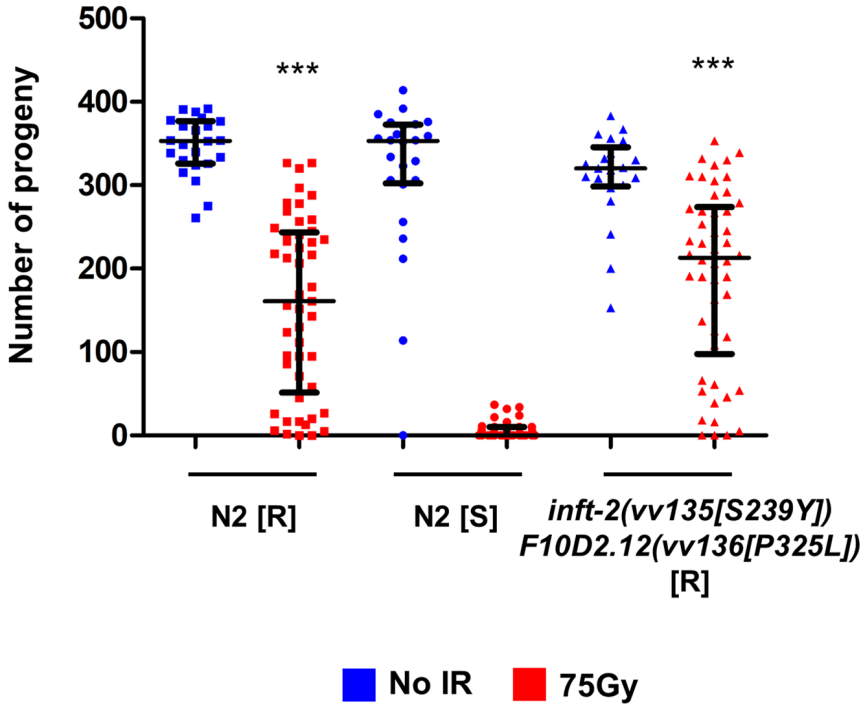

500
400
300
200
100
0
Number of progeny
***
***
N2 [R]
N2 [S]
inft-2(vv135[S239Y])
F10D2.12(vv136[P325L])
[R]
No IR
75Gy

Figure 2.13 - The *inft-2* and *F10D2.12* mutations in the N2 [S] genome are not causative of IR-sensitivity

Total brood size quantification of N2 [R], N2 [S], and *inft-2(vv135) F10D2.12(vv136)* double mutants following irradiation with 75 Gy of IR at the L1 stage. The *inft-2(vv135) F10D2.12(vv136)* double mutants, which carry the N2 [S]-form of the mutations in these two genes but in the N2 [R] genetic background, have a significantly higher post-IR brood size than N2 [S], but not N2 [R] ($p>0.05$). All statistical comparisons shown in the figure are to irradiated N2 [S] (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 23 for unirradiated N2 [R], 49 for irradiated N2 [R], 24 for unirradiated N2 [S], 46 for irradiated N2 [S], 21 for unirradiated *inft-2(vv135) F10D2.12(vv136)* and 50 for irradiated *inft-2(vv135) F10D2.12(vv136).*

IR = ionizing radiation

Gy = Gray (unit)

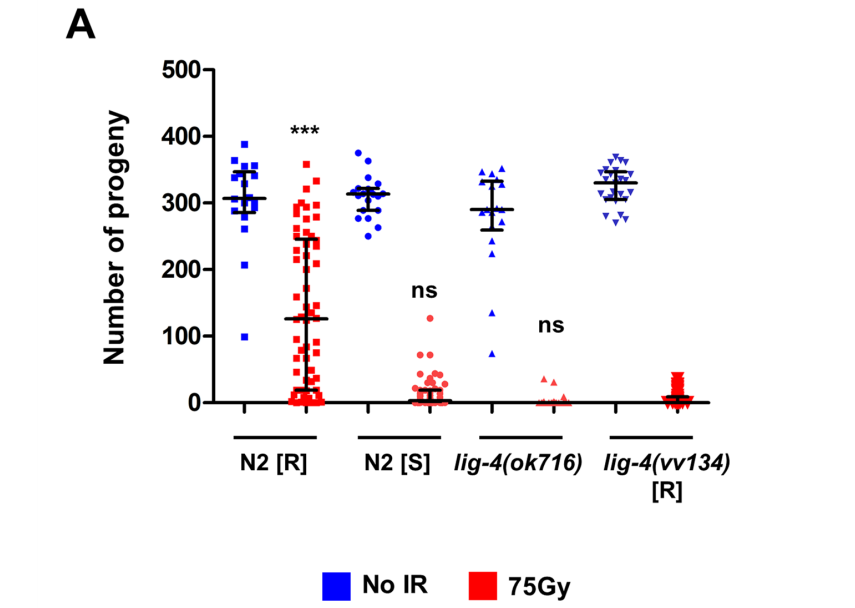
N2 [S] = sensitive N2 strain, derived from the CGC N2

N2 [R] = resistant N2 strain, derived from Andersen lab N2

inft-2(vv135[S239Y]) F10D2.12(vv136[P325L]) = a double mutant harbouring the N2 [S]-form of the *inft-2* and *F10D2.12* sequences, created by CRISPR mutagenesis in the N2 [R] genetic background.

*** = $p<0.001$

Figure 2.14


A
Number of progeny
500
400
300
200
100
0
***
ns
ns
N2 [R]
N2 [S]
lig-4(ok716)
lig-4(vv134) [R]
No IR
75Gy

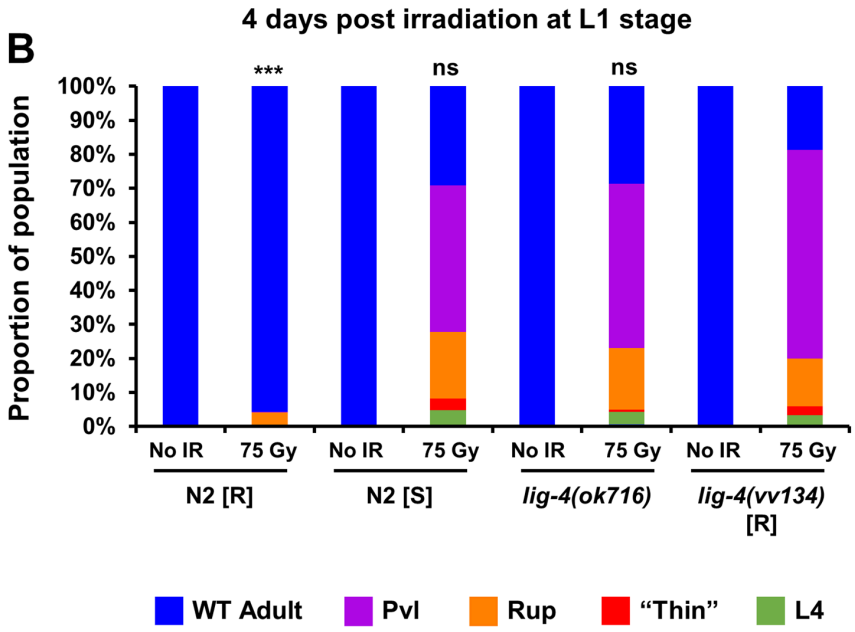

4 days post irradiation at L1 stage
B
Proportion of population
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
***
ns
ns
No IR
75 Gy
N2 [R]
N2 [S]
lig-4(ok716)
lig-4(vv134) [R]
WT Adult
Pvl
Rup
"Thin"
L4

Figure 2.14 - Loss of cNHEJ activity sensitizes the N2 [R] background to IR

(A) Total brood size quantification of N2 [R], N2 [S], and the Ligase IV mutants *lig-4(ok716)* and *lig-4(vv134)* after IR treatment at the L1 stage. Both the published *lig-4(ok716)* deletion mutant (whose genetic background is unknown) and *lig-4(vv134)*, a null mutant the generated by CRISPR in the N2 [R] genetic background, exhibit a post-IR brood size not significantly different than N2 [S] or each other ($p>0.05$ for all comparisons). All three of these genotypes display a significantly reduced post-IR brood size compared to N2 [R] ($p<0.001$ for all comparisons to post-IR N2 [R]). All statistical comparisons shown in the figure are to irradiated *lig-4(vv134)* (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 18 for unirradiated N2 [R], 58 for irradiated N2 [R], 20 for unirradiated N2 [S], 57 for irradiated N2 [S], 18 for unirradiated *lig-4(ok716)*, 56 for irradiated *lig-4(ok716)*, 25 for unirradiated *lig-4(vv134)*, and 50 for irradiated *lig-4(vv134).*

(B) Quantification of post-IR somatic phenotypes four days after IR treatment in the same groups as in **(A)**. N2 [S], *lig-4(ok716)*, and *lig-4(vv134)* animals show a much higher incidence of vulval phenotypes post-IR than do N2 [R] animals ($p<0.001$), while they are not significantly different than each other ($p>0.05$). All statistical comparisons shown in the figure are to *lig-4(vv134)* (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.008$). Sample size (n) is 18 for unirradiated N2 [R], 58 for irradiated N2 [R], 20 for unirradiated N2 [S], 57 for irradiated N2 [S], 18 for unirradiated *lig-4(ok716)*, 56 for irradiated *lig-4(ok716)*, 25 for unirradiated *lig-4(vv134)*, and 50 for irradiated *lig-4(vv134).*

IR = ionizing radiation
Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L4 = larva of the L4 stage
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2

ns = not significant ($p>0.05$ in **(A)**; $p>0.008$ in **(B)**), *** = $p<0.001$

Figure 2.15

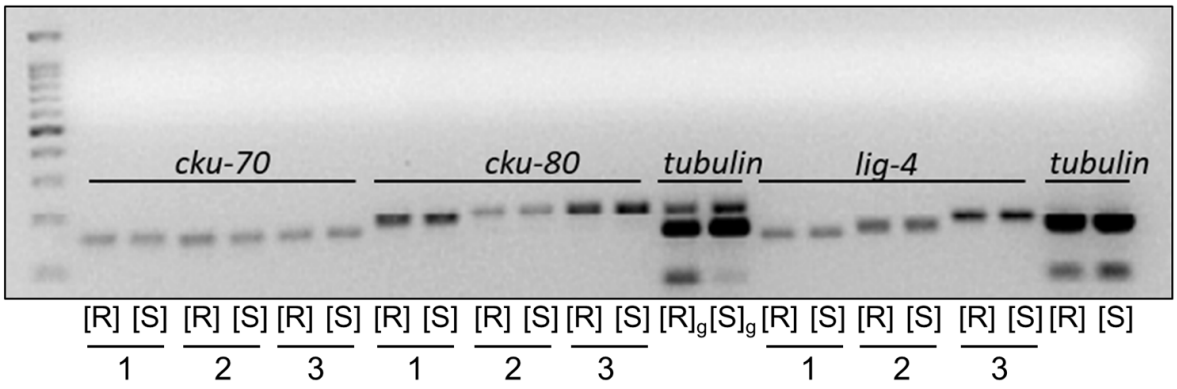


Figure 2.15 - Transcripts of known cNHEJ genes are present in both N2 [R] and N2 [S]

Non-quantitative RT-PCR of transcripts of the three known *C. elegans* cNHEJ genes. Each gene is amplified in three reactions using a different exon-exon junction-overlapping primer set on both N2 [R] and N2 [S] cDNAs. All three cNHEJ gene transcripts are detectable in both N2 [R] and N2 [S] cDNA. A single set of tubulin primers was used as a control for genomic DNA and cDNA quality.

[R] = cDNA from N2 [R] used as the reaction template
[S] = cDNA from N2 [S] used as the reaction template
$[R]_g$ = genomic DNA from N2 [R] used as the reaction template
$[S]_g$ = genomic DNA from N2 [S] used as the reaction template

1 = first primer pair for given gene
2 = second primer pair for given gene
3 = third primer pair for given gene

Chapter III: Characterization of *nhj-1,* a novel member of the canonical non-homologous end joining pathway in *C. elegans*

III.1 Introduction and second round of genetic mapping

In the last chapter, I presented evidence that different lines of the most commonly used *C. elegans* wild type strain, the Bristol isolate N2, exhibit phenotypic diversity in the response to ionizing radiation exposure at the L1 stage. Distinct IR response profiles characterize the sensitive line N2 [S] and the resistant line N2 [R]. I also demonstrated that IR-sensitivity is recessive to IR-resistance, and that it segregates in a classic 3:1 Mendelian ratio, suggesting that it is caused by cryptic variation in the N2 [S] genome. Deep sequencing of the N2 [S] and N2 [R] genomic DNA allowed our bioinformatics collaborators in the group of Dr. Steven Jones to identify 15 mutations unique to the N2 [S] genome affecting protein sequences, which made them the best candidates for the causative mutation. However, while I was able to map the IR-sensitivity-causative locus to chromosome V, I also definitively excluded the only two candidate mutations on that chromosome predicted to cause a change in the protein sequence.

This left 84 N2 [S]-specific homozygous variants on chromosome V located outside coding sequences, of which 16 were determined by the Jones group to be of interest because of proximity to coding genes. Based on the initial mapping result, in which the mutations in *inft-2* and *F10D2.12* were linked to the sensitive phenotype in every successful sequencing reaction, I reasoned that the causative gene was likely located close (in genetic map units) to these two markers. Despite being close together in map units (*F10D2.12* is located at V:0.61 cM and *inft-2* at V:4.81 cM), these two loci are almost 6 million base pairs apart in physical distance (*cf* the entire length of chromosome V at approximately 20 million base pairs). The intervening region contains 9 of the 16 candidate gene-proximal mutations, while the the flanking regions contain two additional candidate mutations which are close enough genetically at V:5.81 cM and V:5.12 cM that I could not exclude them on the basis of the initial mapping results. Considering that the causative mutation may not be among the gene-proximal candidates, I decided to map its location with a higher resolution before commencing to test individual candidates. For this

second round of mapping, I employed a similar strategy as in the first round (**Figure 5.2**), though expanded in scope to 350 N2 [S]/N2 [R] hybrid strains. As molecular markers for this mapping round, I used only the *inft-2* and *F10D2.12* mutations, as they were both far enough apart within the region of interest to provide useful information and amenable to restriction fragment length polymorphism (RFLP)-based genotyping.

Of the 350 hybrid strains, 75 were determined to be phenotypically IR-sensitive, and therefore suitable for genotyping. Of these, 71/75 carried two copies of the N2 [S]-like mutation in *F10D2.12* and *inft-2*, reinforcing the tight linkage between these loci and the causative locus. Among the 4 recombinants, however, two lines were homozygous N2 [R]-like for *F10D2.12* and homozygous N2 [S]-like for *inft-2*, one line was heterozygous at both markers, and one was heterozygous for *F10D2.12* and homozygous N2 [S]-like for *inft-2* (**Figure 5.2C**). These recombination rates indicated that the causative locus is genetically closer to *inft-2*, and most likely within 3 or 4 cM from *F10D2.12*. Depending on the interpretation of the single hybrid line which was homozygous for both marker loci, the causative locus could either be located between *F10D2.12* and *inft-2* (in which case the double heterozygote contains two recombinant chromosomes), or on the far side of *inft-2* (in which case the double heterozygote has one recombinant chromosome).

III.2 An insertion/deletion mutation in N2 [S] disrupts the coding sequence of the uncharacterized gene *nhj-1*/*H19N07.3*

The closer linkage to *inft-2* made it feasible for me to manually parse through the region in the Integrative Genomics Viewer (IGV, (Robinson, Thorvaldsdottir et al. 2011)), during which I discovered a 5 nucleotide deletion and 10 nucleotide duplication in the middle of the third exon in the uncharacterized protein-coding sequence *H19N07.3*, located at V:2.95 cM, which was present in the N2 [S] genome but not the N2 [R] genome. Interestingly, while the *H19N07.3* locus could be PCR-amplified using primers flanking the coding sequence in the N2 [R] genome, it was refractory to PCR-amplification in the N2 [S]. This suggested the possibility that the indel reported by the IGV was part of a larger mutation, perhaps a breakpoint of a larger genomic rearrangement. However, by re-aligning the deep sequencing reads without soft-clipping, a

technique which masks sequencing reads which do not align to the genome end-to-end, our collaborators in the Jones group were able to uncover the true sequence of the mutation in *H19N07.3* – a deletion of 5 nucleotides, an insertion of 107 nucleotides, and an 8-nucleotide duplication of a small segment of exon 3 (**Figure 3.1A**). Regardless of the alignment methodology, this indel had eluded discovery during previous analyses because of its size; most single nucleotide variant (SNV) callers are optimized for variants 10 base pairs or less in size, while copy number variant (CNV) callers recognize much larger alterations, typically above 500 base pairs (Pirooznia, Goes et al. 2015, Xu 2018). Intriguingly, the 107 bp insertion is exclusively composed of adenosines and thymidines, and is predicted to form a strong hairpin (**Figure 3.1B**), which may be the primary reason preventing PCR amplification of *H19N07.3* in the N2 [S] genome (Kieleczawa 2006). Aligning one primer in the first half of the hairpin sequence allows for N2 [S]-specific PCR product to be generated (see section **V.15** for more information).

We have given this natural mutation the allele designation *vv148*, and have provisionally named the *H19N07.3* gene *nhj-1* for non-homologous end joining protein 1, reflecting the putative cNHEJ repair deficiency of the N2 [S] strain. The *nhj-1* gene encodes a protein with 168 amino acid residues, which is truncated by the *nhj-1(vv148)* indel to 93 amino acids, the last two of which are missense residues (**Figure 3.1C**). The *nhj-1* locus can also produce a shorter transcript from which a 130 amino acid NHJ-1 protein isoform can be translated using an alternative start codon located in exon 2. However, a search of the modENCODE RNA expression library data revealed that the short isoform transcripts are only detectable in embryos, in contrast to the long isoform transcripts which are detectable at any life stage. For the purposes of experiments described in the present study, all mutations and alterations in the *nhj-1* locus affect both protein isoforms as they occur in or downstream of exon 3.

III.3 Inactivation of *nhj-1* in the N2 [R] background sensitizes it to ionizing radiation

Because of the major effects the *nhj-1(vv148)* indel is predicted to have on the coding sequence, *nhj-1(vv148)* became the best candidate for the IR-sensitivity-causative mutation as soon as it was discovered. I first wanted to test whether inactivation of *nhj-1* in N2 [R] could sensitize this genetic background to IR. I employed CRISPR mutagenesis to create a 7-nucleotide deletion in

the same region of exon 3 where *nhj-1(vv148)* is found, which I have designated *nhj-1(vv144)* (**Figure 3.1A**). This allele is predicted to truncate the protein to 111 amino acids (73 amino acids in the short isoform), of which 22 are missense residues, leaving only 89 sense residues (51 in the short isoform) (**Figure 3.1C**). While the *nhj-1(vv144)* mutation does not impact the brood size of unirradiated control animals ($p>0.05$ against both N2 [R] and N2 [S] unirradiated groups), it reduces the post-IR brood size to a median of 4 progeny (IQR 17.5), not significantly different from the median 0 progeny (IQR 9) produced by post-IR N2 [S] animals ($p>0.05$) but significantly lower than the median 179 progeny (IQR 166.25) produced by irradiated N2 [R] worms ($p<0.001$) (**Figure 3.2A**). Irradiated *nhj-1(vv144)* animals also were not significantly different from N2 [S] in somatic IR effects, with almost all irradiated *nhj-1(vv144)* and all irradiated N2 [S] animals still found in larval stages three days after treatment, when all unirradiated controls and irradiated N2 [R] animals have become adults (**Figure 3.2B**). This experiment shows that disruption of *nhj-1* in the N2 [R] genetic background causes IR-sensitivity; however, it doesn't formally demonstrate that N2 [S] is IR-sensitive because of the *nhj-1(vv148)* mutation, as it is possible that the NHJ-1^{vv144} protein is more severely affected than NHJ-1^{vv148}, since it is two sense residues shorter, and that another mutation on chromosome V in N2 [S] could be the real causative variant.

III.4 The *nhj-1(vv144)* [R] genome does not complement the N2 [S] genome

To test this possibility, I performed a complementation test in *nhj-1(vv144)* [R] and N2 [S] heterozygotes. While *nhj-1(vv144)* [R]/N2 [R] heterozygotes showed a post-IR brood size not significantly different than N2 [R] homozygotes (108 median progeny and an IQR of 153, vs 149 median progeny and an IQR of 188; $p>0.05$), *nhj-1(vv144)* [R]/N2 [S] heterozygotes exhibited a post-IR brood size that was the same as that of N2 [S] homozygotes (median of 2 progeny and an IQR of 16, vs a median of 2 progeny and an IQR of 24.25) (**Figure 3.3A**). The distribution of post-IR somatic phenotypes mirrored the brood size outcomes, with more than 80% of both *nhj-1(vv144)*/N2 [R] heterozygous animals and N2 [R] homozygous animals having developed into phenotypically normal adults, compared to only 20% of animals in the irradiated N2 [S] homozygote group or 19% in the *nhj-1(vv144)*/N2 [S] heterozygotes (**Figure 3.3B**). Since the *nhj-*

1(vv144) [R] genome is not able to complement the IR-sensitivity of the N2 [S] genome, I could conclude that the IR-sensitivity in N2 [S] is indeed caused by the *nhj-1(vv148)* mutation.

III.5 NHJ-1 functions in the canonical non-homologous end joining pathway

The observation that the loss of cNHEJ activity can sensitize the N2 [R] background to IR (**Figure 2.14**) raised the intriguing possibility that the N2 [S] background may be deficient in a novel component of the cNHEJ pathway, because the three known *C. elegans* cNHEJ factors are all encoded on chromosome III while the sensitizing mutation, now identified as *nhj-1(vv148)*, is on chromosome V and doesn't abolish transcription of the three known cNHEJ factors (**Figure 2.15**). To test the possibility that *nhj-1* acts in the cNHEJ pathway, I created two double mutants: 1) *lig-4(vv141[R18STOP])*, a premature stop codon allele introduced by CRISPR mutagenesis in the same position as the *lig-4(vv134)* allele (see section **II.10** and **Figure 2.14**), but in the N2 [S], *ie*, *nhj-1(vv148)* background; and 2) the *nhj-1(vv144)* deletion allele crossed into the *cku-80(tm1203)* background, which contains a large deletion in the Ku ring component *cku-80*, and has been previously characterized as IR-sensitive (Clejan, Boerckel et al. 2006). If *nhj-1* is part of the same DNA repair pathway as *cku-80* and *lig-4*, the double mutants are predicted to exhibit an IR response of the same severity as either single mutant alone. If, however, *nhj-1* belongs to a different DNA repair pathway that merely mirrors the phenotypes exhibited by cNHEJ mutants, the double mutants would be expected to have an additively more severe post-IR phenotype than the single mutants. Because the brood size and somatic phenotypes of the sensitive lines are already very severe at 75 Gy, the dose that I most commonly employed in this project, I decided to test the additivity hypothesis at 37.5 Gy, which represents half the regular dose but is expected to produce a significantly reduced brood size in the single mutants (see **Figure 2.5A**).

While at this dose the post-IR brood size of N2 [R] animals (median of 290 progeny, IQR 64.25) was not significantly different than that of unirradiated controls (median of 297 progeny, IQR 48; $p>0.05$), it was significantly reduced in N2 [S] animals (median of 188 progeny, IQR of 145.5 vs a median of 307 progeny, IQR 43 in unirradiated controls; $p<0.001$) and in *lig-4(vv134)* [R] animals (median of 102 progeny, IQR 132.5, compared to a median of 287 progeny, IQR 39.5 in unirradiated controls; $p<0.001$) (**Figure 3.4A**). Crucially, while the post-IR brood size of *lig-*

4(vv141); nhj-1(vv148) [S] double mutant animals was significantly lower than in unirradiated controls (median of 170.5 progeny, IQR 186.75, versus a median of 297.5 progeny, IQR 30; $p<0.001$), it was not significantly different than that of post-IR N2 [S] or *lig-4(vv134)* [R] animals ($p>0.05$ for both comparisons) (**Figure 3.4A)**. Similarly, the post-IR incidence of somatic phenotypes in *lig-4(vv141); nhj-1(vv148)* [S] double mutants was not significantly different than those of *lig-4(vv134)* [R] animals at 37.5 Gy ($p>0.05$) (**Figure 3.4B**). Although both *lig-4* [R] single mutants and *lig-4(vv141); nhj-1(vv148)* [S] double mutants exhibited a higher incidence of somatic phenotypes than N2 [S] at 37.5 Gy ($p<0.001$ for both comparisons), at 75 Gy there was no significant difference between the three groups ($p>0.05$) (**Figure 3.4B**). Since the phenotypes of *lig-4* and *nhj-1* were not additive, I concluded that *nhj-1* falls in the same pathway as *lig-4*, which would make it a fourth member of the *C. elegans* cNHEJ system.

This conclusion was further corroborated by data from *cku-80(tm1203); nhj-1(vv144)* double mutants. After treatment with 37.5 Gy of IR, *cku-80(tm1203); nhj-1(vv144)* animals exhibit a brood size (median of 40.5 progeny, IQR 67.75) not significantly smaller than either *cku-80(tm1203)* (median of 55 progeny, IQR 113.5) or *nhj-1(vv144)* (median of 70 progeny, IQR 138.75; $p>0.05$ for both comparisons) (**Figure 3.4C**). The three genotypes are not significantly different in the incidence of somatic phenotypes at either 37.5 Gy or 75 Gy (**Figure 3.4D**), strongly suggesting that *cku-80* and *nhj-1* belong to the same genetic pathway.

III.6 The IR-sensitivity of *nhj-1(vv144)* is rescued by extrachromosomal NHJ-1

Although the experiments described in sections **III.3** and **III.4** provide evidence that loss of *nhj-1* results in radiation sensitivity, I wanted to provide further support for that conclusion by testing whether a mutation in *nhj-1* can be rescued by an exogenously provided *nhj-1* sequence. For this I made use of a fosmid containing GFP-tagged *nhj-1*, generated as part of the *C. elegans* TransgeneOme project which created a library of fosmids containing GFP-tagged versions of 73% of the proteins in the worm proteome (Sarov, Murray et al. 2012). The TransgeneOme fosmids carry a wild type sequence of *unc-119* gene as a selection marker for the presence of the transgenic array, which can otherwise be lost, and are therefore maintained in *unc-119* mutant animals. Through the generosity of Dr. Henrik Bringmann, I obtained an *unc-119(ed3)* strain

rescued by the extrachromosomal array *goeEx386[nhj-1::eGFP::unc-119]* (for the full genotype of the array, see text of **Figure 3.5**). I then crossed the *nhj-1(vv144)* mutation into *unc-119(ed3) goeEx386[nhj-1::eGFP::unc-119]*, and compared the post-IR response of these animals to *nhj-1(vv144)* alone. The post-IR brood size of *nhj-1(vv144)* was significantly lower than that of *nhj-1(vv144); unc-119(ed3) goeEx386[nhj-1::eGFP::unc-119]* animals (median of 2 progeny, IQR 10.75, compared to a median of 131 progeny, IQR 100; $p<0.001$) (**Figure 3.5A**). The post-IR somatic phenotypes were also significantly less prevalent in *nhj-1(vv144); unc-119(ed3) goeEx386[nhj-1::eGFP::unc-119]* animals compared to *nhj-1(vv144)* (**Figure 3.5B**). Thus, the extrachromosomal array is able to rescue both brood size and somatic phenotypes caused by the loss of *nhj-1*.

To exclude the possibility that the observed IR phenotype rescue was due to *goeEx386* rescuing *unc-119* instead of *nhj-1*, and to examine whether extrachromosomal *nhj-1* can also rescue the *nhj-1(vv148)* allele, I compared the post-IR response of *nhj-1(vv148); unc-119(ed3)/+* animals to *nhj-1(vv148); unc-119(ed3)/+ goeEx386[nhj-1::eGFP::unc-119]* animals. Because these animals are heterozygotes for *unc-119*, it was not possible to verify the presence of *goeEx386* in all tested animals of that group. However, *goeEx386* was present in enough animals to significantly rescue the post-IR brood size of *nhj-1(vv148); unc-119(ed3)/+ goeEx386* to a median of 109.5 progeny (IQR 150.5), compared to the *nhj-1(vv148); unc-119(ed3)/+* post-IR brood size of 12.5 median progeny (IQR 27.5; $p<0.05$) (**Figure 3.5C**). The presence of the extrachromosomal array was also able to significantly reduce the prevalence of somatic phenotypes three days after IR treatment in *nhj-1(vv148); unc-119(ed3)/+ goeEx386* ($p<0.001$, compared to *nhj-1(vv148); unc-119(ed3)/+*) (**Figure 3.5D**), further supporting the conclusion that exogenous *nhj-1* can rescue the IR-sensitivity caused by loss of endogenous *nhj-1* function.

III.7 NHJ-1 is expressed in most nuclei in the L1 larva

Because TransgeneOme constructs are tagged with GFP, possession of *goeEx386* allowed me to examine the subcellular localization of the NHJ-1 protein. The localization of NHJ-1::GFP in adult animals has already been examined in a paper published by the Bringmann group, where it was used merely as a negative control for a transcription control assay. In their work, they reported

NHJ-1::GFP expression in all somatic nuclei of the adult animal (Turek, Besseling et al. 2016), which is a pattern consistent with the role of NHJ-1 in cNHEJ, the primary modality of DSB repair in the soma (Clejan, Boerckel et al. 2006).

With this tool in hand, I wanted to examine the localization of NHJ-1::GFP in the L1 larva, and test whether its localization is affected by IR treatment. In unirradiated L1 larvae, NHJ-1::GFP can be seen in nuclei spanning the entire body of the animal (**Figure 3.6A**). Closer examination revealed two salient features of the expression pattern: 1) while many nuclei expressed NHJ-1, not every nucleus did; and 2) NHJ-1 was not detectable in the primordial germ cells, marked by HTP-3 (**Figure 3.6B**). Both of these features could be artifacts of expression from an extrachromosomal array, as not every cell is guaranteed to receive a copy since they are segregated randomly at mitosis; furthermore, extrachromosomal arrays are often repetitive and are recognized as foreign DNA and silenced in the germline (Kelly, Xu et al. 1997). These caveats necessitated an investigation of the endogenous NHJ-1 localization pattern. However, the lack of expression of NHJ-1 from *goeEx386* in the PGCs provides further support for the idea that the loss of *nhj-1* in the soma is a major cause of the post-IR brood size defects, as they can both be rescued by somatic NHJ-1 expression (**Figure 3.5** and **Figure 3.6**). In addition, IR treatment did not affect the subcellular localization of NHJ-1 (**Figure 3.6B**), perhaps because any additional signal or localized rearrangement may be masked by overexpression from the multi-copy extrachromosomal array.

To probe the endogenous NHJ-1 localization pattern, I employed CRISPR mutagenesis to tag the C-terminus of the *nhj-1* locus with OLLAS, a small epitope tag which can be used with commercially available antibodies and has very high sensitivity across several applications, including immunohistochemistry (Park, Cheong et al. 2008). Immunostaining of L1 larvae revealed that NHJ-1::OLLAS expressed from the endogenous locus has a localization pattern highly similar to that of NHJ-1::GFP expressed from *goeEx386* (**Figure 3.7A-B**). Notably, NHJ-1 is not detectable in the PGCs (**Figure 3.7A**), consistent with the well understood fact that HR, and not cNHEJ, is the primary DSB repair pathway in the germline (Clejan, Boerckel et al. 2006, Lemmens and Tijsterman 2011). Like the extrachromosomally-expressed NHJ-1::GFP, the endogenous NHJ-1::OLLAS localization pattern is not affected by ionizing radiation treatment (**Figure 3.7A-B**). This contrasts with the HR strand invasion factor RAD-51, a commonly used

marker for HR repair in *C. elegans*, which is robustly cytologically detectable as discrete foci after irradiation in the tissues where HR predominates, such as the adult germline (Alpi, Pasierbek et al. 2003, Hayashi, Chin et al. 2007) or the PGCs (Butuci, Williams et al. 2015). Although nothing is known about the subcellular localization of cNHEJ factors in *C. elegans*, the ones whose localization has been probed in cultured mammalian cells (Ku80, LIG4, DNA-PKcs, XRCC4 and PAXX) are predominantly nuclear, and localize along the path of a laser in microirradiation experiments (Mari, Florea et al. 2006, Ochi, Blackford et al. 2015, Xing, Yang et al. 2015). In contrast to the laser, IR treatment in an X-ray irradiator induces spatiotemporally diffuse DSBs, which may preclude detection of distinct foci of cNHEJ factor accumulation (see **Chapter IV**).

The localization of some cNHEJ proteins may also depend on the presence of others, as is the case for XRCC4, which depends on its binding partner LIG4 for nuclear localization, and is cytoplasmic in absence of LIG4 (Francis, Kozlov et al. 2014). I therefore wanted to examine whether the status of the Ku ring or LIG-4 influences the localization of NHJ-1. However, neither the nuclear localization nor the intranuclear appearance of NHJ-1::OLLAS was affected in *cku-80(tm1203)* or *lig-4(vv134)* mutants, with or without IR treatment (**Figure 3.7A-B**), indicating that NHJ-1 localization is independent from the other known cNHEJ factors. To express this finding in quantitative terms, I scored the proportion of NHJ-1::OLLAS positive nuclei among all nuclei in the larval mid-section around the PGCs. In unirradiated [R] background animals, 74% of scored nuclei visibly express NHJ-1::OLLAS, compared to 65% in irradiated [R] background animals, a difference which isn't statistically significant ($p>0.05$) (**Figure 3.7C**). In unirradiated *lig-4(vv134)* animals, 66% of scored nuclei are positive for NHJ-1::OLLAS, as are 67% of scored nuclei in irradiated *lig-4(vv134)*, proportions which are not significantly different either from each other or the [R] background ($p>0.05$ for all comparisons) (**Figure 3.7C**). Similarly, NHJ-1::OLLAS is detectable in 72% of unirradiated and 66% of irradiated *cku-80(tm1203)* nuclei, which is likewise not significantly different compared to the unirradiated or irradiated [R] background or *lig-4(vv134)* mutants, nor each other ($p>0.05$) (**Figure 3.7C**).

III.8 NHJ-1 is expressed in late meiotic prophase I in the adult germline

Repair of DSBs in the *C. elegans* germline favors HR both for reasons of maintaining sequence fidelity and, in the meiotic context, for the formation of crossover events that guarantee both genetic diversity and proper segregation of homologous chromosomes in meiosis I (Lemmens and Tijsterman 2011, Lui and Colaiacovo 2013). However, cNHEJ can operate in the adult germline in absence of HR (Adamo, Collis et al. 2010, Yin and Smolikove 2013). Under normal meiotic conditions, cNHEJ activity is blocked by the activity of COM-1, which prevents the engagement of Ku with free DNA ends that would otherwise rapidly occur after SPO-11-mediated DSB induction (Lemmens, Johnson et al. 2013). However, that at least some cNHEJ activity persists in the germline is supported by evidence that HR-mediated CRISPR knock-in efficiency in injected gonads is increased in *cku-80(RNAi)* knockdown conditions (Ward 2015). I therefore wanted to investigate whether NHJ-1 is present in the adult germline.

Immunostaining of dissected gonads of [R] background animals revealed that NHJ-1::OLLAS is not present (*ie* does not exhibit a signal beyond background levels) in the mitotic zone, nor in meiotic cells until the diplotene stage of prophase I (**Figure 3.8A**). In diplotene and diakinesis, NHJ-1::OLLAS exhibited a diffuse punctate localization pattern within nuclei, with no obvious localization bias toward the condensing (or condensed) chromatin. This expression pattern is consistent with the role of cNHEJ as a backup DSB repair pathway in the germline, as NHJ-1::OLLAS accumulation starts after the pachytene stage, when RAD-51-mediated inter-homolog HR repair is most prominent (Colaiacovo, MacQueen et al. 2003). Like in the L1 larvae, the general pattern of expression of NHJ-1::OLLAS is not affected by the loss of *cku-80* or *lig-4*, with NHJ-1::OLLAS remaining detectable above background levels in diplotene and diakinesis nuclei of *cku-80(tm1203)* and *lig-4(vv134)* animals (**Figure 3.8B-C**). The expression of NHJ-1::OLLAS in adult intestinal nuclei, where it is also prominently expressed, is likewise unaffected by *cku-80* or *lig-4* status (**Figure 3.8A-C**).

III.9 LIG-4 is most prominently expressed in intestinal nuclei at the L1 stage

I next wanted to investigate the subcellular localization of LIG-4, since this enzyme performs the terminal ligation step in the cNHEJ pathway, repairing the DSB and restoring chromosomal

integrity. For this, I introduced a C-terminal OLLAS tag in the endogenous *lig-4* locus in the N2 [R] genetic background. In contrast to NHJ-1::OLLAS, which is detectable in the majority of L1 nuclei, immunostaining revealed that LIG-4::OLLAS is strongly visible only in a small subset of cells located along the length of the larval body (**Figure 3.9A-B**). The pattern of LIG-4 expressing nuclei appeared most similar to that of intestinal cell nuclei. To test the hypothesis that LIG-4 expressing nuclei belong to intestinal cells, I crossed the OLLAS-tagged *lig-4* into a genetic background containing an integrated transgene bearing GFP-tagged *elt-2*. ELT-2 is a GATA transcription factor that is essential for the development and maintenance of the *C. elegans* intestine, and is expressed only in this tissue (Wiesenfahrt, Berg et al. 2016). Staining of *lig-4::OLLAS; elt-2::GFP* L1 larvae revealed that LIG-4::OLLAS was present in nuclei expressing ELT-2::GFP (**Figure 3.9C**), confirming the hypothesis that endogenous LIG-4 expression is predominantly intestinal. Like for NHJ-1::OLLAS, neither radiation treatment nor a lack of *cku-80* or *nhj-1* had a noticeable impact on the localization of LIG-4::OLLAS in the L1 larva (**Figure 3.9A-B**). Although the same caveats apply as with NHJ-1::OLLAS with regard to the lack of IR-dependent changes (see discussion in **Chapter IV**), these results demonstrate that LIG-4 does not depend on other cNHEJ factors for its subcellular localization.

III.10 LIG-4 is expressed in mid-to-late meiotic prophase I in the adult germline

The striking difference between the expression pattern of endogenously-expressed NHJ-1 and LIG-4 in the L1 prompted me to investigate LIG-4 localization in the adult germline. In this tissue, LIG-4::OLLAS localization is much more reminiscent of NHJ-1::OLLAS localization, being dispersed throughout the nucleoplasm in a punctate pattern and especially prominent in diakinesis and diplotene, although it becomes detectable earlier than NHJ-1::OLLAS; while some enrichment above background is seen in transition zone nuclei, the LIG-4::OLLAS signal is reliably seen beginning in the pachytene stage (**Figure 3.10A**). Like NHJ-1, LIG-4 expression in the adult germline is not affected by the loss of other cNHEJ factors, including *cku-80* (**Figure 3.10B**) and *nhj-1* (**Figure 3.10C**). Although not as extreme a difference as that observed in the L1 larva, it nevertheless raises the question of why NHJ-1 and LIG-4 would not have an overlapping pattern of expression within a tissue, given that both act in the same pathway.

