UNIVERSITY OF CALIFORNIA SAN DIEGO

Investigating the use of DNA repair as a strategy to enhance production stability in Chinese Hamster Ovary (CHO) cell lines

A Thesis submitted in partial satisfaction of the requirements for the degree of Master of Science

in

Bioengineering

by

Jamie Lee

Committee in Charge:

Professor Nathaniel Lewis, Chair
Professor Christian Metallo, Co-Chair
Professor Peter Wang

2020

The Thesis of Jamie Lee is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

Co-Chair

_____

Chair

University of California San Diego

2020

DEDICATION

For their unconditional support and love, I would like to dedicate this thesis to my loving family, Maurice Lee, Martha Lee, and Deandra Lee. Thank you for believing in me and allowing me to grow into the person that I am today. To my friends, you know who you are, thank you for looking out and showing me what it means to be a true friend. You have all inspired me in many ways and pushed me to be a better version of myself.

EPIGRAPH

"Great minds discuss
ideas; average minds
discuss events; small
minds discuss people."

– Eleanor Roosevelt

"Our deepest fear is not that we are inadequate. Our deepest fear is that we are powerful beyond measure. It is our light, not out darkness that most frightens us…And as we let out own light shine, we unconsciously give other people permission to do the same. As we're liberated from our own fear, our presence automatically liberates others."

– Marianne Williamson

TABLE OF CONTENTS

Signature Page..... iii

Dedication iv

Epigraph v

Table of Contents..... vi

List of Abbreviations vii

List of Figures..... viii

List of Tables..... ix

Acknowledgements..... x

Abstract of the Thesis..... xiv

1 Introduction..... 1

2 Background..... 4

2.1 Double strand break repair..... 4
2.2. Genomic evidence of deficient genomic DNA repair 5
2.3 Targeting DSB repair for increase product stability..... 8

3 Materials & Methods..... 9

3.1 Construction of DNA repair gene plasmids..... 9
3.2. Genomic evidence of deficient genomic DNA repair 13
3.3 Targeting DSB repair for increase product stability..... 16
3.4 Double Strand Break (DSB) repair assay..... 17
3.5 Long term study..... 18

4 Results & Discussion..... 19

5 Conclusion..... 31

Appendix..... 33

References..... 35

LIST OF ABBREVIATIONS

Escherichia Coli- E.Coli

Chinese Hamster Ovary- CHO

Insertion & deletions- InDels

Short nucleotide polymorphism- SNP

Double strand break- DSB

Secreted alkaline phosphate- SEAP

DNA methyltransferase- Dnmt3b

Human cytomegalovirus- hCMV-MIE

Nonhomologous end joining- NHEJ

Homologous Recombination- HR

DNA dependent protein kinase catalytic subunit- DNA-PKcs

DNA protein kinase- DNA-PK

Loss of heterozygosity- LOH

Ethanol- EtOH

Fluorescence-activated cell sorting- FACS

LIST OF FIGURES

Figure 1: Double Strand Break (DSB) repair pathway..... 5

Figure 2: Identification of mutant DNA repair genes..... 7

Figure 3: Double-strand break (DSB) reporter system..... 9

Figure 4: Construction of unmutated DSB repair gene in recombinant lentiviral plasmids..... 22

Figure 5: Integration of plasmid in CHO-SEAP 25

Figure 6: Quantitative PCR (qPCR) comparison of Wrn cell lines 26

Figure 7: Quantitative PCR (qPCR) comparison of Rad1 cell lines 27

Figure 8: DSB repair efficiency and SEAP titer measurements of engineered CHO cell lines..... 30

Supplementary Figure 1: Rad1 colony PCR..... 33

Supplementary Figure 2: Blasticidin kill curve used to determine ideal concentration for selection..... 33

Supplementary Figure 3: Bleocin kill curve used to determine ideal concentration for selection..... 34

LIST OF TABLES

Table 1: PCR program parameters …… 11

Table 1.1: PCR program distinction between repair genes ……..…. 11

Table 2: Engineered CHO-SEAP cell lines ……...…… 15

Supplementary Table 1: Primer list…… 33

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Nathan Lewis for bringing me into his lab and being a great advisor and mentor. He allowed me the opportunity and the resources to investigate this project and for that I am grateful.

I would also like to acknowledge Dr. Philipp Spahn of the Lewis lab, without whom my research would have been incredibly more difficult to execute. I am especially thankful for his incredible mentorship and patience with me in the past year. Lastly, I would like to thank him for allowing me to use some of his figures for the background chapter of this thesis.

Overall, a great thanks to the members of the Lewis lab for their interest in all things CHO cells. Without which this project would most likely not have been conducted.

ABSTRACT OF THE THESIS

Investigating the use of DNA repair as a strategy to enhance production stability in Chinese Hamster Ovary (CHO) cell lines

by

Jamie Lee

Master of Science in Bioengineering

University of California San Diego, 2020

Professor Nathaniel Lewis, Chair
Professor Christian Metallo, Co-Chair

Chinese hamster ovary (CHO) cells are the preferred and most widely used cell line for biopharmaceutical companies in terms of manufacturing mass amounts of recombinant therapeutic proteins. This is due to CHO cells having various characteristics that are vital to the production of therapeutic proteins that sets it apart from other cell lines. One advantage to CHO cells is their increased ability to adapt. This adaptability, although beneficial, has its drawbacks, as it is owed to the inherent genetic instability of CHO cells. It has been hypothesized that the genetic instability of CHO cells comes from the cells inability to

efficiently fix double strand breaks (DSBs) due to various SNPs in DSB related genes. In this study, four different CHO secreting alkaline phosphatase (SEAP) cell lines expressing an unmutated DSB related gene (Wrn, Wrn +Xrcc6, Rad1, or Rad1 + Xrcc6) was compared to CHO-SEAP WT. The cell line containing solely Wrn performed similarly to WT in terms of DSB repair, while the Rad1 cell line repaired DSBs more inefficiently. On the other hand, it was observed that cell lines expressing double DSB genes had opposite effects, as DSB repair worsened in Wrn + Xrcc6 and improved in Rad1+Xrcc6. This observation was further validated when comparing the SEAP titer of these cell line cultures over a period of about 8 weeks. The cell lines expressing Xrcc6 had lower SEAP titer per cell. However, once the cells were irradiated, these cell lines outperformed the cell lines containing only one DNA repair gene (Wrn or Rad1). This suggests the vital role that Xrcc6 plays in the CHO DSB repair system.

1 Introduction

Recombinant proteins have been essential to advancing the biomedical and pharmaceutical industry. Their use as drugs to treat various major human diseases such as diabetes and certain cancers, has made them incredibly competitive in terms of therapeutics (Pham 2018). The process of generating recombinant proteins is done by cloning an isolated gene into an expression vector. Most commonly, they are expressed in organisms, such as *Escherichia Coli* (*E.coli*), yeast, and cultivated mammalian cells (Clark and Pazdernik 2016). Throughout the years, cultivated mammalian cells have become the dominant host expression system, with Chinese Hamster Ovary (CHO) cells being the preferred cell line for mass production of recombinant therapeutic proteins. This is due to their superior ability to not only facilitate human compatible post translational modifications, but also to assemble and fold complex polypeptides (Wurm 2004)); (Jia et al. 2018). Moreover, CHO cells have excellent manufacturing adaptability and plasticity, in terms of the ability to conform to many culture conditions (Jayakpal et al.2007).

