Spring 5-2021

Mapping and Characterization of Mutagen Sensitivity Genes in Drosophila Malanogaster
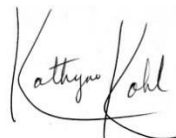
Alexis Nystrom
Winthrop University, alexis.nystrom44@gmail.com

Recommended Citation
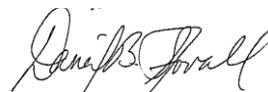
May, 2021

To the Dean of the Graduate School:
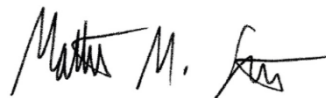
We are submitting a thesis written by Alexis Nystrom entitled Mapping and

Characterization of Mutagen Sensitivity Genes in *Drosophila melanogaster*. We

recommend acceptance in partial fulfillment of the requirements for the degree of Master

of Science in Biology.

_____   _____

Kathryn P. Kohl, Thesis Adviser

_____   _____

Daniel B. Stovall, Committee Member

_____   __

Matthew M. Stern, Committee Member

_____

Takita Sumter, Dean, College of Arts & Sciences

_____

Jack E. DeRochi, Dean, Graduate School

MAPPING AND CHARACTERIZATION OF MUTAGEN SENSITIVITY GENES IN

DROSOPHILA MELANOGASTER

A Thesis

Presented to the Faculty

Of the

College of Arts and Science

In Partial Fulfillment

Of the

Requirements for the Degree

Of

Master of Science

In Biology

Winthrop University

May, 2021

By

Alexis Nystrom

Abstract

The information contained within DNA is vital to directing all biological processes. All organisms have repair mechanisms in place to repair DNA damage quickly and efficiently. Without these repair pathways, DNA can acquire harmful mutations that can compromise the survival of an organism. Studies of DNA repair in *Drosophila melanogaster* have focused on mutagen sensitive (*mus*) mutants, each of which contain a mutation that renders them incapable of performing DNA repair. Since a majority of these *mus* genes are unmapped, the goal of this project was to determine what genes in the *Drosophila melanogaster* genome are *mus106* and *mus108*. Presence of each mutation was confirmed by conducting mutagen sensitivity assays on homozygous crosses. It was determined that the *mus106* mutation is no longer present in its corresponding stock, but that the *mus108* mutation is still present in its stock. After analyzing publically available genome data, we suggest potential candidate genes for *mus106* and *mus108* to be *DNA Ligase 4* and *XRCC1*, respectively. Since *XRCC1* has not been previously studied in *Drosophila melanogaster,* there are no known alleles of this gene. However, we conducted additional mutagen sensitivity tests on a transposon and an RNAi stock designed to target reduction of *XRCC1*. The results from these experiments are inconclusive until *XRCC1* knockdown is confirmed. Further characterization of *mus108* to other mutagens is in progress to better understand what DNA repair pathway MUS108 is involved in. This work can help researchers learn more about DNA repair pathways and fill in gaps of knowledge on gene function in *D. melanogaster*.

Table of Contents

List of Tables

List of Illustrations

Research Question: What genes in the *Drosophila melanogaster* genome are *mus106* and

mus108?

Introduction

1.1 DNA Repair

The genomes of all living things are constantly exposed to numerous substances that have

the potential to cause DNA damage. These sources of damage encompass both exogenous and

endogenous elements. Exogenous sources, also called mutagens, include ultraviolet light (UV),

ionizing radiation (IR), and environmental and chemical pollutants, while endogenous sources

result from normal metabolic processes (Hakem 2008; Torgovnick 2015). Each of these agents

creates a different type of damage which can cause problems like incorrect base pairing, single

and double stranded breaks, and formation of DNA adducts (Hakem 2008; Torgovnick 2015).

One example of a mutagen is the alkylating agent methyl methanesulfonate (MMS) which adds a

methyl group to either adenine or guanine (Beranek 1990) and results in stalled replication forks

(Lundin 2005). Damage that is left unrepaired can cause accumulation of harmful mutations that

can affect the viability of an organism. To combat this, there are numerous DNA repair pathways

that work to fix damaged DNA and maintain genomic integrity. The major eukaryotic repair

pathways are non-homologous end-joining, homologous recombination, nucleotide excision

repair, base excision repair, and single strand break repair (Hakem 2008). A deficiency in any

protein involved in these repair pathways primes cells for apoptosis or retaining mutations that

can trigger the development of cancer (Torgovnick 2015).

There are several diseases in humans that develop as a result of deficiencies in crucial

DNA repair components, such as xeroderma pigmentosum, spinocerebellar ataxia-26, and Bloom

Syndrome (Torgovnick 2015; O'Driscoll 2012; Hoch 2017). For example, spinocerebellar ataxia is caused by mutations in the X-ray repair cross complementing 1, known as XRCC1, gene. XRCC1 is specifically involved in two repair pathways: base excision repair and single-strand break repair, which are described more fully below.

The Base Excision Repair (BER) pathway is most active in the G1 phase of the cell cycle and is responsible for fixing non-bulky lesions that result from oxidation, alkylation, and deamination damage (Chatterjee 2017; Lee 2019). Many genes involved in this pathway are highly conserved indicating the importance of maintaining this type of repair (Lee 2019). BER follows three main steps to repair damaged bases: recognition, excision, and religation. Damaged bases are recognized by a DNA glycosylase which cleaves one side of the phosphodiester bond to create an AP (apurinic/apyrimidinic) site (Chatterjee 2017; Lee 2019). This AP site is then recognized by AP endonuclease 1 (APE1) and breaks the bond to fully release the damaged base (Lee 2019). The size of the repair determines if short patch or long patch repair will follow. Short patch repair replaces a single nucleotide using polymerase beta and religates the break with ligase 3 and XRCC1 (Chatterjee 2017; Lee 2019). Long patch repair can replace 2 to 13 nucleotides using polymerase gamma/epsilon and religates with ligase 1 and flap endonuclease (FEN1) (Chatterjee 2017; Lee 2019).

Single strand break repair (SSBR) is closely associated with the BER pathway but there are differences in the proteins involved. This repair pathway can be triggered by abasic sites, oxidative damage, or problems from the topoisomerase I (TopI) enzyme (Chatterjee 2017). Similarly to BER there are short and long patch repair pathways in SSBR. The short patch version follows identical to that of BER where APE1 recognizes single stranded damage. Long patch SSBR is initiated by interaction of PARP1 with the site of damage (Chatterjee 2017). The

ends of the break are processed by APE1, polynucleotide kinase 3'-phosphate (PNKP), and

aprataxin (APTX) (Chatterjee 2017). FEN1 will then proceed to remove the 5' end of the

damaged base, resulting in a single strand break. Polymerases will fill in the required nucleotides

and ligation is completed by ligase 1 with proliferating cell nuclear antigen (PCNA) and XRCC1

(Chatterjee 2017). Single strand breaks induced by TopI mimic the long patch version of SSBR,

but the end processing is completed by tyrosyl-DNA phosphodiesterase 1 (TDP1) to remove

TopI from the DNA break (Chatterjee 2017).

1.2 Drosophila as a model system

Scientists use model organisms to gain additional knowledge on processes like DNA

repair. The fruit fly, also known as *Drosophila melanogaster*, has been an important model

organism in the field of genetics. Fruit flies are incredibly useful because they are easy to

genetically modify, easy to keep in a lab setting, have short life spans, and produce numerous

offspring (Hales 2015). Many of the DNA repair pathways are conserved across a wide range of

species, and in *Drosophila* these repair pathways include BER, SSBR, NER and DSBR

(Sekelsky 2017). Additionally, an analysis that compared human disease genes from OMIM to

the *Drosophila* genome found that 77% of the human disease genes evaluated contained a

distinct *D. melanogaster* ortholog (Reiter 2001). For example, *D. melanogaster* contain a gene

for Bloom syndrome helicase that is orthologous to the Bloom helicase found in humans, making

fruit flies a great tool to further study DNA repair mechanisms (Kusano 2001).

1.3 Mutagen Sensitivity

Researchers originally used DNA damaging agents to induce random mutations in fruit flies to study basic genetic concepts but later focused on how the mutations formed from various treatments (Sekelsky 2017). Thus, a series of genetic screens identified *D. melanogaster* mutants that are unable to properly perform DNA repair. The screens were conducted decades ago (Boyd 1976; Boyd 1982; Gatti 1979; Oliveri 1989) however, as of 2017 only 14 of the 58 mutagen-sensitivity (*mus*) genes underlying these DNA repair defects had been identified molecularly (Sekelsky 2017). The mutagen sensitivity genes were identified by screening for mutant flies that exhibited sensitivity to the DNA damaging agent MMS in a mutagen-sensitivity assay (Sekelsky 2017). Briefly, a mutagen-sensitivity assay is conducted as follows: Flies with the mutation of interest are set up in a vial to mate and lay eggs. The same adults are moved to second vial to lay an additional round of eggs and are then discarded. Larvae are treated by exposure to either a control substance, usually water, or the damaging agent of interest, like MMS. The control substance and mutagen is added to the food when larvae are chewing through the food. The larvae defective in DNA repair will experience higher amounts of cell death when exposed to the damaging agent. The relative survival of mutagen treated flies compared to control treated flies will determine the sensitivity of the mutation being studied. Since the initial mutagen sensitivity screens were designed to identify as many *mus* genes as possible, most of the genes were only minimally characterized beyond noting sensitivity to MMS. It is important to match each *mus* line to a mutation in a particular gene to learn more about DNA repair pathways in general and why each line expresses mutagen sensitivity.

To further our understanding of DNA repair pathways in *D. melanogaster*- and by extension, the DNA repair pathways in other organisms- the work described here aimed to study

the genes *mus106* and *mus108* in further detail. Both genes were selected for further study because the FlyBase records for each gene included unpublished mapping data (Laurencon 2001), which increases the likelihood of successfully mapping the gene by narrowing the probable gene regions. Additionally, an allele of each gene is publicly available at the Bloomington Drosophila Stock Center, which is a repository of unique Drosophila stocks used by researchers around the world.

The most recent reference with data on the *mus106^{D1}* and *mus108^{A1}* mutations is an unpublished, personal communication to FlyBase from Laurencon (2001) (Thurmond 2019). The data was part of a small-scale attempt to further map the genes using deficiency mapping. This process involves using chromosomes that contain a deficiency, or a deleted region of genes, to help identify the specific location of a gene of interest. The communication refers to another publication for the treatment protocol, but it is unclear how many replicates were tested for each condition. The communication also does not provide the raw or vial-level data for each cross and treatment. For both mutations, Laurencon (2001) provides data on the homozygous cross in addition to each deficiency cross at 0.08% MMS, 0.1% MMS, and a control. The data provides a great starting point to follow up on to further map *mus106^{D1}* and *mus108^{A1}*.

1.4 *mus106*

The *mus106* gene is X-linked and represented by one recessive allele: *mus106^{D1}* (Boyd 1976; Hawley 1985; Boyd 1990). Females homozygous for the mutation are sterile (Boyd 1976; Boyd 1982; Gatti 1979). Data collected from Boyd (1976) showed that *mus106* mutants exhibit a weak sensitivity to both MMS and gamma radiation, and do not exhibit sensitivity to nitrogen

mustard, AAF, or UV. In addition, *mus106* flies are moderately sensitive to high doses of X-ray exposure (Oliveri 1990).

Using recombination mapping, Boyd (1976) loosely assigned *mus106* to be located near the *forked* gene. Later summary data roughly mapped *mus106* to a region on the X chromosome corresponding to map positions 36-44 (Gatti 1979; Smith 1980). The most recent mapping data (unpublished) from Laurencon (2001) maps *mus106* to be between the genes for *up* and *garnet*. The *mus101* gene is also located within this region, however, complementation tests between *mus101* and *mus106* show they are not two alleles of the same gene (Boyd 1976; Laurencon 2001). A complementation cross is useful to determine if two *mus* mutations are located within the same gene (two different alleles) or in two separate genes. This means that when *mus101* and *mus106* were crossed together, flies of the *mus101/mus106* genotype did not exhibit sensitivity to a particular mutagen, meaning they complemented each other and are located in two separate genes.

1.5 *mus108*

The mus108 gene is also X-linked and represented by one recessive allele: *mus108*^{A1} (Boyd 1976; Hawley 1985; Boyd 1990). Female homozygotes are sterile (Boyd 1976; Boyd 1982; Gatti 1979). *mus108*^{A1} flies are sensitive to MMS, show reduced ability for DNA synthesis after exposure to AAF and exhibit a moderate sensitivity to high doses of X-ray exposure (Smith 1980; Oliveri 1990; Boyd 1982). Additionally, *mus108* mutant flies are neither sensitive to ether anesthesia nor gamma radiation (Gamo 1989). However, there are noticeable gaps in the available information for *mus108*. For example, there is no published data describing the level of

sensitivity that *mus108* mutants exhibit when exposed to MMS. Likewise, it is unknown if *mus108* mutants are sensitive to UV or nitrogen mustard (Smith 1980; Buendia 1998), and there is no published data on any complementation tests involving *mus108*.