III.11 NHJ-1 acts downstream of Ku in the adult germ cell cNHEJ

That NHJ-1 is part of the cNHEJ pathway is well supported by genetic evidence presented in section **III.5** (**Figure 3.4**). These observations, however, are not informative about the potential role which NHJ-1 could play in the cNHEJ process. The most obvious hypothesis would be that it participates directly in one of the three steps of cNHEJ (see **Chapter I Part II**). It could act together with the Ku ring in the initial DSB detection and tethering step; or regulate the activity of one or more of the processing factors about which nothing is known in *C. elegans*; or regulate the binding or ligation activity of LIG-4, even if it doesn't influence its subcellular localization. Beyond participating in the cNHEJ repair mechanism directly, NHJ-1 could act to provide the appropriate chromatin context for the initiation of cNHEJ, or to restore the chromatin or DNA configuration post-repair, such as removing the Ku rings trapped on DNA after ligation, which may otherwise inhibit the restoration of the proper chromatin structures or the progression of replication and transcription (Postow 2011).

The regulated activity of the cNHEJ pathway in the adult germline provided me with a system in which I could test two distinct hypotheses about the role of NHJ-1 in the cNHEJ process, with the caveat that they may apply only in this tissue context and not generally. As has been mentioned above, the protein COM-1 plays a critical role in antagonizing cNHEJ in the adult germline. In animals deficient for *com-1*, the current understanding is that the Ku ring is not prevented from engaging with SPO-11-generated DSBs and initiating repair via cNHEJ, which precludes HR-based repair and therefore crossing over (Lemmens, Johnson et al. 2013). The cytologically visible outcome of *com-1* mutations is a lack of six well-defined DAPI-staining bodies in diakinesis, which are normally indicative of properly formed bivalents (*ie*, homologous chromosomes held together by chiasmata). The DNA of *com-1* mutants is instead visible in diakinesis nuclei as disorganized masses, often fewer than 6 in number, which results in extremely low embryonic survival (<2% embryos hatching) (Lemmens, Johnson et al. 2013). This phenotype is significantly rescued in *com-1 cku-80* double mutants, where 30-40% of eggs hatch and become viable progeny, although the fact that the rescue is not complete points to an efficient operation of a different non-HR repair pathway in the absence of *com-1* (Lemmens, Johnson et al. 2013). In contrast to the *cku-80* rescue, the removal of *lig-4* in the *com-1* background does not rescue

embryonic lethality, which has been interpreted as a consequence of the Ku ring engaging with meiotic DSBs and blocking HR even in the absence of the terminal cNHEJ effector (Lemmens, Johnson et al. 2013). The loss of *lig-4* in a *com-1* background does result in an increase of small DNA fragments in diakinesis nuclei, with more than 40% of nuclei exhibiting at least one fragment; this observation has been interpreted as the outcome of incomplete cNHEJ repair in the absence of LIG-4 (Lemmens, Johnson et al. 2013). By introducing an *nhj-1* mutation in a *com-1* background, I could therefore test whether: 1) NHJ-1 acts together with or upstream of CKU-80, in which case the loss of *nhj-1* should rescue the embryonic lethality of *com-1* in a manner similar to *cku-80 com-1* double mutants; and 2) NHJ-1 acts downstream of Ku but in the cNHEJ repair process *per se*, in which case the loss of *nhj-1* in the *com-1* background would mimic the loss of *lig-4* and show a high incidence of DNA fragments.

To test these predictions, I constructed *com-1(t1626); nhj-1(vv144)* and *com-1(t1626) lig-4(vv134)* double mutants. Instead of a genetic *cku-80* mutant, I employed RNAi against *cku-80*, which I performed in all three of the above genotypes both in order to verify that I could replicate the published effects of loss of *cku-80*, as well as to test whether the loss of *nhj-1* might affect the *cku-80* knockdown-mediated rescue of *com-1*. I observed no rescue of embryonic survival in *com-1 lig-4* double mutants (hatching rate of 0.5%) and *com-1 nhj-1* double mutants (0% hatching rate), compared to *com-1* single mutants (0.3% hatching rate), but a significant rescue in *com-1* mutants fed with *cku-80(RNAi)*, which had a hatching rate of 26.4% ($p<0.001$) (**Figure 3.11A**). This result suggests that NHJ-1 is not required for the Ku ring to engage the free DNA ends, according to the model proposed in (Lemmens, Johnson et al. 2013). That the rate of rescue by *cku-80(RNAi)* was not different between *com-1* single mutants (26.4% of embryos hatching) and *com-1 nhj-1* double mutants (27.7% hatching embryos; $p>0.05$) (**Figure 3.11A**) also argues against the possibility that NHJ-1 may engage some of the free DSBs independently of Ku, in which case a stronger rescue would be expected in the double mutant. Thus NHJ-1 appears to act downstream of Ku, at least in the context of the adult germline. While DAPI-staining bodies in diakinesis nuclei of *com-1, com-1 lig-4*, and *com-1; nhj-1* mutants displayed a range of aberrant morphologies, ranging from single aggregates to presence of more than six distinct bodies (**Figure 3.11B**), I was unable to verify the presence of DNA fragment in either *com-1; nhj-1* mutants or *com-1 lig-4*

mutants. The failure to replicate this finding from (Lemmens, Johnson et al. 2013) could have been due to differences in the imaging systems used, but it made the comparison based on DNA fragmentation impossible. However, I did observe that the total number of DAPI-staining bodies in the diakinesis nuclei of *com-1* mutants (median of 3 bodies, IQR 1) was significantly lower than that of *com-1 lig-4* double mutants (median of 4 bodies, IQR 3; $p<0.05$) and *com-1; nhj-1* double mutants (median of 5 bodies, IQR 3; $p<0.001$) (**Figure 3.11C**), suggesting a higher capacity for chromosome aggregation in *com-1* single mutants. This finding is consistent with impaired cNHEJ in *com-1 lig-4* and *com-1; nhj-1* double mutants, and represents further evidence that cNHEJ repair cannot occur without NHJ-1.

III.12 Conclusion

In this chapter, I presented the story of the discovery of a mutation in *nhj-1,* a previously uncharacterized gene encoding two short protein isoforms, as the cause of IR-sensitivity observed in N2 [S] animals which has been detailed in the previous chapter. I have also presented data demonstrating that *nhj-1* is a component of the canonical non-homologous end joining DSB repair pathway, and that it acts downstream of the Ku ring. Finally, I have provided a description of the distinct expression patterns of endogenous NHJ-1 and LIG-4 in L1 larvae and the adult germline, which raise further questions about the tissue-level regulation of cNHEJ and the role of NHJ-1. In the following chapter, the implications of these findings and those presented in **Chapter II** are discussed, as well as the directions future research may take in pursuing some of the unsolved questions this work has raised.

Figures

Figure 3.1

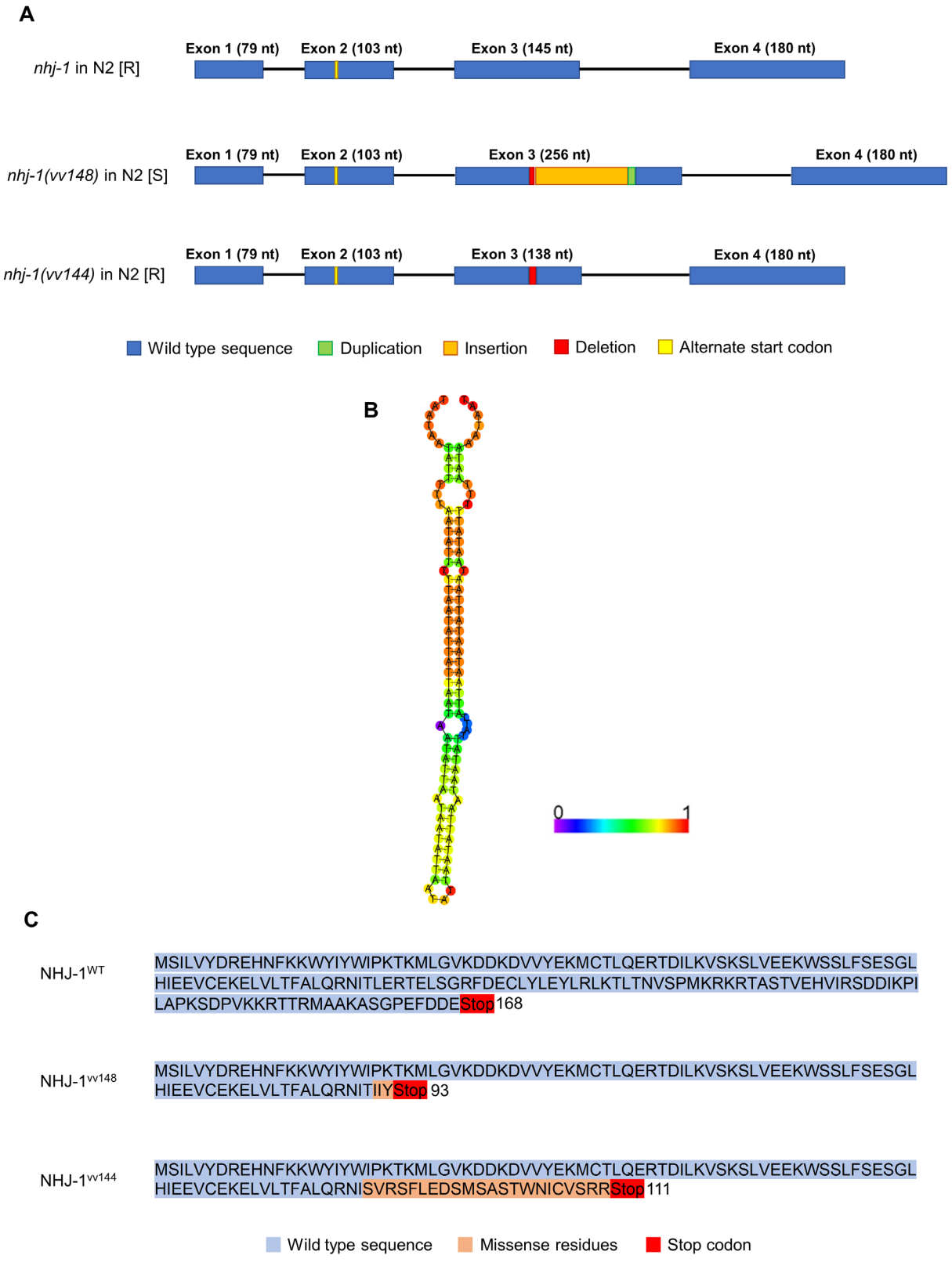


Figure 3.1 - The structure of NHJ-1

(A) The structure of the coding region of *nhj-1*/*H19N07.3.* The uncharacterized gene *H19N07.3*, which I have named *nhj-1* (non-homologous end joining 1), is composed of four exonic regions and three introns. A shorter protein isoform can be translated from an alternate start codon in exon 2. In the N2 [S] background, exon 3 of *nhj-1* has been disrupted by a deletion of 5 nucleotides, and a 115 bp insertion composed of 107 nucleotides of unknown origin (see (**B)**) and 8 nucleotides duplicated from the exonic sequence. I have designated this mutation *nhj-1(vv148).* To test the role of *nhj-1* in IR-sensitivity, I used CRISPR mutagenesis to delete 7 nucleotides from Exon 3 and create the *nhj-1(vv144)* allele.

(B) The predicted secondary structure of the 107 bp insertion in the *nhj-1(vv148)* allele. This sequence of unknown origin is predicted to form a hairpin by the RNAfold tool of the ViennaRNA Package.

(C) The predicted protein sequences of NHJ-1. The wild-type long isoform of the NHJ-1 protein is 168 amino acids long, with no conserved domains. The shorter isoform is 130 residues in length. The *nhj-1(vv148)* indel results in truncated protein products of $93^{long}/55^{short}$ amino acids in total length, with a frameshift producing 3 missense residues after residue $90^{long}/52^{short}$. The *nhj-1(vv144)* deletion results in a frameshift after residue $89^{long}/51^{short}$, which creates a downstream sequence of 22 missense residues before terminating in a stop codon, and produces final products $111^{long}/73^{short}$ amino acids long.

bp = base pair
nt = nucleotide
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
nhj-1(vv144) [R] = an *nhj-1* deletion mutant, generated by CRISPR in the N2 [R] genetic background
nhj-1(vv148) [S] = an *nhj-1* mutant containing a indel that occurred naturally in the N2 [S] background

Figure 3.2

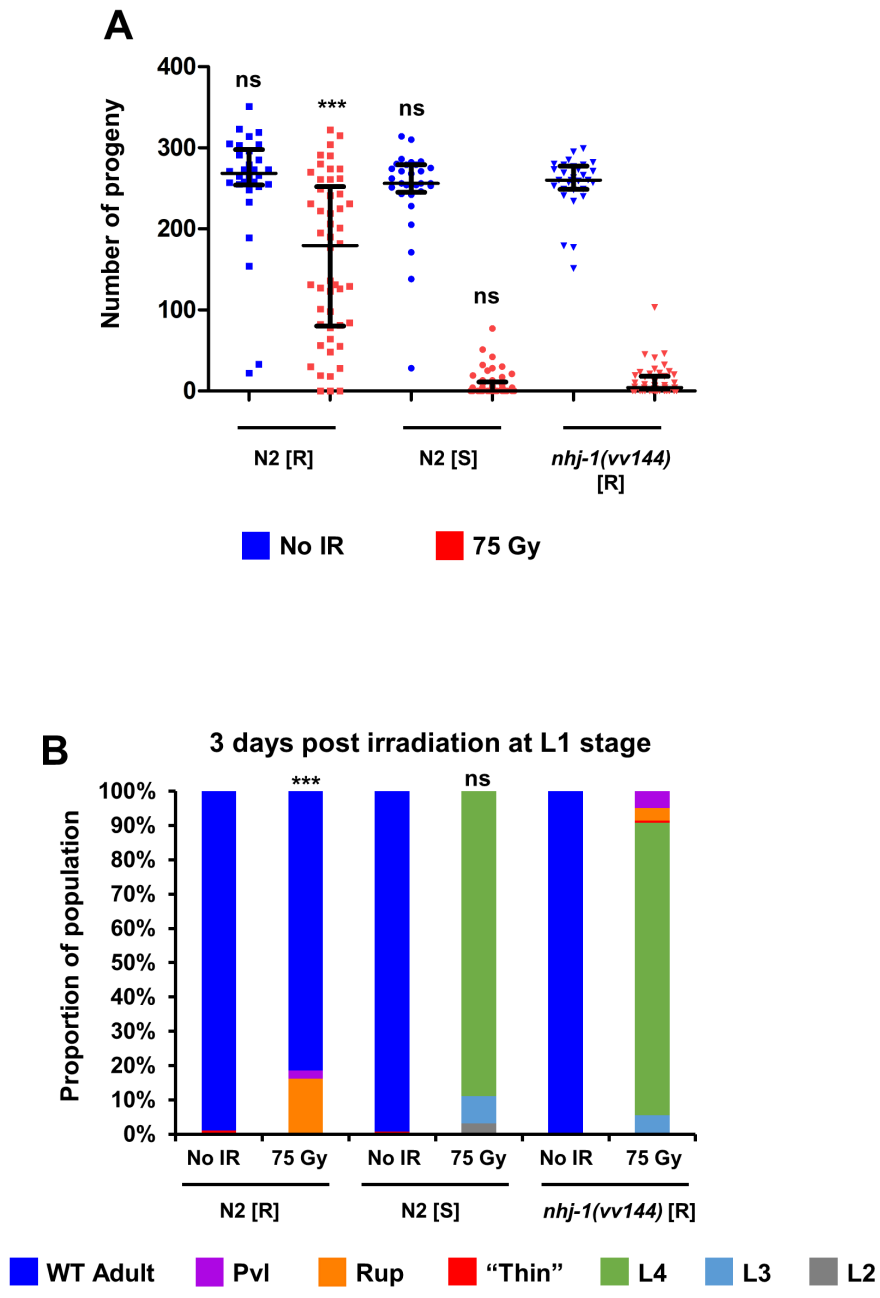

A
Number of progeny
400
300
200
100
0
ns
***
ns
ns
N2 [R]
N2 [S]
nhj-1(vv144) [R]
No IR
75 Gy
B
3 days post irradiation at L1 stage
Proportion of population
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
***
ns
No IR
75 Gy
No IR
75 Gy
No IR
75 Gy
N2 [R]
N2 [S]
nhj-1(vv144) [R]
WT Adult
Pvl
Rup
"Thin"
L4
L3
L2

Figure 3.2 - Loss of *nhj-1* in the N2 [R] background results in IR-sensitivity

(A) Total brood size quantification of N2 [R], N2 [S], and *nhj-1(vv144)*. While the *nhj-1(vv144)* deletion has no effect on untreated brood size ($p>0.05$ against both N2 [R] and N2 [S]), it significantly reduces the post-IR brood size of N2 [R] background compared ($p<0.001$ vs N2 [R] post-IR) to the same level as that of N2 [S] ($p>0.05$). All statistical comparisons shown in the figure are to *nhj-1(vv144)* from the corresponding treatment group (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 30 for all unirradiated groups, 50 for irradiated N2 [R] and irradiated *nhj-1(vv144)*, and 49 for irradiated N2 [S].

(B) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(A)**. Three days after IR treatment, *nhj-1(vv144)* mutants show a strong Gro phenotype, with almost all animals still in the L4 stage, like in N2 [S] ($p>0.05$), but significantly different than N2 [R] ($p<0.001$) in which all animals have molted into adults. All statistical comparisons shown in the figure are to irradiated *nhj-1(vv144)* (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.01$). Sample size (n) is 195 for unirradiated N2 [R], 161 for irradiated N2 [R], 252 for unirradiated N2 [S], 163 for irradiated N2 [S], 240 for unirradiated *nhj-1(vv144)*, and 163 for irradiated *nhj-1(vv144)*.

IR = ionizing radiation
Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L4/L3/L2 = larva of the L4/L3/L2 stage
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
nhj-1(vv144) [R] = an *nhj-1* deletion mutant, generated by CRISPR in the N2 [R] genetic background

ns = not significant ($p>0.05$ in **(A)**; $p>0.01$ in **(B)**), *** = $p<0.001$

Figure 3.3

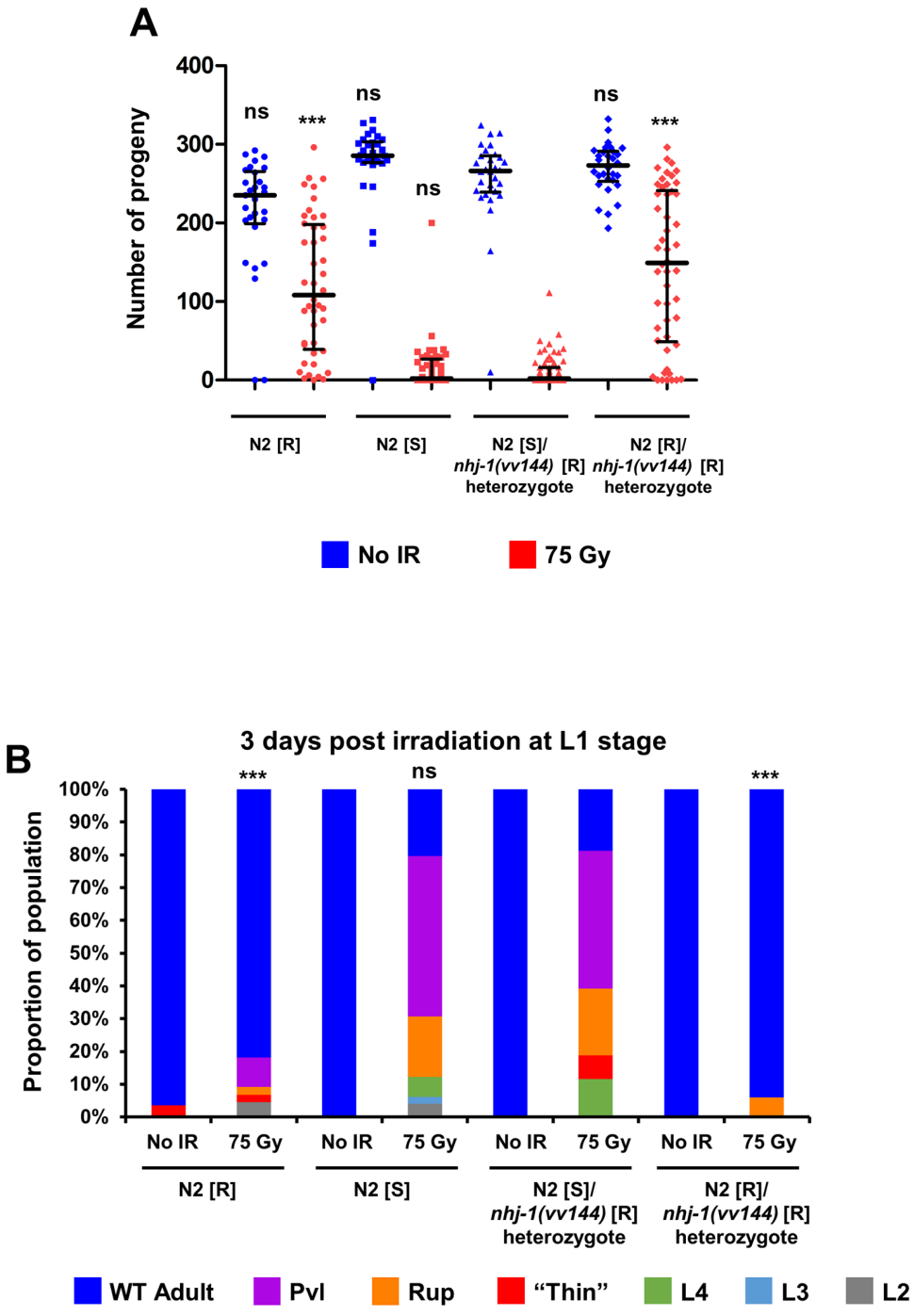

A
Number of progeny
400
300
200
100
0
ns
***
ns
ns
ns
***
N2 [R]
N2 [S]
N2 [S]/ nhj-1(vv144) [R] heterozygote
N2 [R]/ nhj-1(vv144) [R] heterozygote
No IR
75 Gy
B
3 days post irradiation at L1 stage
Proportion of population
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
***
ns
***
No IR
75 Gy
N2 [R]
N2 [S]
N2 [S]/ nhj-1(vv144) [R] heterozygote
N2 [R]/ nhj-1(vv144) [R] heterozygote
WT Adult
Pvl
Rup
"Thin"
L4
L3
L2

Figure 3.3 - The N2 [S] genome does not complement *nhj-1(vv144)*

(A) Total brood size quantification of N2 [R], N2 [S], and N2 [S]/*nhj-1(vv144)* and N2 [R]/*nhj-1(vv144)* heterozygotes. The post-IR brood size of N2 [S]/*nhj-1(vv144)* heterozygotes is not significantly different than that of post-IR N2 [S] animals ($p>0.05$), while both are significantly reduced compared to the post-IR brood size of either N2 [R] animals or N2 [R]/*nhj-1(vv144)* heterozygotes ($p<0.001$ for both comparisons), indicating that the IR-sensitivity of the N2 [S] line is caused by a loss of function in *nhj*-1. All statistical comparisons shown in the figure are to N2 [S]/*nhj-1(vv144)* heterozygotes from the corresponding treatment group (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 29 for unirradiated N2 [R], 44 for irradiated N2 [R], 30 for unirradiated N2 [S], 48 for irradiated N2 [S], 30 for unirradiated N2 [S]/*nhj-1* heterozygote, 68 for irradiated N2 [S]/*nhj-1(vv144)* heterozygote, 30 for unirradiated N2 [R]/*nhj-1* heterozygote, and 50 for irradiated N2 [R]/*nhj-1(vv144)* heterozygote**.**

(B) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(A)**. The incidence of Gro and vulval phenotypes is not significantly different between N2 [S]/*nhj-1(vv144)* heterozygotes and N2 [S] animals after irradiation ($p>0.05$), while these phenotypes are significantly less common in post-IR N2 [R] animals and N2 [R]/*nhj-1(vv144)* heterozygotes ($p>0.001$ against both groups). All statistical comparisons shown in the figure are to N2 [S]/*nhj-1(vv144)* heterozygotes from the corresponding treatment group (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.008$). Sample size (n) is 28 for unirradiated N2 [R], 44 for irradiated N2 [R], 29 for unirradiated N2 [S], 49 for irradiated N2 [S], 30 for unirradiated N2 [S]/*nhj-1* heterozygote, 69 for irradiated N2 [S]/*nhj-1(vv144)* heterozygote, 30 for unirradiated N2 [R]/*nhj-1* heterozygote, and 50 for irradiated N2 [R]/*nhj-1(vv144)* heterozygote**.**

IR = ionizing radiation
Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L4/L3/L2 = larva of the L4/L3/L2 stage
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
nhj-1(vv144) [R] = an *nhj-1* deletion mutant, generated by CRISPR in the N2 [R] genetic background

ns = not significant ($p>0.05$ in **(A)**; $p>0.008$ in **(B)**), *** = $p<0.001$

Figure 3.4

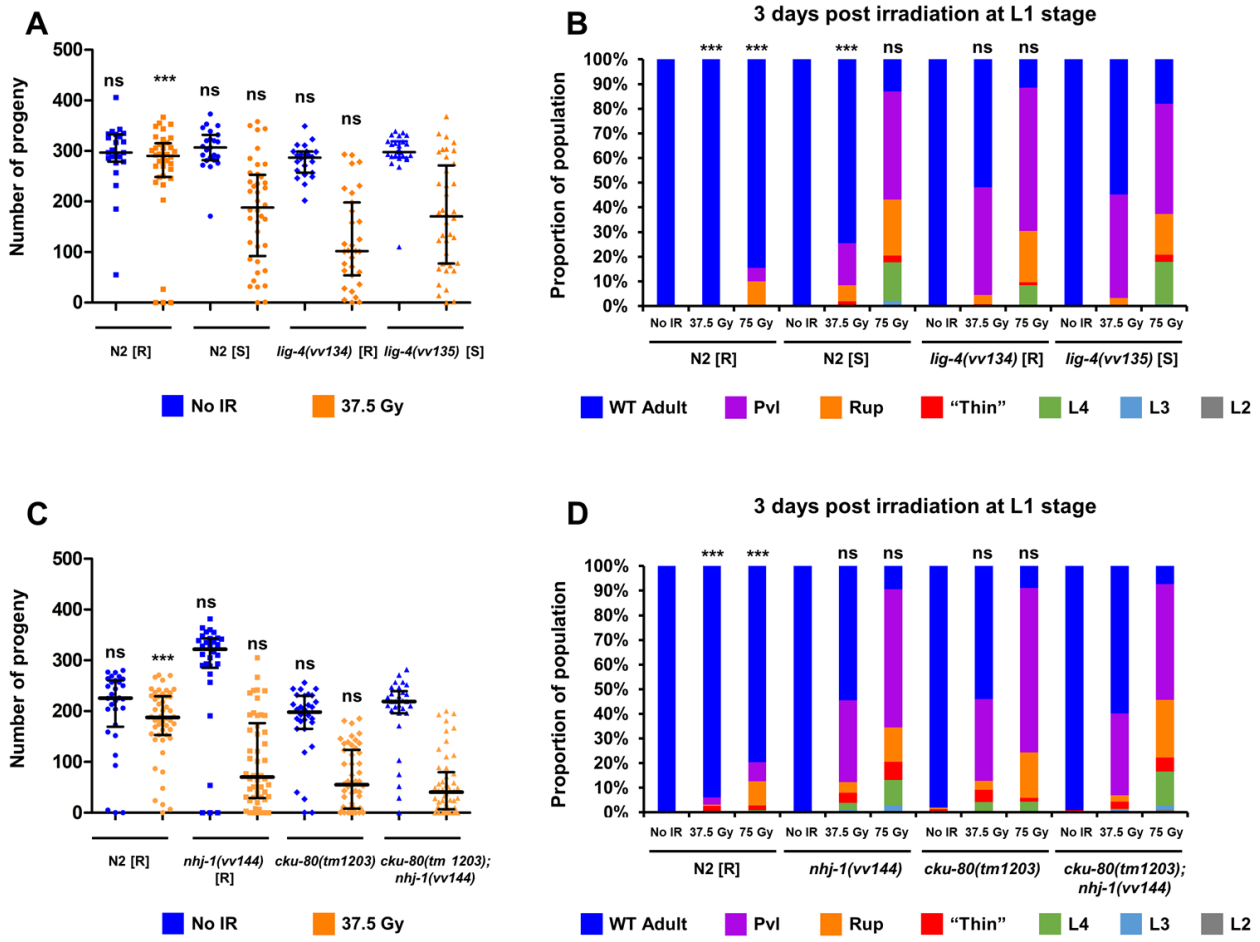


Figure 3.4 - NHJ-1 acts in the cNHEJ pathway

(A) Total brood size quantification of N2 [R], N2 [S], *lig-4(vv134)* [R], and *lig-4(vv141)* [S] hermaphrodites. The post-IR brood size of *lig-4(vv141)* [S] animals, which harbour the same inactivating mutation as *lig-4(vv134)* [R] animals except in the sensitive genetic background, is not significantly different than either N2 [S] or *lig-4(vv134)* [R] animals ($p>0.05$). The lack of additive IR-sensitivity strongly suggests that N2 [S] is IR-sensitive because of a loss of cNHEJ activity. All statistical comparisons shown in the figure are to *lig-4* [S] animals from the corresponding treatment group (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 23 for unirradiated N2 [R], 36 for irradiated N2 [R], 23 for unirradiated N2 [S], 40 for irradiated N2 [S], 23 for unirradiated *lig-4* [R], 31 for irradiated *lig-4* [R], 24 for unirradiated *lig-4* [S], and 40 for irradiated *lig-4* [S].

(B) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(A)**. The incidence of somatic phenotypes in *lig-4* [S] is not significantly different from either *lig-4* [R] or N2 [S] following either 37.5 Gy or 75 Gy of IR ($p>0.05$ for all comparisons), showing that the *lig-4* mutation

Figure 3.4 - NHJ-1 acts in the cNHEJ pathway (*continued*)

(B, *continued*) and the N2 [S] background are not additive with respect to IR-associated somatic phenotypes. All statistical comparisons shown in the figure are to *lig-4* [S] animals from the corresponding treatment group (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.008$). Sample size (n) is 146/201/181 for N2 [R] No IR/37.5 Gy/75 Gy, 188/106/146 for N2 [S] No IR/37.5 Gy/75 Gy, 208/131/167 for *lig-4* [R] No IR/37.5 Gy/ 75 Gy, and 177/134/125 for *lig-4* [S] No IR/37.5 Gy/ 75 Gy.

(C) Total brood size quantification of N2 [R], *nhj-1(vv144)* [R], *cku-80(tm1203)*, and *cku-80(tm1203); nhj-1(vv144)*. Double mutants of *cku-80(tm1203)* and *nhj-1(vv144)* do not exhibit a significantly different post-IR brood size than either single mutant ($p>0.05$ for both), showing that NHJ-1 functions in the same pathway as CKU-80. All statistical comparisons shown in the figure are to *cku-80(tm1203); nhj-1(vv144)* animals from the corresponding treatment group (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 23 for unirradiated N2 [R], 36 for irradiated N2 [R], 23 for unirradiated N2 [S], 40 for irradiated N2 [S], 23 for unirradiated *lig-4* [R], 31 for irradiated *lig-4* [R], 24 for unirradiated *lig-4* [S], and 40 for irradiated *lig-4* [S].

(D) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(C)**. Vulval and slow growth phenotypes do not have a significantly different incidence in the double mutant and either single mutant ($p>0.05$ for all comparisons), supporting the conclusion of CKU-80 and NHJ-1 acting in the same pathway. All statistical comparisons shown in the figure are to *cku-80(tm1203); nhj-1(vv144)* animals from the corresponding treatment group (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.008$). Sample size (n) is 181/194/103 for N2 [R] No IR/37.5 Gy/75 Gy, 156/211/107 for *nhj-1(vv144)* [R] No IR/37.5 Gy/75 Gy, 156/163/136 for *cku-80(tm1203)* No IR/37.5 Gy/75 Gy, and 225/274/175 for *cku-80(tm1203); nhj-1(vv144)* No IR/37.5 Gy/75 Gy.

IR = ionizing radiation; Gy = Gray (unit)
Pvl = protruding vulva phenotype; Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva; L4/L3/L2 = larva of the L4/L3/L2 stage
N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
nhj-1= *nhj-1(vv144)* [R], a deletion mutant generated by CRISPR in the N2 [R] genetic background
lig-4 [S] and *lig-4* [R] = *lig-4(vv141[R18STOP])* [S] and *lig-4[vv134[R18STOP])*, premature stop codon mutations in the *lig-4* sequence introduced in the N2 [S] or N2 [R] backgrounds
cku-80 = *cku-80(tm1203)*, a published deletion allele of *cku-80*.

ns = not significant ($p>0.05$ in **(A)**; $p>0.008$ in **(B)**), *** = $p<0.001$

Figure 3.5

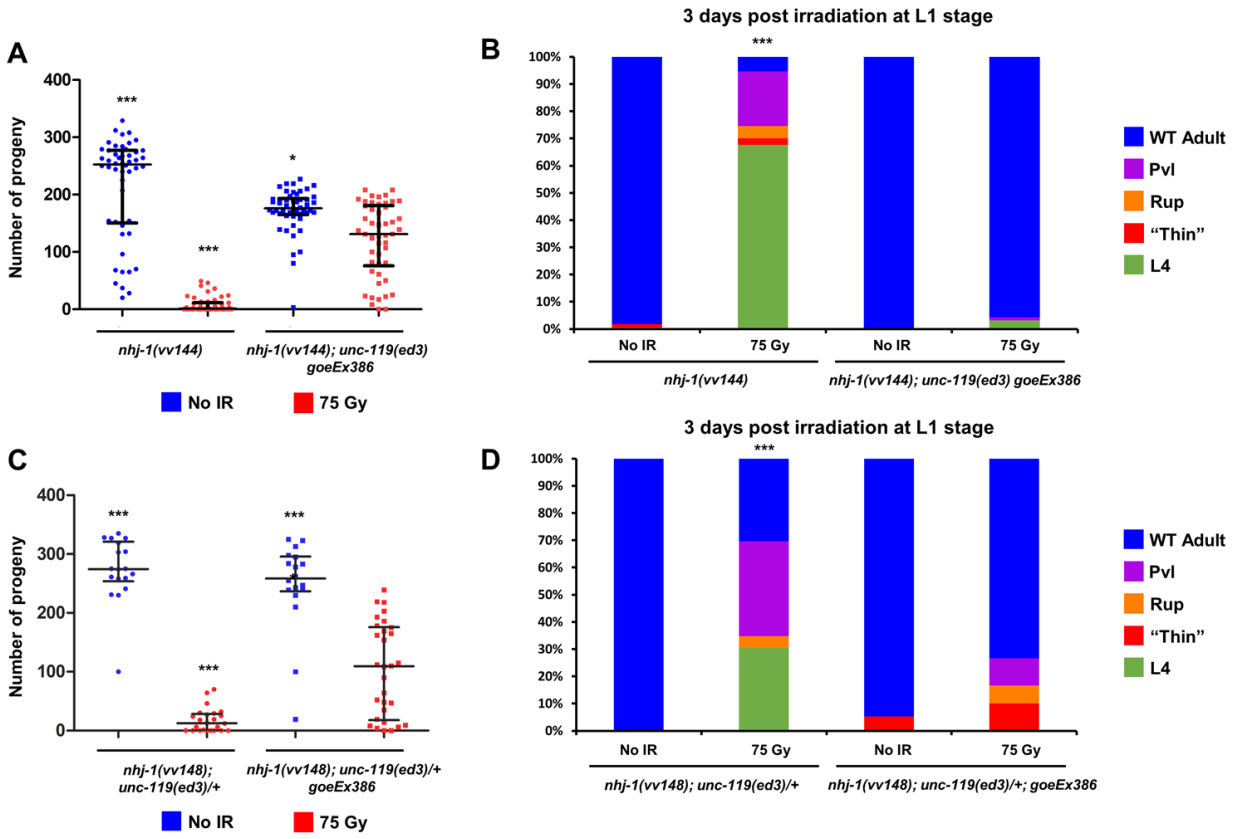


Figure 3.5 - Extrachromosomal *nhj-1* rescues both brood size and somatic IR phenotypes of *nhj-1(vv144)*

(A) Total brood size quantification of *nhj-1(vv144)* and *nhj-1(vv144); unc-119(ed3) goeEx386*. The post-IR brood size of *nhj-1(vv144); unc-119(ed3) goeEx386* animals, in which the *nhj-1(vv144)* allele and the *unc-119(ed3)* allele have been rescued by and extrachromosomal array carrying a GFP-tagged copy of wild type *nhj-1* and a wild type copy of *unc-119*, is significantly higher than that of *nhj-1(vv144)* animals ($p<0.005$), and is also reduced compared to unirradiated *nhj-1(vv144); unc-119(ed3) goeEx386* controls. Extrachromosomal *nhj-1* is thus able to rescue the post-IR brood size phenotype of *nhj-1(vv144).* All statistical comparisons shown in the figure are to irradiated *nhj-1(vv144); unc-119(ed3) goeEx386* animals (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 50 for all groups.