Paramount in the overall productivity of a cell line is its ability to maintain production stability, where production stability refers to the cell lines ability to maintain desired titer and quality over multiple generations. While CHO cells have contributed to more than 70% of recombinant protein production originating from mammalian cells, their exceptional adaptability is simultaneously, and perhaps ironically, a result of a major disadvantage of CHO cells in general: the inherent genetic instability of CHO cells (Vcelar et al. 2018). This results in an expression system which gradually loses production stability as the cells are continuously cultured. This complication is a major concern for many pharmaceutical companies as it can render the cell line useless, increase production cost, and as a result lead to increases in drug prices. Although much research and efforts have been made to increase production stability and investigate the molecular mechanisms that contribute to this

phenomenon there still isn't a clear understanding. Many studies indicate that production instability is linked to two primary mechanisms: (i) the decline of transgene copy numbers in the genome (Kim et al. 2011; Nematpour et al. 2017) and (ii) the transcriptional repression of transgenes via epigenetic events such as DNA methylation and histone modification (Kim et al. 2011). Other studies have stated that production instability is owed to production bottlenecks (Reinhart et al. 2014) or the generation of high amounts of unwanted metabolites (Altamirano et al. 2013).

Productivity is impacted by genome stability. Although the adaptability of CHO cells to various genetic mutations is highly efficient, it results in the generation of widespread mutations in the genome (Dahodwala and Lee 2019). Some of these mutations include: insertions and deletions (InDels), short nucleotide polymorphisms (SnPs), and chromosomal rearrangements, such as translocations, inversions, etc. The accumulation of these various types of mutations are a main cause of genomic instability in CHO cells, with a large fraction of that instability owing itself to chromosomal aberrations; This is mostly due to the nature of chromosomal aberrations, as it randomly integrates itself to different sites that may struggle to support gene expression (Li et al. 2016). It has been suggested that the formation of chromosomal aberrations are largely a result of improper repair of double strand breaks (DSBs), a type of mutation that is formed by ionizing radiation, collapsed DNA forks, and other factors (Varga and Aplan 2005). Due to the genomes propensity for mutations, Eukaryotes are luckily equipped with a DNA repair system that operates to minimize the loss of genetic information (López-Camarillo et al. 2009). However, CHO cells are insufficiently repairing these DSBs due to mutations in key repair genes, resulting in transgene loss.

With the current understanding of the underlying issues that cause production instability, many strategies have been put forth to find alternative ways to boost and or

maintain production stability. Some focus on predicting CHO cell instability, improving protein production in general, or trying to avoid transgene loss altogether. To reduce production instability some studies have shown that certain gene knock-outs can improve transgene expression or increase product titer (Lee et al. 2013); (Matsuyama et al. 2017). For instance, in a study done by Yan et al. CRISPR/Cas9 genome editing was implemented to modify a CHO cell line by performing a gene knockout of DNA Methyltransferase (Dnmt3b) gene, improving the cell line stability as well as transgene expression. Moreover, other studies have indicated that by inducing a mutation in the transcription factor start site (C-179) within human cytomegalovirus (hCMV-MIE), the cells propensity to transgene loss was reduced production of recombinant proteins stabilized (Moritz, Becker, and Göpfert 2015). Lastly, other studies have focused on generating models to enhance the cell line productivity, through testing a variety of factors such as culture temperature and treatments to the culture itself (Xu et al. 2019); (Kuo et al. 2018). However, the ability of these experiments to hold over long-term culture is still undefined or does not actually fix the underlying problem.

The objective of this study is to implement a novel strategy to increase production stability of CHO cells through the restoration of their innate DNA repair capabilities. Using genomic evidence, several CHO DNA repair genes that contain mutations that detrimentally affect protein function were identified. Thereafter, the functionality of the genes that are involved in DSB were restored through knock-in of original Chinese hamster cDNA and tested with a newly implemented DSB reporter system which demonstrated that CHO-SEAP cell lines expressing single (Wrn or Rad1) or double DSB genes (Wrn +Xrcc6 or Rad1 + Xrcc6) performed similarly or worse than the WT. However, when this observation was further analyzed by comparing several SEAP titer measurements over a period of 8 weeks, the cell lines expressing double DSB genes had higher titer than the cell lines expressing single DSB repair genes, suggesting that Xrcc6 plays a key role in the DNA repair pathway.

2 Background

2.1 Double strand break repair

DNA damage often occurs in the genome with numerous causes. A class of DNA damage or mutation, known as double strand breaks (DSB) are formed when the two strands in the DNA double helix are severed (Jackson 2002). They can be induced by several different external DNA-damaging agents. Some of these external agents include ionizing radiation, dysfunctional replication fork processing, or telomere deprotection (Scully et al. 2019). The insufficient repair of DSBs can cause chromosomal translocations, loss of genetic information, and even apoptosis if left unrepaired (Jeggo and Löbrich 2007). Although there are several mechanisms to repair DSBs, the repair pathway that is chosen is dependent on the cellular context of the damage, a strategy the cell has developed to minimize the cell's DNA damage and efficiently repair the lesion (Scully et al. 2019).

Two major DSB repair pathways are nonhomologous end joining (NHEJ) and homologous recombination (HR), shown in *Figure 1*. The two pathways differ in the way that they repair DSBs, mostly in terms of homology. While NHEJ repairs by modifying the broken ends and ligating them together regardless of homology, HR repairs the break by utilizing an undamaged DNA template to exchange nucleotides and reconstruct the damaged sequence (Mao et al. 2008). There are multiple genes that play a role in the two distinct DNA repair pathways. For instance, in NHEJ the Ku heterodimer, made up of Xrrcc6/Xrrcc5 (Ku70/Ku80) binds to the damaged DNA ends and recruits DNA dependent protein kinase catalytic subunit (DNA-PKcs) to make up activated DNA dependent protein kinase (DNA-PK) complex. With the help of ARTEMIS the ends are processed and either Pol μ or λ fills the ends if necessary. Lastly, ligation is initiated by recruitment of DNA ligase IV, Xrcc4, and XLF to the processed ends (Sharma and Raghavan 2016). Although these core components are essential for NHEJ to succeed, there are many other factors that also contribute to this

process. Therefore, if genes that are required for this process to function efficiently are riddled with mutations, it could drastically lower the rate of repair and/or cause genetic instability.

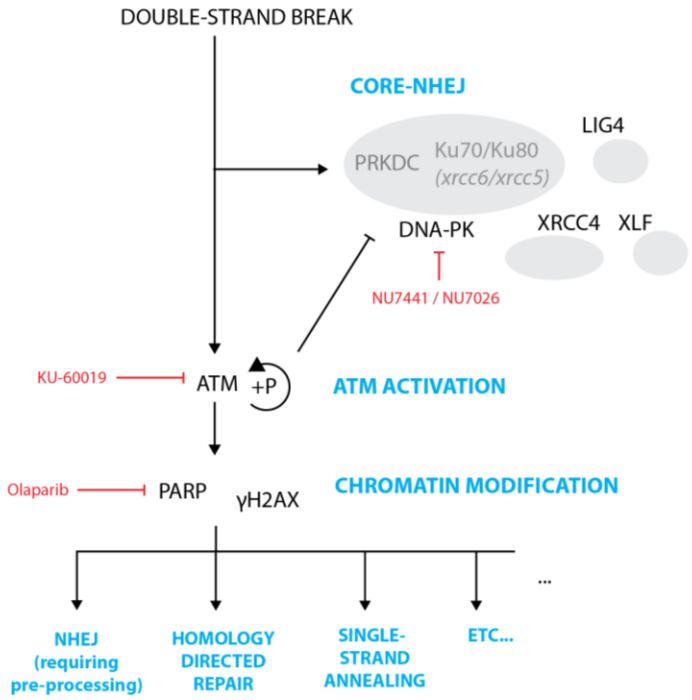


Figure 1: Double Strand Break (DSB) repair pathway. There are several ways to repair DSBs. Depending on the severity of the break the cell will choose which path works best for them. Used with permission from Philipp Spahn.