An interesting aspect of the *mus108* mutation is that it suppresses the magnification of rDNA (rRNA-encoding genes) in male flies. This means when the number of rDNA tandem repeats are reduced, *mus108* mutants are unable to revert to wild-type levels of rDNA (Hawley 1985). These magnification events are blocked at both the pre-meiotic and meiotic stages (Hawley 1985). This phenotype has only been recorded in *mus108* mutants.

mus108 has been roughly mapped to a region on the X chromosome corresponding to the map position 10.8 (Smith 1980). Laurencon (2001) further mapped the mutation to be within the cytological region 4E1-4EF2.

1.6 Characterization of *mus106* and *mus108*

In addition to mapping the location of the *mus* mutations, it is also important to further characterize them to better understand their predicted role in disrupting critical DNA pathways. Different mutagens cause different types of DNA damage, which can influence the type of repair pathway that is triggered in a cell (Chatterjee 2017). To learn more about the repair pathways that both mutations are involved in, we conducted a sensitivity test with two mutagens that the mutations have not yet been tested against: hydroxyurea (HU) and camptothecin (CPT).

Hydroxyurea is a non-alkylating mutagen that has been used in treating certain cancer types and sickle cell anemia (Singh 2016). Depending on the concentration, length of exposure,

and cell sensitivity, exposure to HU can be cytotoxic. In general, it inhibits DNA synthesis which results in DNA damage, stops cytokinesis and produces oxidative stress (Singh 2016). Specifically, HU targets ribonucleotide reductase (RNR), the enzyme responsible for reducing ribonucleoside diphosphate to deoxyribonucleotide which is required for DNA replication and repair (Singh 2016). Cells will pause in S phase if there are not enough dNTP's to continue synthesizing DNA. As DNA polymerase slows down, a replication checkpoint is activated, causing an increase in RNR production to generate more dNTP's so that DNA synthesis can continue (Singh 2016). However, if the replication checkpoint is not activated, then replication forks will become unstable in the presence of HU and can collapse, cause strand breaks and oxidative stress, and lead to cell death.

Camptothecin is an alkaloid compound used in anticancer treatments by inducing DNA damage in rapidly dividing cancer cells (Mei 2019). CPT specifically inhibits TopI, an enzyme that helps to relieve DNA supercoiling as replication occurs by producing single strand breaks (SSBs) to do so. In normal circumstances, TopI will immediately release from DNA so the strand can be religated, but in the presence of CPT, TopI is trapped on the DNA preventing the religation step (Mei 2019). When this happens, a pathway is triggered so that TopI is degraded and the SSB is repaired through the single strand break repair (SSBR) pathway (Mei 2019). If the breaks are not repaired quickly enough, it can lead to the formation of double stranded breaks (DSBs) and replication failure.

Materials and Methods

2.1 Fly husbandry

Fly stocks were obtained from the Bloomington Drosophila Stock Center (BDSC) at Indiana University. These stocks were: *mus106^{D1}/C(1)DX, y^1 f^1/Dp(1;Y)y$^+$* (BDSC #2318); *mus108^{A1}/C(1)DX, y^1 f^1* (BDSC #1490); *Df(1)ED7217, w^{1118} P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED7217/FM7h* (BDSC #8952); *Df(1)JC70/FM7c* (BDSC #944); *Df(1)BSC823, w^{1118}/Binsinscy* (BDSC #27584); *Df(1)ED6727, w^{1118} P{w[+mW.Scer\FRT.hs3]=3'.RS5+3.3'}ED6727/FM7h* (BDSC #8956); *w^{1118}PBac{w[+mC]=WH}XRCC1^{f03685}* (BDSC #18682); *y^1 v^1; P{y[+t7.7] v[+t1.8]=TRiP.HMJ23251}attP40* (BDSC #61359); *y^1 w*; P{w[+mC]=tubP-GAL4}LL7/TM3, Sb^1 Ser^1* (BDSC #5138) and DGRP-59 (BDSC #28129). The *FM7w* stock was obtained from Jeff Sekelsky at the University of North Carolina at Chapel Hill. Each stock is referred to using the following abbreviations, listed in the same order as above: *mus106^{D1}, mus108^{A1}, Df(1)ED7217, Df(1)JC70, Df(1)BSC823, Df(1)ED6727, PBacXRCC1, TRiP, GAL4, wild-type,* and *FM7w*.

All flies were maintained in bottles on standard corn syrup/soy food (Archon Scientific) in a 25°C incubator on a 12-hour dark; 12-hour light cycle. For each fly cross, stocks were placed in bottles ten days prior to virgining and cleared three to five days later to avoid overcrowding. Kimwipes were added to the bottles as needed to absorb moisture and provide newly-eclosed flies a solid surface to walk on. For each cross, virgin females of the appropriate genotype were collected across four days and kept in separate vials by day. On the following day, crosses were set up between the virgin females and males of the appropriate genotype.

2.2 Determining Mutagen Sensitivity

To determine mutagen sensitivity, crosses were set up in bottles each containing around 20 virgin *FM7w* females and 15 mutagen sensitive males (*mus106^{D1}* or *mus108^{A1}*) (Figure 1A). Ten days later, virgin females of the *mus106^{D1}/FM7w* or *mus108^{A1}/FM7w* genotype were collected for a total of four days. On the fifth day, crosses were set up into ten vials each with five virgin *mus/FM7w* females and five of the corresponding mutagen sensitive males. The vials were put through a mutagen sensitivity assay, which is described below:

Ten vials were set up with five females and five males for each cross, creating Brood 1 (Day zero). On Day three, the parents were flipped into new vials to create Brood 2. On Day four, each vial of Brood 1 was treated with 250μL of ddH$_2$O. Brood 2 vials were cleared out on Day five and treated with 250μL of 0.08% or 0.1% MMS (Sigma) dissolved in water on Day six. Brood 1 flies were frozen on Day 14 and Brood 2 flies frozen on Day 17. Relative survival was calculated as the ratio of mutant to non-mutant flies in the mutagen-treated (Brood 2) vials, normalized to the same ratio in the control (Brood 1) vials. Only vials that contained 20 or more flies were used for relative survival calculations (as in Romero et al. 2016). A low relative survival value indicates strong sensitivity to the mutagen. Flies were scored depending on the sex and eye phenotype, resulting in four potential phenotypes: female with wild-type eyes, males with wild-type eyes, females with Bar eyes, and males with Bar eyes (Figure 1B).
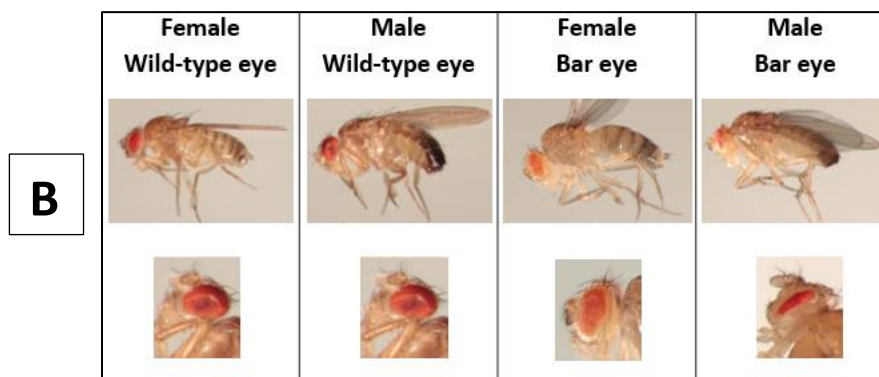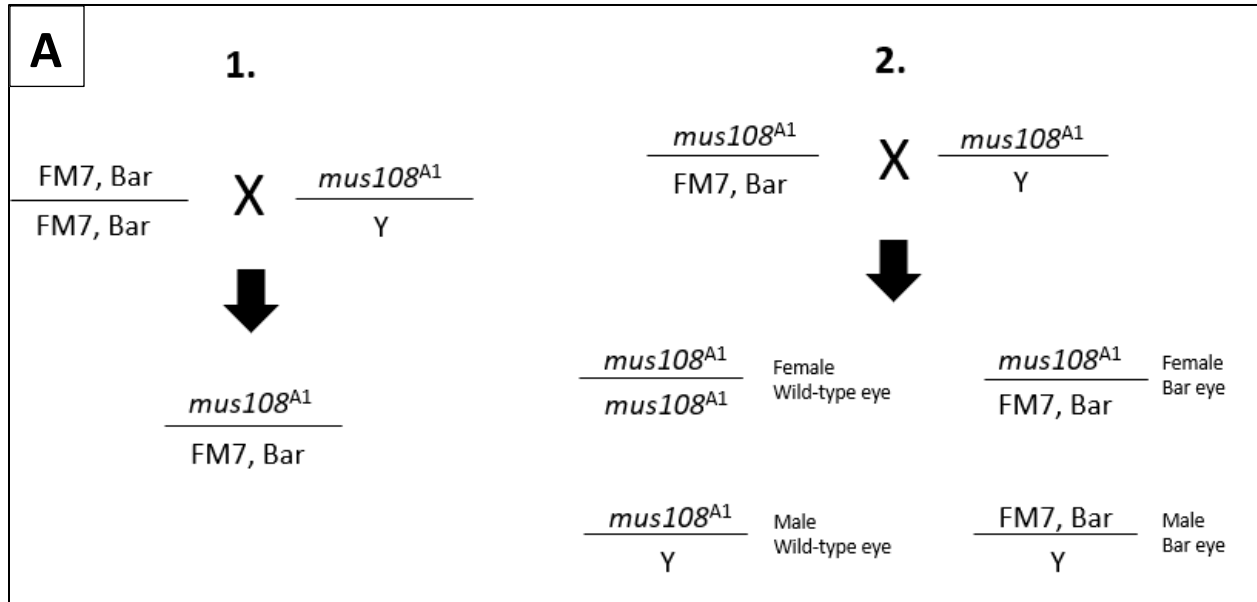
Figure 1. Cross to determine mutagen sensitivity. A) Two step cross to determine mutagen sensitivity in homozygous *mus* mutants. Cross applies to *mus106^{D1}* as well. B) The eye phenotype and sex of the potential offspring that result in the mutagen sensitivity assay. Both *mus* crosses result in the same phenotypes.

2.3 Deficiency Mapping

Deficiency stocks were chosen based on unpublished data from Laurencon's (2001) communication to FlyBase. In this communication, it was suggested that *mus106* is located

between the genes *up* and *garnet.* The deficiency Laurencon found to disrupt *mus106* was originally sought. However, there was a difference in nomenclature on FlyBase, making it unclear which deficiency was being referred to. Upon gathering further information, it was later determined that *Df(1)g* was the deficiency originally used by Laurencon (green bar in Figure 2). In place of *Df(1)g,* the deficiency *Df(1)ED6727* was chosen since it was smaller, covered the undefined region found in *Df(1)g*, and covered the area between *up* and *garnet* (Figure 2). For mapping *mus108,* one deficiency used by Laurencon (2001), *Df(1)JC70*, was obtained in addition to two other stocks. *Df(1)ED6727* covers the undefined portion at the beginning of *Df(1)JC70,* and *Df(1)BSC823* sits near the end of *Df(1)JC70* covering the other undefined region (Figure 3).