(B) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(A)**. The Gro and vulval phenotypes are significantly less prevalent in irradiated *nhj-1(vv144); unc-119(ed3) goeEx386* animals compared to the irradiated *nhj-1(vv144)* group (<0.001), showing that exogenous *nhj-1* rescues the post-IR defects of *nhj-1(144)*. The statistical comparison shown in the figure is to irradiated

Figure 3.5 - Extrachromosomal *nhj-1* rescues both brood size and somatic IR phenotypes of *nhj-1(vv144)* (*continued*)

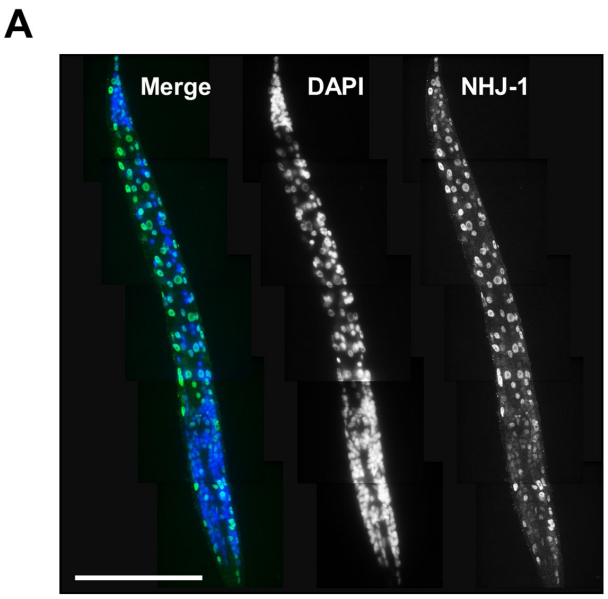
(B, *continued*) *nhj-1(vv144); unc-119(ed3) goeEx386* animals (Chi-squared test, Bonferroni corrected for multiple comparisons to α = 0.008). Sample size (n) is 220 for unirradiated *nhj-1(vv144)* NR, 130 for irradiated *nhj-1(vv144)*, 70 for unirradiated *nhj-1(vv144); unc-119(ed3) goeEx386*, and 93 for irradiated *nhj-1(vv144); unc-119(ed3) goeEx386.*

(C) Total brood size quantification of *nhj-1(vv148); unc-119(ed3)/+* and *nhj-1(vv148); unc-119(ed3)/+ goeEx386*. In *nhj-1(vv148); unc-119(ed3)/+ goeEx386* animals, the post-IR brood size is significantly rescued compared to *nhj-1(vv148); unc-119(ed3)/+* animals which do not carry the rescuing transgene ($p<0.05$), corroborating the conclusion that extrachromosomal *nhj-1* can rescue a lack of endogenous *nhj-1.* All statistical comparisons shown in the figure are to irradiated *nhj-1(vv148); unc-119(ed3)/+ goeEx386* animals (Kruskal-Wallis test, followed by Dunn's post-hoc test). Error bars represent the median and interquartile range. Sample size (n) is 18 for both unirradiated groups, 24 for irradiated *nhj-1(vv148); unc-119(ed3)/+*, and 30 for irradiated *nhj-1(vv148); unc-119(ed3)/+ goeEx386*.

(D) Quantification of post-IR somatic phenotypes three days after IR treatment in the same groups as in **(C)**. Vulval phenotypes and slow growth have a lower incidence in irradiated *nhj-1(vv148); unc-119(ed3)/+ goeEx386* animals compared to irradiated *nhj-1(vv148); unc-119(ed3)/+* animals (<0.001), in line with the brood size results. The statistical comparison shown in the figure is to irradiated *nhj-1(vv148); unc-119(ed3)/+; goeEx386* animals (Chi-squared test, Bonferroni corrected for multiple comparisons to α = 0.008). Sample size (n) is 18 for unirradiated *nhj-1(vv148); unc-119(ed3)/+*, 19 for unirradiated *nhj-1(vv148); unc-119(ed3)/+ goeEx386*, 23 for irradiated *nhj-1(vv148); unc-119(ed3)/+*, and 30 for irradiated *nhj-1(vv148); unc-119(ed3)/+ goeEx386*.

IR = ionizing radiation; Gy = Gray (unit)
Pvl = protruding vulva phenotype
Rup = ruptured through vulva phenotype
"Thin" = thin, whitish L3-like larva
L4 = larva of the L4 stage
nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background
goeEx386 = an extrachromosomal array carrying a tagged version of *nhj-1/H19N07.3* and a wild type copy of *unc-119*, with the genotype *[WRM0635D_B04(pRedFlp-Hgr) (H19N07.3[21364]::S0001_pR6K_Amp_2xTY1ce_EGFP_FRT_rpsl_neo_FRT_3xFlag)dFRT::unc-119-Nat].*
ns = not significant ($p>0.05$ in **(A)**; $p>0.008$ in **(B)**), * = $p<0.05$, *** = $p<0.001$

Figure 3.6


A
Merge
DAPI
NHJ-1
Scale bar = 50 µm

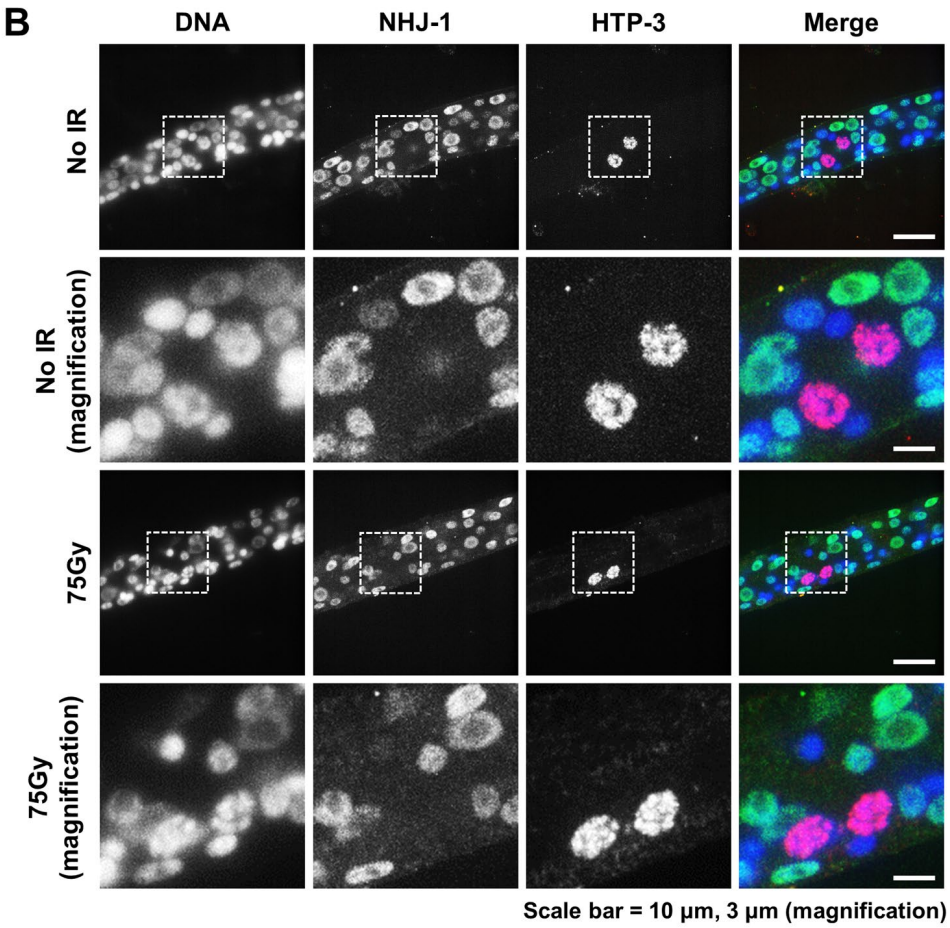

B
DNA
NHJ-1
HTP-3
Merge
No IR
No IR (magnification)
75Gy
75Gy (magnification)
Scale bar = 10 µm, 3 µm (magnification)

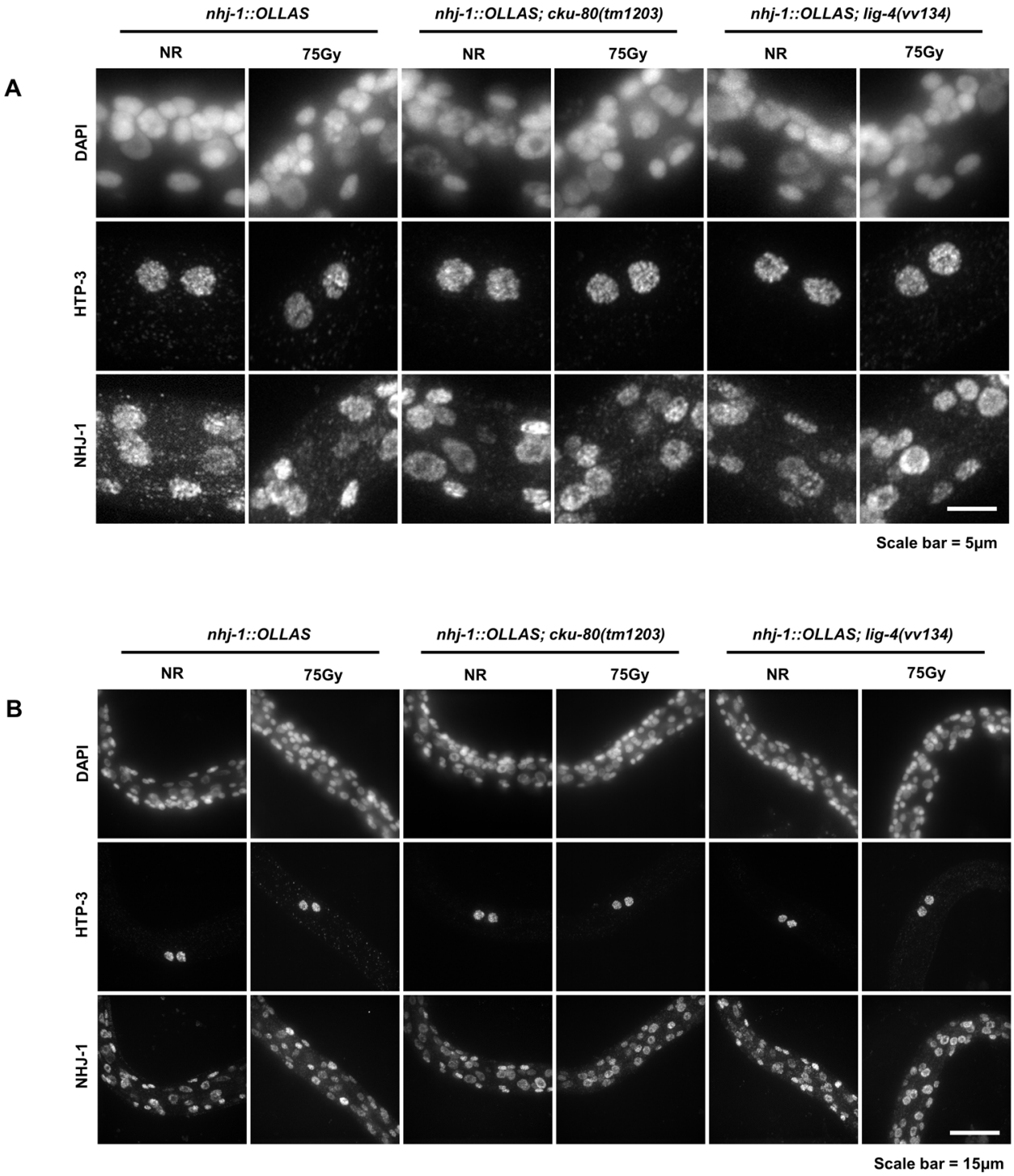
Figure 3.6 - Localization of extrachromosomal NHJ-1::GFP in the L1 larva

(A) A representative image showing the localization of NHJ-1::GFP in nuclei along the entire body of the L1 larva.

(B) Representative micrographs showing the subcellular localization of NHJ-1::GFP expressed from *goeEx386* in control and irradiated L1 larvae. Expression of NHJ-1::GFP is nuclear, as demonstrated by the overlap with the DNA stain DAPI, but not every cell expresses the protein. Notably, it is not expressed in the primordial germ cells (PGCs), which are marked by HTP-3. Ionizing radiation treatment does not alter the expression pattern of NHJ-1::GFP. Area of digital magnification is indicated by dotted line squares.

goeEx386 = an extrachromosomal array carrying a tagged version of *nhj-1/H19N07.3* and a wild type copy of *unc-119*, with the genotype *[WRM0635D_B04(pRedFlp-Hgr) (H19N07.3[21364]::S0001_pR6K_Amp_2xTY1ce_EGFP_FRT_rpsl_neo_FRT_3xFlag)dFRT::unc-119-Nat].*
DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody
NHJ-1 = signal from an anti-GFP antibody

Figure 3.7


A
nhj-1::OLLAS
nhj-1::OLLAS; cku-80(tm1203)
nhj-1::OLLAS; lig-4(vv134)
NR
75Gy
NR
75Gy
NR
75Gy
DAPI
HTP-3
NHJ-1
Scale bar = 5μm
B
nhj-1::OLLAS
nhj-1::OLLAS; cku-80(tm1203)
nhj-1::OLLAS; lig-4(vv134)
NR
75Gy
NR
75Gy
NR
75Gy
DAPI
HTP-3
NHJ-1
Scale bar = 15μm

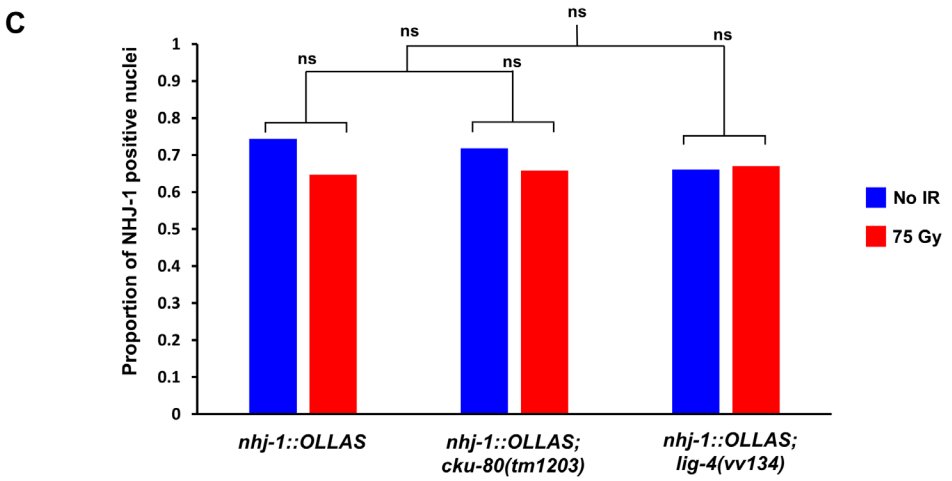


Figure 3.7 - Endogenous NHJ-1 localization in the L1 larva

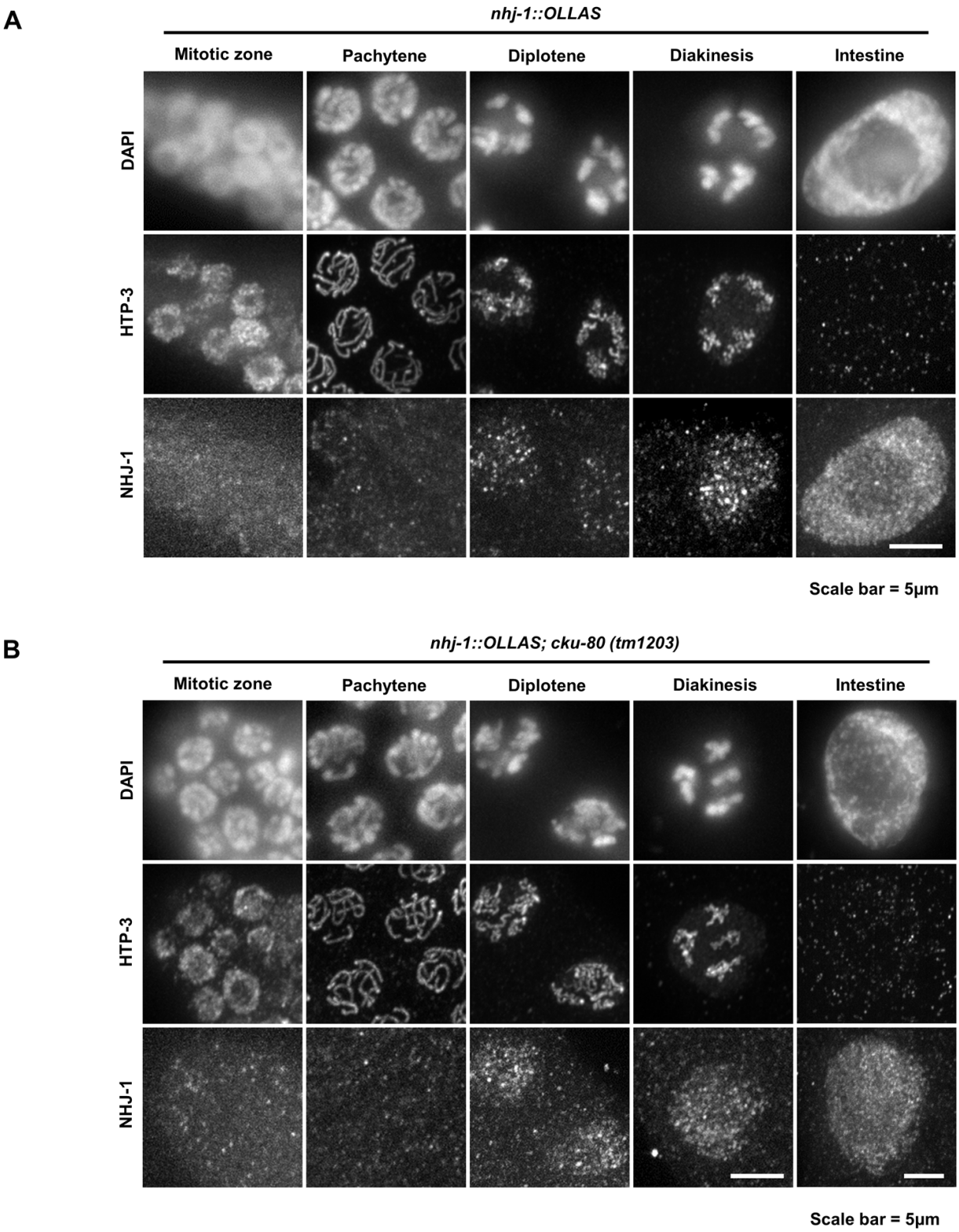
(A) Representative micrographs showing the subcellular localization of NHJ-1::OLLAS from the endogenous locus, together with DNA staining (DAPI) and the germline marker HTP-3, in the L1 larva. The expression of NHJ-1 from the endogenous locus is similar to that from the *goeEx386* array (**Figure 3.6**). The loss of *cku-80* or *lig-4* does not detectably affect the localization of NHJ-1::OLLAS, and neither does the radiation treatment in either the control or *cku-80* or *lig-4* mutant backgrounds.

(B) Representative micrographs showing the subcellular localization of NHJ-1::OLLAS in L1 larvae in the same genotypes and conditions as in **(A)**, showing a wider field of view for comparison.

(C) Quantification of the proportion of cells that express NHJ-1::OLLAS from the endogenous locus in the L1. Irradiation has no significant effect on the proportion of cells expressing NHJ-1::OLLAS, and neither does an absence of CKU-80 or LIG-4 ($p>0.3$ for all comparisons; Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.01$). Sample size is 187 nuclei in unirradiated *nhj-1::OLLAS*, 190 nuclei in irradiated *nhj-1::OLLAS*, 188 nuclei in unirradiated *nhj-1::OLLAS; cku-80,* 196 nuclei in irradiated *nhj-1::OLLAS; cku-80,* 177 nuclei in unirradiated *nhj-1::OLLAS; lig-4*, and 185 nuclei in irradiated *nhj-1::OLLAS; lig-4.* The nuclei were scored in 3 L1 larvae in all genotypes and conditions.

nhj-1::OLLAS = a C-terminal tag of the endogenous *nhj-1* locus with one copy of the OLLAS epitope
cku-80 = *cku-80(tm1203)*, a deletion allele of *cku-80*
lig-4 = *lig-4(vv134[R18STOP])* [R], a predicted null allele of *lig-4* generated in the N2 [R] genetic background
DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody
NHJ-1 = signal from an anti-OLLAS antibody
ns = not significant ($p>0.01$ in **(C)**)

Figure 3.8


A
nhj-1::OLLAS
Mitotic zone
Pachytene
Diplotene
Diakinesis
Intestine
DAPI
HTP-3
NHJ-1
Scale bar = 5μm
B
nhj-1::OLLAS; cku-80 (tm1203)
Mitotic zone
Pachytene
Diplotene
Diakinesis
Intestine
DAPI
HTP-3
NHJ-1
Scale bar = 5μm

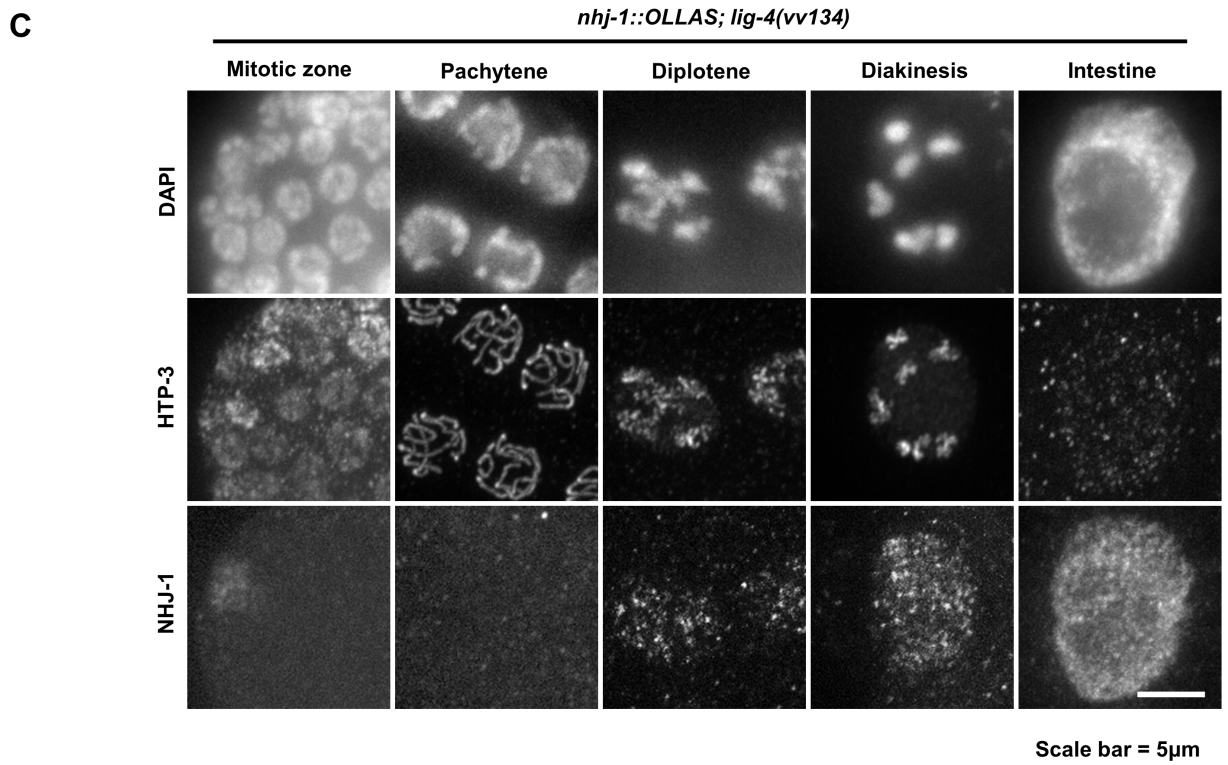


Figure 3.8 - Endogenous NHJ-1 localization in the adult germline

(A) Representative micrographs of NHJ-1::OLLAS expression from the endogenous locus in adult germline cells. Punctate nuclear expression of NHJ-1::OLLAS becomes reliably visible in diplotene, but is not chromatin associated and remains detectable in diakinesis. Adult intestinal cell shown for comparison.

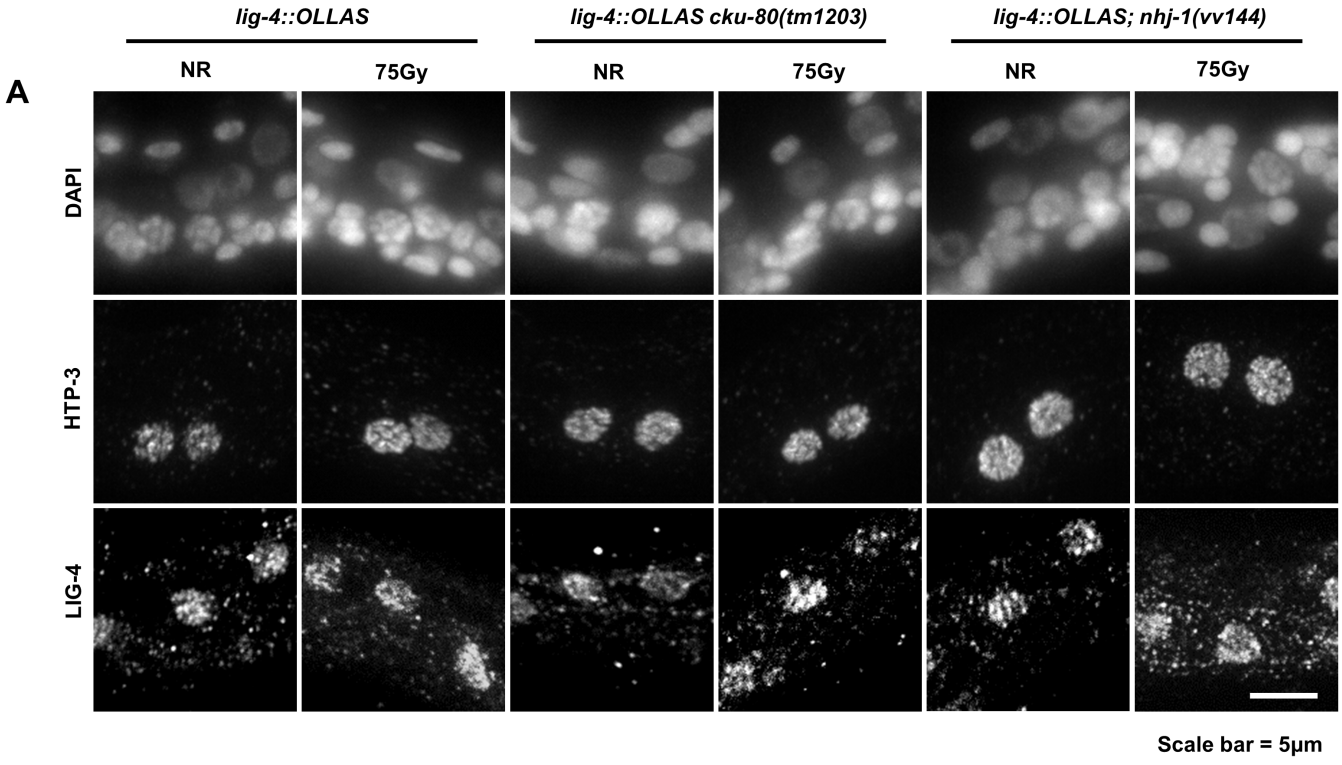
(B) Representative micrographs of NHJ-1::OLLAS expression from the endogenous locus in adult germline cells in animals deficient for *cku-80.* The loss of CKU-80 does not perturb the localization of NHJ-1::OLLAS either in the germline or in the intestine.

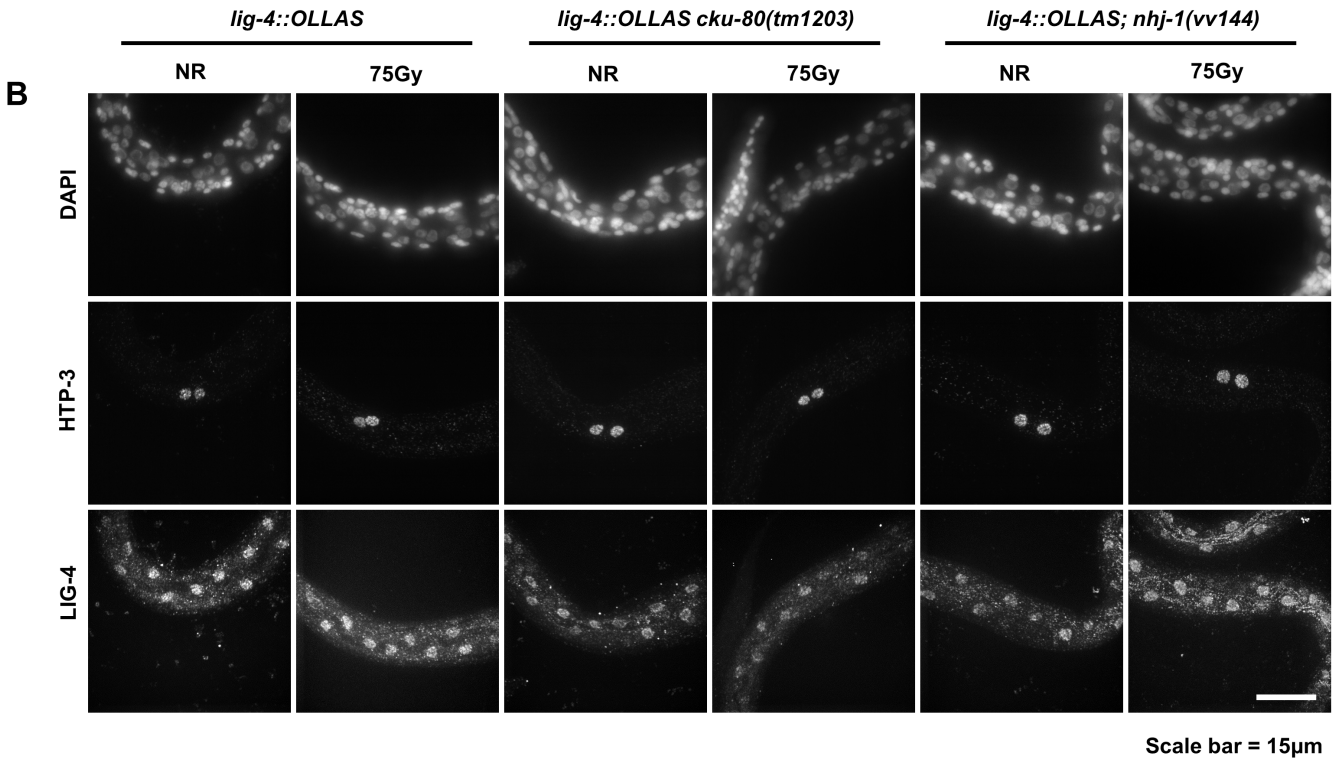
(C) Representative micrographs of NHJ-1::OLLAS expression from the endogenous locus in adult germline cells in animals deficient for *lig-4.* Like the loss of CKU-80, the loss of LIG-4 does not affect the pattern of NHJ-1::OLLAS expression either in the germ cells or intestinal cells.

nhj-1::OLLAS = a C-terminal tag of the endogenous *nhj-1* locus with one copy of the OLLAS epitope
cku-80 = *cku-80(tm1203)*, a deletion allele of *cku-80*
lig-4 = *lig-4(vv134[R18STOP])* [R], a predicted null allele of *lig-4* generated in the N2 [R] genetic background

DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody
NHJ-1 = signal from an anti-OLLAS antibody

Figure 3.9


A
lig-4::OLLAS
lig-4::OLLAS cku-80(tm1203)
lig-4::OLLAS; nhj-1(vv144)
NR
75Gy
NR
75Gy
NR
75Gy
DAPI
HTP-3
LIG-4
Scale bar = 5μm


B
lig-4::OLLAS
lig-4::OLLAS cku-80(tm1203)
lig-4::OLLAS; nhj-1(vv144)
NR
75Gy
NR
75Gy
NR
75Gy
DAPI
HTP-3
LIG-4
Scale bar = 15μm

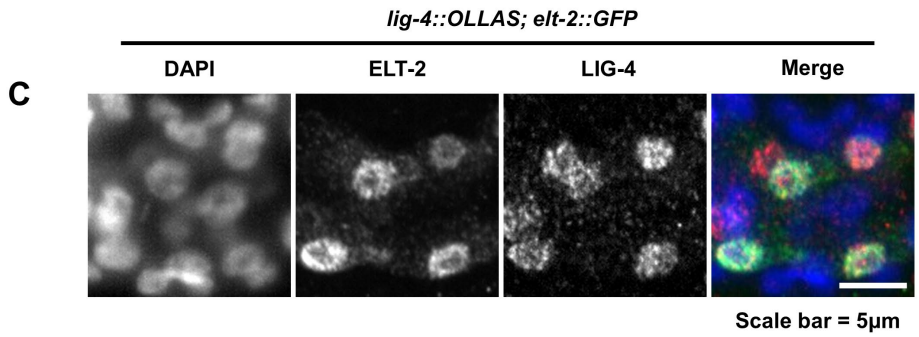


Figure 3.9 - Endogenous LIG-4 localization in the L1

(A) Representative micrographs showing the subcellular localization of LIG-4::OLLAS from the endogenous locus, together with DNA staining (DAPI) and the germline marker HTP-3, in the L1 larva. The LIG-4 signal is detectable beyond background levels only in a row of nuclei along the anterior-posterior axis (see **(B)**). No LIG-4 signal is detected in the PGCs.

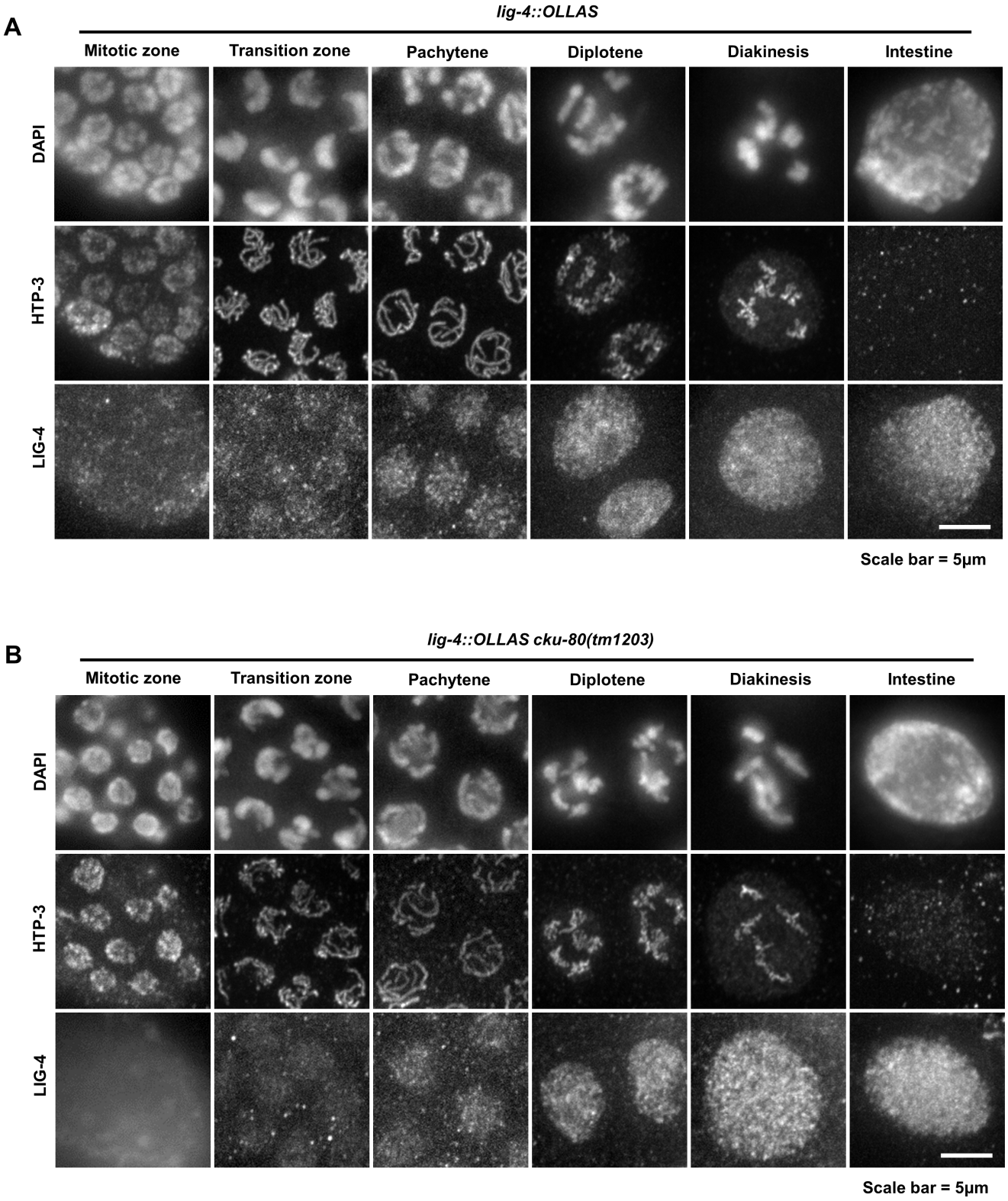
(B) Representative micrographs showing the subcellular localization of LIG-4::OLLAS, HTP-3, and DNA in the same genotypes and conditions as in **(A)**, but in a wider field of view, showing the enrichment in a longitudinal row of nuclei.

(C) Representative micrograph showing the nuclear co-localization of LIG-4::OLLAS and ELT-2::GFP, an intestinal cell marker. The nuclei which most strongly express LIG-4 also express the intestinal marker ELT-2::GFP, suggesting that LIG-4 is enriched in the intestine.

lig-4::OLLAS = a C-terminal tag of the endogenous *lig-4* locus with one copy of the OLLAS epitope
elt-2::GFP = a C-terminal transgenic GFP tag of *elt-2*, integrated by particle bombardment
cku-80 = *cku-80(tm1203)*, a deletion allele of *cku-80*
nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background

DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody
LIG-4 = signal from an anti-OLLAS antibody
ELT-2 = signal from an anti-GFP antibody

Figure 3.10


A
lig-4::OLLAS
Mitotic zone
Transition zone
Pachytene
Diplotene
Diakinesis
Intestine
DAPI
HTP-3
LIG-4
Scale bar = 5μm
B
lig-4::OLLAS cku-80(tm1203)
Mitotic zone
Transition zone
Pachytene
Diplotene
Diakinesis
Intestine
DAPI
HTP-3
LIG-4
Scale bar = 5μm

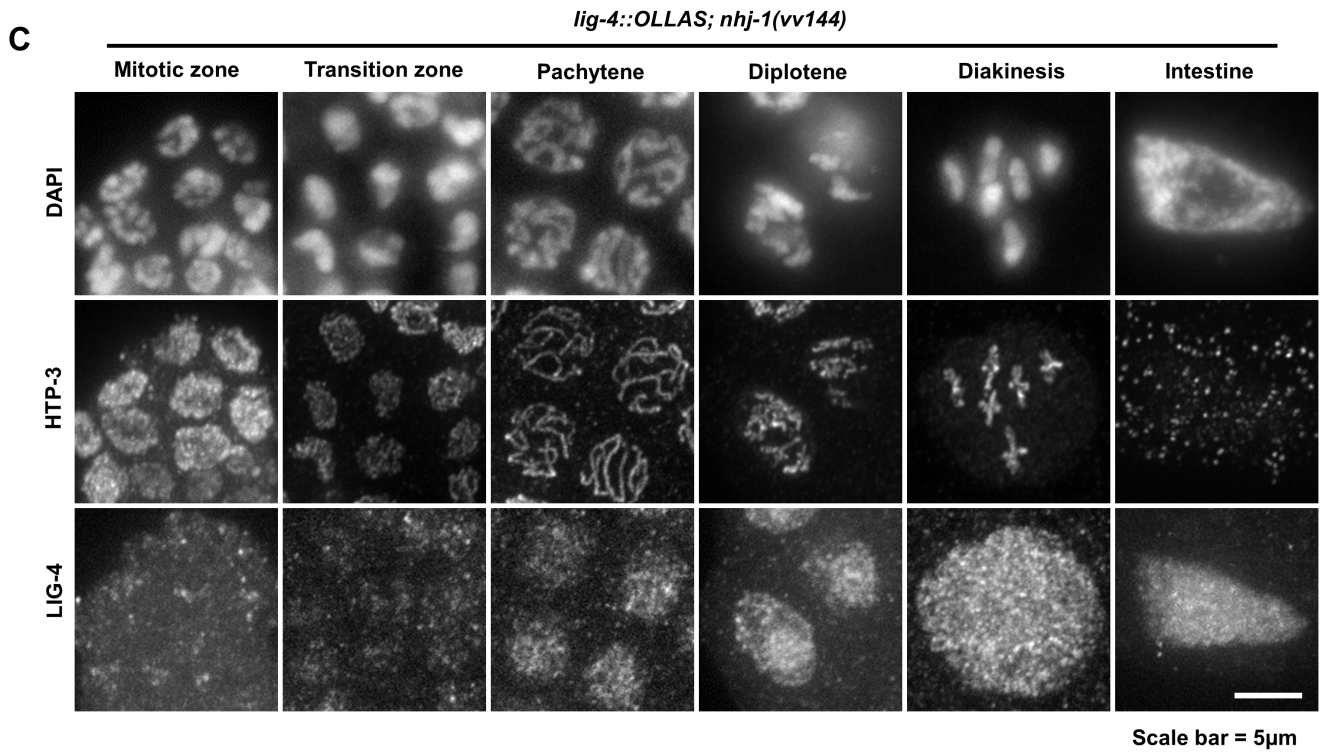


Figure 3.10 - Endogenous LIG-4 localization in the adult germline

(A) Representative micrographs of LIG-4::OLLAS expression from the endogenous locus in adult germline cells. The expression of LIG-4::OLLAS becomes reliably visible in pachytene, and is nuclear, punctate, and not chromatin associated. An adult intestinal cell, where LIG-4 is also strongly expressed, is shown for comparison.