2.2 Genomic evidence of deficient genomic DNA repair

To combat production stability pending from DSBs, this experiment will utilize DNA repair as an approach to scale back transgene loss, and in turn suppress the resulting production instability of CHO. To do so, DNA repair genes in CHO that contain highly detrimental mutations were identified by Drs. Shangzhong and Hooman Hefzi in the Lewis lab (Spahn et al. in preparation). The process to find these genes was done by aligning of whole-genome sequencing data from 11 CHO cell lines to the recent Chinese hamster genome assembly (Rupp et al. 2018). From there, the data were analyzed and nonsynonymous SNPs were found in genes that are linked to DNA repair (Fig. 2). Sequencing analysis revealed 157 SNPs in DNA repair genes spanning 11 major CHO cell lines and 14 ontology categories associated with DNA repair (Fig 2A). It was found that, of the 157 SNPs 62 of them

demonstrated a loss of heterozygosity (LOH) (Fig 2B) and 19 were classified as detrimental (Fig 2B, dashed line). The number of SNPs were further narrowed down the by filtering for SNPs that have gone through LOH and are also associated with DSB repair. Of the remaining variant, only 10 were classified detrimental according to PROVEAN. The deleteriousness of the SNP on protein function was predicted by its PROVEAN score (Choi et al. 2012), allowing for ranking and quantification of these mutations in CHO DNA repair genes.

Based on the variant analysis data, we identified a variety of gene targets that, if repaired, had the possibility to improve the DSB repair and essentially increase production stability. With that said, we chose to test our strategy by repairing three DSB-related repair genes (Rad1, Wrn, and Xrcc6) and potentially restore their DSB repair capabilities. Based on SNP analysis Rad1 checkpoint DNA exonuclease (Rad1) was shown to be positive in 11 cell lines and contained a PROVEAN score of -6.383, one of the highest scoring genes in terms of deleteriousness. Rad1 is a protein coding gene that encodes the 9-1-1 cell cycle checkpoint response complex, a heterotrimeric cell cycle checkpoint complex that plays a major role in DNA repair (Wang et al. 2004). SNP analysis also revealed that Werner syndrome RecQ like helicase (Wrn) was also positive in all 11 cell lines and contained two mutations Wrn V1096A and Wrn R879Q with PROVEAN scores of -3.654 and -2.478, respectively. Wrn is a gene that encodes a subfamily of DNA helicase proteins, RecQ. It plays a role in the DSB repair pathway by physically interacting with the ku heterodimer and DNA-PKcs. Wrn is then phosphorylated by DNA-PKcs, stimulating Wrn enzymatic activity to aid in processing the ends of double strand breaks prior to ligation (Yannone et al. 2001). Lastly, X-ray repair cross complementing 6 (Xrcc6) is a protein coding gene that makes up part of the ku heterodimer/complex previously mentioned. This ku complex directly interacts with the DSB by binding to the ends of the DSB to allow for NHEJ (Wang et al. 2013).

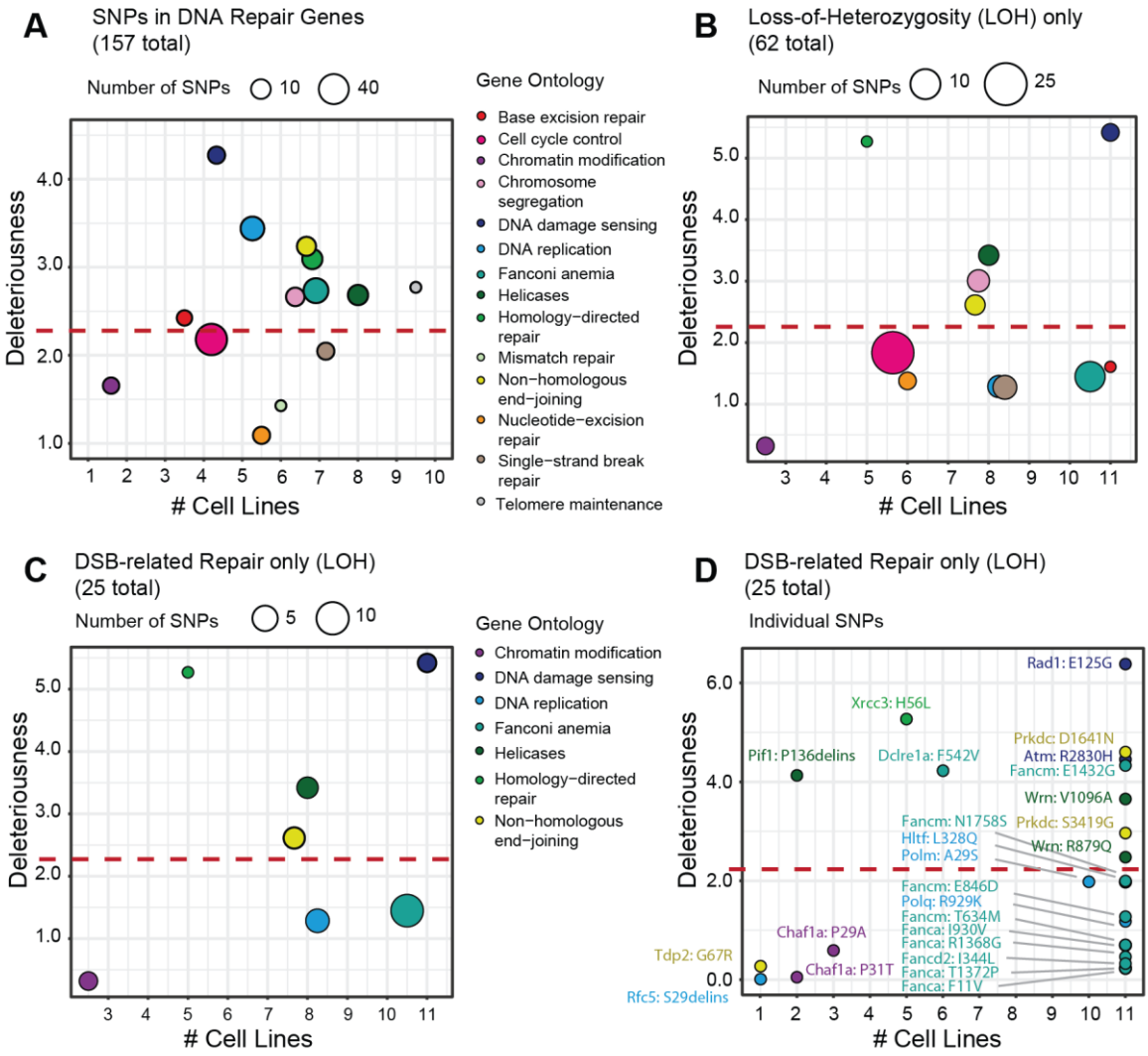


Figure 2: Identification of mutant DNA repair genes. A) 157 SNPs linked to a wide range of DNA repair categories were identified from whole-genome sequencing data analysis of 11 major CHO cell lines. In each category, the number of CHO lines affected (x-axis) and SNP deleteriousness (y-axis: Negative PROVEAN score) was averaged for all mutations. The Dashed line represents the recommended threshold to distinguish detrimental SNPs from neutral SNPs (Choi et al. 2012). **B):** SNPs that have undergone loss of heterozygosity (LOH). **C)** SNPs that have undergone LOH in genes linked to double-strand break (DSB) repair. **D)** view of Data from (C) with individual SNPs. Used with permission from Philipp Spahn.