Five virgin females of deficiency stock *Df(1)ED6727* were crossed to five *mus106^{D1}* males in ten vials for two separate rounds. Five virgin females of the deficiency stocks *Df(1)ED6727, Df(1)JC70,* and *Df(1)BSC823* were each crossed to five *mus108^{A1}* males in ten vials each. For comparison, five *wild-type/FM7w* females were crossed to five *wild-type* males in ten vials each. The MMS sensitivity assay was then conducted as described above in "Determining Mutagen Sensitivity".
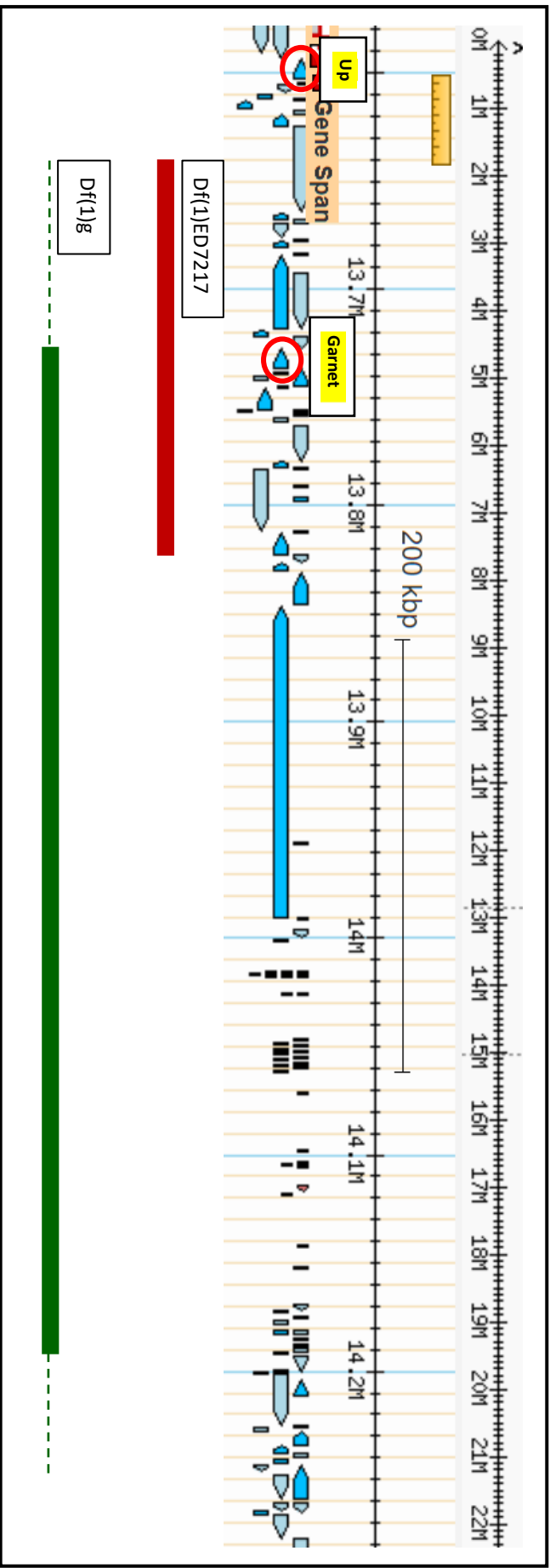
Figure 2. Genomic location of the deficiency used for mapping *mus106*. Blue arrowheads represent genes of various sizes and the direction in which they are transcribed (derived from FlyBase, Thurmond 2019). The solid red bar shows the defined region that is deleted from the genome in the *Df(1)ED7217* deficiency stock. The green bar shows the region that is deleted from the genome in the *Df(1)g* deficiency stock originally used by Laurencon (2001). Dotted green lines are areas of *Df(1)g* that do not have a defined end point. Red circles indicate the location of the genes *up* and *garnet*.
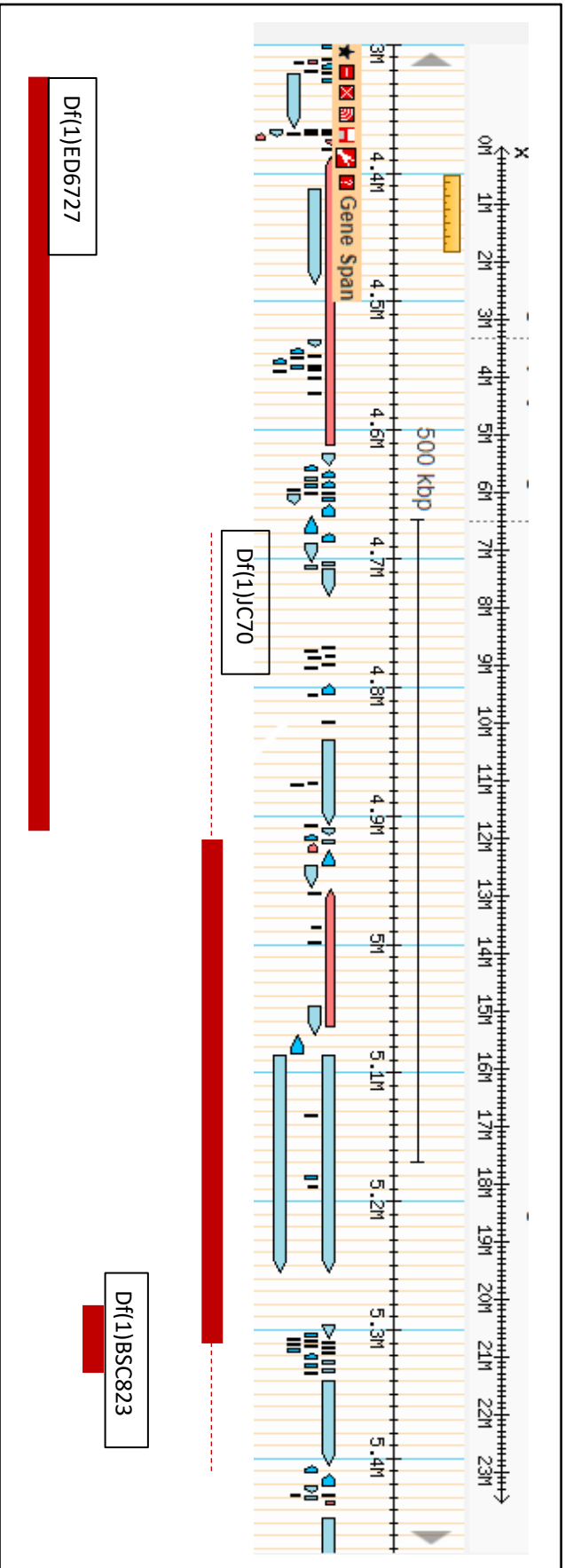
Figure 3. Genomic location of each deficiency used for mapping *mus108*. Blue arrowheads represent genes of various sizes and the direction in which they are transcribed (derived from FlyBase, Thurmond 2019). The solid red bars show defined regions that are deleted from the genome under each respective deficiency stock. Dotted red lines are areas of *Df(1)JC70* that do not have a defined end point.

2.4 Selection of Candidate Gene for *mus108*

FlyBase (Thurmond 2019) was used to determine the biological functions of all genes uncovered by *Df(1)JC70* and a candidate gene was chosen from among these.

2.5 Sequencing *XRCC1*

DNA was extracted from a single *mus108[A1]* male using the "Fly Squish protocol" as in Gloor et al. 1993. Forward and reverse XRCC1 PCR primers (Table 1; Integrated DNA Technologies) were designed using the FlyBase reference genome sequence (Thurmond 2019) and PCR was used to amplify the candidate gene, *XRCC1*. Optimization of PCR product was made by altering $MgCl_2$ concentrations, type of polymerase enzyme, extension time, and annealing conditions. The following conditions produced the desired product: 1μL template DNA (derived from Fly Squish protocol above), 0.63μL dNTP (10mM each NTP; Promega), 0.63μL 50pmol/μL primer -72, 0.63μL 50pmol/μL primer 2356a, 5μL 5x iProof (Bio-Rad) standard buffer, 0.75μL 50mM $MgCl_2$ (final concentration: 1.5mM), 16.13μL H_2O, and 0.25μL iProof (Bio-Rad) high-fidelity DNA polymerase with an annealing temperature of 58.4°C and a one minute and 30 second extension time.

PCR samples were then run on 0.7% agarose gel with 1x TAE buffer solution. After optimizing PCR conditions to obtain the desired product, the DNA band was excised from the gel, then purified using the GeneJet Gel Extraction kit (Thermo Scientific) according to manufacturer instructions and eluted in 50μL ddH_2O. The purified DNA and primers were then sent to Eurofins Genomics for Sanger sequencing.

Sequencing data was uploaded into the Sequencher program (Gene Codes Corporation) to align the samples to the wild-type *XRCC1* sequence on FlyBase (Thurmond 2019). Any nucleotide differences between the reference sequence and the *mus108^{A1}* sequence were recorded along with the amino acid encoded by those nucleotides. All observed mutations were confirmed in a second fly prep and sequencing reaction.

Primer Name	Primer Sequence
XRCC1 -72	5'-GGCGCAACTGTCGGCAAAG-3'
XRCC1 440	5'-GCGGTGGGAGATGAATGTCG-3'
XRCC1 502	5'-GCATGTAGTGGCTGCTGCGG-3'
XRCC1 1063	5'-GCACAATTCGTCTCCAGAG-3'
XRCC1 1260a	5'-CAGATAGTCTGCACACAGCC-5'
XRCC1 1632	5'-GATTGTGACACGCAGTTGG-3'
XRCC1 2192a	5'-CGAGCATATTCGATTAAC-3'
XRCC1 2356a	5'-GCTAGAGATGGACCATCG-3'

Table 1: PCR and Sequencing Primers. Primers are numbered according to their location with respect to the ATG of *XRCC1* (-72 was 72 nucleotides before ATG). The lowercase 'a' refers to the antisense strand used for reverse primers. Primers -72 and 2356a were used to isolate the full *XRCC1* gene from *mus108^{A1}* males. The remaining primers were distributed throughout the coding sequence to be used for Sanger sequencing. All primers were used for sequencing.

2.6 XRCC1 Alignments

The following protein orthologs of XRCC1 were obtained from NCBI (National 1988): *Drosophila melanogaster* (accession: NP_572217) , *Danio rerio* (accession: NP_0010033988), *Xenopus tropicalis* (accession:XP_031761618), *Homo sapiens* (accession:NP_006288.2), *Mus musculus* (accession:NP_033558). The sequences were aligned using default parameters in the Clustal Omega program (Sievers 2011), then viewed using GeneDoc (Nicholas 1997).

2.7 RNAi Knockdown of *XRCC1*

Crosses were set up in ten vials each with five virgin *TRiP* females and five *GAL4* males (Figure 4). The same MMS sensitivity assay was conducted as described above in "Determining Mutagen Sensitivity". The offspring showed the following phenotypic combinations: wild-type bristles and wild-type wings, stubble bristles and wild-type wings, wild-type bristles and curly wings, and stubble bristles and curly wings. Each phenotypic combination was seen in both sexes of flies.
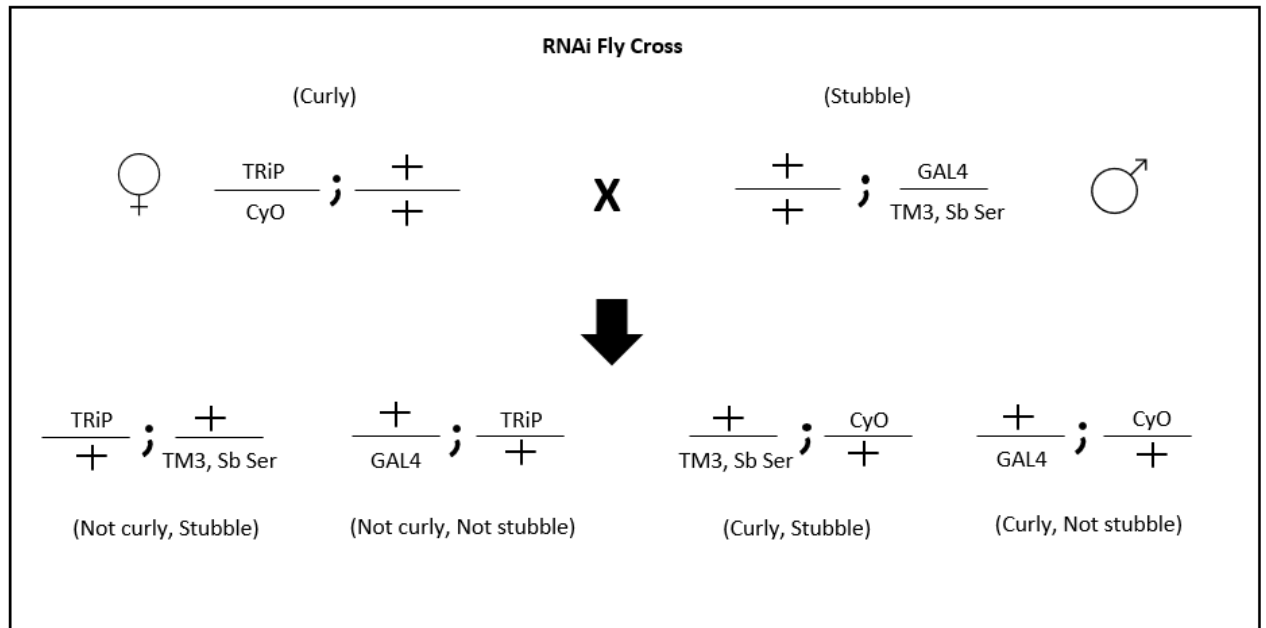
Figure 4. Cross for RNAi knockdown of *XRCC1*. Cross between females with the RNAi construct balanced over *CurlyO* (balancer for the second chromosome) to males with *GAL4* balanced over *TM3, Sb, Ser* (balancer for the third chromosome). *+: wild-type; CyO: CurlyO; Sb: Stubble; Ser: Serrate*.