(B) Representative micrographs of LIG-4::OLLAS expression from the endogenous locus in adult germline cells in animals deficient for *cku-80.* Similar to NHJ-1::OLLAS, the loss of CKU-80 does not alter the localization of LIG-4::OLLAS either in the germline or in the intestine.

(C) Representative micrographs of LIG-4::OLLAS expression from the endogenous locus in adult germline cells in animals deficient for *nhj-1.* The absence of NHJ-1 does not affect the localization pattern of LIG-4.

lig-4::OLLAS = a C-terminal tag of the endogenous *lig-4* locus with one copy of the OLLAS epitope
cku-80 = *cku-80(tm1203)*, a deletion allele of *cku-80*
nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background

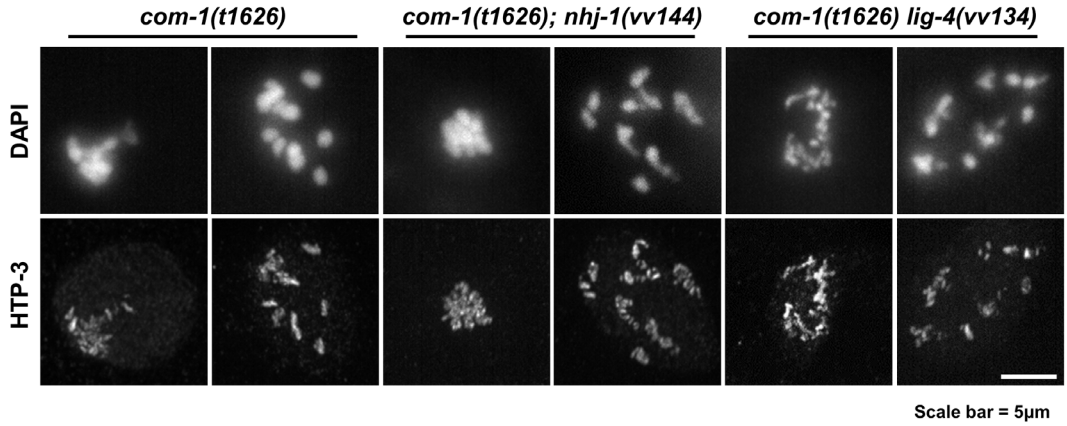
DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody
LIG-4 = signal from an anti-OLLAS antibody

Figure 3.11

A

	Control RNAi (empty vector)			*α-cku-80* RNAi		
Genotype	*com-1(t1626)*	*com-1(t1626); nhj-1(vv144)*	*com-1(t1626) lig-4(vv134)*	*com-1(t1626)*	*com-1(t1626); nhj-1(vv144)*	*com-1(t1626) lig-4(vv134)*
Eggs laid	982	1106	1123	1302	1157	1343
Eggs hatched	3 (0.31%)	0 (0.00%)	6 (0.53%)	344 (26.42%)***	321 (27.74%)***	360 (26.81%)***

B



C

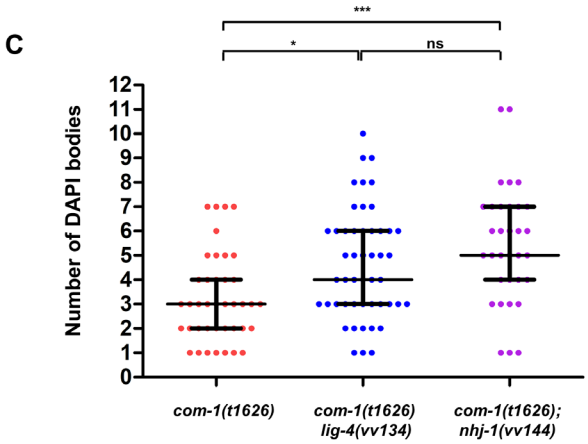


Figure 3.11 - NHJ-1 acts downstream of the Ku ring in the adult germline

(A) Table showing the proportion of eggs hatching in *com-1, com-1; nhj-1,* and *com-1 lig-4* mutants treated with *cku-80(RNAi)* and controls. In control conditions, only a small fraction (<1%) of eggs laid in all three genotypes hatch. With RNAi against *cku-80*, the proportion of hatching eggs is significantly increased ($p<0.001$ versus RNAi control) in all three genotypes. All statistical comparisons shown in the figure are to the RNAi control group within the same genotype (Chi-squared test, Bonferroni correction for multiple comparisons to = 0.008).

(B) Example micrographs showing the diverse DNA morphologies in diakinesis nuclei with low and high numbers of DAPI-staining entities in *com-1, com-1; nhj-1*, and *com-1 lig-4* mutants.

(C) Quantification of DAPI-staining bodies in *com-1*, *com-1 lig-4*, and *com-1; nhj-1* mutants. The number of DAPI-staining bodies is significantly higher in *com-1 lig-4* ($p<0.05$) and *com-1; nhj-1* ($p<0.001$) double mutants is significantly higher than that of *com-1* single mutants, while the two double mutants are not significantly different from each other ($p>0.05$). Sample size (n) is 40 for *com-1*, 49 for *com-1 lig-4*, and 33 for *com-1; nhj-1*.

nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background
lig-4 = *lig-4(vv134[R18STOP])* [R], a predicted null allele of *lig-4* generated in the N2 [R] genetic background
com-1 = *com-1(t1626)*, a published loss of function allele causing a premature stop codon in exon 3 of *com-1*

DAPI = the DNA stain 4',6-diamidino-2-phenylindole
HTP-3 = signal from an anti-HTP-3 antibody

Chapter IV: Discussion

IV.1 Introduction

For this thesis, I had initially set out to investigate how genomic integrity of the early germline in newly hatched *C. elegans* is assured, aiming to identify novel factors which protect the primordial germ cells against the genotoxic effects of ionizing radiation. Entirely serendipitously, this project instead led me to the discovery of IR response heterogeneity within the ostensibly isogenic wild type strain N2, caused by cryptic genetic diversity. Further investigation revealed that the difference in the IR response was caused by a mutation in *nhj-1*, an uncharacterized gene which I have shown to be an essential regulator of the canonical non-homologous end joining DNA repair pathway, the fourth member of the comparatively reduced cNHEJ system in *C. elegans*. In this chapter, I summarize my findings, discuss them and their implications in the context of contribution to wider knowledge, and examine the remaining questions not addressed by the work in this thesis with a prospective lens towards finding the answers.

IV.2 Summary of results

I began my study of the L1 response to ionizing radiation with the primary goal of investigating whether HTP-3, a critical regulator of meiotic processes which is also expressed in the pre-meiotic germline throughout development, plays a role in protecting the PGCs from genotoxic stress. Towards this aim I designed an RNAi screen for chromatin factors which may sensitize or protect the germline at the L1 stage, using adult brood size as the readout. I conducted the screen in the wild type strain N2, as well as the somatic-RNAi deficient *rrf-1(pk1417)* mutant and an *htp-3* mutant with a defective DDR response in the adult germline, *htp-3(vc75)*. While I was able to identify 33 genes whose knockdown modulates post-IR brood size (**Figure 2.2**), none had a differential response in the three genotypes. However, I noticed that N2 had a significantly lower brood size than either *rrf-1(1417)* or *htp-3(vc75)* (**Figure 2.1**), prompting me to investigate the possibility that the N2 line used in our lab may carry an unannotated genetic mutation sensitizing it to IR. Though it initially reassured me that the N2 from the Caenorhabditis Genetics Center had the same response as the N2 in our lab (**Figure 2.3**), testing of a wider sample of laboratory N2

lines, 18 geographically and genetically diverse *C. elegans* wild type isolates, and two *C. briggsae* strains, revealed that the IR-sensitive phenotype was restricted to only a small subset of laboratory N2 lines (**Figure 2.4** and **Table 2.1**). This observation left a recent evolution of IR-sensitivity in N2 as the most parsimonious explanation. To further characterize the differences between the IR-resistant and IR-sensitive strains, I derived a resistant line, N2 [R], and a sensitive line, N2 [S], from single individuals. I showed that in both N2 [R] and N2 [S] the brood size decreases with increasing IR dose, but always to a greater extent in N2 [S], and that this differential brood size response is not apparent with IR treatment at the L4 stage, nor with treatment with other genotoxic agents at the L1 stage, including ENU and UV radiation (**Figure 2.5**), suggesting a developmentally modulated sensitivity to DNA DSBs. L1 radiation exposure does not result in cytologically visible phenotypes in the adult germlines of either N2 [S] or N2 [R] (**Figure 2.6**). Surprisingly, while the embryonic lethality of the progeny of irradiated N2 [S] animals is not different than that of unirradiated controls, that of N2 [R] progeny is slightly but significantly increased (**Table 2.2**), suggestive of mutagenic DNA repair in the germline of this background. Mating of irradiated N2 [R] and N2 [S] hermaphrodites with unirradiated males failed to increase the brood size (**Figure 2.7**), showing that the brood size of irradiated animals of both genotypes is oocyte-limited. In addition to the lowered brood size, IR treatment at the L1 stage also resulted in a much higher frequency of two distinct somatic developmental outcomes in N2 [S]: slow growth, and abnormal vulval phenotypes (protruding vulva, ruptured through vulva, and "bagging") (**Figure 2.8**), which prompted me to hypothesize that the sensitivity of N2 [S] may be related to a deficiency in a somatically active DNA repair pathway, and that the brood size phenotypes may be, at least in part, secondary to the somatic ones. The idea of a loss-of-function mutation in the N2 [S] background was supported by genetic evidence that IR-sensitivity is recessive to IR-resistance (**Figure 2.9**). After eliminating a mutation in *dcr-1* which is present in one of IR-sensitive non-N2 strains as causative of the sensitivity (**Figure 2.10**), I sent N2 [R] and N2 [S] genomic DNA for deep sequencing with the goal of identifying the causative mutation. The sequencing revealed that the majority of fixed variants found in N2 [R] and N2 [S] are shared, with a smaller set of unique variants in each genome (**Figure 2.11**), suggesting that N2 [S] and N2 [R] diverged from each other only after their common lineage split from the reference genome.

With the help of our bioinformatics collaborators in the group of Dr. Steven Jones, 15 fixed variants unique to the N2 [S] genome and predicted to alter protein sequences were identified as the most likely candidates for the causative mutations (**Table 2.3**). Sequencing these candidate mutations in a number of hybrid lines derived from crosses between N2 [R] and N2 [S] animals, I determined that the IR-sensitive phenotype segregates with mutations in *F10D2.12* and *inft-2*, which are both located on the fifth autosome (**Figure 2.12**). Conversion of the *F10D2.12* and *inft-2* sequences into an N2 [S]-like form did not sensitize the N2 [R] background to IR, however, thus eliminating these mutations as causative (**Figure 2.13**). Noticing the similarity between the somatic post-IR N2 [S] phenotypes and those described in cNHEJ pathway mutants (Clejan, Boerckel et al. 2006), I tested whether inactivation of *lig-4* would sensitize the N2 [R] background to IR, which proved to be the case (**Figure 2.14**). However, all three known cNHEJ genes in *C. elegans* are encoded on chromosome III, and their transcripts were expressed in the N2 [S] background (**Figure 2.15**), which together with the other observations suggested that a mutation in a novel cNHEJ factor, located on chromosome V, may be responsible for IR-sensitivity.

Using the mutations in *F10D2.12* and *inft-2* as molecular markers in a large set of hybrid strains, I was able to map the causative locus with greater precision to the neighborhood of *inft-2*, and through a manual search in the N2 [S] genome I located an indel in the third exon of a previously uncharacterized protein-coding gene *H19N07.3*, which I have named *nhj-1* for its potential role in cNHEJ. This indel is predicted to have a severe effect on the two protein isoforms encoded by *nhj-1*, substantially shortening the resulting proteins (**Figure 3.1**). To test whether *nhj-1* plays a role in IR-resistance, I introduced a small deletion in exon 3 in the N2 [R] background, creating the allele *nhj-1(vv144)*, which is predicted to result in a truncation similar to the N2 [S] allele *nhj-1(vv148)* (**Figure 3.1**). This sensitized the N2 [R] background to IR (**Figure 3.2**), strongly suggesting that N2 [S] is sensitive because of a loss of function in *nhj-1.* To formally demonstrate this to be the case, I showed that the IR-sensitivity of the *nhj-1(vv148)* [S] cannot be complemented by the CRISPR-generated *nhj-1(vv144)* allele (**Figure 3.3**). I then proceeded to test the hypothesis that *nhj-1* belongs to the cNHEJ pathway by comparing the severity of post-IR phenotypes in *nhj-1(vv148); lig-4* double mutants and *nhj-1(vv144); cku-80* double mutants with either of the single mutants alone. Since the double mutants did not exhibit a worse post-IR phenotype than either

of the single mutants (**Figure 3.4**), I concluded that *nhj-1* is indeed a novel fourth member of cNHEJ in *C. elegans*. I next showed that extrachromosomal *nhj-1* can rescue both the somatic phenotypes and the brood size reduction in both *nhj-1(vv144)* and *nhj-1(vv148)* [S] (**Figure 3.5**), further supporting the conclusion that IR-sensitivity arises because of the loss of NHJ-1 protein function, rather than a loss of a cis-regulatory sequence within the *nhj-1* locus. The extrachromosomal *nhj-1* construct was also translationally fused to GFP, allowing me to probe the intracellular localization of NHJ-1, which was revealed to be present in many somatic nuclei in the L1 larva, but not in the PGCs (**Figure 3.6**). The lack of expression of extrachromosomal NHJ-1::GFP in the PGCs, together with the somatic and brood size rescue, argued that the observed brood size phenotype derive predominantly from a somatic *nhj-1* deficiency. I next investigated the endogenous localization pattern of NHJ-1, employing a C-terminal OLLAS tag generated by CRISPR, and found that endogenous NHJ-1::OLLAS has the same expression pattern as the exogenous NHJ-1::GFP, with NHJ-1 detectable in the majority of somatic nuclei but not in the PGCs (**Figure 3.7**). This pattern was not perturbed in backgrounds deficient for *cku-80* or *lig-4*, nor following IR treatment (**Figure 3.7**). I also examined the endogenous NHJ-1::OLLAS localization pattern in the adult germline, a tissue where cNHEJ functions only as a backup repair pathway. I observed the NHJ-1::OLLAS signal only in diplotene and diakinesis nuclei, which, like in the L1 larva, was not altered in *cku-80* or *lig-4* deficient backgrounds (**Figure 3.8**). I also OLLAS-tagged the endogenous locus of *lig-4*, the terminal effector of the cNHEJ pathway, and investigated its localization pattern and whether it is affected by the status of *cku-80* or *nhj-1.* In contrast to NHJ-1, LIG-4 prominently localizes only to intestinal cell nuclei in the L1 larva (**Figure 3.9**), while in the adult gonad it can be detected earlier than NHJ-1, starting in pachytene and persisting in diplotene and diakinesis (**Figure 3.10**). The localization pattern of LIG-4 is not affected by loss of *cku-80* or *nhj-1* either in the L1 larva or the adult germline. Finally, to position *nhj-1* within the cNHEJ pathway, I tested whether the loss of *nhj-1* can rescue the embryonic lethality of *com-1* mutants, as the loss of *cku-80* has been shown to be partially capable of doing (Lemmens, Johnson et al. 2013). However, *com-1; nhj-1* double mutants exhibit the nearly complete embryonic lethality characteristic of *com-1* mutants, and like *com-1 lig-4* double mutants, they also show increased numbers of DAPI bodies in diakinesis (**Figure 3.11**). These

results suggest that, in the absence of NHJ-1, the Ku ring is capable of engaging DNA ends but repair cannot efficiently occur, arguing that NHJ-1 acts downstream of Ku and upstream of LIG-4 in the pathway.

IV.3 Scientific contribution and implications of findings

The most significant contribution of the present study to the body of scientific knowledge is the discovery of *nhj-1*, an entirely novel player in the cNHEJ pathway in *C. elegans*, which provides an affirmative answer to the question of whether the worm possesses cNHEJ components other than the core kit of *cku-70*, *cku-80*, and *lig-4*. However, the way that led to that discovery also raises further questions and has implications for the field of *C. elegans* research and biology more generally.

IV.3.1 The non-isogeneity of N2

One such implication is that there may be more genetic variability in the strains of animals that experimental scientists work with than is generally acknowledged. As I have demonstrated in this study, the IR-sensitivity caused by the spontaneous mutation *nhj-1(vv148)* is visible only under the specific conditions of exposure to high doses (≥25 Gy) of ionizing radiation, and only early on in development. Under normal conditions, N2 [S] is phenotypically indistinguishable from N2 [R], and the animals give no indication of a deficiency in a major DNA repair pathway. This illustrates the disquieting possibility that other cryptic mutations may be present either in the ostensibly wild type strains commonly used as controls, or in one or more of the thousands of mutant strains used within the *C. elegans* research community. The presence and danger of genetic drift in *C. elegans* has been noted before (Flibotte, Edgley et al. 2010). Such cryptic mutations may affect the response studied and the conclusions drawn from the studies, as was the case during the early stages of my own research, when the IR-sensitivity of the Zetka N2 strain and the apparent IR-resistance of the endo siRNA and some chromatin mutants misled me into an erroneous interpretation of the data.

As the deep sequencing has revealed, in addition to *nhj-1(vv148)*, the N2 [S] genome carries at least an additional 15 single nucleotide variants which have a direct impact on the sequences of translated proteins. The N2 [R] genome contains 13 such single nucleotide mutations which

affect protein sequences, and these are just among the variants that are unique to N2 [S] and N2 [R]. While the variants shared between the two N2 lines sequenced for this study do not play a role in the phenotypic differences observed and described here, they may influence other processes and cause phenotypes conditional to other environmental or genetic factors. Evolution and genetic drift cannot be stopped in *C. elegans* any more than in other living organisms, and the importance of maintaining isogenic or near-isogenic lines has long been recognized in the worm research community. This is greatly aided by the ability to cryogenically preserve strains in liquid nitrogen (Lewis and Fleming 1995), which prevents random mutations and changes in allele frequency. Despite this, cryptic genetic variation can and does spread, as this study has shown, especially if facilitated by distribution centers like the CGC. My analysis has confirmed the presence of the *nhj-1(vv148)* mutation in only five strains - the Zetka lab N2, the CGC N2, YY470 *dcr-1(mg375)*, HBR1099 *unc-119(ed3)*; *goeEx386*, and YY186 *nrde-2(gg91)*. However, the allele may be more widely distributed, and with the development of a PCR-based genotyping assay (see section **V.15**), the wider community now has a readily employable genotyping option.

The course of this study also illustrates that new technological solutions can help overcome the more troubling implications of the non-isogeneity among experimental strains. First, whole-genome sequencing has become both affordable and rapid (Levy and Myers 2016), allowing individual research groups the opportunity to reveal the full spectrum of variations in their strains. Second, the advent of CRISPR-Cas9 mutagenesis (Hsu, Lander et al. 2014, Dickinson and Goldstein 2016) has enabled researchers to introduce almost any desired alteration within a known genetic background, making possible an “apples-to-apples” comparison from which to interpret the effects of the mutation.

IV.3.2 NHJ-1 is a novel member of the cNHEJ pathway in *C. elegans*

The identification of NHJ-1 as a critical regulator of cNHEJ in *C. elegans* represents the major finding of the present work (**Figure 4.1**). Before this study, it has been hypothesized that *C. elegans* either possesses a minimal cNHEJ system, composed of only the Ku ring and LIG-4, or that other cNHEJ components have yet to be identified because no saturated screen for cNHEJ factors has yet been done (Lemmens and Tijsterman 2011). Such unidentified cNHEJ components

have been postulated to be nucleases or kinases (Lemmens and Tijsterman 2011), which would act in the processing of free DNA ends, or signal to coordinate the enzymatic activity during cNHEJ. The evidence of *cku-80; nhj-1* and *lig-4; nhj-1* double mutants shows that NHJ-1 functions in the cNHEJ pathway, and the lack of embryonic lethality rescue as well as an increased number of diakinetic DAPI bodies in *com-1; nhj-1* double mutants compared to *com-1* single mutants suggests that it acts downstream of Ku and upstream of *lig-4*. The roles NHJ-1 may play downstream of Ku binding include: 1) DNA end processing; 2) Signaling to coordinate the activity or assembly of the cNHEJ complex; 3) Promoting the activity of other cNHEJ pathway components as a cofactor; and 4) Acting as a structural scaffold to organize and coordinate other cNHEJ factors. NHJ-1 contains no conserved domains. Its structure is predicted to contain an alternating sequence of alpha helices and beta sheets, while the last 63 amino acids of both the long and the short isoform are predicted to be intrinsically unstructured (**Figure 4.2**). Taking this together with the relatively small size of the protein (with even the longer isoform being only 168 amino acids long), I consider the possibility that NHJ-1 is an enzyme unlikely. Signaling and processing enzymes with active roles in cNHEJ tend to be much larger, with the ~4,000 amino acid-long DNA-PKcs at the higher end of the spectrum, and the ~500 amino acid-long nuclease APLF at the lower end (Chang, Pannunzio et al. 2017). By contrast, the structural proteins XRCC4, XLF, and PAXX, are of much more modest size, ranging from 201 aa (PAXX) to 334 aa (XRCC4) in *H. sapiens* (Chang, Pannunzio et al. 2017), although XRCC4 and XLF have been shown to oligomerize into much larger filaments that support other cNHEJ machinery (Hammel, Rey et al. 2011, Mahaney, Hammel et al. 2013). NHJ-1 could act in an analogous manner in *C. elegans*, even though it shares no sequence homology with XRCC4 or its homologs. However, a role in an enzyme-driven step of cNHEJ cannot be definitively excluded, as NHJ-1 could act to promote enzymatic activity; furthermore, unlikely though it is, NHJ-1 could possess enzymatic activity itself, particularly since smaller enzymes have been described, such as the 63 amino acid-long bacterial protein 4-Oxalocrotonate tautomerase, which oligomerizes to form the active site (Chen, Kenyon et al. 1992).

The scope of the evolutionary conservation of NHJ-1 is limited. Proteins with high identity with the NHJ-1 long isoform exist in several species of the genus *Caenorhabditis*, including *C. brenneri*

(90% identity), *C. briggsae* (88% identity), *C. remanei* (84% identity), and *C. latens* (84% identity). In the family Rhabditidae, which includes the genus *Caenorhabditis,* there are two homologs in the asexual worm *Diploscapter pachys* (33% and 32% identity). The only other proteins with homology belong to two parasitic hookworms in the family Ancylostomatidae, *Necator americanus* (24% identity), and *Ancylostoma duodenale* (22% and 20% identity). What roles the homologs of NHJ-1 play in the other nematodes is not known. Given the relatively high sequence conservation within *Caenorhabditis*, NHJ-1 homologs in the other species of this genus may also participate in cNHEJ. The nematode family Rhabditidae, or more speculatively if *N. americanus* and *A. duodenale* are included, the order Rhabditida, thus appears to have evolved a novel regulator of the nearly universally conserved (Gu and Lieber 2008) cNHEJ pathway. The restructuring of cNHEJ by incorporation of a completely novel factor is rare in eukaryotes, with perhaps the closest example being the cooption of the MRX complex into the cNHEJ mechanism in *S. cerevisiae* (Daley, Palmbos et al. 2005, Emerson and Bertuch 2016). This nevertheless illustrates the evolutionary plasticity of even the most ancient pathways. The lack of sequence conservation between NHJ-1 and the known cNHEJ factors in other phyla also points to the possibility that the cNHEJ toolkit in *C. elegans* may be much larger, and novel functional analogs to other cNHEJ factors may yet be discovered.

This study also examined the subcellular localization of NHJ-1 in L1 larvae and adult gonad and intestinal tissue, and provided the first description of LIG-4 localization in *C. elegans*, both of which raised interesting questions about the regulation of cNHEJ in specific tissue contexts. The localization of cNHEJ components, the regulation of their recruitment to sites of DNA damage, and their dependence on other cNHEJ factors for nuclear recruitment has primarily been studied in the context of cultured mammalian cells; see, among others, (Koike, Awaji et al. 1999, Nilsson, Sirzen et al. 1999, Koike, Shiomi et al. 2001, Girard, Kysela et al. 2004, Mari, Florea et al. 2006). As expected for DNA repair factors, Ku, DNA-PKcs, XRCC4, XLF, PAXX, and LIG4 are predominantly nuclear, with Ku and possibly others excluded from the nucleus only during mitosis (Nilsson, Sirzen et al. 1999, Koike, Shiomi et al. 2001, Girard, Kysela et al. 2004, Yurchenko, Xue et al. 2006, Koike, Yutoku et al. 2015, Ochi, Blackford et al. 2015). By contrast, few published studies have examined the localization of cNHEJ factors in the tissue or organ context. In healthy human colon

tissue, Ku70 is detectable by immunohistochemistry in 74% of nuclei, in contrast to Ku80 which can be seen in only 32% of nuclei (Mazzarelli, Parrella et al. 2005). Similarly, many but not all cells in the crypts of human and murine small intestine express LIG4, which is detectable in both nuclei and the cytoplasm in the cells that express it (Jun, Jung et al. 2016). Another tissue which has been examined is the mouse testis, where Ku70 localizes to the nuclei of the somatic Sertoli cells, spermatogonia, late (post-pachytene) spermatocytes I, spermatocytes II, and spermatids (Ahmed, Sfeir et al. 2013).

These observations accord in broad outline with the NHJ-1 and LIG-4 localization data from this study. In *C. elegans*, both NHJ-1 and LIG-4 are nuclearly localized in both the L1 larva and the adult gonad and intestine. In the adult germline, the two proteins have a similar expression pattern, with strongest expression in diplotene and diakinesis, although LIG-4 becomes visible in pachytene. This is in line with the role of cNHEJ as a backup DNA repair pathway during meiotic prophase I, when inter-homolog HR repair is heavily favored (Clejan, Boerckel et al. 2006, Smolikov, Eizinger et al. 2007, Adamo, Collis et al. 2010, Lemmens and Tijsterman 2011). This expression pattern is also reminiscent of the localization of Ku70 in the mouse seminiferous tubules (Ahmed, Sfeir et al. 2013), except that some Ku70 expression is seen in spermatogonia while neither NHJ-1 nor LIG-4 are visible in the mitotic zone of *C. elegans*. The tissue-level expression pattern of NHJ-1 and LIG-4 in the L1 larva is markedly different, however. While NHJ-1 localizes to the majority (60-75%) of somatic nuclei, LIG-4 is detectable primarily in the intestine, raising the question of why this should be the case if both proteins act in the same pathway. Several possibilities exist that could explain the observed discordance. Perhaps NHJ-1 is pleiotropic and is serving a cNHEJ-independent function in non-intestinal cells. I consider this explanation unlikely because it suggests that cNHEJ operates only in the intestinal cells at the L1 stage. Since the post-L1-IR vulval phenotypes have been found to result from necrotic loss of vulval precursor cells (VPCs) at the L4 stage (Weidhaas, Eisenmann et al. 2006), if cNHEJ were operating only in the intestine at L1, the loss of VPCs would have to result non-cell-autonomously from unrepaired or deleteriously repaired DNA damage in the intestine. Another possibility is that the expression of LIG-4 may be below the detection threshold in non-intestinal nuclei, but with LIG-4 nevertheless present, which would suggest a relative enrichment in intestinal cells

compared to other tissues. Although intestinal cell nuclei are diploid in the early L1 larva, their ploidy doubles with each larval stage to the final number of 32 copies of each chromosome in adult (Hedgecock and White 1985), suggesting that the LIG-4 enrichment may reflect a greater need for cNHEJ in this tissue. If this interpretation is correct, however, it raises the question of why LIG-4 should be "pre-loaded" in diploid L1 intestinal nuclei, rather than increase with ploidy over the course of development, and why a degree of enrichment is not also seen in the cells of the lateral hypodermis, which become tetraploid during later larval development (Hedgecock and White 1985). In addition to increased ploidy, an increased requirement for cNHEJ in the intestinal cells may result from the fact that these cells are the ones most likely to be directly exposed to toxins produced by pathogenic bacteria and other microbiota which can colonize the intestinal lumen (Jiang and Wang 2018). However, NHJ-1 is not enriched in intestinal nuclei compared to other somatic nuclei, suggesting that a general enrichment of cNHEJ factors is not sufficient to explain the LIG-4 pattern. Future experiments, some of which are discussed in section **IV.4**, will be required both to refine these possibilities and test their predictions.

IV.3.3 The IR-dependent brood size reduction is primarily dependent on somatic repair

Decreased fertility following IR exposure during early development has been observed both in wild type animals (Weidhaas, Eisenmann et al. 2006) and in cNHEJ mutants (Clejan, Boerckel et al. 2006). Because of the egg-laying phenotypes associated with IR treatment early in development (Clejan, Boerckel et al. 2006), the fertility of cNHEJ mutants has not been scored at the level of the total brood size before this study. Two lines of evidence suggest that the brood size reduction differential between N2 [R] and N2 [S] after IR treatment at the L1 stage is primarily mediated by somatic effects. First, there appears to be no germline DNA damage in N2 [S], according to both a lack of cytological abnormalities and a lack of increased embryonic lethality among the progeny. Second, the expression of NHJ-1 from an extrachromosomal array, which are often silenced in the germline (Kelly, Xu et al. 1997), substantially rescues the brood size phenotype of *nhj-1* mutants.

However, some evidence suggests that NHJ-1 or cNHEJ does play a role in the IR response of the germline tissue. First, the significantly increased embryonic lethality among the progeny of N2

[R] animals irradiated at the L1 stage is consistent with the existence of a cNHEJ-mediated mutagenic repair option in this genetic background, but not in N2 [S]. That the increase in embryonic lethality in N2 [R] is caused by cNHEJ activity rather than another mutation in the genetic background is supported by the lack of increased embryonic lethality among the progeny of irradiated *nhj-1(vv144)* and *lig-4(vv134)* mutants (**Table 4.1**), both of which have been generated by CRISPR in the N2 [R] background. A collection of observations relating to the sensitive and resistant IR response and their potential epigenetic regulation also deserves brief mention here, in the context of a potential germline contribution to the IR response phenotype.

Before we understood the IR-sensitivity of N2 [S] as a loss of cNHEJ activity, we worked under the paradigm that the post-IR brood size decrease in N2 [S] may come as a consequence of IR-dependent changes that happen in the PGCs, and which are subsequently transmitted to the adult, reproductively-capable germline. Since we were interested in chromatin changes, we quantified the levels of trimethylated lysine 9 on histone 3 (H3K9me3) and of acetylated lysine 9 on histone 3 (H3K9Ac), histone marks associated with transcriptional repression and activation, respectively (Lawrence, Daujat et al. 2016) in the PGCs in irradiated L1s and unirradiated controls. Intriguingly, we noticed a divergent response of H3K9me3 to IR in N2 [R] and N2 [S], with levels of this mark being reduced in the PGC of N2 [R] following IR exposure ($p<0.001$ vs unirradiated controls), while being increased following IR in the N2 [S] background ($p<0.001$ vs unirradiated controls) (**Figure 4.3A**). Conversely, the levels of H3K9Ac increase following IR in N2 [S] ($p<0.001$ vs unirradiated controls), and do not change in N2 [R] ($p>0.05$ vs unirradiated controls) (**Figure 4.3B**). The post-IR increase in H3K9me3 and the decrease in H3K9Ac in the primordial germline are thus correlated to the decreased brood size in N2 [S]. We next asked whether the resetting of histone marks could affect the post-IR brood size outcome. To test this, we passed irradiated N2 [S] animals through the dauer stage, which has been shown to result in a global resetting of the histone mark landscape (Hall, Beverly et al. 2010). Passage through dauer slightly but significantly increased the brood size of irradiated N2 [S] animals (median brood size 46.5 progeny and IQR 118.75, compared to the median brood size of 4 progeny and an IQR of 39.5 in irradiated N2 [S] animals that haven't been passed through dauer; $p<0.001$) (**Figure 4.3C**). This observation is consistent with a partial control of post-IR brood size via histone marks in the

germline. Perhaps an N2 [S]-specific epigenetic signal, set up during or after IR exposure in the L1, is partially responsible for the reduction in the adult brood size. However, since I did not assay what effects dauer passage has on the somatic phenotypes, the dauer-passage brood size rescue may simply reflect a partial somatic rescue.

Finally, we investigated whether the comparatively greater post-IR brood size reduction of the sensitive lines is transmitted to the progeny. To our surprise, we noticed that the F1 progeny of irradiated N2 [R] animals, which have never directly experienced IR, have a significantly lower brood size (median of 179 progeny, IQR 93.5) than either the F1 progeny of unirradiated N2 [R] (median of 251 progeny, IQR 65; $p<0.05$) or the F1 progeny of irradiated N2 [S] (median of 290 progeny, IQR 70.25; $p<0.001$) (**Figure 4.3D**). The reduced brood size in N2 [R] F1s from irradiated mothers is rescued by both inactivation of *nhj-1* (median 283.5 progeny, IQR 60.75; $p<0.001$ vs N2 [R]) and of *lig-4* (median 266 progeny, IQR 70; $p<0.01$ vs N2 [R]) (**Figure 4.3D**), suggesting that the reduction in brood size in F1 animals is a consequence of cNHEJ activity in the maternal germline. A major caveat of this interpretation is that effects on the F1 phenotype cannot be properly considered transgenerational since the early development of F1 progeny takes place within the body of the directly IR-affected mothers, and may therefore be germline non-autonomous. However, that it is N2 [R] rather than the cNHEJ-deficient strains that exhibits the F1 phenotype argues against a soma-mediated effect because somatic tissues are much less severely affected by IR in N2 [R] than in cNHEJ mutants. Several avenues for future experimentation, discussed in the following section, could shed further light on the question of cNHEJ involvement in the germline.

IV.4 Remaining questions and future research

The data presented in this thesis suggest a number of distinct questions which can be tested in the near future both to strengthen the current understanding of NHJ-1 and its role in cNHEJ and beyond, as well as reveal more about the cNHEJ process in the worm.

IV.4.1 What is the cause of brood size reduction following IR treatment?

The first group of questions concern the role of the germline in the L1 IR response, both in the wild type, “resistant” lines and in the cNHEJ deficient mutants. First, it remains unknown what

the proximal cause of the differentially decreased brood size in the sensitive lines is. The two most obvious possibilities are a decreased overall number of germ cells, or increased apoptotic cell death in the germline. The number of germ cells can be readily scored with DAPI and HTP-3 immunostaining, while apoptosis can be scored with acridine orange or SYTO 12, dyes which selectively stain apoptotic cells (Lant and Derry 2014). Increased germ cell apoptosis is a hallmark of IR exposure at the L4 stage (Gartner, Milstein et al. 2000), and is mediated by activity of the *C. elegans* p53 ortholog *cep-1* (Derry, Putzke et al. 2001). If the post-L1-IR brood size decrease is mediated by this apoptotic pathway, the brood size of IR-sensitive lines should be rescued in *cep-1* mutants. Or, since *cep-1* has a role in antagonizing cNHEJ in the germline which may complicate analysis (Mateo, Kessler et al. 2016), mutants of the core apoptotic pathway like the caspase *ced-3* (Gartner, Boag et al. 2008) could be used instead. Second, is the post-IR H3K9me3 enrichment and the depletion of H3K9Ac in the PGCs of N2 [S] but not N2 [R] animals caused by a difference in cNHEJ activity, a difference in the status of NHJ-1, or something else in the genetic background? This can be directly addressed by testing whether the histone mark changes observed in N2 [S] also occur in *nhj-1* and in another cNHEJ mutant, such as *lig-4*. If they do not, the difference can be ascribed to another background mutation. If the phenotypes are recapitulated in *nhj-1* mutants, NHJ-1 can be concluded to be the cause, and if they are also visible in *lig-4* mutants, repair by cNHEJ, whether residually in the PGCs or in somatic tissues, would be the likeliest cause. Another remaining question is whether the dauer rescue is mediated by the soma or the germline, which can be investigated by scoring whether dauer passage also rescues the observed somatic phenotypes, and whether a partial rescue is also possible in the N2 [R] strain, where the post-IR chromatin marks in the PGCs behave antithetically to N2 [S]. If dauer passage also rescues somatic phenotypes, the rescue could be a systemic rather than a germline-specific event, and if it is also observed in N2 [R], the correlation between the acquisition or loss of specific marks and the physiological IR outcomes would be weakened. A final group of questions on the germline involvement of cNHEJ concerns the decreased brood size of F1s originating from irradiated N2 [R] mothers. What is causing the brood size decrease? The first possibility to test would be increased embryonic lethality, which would accord with the paradigm of mutagenic cNHEJ activity in the germline. Furthermore, if the brood size decrease is caused by mutations, they

should breed true in every following generation, which can be easily tested. The incidence of non-lethal mutations such as Dpy (dumpy/shorter and stouter body) and Unc (uncoordinated locomotion) would also be expected to be higher in this case, which can be readily assayed.