2.3 Targeting DSB repair for increase product stability

It has been established that transgene loss is largely due to the chromosomal instability of CHO cells. Therefore, we predict that by restoring the DNA repair capabilities of genes involved in DSB repair we can enhance production stability and increase protein titer. With the restoration of key genes identified from the previous SNP analysis, a DSB reporter system (based on the EJ5-GFP tool provided in (Bennardo et al. 2008) was implemented in the CHO-SEAP cell lines (Hayduk and Lee, 2005) that include the restored DSB repair genes. The EJF-GFP assay (DSB-repair assay) works by inducing two DSBs via two sgRNAs at the 5' and 3' end of a 2kb spacer that sits between a promoter and a GFP reading frame (Fig 3A & B). The idea is that if the cell line can successfully restore the DSB, then GFP expression would stay negative. In turn, if the cell line is unable to fix the DSB, the spacer would be lost and it would bring the GFP closer to the promoter establishing a positive GFP expression. Overall, this assay provides a quantitative analysis of how well the cells are able to repair DSB after restoration of mutant DSB repair genes.

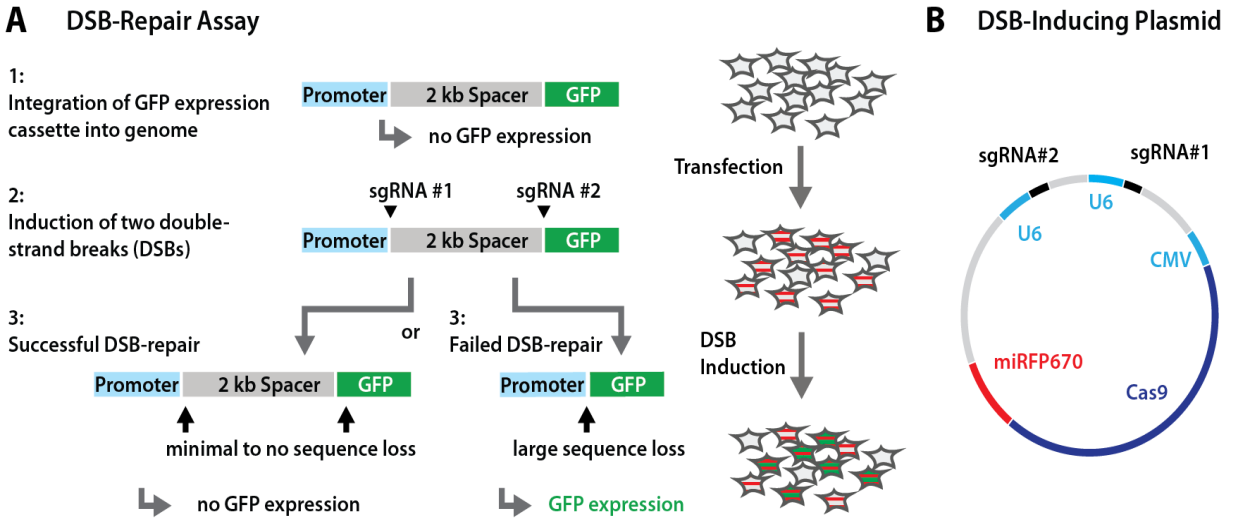


Figure 3: Double-strand break (DSB) reporter system. A) Step 1: Integration of GFP expression cassette, comprising of a promoter, a 2kb spacer, and a GFP reading frame, into genome of target cell line. Expression of GFP from the promoter is blocked by the spacer. Step 2: Transfection of DSB-inducing plasmid (B) into the cell line and induces two DSBs at the 5' and 3' ends of the spacer via two sgRNAs. Cells that have been transfected successfully is illuminated by the far-red fluorescence from the miRFP670 that is fused to Cas9 (B). Step 3: Transfected cells that can repair both DSBs efficiently will remain keep GFP-negative, due to the spacer being held in place. Meanwhile, transfected cells that are unable to repair the DSBs efficiently will lose the spacer and the promoter will move closer to the GFP and drive GFP. Assay modified after (Bennardo and Stark, 2010). **B)** DSB-inducing plasmid used in assay. Holds two sgRNAs that target the 2kb spacer ends, and a far-red fluorescence protein (miRFP670) fused to a Cas9 reading frame. Used with permission from Philipp Spahn.

3 Materials & Methods

3.1 Construction of DNA repair gene plasmids

Preparation of lentiviral backbone

Plasmids pLJM1-EGFP (addgene plasmid #19319) and lentiCas9-Blast (addgene plasmid #52962) were obtained from Addgene. Lentiviral backbone pLJM1-EGFP with corresponding selection markers (bleomycin or blasticidin) were made in preparation of lentiviral particles for transduction. First, double restriction digests of pLJM1-EGFP were done using restriction enzymes BamHI & KpnI to isolate original selection marker—puromycin. Puromycin was replaced through ligation of selection markers blasticidin or bleomycin from PCR of lentiCas9-Blast plasmid.

In preparation for cDNA knock-in of working DNA repair gene, a double digest of

pLJM1.EGFP.Blasticidin and pLJM1.EGFP.Bleomycin was conducted using restriction enzymes NheI-HF & BstBI, following New England Biolabs (NEB) protocols. The digested samples were run on a 1% Agarose gel and Lentiviral Backbone was isolated from EGFP by gel extraction. Samples were then purified following *Promega Wizard purification kit* protocol.

cDNA Amplification and Purification

Total cDNA from primary Chinese hamster lung fibroblasts was prepared following *PureLink PCR purification kit* (Invitrogen), and single cDNAs Rad1(XM_007626026) and Wrn (XM_007612274) was amplified via PCR for ligation into digested lentiviral backbones pLJM1.Bleomycin (pLJM1.Bleo) or pLJM1. Blasticidin (pLJM1.Blast). PCR amplification was done using two primers designed specifically for Rad1 or Wrn cDNA with included overhangs for digestion with NheI-HF & BstBI. Utilizing *Q5 High-Fidelity DNA Polymerase kit* (New England Biolabs), master mixes were made with the following reaction concentrations: 1X of 5X Q5 Reaction Buffer, 200 µM of 10 mM dNTPs, 1 µM of 10 µM primer pairs (Supplementary Table 1) Wrn_Nhe_F2/Wrn_Bst_R2 or Rad1_Nhe_F/Rad1_Bstb1_R, ~150 ng of cDNA, 0.02 U/µl of Q5 DNA polymerase, and nuclease-free H20. Master mix was made for a total of eight reactions for Rad1 and six reactions for Wrn, each with a total of 25 µl. The PCR reactions were placed into a thermocycler and adhered to the PCR parameters shown in *Table 1*. The PCR samples were then run on a 1% agarose gel at 100V until the bands reached 3/4 of the way down the gel. All Wrn PCR samples were run on an agarose gel, then gel extracted and purified following *Wizard SV Gel and PCR Clean-Up system* (Promega). Only two of the Rad1 PCR reactions were run on a gel for verification of amplification, while the other samples were purified following *PureLink PCR purification kit* (Invitrogen).

Table 1: PCR Program Parameters

Step	PCR Temp. (°C)	Time	Cycle
Initial Denaturation	98	2 min	1x
1. Denaturation	98	10 sec	35x
2. Annealing	Table 1.1	20 sec	
3. Extension	72	Table 1.1	
Final Extension	72	2 min	1x
Storage	4	∞	_

Table 1.1: Program Distinction Between Repair Genes

Target cDNA/gene	Annealing Temp. (°C)	Extension Time
Rad1	60	30 sec
Wrn	62-70	2 min

Restriction Digest and Ligation into pLJM1.Blasticidin/Bleomycin Plasmid

Isolated Rad1 and Wrn PCR product (inserts) were digested with corresponding restriction enzymes NheI-HF & BstBI following New England Biolab's RE digestion protocol, for ligation into previously digested lentiviral plasmid. To do this, 50 µl reactions were created with the following reagents: 300 ng of insert DNA (Rad1 or Wrn), 1X of 10X cutsmart buffer, 20 U/µl of restriction endonuclease NheI-HF, and nuclease free water. The digests were then incubated in a 37°C heat block for 20 min. Following the first incubation, 20 U/µl of restriction endonuclease BstBI was added and incubated at 65°C for 20 min. Lastly, digests were purified using *PureLink PCR purification kit* (Invitrogen). Note that only 35 µl was used during the elution step.