2.8 *mus108* Complementation Test (with transposon in *XRCC1*)

One tool available for exploring *XRCC1* was a piggyBac stock that contains a transposable element located in the 5'UTR of the *XRCC1* transcript (Figure 5). Crosses were set up in bottles each containing around 20 virgin *PBacXRCC1* females and 15 *FM7w* males (Figure 6). Ten days later, virgin females of the *PBacXRCC1/FM7w* genotype were collected for a total of four days. On the fifth day, crosses were set up into ten vials each with five virgin *PBacXRCC1/FM7w* females and five *mus108^{A1}* males. The same MMS sensitivity assay was conducted as described above in "Determining Mutagen

Sensitivity". The offspring scored showed the same sex and phenotype combinations as seen in Figure 1.
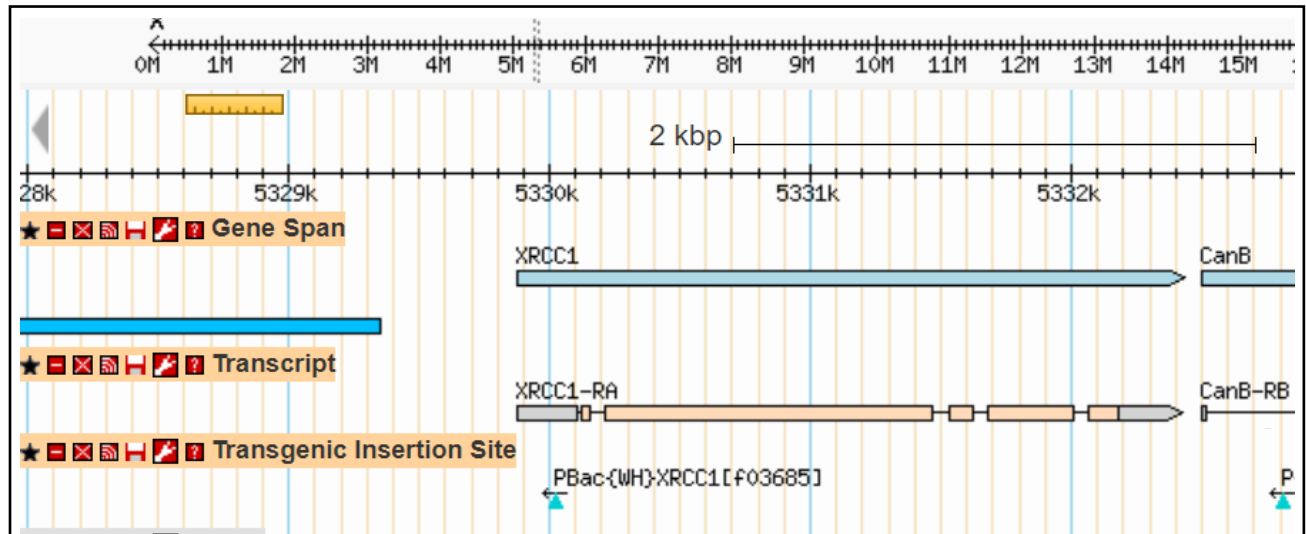
Figure 5. Genomic location of *XRCC1* gene, *XRCC1* transcript, and *PBacXRCC1* element. Blue arrowhead represents the gene and the direction in which it is transcribed. Gray regions on the left and right side of the *XRCC1* transcript represent 5' and 3' UTRs, respectively. Tan regions represent exons and the lines in between tan sections represent introns. Blue triangle indicates the transgenic insertion site of the *PBacXRCC1* transposon (derived from FlyBase, Thurmond 2019).
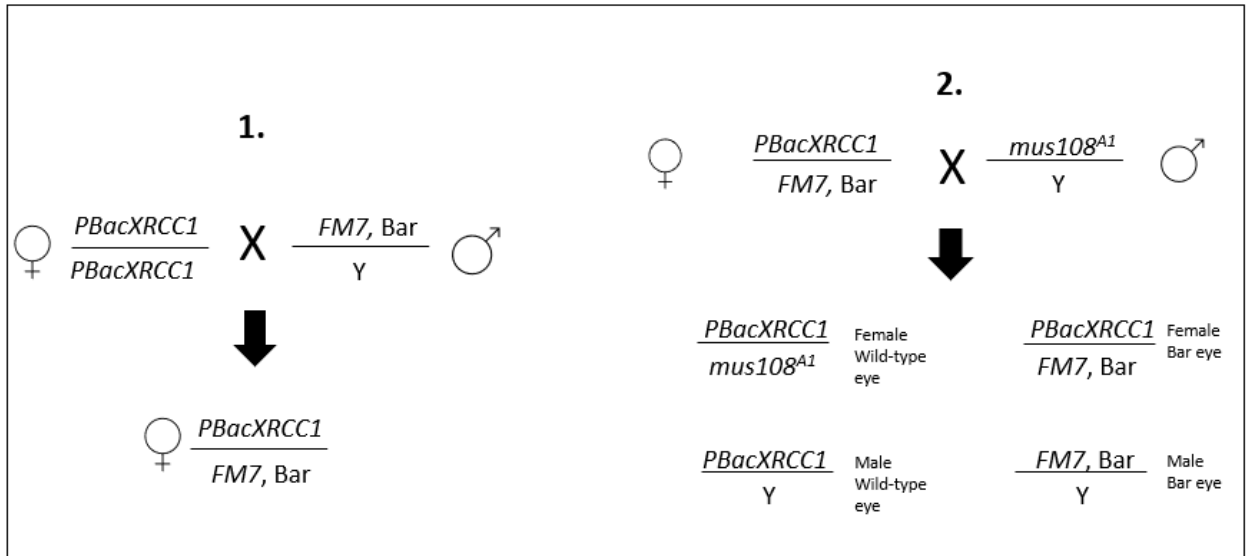
Figure 6. Complementation cross (with transposon in *XRCC1*). Two step cross to determine if *mus108^{A1}* is allelic to *XRCC1*.

2.9 Additional Characterization of *mus108^{A1}*

To test the sensitivity of *mus108^{A1}* flies to two additional mutagens, the same cross set up and mutagen sensitivity assay was conducted as described in "Determining Mutagen Sensitivity" with differences in mutagens used. To test sensitivity to hydroxyurea, Brood 2 vials were treated with 250µL of 80mM hydroxyurea (Arcos) dissolved in water. To test sensitivity to camptothecin , a stock solution of 10mg/mL camptothecin (Arcos) in DMSO was made and diluted into a 1% Tween-20/5% ethanol solution to create a 50µM solution. Brood 1 vials were treated with 250µL of 1% Tween-20/5% ethanol solution, and Brood 2 vials were treated with 250µL of 50µM camptothecin.

Results

3.1 Determining Mutagen Sensitivity

To verify the presence of a mutagen-sensitivity allele in the *mus106^{D1}* and *mus108^{A1}* stocks, an MMS sensitivity assay was conducted on homozygous individuals from each *mus* stock, and the average relative survival was calculated. Flies homozygous for *mus106^{D1}* showed an overall relative survival of 0.94 ±0.47 when exposed to 0.08% MMS, and an overall relative survival of 0.72 ±0.28 when exposed to 0.1% MMS (Figure 7). Flies homozygous for *mus108^{A1}* showed an overall relative survival of 0.47 ±0.25 when exposed to 0.08% MMS, and an overall relative survival of 0.26 ±0.13 when exposed to 0.1% MMS (Figure 8). Both mutagen sensitive stocks showed a decrease in relative survival with an increase in the MMS concentration. Overall, the *mus108^{A1}* flies showed a lower relative survival than the *mus106^{D1}* flies.

Figure 7. Sensitivity of *mus106^{D1}* flies to MMS. Each point represents one vial. Large

horizontal line is the mean, while the upper and lower lines show standard deviation.

Graph includes two rounds of sensitivity data. Only vials with 20+ offspring were
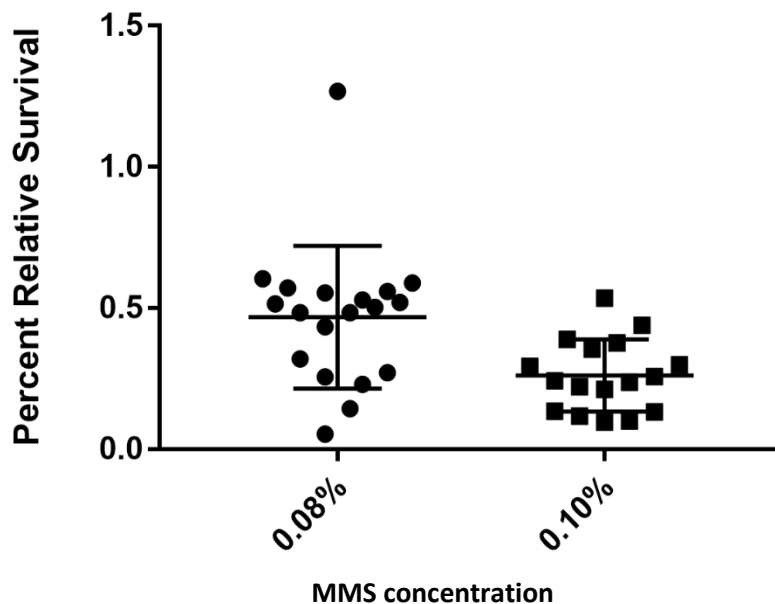
included.

Figure 8. Sensitivity of *mus108^{A1}* flies to MMS. Each point represents one vial. Large

horizontal line is the mean, while the upper and lower lines show standard deviation.

Graph includes two rounds of sensitivity data. Only vials with 20+ flies were included.

3.2 Deficiency Mapping

Crosses were conducted between the mutagen sensitivity and deficiency stocks to

narrow the location of each *mus* gene. Each deficiency stock contained a deletion of

varying length and location on the X chromosome. The mutagen sensitivity assay was

performed on these crosses and the relative survival calculated.

The overall relative survival value for *mus106^{D1}/Df(1)ED7217* was 1.60 ±0.83

when exposed to 0.08% MMS (Figure 9). The overall relative survival values for

mus108^{A1}/Df(1)ED6727 was 0.99 ±0.38, *mus108^{A1}/Df(1)JC70* was 1.00 ±0.41, and

mus108^{A1}/Df(1)BSC823 was 0.95 ±0.53 when exposed to 0.08% MMS (Figure 10). The

relative survival values for these deficiencies were very close and have a somewhat large

standard deviation, although they are not as far spread as the *mus106* deficiency data. An

ANOVA test was performed on the relative survival values of the three *mus108^{A1}*

deficiencies and they were not found to be statistically different (p=0.9727). A cross of

homozygous wild-type flies was also conducted as a comparison. The average relative

survival for the wild-type cross was 0.97 ±0.17 when exposed to 0.08% MMS (Figure

10). The *mus108* deficiency data was compared to the wild-type data and still presented

with no statistical difference (ANOVA; p=0.9941).
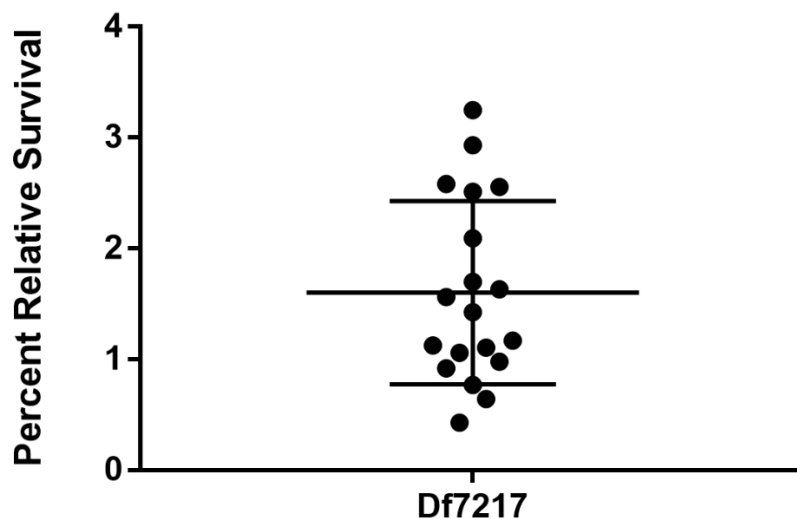
Figure 9. Sensitivity of *mus106^{D1}/Df(1)ED7217* flies to 0.08% MMS. Large horizontal

line is the mean, while the upper and lower lines show standard deviation. Graph includes