IV.4.2 What role does NHJ-1 play in *C. elegans* cNHEJ and beyond?

The second group of questions raised by this work concerns the protein NHJ-1, the mechanism of its role in cNHEJ, and the composition of the cNHEJ pathway in *C. elegans*. The lack of rescue of embryonic survival in *com-1; nhj-1* double mutants suggests that NHJ-1 acts downstream of Ku, and the increase in the number of DAPI-staining bodies in *com-1; nhj-1* double mutants implies that ligation cannot happen without NHJ-1. Another way to test the role of NHJ-1 in repair is a system developed in the Tijsterman lab, in which a single DSB can be induced by heat shock treatment, which causes the restriction enzyme SceI to cut at a precise location in the genome (Pontier and Tijsterman 2009). The locus containing the SceI cut site can be PCR amplified and the repair products analyzed, with a distinct spectrum of products generated by cNHEJ-proficient and cNHEJ-deficient backgrounds. This method could provide molecular evidence of the role of NHJ-1 to supplement the cytological data. At the time of writing, I have crossed *nhj-1(vv144)* into this system, as well as *lig-4(vv134)* to serve as a positive control. This will allow me to test whether *nhj-1* and *lig-4* deficient backgrounds produce similar repair products, which would further support a role for *nhj-1* in repair. A more precise dissection of the mechanistic role of NHJ-1 would require a further characterization of its two isoforms, including a generation of an allelic series, investigation of potential post-translational modifications, and a search for interacting partners. Identifying the interacting partners of NHJ-1, which could be done by immunoprecipitation of the OLLAS-tagged version followed by mass spectrometry, or by a yeast two-hybrid screen, could also act as a potential springboard from which to identify other *C. elegans* cNHEJ pathway members.

Several remaining questions are raised by the contrasting localization patterns of LIG-4 and NHJ-1, and some of these can be addressed by straightforward experiments. First, is the localization of LIG-4 enriched in the intestine because this is the only physiologically relevant tissue for cNHEJ? This can be tested by investigating whether a *lig-4* sequence driven by an intestine-specific promoter like *elt-2p*, could fully rescue the IR-sensitivity phenotype of *lig-4* mutants.

Similarly, an *elt-2p::nhj-1* construct can be assayed for rescue of IR-sensitivity in *nhj-1* mutants. If the answer is negative, the most likely explanation for LIG-4 being visible only in intestinal nuclei would be that LIG-4 is present below detectable levels in other tissues. Determining the localization pattern of CKU-70 and CKU-80 would further inform this question. Whether the Ku ring components are more akin to NHJ-1 or LIG-4 could suggest which localization pattern is more relevant for cNHEJ. Although *cku-70* and/or *cku-80* can be tagged with an epitope tag like *nhj-1* and *lig-4* were in the present study, a translational fusion with a fluorescent protein, if not functionally disruptive, may be more informative. A fluorescently-tagged *nhj-1* and *lig-4,* in addition to a Ku ring protein, would enable not only direct co-localization studies, but open the possibility of using laser microirradiation to observe the formation of cNHEJ complexes in real time, as has been done in mammalian cell culture (see (Mari, Florea et al. 2006) and others), but was not possible with immunostaining against epitope tags in this study.

If NHJ-1 does also serve a non-cNHEJ function, what might that be? One possibility is that it participates in another DNA repair pathway. DNA repair proteins shared across multiple pathways include the MRX complex, which contributes to both HR and cNHEJ in yeast (Daley, Palmbos et al. 2005, Emerson and Bertuch 2016), the XRCC1 scaffold participating in mammalian BER, NER, SSBR, and MMEJ (Dianov and Hubscher 2013, Caldecott 2014, Chatterjee and Walker 2017, Sallmyr and Tomkinson 2018), and the *C. elegans* endonuclease XPF-1, which is involved in NER (Lans and Vermeulen 2011), SSA (Pontier and Tijsterman 2009), and dHJ resolution (Agostinho, Meier et al. 2013, O'Neil, Martin et al. 2013, Saito, Lui et al. 2013). The lack of UV and ENU sensitivity displayed by *nhj-1(vv148)* mutants argues against a role for NHJ-1 in NER or BER, and the lack of evidence for germline DNA damage argues against a role in the dHJ pathway of HR. However, a role in Alt-EJ, SSA, or another mutagenic pathway cannot at present be excluded. Recently published work has shown that in the absence of *rad-51*, the loss of cNHEJ, TMEJ, SSA, or a combination of these repair pathways, results in distinct outcomes in terms of the number of DAPI stained bodies (Macaisne, Kessler et al. 2018). If *nhj-1* were to be crossed into a *rad-51* deficient background with or without *xpf-1* (SSA) or the polymerase θ gene *polq-1,* these outcomes can be used to genetically test whether NHJ-1 contributes to SSA or TMEJ in addition to cNHEJ.

IV.4.3 Screen follow up

Finally, the RNAi screen for chromatin factors mediating the L1 IR response yielded 33 candidates which represent a wellspring of potential new directions for research. As a first line approach to quantitatively validate the qualitative observations from the screen, the brood size and somatic phenotypes of each RNAi candidate should be scored following knockdown and irradiation, and the most promising candidates followed through. Of particular interest are the candidates which increase the post IR brood size, as this outcome suggests they may play a role in antagonizing a DNA repair pathway.

IV.5 Conclusion

In conclusion, my work has uncovered a critical novel regulator of canonical non-homologous end joining in *C. elegans*, *nhj-1*, and shown that a spontaneously generated allele of this gene has propagated in several laboratory strains of *C. elegans*, including the most commonly used wild type strain, N2. The major implications of these discoveries are that *C. elegans* has restructured an ancient and conserved DNA repair pathway, which hints at the possibility of a larger cNHEJ toolkit in the worm, and that cryptic genetic variation of conditional functional importance exists within the ostensibly wild type line and may confound interpretation of experimental results unless accounted for. Using these findings as a stepping stone, future investigations will shed more light on the mechanism and regulation of canonical non-homologous end joining in the worm, as well as advance general knowledge about this DNA repair pathway through comparative examination of the more conserved cNHEJ systems.

Tables and Figures

Table 4.1

	Control L1 unhatched/total (%)	Irradiated L1 (75 Gy) unhatched/total (%)
lig-4(vv134) [R]	9/1505 (0.60%)^{ns}	2/672 (0.30%)^{ns}
nhj-1(vv144) [R]	2/1322 (0.16%)	4/627 (0.64%)
	Control L4 unhatched/total (%)	**Irradiated L4 (75 Gy)** unhatched/total (%)
lig-4(vv134) [R]	8/972 (0.82%)^{ns}	627/1244 (50.40%)^{ns}
nhj-1(vv144) [R]	7/851 (0.82%)	653/1288 (50.70%)

Table 4.1 - Embryonic lethality is not increased following L1 IR in *nhj-1(vv144)* and *lig-4(vv134)*

The progeny of *nhj-1(vv144)* or *lig-4(vv134)* mothers irradiated with 75 Gy at the L1 stage do not exhibit an increased embryonic lethality ($p>0.05$ for both comparisons) compared to the progeny of unirradiated mothers showing a lack of mutagenic DNA repair in these genotypes. The two genotypes are not significantly different from each other in either the untreated or IR treated groups, at both L1 and L4 ($p>0.05$ for all comparisons). All statistical comparisons shown in the figure are to *nhj-1(vv144)* at the equivalent IR dose and stage (Chi-squared test, Bonferroni corrected for multiple comparisons to $\alpha = 0.01$).

Gy = Gray (unit)
nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background
lig-4 = *lig-4(vv134[R18STOP])* [R], a predicted null allele of *lig-4* generated in the N2 [R] genetic background

ns = not significant ($p>0.01$)

Figure 4.1

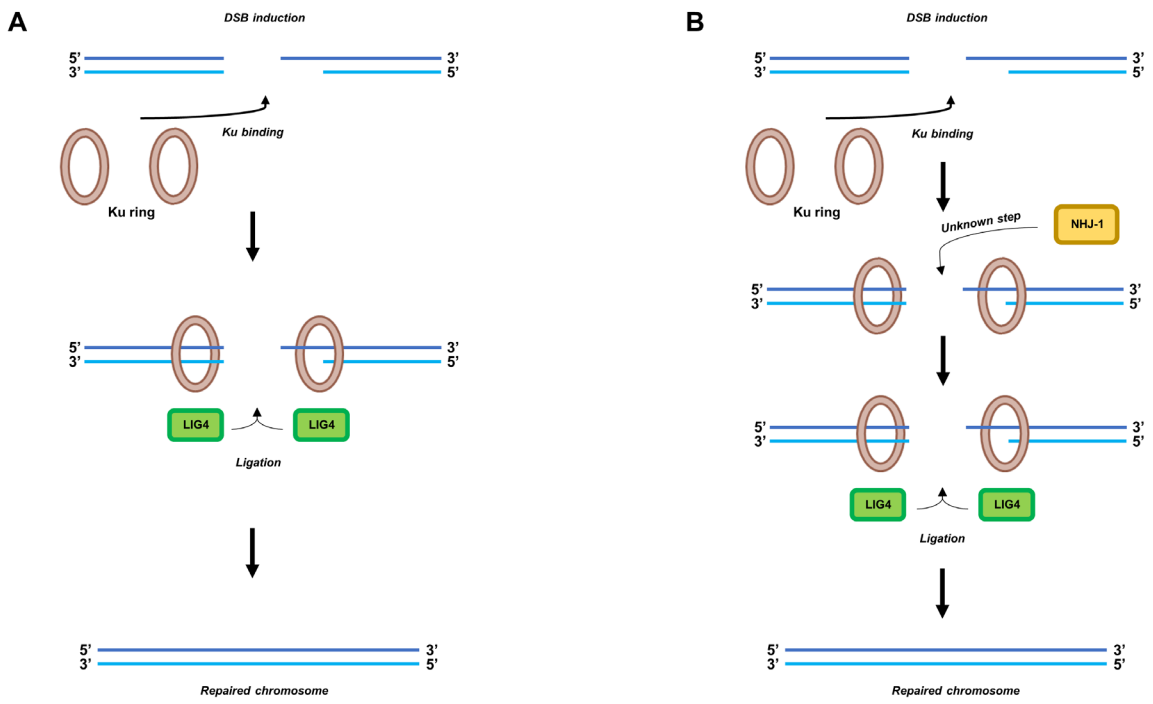


Figure 4.1 - NHJ-1 is a fourth member of the canonical non-homologous end joining pathway in *C. elegans*

(A) Prior to this study, the *C. elegans* canonical non-homologous end joining system was believed to consist only of the universally conserved core factors, the Ku ring and LIG-4, or that it contained other factors which were not homologous to the other known eukaryotic cNHEJ proteins (Lemmens and Tijsterman 2011). How the steps of cNHEJ between Ku binding and terminal ligation by LIG-4 occur was unknown.

(B) My work has shown that the *C. elegans* cNHEJ system includes the protein NHJ-1, which is taxonomically restricted only to *C. elegans* and related rhabditids. NHJ-1 is a critical regulator of cNHEJ, and it acts downstream of Ku, and presumably upstream of LIG-4 to effect repair. Its mechanistic role in the process remains unknown.

Figure 4.2

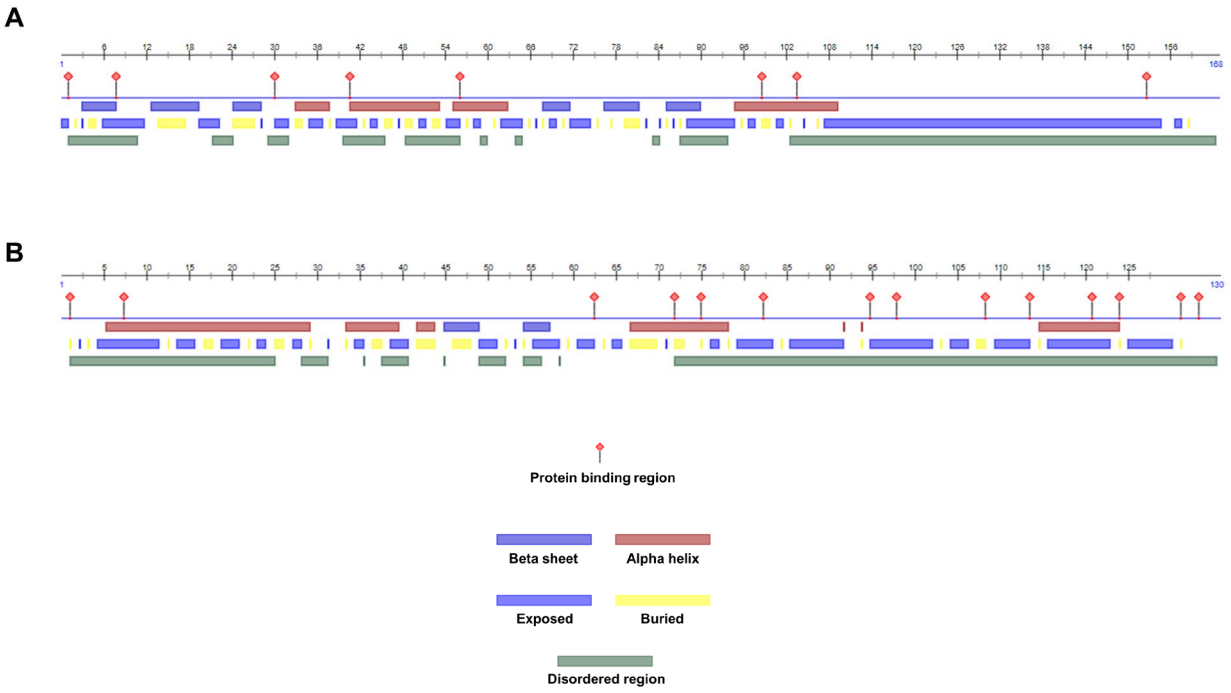


Figure 4.2 - Predicted structure of NHJ-1 protein isoforms

(A) Bioinformatically predicted structure of the long isoform of NHJ-1, showing predicted protein-protein interaction sites, secondary structures (beta sheets and alpha helices), water-exposed and buried regions, and disordered regions. The structures were predicted on the PredictProtein server. Notably, the last third of the protein is predicted as disordered.

(B) Same prediction as in (A), but for the short isoform of the NHJ-1 protein. The final 62 amino acids are predicted to be disordered.

Figure 4.3

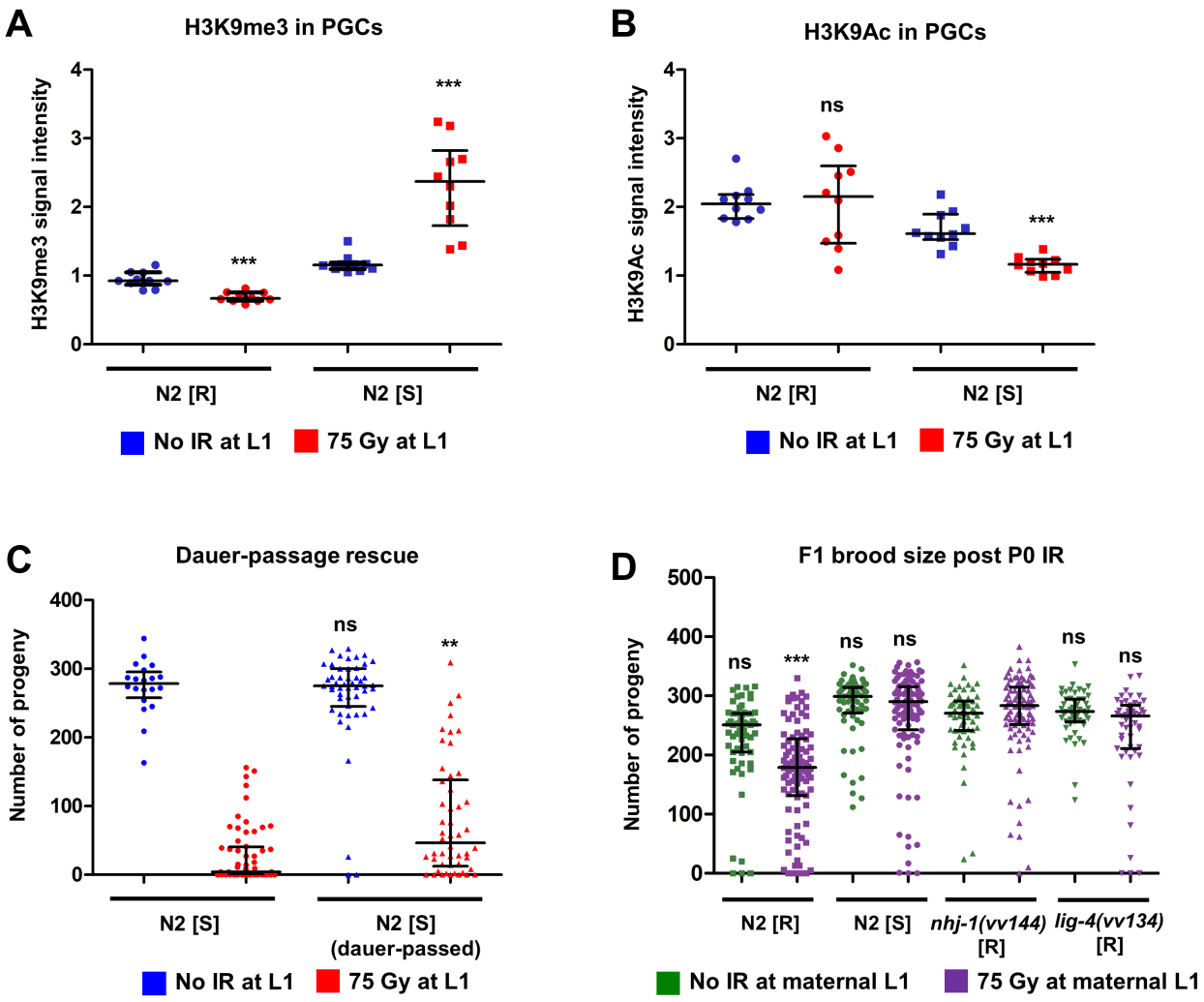


Figure 4.3 - Germline epigenetic phenotypes in N2 [S] and transgenerational fertility effects

(A) Quantification of H3K9me3 antibody signal in the PGCs of irradiated and control N2 [R] and N2 [S] animals. Compared to unirradiated controls, H3K9me3 levels are elevated in irradiated N2 [S] animals and reduced in irradiated N2 [R] animals ($p<0.001$). H3K9me3 antibody signal is normalized to HTP-3 antibody signal. All statistical comparisons are to the unirradiated group of the same genotype (Mann-Whitney test). Sample size (n) is 10 individual PGCs in 10 L1 larvae for all groups.

Figure 4.3 - Germline epigenetic phenotypes in N2 [S] and transgenerational fertility effects (*continued*)

(B) Quantification of H3K9Ac antibody signal in the PGCs of irradiated and control N2 [R] and N2 [S] animals. Compared to unirradiated controls, H3K9Ac levels are reduced in irradiated N2 [S] animals ($p<0.001$) and unchanged N2 [R] animals ($p>0.05$). H3K9Ac antibody signal is normalized to HTP-3 antibody signal. All statistical comparisons are to the unirradiated group of the same genotype (Mann-Whitney test). Sample size (n) is 10 individual PGCs in 10 L1 larvae for all groups.

(C) Total brood size quantification in irradiated and dauer-passed N2 [S] animals. Dauer passage is able to slightly but significantly rescue the brood size of post-irradiated N2 [S] animals ($p<0.01$). All comparisons shown in the figure are to non dauer-passed animals of the corresponding treatment group (Mann-Whitney test, Bonferroni corrected for multiple comparisons to $\alpha = 0.01$). Sample size (n) is 20 in the non dauer-passed unirradiated group, 60 in the non dauer-passed irradiated group, and 50 for both dauer-passed groups.

(D) Total brood size quantification in N2 [R], N2 [S], *lig-4(vv134)* [R] and *nhj-1(vv144)* [R] F1 animals from irradiated and control P0 mothers. The F1 progeny of the irradiated sensitive genotypes N2 [S], *lig-4(vv134)* [R], and *nhj-1(vv144)* [R] exhibit a brood size not significantly different from the F1 progeny of unirradiated controls of the same genotype ($p>0.05$ for all three genotypes). The F1 progeny of irradiated N2 [R], by contrast, show a significantly lower brood size than the F1 progeny of unirradiated N2 [R] control animals ($p<0.05$). All statistical comparisons shown in the figure are to *nhj-1(vv144)* [R] animals of the corresponding treatment group (Kruskal-Wallis test, followed by Dunn's post-hoc test). Sample size (n) is 59 and 94 for F1s from unirradiated and irradiated N2 [R] P0s, respectively; 61 and 98 for F1s from unirradiated and irradiated N2 [S] P0s, respectively; 56 and 94 for F1s from unirradiated and irradiated *nhj-1* [R] P0s, respectively; and 50 and 51 F1s from unirradiated and irradiated *lig-4* [R] P0s, respectively.

IR = ionizing radiation; Gy = Gray (unit)

N2 [S] = sensitive N2 strain, derived from the CGC N2
N2 [R] = resistant N2 strain, derived from Andersen lab N2
nhj-1(vv144) = a deletion mutant generated by CRISPR in the N2 [R] genetic background
lig-4 = *lig-4(vv134[R18STOP])* [R], a predicted null allele of *lig-4* generated in the N2 [R] genetic background

ns = not significant ($p>0.05$ in **(B)** and **(D)**; $p>0.01$ in **(C)**), ** = $p<0.01$; *** = $p<0.001$

Chapter V: Materials and Methods

V.1 Introduction

This chapter details all of the materials and methods relevant for understanding the experiments described in the present study. It contains three supporting figures and one supporting table.

V.2 *Caenorhabditis* strain maintenance

All *C. elegans* and *C. briggsae* strains have been maintained under standard conditions, at 20°C on 60mm x 15 mm petri dishes (Thermo Fisher Scientific) containing Nematode Growth Medium (NGM) with the *E. coli* strain OP50 as a food source (Brenner 1974, Stiernagle 2006). These techniques apply to all figures presented in **Chapters II-V**. For a list of strains used in this study, see **Table 5.1**.

V.3 *C. elegans* mating

All *C. elegans* mating, whether for crosses or the mating assay, was done according to prior published protocol (Fay 2006), with the exception that the male:hermaphrodite ratio was at least 7:1 per mating plate. This methodology applies to results from **Figures 2.7**, **2.9**, **2.12**, and **3.3**, as well as to any crosses done in strain creation (see **Table 5.1**).

V.4 Ionizing radiation treatment

Animals were treated with ionizing radiation either with a radioisotopic source (Cesium-137 in a Gammacell 40, Best Theratronics Ltd) at the rate of 2.5 Gray per minute, or, after the Gammacell 40 was decommissioned, with X-rays (RS 2000 small animal X-ray irradiator, Rad Source Technologies Inc) at the rate of 2.34 Gray per minute. Control animals were kept next to the IR source during the irradiation. The Cesium-137 source was used for IR irradiation in the RNAi screen and early experiments (**Figures 2.1** and **2.2**), while the X-ray irradiator was used for IR treatment in **Figures 2.3 - 2.10**, **2.13**, **2.14**, **Tables 2.1** and **2.2**, and **Figures 3.3 - 3.7**, **3.9**, **4.3**.

V.4.1 IR irradiation of L1 animals

For homozygous L1 animals, irradiation was performed in M9 buffer in 1.5 ml microcentrifuge tubes. Synchronized L1 animals were obtained by hypochlorite treatment as previously described (Porta-de-la-Riva, Fontrodona et al. 2012), with the following modifications: washes were done in water instead of M9 buffer, the animals were treated with hypochlorite for 10-12 minutes, and the hypochlorite solution recipe used was 3.3 ml water, 1.2 ml sodium hypochlorite (4%), 0.5 sodium hydroxide (0.5M). After hypochlorite treatment, the eggs were left overnight in M9 buffer in 15 ml centrifuge tubes to hatch. Two hours before irradiation, a concentrated culture of OP50 *E. coli* in the amount totaling 10% of the M9 buffer volume was added to the 15 ml centrifuge tubes as a food source for the L1-arrested larvae. Immediately before irradiation, the larvae were transferred to 1.5 ml microcentrifuge tubes, and irradiated as described above. Following irradiation, the animals were transferred by glass Pasteur pipettes to fresh NGM plates. This method of L1 irradiation was used in **Figures 2.1**, **2.3 - 2.8**, **2.10**, **2.13**, **3.2**, **3.5A-B, 3.6 - 3.10**, **4.1**, and **Table 2.2**.

V.4.2 IR irradiation of L1 cross-progeny

For heterozygous and homozygous L1 cross-progeny, irradiation was performed on NGM plates. Near-synchronized cross-progeny L1s were obtained by isolating mated hermaphrodites on NGM plates, allowing them to lay eggs for 4 hours, and then removing them from the plates. The plates containing hatched L1 larvae were irradiated 12-14 hours after the removal of the mothers. This method of L1 irradiation was used in **Figures 2.9**, **3.3**, and **3.5C-D**.

V.4.3 IR irradiation of L4 animals

For L4 animals, irradiation was performed on NGM plates. Animals were synchronized using hypochlorite treatment as described above (section **V.4.1**) and dispensed onto NGM plates. After 48 hours, the L4 animals were irradiated directly on the plate. This method of irradiation was used in **Figure 2.5B**, **Table 2.2**, and **Table 4.1**.

V.5 ENU treatment

For ENU treatment, synchronized L1 larvae were obtained by hypochlorite bleaching and provided with OP50 as described above (section **V.4.1**). After allowing 2 hours for feeding, the larvae were incubated in a 15 ml centrifuge tube with the working solution of ENU as described in (Kutscher and Shaham 2014), except: L1 larvae were used instead of L4 larvae, and the working concentration of ENU used was 5 mM and 10 mM. Control animals were kept in tubes containing only M9, next to the ENU tubes. This method was used in **Figure 2.5C**.

V.6 UV treatment

For UV irradiation, synchronized L1 larvae were likewise obtained by the hypochlorite bleaching method described in section **V.4.1**. Following synchronization, the L1 larvae were dispensed on NGM plates with a glass Pasteur pipette, and irradiated on plates with 50 J/m^2 or 100J/m^2 of UV-C in a Stratalinker 1800 UV crosslinker (Stratagene California). Control animals were kept on plates next to the crosslinker. This method was used in **Figure 2.5D**.

V.7 RNAi screen

The RNAi screen was performed using the feeding RNAi method (Timmons and Fire 1998), on a published RNAi sublibrary of chromatin-associated genes (Tursun, Patel et al. 2011) including only the clones from the Ahringer RNAi library (Kamath, Fraser et al. 2003). Synchronized L1 larvae were obtained as described in section **V.4.1**, and dispensed on 12-well plates, irradiated, and the population scored 4 days later for the combinatorial effects of RNAi and IR treatment as described in **Figure 5.1**.

V.8 Brood size scoring

Following IR irradiation, ENU treatment, or UV irradiation, effects on brood size were assessed by one of the methods described below.

V.8.1 Brood size scoring, total brood

Animals were isolated on separate NGM plates immediately after treatment and allowed to develop for 72-76 hours (in the case of treated L1s) or 24-28 hours (for treated L4s), at which

point each animal was switched to a fresh plate, and the original plate kept. This procedure was repeated 24-28 hours after the first transfer, and the animal was then left on the third plate. The number of animals on the first plate (brood 1) was scored with a handheld lap counter two days after the first transfer, and the number of animals on the second and third plate (brood 2 and 3) was scored two days after the second transfer. The total brood size was then calculated by combining all three broods. This method of brood size scoring was used in **Figures 2.1C**, **2.3**, **2.5**, **2.9**, **2.10**, **2.13**, **3.2A**, **3.3A**, **3.4A**, **3.4C**, **3.5A**, and **3.5C**. For **Figure 2.7**, this method was employed with the following modification: hermaphrodites were isolated at the L4 stage, mated overnight as in section **V.3,** and then moved to a fresh plate while males were removed from the first brood plate. A post-mating brood size was scored only if there was a high incidence of males in the progeny, with plates with no progeny discounted. For **Figure 4.3C**, IR-treated and control animals were first induced into dauer on dauer pheromone plates at 25°C (Ailion and Thomas 2000). Two days following treatment, dauer larvae were identified by morphology and put on individual plates to resume development, and the brood size was scored as above. For **Figure 4.3D**, F1 animals from treated and untreated groups were isolated as L1 larvae, and the above method followed.

V.8.2 Brood size scoring, category binned

Category binned brood size scoring was performed as in section **V.8.1**, with the exception that the brood size for each plate was scored only up to 51 progeny, and the brood sizes then divided into one of four categories: 0 progeny, 1-10 progeny, 11-50 progeny, and more than 50 progeny. This method was used in **Figure 2.1A** and **2.1B**.

V.9 Scoring of somatic phenotypes

Following IR, UV, or ENU exposure, animals were transferred onto NGM plates (or left on the plate if treated on plates), and left to develop for three days (72-76 hours) or four days (96-100 hours). At those time points, the incidence of somatic phenotypes was assessed in the following way. Protruding vulva, ruptured through vulva, and larvae were scored directly on the plate using a Leica MS5 stereomicroscope. This scoring method was used to obtain the data in **Figures 2.8A-B**, **3.2B**, **3.3B**, **3.4B** and **3.4D**, and **3.5B**. A variant method was used for the cross progeny in **2.9B**

and **3.5D**, where the animals were isolated to individual plates as L1 following treatment, and the somatic phenotype scored for each animal at three and four days after treatment.

The IR-resistance/sensitivity of the strains tested in **Table 2.1** was assessed by visual inspection three days after irradiation with 75 Gy of IR, with strains scoring Gro and high incidence of vulval phenotypes scored as sensitive and strains showing no Gro and no or low incidence of vulval phenotypes scored as resistant.

The proportion of worms showing a bag of worms phenotype was scored by dissecting individual animals under a Leica MS5 stereomicroscope using hypodermal injection needles (Becton, Dickinson and Company). Animals were scored as bagging if they contained one or more hatched larvae within their body at time of dissection. This method was used to obtain the data in **Figure 2.8D**.

V.10 Scoring of embryonic lethality

Groups of IR-treated and control animals were moved to a fresh NGM plate 24 hours after the L4 stage, and moved to a fresh plate two times after that in 8 hour intervals. After 8 hours on the last plate, the animals were removed, leaving three brood plates. The number of eggs was scored on each plate following the transfer or removal of the animals, and the number of hatched larvae was counted on each plate ~24 hours after the transfer or removal of the mothers. For L1 treated animals, 10-12 animals per plate were used for unirradiated and IR-treated resistant groups, and 20 animals per plate for IR-treated sensitive groups. For animals treated at L4, 10 animals per plate were used for unirradiated controls and 12-15 animals were used for IR-treated groups. This method was used to obtain the data in **Table 2.2** and **Table 4.1**.

V.11 CRISPR-Cas9 mutagenesis

CRISPR-Cas9 mutagenesis was performed using the *in vitro* assembled ribonucleoprotein complex as described (Paix, Folkmann et al. 2015). Young adult animals were microinjected the Cas9/tracrRNA/crRNA/DNA repair template mixture 1 day after L4 stage. The *dpy-10(cn64)* allele was used as a co-conversion marker (Arribere, Bell et al. 2014). Heterozygous *dpy-10(cn64/+)* rollers were isolated from the progeny of injected animals, allowed to lay progeny, then lysed

and genotyped for the mutation of interest. If positive, wild-type moving F2 progeny was isolated and genotyped for homozygosity of the mutation of interest. TracrRNA and crRNAs were synthesized by GE Healthcare Dharmacon, Inc. The repair template DNA oligonucleotides were synthesized by Integrated DNA Technologies, Inc. The Cas9 endonuclease was purchased from PNA Bio Inc. The crRNAs and repair templates used to create mutations used in this study are listed below.

dcr-1 crRNA (DNA target): GAGATCATACGTTCAGTCAA

dcr-1(mg375[G492R)] repair template:

CAACAGTGATAACATATCTTGATCCAGCTCGCCGAGCTCTTCTCTTTGACTGAACGTATGATCTCATGTCA
AGTGGTCGATCAAATTTAATCACAAGATTACATTGTTTAACATC

lig-4 N-term crRNA (DNA target): TTGACGTCTTCAACAAGATT

lig-4(vv134[R18STOP]) repair template:

ATGGCGTCAGATGTGATCTTCGACGAAGTAGTTGACGTCTTCAACAAGATTTGACGGACTTCAAATGTG
AAATCAAAGCAAGCAACCTTTCAGAAAAACTTTGAATCATGGAAAG

lig-4 C-term crRNA (DNA target): CGAAGGTGGATTCGAGATTC

lig-4(vv145[lig-4::OLLAS]) repair template:
TGGTTGCCTTCTGATGTGTTTCATGCCATCGAAGGTGGATTCGAGATTCAGGAATACCCATATGATGTCC
CGGATTACGCTTAATTTACTAATTTCGATTATATGTGATATCGCTCTTTATTTCCTTTTT

F10D2.12 crRNA (DNA target): TTACCCAAAAGTGCAGTTTG

F10D2.12(vv136[P325L]) repair template:
TGACATTGAGGACTTTGCACGAGGAATTCAAAATATTCATTTTGTAAAATGGGTACTACAAACTGCACTT
TTGGGTAAGTTTTTATTCCGTATTTCAAGTTACCGTATTTTCTAT

inft-2 crRNA (DNA target): TTCTGAAACTCAGACAGATG

inft-2(vv135[S239Y]) repair template:

TGCAAGTATTTCCGGGCTTCGTGTCTCCTGTTCTTCGAATGTTCTCACAAAAGATTATGAAACTCAGACAG
ATGTAGTCCAGCGGAAATTGAACAAGGAAAGGAGGGAGCCTTCT

nhj-1 crRNA (DNA target): CTAGAGCGTACGGAGCTTTC

nhj-1(144) repair template:

TTCCCTCTTCTCTGAAAGTGGCCTTCATATCGAAGAAGTTTGTGAGAAGGAGCTTGTGCTCACGTTTGCA
CTTCAAAGAAATATTAGCGTACGGAGCTTTCTGGAAGATTCGATGAGTGCCTCTACTTGGAATATTTGGT
GGGTTAAAAAGTT

nhj-1 C-term crRNA (DNA target): ATCGTCAAACTCTGGTCCAC

nhj-1(vv147[nhj-1::OLLAS]) repair template:

TATGGCTGCCAAGGCCAGTGGACCAGAGTTTGACGATGAATCTGGATTCGCTAACGAGCTTGGACCACG
CCTTATGGGAAAGTAATTTCACAATTAATTAACCCCATCTTTCTTGTTCCATG

V.12 Genomic DNA preparation and sequencing

Genomic DNA for deep sequencing was prepared using the Schedl lab protocol, which can be found in full at the following URL: http://genetics.wustl.edu/tslab/protocols/genomic-stuff/worm-genomic-dna-prep. Briefly, 5 medium-sized (10 cm in diameter) NGM plates were seeded with 15 L4 animals and allowed to starve over the course of a week, resulting in a large number of arrested L1 larvae. The L1 larvae were then grown in liquid NGM 3-4 days until they developed into adult animals, which were then treated with a 30% sucrose float to remove food contamination and separated into 500 µl aliquots which were then frozen at -80°C. An aliquot was then transferred to a 15 ml centrifuge tube, 4.5 ml of worm lysis buffer (0.1M Tris-Cl pH 8.5, 0.1M NaCl, 50 mM EDTA pH 8.0, 1% SDS) and 200 µl of Protease K (20 mg/ml in TE pH 8.0) added, and the worms vortexed. The mixture was incubated for 1 hour at 62°C, with intermittent vortexing. Then, 800 µL of 5M NaCl and was added and the tube mixed by inversion, after which 800 µL of CTAB solution (10 % CTAB in 0.7M NaCl) was added and the tube incubated for 10 minutes at 37°C. Following this, 7 ml of chloroform was added and the tube mixed and spun, the aqueous phase recovered, and the step repeated with 7 ml phenol/chloroform/isoamyl alcohol.

Next, 0.6 volume of -20°C isopropyl alcohol was added, and mixed, and the DNA spun at 4°C for 5 minutes. The DNA pellet was washed in 70% ethanol, dried, and resuspended in 340 µl of TE buffer. Next, 10 µl of RNase A (10 mg/ml) was added and the tube incubated for 2 hours at 42°C, following which 20 µl of 20 % SDS, 10 µl of 0.5 M EDTA pH 8.0, and 20 µl of Protease K was added and the tube incubated for 2 hours at 65°C. Then, 40 µl of 10 M Ammonium Acetate was added, the DNA extracted twice with phenol/chloroform/isoamyl alcohol and once with chloroform, 1 ml of ethanol added, and the DNA spun down at 4°C for 10 minutes. The DNA was washed twice with 70% ethanol, dried, and resuspended in 200 µl of TE buffer.

Sample paired-end tag libraries were prepared by Canada's Michael Smith Genome Sciences Centre, and the samples were sequenced with Illumina HiSeq 2500 (125 bp read length) to a coverage of 100X for N2 [S] and 200X for N2 [R]. The Genome Sciences Centre also provided the binary alignment (bam) files for both genomes.

V.13 Bioinformatics

Sequence variants in N2 [S] and N2 [R] were called with SAMtools (Li, Handsaker et al. 2009), using the mpileup function against the WS249_cel235.fa reference genome. Filtering was performed and strain-specific variants determined using the somatic variation function in the small variant caller Strelka2 (Kim, Scheffler et al. 2018), except the *nhj-1(vv148)* mutation, which was identified by manual parsing through variants called by SAMtools mpileup. The full sequence of the *nhj-1(vv148)* indel was identified by N2 [S] genome reassembly with ABySS 2.0 (Jackman, Vandervalk et al. 2017) from the sorted bam file.

The search for protein sequences homologous to NHJ-1 was conducted with DELTA-BLAST on the NCBI online tool, https://blast.ncbi.nlm.nih.gov/Blast.cgi (Boratyn, Schaffer et al. 2012). NHJ-1 isoform sequences were analyzed for domain conservation by SMART (Simple Molecular Architecture Research Tool), http://smart.embl-heidelberg.de (Letunic and Bork 2018). The hairpin in *nhj-1(vv148)* insertion was predicted using the ViennaRNA package 2.0 (Lorenz, Bernhart et al. 2011).

NHJ-1 protein structure was predicted with PredictProtein https://open.predictprotein.org (Rost, Yachdav et al. 2004), which integrates several bioinformatic tools (Rost, Fariselli et al. 1996,

Bigelow, Petrey et al. 2004, Ofran and Rost 2007, Schlessinger, Punta et al. 2009, Hamp, Kassner et al. 2013) to generate predictions about protein and nucleic acid interaction, secondary structures, and disordered regions.