Ligation of digested and purified Rad1 and Wrn inserts into their respective lentiviral backbones pLJM1.Bleo and pLJM1.Blast was performed following the *Quick Ligation* protocol (New England Biolabs). Vector concentration and insert concentration were both adjusted accordingly to obtain a 3 (insert): 1 (vector) ratio. For each insert sample (Rad1 or Wrn), three samples were created; one with the insert (+), one with no insert (-), and one

without ligase (--). Samples were prepared with the following components: 21 ng/µl Rad1 or 2.7 ng/µl Wrn insert, 50 ng/µl pLJM1.Bleo or 20 ng/µl pLJM1.Blast lentiviral vector, distilled H_20, 1X of 2X Quick Ligase Buffer, and 1 µl Quick Ligase. Samples were then put on ice for 5min.

Transformation of E.coli

Transformation of Rad1 and Wrn ligation samples (+, -, --) into *E.coli* competent cells was performed following *High Efficiency Transformation for NEB Stable Competent E.coli* (C3040H) protocol (New England Biolabs), to amplify the samples. The three samples were transformed by gently pipetting 2 µl of chilled Rad1 and chilled Wrn ligation samples into three separate premade tubes containing 50 µl of thawed *E.coli* (C3040H) cells. Samples were gently mixed and immediately put on ice for a 30 min incubation period. Following incubation, samples were heat shocked in a 42ºC water bath for 30 sec, then placed in ice for a 5 min incubation period. Next, 950 µl of room temperature SOC media was added to each sample and placed in a 37ºC shaker for 60 min at 220 rpm. Lastly, 100 µl of each sample was plated onto warmed Luria Broth (LB) plates with 100 µl/mL Ampicillin (Amp) and placed in an incubator at 30ºC for 24 hr incubation.

Selection & Screening (Colony PCR)

After colony formation, colonies were picked from positive Rad1 and Wrn plates and resuspended in 50 µl of warmed LB (100 µl/mL Amp) media. A total of 13 (Rad1) and 3 (Wrn) colonies were picked and used for colony PCR to confirm successful ligation of insert into lentiviral vector pLJM1.Bleo or pLJM1.Blast. Reactions were made following the same protocol as initial PCR and adhered to the same program as *Table 1*, with an annealing temperature of 60-66ºC for Wrn. The colony PCR for Wrn and Rad1 also used the same primers (Supplementary Table 1). Colony PCR products were run on a 1% agarose gel and analyzed. After validating success of ligation, positive resuspended colonies were inoculated

into 1 mL of warmed LB (100 µl/mL Amp) media and placed in a 30ºC shaker overnight at 220 rpm. Next, 50 µl of preculture was used to inoculate a 50 mL culture, which was also placed in a 30ºC shaker for 24 hrs at 220 rpm.

DNA isolation and purification/Restriction digest

Following *GenElute HP Plasmid Midiprep Kit* (Sigma Aldrich) pLJM1.Rad1.Bleo and pLJM1.Wrn.Blast plasmids were harvested from the 50 mL *E.coli* culture. DNA was eluted with 500 µl of elution buffer as opposed to 1 mL elution buffer. To validate the successful ligation of inserts Rad1 and Wrn into their respective pLJM1 plasmid, a double restriction digest was done using restriction enzymes NheI-HF & BstB1 following the previously mentioned method. The undigested plasmid and digested sample were then run on a 1% agarose gel and analyzed.

Sequencing

For more validation of the engineered plasmids, Sanger sequencing was done using multiple primers. Primers were designed to target Rad1 and Wrn insert DNA as well as their selection markers; primers are listed in (Supplementary Table 1). Multiple reaction samples were prepared, each with about 20 ng of Wrn or Rad1 template (10 µl) and 5 µM of forward or reverse primer (5 µl) to make a total of 15 µl per reaction.

3.2 Cell Line Generation

Lentiviral Generation

Hek293T cells were infected with lentiviral plasmids for lentiviral generation. HEK293T cells were seeded at a density of $2.5x10^5$ cells/well in pre-warmed DMEM (+10%FBS, +1% pen/strep) in each well of a 6-well plate and placed in the incubator at 37°C. Once the cells reached 70-90% confluency the cells were infected with either lentiviral plasmid pLJM1.Rad.Bleo or pLJM1.Wrn.Blast following the protocol for Lipofectamine LTX transfection. To do this, a master transfection mix was composed so that each well contained

each of the following components: 2.5 µg of DNA, 0.83 µg of psPAX2 (Packaging plasmid), 0.83 µg of pMD2.g (Envelope plasmid), 0.83 µg of pLJM1.Rad.Bleo or pLJM1.Wrn.Blast (gene of interest) or pLJM1.EGFP (positive control) plasmid, and 500 µL opti-MEM media. Next, 8.75 µl per well of Lipofectamine LTX was mixed into the master mix and incubated for 25 min. DMEM media from the seeded 6 well plate was aspirated and carefully washed with PBS and replaced with 1.5 mL of prewarmed opti-MEM. Next, 500 µl of transfection mix was added (drop wise) to each well after incubation time and mixed in by gently swirling the plate. The plate was then placed into an incubator overnight at 37°C.

The next day, the positive control was checked for green fluorescence of GFP with a microscope. After verification of glowing positive control, the media was replaced with full DMEM and placed into an incubator at 37°C for the next day. Following incubation, the medium containing the viral particles was collected into a 15 mL centrifuge tube and pelleted for 4 minutes at 2000xg (4°C). The supernatant was filtered through a 0.45 µm syringe filter into 1.5 microcentrifuge tubes in 1 mL aliquots and stored in 4°C (stable for up to a week) or 80°C for long-term storage.

Transduction of CHO-SEAP cells

Harvested lentivirus of Rad1, Wrn, and EGFP (positive control) was transduced into CHO-SEAP WT and CHO-SEAP Xrcc6 cells to generate our desired cell lines, listed in *Table 2*. First, CHO-SEAP cells were seeded at ~100k cells/well in 6-well plates. Polybrene was added to the harvested lentiviral media to get a final concentration of 8 µg/mL, then the viral medium was added to the seeded wells after the cells reached 50-70% confluency using the following set up : 1:1 1 mL of IMDM medium (10% Fetal Bovine Serum, 1% pen/strep, 7 µg/mL Puromycin, 300µg/mL Hydromycin) + 1 mL virus, 1:10 1.9 mL IMDM medium + 100 µl virus, negative control 2 mL IMDM medium + no virus. The plate was put to incubate 37°C for 36hrs.

During the incubation period, the plate was put onto a microscope to verify GFP fluorescence of the positive control. After incubation, the viral particle-containing media was removed, and the cells were split such that the confluency was ~30% the next day and placed back into the incubator for overnight incubation. Selection toxin (75ng/mL Bleomycin or 10ng/mL Blasticidin) was added to the culture medium based on previous toxin kill curve (Supplementary figure 2 and 3) to select for cells that stably integrated the lentivirus. Cells were split every other day and toxin was added to fresh IMDM media until control (no virus) morphology appeared widespread and detached. Surviving cells of Rad1 and Wrn were expanded into a regular T75 flask with regular IMDM media (no toxin). During expansion, cells were split 1:2 every other day and selection toxin was re-added to media after first passage. Some cells were then reserved, frozen down, and stored in the liquid nitrogen storage container for long-term storage.