two rounds of sensitivity data. Each point represents one vial. Only vials with 20+
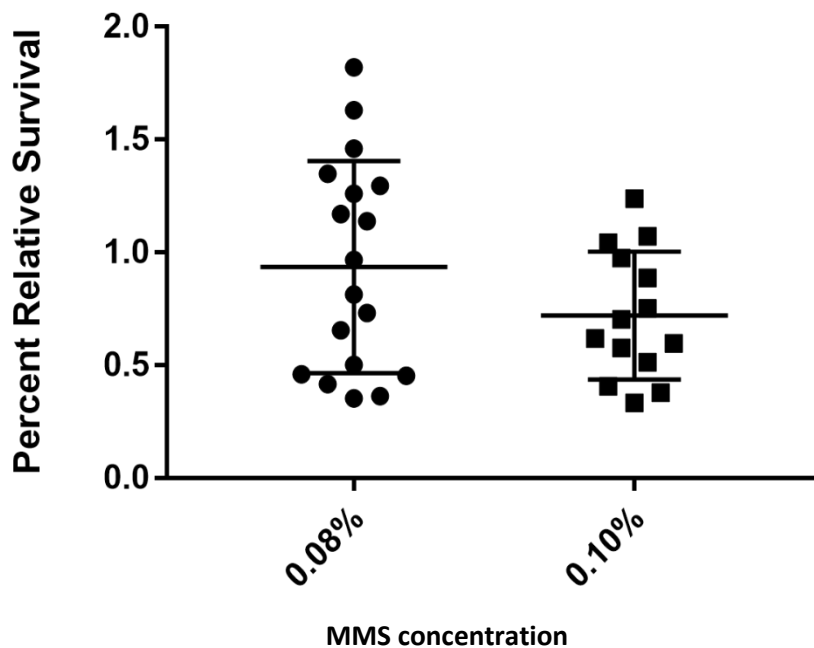
offspring were included.
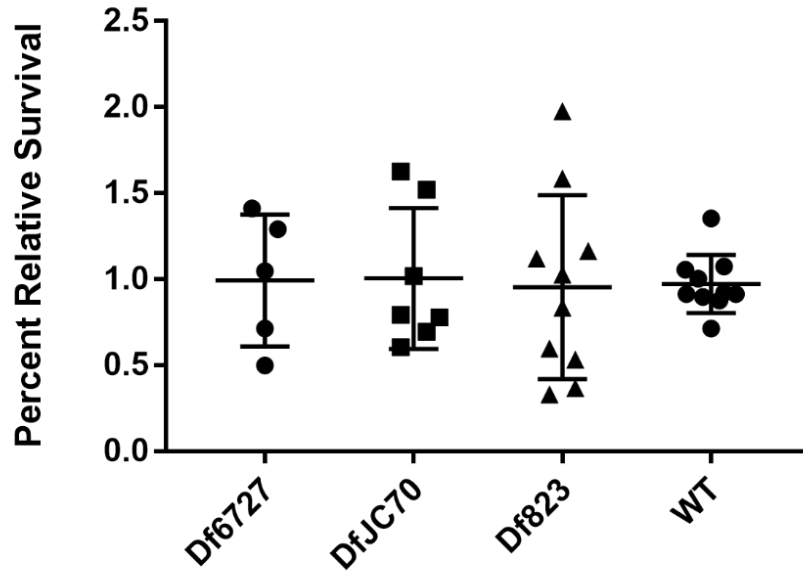
Figure 10. Sensitivity of *mus108^{A1}/Df* flies to 0.08% MMS. Each point represents one

vial. Large horizontal line is the mean, while the upper and lower lines show standard

deviation. Only vials with 20+ offspring were included. There was no significant

difference between the mean relative survival values of the three deficiencies (one-way

ANOVA; p=0.9727), or when compared to the wild-type cross (one-way ANOVA;

p=0.9941).

3.3 Prediction of Candidate Gene for *mus108*

Although the relative survival values for the three *mus108^{A1}* deficiency crosses

were similarly close to 1, previous data suggested *Df(1)JC70* as a possible location for

mus108^{A1} (Laurencon 2001). A candidate gene was sought by investigating all genes

uncovered by *Df(1)JC70* (Table 2). Three genes with a molecular function involved with DNA were *Klf15* (Kruppel-like Factor 15), *MCM3* (Minichromosome maintenance 3) and *XRCC1* (X-ray repair cross complementing 1) which are involved in sequence-specific DNA binding, DNA replication origin binding, and damaged DNA binding, respectively.

Symbol	Name	Molecular Function
Pp2C1	Protein phosphatase 2C	Protein serine/threonine phosphatase activity
Ctp	Cut up	Protein binding; Dynein light intermediate chain bindings
Pdha	Pyruvate dehydrogenase E1 alpha subunit	Pyruvate dehydrogenase activity
CG7024	n/a	Dehydrogenase activity
Proc-R	Proctolin receptor	Proctolin receptor activity; Neuropeptide receptor activity
Klf15	Kruppel-like Factor 15	Sequence-specific DNA binding; DNA binding transcription factor activity
CG6978	n/a	Transmembrane transporter activity
CG42594	n/a	Potassium ion leak channel activity
CG6927	n/a	pH-gated chloride channel activity; Neurotransmitter receptor activity
CG32772	n/a	RNA polymerase II cis-regulatory region sequence specific DNA binding

CG4041	n/a	ATP binding; GTPase activator activity
CG6903	n/a	Heparan-alpha-glucosaminide N-acetyltransferase activity
Ptp4E	Protein tyrosine phosphatase 4E	Protein tyrosine phosphatase activity
Ovo	n/a	RNA polymerase II cis-regulatory region sequence specific DNA binding
CG32767	n/a	RNA polymerase II cis-regulatory region sequence specific DNA binding
Rg	Rugose	Protein kinase binding; Protein kinase A binding
Rpn13R	Regulatory particle non-ATPase 13-related	Proteasome binding; Ubiquitin binding
Cdk7	Cyclin-dependent kinase 7	Serine/threonine kinase activity; ATP binding
snf	Sans fille	Protein of U1 & U2 snRNPs; Assemble spliceosome
TrxT	Thioredoxin T	Disulfide oxidoreductase activity
dhd	Deadhead	Disulfide oxidoreductase activity
Rnp4F	RNA-binding protein 4F	mRNA binding
Mcm3	Minichromosome maintenance 3	DNA replication origin binding

		Rab guanyl-nucleotide exchange factor
CG3309	n/a	activity
XRCC1	XRCC1	Damaged DNA binding
CanB	Calcineurin B	Calcium ion binding
SK	Small conductance calcium-activated potassium channel	Photoreceptor activity; Calmodulin binding
NAAT1	Nutrient amino acid transporter 1	Transmembrane transporter activity

Table 2. Abbreviated list of genes uncovered by the *Df(1)JC70* deletion. All genes with an unknown molecular function (n = 24) were removed from this abbreviated list (See Table5 in Appendix for complete list of genes uncovered by *Df(1)JC70*). Genes are listed in the same order as they appear on the *D. melanogaster* X chromosome (Thurmond 2019).

3.4 Sequencing *XRCC1*

Prior to sequencing the *mus108* candidate gene *XRCC1*, multiple PCR reactions were run to optimize gene amplification. Each set of reaction conditions was tested on two primer combinations (XRCC1 -72 with XRCC1 2192a and XRCC1 -72 with XRCC1 2356a), which were predicted to produce amplicons of 2264 base pairs and 2428 base pairs, respectively. An initial PCR reaction was run using a touchdown protocol where

the annealing temperature decreased by 0.5°C each cycle, from 62°C to 54°C, and

2.5mM $MgCl_2$. This reaction produced faint or non-existent bands at the predicted

amplicon sizes (data not shown). Optimization steps included using a PCR gradient to

test multiple annealing temperatures, using *Taq* versus iProof polymerase, and using

iProof HF versus GC buffers. Each of these conditions again produced extremely faint or

non-existent bands (data not shown). A gradient PCR using the -72 and 2356a primers

and a decreased $MgCl_2$ concentration of 1.5mM produced a clearly defined band at an

annealing temperature of 58.4°C (Figure 11). This band was excised and purified for

sequencing using all of the primers listed in Table 1.
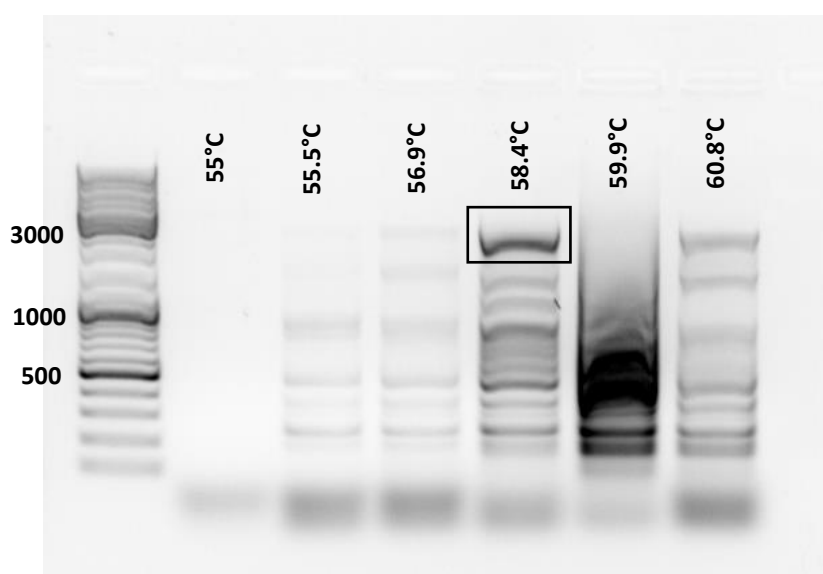
Figure 11. Isolation of XRCC1 from *mus108^{A1}* males. Representative gel from attempts

to optimize conditions for amplification of XRCC1. Lanes differ by annealing

temperature. Lane one contains the O'GeneRuler DNA Ladder Mix (Thermo Scientific).

Remaining lanes contain amplicons generated from *mus108^{A1}* males using primers -72

and 2356a using conditions as described in Materials and Methods. Black box identifies

the band excised for sequencing, ~2250 bp.

3.5 XRCC1 Alignments

Following Sanger sequencing of the -72 and 2356a *XRCC1* amplicon, the following missense mutations between the *Drosophila melanogaster* reference genome sequence and the *mus108^{A1}* sequence were identified: Q127K, Q195E, T196M, I203V, and S204T. Four silent mutations were also found (data not shown). To assess the degree of conservation for each of these residues, an alignment of XRCC1 protein sequences from five species was generated as shown in Figure 12.

Figure 12. Alignment of XRCC1 orthologs. Highlighted regions indicate the domain regions in the human XRCC1 sequence as defined by (London 2015). Yellow= N-Terminal Domain; Blue= Linker 1 containing a Rev1-interacting region (RIR) and a Nuclear localization signal (NLS); Green= BRCTa domain; Pink= Linker 2 containing a phosphorylated FHA binding sequence (PFBS); Orange= BRCTb domain. Red stars indicate missense mutations found in the *mus108*[A1] sequences as noted in the text. Abbreviations are as follows: Dmel: *Drosophila melanogaster*; Drer: *Danio rerio*; Xtro: *Xenopus tropicalis*; Hsap: *Homo sapiens*; Mmus: *Mus musculus*.

3.6 RNAi Knockdown of *XRCC1*

A cross between *TRiP* and *GAL4* flies was conducted to determine if a knockdown of *XRCC1* would express a mutagen sensitive phenotype after exposure to MMS. The overall relative survival value for the *TRiP/GAL4* cross was 1.16 ±0.33 (Figure 13).
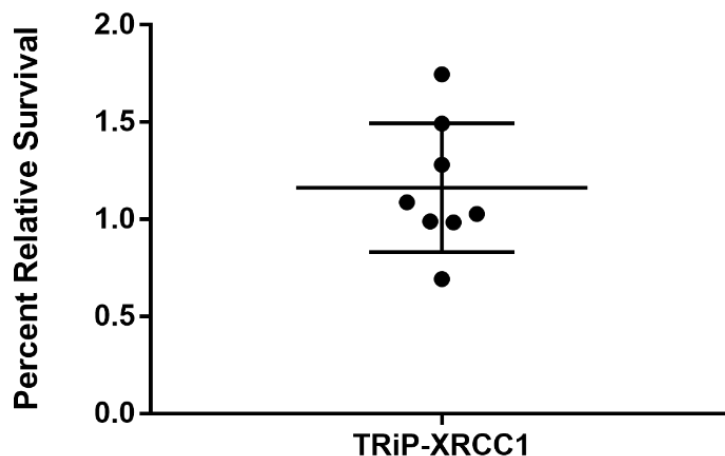
Figure 13. Sensitivity of *TRiP/GAL4* flies to 0.08% MMS. Each point represents one vial. Large

horizontal line is the mean, while the upper and lower lines show standard deviation. Only vials

with 20+ offspring were included.

3.7 *mus108* Complementation Test (with transposon in *XRCC1*)

To determine if *XRCC1* is *mus108^{A1}*, a complementation cross between *PBacXRCC1*,

which contains a transposable element in the 5' UTR of *XRCC1*, and *mus108^{A1}* was conducted.