V.14 Mapping

The mapping of the IR-sensitivity-causative locus (*nhj-1*) was done in two rounds, using N2 [S]- and N2 [R]-specific molecular markers identified through deep sequencing to assay marker segregation in F2 hybrid strains originating from an N2 [S] X N2 [R] cross. In the first, low-resolution mapping round, N2 [S]-specific candidate loci were PCR amplified and Sanger sequenced at the Genome Quebec center at McGill University Campus. In the second, high-resolution mapping round, *F10D2.12* and *inft-2* were PCR amplified and genotyped using RFLPs, with KpnI cutting the [R]-form of *F10D2.12* and TfiI cutting the [R]-form of *inft-2*, but neither enzyme being able to digest the [S]-form of either gene. For more information on the pipeline and recombination outcomes, see **Figure 5.2**.

V.15 Genotyping of the *nhj-1* locus

The wild type *nhj-1* locus and *nhj-1(vv144)* were amplified using the following gene flanking primers: TTGTGTTGAAACTGTACCGTCT and CAAAGTAGTCCCCCTAATCGCA. Digestion of the resulting product with XbaI yields two bands on *nhj-1(+)* but does not cut *nhj-1(vv144)*. The flanking primer pair does not yield a product with *nhj-1(vv148)*, so this allele was amplified with the primers: TTGTGTTGAAACTGTACCGTCT and TAATAATATTTTTAATAAATAATAGTAATAT, where the second primer partially anneals in the hairpin insertion of *nhj-1(vv148)*.

V.16 RT-PCR

Total RNA was extracted from N2 [S] and N2 [R] using TRIzol reagent (Thermo Fisher Scientific) per the manufacturer's instructions. cDNA was synthesized using the iScript reverse transcription kit (Bio-Rad Laboratories) following the manufacturer's instructions. Qualitative PCR was performed on the cDNA with the following primer pairs:

cku-70:

F1: CAACAAGCGATCGTGTGCAA and R1: GGCGGAGCAAATCAGGTTTC

F2: GAAACCTGATTTGCTCCGCC and R2: GGCTCCGTGAGCACATTTTG

F3: TCGCGGAGCCTTAACAATCA and R3: GGAGGTTTTCTTCCTCCGCA

cku-80:

F1: TCAAAGGTGTTCTGGGAGCC and R1: CGATTCCAATTATCATCAAATCTGC

F2: TGACAAATGGATTAAACGAGAACA and R2: AGCATCTTTGCGACGTTTTGAA

F3: CTGACAAATGGATTAAACGAGAACA and R3: GCATCTTTGCGACGTTTTGAA

lig-4:

F1: TGTGGCGAAACATGAGGAGG and R1: AGCGCTTCTTTGTCCTTTCAC

F2: ACGTGAAAGGACAAAGAAGCG and R2: AGCCGCAAATCTGTAGAGCG

F3: TGGATGATGGGAACAAACGAGA and R3: AGCACAAGCAGAAGGCGATA

V.17 Immunostaining

Gonad and intestinal immunohistochemistry was performed on dissected organs as described in (Martinez-Perez and Villeneuve 2005), with the following adjustments: dissection was performed in M9 buffer, four washes with PBST preceded the blocking, blocking was done with 1% BSA in PBST, slides were incubated with primary antibodies at 4°C, slides were washed with 1% BSA PBST four times before addition of secondary antibodies, incubated with secondary antibodies for 2 hours at room temperature, and washed again four times with PBST before addition of DAPI in Vectashield mounting medium (Vector Laboratories Inc). This method was used to obtain to prepare the tissues imaged in **Figures 2.6**, **3.8A-C**, **3.10A-C**, and **3.11B**.

Immunohistochemistry of L1 larvae was performed using the freeze-crack method as described in (Butuci, Williams et al. 2015), with the following modifications: slides were left in -20°C for 1 minute, fixed with 1% formaldehyde in PBST for 5 minutes, washed four times with PBST before

blocking with 1% BSA in PBST, washed four times with 1% BSA in PBST before secondary antibody incubation, and washed four times with PBST before addition of DAPI in Vectashield. This method was used to prepare the larvae imaged in **Figures 3.6**, **3.7A-B**, **3.9A-C**, and **4.3A-B**.

The primary antibodies and concentrations used in this study are as follows: guinea pig α-HTP-3 (1:500) (Goodyer, Kaitna et al. 2008), rat α-OLLAS (1:200) (Novus Biologicals, Inc), mouse α-GFP (1:200) (Abcam), rabbit α-H3K9me3 (1:500) (Cell Signaling Technology, Inc), and rabbit α-H3K9Ac (1:200) (Cell Signaling Technology, Inc).

Secondary antibodies used in this study include: Alexa 488-conjugated α-guinea pig, Alexa 488-conjugated α-mouse, Alexa 555-conjugated α-rabbit, and Alexa 555-conjugated α-rat. All secondary antibodies were purchased from Molecular Probes Inc, and used at a concentration of 1:1000.

V.18 Microscopy

All worm manipulations, transfers, and crosses (section **V.3**), as well as brood size scoring (section **V.8**), somatic phenotype scoring (section **V.9**), embryonic lethality scoring (section **V.10**), and worm dissection (section **V.9** and **V.17**) was performed on Leica MS5 stereomicroscopes.

Example somatic phenotypes shown in **Figure 2.8C** were imaged with a 12-bit QICAM digital camera (QImaging and Photometrics) on a Leica MZ8 stereomicroscope.

Micrographs shown in **Figure 3.6** and those from which signal intensity was quantified in **Figure 4.3A-B** were acquired with a DeltaVision Image Restoration System (Applied Precision) on an Olympus IX70 florescence microscope (Olympus Scientific Solutions Americas Corp), in stacks of 15-25 Z-planes acquired in increments of 0.2 µm. Deconvolution and stack projection was performed in SoftWoRx 3.0 (Applied Precision). Fluorescence quantification was performed on stacks in SoftWoRx 3.0, and HTP-3 signal was used to normalize the H3K9me3 and H3K9Ac signal.

Micrographs shown in **Figures 2.6**, **3.7A-B**, **3.8A-C**, **3.9A-C**, **3.10A-C**, and **3.11B** were acquired with a Leica DMI 6000B inverted microscope and EM CCD camera C1900 (Hamamatsu Photonics KK). The DAPI signal was acquired with wide-field X-Cite 120 florescence illumination system (Excelitas Technologies), while the Alexa-488 and Alexa-555 conjugated antibody signals were

acquired with a Quorum WaveFX spinning disc confocal system (Quorum Technologies), both integrated with the Leica DMI 6000B microscope. Images were acquired in stacks of 15-40 Z-planes in increments of 0.2 μm. Stack projections and contrast and brightness adjustments were performed in ImageJ (National Institutes of Health and Laboratory for Optical and Computational Instrumentation).

V.19 RNAi of *cku-80*

RNAi knockdown of *cku-80* was done according to the standard feeding protocol (Conte, MacNeil et al. 2015). Heterozygous *com-1* animals were put on plates containing α-*cku-80* bacteria at the L4 stage, and the F1 progeny put onto individual α-*cku-80* plates for scoring of embryonic lethality (see section **V.10**). Control animals were fed bacteria expressing the empty vector L4440.

V.20 Scoring of DAPI bodies

The number of DAPI-staining bodies in **Figure 3.11C** was determined in immuno-stained gonads (**V.17**) by imaging diakinesis nuclei of -1, -2, and -3 oocytes in stacks of 20-40 Z-slices 0.2 micron thick using the EM CCD camera C1900 (Hamamatsu Photonics KK) under the Leica DMI 6000B inverted microscope (**V.18**), and counting the number of discrete entities in the stacks.

V.21 Statistical analyses, descriptive statistics, and data presentation

In every experiment where brood size was counted, at least one experimental group did not pass the Shapiro-Wilk test for normality (Shapiro and Wilk 1965). These include Non-normal distributions were also observed in the number of DAPI bodies (**Figure 3.11C**) and fluorescence signals (**Figure 4.3A-B**). The statistical significance of brood size, DAPI body number, and fluorescence intensity is therefore analyzed by several non-parametric tests, which do not assume normality. The most commonly employed such test in the present study is the Kruskal-Wallis H-test (Kruskal and Wallis 1952), considered a "non-parametric ANOVA". Because the Kruskal-Wallis test does not perform pairwise comparisons between individual groups, Dunn's post-hoc test is used after a significant difference has been found by the Kruskal-Wallis tests in at least one group (Dunn 1964), or a Mann-Whitney U-test (Mann and Whitney 1947) is used for each individual comparison, and a Bonferroni correction (the division of the α-value of the test

by the number of individual comparisons (Shaffer 1995)) is applied. The same tests were used for binned brood size data as in **Figure 1A-B**.

Categorical data, including the incidence of post-IR somatic phenotypes and embryonic lethality, is analyzed by the venerable Pearson's Chi-squared test (Pearson 1900), and in cases of multiple comparisons compared against a Bonferroni-corrected α-value.

Because of the non-normal distribution of the brood size data and the non-parametric tests used to determine significance, the descriptive statistical metrics used both as error bars in the figures and reported in the text are the median and the interquartile range, rather than the mean and standard deviation.

All statistical tests were performed in GraphPad Prism 5 (GraphPad Software Inc). Vertical scatter plots were generated in GraphPad Prism 5, and 100% stacked column bar graphs were generated in Microsoft Excel (Microsoft Corporation).

Tables and Figures

Table 5.1

Strain name	Genotype	Source
N2 (Zetka)	Wild type isolate	Obtained from the CGC
N2 (CGC)	Wild type isolate	Obtained from the CGC
N2 (Roy)	Wild type isolate	Generously provided by Dr. Richard Roy
N2 (Hekimi)	Wild type isolate	Generously provided by Dr. Siegfried Hekimi
N2 (Andersen)	Wild type isolate	Generously provided by Dr. Eric Andersen
N2 (Zhen)	Wild type isolate	Generously provided by Dr. Mei Zhen
N2 (NBRP)	Wild type isolate	Obtained from the NBRP of Japan
QX1211	Wild type isolate	Obtained from the CGC
JU775	Wild type isolate	Obtained from the CGC
CB4856	Wild type isolate	Obtained from the CGC
DL238	Wild type isolate	Obtained from the CGC
MY16	Wild type isolate	Obtained from the CGC
MY23	Wild type isolate	Obtained from the CGC
EG4724	Wild type isolate	Obtained from the CGC
JU258	Wild type isolate	Obtained from the CGC
LKC34	Wild type isolate	Obtained from the CGC
JU1088	Wild type isolate	Obtained from the CGC
ED3073	Wild type isolate	Obtained from the CGC
DL226	Wild type isolate	Obtained from the CGC
AB4	Wild type isolate	Obtained from the CGC
JU1171	Wild type isolate	Obtained from the CGC
AB1	Wild type isolate	Obtained from the CGC
JU1652	Wild type isolate	Obtained from the CGC
JU1896	Wild type isolate	Obtained from the CGC
VT847	*C. briggsae* wild type isolate	Obtained from the CGC
PB800	*C. briggsae* wild type isolate	Obtained from the CGC
NL2098	*rrf-1(pk1417)* I	Obtained from the CGC
EZ208	*htp-3(vc75)* I	EMS mutant from TILLING screen
NL2099	*rrf-3(pk1426)* II	Obtained from the CGC
YY470	*dcr-1(mg375)* III	Obtained from the CGC
WM158	*ergo-1(tm1860)* V	Obtained from the CGC
NL2550	*ppw-1(pk2505)* I	Obtained from the CGC
NL5117	*ppw-2(pk1673)* I	Obtained from the CGC
WM160	*sago-1(tm1195)* V	Obtained from the CGC
WM154	*sago-2(tm894)* I	Obtained from the CGC
MT13649	*nurf-1(n4295)* II	Obtained from the CGC
RB1831	*set-22(ok2370)* V	Obtained from the CGC
EU1481	*orIs20* [*pie-1p::GFP::MOE*]	Generously provided by Dr. Richard Roy

Table 5.1 (*continued*)

Strain name	Genotype	Source
EZ412/N2 [S]	Wild type	Generated from a single **N2 (CGC)** animal
EZ413/N2 [R]	Wild type	Generated from a single **N2 (Andersen)** animal
EZ421	*dcr-1(vv121)* III [S]	Generated by CRISPR mutagenesis in **N2 [S]**
EZ422	*dcr-1(vv122)* III [R]	Generated by CRISPR mutagenesis in **N2 [R]**
EZ437	*inft-2(vv135) F10D2.12(vv136)* V [R]	Generated by CRISPR mutagenesis in **N2 [R]**
RB873	*lig-4(ok716)* III	Obtained from the CGC
EZ436	*lig-4(vv134)* III [R]	Generated by CRISPR mutagenesis in **N2 [R]**
EZ443	*lig-4(vv141)* III [S]	Generated by CRISPR mutagenesis in **N2 [S]**
EZ454	*nhj-1(vv144)* V [R]	Generated by CRISPR mutagenesis in **N2 [R]**
FX01203	*cku-80(tm1203)* III	Obtained from the NBRP of Japan
HBR1099	*unc-119(ed3)* III; *goeEx386 [WRM0635D_B04(pRedFlp-Hgr) (H19N07.3[21364]::S0001_pR6K_Amp_2xTY1c e_EGFP_FRT_rpsl_neo_FRT_3xFlag) dFRT::unc-119-Nat]*	Generously provided by Dr. Henrik Bringmann
EZ477	*unc-119(ed3)* III; *nhj-1(vv144)* V*; goeEx386*	Generated by crossing **HBR1099** with **EZ454**
EZ463	*cku-80(tm1203)* III; *nhj-1(vv144)* V	Generated by crossing **EZ454** with **FX01203**
EZ459	*nhj-1(vv147[nhj-1::OLLAS])* V [R]	Generated by CRISPR mutagenesis in **N2 [R]**
EZ461	*cku-80(tm1203)* III; *nhj-1(vv147[nhj-1::OLLAS])* V	Generated by crossing **YA942** with **EZ459**
EZ460	*lig-4 (vv134)* III; *nhj-1(vv147[nhj-1::OLLAS])* V	Generated by crossing **EZ436** with **EZ459**
EZ457	*lig-4(vv145[lig-4::OLLAS])* III [R]	Generated by CRISPR mutagenesis in **N2 [R]**
EZ464	*cku-80(tm1203) lig-4(vv145[lig-4::OLLAS])* III	Generated by crossing **YA942** with **EZ457**
EZ465	*lig-4(vv145[lig-4::OLLAS])* III; *nhj-1(vv144)* V	Generated by crossing **EZ454** with **EZ457**
MR156	*unc-119(ed3)* III*; rrIs01* [*elt-2::GFP unc-119(+)*] X	Generously provided by Dr. Richard Roy
EZ466	*unc-119(ed3)* III*; nhj-1(vv144)* V; *rrIs01* [*elt-2::GFP unc-119(+)*] X	Generated by crossing **MR156** with **EZ454**
GE4132	*unc-32(e189) com-1(t1626)/*qC1 [*dpy-19(e1259) glp-1(q339)*] III	Obtained from the CGC
EZ467	*unc-32(e189) com-1(t1626)/*qC1 [*dpy-19(e1259) glp-1(q339)*] III; *nhj-1(vv144)* V	Generated by crossing **GE4132** with **EZ454**
EZ476	*unc-32(e189) com-1(t1626) lig-4(vv134)/*qC1 [*dpy-19(e1259) glp-1(q339)*] III	Generated by CRISPR mutagenesis in **GE4132**

Table 5.1 - List of strains used in this study

This table presents the list of all strains used in this study, listing the strain name, genotype (and genetic background, if known), and the source from which the strain was obtained.

Figure 5.1

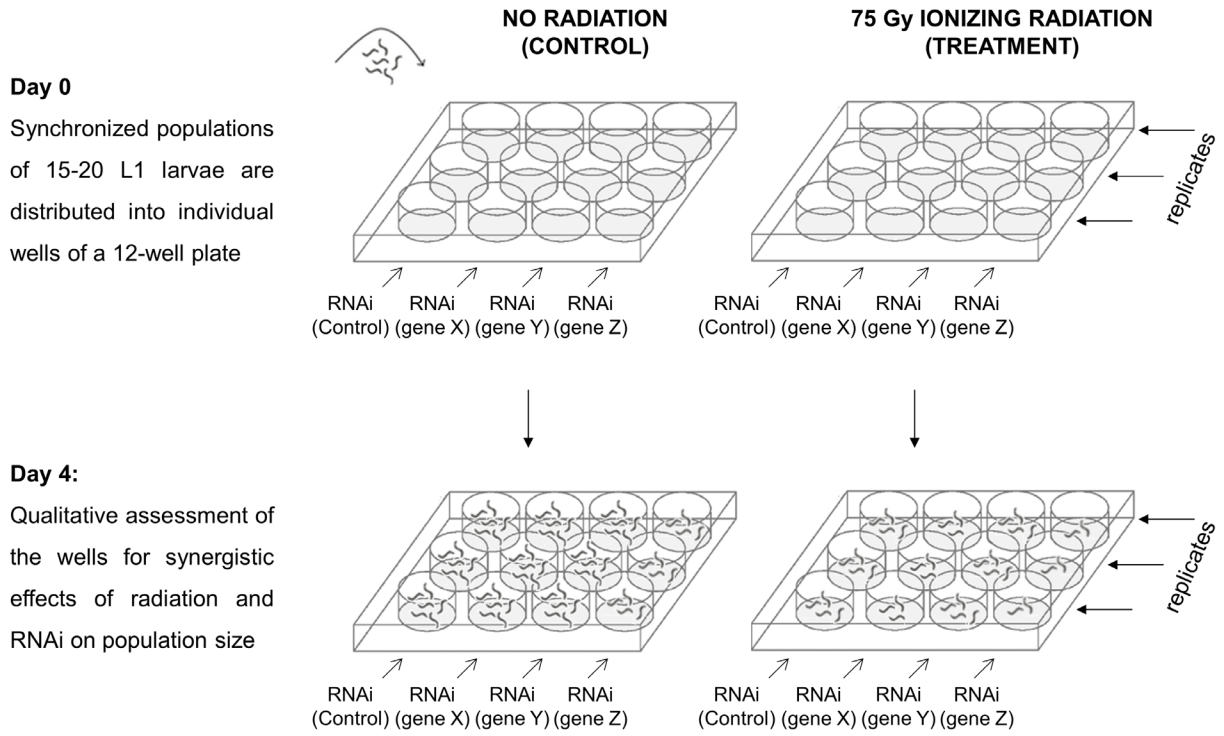
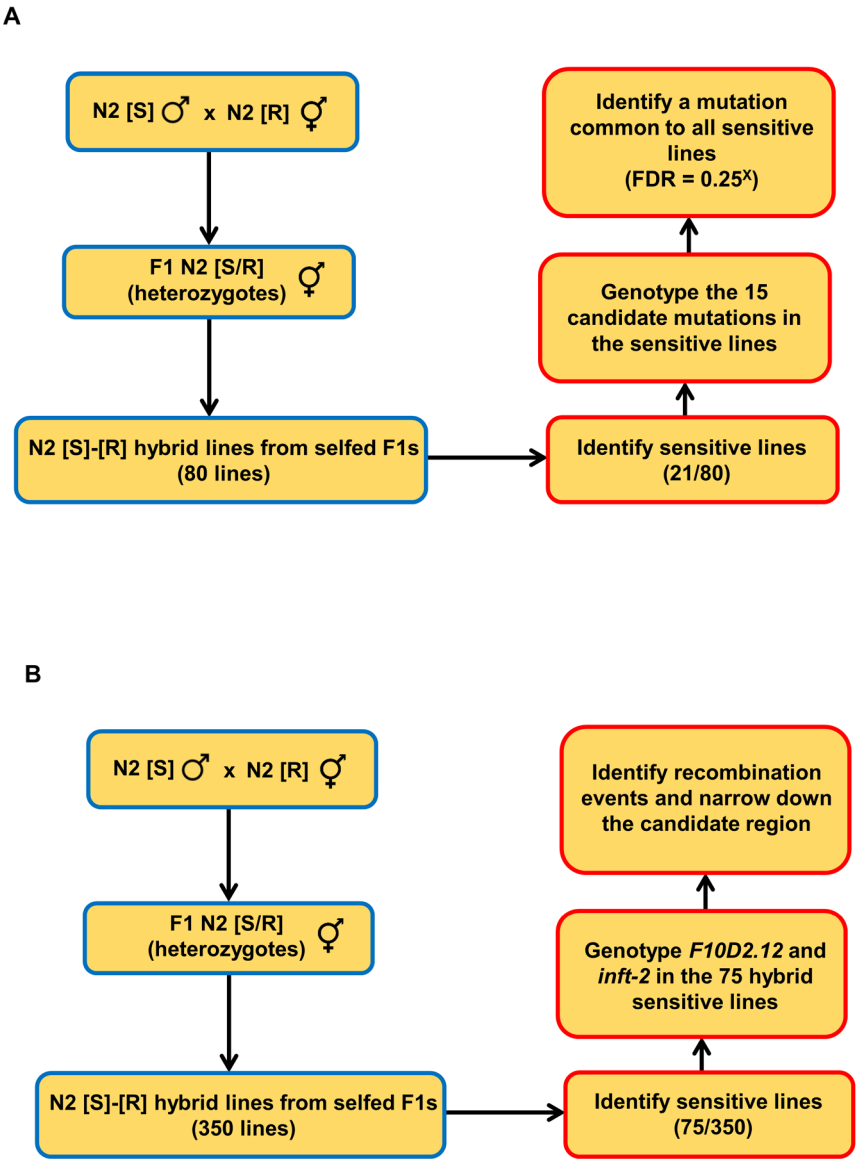


Figure 5.1 - L1 IR response RNAi screen method

This figure shows the screening strategy employed to identify candidate chromatin-associated genes affecting the IR response at the L1 stage. Twelve-well NGM plates (Fisher Scientific) were seeded with *E. coli* strain HB101 carrying the empty RNAi vector L4440 (Timmons and Fire 1998) in the first column of each plate, and clones from the chromatin factor RNAi sublibrary (Tursun, Patel et al. 2011) in the other three columns. At day 0, synchronized populations of L1 animals were dispensed (15-20 per well) into the wells, allowed to feed for 4 hours, and then irradiated with 75 Gy on the treatment plate or left unirradiated on the control plate. Four days later, each well was qualitatively examined for the combinatorial effects of RNAi treatment and IR exposure, and candidates scored as positive if the population-level amount of progeny was smaller or larger than would be expected from a merely additive effect of the two conditions. The rows in each plate acted as replicates, and each plate was also replicated three times, with hits scored as positive only if they consistently exhibited an effect on brood size.

Figure 5.2


A
N2 [S] ♂ x N2 [R] ♀
F1 N2 [S/R] (heterozygotes) ♂
N2 [S]-[R] hybrid lines from selfed F1s (80 lines)
Identify sensitive lines (21/80)
Genotype the 15 candidate mutations in the sensitive lines
Identify a mutation common to all sensitive lines (FDR = 0.25X)
B
N2 [S] ♂ x N2 [R] ♀
F1 N2 [S/R] (heterozygotes) ♂
N2 [S]-[R] hybrid lines from selfed F1s (350 lines)
Identify sensitive lines (75/350)
Genotype F10D2.12 and inft-2 in the 75 hybrid sensitive lines
Identify recombination events and narrow down the candidate region

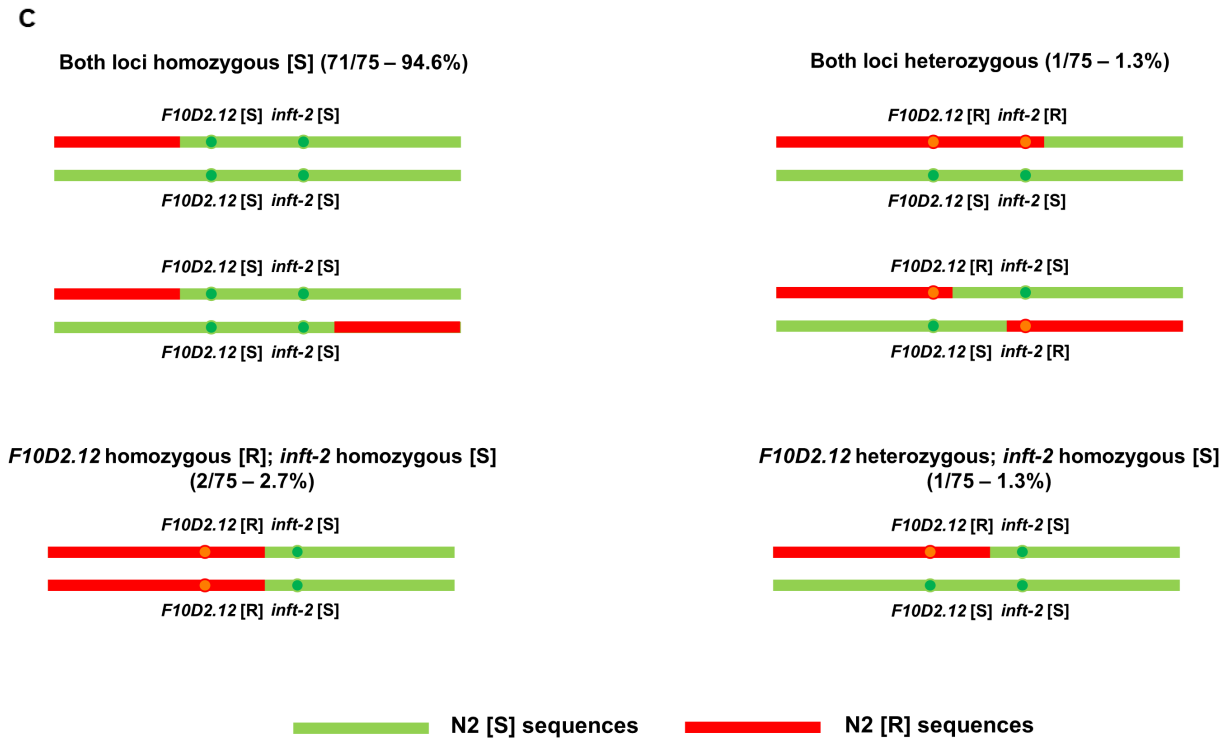


Figure 5.2 - Mapping of the IR-sensitivity-causative locus

(A) This pipeline shows the first line strategy in mapping. Twenty-one phenotypically IR-sensitive lines were derived from an N2 [S] X N2 [R] cross as shown here, and the 15 N [S]-specific, protein sequence-altering single nucleotide mutations sequenced in the hybrid IR-sensitive lines (**Figure 2.12**).

(B) For higher resolution mapping, another 75 phenotypically sensitive hybrid lines were derived, and *F10D2.12* and *inft-2* genotyped to identify recombination events between the causative locus and these two markers.

(C) This panel shows the possible recombination outcomes that explain the observed genotyping outcomes. The causative locus recombined away from *F10D2.12* in 4/75 strains and from *inft-2* in 1/75 strains, showing that it is located closer to *inft-2.*

FDR = false discovery rate; X = number of hybrid lines sequenced for each gene

Bibliography

Abraham, R. T. (2004). "PI 3-kinase related kinases: 'big' players in stress-induced signaling pathways." DNA Repair (Amst) **3**(8-9): 883-887.

Acevedo-Arozena, A., S. Wells, P. Potter, M. Kelly, R. D. Cox and S. D. Brown (2008). "ENU mutagenesis, a way forward to understand gene function." Annu Rev Genomics Hum Genet **9**: 49-69.

Adachi, N., H. Suzuki, S. Iiizumi and H. Koyama (2003). "Hypersensitivity of nonhomologous DNA end-joining mutants to VP-16 and ICRF-193: implications for the repair of topoisomerase II-mediated DNA damage." J Biol Chem **278**(38): 35897-35902.

Adamo, A., S. J. Collis, C. A. Adelman, N. Silva, Z. Horejsi, J. D. Ward, E. Martinez-Perez, S. J. Boulton and A. La Volpe (2010). "Preventing nonhomologous end joining suppresses DNA repair defects of Fanconi anemia." Mol Cell **39**(1): 25-35.

Agostinho, A., B. Meier, R. Sonneville, M. Jagut, A. Woglar, J. Blow, V. Jantsch and A. Gartner (2013). "Combinatorial Regulation of Meiotic Holliday Junction Resolution in C. elegans by HIM-6 (BLM) Helicase, SLX-4, and the SLX-1, MUS-81 and XPF-1 Nucleases." Plos Genetics **9**(7).

Aguinaldo, A. M., J. M. Turbeville, L. S. Linford, M. C. Rivera, J. R. Garey, R. A. Raff and J. A. Lake (1997). "Evidence for a clade of nematodes, arthropods and other moulting animals." Nature **387**(6632): 489-493.

Ahel, I., U. Rass, S. F. El-Khamisy, S. Katyal, P. M. Clements, P. J. McKinnon, K. W. Caldecott and S. C. West (2006). "The neurodegenerative disease protein aprataxin resolves abortive DNA ligation intermediates." Nature **443**(7112): 713-716.

Ahmed, E. A., A. Sfeir, H. Takai and H. Scherthan (2013). "Ku70 and non-homologous end joining protect testicular cells from DNA damage." J Cell Sci **126**(Pt 14): 3095-3104.

Ahnesorg, P., P. Smith and S. P. Jackson (2006). "XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining." Cell **124**(2): 301-313.

Ailion, M. and J. H. Thomas (2000). "Dauer formation induced by high temperatures in Caenorhabditis elegans." Genetics **156**(3): 1047-1067.

Alpi, A., P. Pasierbek, A. Gartner and J. Loidl (2003). "Genetic and cytological characterization of the recombination protein RAD-51 in Caenorhabditis elegans." Chromosoma **112**(1): 6-16.

Ames, B. N., M. Profet and L. S. Gold (1990). "Dietary pesticides (99.99% all natural)." Proc Natl Acad Sci U S A **87**(19): 7777-7781.

Andersen, E. C., J. P. Gerke, J. A. Shapiro, J. R. Crissman, R. Ghosh, J. S. Bloom, M. A. Felix and L. Kruglyak (2012). "Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity." Nature Genetics **44**(3): 285-U283.

Angelo, G. and M. R. Van Gilst (2009). "Starvation protects germline stem cells and extends reproductive longevity in C. elegans." Science **326**(5955): 954-958.

Ankeny, R. A. (2001). "The natural history of Caenorhabditis elegans research." Nat Rev Genet **2**(6): 474-479.

Aravind, L. and E. V. Koonin (2000). "SAP - a putative DNA-binding motif involved in chromosomal organization." Trends Biochem Sci **25**(3): 112-114.

Aravind, L. and E. V. Koonin (2001). "Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system." Genome Res **11**(8): 1365-1374.

Arribere, J. A., R. T. Bell, B. X. Fu, K. L. Artiles, P. S. Hartman and A. Z. Fire (2014). "Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans." Genetics **198**(3): 837-846.

Audebert, M., B. Salles, M. Weinfeld and P. Calsou (2006). "Involvement of polynucleotide kinase in a poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway." J Mol Biol **356**(2): 257-265.

Bailey, S. M. and J. P. Murnane (2006). "Telomeres, chromosome instability and cancer." Nucleic Acids Res **34**(8): 2408-2417.

Baille, D., A. Barriere and M. A. Felix (2008). "Oscheius tipulae, a widespread hermaphroditic soil nematode, displays a higher genetic diversity and geographical structure than Caenorhabditis elegans." Mol Ecol **17**(6): 1523-1534.

Bakkenist, C. J. and M. B. Kastan (2004). "Initiating cellular stress responses." Cell **118**(1): 9-17.

Barnes, G. and D. Rio (1997). "DNA double-strand-break sensitivity, DNA replication, and cell cycle arrest phenotypes of Ku-deficient Saccharomyces cerevisiae." Proc Natl Acad Sci U S A **94**(3): 867-872.

Barriere, A. and M. A. Felix (2005). "High local genetic diversity and low outcrossing rate in Caenorhabditis elegans natural populations." Curr Biol **15**(13): 1176-1184.

Barriere, A. and M. A. Felix (2007). "Temporal dynamics and linkage disequilibrium in natural Caenorhabditis elegans populations." Genetics **176**(2): 999-1011.

Barriere, A. and M. A. Felix (2014). "Isolation of C. elegans and related nematodes." WormBook: 1-19.

Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev and K. Zhao (2007). "High-resolution profiling of histone methylations in the human genome." Cell **129**(4): 823-837.

Bau, D. T., Y. C. Mau, S. L. Ding, P. E. Wu and C. Y. Shen (2007). "DNA double-strand break repair capacity and risk of breast cancer." Carcinogenesis **28**(8): 1726-1730.

Baugh, L. R. (2013). "To grow or not to grow: nutritional control of development during Caenorhabditis elegans L1 arrest." Genetics **194**(3): 539-555.

Beall, E. L., A. Admon and D. C. Rio (1994). "A Drosophila protein homologous to the human p70 Ku autoimmune antigen interacts with the P transposable element inverted repeats." Proc Natl Acad Sci U S A **91**(26): 12681-12685.

Bhargava, R., D. O. Onyango and J. M. Stark (2016). "Regulation of Single-Strand Annealing and its Role in Genome Maintenance." Trends Genet **32**(9): 566-575.

Bigelow, H. R., D. S. Petrey, J. Liu, D. Przybylski and B. Rost (2004). "Predicting transmembrane beta-barrels in proteomes." Nucleic Acids Res **32**(8): 2566-2577.

Blaxter, M. (2011). "Nematodes: the worm and its relatives." PLoS Biol **9**(4): e1001050.

Blaxter, M. and G. Koutsovoulos (2015). "The evolution of parasitism in Nematoda." Parasitology **142 Suppl 1**: S26-39.

Bleuyard, J. Y., M. E. Gallego and C. I. White (2006). "Recent advances in understanding of the DNA double-strand break repair machinery of plants." DNA Repair (Amst) **5**(1): 1-12.

Blier, P. R., A. J. Griffith, J. Craft and J. A. Hardin (1993). "Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks." J Biol Chem **268**(10): 7594-7601.

Blunt, T., N. J. Finnie, G. E. Taccioli, G. C. Smith, J. Demengeot, T. M. Gottlieb, R. Mizuta, A. J. Varghese, F. W. Alt, P. A. Jeggo and S. P. Jackson (1995). "Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation." Cell **80**(5): 813-823.

Boisvert, M. E. and M. J. Simard (2008). "RNAi pathway in C. elegans: the argonautes and collaborators." Curr Top Microbiol Immunol **320**: 21-36.

Boiteux, S. and M. Guillet (2004). "Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae." DNA Repair (Amst) **3**(1): 1-12.

Boratyn, G. M., A. A. Schaffer, R. Agarwala, S. F. Altschul, D. J. Lipman and T. L. Madden (2012). "Domain enhanced lookup time accelerated BLAST." Biol Direct **7**: 12.

Botchan, M., J. Stringer, T. Mitchison and J. Sambrook (1980). "Integration and excision of SV40 DNA from the chromosome of a transformed cell." Cell **20**(1): 143-152.

Boulton, S. J. and S. P. Jackson (1998). "Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing." EMBO J **17**(6): 1819-1828.

Bouwman, P. and J. Jonkers (2012). "The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance." Nat Rev Cancer **12**(9): 587-598.

Bowater, R. and A. J. Doherty (2006). "Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining." PLoS Genet **2**(2): e8.

Bowater, R. P. and Z. A. Waller (2014). DNA Structure. eLS, John Wiley & Sons.

Braithwaite, E., X. Wu and Z. Wang (1998). "Repair of DNA lesions induced by polycyclic aromatic hydrocarbons in human cell-free extracts: involvement of two excision repair mechanisms in vitro." Carcinogenesis **19**(7): 1239-1246.

Brenner, S. (1974). "The genetics of Caenorhabditis elegans." Genetics **77**(1): 71-94.

Buck, D., L. Malivert, R. de Chasseval, A. Barraud, M. C. Fondaneche, O. Sanal, A. Plebani, J. L. Stephan, M. Hufnagel, F. le Deist, A. Fischer, A. Durandy, J. P. de Villartay and P. Revy (2006). "Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly." Cell **124**(2): 287-299.

Buckland, R. J., D. L. Watt, B. Chittoor, A. K. Nilsson, T. A. Kunkel and A. Chabes (2014). "Increased and imbalanced dNTP pools symmetrically promote both leading and lagging strand replication infidelity." PLoS Genet **10**(12): e1004846.

Burton, N. O., K. B. Burkhart and S. Kennedy (2011). "Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans." Proc Natl Acad Sci U S A **108**(49): 19683-19688.

Butuci, M., A. B. Williams, M. M. Wong, B. Kramer and W. M. Michael (2015). "Zygotic Genome Activation Triggers Chromosome Damage and Checkpoint Signaling in C. elegans Primordial Germ Cells." Dev Cell **34**(1): 85-95.

Byerly, L., R. C. Cassada and R. L. Russell (1976). "The life cycle of the nematode Caenorhabditis elegans. I. Wild-type growth and reproduction." Dev Biol **51**(1): 23-33.

Cadet, J. and J. R. Wagner (2013). "DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation." Cold Spring Harb Perspect Biol **5**(2).

Cahoon, C. K. and R. S. Hawley (2016). "Regulating the construction and demolition of the synaptonemal complex." Nat Struct Mol Biol **23**(5): 369-377.

Caldecott, K. W. (2007). "Mammalian single-strand break repair: mechanisms and links with chromatin." DNA Repair (Amst) **6**(4): 443-453.

Caldecott, K. W. (2008). "Single-strand break repair and genetic disease." Nat Rev Genet **9**(8): 619-631.

Caldecott, K. W. (2014). "DNA single-strand break repair." Exp Cell Res **329**(1): 2-8.

Cantacessi, C., M. Mitreva, A. R. Jex, N. D. Young, B. E. Campbell, R. S. Hall, M. A. Doyle, S. A. Ralph, E. M. Rabelo, S. Ranganathan, P. W. Sternberg, A. Loukas and R. B. Gasser (2010). "Massively parallel sequencing and analysis of the Necator americanus transcriptome." PLoS Negl Trop Dis **4**(5): e684.

Caplin, N. and N. Willey (2018). "Ionizing Radiation, Higher Plants, and Radioprotection: From Acute High Doses to Chronic Low Doses." Front Plant Sci **9**: 847.

Castel, S. E. and R. A. Martienssen (2013). "RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond." Nat Rev Genet **14**(2): 100-112.

Ceccaldi, R., B. Rondinelli and A. D. D'Andrea (2016). "Repair Pathway Choices and Consequences at the Double-Strand Break." Trends Cell Biol **26**(1): 52-64.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward and D. C. Prasher (1994). "Green fluorescent protein as a marker for gene expression." Science **263**(5148): 802-805.