Table 2: Engineered CHO-SEAP cell lines

Cell line	Description	Selection
CHO-SEAP WT LN#66	WT version of Chinese hamster ovary cells	
CHO-SEAP Wrn	CHO-SEAP WT (LN#66), transduced with pLJM1.Wrn.Blast (PL#79)	Blasticidin
CHO-SEAP Wrn + Xrcc6	CHO-SEAP Xrcc6 (LN#70), transduced with pLJM1.Wrn.Blast (PL#79)	Blasticidin
CHO-SEAP Rad1	CHO-SEAP WT (LN#66), transduced with pLJM1.Rad1.Bleo (PL#56)	Bleocin
CHO-SEAP Rad1 + Xrcc6	CHO-SEAP Xrcc6 (LN#70), transduced with pLJM1.Rad1.Bleo (PL#56)	Bleocin

3.3 Validation of engineered cell line

RNA extraction and cDNA synthesis

RNA extraction was also done on the manufactured cell lines in preparation for cDNA synthesis. Cells were harvested as a pellet and RNA extraction was performed following the *Qiagen's Quick -Start Protocol RNeasy Mini Kit*. Optional DNase digestion was performed following steps 1-4 of "On Column DNase digestion" in the *Quick -Start Protocol RNeasy Mini Kit, Part 2*. RNA was eluted with 30 µl of RNase free water. Reverse transcription of the RNA extracted engineered cell lines was performed to synthesize a cDNA library for qPCR following Invitrogen's *SuperScript III First-Strand cDNA Synthesis System for RT-PCR* protocol. First, 1 µg of RNA (WT, Wrn, Wrn+Xrcc6, Rad1, Rad1+Xrcc6) was mixed in with primer (50 µM oligo(dT)), 10mM dNTP mix, and DEPC-treated water in a PCR tube to make 10 µl RNA/primer mixers. Each tube was placed in a heat block and incubated at 65C for 5 min then transferred onto ice. Next, a cDNA synthesis Mix was made using the following regents: 2X of 10X RT buffer, 10 mM of 25 mM $MgCl_2$, 20 mM of 0.1 M DTT, 4 U/µl of 40 U/µl RNaseOUT, and 20 U/µl of 200 U/ µl of SuperScript III RT. 10 µl of cDNA synthesis mix was gently mixed into each RNA/primer mixture, briefly centrifuged, then incubated for 50min at 50ºC. Step 5 Random hexamer incubation period was skipped, and the reactions were terminated by incubating for 5 min at 85ºC, chilling on ice then briefly centrifuging. Lastly, 1 µL of RNase H was added to each tube and incubated for 20 min at 37ºC.

Ethanol (EtOH) precipitation was done on each cDNA sample (WT, Wrn, Wrn+Xrcc6, Rad1, Rad1+Xrcc6). First, 1/10 of sodium acetate was added to each sample. Then, 2.5x of 100% EtOH was added and incubated at room temperature for 30 min. Samples were then centrifuged at 15,000g for 30 min. The supernatant was carefully decanted and discarded, and the samples were rinsed with 500 µl of 70% EtOH and centrifuged again for 15 min at 15,000g. The supernatant

was discarded again, and the pellet was air dried for 2-3 hrs before being resuspended in 50 µl of Tris.

Polymerase Chain Reaction

Successful integration of plasmid into the genome was validated via PCR. The forward primer (Supplementary table 1) was designed specifically to target the construct, while the reverse primer was designed to target the DNA of our DNA repair gene of interest (Wrn or Rad1) as shown in *Figure 6A*. PCR was done following the previously mentioned protocol. However, this PCR adhered to a different program setting of: 98ºC for 5 min, followed by 35 cycles of 98ºC for 10 min, 60-65ºC for 20 sec, and 72C for 30 sec, lastly a final extension of 2 min at 72ºC followed by a 4ºC hold. These samples were then purified following *PureLink PCR purification kit* (Invitrogen) and sequenced as in the previously explained method.

Quantification polymerase chain reaction (qPCR)

Transcription levels of DNA repair genes are determined by qPCR. To do so, synthesized cDNA is run with both target and reference primers, which are shown in *Supplementary Table 1* , in 5 different samples (WT, Wrn, Wrn+Xrcc6, Rad1, Rad1+Xrcc6), each in triplicate. Six samples are created for the WT sample, three containing Wrn specific primers and three with Rad1 specific primers. Each sample is prepared using the following reagents: 2 µM of 10µM forward primer, 2 µM of 10 µM Reverse Primer, 1X of 2X SYBR Green, nuclease free water, and 1µl cDNA. Controls are prepared similarly for each cDNA sample, apart from the added cDNA. The 96-well plate is then placed into a quantitative PCR machine; adhering to the following program: 95ºC for 5 min, then a 40-cycle sequence of: 95ºC (30 sec) and 65ºC (1 min).

3.4 Double strand break (DSB) repair assay

Transfection of CHO-SEAP (WT, Wrn, Wrn+Xrcc6, Rad1, Rad1+Xrcc6)

The different CHO-SEAP cell lines that were made were transfected with EJ5-GFP

reporter plasmid (Bennardo and Stark, 2010) (addgene #44026) to run the DSB repair assay (EJF-GFP assay) as explained in *Figure 3*. To do this, each cell line was seeded in triplicate at 100k cells/well in a 12 well plate. Note that the WT was seeded in quadruplicate to include a control. The next day, two mixes were made A and B. Mix A per reaction contained 100 µl of opti-MEM and 2.5 µl Lipofectamine LTX. Mix B per reaction contained 100 µl of opti-MEM, 1.0 µg of EJ5-GFP plasmid, and 2 µl of plus reagent. Mix A and B were gently combined and incubated for 25 min at room temperature. During incubation time, media were removed from the previously seeded 12 well plate and washed with Dulbecco's phosphate-buffered saline (DPBS). Prewarmed opti-MEM media was then added to each well at 800 µl. Following the incubation period, the mix was added dropwise to each well (except control) and rocked back and forth gently. Plasmid expression was then checked after 36hrs. The cells were lifted and resuspended in IMDM medium and placed into 1.5 ml tubes and centrifuged at 300g for 5min. The supernatant was discarded, and the cells were resuspended in 300 µl of PBS and placed in flow cytometry tubes to run Fluorescence-activated cell sorting (FACS).

Fluorescence-activated cell sorting

Quantification of Fluorescent protein expression was done on a FACS Canto II (BD) with 50,000 cells per sample. Gates for FSC, SSC, and far-red fluorescence were set and defined to select for viable cells expressing the DSB inducer. The gates that were chosen were used to relate GFP expressing cells to non-GFP expressing cells. Quantification of unrepaired DSBs is carried out by first filtering for live cells (SSH/FSC gating) and then relating the fraction of both far-red positive and GFP positive cells to the total fraction of far-red positive cells.

3.5 Long-term study

Phospha-light assay

In 12 well plates each cell line from Table 2 was seeded in triplicate at exactly 100k

cells/well, including the conditions CHO-SEAP WT + Methotrexate (MTX) and CHO-SEAP WT + Ku 60019 inhibitor. In the first week each cell line/condition was passaged three times and was supplemented with 5 µM of MTX. From time stamp T_0 onward, the cells were passaged 3 times per week for a period of about 8 weeks, supplementing Blasticidin and Bleomycin at the beginning of each week to their corresponding condition. Alternatively, 5 µM MTX or 5 µM Ku was supplemented after each passage to the corresponding condition.