The overall relative survival for the *PBacXRCC1* cross was 0.87 ±0.29 (Figure 14). However, it

should be noted that these data were generated from only six vials (replicates) due to a lack of

PBacXRCC1/*FM7* females collected during the virgining process.
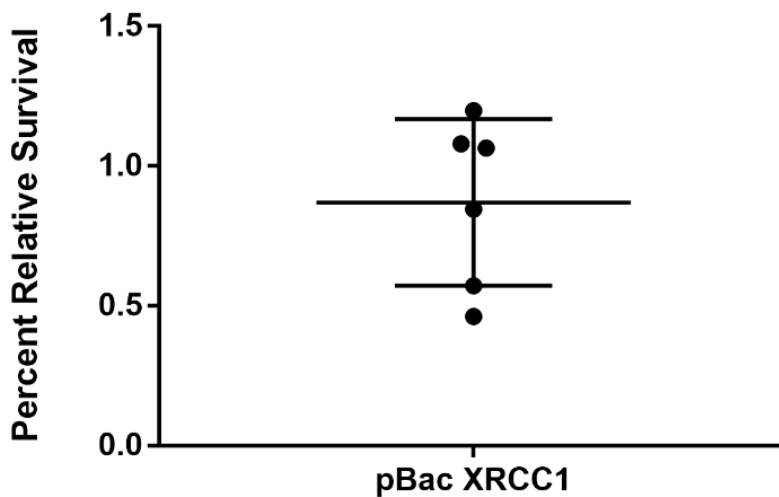
Figure 14. Sensitivity of *PBacXRCC1*/*mus108^{A1}* flies to 0.08% MMS. Each point represents one

vial. Large horizontal line is the mean, while the upper and lower lines show standard deviation.

Only vials with 20+ offspring were included.

Discussion

4.1 *mus106*

Previous research has shown that the *mus106^{D1}* mutation is recessive, located on the X chromosome, and it is the only identified allele of *mus106* (Boyd et al 1976; Hawley 1985). *mus106^{D1}* mutants exhibit sensitivity to MMS, gamma radiation, and X-rays (Boyd et al 1976; Oliveri 1990). The most recent data on *mus106^{D1}* was collected in 2001 in an attempt to further narrow down the location of the mutation (Laurencon). In that study, Laurencon identified one deficiency that failed to complement *mus106^{D1}* in MMS sensitivity assays at both 0.08% and 0.1% MMS (2001). Now, 20 years later, before further mapping efforts were made, we needed to determine whether the mutant allele was still present in the stock by assessing sensitivity of *mus106^{D1}* homozygotes to MMS. As shown in Figure 7, the relative survival values were 0.94 and 0.72 at 0.08% and 0.1% MMS, respectively. Even though *mus106^{D1}* is considered to have a weak sensitivity to MMS (as compared to other *mus* alleles (Boyd 1976)), an average relative survival of 0.94 at 0.08% MMS is much higher than would be expected for any "weak" sensitivity allele and is higher than the relative survival observed for *mus106^{D1}* when it was initially discovered (~20%) (Boyd 1976). An increase in MMS concentration however did further reduce relative survival (Figure 7), a feature which has been observed with most *mus* genes, including *mus106* (Boyd 1976).

Nonetheless, the current 0.08% MMS data has a relatively large standard deviation (±0.47), indicating many of the data points lie far from the average (Figure 7). For the 0.1% MMS data the standard deviation was only ±0.28 (Figure 7). With such a high standard deviation for the 0.08% MMS data, it is difficult to decide anything with certainty. It is unclear what caused such a large value given that both rounds of the assay were conducted by the same person

who followed the same protocol for every vial, including using the same commercially-prepared food, same Drosophila incubator, and same methods over a two-week span.

The deficiency data for *mus106^{D1}* was also very spread out and had an average relative survival over 1 (Figure 9). A value of 1 represents that there is no difference in the ratio of *mus106^{D1}/Df(1)ED7217* flies to their balanced siblings between the MMS treatment and the control. If Laurencon's prediction is correct and *mus106^{D1}* is between the genes for *up* and *garnet*, we would have seen much stronger sensitivity using the *Df(1)ED7217* deficiency than we did. It is possible that *mus106^{D1}* simply does not lie within the deficiency *Df(1)ED7217*, considering the deficiency that Laurencon (2001) observed to disrupt *mus106^{D1}* (*Df(1)g*) was much longer (606,616 bp) than *Df(1)ED7217* (180,238 bp) (Figure 2). However, it would not be expected for numerous data points in the deficiency cross to be above a relative survival of 2 or 3. Relative survival values in this range would indicate *mus106^{D1}/Df(1)ED7217* flies are surviving in a greater ratio to their balanced siblings. This is not expected since flies with a deficiency have a region of their genome missing and typically do not survive better than those without such a deletion (Cook 2012). *mus106^{D1}* could be located within a different portion of the *Df(1)g* deficiency. We only pursued a deficiency that covered the beginning of *Df(1)g* because of the undefined start point in *Df(1)g* and because *Df(1)ED7217* covered most of the genes between *up* and *garnet*. To confirm if *mus106^{D1}* is located within a different part of *Df(1)g*, additional deficiencies that cover the remaining portion of *Df(1)g* could be tested with *mus106^{D1}*.

Based on the data collected on *mus106*, it is possible that the *mus106^{D1}* mutation is no longer present in the stock. It has been almost 20 years since this *mus* allele has been worked with (and 45 years since its original discovery) and it is possible that over time the mutation has been lost throughout numerous generations of maintaining the stock. A loss of the mutation

could explain why the homozygous cross and deficiency cross showed high average relative survival values. If the mutation is no longer present then DNA damage can be repaired properly, and the "mutant" flies would be surviving as well as their balanced siblings.

Since Laurencon narrowed down the potential location of *mus106* quite specifically and again determined that *mus106* and *mus101* are not alleles of the same gene, we investigated the molecular function of genes between *up* and *garnet* to choose a potential candidate gene (Table 3) (Thurmond 2019). We propose DNA ligase 4 (*DNAlig4*) as a potential candidate gene for *mus106^{D1}*. Out of the genes seen in Table 3, *DNAlig4* is the only gene that is involved in DNA replication and repair, consistent with the observation that mutagen sensitivity occurs when proteins involved in DNA repair are defective or mutated.

There is no way to confirm that the *mus106^{D1}* mutation discovered by Boyd (1976) was an allele of *DNAlig4* if the *mus106^{D1}* mutation is truly gone, since we cannot perform complementation analysis with only one allele of *mus106*. However, we could do a series of mutagen sensitivity assays with alleles of the predicted candidate gene, *DNAlig4*, to see if it presents similar mutagen sensitivity to previous data seen on *mus106*. If so, this would further suggest that *mus106* was *DNAlig4*.

Symbol	Name	Molecular Function
Up	Upheld	Tropomyosin binding; calcium ion binding
Ndc80	Ndc80	Protein binding
NFAT	Nuclear factor of activated T-cells	DNA-binding transcription factor activity; chromatin binding activity

DNAlig4	DNA ligase 4	DNA ligase (ATP) activity; DNA metabolism (response to IR)
CG11164	n/a	RNA-DNA hybrid ribonuclease activity
CG15760	n/a	localize integral component of organelle membrane
CG11162	n/a	C-4 methylsterol oxidase activity
Nadsyn	NAD synthetase	NAD+ synthase activity; glutaminase activity
Nna1	Nna1 carboxypeptidase	metallocarboxypeptidase activity; zinc ion binding activity
CG9941	n/a	ubiquitin protein ligase activity; zinc ion binding
G	Garnet	cargo adaptor activity

Table 3. Abbreviated list of genes between *up* and *garnet*. All genes with an unknown molecular function (8) were removed from this abbreviated list (See Table 6 in Appendix for complete list of genes between *up* and *garnet)*. Genes are listed in the same order as they appear on the *D. melanogaster* X chromosome (Thurmond 2019).

4.2 *mus108*

Previous research has shown that the *mus108^{A1}* mutation is recessive, located on the X chromosome, and it is the only identified allele of *mus108* (Boyd et al 1976; Hawley 1985). *mus108^{A1}* mutants exhibit sensitivity to MMS, AAF, and X-rays (Smith 1980; Oliveri 1990). Like *mus106^{D1}*, the most recent data on *mus108^{A1}* was collected in 2001 in an attempt to further

narrow down the location of the mutation (Laurencon). In Laurencon's study, four deficiencies failed to complement $mus108^{A1}$ in an MMS sensitivity assay at both 0.08% and 0.1% (2001).

Again, it was important to determine whether the mutant allele was still present in the stock by assessing sensitivity of $mus108^{A1}$ homozygotes to MMS. $mus108^{A1}$ flies exhibited a moderate level of sensitivity when exposed to 0.08% MMS, and an increase in MMS concentration to 0.1% further lowered relative survival. As shown in Figure 8, the relative survival values were 0.47 and 0.26 at 0.08% and 0.1% MMS, respectively. The standard deviation is also smaller (±0.25 at 0.08%; ±0.13 at 0.1%) than what was seen with the $mus106$ data, which more confidently suggests that flies homozygous for $mus108^{A1}$ are still sensitive to MMS. From the previous studies on $mus108^{A1}$ there are no data that indicates what level of sensitivity this stock should show with exposure to MMS, only that it is sensitive (Smith & Dunsberry 1980). Although we cannot directly compare this to what was seen ~40 years ago, $mus108^{A1}$ exhibits a dose-dependent response in sensitivity, and presents a moderate level of sensitivity based on what has been seen for other mus genes (Boyd 1976). For example, $mus101^{D1}$ is exhibits a moderate level of sensitivity with a relative survival of about 20% at 0.08% MMS (Boyd 1976).

After confirming that the $mus108^{A1}$ mutation was still present in the stock, we then used deficiency mapping to narrow the location of the gene within the fly genome. There were four deficiency stocks that Laurencon showed to disrupt $mus108^{A1}$, but only one deficiency from this list was chosen to use in the deficiency mapping process: $Df(1)JC70$. The other three deficiencies either did not have stocks available ($Df(1)ovoG6$ and $Df(1)ovoG7$) or were only found in stocks that contained genomic deletions and duplications simultaneously ($Df(1)A113$). We used $Df(1)JC70$ as our starting point and chose two additional deficiencies ($Df(1)ED6727$

and *Df(1)BSC823*) that overlapped the undefined breakpoints of *Df(1)JC70* to continue with the deficiency mapping.

The rationale for our mapping assay using these deficiencies is as follows: *Df(1)ED6727* deletes a section before *Df(1)JC70* begins up through the undefined start point of *Df(1)JC70* (Figure 3). If both *Df(1)JC70* and *Df(1)ED6727*, for example, had shown a similar decrease in relative survival that was significantly different from *Df(1)BSC823* – i.e. a failure to complement the *mus108^{A1}* allele – then we would have looked for a potential candidate gene within the deleted area that *Df(1)JC70* and *Df(1)ED6727* overlap. *Df(1)BSC823* is a smaller deficiency that covers the undefined end region of *Df(1)JC70,* and the same strategy can be applied to this deficiency and *Df(1)JC70*. Finally, it was possible that *mus108* is located in a region that is specific to only *Df(1)JC70*, in which case we would have expected to see decreased relative survival for this deficiency only. In particular, Laurencon saw ~50% survival with this deficiency, so we would have anticipated seeing similar values in our assay.

However, there was no significant difference in relative survival for any of the deficiencies as compared to the wild type (ANOVA; $p=0.9941$) (Figure 10). The average relative survival values for the deficiencies were 0.99, 1.00, and 0.95 for *Df(1)ED6727, Df(1)JC70,* and *Df(1)BSC823* respectively. This suggests that none of the deficiencies fail to complement *mus108^{A1}* (Figure 10).

There was a large standard deviation seen in the deficiency crosses (*Df(1)ED6727*: ±0.38; *Df(1)JC70*: ±0.41; *Df(1)BSC823*: ±0.53). It is possible that the large standard deviations are a result of a modifier segregating in the stock. A modifier is a genetic variant that can modify the overall phenotype of the primary variant or mutation. This could cause a large standard deviation if all the X chromosomes were not identical, with the level of sensitivity changing between

variants of the X chromosome. One way to check for this would be to restart the *mus108^{A1}* stock

from a single male, ensuring that all the X chromosomes in the stock are identical. Since we

would not know which males carry the modifier, several *mus108^{A1}* lines could be set up, each

starting from a single male. Mutagen sensitivity would be tested on each line and then compared.

If the sensitivities differ between each line, it could suggest that the stock is indeed segregating

certain sequences, whether that is *mus108^{A1}* or a modifier, which could possibly explain the large

standard deviation seen in each deficiency cross.

Another potential explanation for the lack of sensitivity seen in the deficiency crosses is

that since the *mus108^{A1}* chromosome was originally created by exposure to a mutagen, the

sensitivity we saw in the *mus108^{A1}*/*mus108^{A1}* flies may be due, in part, to homozygosity for an

additional mutation(s) on the chromosome.