Chan, S. H., A. M. Yu and M. McVey (2010). "Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila." PLoS Genet **6**(7): e1001005.

Chandra, V., A. Bortnick and C. Murre (2015). "AID targeting: old mysteries and new challenges." Trends Immunol **36**(9): 527-535.

Chang, H. H. Y., N. R. Pannunzio, N. Adachi and M. R. Lieber (2017). "Non-homologous DNA end joining and alternative pathways to double-strand break repair." Nat Rev Mol Cell Biol **18**(8): 495-506.

Chapman, J. R., M. R. Taylor and S. J. Boulton (2012). "Playing the end game: DNA double-strand break repair pathway choice." Mol Cell **47**(4): 497-510.

Charbonnel, C., M. E. Gallego and C. I. White (2010). "Xrcc1-dependent and Ku-dependent DNA double-strand break repair kinetics in Arabidopsis plants." Plant J **64**(2): 280-290.

Chatterjee, N. and G. C. Walker (2017). "Mechanisms of DNA damage, repair, and mutagenesis." Environ Mol Mutagen **58**(5): 235-263.

Chen, L. H., G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour and C. P. Whitman (1992). "4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer." J Biol Chem **267**(25): 17716-17721.

Ciccia, A. and S. J. Elledge (2010). "The DNA damage response: making it safe to play with knives." Mol Cell **40**(2): 179-204.

Clejan, I., J. Boerckel and S. Ahmed (2006). "Developmental modulation of nonhomologous end joining in Caenorhabditis elegans." Genetics **173**(3): 1301-1317.

Clements, P. M., C. Breslin, E. D. Deeks, P. J. Byrd, L. Ju, P. Bieganowski, C. Brenner, M. C. Moreira, A. M. Taylor and K. W. Caldecott (2004). "The ataxia-oculomotor apraxia 1 gene product has a role distinct from ATM and interacts with the DNA strand break repair proteins XRCC1 and XRCC4." DNA Repair (Amst) **3**(11): 1493-1502.

Clouaire, T. and G. Legube (2015). "DNA double strand break repair pathway choice: a chromatin based decision?" Nucleus **6**(2): 107-113.

Cobb, N. A. (1915). "Nematodes and their relationships." U S Department of Agriculture Yearbook **1914**: (157-490).

Colaiacovo, M. P., A. J. MacQueen, E. Martinez-Perez, K. McDonald, A. Adamo, A. La Volpe and A. M. Villeneuve (2003). "Synaptonemal complex assembly in C-elegans is dispensable for loading strand-exchange proteins but critical for proper completion of recombination." Developmental Cell **5**(3): 463-474.

Compe, E. and J. M. Egly (2012). "TFIIH: when transcription met DNA repair." Nat Rev Mol Cell Biol **13**(6): 343-354.

Consortium, C. e. S. (1998). "Genome sequence of the nematode C. elegans: a platform for investigating biology." Science **282**(5396): 2012-2018.

Conte, D., Jr., L. T. MacNeil, A. J. Walhout and C. C. Mello (2015). "RNA Interference in Caenorhabditis elegans." Curr Protoc Mol Biol **109**: 26 23 21-30.

Cosman, M., C. de los Santos, R. Fiala, B. E. Hingerty, S. B. Singh, V. Ibanez, L. A. Margulis, D. Live, N. E. Geacintov, S. Broyde and et al. (1992). "Solution conformation of the major adduct between the carcinogen (+)-anti-benzo[a]pyrene diol epoxide and DNA." Proc Natl Acad Sci U S A **89**(5): 1914-1918.

Couteau, F. and M. Zetka (2011). "DNA damage during meiosis induces chromatin remodeling and synaptonemal complex disassembly." Dev Cell **20**(3): 353-363.

Craxton, A., J. Somers, D. Munnur, R. Jukes-Jones, K. Cain and M. Malewicz (2015). "XLS (c9orf142) is a new component of mammalian DNA double-stranded break repair." Cell Death and Differentiation **22**(6): 890-897.

Creer, S., V. G. Fonseca, D. L. Porazinska, R. M. Giblin-Davis, W. Sung, D. M. Power, M. Packer, G. R. Carvalho, M. L. Blaxter, P. J. Lambshead and W. K. Thomas (2010). "Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises." Mol Ecol **19 Suppl 1**: 4-20.

Cui, X., Y. Yu, S. Gupta, Y. M. Cho, S. P. Lees-Miller and K. Meek (2005). "Autophosphorylation of DNA-dependent protein kinase regulates DNA end processing and may also alter double-strand break repair pathway choice." Mol Cell Biol **25**(24): 10842-10852.

Cutter, A. D., M. A. Felix, A. Barriere and D. Charlesworth (2006). "Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of Caenorhabditis briggsae." Genetics **173**(4): 2021-2031.

Daley, J. M., P. L. Palmbos, D. Wu and T. E. Wilson (2005). "Nonhomologous end joining in yeast." Annu Rev Genet **39**: 431-451.

Daley, J. M. and T. E. Wilson (2005). "Rejoining of DNA double-strand breaks as a function of overhang length." Mol Cell Biol **25**(3): 896-906.

Davis, A. J. and D. J. Chen (2013). "DNA double strand break repair via non-homologous end-joining." Transl Cancer Res **2**(3): 130-143.

Davis, A. J., K. J. Lee and D. J. Chen (2013). "The N-terminal region of the DNA-dependent protein kinase catalytic subunit is required for its DNA double-stranded break-mediated activation." J Biol Chem **288**(10): 7037-7046.

de Villartay, J. P. (2015). "Congenital defects in V(D)J recombination." Br Med Bull **114**(1): 157-167.

de Villartay, J. P. (2015). "When natural mutants do not fit our expectations: the intriguing case of patients with XRCC4 mutations revealed by whole-exome sequencing." EMBO Mol Med **7**(7): 862-864.

Decottignies, A. (2013). "Alternative end-joining mechanisms: a historical perspective." Front Genet **4**: 48.

Deem, A., A. Keszthelyi, T. Blackgrove, A. Vayl, B. Coffey, R. Mathur, A. Chabes and A. Malkova (2011). "Break-induced replication is highly inaccurate." PLoS Biol **9**(2): e1000594.

Deriano, L. and D. B. Roth (2013). "Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage." Annu Rev Genet **47**: 433-455.

Derry, W. B., A. P. Putzke and J. H. Rothman (2001). "Caenorhabditis elegans p53: role in apoptosis, meiosis, and stress resistance." Science **294**(5542): 591-595.

Desouky, O., N. Ding and G. M. Zhou (2015). "Targeted and non-targeted effects of ionizing radiation." Journal of Radiation Research and Applied Sciences **8**(2): 247-254.

di Fagagna, F. D., G. R. Weller, A. J. Doherty and S. P. Jackson (2003). "The Gam protein of bacteriophage Mu is an orthologue of eukaryotic Ku." Embo Reports **4**(1): 47-52.

Dianov, G. L. and U. Hubscher (2013). "Mammalian base excision repair: the forgotten archangel." Nucleic Acids Res **41**(6): 3483-3490.

Dickinson, D. J. and B. Goldstein (2016). "CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering." Genetics **202**(3): 885-901.

Dobbs, T. A., J. A. Tainer and S. P. Lees-Miller (2010). "A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation." DNA Repair (Amst) **9**(12): 1307-1314.

Doherty, A. J., S. P. Jackson and G. R. Weller (2001). "Identification of bacterial homologues of the Ku DNA repair proteins." FEBS Lett **500**(3): 186-188.

Dolan, L. C., R. A. Matulka and G. A. Burdock (2010). "Naturally occurring food toxins." Toxins (Basel) **2**(9): 2289-2332.

Dolgin, E. S., M. A. Felix and A. D. Cutter (2008). "Hakuna Nematoda: genetic and phenotypic diversity in African isolates of Caenorhabditis elegans and C. briggsae." Heredity (Edinb) **100**(3): 304-315.

Donya, M., M. Radford, A. ElGuindy, D. Firmin and M. H. Yacoub (2014). "Radiation in medicine: Origins, risks and aspirations." Glob Cardiol Sci Pract **2014**(4): 437-448.

Douki, T., D. Perdiz, P. Grof, Z. Kuluncsics, E. Moustacchi, J. Cadet and E. Sage (1999). "Oxidation of guanine in cellular DNA by solar UV radiation: biological role." Photochem Photobiol **70**(2): 184-190.

Downs, J. A. and S. P. Jackson (2004). "A means to a DNA end: the many roles of Ku." Nat Rev Mol Cell Biol **5**(5): 367-378.

Duerr, J. S. (2006). "Immunohistochemistry." WormBook: 1-61.

Dunn, O. J. (1964). "Multiple Comparisons Using Rank Sums." Technometrics **6**(3): 241-&.

El-Khamisy, S. F., E. Hartsuiker and K. W. Caldecott (2007). "TDP1 facilitates repair of ionizing radiation-induced DNA single-strand breaks." DNA Repair (Amst) **6**(10): 1485-1495.

Emerson, C. H. and A. A. Bertuch (2016). "Consider the workhorse: Nonhomologous end-joining in budding yeast." Biochem Cell Biol **94**(5): 396-406.

Ermolaeva, M. A., A. Segref, A. Dakhovnik, H. L. Ou, J. I. Schneider, O. Utermohlen, T. Hoppe and B. Schumacher (2013). "DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance." Nature **501**(7467): 416-420.

Ewa, B. and M. S. Danuta (2017). "Polycyclic aromatic hydrocarbons and PAH-related DNA adducts." J Appl Genet **58**(3): 321-330.

Ewald, C. Y., J. I. Castillo-Quan and T. K. Blackwell (2018). "Untangling Longevity, Dauer, and Healthspan in Caenorhabditis elegans Insulin/IGF-1-Signalling." Gerontology **64**(1): 96-104.

Fagbemi, A. F., B. Orelli and O. D. Scharer (2011). "Regulation of endonuclease activity in human nucleotide excision repair." DNA Repair (Amst) **10**(7): 722-729.

Falck, J., J. Coates and S. P. Jackson (2005). "Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage." Nature **434**(7033): 605-611.

Fay, D. (2006). "Genetic mapping and manipulation: chapter 1--Introduction and basics." WormBook: 1-12.

Felix, M. A. and C. Braendle (2010). "The natural history of Caenorhabditis elegans." Curr Biol **20**(22): R965-969.

Fire, A. (1986). "Integrative transformation of Caenorhabditis elegans." EMBO J **5**(10): 2673-2680.

Fire, A. (1994). "A four-dimensional digital image archiving system for cell lineage tracing and retrospective embryology." Comput Appl Biosci **10**(4): 443-447.

Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver and C. C. Mello (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans." Nature **391**(6669): 806-811.

Flibotte, S., M. L. Edgley, I. Chaudhry, J. Taylor, S. E. Neil, A. Rogula, R. Zapf, M. Hirst, Y. Butterfield, S. J. Jones, M. A. Marra, R. J. Barstead and D. G. Moerman (2010). "Whole-genome profiling of mutagenesis in Caenorhabditis elegans." Genetics **185**(2): 431-441.

Fousteri, M., W. Vermeulen, A. A. van Zeeland and L. H. Mullenders (2006). "Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo." Mol Cell **23**(4): 471-482.

Francis, D. B., M. Kozlov, J. Chavez, J. Chu, S. Malu, M. Hanna and P. Cortes (2014). "DNA Ligase IV regulates XRCC4 nuclear localization." DNA Repair (Amst) **21**: 36-42.

Friedberg, E. C. (2005). "Suffering in silence: the tolerance of DNA damage." Nat Rev Mol Cell Biol **6**(12): 943-953.

Frit, P., N. Barboule, Y. Yuan, D. Gomez and P. Calsou (2014). "Alternative end-joining pathway(s): bricolage at DNA breaks." DNA Repair (Amst) **17**: 81-97.

Fu, D., J. A. Calvo and L. D. Samson (2012). "Balancing repair and tolerance of DNA damage caused by alkylating agents." Nat Rev Cancer **12**(2): 104-120.

Fukagawa, T. and W. C. Earnshaw (2014). "The centromere: chromatin foundation for the kinetochore machinery." Dev Cell **30**(5): 496-508.

Gartner, A., P. R. Boag and T. K. Blackwell (2008). "Germline survival and apoptosis." WormBook: 1-20.

Gartner, A., S. Milstein, S. Ahmed, J. Hodgkin and M. O. Hengartner (2000). "A conserved checkpoint pathway mediates DNA damage--induced apoptosis and cell cycle arrest in C. elegans." Mol Cell **5**(3): 435-443.

Gell, D. and S. P. Jackson (1999). "Mapping of protein-protein interactions within the DNA-dependent protein kinase complex." Nucleic Acids Res **27**(17): 3494-3502.

Gent, J. I., A. T. Lamm, D. M. Pavelec, J. M. Maniar, P. Parameswaran, L. Tao, S. Kennedy and A. Z. Fire (2010). "Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma." Mol Cell **37**(5): 679-689.

Georgakilas, A. G., P. O'Neill and R. D. Stewart (2013). "Induction and repair of clustered DNA lesions: what do we know so far?" Radiat Res **180**(1): 100-109.

Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M. S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dose, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A. Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M. Jensen, M. Kato, W. J. Kent, E. Kephart, V. Khivansara, E. Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecenas, G. Merrihew, D. M. Miller, 3rd, A. Muroyama, J. I. Murray, S. L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Ratsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, C. Slightam, R. Smith, W. C. Spencer, E. O. Stinson, S. Taing, T. Takasaki, D. Vafeados, K. Voronina, G. Wang, N. L. Washington, C. M. Whittle, B. Wu, K. K. Yan, G. Zeller, Z. Zha, M. Zhong, X. Zhou, E. C. mod, J. Ahringer, S. Strome, K. C. Gunsalus, G. Micklem, X. S. Liu, V. Reinke, S. K. Kim, L. W. Hillier, S. Henikoff, F. Piano, M. Snyder, L. Stein, J. D. Lieb and R. H. Waterston (2010). "Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project." Science **330**(6012): 1775-1787.

Gilchrist, E. J., N. J. O'Neil, A. M. Rose, M. C. Zetka and G. W. Haughn (2006). "TILLING is an effective reverse genetics technique for Caenorhabditis elegans." BMC Genomics **7**: 262.

Gimenez, E. and F. Manzano-Agugliaro (2017). "DNA Damage Repair System in Plants: A Worldwide Research Update." Genes (Basel) **8**(11).

Girard, P. M., B. Kysela, C. J. Harer, A. J. Doherty and P. A. Jeggo (2004). "Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms." Hum Mol Genet **13**(20): 2369-2376.

Gloria-Soria, A. and R. B. Azevedo (2008). "npr-1 Regulates foraging and dispersal strategies in Caenorhabditis elegans." Curr Biol **18**(21): 1694-1699.

Goodman, M. F. and R. Woodgate (2013). "Translesion DNA polymerases." Cold Spring Harb Perspect Biol **5**(10): a010363.

Goodyer, W., S. Kaitna, F. Couteau, J. D. Ward, S. J. Boulton and M. Zetka (2008). "HTP-3 links DSB formation with homolog pairing and crossing over during C. elegans meiosis." Dev Cell **14**(2): 263-274.

Gorski, M. M., J. C. Eeken, A. W. de Jong, I. Klink, M. Loos, R. J. Romeijn, B. L. van Veen, L. H. Mullenders, W. Ferro and A. Pastink (2003). "The Drosophila melanogaster DNA Ligase IV gene plays a crucial role in the repair of radiation-induced DNA double-strand breaks and acts synergistically with Rad54." Genetics **165**(4): 1929-1941.

Gottlieb, T. M. and S. P. Jackson (1993). "The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen." Cell **72**(1): 131-142.

Grabarz, A., A. Barascu, J. Guirouilh-Barbat and B. S. Lopez (2012). "Initiation of DNA double strand break repair: signaling and single-stranded resection dictate the choice between homologous recombination, non-homologous end-joining and alternative end-joining." Am J Cancer Res **2**(3): 249-268.

Grawunder, U., M. Wilm, X. Wu, P. Kulesza, T. E. Wilson, M. Mann and M. R. Lieber (1997). "Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells." Nature **388**(6641): 492-495.

Gregory, C. D. and A. E. Milner (1994). "Regulation of cell survival in Burkitt lymphoma: implications from studies of apoptosis following cold-shock treatment." Int J Cancer **57**(3): 419-426.

Grundy, G. J., S. L. Rulten, Z. Zeng, R. Arribas-Bosacoma, N. Iles, K. Manley, A. Oliver and K. W. Caldecott (2013). "APLF promotes the assembly and activity of non-homologous end joining protein complexes." EMBO J **32**(1): 112-125.

Gu, J. and M. R. Lieber (2008). "Mechanistic flexibility as a conserved theme across 3 billion years of nonhomologous DNA end-joining." Genes Dev **22**(4): 411-415.

Gu, J. Y., Q. Wang, M. Cui, B. Han, H. J. Guo, L. S. Zhao, Y. D. Xie, X. Y. Song and L. X. Liu (2014). "Cloning and characterization of Ku70 and Ku80 homologues involved in DNA repair process in wheat (Triticum aestivum L.)." Plant Genetic Resources-Characterization and Utilization **12**: S99-S103.

Gu, S. G., J. Pak, S. Guang, J. M. Maniar, S. Kennedy and A. Fire (2012). "Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint." Nat Genet **44**(2): 157-164.

Gupta, S. and K. Meek (2005). "The leucine rich region of DNA-PKcs contributes to its innate DNA affinity." Nucleic Acids Research **33**(22): 6972-6981.

Haber, J. E. (2012). "Mating-type genes and MAT switching in Saccharomyces cerevisiae." Genetics **191**(1): 33-64.

Halanych, K. M. (1995). "Evidence from 18s Ribosomal DNA That the Lophophorates Are Protostome Animals (Vol 267, Page 1641, 1995)." Science **268**(5210): 485-485.

Hall, S. E., M. Beverly, C. Russ, C. Nusbaum and P. Sengupta (2010). "A cellular memory of developmental history generates phenotypic diversity in C. elegans." Curr Biol **20**(2): 149-155.

Hammel, M., M. Rey, Y. Yu, R. S. Mani, S. Classen, M. Liu, M. E. Pique, S. Fang, B. L. Mahaney, M. Weinfeld, D. C. Schriemer, S. P. Lees-Miller and J. A. Tainer (2011). "XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair." J Biol Chem **286**(37): 32638-32650.

Hammons, G. J., D. Milton, K. Stepps, F. P. Guengerich, R. H. Tukey and F. F. Kadlubar (1997). "Metabolism of carcinogenic heterocyclic and aromatic amines by recombinant human cytochrome P450 enzymes." Carcinogenesis **18**(4): 851-854.

Hamp, T., R. Kassner, S. Seemayer, E. Vicedo, C. Schaefer, D. Achten, F. Auer, A. Boehm, T. Braun, M. Hecht, M. Heron, P. Honigschmid, T. A. Hopf, S. Kaufmann, M. Kiening, D. Krompass, C. Landerer, Y. Mahlich, M. Roos and B. Rost (2013). "Homology-based inference sets the bar high for protein function prediction." BMC Bioinformatics **14 Suppl 3**: S7.

Harris, T. W., I. Antoshechkin, T. Bieri, D. Blasiar, J. Chan, W. J. Chen, N. De La Cruz, P. Davis, M. Duesbury, R. H. Fang, J. Fernandes, M. Han, R. Kishore, R. Lee, H. M. Muller, C. Nakamura, P. Ozersky, A. Petcherski, A. Rangarajan, A. Rogers, G. Schindelman, E. M. Schwarz, M. A. Tuli, K. Van Auken, D. Wang, X. D. Wang, G. Williams, K. Yook, R. Durbin, L. D. Stein, J. Spieth and P. W. Sternberg (2010). "WormBase: a comprehensive resource for nematode research." Nucleic Acids Research **38**: D463-D467.

Hartlerode, A. J., Y. Guan, A. Rajendran, K. Ura, G. Schotta, A. Xie, J. V. Shah and R. Scully (2012). "Impact of histone H4 lysine 20 methylation on 53BP1 responses to chromosomal double strand breaks." PLoS One **7**(11): e49211.

Hartman, H. and T. F. Smith (2014). "The evolution of the ribosome and the genetic code." Life (Basel) **4**(2): 227-249.

Hayashi, M., G. M. Chin and A. M. Villeneuve (2007). "C. elegans germ cells switch between distinct modes of double-strand break repair during meiotic prophase progression." Plos Genetics **3**(11): 2068-2084.

Hedgecock, E. M. and J. G. White (1985). "Polyploid tissues in the nematode Caenorhabditis elegans." Dev Biol **107**(1): 128-133.

Heeres, J. T. and P. J. Hergenrother (2007). "Poly(ADP-ribose) makes a date with death." Curr Opin Chem Biol **11**(6): 644-653.

Heflich, R. H. and R. E. Neft (1994). "Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites." Mutat Res **318**(2): 73-114.

Henle, E. S. and S. Linn (1997). "Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide." J Biol Chem **272**(31): 19095-19098.

Hodgkin, J. and T. M. Barnes (1991). "More is not better: brood size and population growth in a self-fertilizing nematode." Proc Biol Sci **246**(1315): 19-24.

Hodgkin, J. and T. Doniach (1997). "Natural variation and copulatory plug formation in Caenorhabditis elegans." Genetics **146**(1): 149-164.

Holliday, R. (2007). "A mechanism for gene conversion in fungi (Reprinted)." Genetics Research **89**(5-6): 285-307.

Hong, R. L. and R. J. Sommer (2006). "Pristionchus pacificus: a well-rounded nematode." Bioessays **28**(6): 651-659.

Houtgraaf, J. H., J. Versmissen and W. J. van der Giessen (2006). "A concise review of DNA damage checkpoints and repair in mammalian cells." Cardiovascular Revascularization Medicine **7**(3): 165-172.

Hsu, P. D., E. S. Lander and F. Zhang (2014). "Development and applications of CRISPR-Cas9 for genome engineering." Cell **157**(6): 1262-1278.

Hu, P. J. (2007). "Dauer." WormBook: 1-19.

Huettel, R. N. and A. M. Golden (1991). "Cobb,Nathan,Augustus - the Father of Nematology in the United-States." Annual Review of Phytopathology **29**: 14-26.

Hunter, S. E., M. A. Gustafson, K. M. Margillo, S. A. Lee, I. T. Ryde and J. N. Meyer (2012). "In vivo repair of alkylating and oxidative DNA damage in the mitochondrial and nuclear genomes of wild-type and glycosylase-deficient Caenorhabditis elegans." DNA Repair (Amst) **11**(11): 857-863.

Hutchinson, F. (1985). "Chemical-Changes Induced in DNA by Ionizing-Radiation." Progress in Nucleic Acid Research and Molecular Biology **32**: 115-154.

Iles, N., S. Rulten, S. F. El-Khamisy and K. W. Caldecott (2007). "APLF (C2orf13) is a novel human protein involved in the cellular response to chromosomal DNA strand breaks." Molecular and Cellular Biology **27**(10): 3793-3803.

Iliakis, G. (1991). "The Role of DNA Double Strand Breaks in Ionizing Radiation-Induced Killing of Eukaryotic Cells." Bioessays **13**(12): 641-648.

Iliakis, G., T. Murmann and A. Soni (2015). "Alternative end-joining repair pathways are the ultimate backup for abrogated classical non-homologous end-joining and homologous recombination repair: Implications for the formation of chromosome translocations." Mutation Research-Genetic Toxicology and Environmental Mutagenesis **793**: 166-175.

International Commission on Radiation, U. and Measurements (2011). "Report 85: Fundamental quantities and units for ionizing radiation." J ICRU **11**(1): 1-31.

Jackman, S. D., B. P. Vandervalk, H. Mohamadi, J. Chu, S. Yeo, S. A. Hammond, G. Jahesh, H. Khan, L. Coombe, R. L. Warren and I. Birol (2017). "ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter." Genome Res **27**(5): 768-777.

Jacoby, D. B. and P. C. Wensink (1996). "DNA binding specificities of YPF1, a Drosophila homolog to the DNA binding subunit of human DNA-dependent protein kinase, Ku." Journal of Biological Chemistry **271**(28): 16827-16832.

Jha, V., C. Bian, G. Xing and H. Ling (2016). "Structure and mechanism of error-free replication past the major benzo[a]pyrene adduct by human DNA polymerase kappa." Nucleic Acids Res **44**(10): 4957-4967.

Jiang, H. and D. Wang (2018). "The Microbial Zoo in the C. elegans Intestine: Bacteria, Fungi and Viruses." Viruses **10**(2).

Jiricny, J. (2006). "The multifaceted mismatch-repair system." Nat Rev Mol Cell Biol **7**(5): 335-346.

Jiricny, J. (2013). "Postreplicative mismatch repair." Cold Spring Harb Perspect Biol **5**(4): a012633.

Johnston, W. L. and J. W. Dennis (2012). "The eggshell in the C. elegans oocyte-to-embryo transition." Genesis **50**(4): 333-349.

Jun, S., Y. S. Jung, H. N. Suh, W. Wang, M. J. Kim, Y. S. Oh, E. M. Lien, X. Shen, Y. Matsumoto, P. D. McCrea, L. Li, J. Chen and J. I. Park (2016). "LIG4 mediates Wnt signalling-induced radioresistance." Nat Commun **7**: 10994.

Kaczanowski, S., M. Sajid and S. E. Reece (2011). "Evolution of apoptosis-like programmed cell death in unicellular protozoan parasites." Parasit Vectors **4**: 44.

Kamath, R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin, M. Gotta, A. Kanapin, N. Le Bot, S. Moreno, M. Sohrmann, D. P. Welchman, P. Zipperlen and J. Ahringer (2003). "Systematic

functional analysis of the Caenorhabditis elegans genome using RNAi." Nature **421**(6920): 231-237.

Kanno, S., H. Kuzuoka, S. Sasao, Z. Hong, L. Lan, S. Nakajima and A. Yasui (2007). "A novel human AP endonuclease with conserved zinc-finger-like motifs involved in DNA strand break responses." EMBO J **26**(8): 2094-2103.

Kantidze, O. L., A. K. Velichko, A. V. Luzhin and S. V. Razin (2016). "Heat Stress-Induced DNA Damage." Acta Naturae **8**(2): 75-78.

Karam, P. A. and S. A. Leslie (1999). "Calculations of background beta-gamma radiation dose through geologic time." Health Phys **77**(6): 662-667.

Kelly, W. G., S. Xu, M. K. Montgomery and A. Fire (1997). "Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene." Genetics **146**(1): 227-238.

Kenney, E. and I. Eleftherianos (2016). "Entomopathogenic and plant pathogenic nematodes as opposing forces in agriculture." International Journal for Parasitology **46**(1): 13-19.

Kiefer, J. (2007). Effects of Ultraviolet Radiation on DNA. Chromosomal Alterations: Methods, Results and Importance in Human Health. G. Obe and Vijayalaxmi. Berlin, Heidelberg, Springer Berlin Heidelberg**:** 39-53.

Kieleczawa, J. (2006). "Fundamentals of sequencing of difficult templates--an overview." J Biomol Tech **17**(3): 207-217.

Kim, S., K. Scheffler, A. L. Halpern, M. A. Bekritsky, E. Noh, M. Kallberg, X. Chen, Y. Kim, D. Beyter, P. Krusche and C. T. Saunders (2018). "Strelka2: fast and accurate calling of germline and somatic variants." Nat Methods **15**(8): 591-594.

Kiontke, K. and D. H. Fitch (2005). "The phylogenetic relationships of Caenorhabditis and other rhabditids." WormBook: 1-11.

Kiontke, K. and W. Sudhaus (2006). "Ecology of Caenorhabditis species." WormBook: 1-14.

Kipreos, E. T. (2005). "C. elegans cell cycles: invariance and stem cell divisions." Nat Rev Mol Cell Biol **6**(10): 766-776.

Kirchgessner, C. U., C. K. Patil, J. W. Evans, C. A. Cuomo, L. M. Fried, T. Carter, M. A. Oettinger and J. M. Brown (1995). "DNA-dependent kinase (p350) as a candidate gene for the murine SCID defect." Science **267**(5201): 1178-1183.

Koch, C. A., R. Agyei, S. Galicia, P. Metalnikov, P. O'Donnell, A. Starostine, M. Weinfeld and D. Durocher (2004). "Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV." EMBO J **23**(19): 3874-3885.

Koike, M., T. Awaji, M. Kataoka, G. Tsujimoto, T. Kartasova, A. Koike and T. Shiomi (1999). "Differential subcellular localization of DNA-dependent protein kinase components Ku and DNA-PKcs during mitosis." Journal of Cell Science **112**(22): 4031-4039.

Koike, M., T. Shiomi and A. Koike (2001). "Dimerization and nuclear localization of ku proteins." J Biol Chem **276**(14): 11167-11173.

Koike, M., Y. Yutoku and A. Koike (2015). "Dynamic changes in subcellular localization of cattle XLF during cell cycle, and focus formation of cattle XLF at DNA damage sites immediately after irradiation." J Vet Med Sci **77**(9): 1109-1114.

Koole, W., R. van Schendel, A. E. Karambelas, J. T. van Heteren, K. L. Okihara and M. Tijsterman (2014). "A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites." Nat Commun **5**: 3216.

Kriek, E. (1992). "Fifty years of research on N-acetyl-2-aminofluorene, one of the most versatile compounds in experimental cancer research." J Cancer Res Clin Oncol **118**(7): 481-489.

Krokan, H. E. and M. Bjoras (2013). "Base excision repair." Cold Spring Harb Perspect Biol **5**(4): a012583.

Kruskal, W. H. and W. A. Wallis (1952). "Use of Ranks in One-Criterion Variance Analysis." Journal of the American Statistical Association **47**(260): 583-621.

Kumsta, C. and M. Hansen (2012). "C. elegans rrf-1 mutations maintain RNAi efficiency in the soma in addition to the germline." PLoS One **7**(5): e35428.

Kunkel, T. A. (2009). "Evolving views of DNA replication (in)fidelity." Cold Spring Harb Symp Quant Biol **74**: 91-101.

Kunkel, T. A. (2011). "Balancing eukaryotic replication asymmetry with replication fidelity." Curr Opin Chem Biol **15**(5): 620-626.

Kutscher, L. M. and S. Shaham (2014). "Forward and reverse mutagenesis in C. elegans." WormBook: 1-26.

Lambshead, P. J. D. and G. Boucher (2003). "Marine nematode deep-sea biodiversity - hyperdiverse or hype?" Journal of Biogeography **30**(4): 475-485.

Lans, H. and W. Vermeulen (2011). "Nucleotide Excision Repair in Caenorhabditis elegans." Mol Biol Int **2011**: 542795.

Lans, H. and W. Vermeulen (2015). "Tissue specific response to DNA damage: C. elegans as role model." DNA Repair (Amst) **32**: 141-148.

Lant, B. and W. B. Derry (2014). "Fluorescent visualization of germline apoptosis in living Caenorhabditis elegans." Cold Spring Harb Protoc **2014**(4): 420-427.

Lawrence, M., S. Daujat and R. Schneider (2016). "Lateral Thinking: How Histone Modifications Regulate Gene Expression." Trends in Genetics **32**(1): 42-56.

Leesmiller, S. P., Y. R. Chen and C. W. Anderson (1990). "Human-Cells Contain a DNA-Activated Protein-Kinase That Phosphorylates Simian Virus-40 T-Antigen, Mouse P53, and the Human Ku-Autoantigen." Molecular and Cellular Biology **10**(12): 6472-6481.

Leesmiller, S. P., R. Godbout, D. W. Chan, M. Weinfeld, R. S. Day, G. M. Barron and J. Allalunisturner (1995). "Absence of P350 Subunit of DNA-Activated Protein-Kinase from a Radiosensitive Human Cell-Line." Science **267**(5201): 1183-1185.

Lemmens, B. B. L. G., N. M. Johnson and M. Tijsterman (2013). "COM-1 Promotes Homologous Recombination during Caenorhabditis elegans Meiosis by Antagonizing Ku-Mediated Non-Homologous End Joining." Plos Genetics **9**(2).

Lemmens, B. B. L. G. and M. Tijsterman (2011). "DNA double-strand break repair in Caenorhabditis elegans." Chromosoma **120**(1): 1-21.

Lempiainen, H. and T. D. Halazonetis (2009). "Emerging common themes in regulation of PIKKs and PI3Ks." Embo Journal **28**(20): 3067-3073.

Letunic, I. and P. Bork (2018). "20 years of the SMART protein domain annotation resource." Nucleic Acids Res **46**(D1): D493-D496.

Levy, S. E. and R. M. Myers (2016). "Advancements in Next-Generation Sequencing." Annu Rev Genomics Hum Genet **17**: 95-115.

Lewis, J. A. and J. T. Fleming (1995). "Basic culture methods." Methods Cell Biol **48**: 3-29.

Li, G. M. (2014). "New insights and challenges in mismatch repair: Getting over the chromatin hurdle." DNA Repair **19**: 48-54.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Li, H., H. Vogel, V. B. Holcomb, Y. S. Gu and P. Hasty (2007). "Deletion of Ku70, Ku80, or both causes early aging without substantially increased cancer." Molecular and Cellular Biology **27**(23): 8205-8214.

Li, L. Y., D. Moshous, Y. G. Zhou, J. H. Wang, G. Xie, E. Salido, D. Hu, J. P. de Villartay and M. J. Cowan (2002). "A founder mutation in Artemis, an SNM1-like protein, causes SCID in Athabascan-speaking Native Americans." Journal of Immunology **168**(12): 6323-6329.

Li, S., S. Kanno, R. Watanabe, H. Ogiwara, T. Kohno, G. Watanabe, A. Yasui and M. R. Lieber (2011). "Polynucleotide kinase and aprataxin-like forkhead-associated protein (PALF) acts as both a single-stranded DNA endonuclease and a single-stranded DNA 3' exonuclease and can participate in DNA end joining in a biochemical system." J Biol Chem **286**(42): 36368-36377.

Li, X. R. and R. E. Moses (2003). "The beta-lactamase motif in Snm1 is required for repair of DNA double-strand breaks caused by interstrand crosslinks in S-cerevisiae." DNA Repair **2**(1): 121-129.

Lieber, M. R. (2008). "The mechanism of human nonhomologous DNA end joining." Journal of Biological Chemistry **283**(1): 1-5.

Lieber, M. R. (2010). "The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway." Annual Review of Biochemistry, Vol 79 **79**: 181-211.

Lieber, M. R. (2016). "Mechanisms of human lymphoid chromosomal translocations." Nature Reviews Cancer **16**(6): 387-398.

Lindahl, T. (1979). "DNA glycosylases, endonucleases for apurinic/apyrimidinic sites, and base excision-repair." Prog Nucleic Acid Res Mol Biol **22**: 135-192.

Lindahl, T. (1993). "Instability and Decay of the Primary Structure of DNA." Nature **362**(6422): 709-715.

Loeb, L. A. and R. J. Monnat (2008). "DNA polymerases and human disease." Nature Reviews Genetics **9**(8): 594-604.

Lorenz, R., S. H. Bernhart, C. Honer Zu Siederdissen, H. Tafer, C. Flamm, P. F. Stadler and I. L. Hofacker (2011). "ViennaRNA Package 2.0." Algorithms Mol Biol **6**: 26.

Lowden, M. R., B. Meier, T. W. S. Lee, J. Hall and S. Ahmed (2008). "End joining at Caenorhabditis elegans Telomeres." Genetics **180**(2): 741-754.

Lucas-Lledo, J. I. and M. Lynch (2009). "Evolution of Mutation Rates: Phylogenomic Analysis of the Photolyase/Cryptochrome Family." Molecular Biology and Evolution **26**(5): 1143-1153.

Lui, D. Y. and M. P. Colaiacovo (2013). "Meiotic development in Caenorhabditis elegans." Adv Exp Med Biol **757**: 133-170.

Luoto, K. R., R. Kumareswaran and R. G. Bristow (2013). "Tumor hypoxia as a driving force in genetic instability." Genome Integr **4**(1): 5.

Ma, Y., U. Pannicke, K. Schwarz and M. R. Lieber (2002). "Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination." Cell **108**(6): 781-794.

Ma, Y., K. Schwarz and M. R. Lieber (2005). "The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps." DNA Repair (Amst) **4**(7): 845-851.

Macaisne, N., Z. Kessler and J. L. Yanowitz (2018). "Meiotic Double-Strand Break Proteins Influence Repair Pathway Utilization." Genetics.

Mahaney, B. L., M. Hammel, K. Meek, J. A. Tainer and S. P. Lees-Miller (2013). "XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair." Biochem Cell Biol **91**(1): 31-41.

Mahowald, G. K., J. M. Baron and B. P. Sleckman (2008). "Collateral Damage from Antigen Receptor Gene Diversification." Cell **135**(6): 1009-1012.

Malu, S., V. Malshetty, D. Francis and P. Cortes (2012). "Role of non-homologous end joining in V(D)J recombination." Immunologic Research **54**(1-3): 233-246.

Mann, H. B. and D. R. Whitney (1947). "On a Test of Whether One of 2 Random Variables Is Stochastically Larger Than the Other." Annals of Mathematical Statistics **18**(1): 50-60.

Manova, V. and D. Gruszka (2015). "DNA damage and repair in plants - from models to crops." Frontiers in Plant Science **6**.

Marcello, M. R. and A. Singson (2010). "Fertilization and the oocyte-to-embryo transition in C. elegans." Bmb Reports **43**(6): 389-399.

Mari, P. O., B. I. Florea, S. P. Persengiev, N. S. Verkaik, H. T. Brueggenwirth, M. Modesti, G. Giglia-Mari, K. Bezstarosti, J. A. A. Demmers, T. M. Luider, A. B. Houtsmuller and D. C. van Gent (2006). "Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4." Proceedings of the National Academy of Sciences of the United States of America **103**(49): 18597-18602.

Marteijn, J. A., H. Lans, W. Vermeulen and J. H. Hoeijmakers (2014). "Understanding nucleotide excision repair and its roles in cancer and ageing." Nat Rev Mol Cell Biol **15**(7): 465-481.

Martinez-Perez, E. and A. M. Villeneuve (2005). "HTP-1-dependent constraints coordinate homolog pairing and synapsis and promote chiasma formation during C. elegans meiosis." Genes Dev **19**(22): 2727-2743.

Mateo, A. R., Z. Kessler, A. K. Jolliffe, O. McGovern, B. Yu, A. Nicolucci, J. L. Yanowitz and W. B. Derry (2016). "The p53-like Protein CEP-1 Is Required for Meiotic Fidelity in C. elegans." Curr Biol **26**(9): 1148-1158.

Mateos-Gomez, P. A., F. Gong, N. Nair, K. M. Miller, E. Lazzerini-Denchi and A. Sfeir (2015). "Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination." Nature **518**(7538): 254-257.