The phospha-light assay was performed at T_0 and every other week to measure titer. The assay was run following the detection protocol for secreted placental alkaline phosphatase (Thermofisher, Phospha-Light System Protocol). First, the assay buffer & Reaction Buffer Diluent were brought to room temperature. Next, CSPD substrate was diluted 1:20 in Reaction Buffer Diluent and 5X Dilution buffer was diluted to 1X with water. Sample preparation was done by diluting 50 µl of culture medium in 50 µl of 1x Dilution buffer then heated at 65°C for 30 min then cooled on ice to room temperature. Equal parts of 50 µl diluted sample, assay buffer, and reaction buffer was added into a microplate well, with 5 min and 20 min incubation periods in between the addition of each buffer, respectively. Finally, the plate was placed on an illuminator and measured at 0.5 sec per well. Note that after passage 21 the culture medium was diluted 1:10 in 1MDM media (10% Fetal Bovine Serum, 1% pen/strep, 7 µg/mL Puromycin, 300 µg/mL Hydromycin) and prior to passage 21 cells were irradiated to induce DSBs. Luminescence measurements were normalized to the cell count at moment of culture medium collection.

4 Results and Discussion

4.1 Vector Engineering & Validation

Whole genome sequencing analysis revealed several potential mutated DSB related repair gene candidates (Fig. 2) that if repaired could improve DSB repair. Prior to testing if the restoration of these DNA repair genes in CHO-SEAP can facilitate better DSB repairs and retain higher production titer in long term culture, a total of four CHO-SEAP cell lines

were made, each expressing one or more DSB repair genes. Construction of the desired plasmids carrying an unmutated DNA repair gene of interest (Wrn, Wrn + Xrcc6, Rad1, Rad1 +Xrcc6) was done following the experimental design procedure shown in *Figure 4A*. The restriction double digest of lentiviral backbone pLJM1-EGFP.Blasticidin and pLJM1-EGFP. Bleomycin with restriction endonucleases NheI & BstB1 were visualized on a 1% agarose gel as shown in *Figure 4B & C*, respectively. The bands on each gel are located roughly at about 7,200 bp and 766 bp—with the top tier of dark bands representing the vector backbone pLJM1(Bleomycin or Blasticidin) and the lower set of bands representing EGFP, indicating a successful digestion.

Amplification of CHO-SEAP cDNA for isolation of single cDNAs of unmutated Wrn and Rad1 DNA repair genes were isolated via PCR. PCR results of Wrn cDNA (Fig. 4D) revealed the target Wrn band of interest at ~4000 bp as well as several other unidentified bands in lanes 1-5, due to the various annealing temperatures that were tested. To ensure that nonspecific bands were not cloned into our engineered vector, the 4604 bp bands were gel extracted, purified, and later digested by restriction endonucleases Nhe1 and BstB1. Similarly, Rad1 cDNA was also PCR amplified (Fig. 4E), the expected band of interest for the Rad1 gene of interest is 993 bp, however the bands appeared to be located at ~900 bp in lanes 2-3, which could result from overloading the gel, thus giving the appearance of a lower band. The lighter set of bands, denoted by the arrows in lanes 2-3, are greater than 3000 bp and are possibly due to the use of a longer extension time of 1 min as opposed to a 30 sec extension time.

Following ligation of the Wrn and Rad1 DNA to their respective lentiviral backbone (plJM1.Blast and pLJM1.Bleo), colony PCR was performed to verify correct ligation of insert to the corresponding vector backbone. Visualization of Wrn colony PCR (Fig. 4F) shows a prominent band of roughly 4000 bp in lanes 2-5 (red box). Nonspecific bands are shown in

lanes 3-5—also due to the varying annealing temperatures used. However, these bands appear darker and less scarce in lane 5; which could be due to the sample being a ligation mix and not a colony, therefore it is most likely that the primers annealed to undigested portions of the vector DNA . Rad1 colony PCR (Supplementary Figure 1) was also done and revealed that out of the 13 colonies that were picked only 2 contained a band, a plausible explanation for this occurrence is due to not setting a longer initial duration time during PCR, making full cell lysis limited. Due to the distance of the colonies to the marker, the two colonies (#2 & #11) were rerun on a different gel at different concentrations as shown in *Figure 4G* below. As can be seen from the image, highlighted by the red box colony #11 in lanes 3&4 confirmed the plasmid plJM1.Rad1.Bleo was indeed present.

To validate the two engineered plasmids pLJM1.Wrn.Blast and pLJM1.Rad1.Bleo a final restriction digest was performed (Fig. 5H and Fig. 5I, respectively), as well as sequencing analysis of the isolated plasmid. *Figure 5H* showcases double restriction enzyme digests of plasmid pLJM1.Wrn.Blast. Lane 1 contains the undigested pLJM1.Wrn.Blast plasmid denoted by the thick dark band coming in at ~10,000 bp; the band sits lower than the expected band of 11,847 bp most likely due to the plasmid being supercoiled, thus traveling faster in the gel. A double restriction digest using Nhe1 and Bstb1 are shown in lane 2 with two expected bands of 7,226 bp and 4,621 bp, representing the plJM1. Blast backbone and the Wrn gene, respectively. Moreover, restriction double digest with restriction endonucleases Kpn and Bam1 is shown in lane 3 and results in a band of 588bp and roughly 11,00bp as expected. Rad1 double digests (Fig. 4I) was cut with the same restriction enzymes, however Kpn and Bam1 was not used. Lane 1 contains undigested plasmid of Rad1 (8,134 bp) and lane 2 and 3 contain different concentrations of cut plasmid with Nhe1 and BstB1,

with bands at 993 bp and 7,145 bp.