The data collected in this experiment does not lend support for choosing one deficiency

over another for further study. Since Laurencon originally saw that *Df(1)JC70* failed to

complement *mus108^{A1},* we decided to pursue this deficiency as the target location to search for

potential candidate genes.

4.3 Proposal of a *mus108* candidate gene

The relative survival of the three deficiencies used to map *mus108^{A1}* had no significant

difference in average relative survival, so we looked at the genes covered by *Df(1)JC70*, the

deficiency that Laurencon originally saw to uncover the *mus108* mutation. Looking at the

molecular function of all the genes deleted by *Df(1)JC70,* we found three genes with a molecular

function involved in DNA processes (Table 2). Kruppel-like Factor 15 (*Klf15*) is involved in

sequence-specific DNA binding. Minichromosome maintenance 3 (*MCM3*) forms part of the

MCM2-7 hexamer, which aids specifically in 3' to 5' DNA helicase activity and DNA

replication origin binding. X-ray repair cross complementing protein 1 (*XRCC1*) is involved in

damaged DNA binding. Out of these three genes, we chose *XRCC1* as our candidate gene to

pursue mapping *mus108* because this gene is specifically involved in repair of DNA damage,

which is a key component of the mutagen sensitive classification.

4.4 Candidate gene: *XRCC1*

XRCC1 is a scaffold protein that aids in the recruitment and organization of various

repair enzymes to carry out the complex processes to repair damaged DNA (London 2015/2020,

Caldecott 2019). Scaffold proteins also help reduce the chance of toxic intermediates from being

released during the repair process and causing additional damage (London 2015). The main

repair pathway XRCC1 assists in is single strand break repair (SSBR), but it also supports

interactions in the base-excision repair (BER) pathway and specific ligation issues (London

2015/2020, Caldecott 2019). Notably, alkylation damage, caused by mutagens like MMS, is

repaired by the BER pathway, making *XRCC1* a stronger choice for our candidate gene (London

2015/2020).

The human XRCC1 ortholog is 633 amino acids long and contains several domains

dedicated for specific molecular interactions (Caldecott 2019). Analysis of the human ortholog

shows the domains include an N-terminal domain that binds to DNA polymerase Beta, a BRCT1

domain which binds to poly(ADP-ribose) polymerase and dsDNA, and a C-terminal domain

containing a BRCT2 domain that binds with Ligase 3 (London 2020/2015, Caldecott 2019). The

C-terminal region is not found in *Drosophila melanogaster* (London 2015). There are two linker sequences that sit between the domains. Linker 1 is ~150 amino acids long and contains a nuclear localization signal and a Rev1-interacting region (RIR) sequence that allows Rev-1 to bind and recruit additional polymerases (London 2015). Linker 2 is ~120 amino acids long and contains a motif that interacts with forkhead associated (FHA) domains used to recruit additional repair enzymes (London 2015). In humans, XRCC1 helps protect cells against damage caused by ionizing radiation, alkylation, and UV damage (London 2015). Several mammalian cell lines with mutations in XRCC1 have shown sensitivity to the mutagens MMS, CPT, HU, IR, and UV radiation, and ethyl methanesulfonate (EMS) (Caldecott 2019).

4.5 Alignment of XRCC1 Orthologs

The *XRCC1* gene sequence in the *mus108^{A1}* stock was compared to the reference gene sequence on FlyBase to identify any mutations that exist in the *mus108* candidate gene (Thurmond 2019). We identified five missense mutations, shown in Figure 12 and Table 4, and four silent mutations (data not shown). It was important to determine if any of the mutations were located in a highly conserved region of the gene. Mutations in a highly conserved area could be problematic to protein function because a region that is conserved across organisms indicates importance to the functionality of that protein. The first missense mutation in the *mus108 XRCC1* sequence changes the 127th amino acid from glutamine to lysine, which is the conserved amino acid seen in each of the orthologs examined (Figure 12, Table 4). The first three mutations, at amino acids 127, 195, and 196, all change the chemical properties of the amino acid (e.g. Polar/uncharged amino acid to a hydrophobic amino acid). The last two mutations, at amino acids 203 and 204, change the amino acids, but still retain similar amino

acid properties (i.e. Changing from one hydrophobic amino acid to a different hydrophobic amino acid).

An alignment of XRCC1 orthologs was created to visually assess regions of conservation in the gene. The mutation at amino acid 127 occurs in a conserved region of the N-terminal domain of XRCC1 (Figure 12).The remaining four missense mutations are in a less conserved region of the Linker 1 segment of the gene (Figure 12). None of the missense mutations stand out as being potentially problematic to the XRCC1 protein. The mutation at amino acid 127 would be promising if the amino acid were not changed to the conserved amino acid seen in the other orthologs. Even though amino acid changes at 195 and 196 change amino acid properties, they do not appear problematic because they are not located in a conserved region. However, it is possible that the cluster of mutations (two sets of adjacent mutations), which are only nine amino acids apart (195-204), could influence amino acid interactions and cause a problem with protein structure. Of course, this is just speculation and additional research would need to be conducted to explore this.

Amino Acid location in Drosophila	Amino acid in *XRCC1* reference sequence (amino acid property)	Amino acid in *mus108 XRCC1* sequence (amino acid property)
127	Glutamine (polar/uncharged)	Lysine (positive charge)
195	Glutamine (polar/uncharged)	Glutamic acid (negative charge)
196	Threonine (polar/uncharged)	Methionine (hydrophobic)
203	Isoleucine (hydrophobic)	Valine (hydrophobic)
204	Serine (polar/uncharged)	Threonine (polar/uncharged)

Table 4. Missense mutations in *mus108* candidate gene, *XRCC1*. Five missense mutations seen in *mus108* candidate gene when compared to wild-type reference sequence on FlyBase (Thurmond 2019).

4.6 RNAi Knockdown of *XRCC1*

An ideal experiment to determine if the candidate gene *XRCC1* is *mus108* would be to conduct a complementation test between alleles of both genes. If the alleles failed to complement each other, then that would suggest that *mus108* is *XRCC1*. However, there are no true mutants of *XRCC1* that exist currently (Thurmond 2019) so an RNAi line designed to knockdown *XRCC1* was used to determine if *XRCC1* mutants are mutagen sensitive. Specifically, this RNAi line was created as part of the Transgenic RNAi Project (TRiP) to create a genome wide collection of RNAi stocks that can allow researchers to learn more about gene function across all tissue types

within *Drosophila* (Perkins 2015). The *TRiP* stock chosen for this experiment contains a hairpin that was designed to knock down *XRCC1* under the control of the UAS promoter when expressed with a GAL4 line. The goal of this cross was to establish if *XRCC1* mutants present a mutagen sensitive phenotype after exposure to MMS. Offspring that contain both the *TRiP* and *GAL4* genotype together should knockdown *XRCC1*. If a reduced relative survival is seen, it may suggest *XRCC1* is mutagen sensitive.

The data for determining if a knockdown of *XRCC1* presents a mutagen sensitive phenotype is inconclusive (Figure 13). Looking at just the average relative survival of 1.16 ± 0.33, it suggests that a knockdown of *XRCC1* does not present a mutagen sensitive phenotype. However, it is unknown if *TRiP/GAL4* flies are actually knocking down *XRCC1*. The only way to confirm *XRCC1* knockdown is to perform a RT-qPCR and compare levels of *XRCC1* in those with *TRiP/GAL4* to their siblings (those without *TRiP* or *GAL4* or both). This is possible; however, the mutagen is applied at the larval stage, and the phenotypic markers used in this cross to differentiate the *TRiP* and *GAL4* flies are seen only at the adult stage. Using the adult flies for the RT-qPCR is possible, however this is not ideal because it is not representative of the stage that we are assessing survival at (larval stage). Instead, two new stocks would need to be generated that each contained a larval phenotypic marker on the appropriate chromosome (chromosome 2 for the *TRiP* stock, chromosome 3 for the *GAL4* stock) so that larvae with the *TRiP/GAL4* genotype can be separated from their siblings and then prepared for the RT-qPCR.

4.7 *mus108* Complementation Test (with transposon in *XRCC1*)

      An ideal experiment to determine if *XRCC1* is *mus108* would be to conduct a complementation test between alleles of the two genes. However, this is not currently possible since there are no true mutants of *XRCC1* that exist. Similarly to the RNAi lines designed as part of TRiP, a transposon tool kit was created as another option for gene knockout in *Drosophila melanogaster* (Thibault 2004). FlyBase listed an available *piggyBac* stock that contained a known transposon insertion, derived from Lepidoptera, in the 5' UTR of *XRCC1* (Thibault 2004, Thurmond 2019).

      We conducted a cross between *PBacXRCC1* and *mus108^{A1}* flies as a type of complementation test to determine if *XRCC1* is *mus108*. The average relative survival for the *PBacXRCC1* cross was 0.87 ±0.29 (Figure 14). On its own, the average relative survival suggests *PBacXRCC1/mus108^{A1}* flies exhibit a weak sensitivity to 0.08% MMS. However, only six vials were used for the cross due to difficulty obtaining balanced *PBacXRCC1* flies from the initial cross (Figure 6) and half of the data points lie above a relative survival value of 1, making it difficult to conclude sensitivity.

      This is a similar situation to the RNAi knockdown experiment, where it is unknown if the *piggyBac* transposon is actually disrupting *XRCC1*. *XRCC1* has not been studied in *Drosophila*, and this stock was created in addition to thousands of other *piggyBac* stocks (Thibault 2004, Thurmond 2019). The *PBacXRCC1* stock has not been confirmed to disrupt *XRCC1* gene expression but there are a couple ways to confirm this. The first option is to perform a RT-qPCR and compare levels of *XRCC1* expression in *PBacXRCC1* flies and wild-type flies. If *pBacXRCC1* disrupts gene expression, we would expect there to be a decrease in the amount of *XRCC1* RNA, meaning decreased levels of the protein, in the *piggyBac* flies. A second option is

to conduct a Northern blot to compare the size of the RNA transcript in the *PBacXRCC1* flies to wild-type flies.

4.8 In progress

Additional characterization of *mus108^A1* to other mutagens is currently in progress. The initial round of exposure to camptothecin and hydroxyurea failed due to the age of the food. We are confident the age of food initially used is the reason for this setback because Brood 1 of both assays were given older food and resulted in very few offspring. In the assay involving hydroxyurea as the mutagen, vial ten had newer food and resulted in many offspring compared to vial one of the same brood. Additionally, both Brood 2's resulted in more offspring than the corresponding Brood 1, which should not be the case. Due to time constraints the crosses have been re-started but data collection will be completed after submission of this thesis.

4.9 Future Directions

This work has provided several options for continuing research on *mus108^A1*. The first suggestion for future experiments involving *mus108^A1* would be to restart the stock from a single male fly to ensure that all of the X chromosomes are identical. If a modifier is segregating in the stock there could be multiple versions of the *mus108^A1* X chromosome, which could be skewing the data. Several *mus108^A1* lines would need to be started, each from a single male, and if sensitivities differ between each line it could suggest that the stock is segregating certain sequences, influencing the data obtained.

A second suggestion for experiments using *mus108^{A1}* is to conduct complementation analysis with other *mus* mutations on the X chromosome. It is assumed that previous researchers did complete a complementation analysis to ensure that *mus108^{A1}* is not an allele of another *mus* mutation, but there is no published complementation data for *mus108^{A1}*. This would be a useful step to confirm that *mus108^{A1}* represents a separate complementation group (or not).

A third suggestion is to confirm the RNAi knockdown of *XRCC1* (utilizing the GAL4/UAS system) and confirm the transposon disruption of *XRCC1* (*PBacXRCC1*). Confirmation of these constructs would allow future researchers to study *XRCC1* in greater detail.

A fourth suggestion for future experiments is to repeat the *PBacXRCC1* transposon cross since this project failed to acquire enough replicates for the mutagen sensitivity assay. For the initial cross to get *PBacXRCC1* balanced, more than three bottles should be set up in order to collect enough of the *PBacXRCC1/FM7* females for the second part of the cross.

Finally, future work should create mutant alleles of *XRCC1* to conduct complementation crosses to *mus108^{A1}* and measure mutagen sensitivity. If there were true alleles of *XRCC1*, then it would eliminate the guess work of using non-validated RNAi or *piggyBac* experiments. While this is the better option to conduct a complementation test, it is not a task that can be completed quickly or easily, especially in the time frame for this project.

4.10 Conclusion

Even though the pandemic disrupted our initial timeline, we conducted several experiments and reached several interesting conclusions. First, we determined that the *mus106^{D1}*

allele is no longer present in the fly stock 45 years after its initial discovery. We analyzed data from previous researchers and utilized the publicly-available fly genome sequence to determine a likely candidate gene for *mus106^{D1}* to be DNA ligase 4. We determined that the *mus108^{A1}* mutation is still present in the fly stock and analyzed previous research to determine a likely candidate gene for *mus108^{A1}* to be *XRCC1*. While we still need to validate the disruption of *XRCC1* in the transposon and RNAi lines, we performed initial work with *XRCC1*, a gene that has not been previously studied in *Drosophila melanogaster*.