Mateos-Gomez, P. A., T. Kent, S. K. Deng, S. McDevitt, E. Kashkina, T. M. Hoang, R. T. Pomerantz and A. Sfeir (2017). "The helicase domain of Poltheta counteracts RPA to promote alt-NHEJ." Nat Struct Mol Biol **24**(12): 1116-1123.

Mathieu, A. L., E. Verronese, G. I. Rice, F. Fouyssac, Y. Bertrand, C. Picard, M. Chansel, J. E. Walter, L. D. Notarangelo, M. J. Butte, K. C. Nadeau, K. Csomos, D. J. Chen, K. Chen, A. Delgado, C. Rigal, C. Bardin, C. Schuetz, D. Moshous, H. Reumaux, F. Plenat, A. Phan, M. T. Zabot, B. Balme, S. Viel, J. Bienvenu, P. Cochat, M. van der Burg, C. Caux, E. H. Kemp, I. Rouvet, C. Malcus, J. F. Meritet, A. Lim, Y. J. Crow, N. Fabien, C. Menetrier-Caux, J. P. De Villartay, T. Walzer and A. Belot (2015). "PRKDC mutations associated with immunodeficiency, granuloma, and autoimmune regulator-dependent autoimmunity." J Allergy Clin Immunol **135**(6): 1578-1588 e1575.

Maupas, E. (1900). "Modes et formes de reproduction des Nematodes." Archives de Zoologie Experimentale **(3)**(viii): pp. 463-624.

Mayer, W. E., M. Herrmann and R. J. Sommer (2007). "Phylogeny of the nematode genus Pristionchus and implications for biodiversity, biogeography and the evolution of hermaphroditism." BMC Evol Biol **7**: 104.

Mazzarelli, P., P. Parrella, D. Seripa, E. Signori, G. Perrone, C. Rabitti, D. Borzomati, A. Gabbrielli, M. G. Matera, C. Gravina, M. Caricato, M. L. Poeta, M. Rinaldi, S. Valeri, R. Coppola and V. M. Fazio (2005). "DNA end binding activity and Ku70/80 heterodimer expression in human colorectal tumor." World J Gastroenterol **11**(42): 6694-6700.

McLellan, J. L., N. J. O'Neil, I. Barrett, E. Ferree, D. M. van Pel, K. Ushey, P. Sipahimalani, J. Bryan, A. M. Rose and P. Hieter (2012). "Synthetic lethality of cohesins with PARPs and replication fork mediators." PLoS Genet **8**(3): e1002574.

McVey, M., V. Y. Khodaverdian, D. Meyer, P. G. Cerqueira and W. D. Heyer (2016). "Eukaryotic DNA Polymerases in Homologous Recombination." Annu Rev Genet **50**: 393-421.

Meek, K., V. Dang and S. P. Lees-Miller (2008). "DNA-PK: the means to justify the ends?" Adv Immunol **99**: 33-58.

Mehta, A. and J. E. Haber (2014). "Sources of DNA double-strand breaks and models of recombinational DNA repair." Cold Spring Harb Perspect Biol **6**(9): a016428.

Mimori, T., M. Akizuki, H. Yamagata, S. Inada, S. Yoshida and M. Homma (1981). "Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap." J Clin Invest **68**(3): 611-620.

Mimori, T., J. A. Hardin and J. A. Steitz (1986). "Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders." J Biol Chem **261**(5): 2274-2278.

Mishina, Y., E. M. Duguid and C. He (2006). "Direct reversal of DNA alkylation damage." Chem Rev **106**(2): 215-232.

Mittelman, D. (2013). Stress-induced mutagenesis. New York, NY, Springer.

Mladenov, E. and G. Iliakis (2011). "Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways." Mutat Res **711**(1-2): 61-72.

Moreira, M. C., C. Barbot, N. Tachi, N. Kozuka, E. Uchida, T. Gibson, P. Mendonca, M. Costa, J. Barros, T. Yanagisawa, M. Watanabe, Y. Ikeda, M. Aoki, T. Nagata, P. Coutinho, J. Sequeiros and M. Koenig (2001). "The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin." Nat Genet **29**(2): 189-193.

Morrical, S. W. (2015). "DNA-pairing and annealing processes in homologous recombination and homology-directed repair." Cold Spring Harb Perspect Biol **7**(2): a016444.

Moshous, D., I. Callebaut, R. de Chasseval, B. Corneo, M. Cavazzana-Calvo, F. Le Deist, I. Tezcan, O. Sanal, Y. Bertrand, N. Philippe, A. Fischer and J. P. de Villartay (2001). "Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency." Cell **105**(2): 177-186.

Mu, H., K. Kropachev, L. Wang, L. Zhang, A. Kolbanovskiy, M. Kolbanovskiy, N. E. Geacintov and S. Broyde (2012). "Nucleotide excision repair of 2-acetylaminofluorene- and 2-aminofluorene-(C8)-guanine adducts: molecular dynamics simulations elucidate how lesion structure and base sequence context impact repair efficiencies." Nucleic Acids Res **40**(19): 9675-9690.

Muschiol, D., F. Schroeder and W. Traunspurger (2009). "Life cycle and population growth rate of Caenorhabditis elegans studied by a new method." BMC Ecol **9**: 14.

Nakamura, N., H. Morinaga, M. Kikuchi, S. Yonekura, N. Ishii, K. Yamamoto, S. Yonei and Q. M. Zhang (2008). "Cloning and characterization of uracil-DNA glycosylase and the biological consequences of the loss of its function in the nematode Caenorhabditis elegans." Mutagenesis **23**(5): 407-413.

Neal, J. A., V. Dang, P. Douglas, M. S. Wold, S. P. Lees-Miller and K. Meek (2011). "Inhibition of homologous recombination by DNA-dependent protein kinase requires kinase activity, is titratable, and is modulated by autophosphorylation." Mol Cell Biol **31**(8): 1719-1733.

Nedelcu, A. M., W. W. Driscoll, P. M. Durand, M. D. Herron and A. Rashidi (2011). "On the paradigm of altruistic suicide in the unicellular world." Evolution **65**(1): 3-20.

Neutelings, T., C. A. Lambert, B. V. Nusgens and A. C. Colige (2013). "Effects of mild cold shock (25 degrees C) followed by warming up at 37 degrees C on the cellular stress response." PLoS One **8**(7): e69687.

Nielsen, C. (2017). "Evolution of deuterostomy - and origin of the chordates." Biological Reviews **92**(1): 316-325.

Nigon, V. M. and M. A. Felix (2017). "History of research on C. elegans and other free-living nematodes as model organisms." WormBook **2017**: 1-84.

Nilsson, A., F. Sirzen, R. Lewensohn, N. Wang and S. Skog (1999). "Cell cycle-dependent regulation of the DNA-dependent protein kinase." Cell Prolif **32**(4): 239-248.

Nishizawa-Yokoi, A., S. Nonaka, H. Saika, Y. I. Kwon, K. Osakabe and S. Toki (2012). "Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice." New Phytol **196**(4): 1048-1059.

Niu, W., Z. J. Lu, M. Zhong, M. Sarov, J. I. Murray, C. M. Brdlik, J. Janette, C. Chen, P. Alves, E. Preston, C. Slightham, L. Jiang, A. A. Hyman, S. K. Kim, R. H. Waterston, M. Gerstein, M. Snyder and V. Reinke (2011). "Diverse transcription factor binding features revealed by genome-wide ChIP-seq in C. elegans." Genome Res **21**(2): 245-254.

O'Neil, N. and A. Rose (2006). "DNA repair." WormBook: 1-12.

O'Neil, N. J., J. S. Martin, J. L. Youds, J. D. Ward, M. I. R. Petalcorin, A. M. Rose and S. J. Boulton (2013). "Joint Molecule Resolution Requires the Redundant Activities of MUS-81 and XPF-1 during Caenorhabditis elegans Meiosis." Plos Genetics **9**(7).

O'Sullivan, R. J. and J. Karlseder (2010). "Telomeres: protecting chromosomes against genome instability." Nature Reviews Molecular Cell Biology **11**(3): 171-181.

Ochi, T., A. N. Blackford, J. Coates, S. Jhujh, S. Mehmood, N. Tamura, J. Travers, Q. Wu, V. M. Draviam, C. V. Robinson, T. L. Blundell and S. P. Jackson (2015). "DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair." Science **347**(6218): 185-188.

Odell, I. D., S. S. Wallace and D. S. Pederson (2013). "Rules of engagement for base excision repair in chromatin." J Cell Physiol **228**(2): 258-266.

Ofran, Y. and B. Rost (2007). "ISIS: interaction sites identified from sequence." Bioinformatics **23**(2): e13-16.

Ogi, T., S. Limsirichaikul, R. M. Overmeer, M. Volker, K. Takenaka, R. Cloney, Y. Nakazawa, A. Niimi, Y. Miki, N. G. Jaspers, L. H. Mullenders, S. Yamashita, M. I. Fousteri and A. R. Lehmann (2010). "Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells." Mol Cell **37**(5): 714-727.

Orren, D. K., A. Machwe, P. Karmakar, J. Piotrowski, M. P. Cooper and V. A. Bohr (2001). "A functional interaction of Ku with Werner exonuclease facilitates digestion of damaged DNA." Nucleic Acids Res **29**(9): 1926-1934.

Paix, A., A. Folkmann, D. Rasoloson and G. Seydoux (2015). "High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes." Genetics **201**(1): 47-+.

Palazzo, A. F. and T. R. Gregory (2014). "The case for junk DNA." PLoS Genet **10**(5): e1004351.

Palopoli, M. F., M. V. Rockman, A. TinMaung, C. Ramsay, S. Curwen, A. Aduna, J. Laurita and L. Kruglyak (2008). "Molecular basis of the copulatory plug polymorphism in Caenorhabditis elegans." Nature **454**(7207): 1019-U1066.

Park, S. H., C. Cheong, J. Idoyaga, J. Y. Kim, J. H. Choi, Y. Do, H. Lee, J. H. Jo, Y. S. Oh, W. Im, R. M. Steinman and C. G. Park (2008). "Generation and application of new rat monoclonal antibodies against synthetic FLAG and OLLAS tags for improved immunodetection." Journal of Immunological Methods **331**(1-2): 27-38.

Pawelczak, K. S. and J. J. Turchi (2010). "Purification and characterization of exonuclease-free Artemis: Implications for DNA-PK-dependent processing of DNA termini in NHEJ-catalyzed DSB repair." DNA Repair (Amst) **9**(6): 670-677.

Pazdernik, N. and T. Schedl (2013). "Introduction to germ cell development in Caenorhabditis elegans." Adv Exp Med Biol **757**: 1-16.

Pearson, K. (1900). "On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling." Philosophical Magazine **50**(302): 157-175.

Perry, J. and N. Kleckner (2003). "The ATRs, ATMs, and TORs are giant HEAT repeat proteins." Cell **112**(2): 151-155.

Philippe, H., N. Lartillot and H. Brinkmann (2005). "Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia." Molecular Biology and Evolution **22**(5): 1246-1253.

Pirooznia, M., F. S. Goes and P. P. Zandi (2015). "Whole-genome CNV analysis: advances in computational approaches." Front Genet **6**: 138.

Pitcher, R. S., N. C. Brissett and A. Doherty (2007). "Nonhomologous end joining in bacteria: A microbial perspective." Annual Review of Microbiology **61**: 259-282.

Pitcher, R. S., A. J. Green, A. Brzostek, M. Korycka-Machala, J. Dziadek and A. J. Doherty (2007). "NHEJ protects mycobacteria in stationary phase against the harmful effects of desiccation." DNA Repair **6**(9): 1271-1276.

Pommier, Y., J. A. Barcelo, V. A. Rao, O. Sordet, A. G. Jobson, L. Thibaut, Z. H. Miao, J. A. Seiler, H. Zhang, C. Marchand, K. Agama, J. L. Nitiss and C. Redon (2006). "Repair of topoisomerase I -

Mediated DNA damage." Progress in Nucleic Acid Research and Molecular Biology, Vol 81 **81**: 179-+.

Pontier, D. B. and M. Tijsterman (2009). "A Robust Network of Double-Strand Break Repair Pathways Governs Genome Integrity during C. elegans Development." Current Biology **19**(16): 1384-1388.

Porta-de-la-Riva, M., L. Fontrodona, A. Villanueva and J. Ceron (2012). "Basic Caenorhabditis elegans methods: synchronization and observation." J Vis Exp(64): e4019.

Portin, P. and A. Wilkins (2017). "The Evolving Definition of the Term "Gene"." Genetics **205**(4): 1353-1364.

Postow, L. (2011). "Destroying the ring: Freeing DNA from Ku with ubiquitin." FEBS Lett **585**(18): 2876-2882.

Potenski, C. J. and H. L. Klein (2014). "How the misincorporation of ribonucleotides into genomic DNA can be both harmful and helpful to cells." Nucleic Acids Research **42**(16): 10226-U10798.

Qiu, R. Y., M. Sakato, E. J. Sacho, H. Wilkins, X. D. Zhang, P. Modrichd, M. M. Hingorani, D. A. Erie and K. R. Weninger (2015). "MutL traps MutS at a DNA mismatch." Proceedings of the National Academy of Sciences of the United States of America **112**(35): 10914-10919.

Radhakrishnan, S. K., N. Jette and S. P. Lees-Miller (2014). "Non-homologous end joining: Emerging themes and unanswered questions." DNA Repair **17**: 2-8.

Ramsden, D. A. and K. Asagoshi (2012). "DNA polymerases in nonhomologous end joining: Are there any benefits to standing out from the crowd?" Environmental and Molecular Mutagenesis **53**(9): 741-751.

Rastogi, R. P., Richa, A. Kumar, M. B. Tyagi and R. P. Sinha (2010). "Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair." Journal of Nucleic Acids.

Rathmell, W. K. and G. Chu (1994). "A DNA End-Binding Factor Involved in Double-Strand Break Repair and V(D)J Recombination." Molecular and Cellular Biology **14**(7): 4741-4748.

Reynolds, P., J. A. Anderson, J. V. Harper, M. A. Hill, S. W. Botchway, A. W. Parker and P. O'Neill (2012). "The dynamics of Ku70/80 and DNA-PKcs at DSBs induced by ionizing radiation is dependent on the complexity of damage." Nucleic Acids Research **40**(21): 10821-10831.

Riddle, D. L., T. Blumenthal, B. J. Meyer and J. R. Priess (1997). Introduction to C. elegans. C. elegans II. nd, D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess. Cold Spring Harbor (NY).

Rivera-Calzada, A., L. Spagnolo, L. H. Pearl and O. Llorca (2007). "Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs." Embo Reports **8**(1): 56-62.

Robbiani, D. F., A. Bothmer, E. Callen, B. Reina-San-Martin, Y. Dorsett, S. Difilippantonio, D. J. Bolland, H. T. Chen, A. E. Corcoran, A. Nussenzweig and M. C. Nussenzweig (2008). "AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations." Cell **135**(6): 1028-1038.

Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz and J. P. Mesirov (2011). "Integrative genomics viewer." Nat Biotechnol **29**(1): 24-26.

Rogers, C., V. Reale, K. Kim, H. Chatwin, C. Li, P. Evans and M. de Bono (2003). "Inhibition of Caenorhabditis elegans social feeding by FMRFamide-related peptide activation of NPR-1." Nature Neuroscience **6**(11): 1178-1185.

Rose, A. (2014). "Replication and repair." WormBook: 1-16.

Rost, B., P. Fariselli and R. Casadio (1996). "Topology prediction for helical transmembrane proteins at 86% accuracy." Protein Sci **5**(8): 1704-1718.

Rost, B., G. Yachdav and J. Liu (2004). "The PredictProtein server." Nucleic Acids Res **32**(Web Server issue): W321-326.

Rougvie, A. E. and E. G. Moss (2013). "Developmental transitions in C. elegans larval stages." Curr Top Dev Biol **105**: 153-180.

Roy, S., A. J. de Melo, Y. Xu, S. K. Tadi, A. Negrel, E. Hendrickson, M. Modesti and K. Meek (2015). "XRCC4/XLF Interaction Is Variably Required for DNA Repair and Is Not Required for Ligase IV Stimulation." Molecular and Cellular Biology **35**(17): 3017-3028.

Rual, J. F., J. Ceron, J. Koreth, T. Hao, A. S. Nicot, T. Hirozane-Kishikawa, J. Vandenhaute, S. H. Orkin, D. E. Hill, S. van den Heuvel and M. Vidal (2004). "Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library." Genome Research **14**(10b): 2162-2168.

Sachadyn, P. (2010). "Conservation and diversity of MutS proteins." Mutat Res **694**(1-2): 20-30.

Saito, S., A. Kurosawa and N. Adachi (2016). "Mutations in XRCC4 cause primordial dwarfism without causing immunodeficiency." Journal of Human Genetics **61**(8): 679-685.

Saito, T. T., D. Y. Lui, H. M. Kim, K. Meyer and M. P. Colaiacovo (2013). "Interplay between Structure-Specific Endonucleases for Crossover Control during Caenorhabditis elegans Meiosis." Plos Genetics **9**(7).

Sakashita, T., T. Takanami, S. Yanase, N. Hamada, M. Suzuki, T. Kimura, Y. Kobayashi, N. Ishii and A. Higashitani (2010). "Radiation biology of Caenorhabditis elegans: germ cell response, aging and behavior." J Radiat Res **51**(2): 107-121.

Sallmyr, A. and A. E. Tomkinson (2018). "Repair of DNA double-strand breaks by mammalian alternative end-joining pathways." Journal of Biological Chemistry **293**(27): 10536-10546.

Sarov, M., J. I. Murray, K. Schanze, A. Pozniakovski, W. Niu, K. Angermann, S. Hasse, M. Rupprecht, E. Vinis, M. Tinney, E. Preston, A. Zinke, S. Enst, T. Teichgraber, J. Janette, K. Reis, S. Janosch, S. Schloissnig, R. K. Ejsmont, C. Slightam, X. Xu, S. K. Kim, V. Reinke, A. F. Stewart, M. Snyder, R. H. Waterston and A. A. Hyman (2012). "A genome-scale resource for in vivo tag-based protein function exploration in C. elegans." Cell **150**(4): 855-866.

Scharer, O. D. (2013). "Nucleotide Excision Repair in Eukaryotes." Cold Spring Harbor Perspectives in Biology **5**(10).

Schieber, M. and N. S. Chandel (2014). "ROS function in redox signaling and oxidative stress." Curr Biol **24**(10): R453-462.

Schlessinger, A., M. Punta, G. Yachdav, L. Kajan and B. Rost (2009). "Improved disorder prediction by combination of orthogonal approaches." PLoS One **4**(2): e4433.

Schnabel, R., H. Hutter, D. Moerman and H. Schnabel (1997). "Assessing normal embryogenesis in Caenorhabditis elegans using a 4D microscope: variability of development and regional specification." Dev Biol **184**(2): 234-265.

Schotta, G., R. Sengupta, S. Kubicek, S. Malin, M. Kauer, E. Callen, A. Celeste, M. Pagani, S. Opravil, I. A. De La Rosa-Velazquez, A. Espejo, M. T. Bedford, A. Nussenzweig, M. Busslinger and T. Jenuwein (2008). "A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse." Genes Dev **22**(15): 2048-2061.

Schwertman, P., A. Lagarou, D. H. Dekkers, A. Raams, A. C. van der Hoek, C. Laffeber, J. H. Hoeijmakers, J. A. Demmers, M. Fousteri, W. Vermeulen and J. A. Marteijn (2012). "UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair." Nat Genet **44**(5): 598-602.

Sekelsky, J. (2017). "DNA Repair in Drosophila: Mutagens, Models, and Missing Genes." Genetics **205**(2): 471-490.

Sendoel, A., I. Kohler, C. Fellmann, S. W. Lowe and M. O. Hengartner (2010). "HIF-1 antagonizes p53-mediated apoptosis through a secreted neuronal tyrosinase." Nature **465**(7298): 577-583.

Severson, A. F., L. Ling, V. van Zuylen and B. J. Meyer (2009). "The axial element protein HTP-3 promotes cohesin loading and meiotic axis assembly in C. elegans to implement the meiotic program of chromosome segregation." Genes Dev **23**(15): 1763-1778.

Shaffer, J. P. (1995). "Multiple Hypothesis-Testing." Annual Review of Psychology **46**: 561-584.

Shaltiel, I. A., L. Krenning, W. Bruinsma and R. H. Medema (2015). "The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle." Journal of Cell Science **128**(4): 607-620.

Shapiro, S. S. and M. B. Wilk (1965). "An Analysis of Variance Test for Normality (Complete Samples)." Biometrika **52**: 591-&.

Shatilla, A., A. Leduc, X. Yang and D. Ramotar (2005). "Identification of two apurinic/apyrimidinic endonucleases from Caenorhabditis elegans by cross-species complementation." DNA Repair (Amst) **4**(6): 655-670.

Shatilovich, A. V., A. V. Tchesunov, T. V. Neretina, I. P. Grabarnik, S. V. Gubin, T. A. Vishnivetskaya, T. C. Onstott and E. M. Rivkina (2018). "Viable Nematodes from Late Pleistocene Permafrost of the Kolyma River Lowland." Dokl Biol Sci **480**(1): 100-102.

Shibata, A., S. Conrad, J. Birraux, V. Geuting, O. Barton, A. Ismail, A. Kakarougkas, K. Meek, G. Taucher-Scholz, M. Lobrich and P. A. Jeggo (2011). "Factors determining DNA double-strand break repair pathway choice in G2 phase." EMBO J **30**(6): 1079-1092.

Shirodkar, P., A. L. Fenton, L. Meng and C. A. Koch (2013). "Identification and functional characterization of a Ku-binding motif in aprataxin polynucleotide kinase/phosphatase-like factor (APLF)." J Biol Chem **288**(27): 19604-19613.

Shuman, S. and M. S. Glickman (2007). "Bacterial DNA repair by non-homologous end joining." Nat Rev Microbiol **5**(11): 852-861.

Sibanda, B. L., D. Y. Chirgadze and T. L. Blundell (2010). "Crystal structure of DNA-PKcs reveals a large open-ring cradle comprised of HEAT repeats." Nature **463**(7277): 118-U132.

Sibanda, B. L., S. E. Critchlow, J. Begun, X. Y. Pei, S. P. Jackson, T. L. Blundell and L. Pellegrini (2001). "Crystal structure of an Xrcc4-DNA ligase IV complex." Nature Structural Biology **8**(12): 1015-1019.

Singleton, B. K., M. I. Torres-Arzayus, S. T. Rottinghaus, G. E. Taccioli and P. A. Jeggo (1999). "The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit." Molecular and Cellular Biology **19**(5): 3267-3277.

Sirbu, B. M. and D. Cortez (2013). "DNA Damage Response: Three Levels of DNA Repair Regulation." Cold Spring Harbor Perspectives in Biology **5**(8).

Smider, V., W. K. Rathmell, M. R. Lieber and G. Chu (1994). "Restoration of X-ray resistance and V(D)J recombination in mutant cells by Ku cDNA." Science **266**(5183): 288-291.

Smolikov, S., A. Eizinger, A. Hurlburt, E. Rogers, A. M. Villeneuve and M. P. Colaiacovo (2007). "Synapsis-Defective mutants reveal a correlation between chromosome conformation and the

mode of double-strand break repair during Caenorhabditis elegans meiosis." Genetics **176**(4): 2027-2033.

Spagnolo, L., A. Rivera-Calzada, L. H. Pearl and O. Llorca (2006). "Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair." Molecular Cell **22**(4): 511-519.

Steenkamp, E. T., J. Wright and S. L. Baldauf (2006). "The protistan origins of animals and fungi." Molecular Biology and Evolution **23**(1): 93-106.

Stiernagle, T. (2006). "Maintenance of C. elegans." WormBook: 1-11.

Stinchcomb, D. T., J. E. Shaw, S. H. Carr and D. Hirsh (1985). "Extrachromosomal DNA transformation of Caenorhabditis elegans." Mol Cell Biol **5**(12): 3484-3496.

Stolarek, M., D. Gruszka, A. Braszewska-Zalewska and M. Maluszynski (2015). "Functional analysis of the new barley gene HvKu80 indicates that it plays a key role in double-strand DNA break repair and telomere length regulation." Mutagenesis **30**(6): 785-797.

Sulston, J. E. and S. Brenner (1974). "The DNA of Caenorhabditis elegans." Genetics **77**(1): 95-104.

Sulston, J. E. and H. R. Horvitz (1977). "Post-embryonic cell lineages of the nematode, Caenorhabditis elegans." Dev Biol **56**(1): 110-156.

Sulston, J. E., E. Schierenberg, J. G. White and J. N. Thomson (1983). "The embryonic cell lineage of the nematode Caenorhabditis elegans." Dev Biol **100**(1): 64-119.

Szewczyk, N. J., R. L. Mancinelli, W. McLamb, D. Reed, B. S. Blumberg and C. A. Conley (2005). "Caenorhabditis elegans survives atmospheric breakup of STS-107, space shuttle Columbia." Astrobiology **5**(6): 690-705.

Taccioli, G. E., T. M. Gottlieb, T. Blunt, A. Priestley, J. Demengeot, R. Mizuta, A. R. Lehmann, F. W. Alt, S. P. Jackson and P. A. Jeggo (1994). "Ku80 - Product of the Xrcc5 Gene and Its Role in DNA-Repair and V(D)J Recombination." Science **265**(5177): 1442-1445.

Tadi, S. K., C. Tellier-Lebegue, C. Nemoz, P. Drevet, S. Audebert, S. Roy, K. Meek, J. B. Charbonnier and M. Modesti (2016). "PAXX Is an Accessory c-NHEJ Factor that Associates with Ku70 and Has Overlapping Functions with XLF." Cell Reports **17**(2): 541-555.

Tamura, K., Y. Adachi, K. Chiba, K. Oguchi and H. Takahashi (2002). "Identification of Ku70 and Ku80 homologues in Arabidopsis thaliana: evidence for a role in the repair of DNA double-strand breaks." Plant J **29**(6): 771-781.

Tang, J., N. W. Cho, G. Cui, E. M. Manion, N. M. Shanbhag, M. V. Botuyan, G. Mer and R. A. Greenberg (2013). "Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination." Nat Struct Mol Biol **20**(3): 317-325.

Tijsterman, M., J. Pothof and R. H. Plasterk (2002). "Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans." Genetics **161**(2): 651-660.

Timmons, L. and A. Fire (1998). "Specific interference by ingested dsRNA." Nature **395**(6705): 854-854.

Tsai, A. G., H. H. Lu, S. C. Raghavan, M. Muschen, C. L. Hsieh and M. R. Lieber (2008). "Human Chromosomal Translocations at CpG Sites and a Theoretical Basis for Their Lineage and Stage Specificity." Cell **135**(6): 1130-1142.

Tsai, C. J., S. A. Kim and G. Chu (2007). "Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends." Proc Natl Acad Sci U S A **104**(19): 7851-7856.

Turek, M., J. Besseling, J. P. Spies, S. Konig and H. Bringmann (2016). "Sleep-active neuron specification and sleep induction require FLP-11 neuropeptides to systemically induce sleep." Elife **5**.

Tursun, B., T. Patel, P. Kratsios and O. Hobert (2011). "Direct conversion of C. elegans germ cells into specific neuron types." Science **331**(6015): 304-308.

van der Burg, M., H. Ijspeert, N. S. Verkaik, T. Turul, W. W. Wiegant, K. Morotomi-Yano, P. O. Mari, I. Tezcan, D. J. Chen, M. Z. Zdzienicka, J. J. van Dongen and D. C. van Gent (2009). "A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining." J Clin Invest **119**(1): 91-98.

van Schendel, R., J. van Heteren, R. Welten and M. Tijsterman (2016). "Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End-Joining." PLoS Genet **12**(10): e1006368.

Vasale, J. J., W. Gu, C. Thivierge, P. J. Batista, J. M. Claycomb, E. M. Youngman, T. F. Duchaine, C. C. Mello and D. Conte, Jr. (2010). "Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway." Proc Natl Acad Sci U S A **107**(8): 3582-3587.

Vertessy, B. G. and J. Toth (2009). "Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases." Acc Chem Res **42**(1): 97-106.

Vilenchik, M. M. and A. G. Knudson (2003). "Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer." Proc Natl Acad Sci U S A **100**(22): 12871-12876.

Volk, T., U. Pannicke, I. Reisli, A. Bulashevska, J. Ritter, A. Bjorkman, A. A. Schaffer, M. Fliegauf, E. H. Sayar, U. Salzer, P. Fisch, D. Pfeifer, M. Di Virgilio, H. Cao, F. Yang, K. Zimmermann, S. Keles, Z. Caliskaner, S. U. Guner, D. Schindler, L. Hammarstrom, M. Rizzi, M. Hummel, Q. Pan-Hammarstrom, K. Schwarz and B. Grimbacher (2015). "DCLRE1C (ARTEMIS) mutations causing phenotypes ranging from atypical severe combined immunodeficiency to mere antibody deficiency." Hum Mol Genet **24**(25): 7361-7372.

Wagner, E. J. and P. B. Carpenter (2012). "Understanding the language of Lys36 methylation at histone H3." Nat Rev Mol Cell Biol **13**(2): 115-126.

Walker, J. R., R. A. Corpina and J. Goldberg (2001). "Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair." Nature **412**(6847): 607-614.

Walsh, D. A. and W. F. Doolittle (2005). "The real 'domains' of life." Curr Biol **15**(7): R237-240.

Wang, C. and S. P. Lees-Miller (2013). "Detection and repair of ionizing radiation-induced DNA double strand breaks: new developments in nonhomologous end joining." Int J Radiat Oncol Biol Phys **86**(3): 440-449.

Wang, J. C. (2002). "Cellular roles of DNA topoisomerases: a molecular perspective." Nat Rev Mol Cell Biol **3**(6): 430-440.

Ward, J. D. (2015). "Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair." Genetics **199**(2): 363-377.

Wardman, P. (2009). "The importance of radiation chemistry to radiation and free radical biology (The 2008 Silvanus Thompson Memorial Lecture)." Br J Radiol **82**(974): 89-104.

Weaver, D. T. (1995). "What to Do at an End - DNA Double-Strand-Break Repair." Trends in Genetics **11**(10): 388-392.

Weber, S. (2005). "Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase." Biochimica Et Biophysica Acta-Bioenergetics **1707**(1): 1-23.

Wei, H. and X. Yu (2016). "Functions of PARylation in DNA Damage Repair Pathways." Genomics Proteomics Bioinformatics **14**(3): 131-139.

Weidhaas, J. B., D. M. Eisenmann, J. M. Holub and S. V. Nallur (2006). "A Caenorhabditis elegans tissue model of radiation-induced reproductive cell death." Proc Natl Acad Sci U S A **103**(26): 9946-9951.

Weinfeld, M., R. S. Mani, I. Abdou, R. D. Aceytuno and J. N. Glover (2011). "Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair." Trends Biochem Sci **36**(5): 262-271.

Welker, N. C., D. M. Pavelec, D. A. Nix, T. F. Duchaine, S. Kennedy and B. L. Bass (2010). "Dicer's helicase domain is required for accumulation of some, but not all, C-elegans endogenous siRNAs." Rna-a Publication of the Rna Society **16**(5): 893-903.

Weller, G. R. and A. J. Doherty (2001). "A family of DNA repair ligases in bacteria?" FEBS Lett **505**(2): 340-342.

Weller, G. R., B. Kysela, R. Roy, L. M. Tonkin, E. Scanlan, M. Della, S. K. Devine, J. P. Day, A. Wilkinson, F. d'Adda di Fagagna, K. M. Devine, R. P. Bowater, P. A. Jeggo, S. P. Jackson and A. J. Doherty (2002). "Identification of a DNA nonhomologous end-joining complex in bacteria." Science **297**(5587): 1686-1689.

West, C. E., W. M. Waterworth, Q. Jiang and C. M. Bray (2000). "Arabidopsis DNA ligase IV is induced by gamma-irradiation and interacts with an Arabidopsis homologue of the double strand break repair protein XRCC4." Plant J **24**(1): 67-78.

West, C. E., W. M. Waterworth, G. W. Story, P. A. Sunderland, Q. Jiang and C. M. Bray (2002). "Disruption of the Arabidopsis AtKu80 gene demonstrates an essential role for AtKu80 protein in efficient repair of DNA double-strand breaks in vivo." Plant J **31**(4): 517-528.

Weterings, E., N. S. Verkaik, G. Keijzers, B. I. Florea, S. Y. Wang, L. G. Ortega, N. Uematsu, D. J. Chen and D. C. van Gent (2009). "The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends." Mol Cell Biol **29**(5): 1134-1142.

Whitehouse, C. J., R. M. Taylor, A. Thistlethwaite, H. Zhang, F. Karimi-Busheri, D. D. Lasko, M. Weinfeld and K. W. Caldecott (2001). "XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair." Cell **104**(1): 107-117.

Wiesenfahrt, T., J. Y. Berg, E. Osborne Nishimura, A. G. Robinson, B. Goszczynski, J. D. Lieb and J. D. McGhee (2016). "The function and regulation of the GATA factor ELT-2 in the C. elegans endoderm." Development **143**(3): 483-491.

Williams, D. R., K. J. Lee, J. Shi, D. J. Chen and P. L. Stewart (2008). "Cryo-EM structure of the DNA-dependent protein kinase catalytic subunit at subnanometer resolution reveals alpha helices and insight into DNA binding." Structure **16**(3): 468-477.

Wilson, J. H., P. B. Berget and J. M. Pipas (1982). "Somatic cells efficiently join unrelated DNA segments end-to-end." Mol Cell Biol **2**(10): 1258-1269.

Wilson, T. E., U. Grawunder and M. R. Lieber (1997). "Yeast DNA ligase IV mediates non-homologous DNA end joining." Nature **388**(6641): 495-498.

Wilson, T. E. and M. R. Lieber (1999). "Efficient processing of DNA ends during yeast nonhomologous end joining - Evidence for a DNA polymerase beta (POL4)-dependent pathway." Journal of Biological Chemistry **274**(33): 23599-23609.

Winocour, E. and I. Keshet (1980). "Indiscriminate recombination in simian virus 40-infected monkey cells." Proc Natl Acad Sci U S A **77**(8): 4861-4865.

Wolters, S. and B. Schumacher (2013). "Genome maintenance and transcription integrity in aging and disease." Front Genet **4**: 19.

Woodbine, L., A. R. Gennery and P. A. Jeggo (2014). "The clinical impact of deficiency in DNA non-homologous end-joining." DNA Repair (Amst) **16**: 84-96.

Wu, P. Y., P. Frit, S. Meesala, S. Dauvillier, M. Modesti, S. N. Andres, Y. Huang, J. Sekiguchi, P. Calsou, B. Salles and M. S. Junop (2009). "Structural and Functional Interaction between the Human DNA Repair Proteins DNA Ligase IV and XRCC4." Molecular and Cellular Biology **29**(11): 3163-3172.

Wu, Q., T. Ochi, D. Matak-Vinkovic, C. V. Robinson, D. Y. Chirgadze and T. L. Blundell (2011). "Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions." Biochem Soc Trans **39**(5): 1387-1392, suppl 1382 p following 1392.

Wu, X., T. E. Wilson and M. R. Lieber (1999). "A role for FEN-1 in nonhomologous DNA end joining: the order of strand annealing and nucleolytic processing events." Proc Natl Acad Sci U S A **96**(4): 1303-1308.

Wyatt, M. D. and D. L. Pittman (2006). "Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks." Chemical Research in Toxicology **19**(12): 1580-1594.

Xing, M., M. Yang, W. Huo, F. Feng, L. Wei, W. Jiang, S. Ning, Z. Yan, W. Li, Q. Wang, M. Hou, C. Dong, R. Guo, G. Gao, J. Ji, S. Zha, L. Lan, H. Liang and D. Xu (2015). "Interactome analysis identifies a new paralogue of XRCC4 in non-homologous end joining DNA repair pathway." Nat Commun **6**: 6233.

Xu, C. (2018). "A review of somatic single nucleotide variant calling algorithms for next-generation sequencing data." Comput Struct Biotechnol J **16**: 15-24.

Yang, H., Y. Matsumoto, K. M. Trujillo, S. P. Lees-Miller, M. A. Osley and A. E. Tomkinson (2015). "Role of the yeast DNA repair protein Nej1 in end processing during the repair of DNA double strand breaks by non-homologous end joining." DNA Repair **31**: 1-10.

Yano, K., K. Morotomi-Yano, K. J. Lee and D. J. Chen (2011). "Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF." Febs Letters **585**(6): 841-846.

Yano, K. I., K. Morotomi-Yano, S. Y. Wang, N. Uematsu, K. J. Lee, A. Asaithamby, E. Weterings and D. J. Chen (2008). "Ku recruits XLF to DNA double-strand breaks." Embo Reports **9**(1): 91-96.

Yigit, E., P. J. Batista, Y. Bei, K. M. Pang, C. C. Chen, N. H. Tolia, L. Joshua-Tor, S. Mitani, M. J. Simard and C. C. Mello (2006). "Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi." Cell **127**(4): 747-757.

Yin, Y. Z. and S. Smolikove (2013). "Impaired Resection of Meiotic Double-Strand Breaks Channels Repair to Nonhomologous End Joining in Caenorhabditis elegans." Molecular and Cellular Biology **33**(14): 2732-2747.

Yu, J. H., K. Marshall, M. Yamaguchi, J. E. Haber and C. F. Weil (2004). "Microhomology-dependent end joining and repair of transposon-induced DNA hairpins by host factors in Saccharomyces cerevisiae." Molecular and Cellular Biology **24**(3): 1351-1364.

Yurchenko, V., Z. Xue and M. J. Sadofsky (2006). "SUMO modification of human XRCC4 regulates its localization and function in DNA double-strand break repair." Mol Cell Biol **26**(5): 1786-1794.

Zhang, R., J. Erler and J. Langowski (2017). "Histone Acetylation Regulates Chromatin Accessibility: Role of H4K16 in Inter-nucleosome Interaction." Biophys J **112**(3): 450-459.

Zhang, Z. M., W. D. Hu, L. Cano, T. D. Lee, D. J. Chen and Y. Chen (2004). "Solution structure of the C-terminal domain of Ku80 suggests important sites for protein-protein interactions." Structure **12**(3): 495-502.

Zhang, Z. M., L. Y. Zhu, D. H. Lin, F. Q. Chen, D. J. Chen and Y. Chen (2001). "The three-dimensional structure of the C-terminal DNA-binding domain of human Ku70." Journal of Biological Chemistry **276**(41): 38231-38236.

Zhou, T., J. W. Lee, H. Tatavarthi, J. R. Lupski, K. Valerie and L. F. Povirk (2005). "Deficiency in 3 '-phosphoglycolate processing in human cells with a hereditary mutation in tyrosyl-DNA phosphodiesterase (TDP1)." Nucleic Acids Research **33**(1): 289-297.

Zickler, D. and N. Kleckner (2015). "Recombination, Pairing, and Synapsis of Homologs during Meiosis." Cold Spring Harb Perspect Biol **7**(6).