References

Beranek DT. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. 1990. Mutation Res. 231(1): 11-30.

Boyd JB, Golino MD, Nguyen TD, Green MM. Isolation and characterization of X-linked mutants of *Drosophila melanogaster* which are sensitive to mutagens. 1976. Genet. 84(3): 485-506.

Boyd JB, Shaw KE. Postreplication Repair Defects in Mutants of *Drosophila melanogaster*. 1982. Mol Gen Genet. 186: 289-294.

Caldecott KW. XRCC1 protein; Form and function. 2019. DNA repair. 81(102664).

Chatterjee N, Walker GC. Mechanisms of DNA damage, repair and mutagenesis. 2017. Environ Mol Mutagen. 58(5) 235-263.

Cook KR, Christensen SJ, Deal JA, Cobum RA, Deal ME, Gresens JM, Kaufman TC, Cook KR. The generation of chromosomal deletions to provide extensive coverage and subdivision of the *Drosophila melanogaster* genome. 2012. Genome Bio. 13:R21.

Gatti M. Genetic control of chromosome breakage and rejoining in Drosophila melanogaster: spontaneous chromosome aberrations in X-linked mutants defective in DNA metabolism. 1979. Proc Natl Acad Sci U S A. 76(3):1377-1381.

Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK, Robertson HM, Engels WR. Type I Repressors of *P* Element Mobility. 1993. Gen Soc of Am. 135: 81-95.

Hakem R. DNA-Damage repair; the good, the bad, and the ugly. 2008. The EMBO Journal. 27: 589-605.

Hoch NC, Hanzlikova H, Rulten SL, Tétreault M, Komulainen E, Ju L, Hornyak P, Zeng Z, Gittens W, Rey SA, Staras K, Mancini GMS, McKinnon PJ, Wang Z, Wagner JD, Care4Rare Canada Consortium, Yoon G, Caldecott KW. *XRCC1* mutation is associated with PARP1 hyperactivation and cerebellar ataxia. 2017. Nature. 541: 87-91.

Kusano K, Johnson-Schlitz DM, Engels WR. Sterility of *Drosophila* with Mutations in the Bloom Syndrome Gene- Complementation by *Ku70*. 2001. Science. 291: 2600-2602.

Laurencon A. 2001. Personal communication to FlyBase. http://flybase.org/reports/FBrf0137505#reference

Lee T, Kang T. DNA Oxidation and Excision Repair Pathways. 2019. Int Jrnl Molec Sci. 20(23).

London RE. Strategies for coordinating and assembling a versatile DNA damage response. 2020. DNA Repair. 93(102917).

London RE. The structural basis of XRCC1-mediated DNA repair. 2015. DNA Rep. 30: 90-103.

Lundin C, North M, Erixon K, Walters K, Jenssen D, Goldman ASH, Helleday T. 2005. Nuc Acids Res. 33(12): 3799-3811.

Mei C, Lei L, Tan L, Xu X, He B, Luo C, Yin J, Li X, Zhang W, Zhou H, Liu Z. The role of single strand break repair pathways in cellular responses to Camptothecin induced DNA damage. 2020. Biomed & Pharmacotherapy. 125:109875, 1-14.

National Center for Biotechnology Information (NCBI). 1988. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/

Nicholas KB, Nicholas HB. GeneDoc: A tool for editing and annotating multiple sequence alignments. 1997. Embnet.news.

O'Driscoll M. Diseases Associated with Defective Responses to DNA Damage. 2012. Cold Spring Harb Perspect Biol. 4(12): a012773.

Oliveri DR, Harris PV, Boyd JB. X-Ray sensitivity and single-strand DNA break repair in mutagen-sensitive mutants of *Drosophila melanogaster*. 1989. Mut Res/DNA Rep. 235(1): 25-31.

Perkins LA, Holderbaum L, Tao R, Hu Y, Sopko R, McCall K, Yang-Zhou D, Flockhart I, Binari R, Shim HS, Miller A, Housden A, Foos M, Randkelv S, Kelley C, Namgyal P, Villalta C, Liu LP, Jiang X, Huan-Huan Q, Wang X, Fujiyama A, Toyoda A, Ayers K, Blum A, Czech B, Neumuller R, Yan D, Cavallaro A, Hibbard K, Hall D, Cooley L, Hannon GJ, Lehmann R, Parks A, Mohr SE, Ueda R, Kondo S, Ni JQ, Perrimon N. The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. Genetics. 2015 Nov;201(3):843-52

Reiter LT, Potocki L, Chien S, Gribskov M, Bier E. A Systematic Analysis of Human Disease-Associated Gene Sequences in *Drosophila melanogaster*. 2001. Gen Res. 11:1114-1125.

Sekelsky J. DNA Repair in *Drosophila*: Mutagens, Models, and Missing Genes. 2017. Genetics. 205(2): 471-490.

Sequencher® version 5.4.6 DNA sequence analysis software, Gene Codes Corporation, Ann Arbor, MI USA http://www.genecodes.com

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert
	M, Söding J, Thompson JD, Higgins DG**.** Fast, scalable generation of high-quality
	protein multiple sequence alignments using Clustal Omega. 2011. Mol. Syst. Biol. 7:539

Singh A, Xu Y. The cell killing mechanisms of Hydroxyurea. 2016. Genes. 7(99): 1-15.

Smith PD, Snyder RD, Dusenbery RL. Isolation and characterization of repair0deficient mutants
	of *Drosophila melanogaster*. 1980. Symposium on DNA Repair and Mutagenesis in
	Eukaryotes. 175-188.

Thibault S, Singer M, Miyazaki W, Milash B, Dompe N, Singh C, Buchholz R, Demsky M,
	Fawcett R, Francis-Lang H, Ryner L, Cheung L, Chong A, Erickson C, Fisher W, Greer
	K, Hartouni S, Howie E, Jakkula L, Joo D, Killpack K, Laufer A, Mazzotta J, Smith R,
	Stevens L, Stuber C, Tan L, Ventura R, Woo A, Zakrajsek I, Zhao L, Chen F, Swimmer
	C, Kopczynski C, Duyk G, Winberg M, Margolis J. A complementary transposon tool kit
	for *Drosophila melanogaster* using *P* and *piggyBac*. 2001. Nat Genet. 36: 283–287.

Thurmond J, Goodman JL, Strelets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB,
	Millburn G, Antonazzo G, Trovisco V, Kaufman TC, Calvi BR and the FlyBase
	Consortium. 2019. FlyBase 2.0: the next generation. Nucleic Acids Res. 47(D1) D759–
	D765.

Torgovnick A, Schumacher B. DNA repair mechanisms in caner development and therapy. 2015.
	Front in Gen. 6(157): 1-15.

Appendix

Symbol	Name	Molecular Function
Pp2C1	Protein phosphatase 2C	Protein serine/threonine phosphatase activity
Ctp	Cut up	Protein binding; Dynein light intermediate chain bindings
Pdha	Pyruvate dehydrogenase E1 alpha subunit	Pyruvate dehydrogenase activity
CG7024	n/a	Dehydrogenase activity
Proc-R	Proctolin receptor	Proctolin receptor activity; Neuropeptide receptor activity
CG15472	n/a	Unknown
Klf15	Kruppel-like Factor 15	Sequence-specific DNA binding; DNA binding transcription factor activity
lncRNA:CR45515	Long non-coding RNA:CR45515	Unknown
CG2871	n/a	Unknown
CG15471	n/a	Unknown
CG6978	n/a	Transmembrane transporter activity
CG2861	n/a	Unknown
lncRNA:CR45516	Long non-coding RNA:CR45516	Unknown
CG12682	n/a	Unknown

CG42594	n/a	Potassium ion leak channel activity
lncRNA:CR43495	n/a	Unknown
Boil	Boilerman	Unknown
CG6927	n/a	pH-gated chloride channel activity; Neurotransmitter receptor activity
CG32772	n/a	RNA polymerase II cis-regulatory region sequence specific DNA binding
CG4041	n/a	ATP binding; GTPase activator activity
CG6903	n/a	Heparan-alpha-glucosaminide N-acetyltransferase activity
HpRNA:CR46342	Hairpin RNA:CR46342	Unknown
CG44774	n/a	Unknown
Ptp4E	Protein tyrosine phosphatase 4E	Protein tyrosine phosphatase activity
lncRNA:CR44833	Long non-coding RNA:CR44833	Unknown
CG15468	n/a	Unknown
lncRNA:CR45792	n/a	Unknown
CG12680	n/a	Unknown
Ovo	n/a	RNA polymerase II cis-regulatory region sequence specific DNA binding

		RNA polymerase II cis-regulatory region sequence specific DNA binding
CG32767	n/a	
Rg	Rugose	Protein kinase binding; Protein kinase A binding
CG15465	n/a	Unknown
Rpn13R	Regulatory particle non-ATPase 13-related	Proteasome binding; Ubiquitin binding
CG5062	n/a	Unknown
lncRNA:CR44834	n/a	Unknown
CG42749	n/a	Unknown
CG3323	n/a	Unknown
CG17764	n/a	Unknown
Cdk7	Cyclin-dependent kinase 7	Serine/threonine kinase activity; ATP binding
snf	Sans fille	Protein of U1 & U2 snRNPs; Assemble spliceosome
TrxT	Thioredoxin T	Disulfide oxidoreductase activity
dhd	Deadhead	Disulfide oxidoreductase activity
CG4198	n/a	Unknown
CG15930	n/a	Unknown
Sas10	Something about silencing 10	Unknown
Rnp4F	RNA-binding protein 4F	mRNA binding

Mcm3	Minichromosome maintenance 3	3-5' DNA helicase activity
CG3309	n/a	Rab guanyl-nucleotide exchange factor activity
XRCC1	XRCC1	Damaged DNA binding
CanB	Calcineurin B	Calcium ion binding
SK	Small conductance calcium-activated potassium channel	Photoreceptor activity; Calmodulin binding
NAAT1	Nutrient amino acid transporter 1	Transmembrane transporter activity

Appendix A. List of all genes uncovered by the *Df(1)JC70* deletion. Genes are listed in the same order as they appear on FlyBase (Thurmond 2019).

Symbol	Name	Molecular Function
Up	Upheld	tropomyosin binding; calcium ion binding
lncRNA:CR44654	long non-coding RNA:CR44654	unknown
CG11178	n/a	unknown
Ndc80	Ndc80	protein binding
tth	Toothrin	unknown
Tango2	Transport and golgi organization 2	unknown
CG2691	n/a	unknown
NFAT	Nuclear factor of activated T-cells	DNA-binding transcription factor activity; chromatin binding activity; RNA polymerase II cis-regulatory region sequence-specific DNA binding activity
DNAlig4	DNA ligase 4	DNA ligase (ATP) activity; DNA metabolism (response to IR)
CG11164	n/a	RNA-DNA hybrid ribonuclease activity
CG15760	n/a	localize integral component of organelle membrane
CG11162	n/a	C-4 methylsterol oxidase activity
CG12177	n/a	unknown
CG11158	n/a	unknown
Nadsyn	NAD synthetase	NAD+ synthase activity; glutaminase activity

Nna1	Nna1 carboxypeptidase	metallocarboxypeptidase activity; zinc ion binding activity
CG9941	n/a	ubiquitin protein ligase activity; zinc ion binding
mus101	mutagen-sensitive 101	unknown
G	Garnet	cargo adaptor activity
Grip91	Gamma-tubulin ring protein 91	gamma-tubulin binding; microtubule minus-end binding
CG11134	n/a	methylthioribulose 1-phosphate dehydratase activity
CG11151	n/a	unknown
lncRNA:CR42861	long non-coding RNA:CR42861	unknown
Pdcd4	Programmed cell death 4	unknown
asRNA:CR43833	antisense RNA:CR43833	unknown
CG32625	n/a	unknown
Rtc1	Rtc1	endoribonuclease activity
Yp3	Yolk protein 3	carboxylic ester hydrolase activity
rdgB	retinal degeneration B	ion binding; lipid & phospholipid transporter activity
CtsB	Cathepsin B	cysteine-type endopeptidase activity

CG11103	n/a	unknown
inaE	inactivation no afterpotential E	lipoprotein lipase activity
CG10993	n/a	RNA binding
CG10996	n/a	carbohydrate binding; aldose 1-epimerase activity
CG11095	n/a	hydrolase activity
Clic	Chloride intracellular channel	glutathione peroxidase activity; chloride channel activity; lipid binding

Appendix B. List of all genes located between *up* and *garnet*. Genes are listed in the same order as they appear on FlyBase (Thurmond 2019).