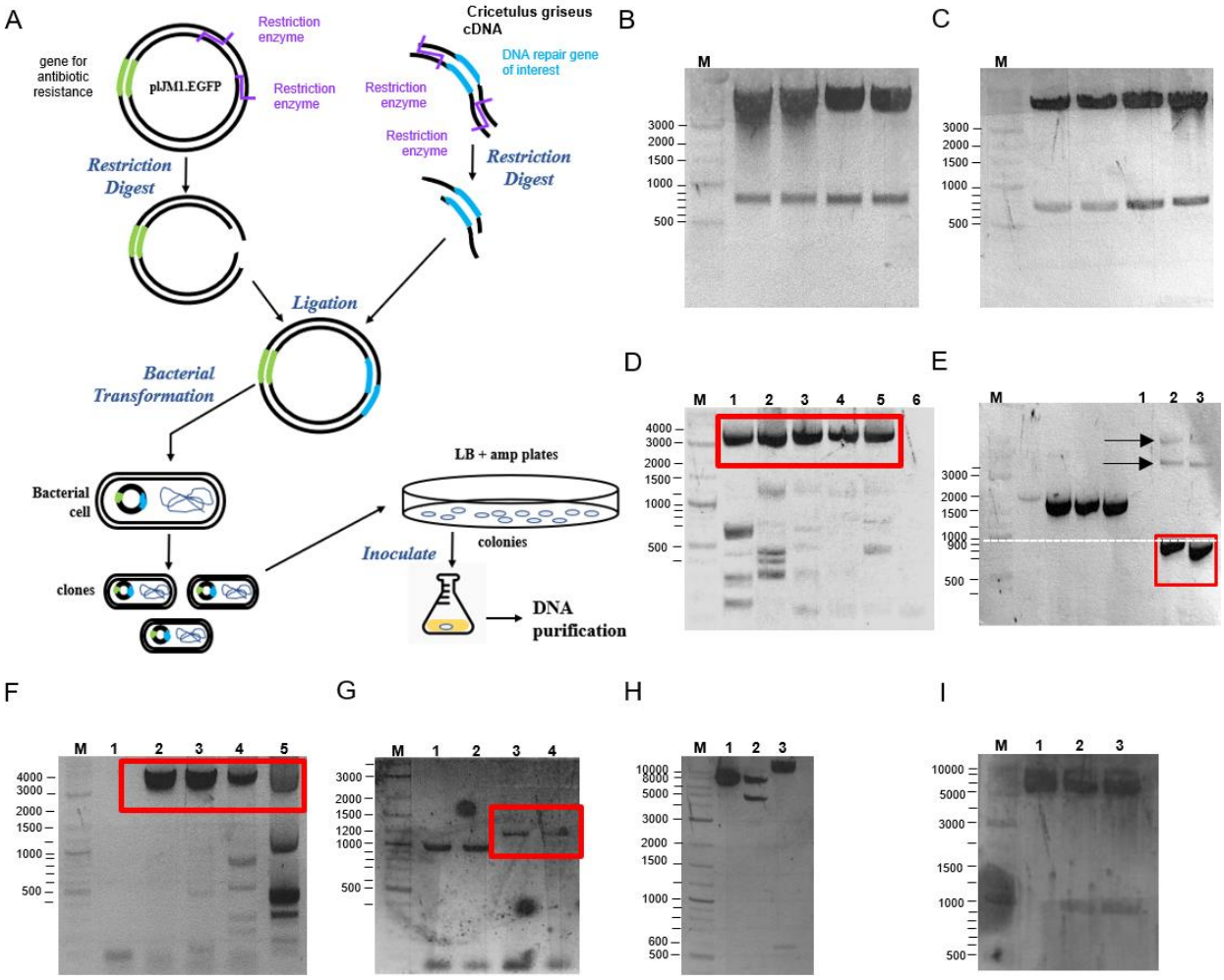


Figure 4: Construction of unmutated DSB repair gene recombinant lentiviral plasmids. A) Step 1: pLJM1.EGFP with Blasicidin or Bleomycin resistance gene is cut with restriction enzymes (RE) Nhe1 and BstB1. Step 2: DNA repair gene Rad1 and Wrn is isolated from Chinese hamster cDNA via PCR and cut with RE's Nhe1 and BstB1. Step 3: vector backbone pLJM1. Blast or Bleo is ligated to Wrn or Rad1 cDNA and transformed into *E.coli* and plated onto LB (100 µg/mL AMP) plates. Step 4: Colonies are screened via colony PCR and inoculated into LB culture overnight and purified. **B)** Restriction digest of pLJM1.Blasticidin. **C)** Restriction digest of pLJM1. Bleomycin. **D)** PCR of Wrn cDNA run on an agarose gel; (1-5) Wrn cDNA tested at different annealing temperatures of 62-70 °C; (6) negative control. **E)** PCR of Rad1 cDNA run on an agarose gel; (1) negative control; (2&3) Rad1 cDNA. **F)** Wrn Colony PCR tested at annealing temperatures of 60-66°C ; (1) negative control; (2-4) Wrn colonies, #1-3 at annealing temperatures of; (5) Ligation mix. **G)** Rad1 Colony PCR; (1-2) colony #2; (2-4) colony #11. **H)** Restriction digest of pLJM1.Wrn.Blasticidin; (1) undigested; (2) cut with Nhe1 and BstB1; (3) cut with Kpn and Bam1. **I)** Restriction digest of pLJM1.Rad1.Bleomycin; (1) undigested; (2-3) cut with Nhe1 and BstB1.

4.2 Cell line generation and Validation

Upon generating the four engineered cell lines validation of plasmid integration into the CHO-SEAP genome was done via PCR by targeting the plasmid construct and the DNA of Wrn or Rad1 as shown in *Figure 5A.* Results of the Wrn and Rad1 PCR are shown in *Figure 5B*. Lanes 1-4 represent the PCR samples that have been tested with the Wrn_gPCR primer set, while lanes 5-8 represent samples tested with the Rad1_gPCR primer set. CHO-SEAP Wrn (3), CHO-SEAP Wrn + Xrcc6 (4), CHO-SEAP Rad1 (7), and CHO-SEAP Rad1 + Xrcc6 (8) all produce a dark band at roughly 800 bp, with the band in lanes 7 and 8 sitting slightly lower than lanes 3 and 4, indicating a smaller amplicon. Most importantly, the dark bands correspond to the expected amplicon sizes of both Wrn (802 bp) and Rad1 (769 bp) PCRs . As a negative control, both the Wrn and Rad1 primers were used against WT CHO-SEAP (lanes 2 and 6, respectively), which indicated no band at the corresponding length of the amplicon. However, lane 6 contains a faint band at ~925 bp, which could be due to non-specific priming and extension with the Rad1 primers.

A CHO-SEAP (Wrn, Wrn +Xrcc6, Rad1, Rad1 + Xrcc6, & WT) cDNA library was made and qPCR was performed to ensure expression of the plasmids (Wrn, Wrn +Xrcc6, Rad1, Rad1 + Xrcc6) in their respective CHO-SEAP cell line. qPCR was done utilizing two different primer sets Wrn or Rad1_qPCR primers and GAPDH primers and compared to WT. The average Cq of Wrn (26.32) was smaller than the average Cq of Wrn + Xrcc6 (26.7), with the average Cq of the WT (Wrn_qPCR_fw/rev) being the largest (29.08). Meanwhile, the average Cq of Rad1 (18.96) was also smaller when compared to both the double repair gene cell line, Rad1 + Xrrcc6 (21.72) and the average Cq of WT (Rad1_qPCR_fw/rev). Utilizing the Livak method the average ΔΔCq for Wrn, Wrn +Xrcc6, WT was calculated and normalized to GADPH, and the relative fold expression was plotted in *Figure 6A* along with the melt curves in *Figure 6B*. From the results, expression of the Wrn gene in CHO-SEAP Wrn was greater

than CHO-SEAP WT, however although the double repair gene cell line (CHO-SEAP Wrn+Xrcc6) demonstrated a greater expression than WT as was expected, the expression of the Wrn gene was lower than CHO-SEAP Wrn. Gel analysis of the qPCR utilizing Wrn_qPCR (lanes 1-7 and 1-3 of Figure 6B&C, respectively) primer sets show that WT (lanes 2-4), Wrn (lanes 5-7) and Wrn + Xrcc6 (lanes 1-3 of Fig. 6C) show a band at roughly 300 bp, that correspond to the qPCR amplicon. Meanwhile, qPCR using the reference primers GAPDH_qPCR_fw/rev (lanes 8-14 and 4-6 of Fig. 6B&C, respectively) show a corresponding band of ~400 bp.

The average ΔΔCq for Rad1, Rad1 +Xrcc6, WT was also calculated normalized to GADPH and the relative expression (Fig. 7A) data demonstrated a similar pattern of results as the Wrn qPCR. Rad1 gene expression was 6.8 folds greater than WT and 3.6 folds greater than Rad1 + Xrcc6. Once again, the double repair gene cell line showed a much smaller expression than the single repair gene cell line. Similarly, transcription levels of Rad1 gene with Rad1_qPCR primer set (lanes 1-7 and 1-3 of Fig. 7B&C, respectively) and GADPH primer set (lanes 1-3 of Fig 7C) was analyzed and revealed a band at ~300bp for Rad1_qPCR primer set and a band at 400 bp for GADPH. CHO-SEAP WT (lanes 2-4) contained relatively fainter bands than the other Rad1 cell lines.

Overall, based on the *Figure 6* and *Figure 7,* the four pLJM1 carrying DNA repair gene plasmids were successfully integrated into each cell line and were able to overexpress the DNA repair gene of choice.

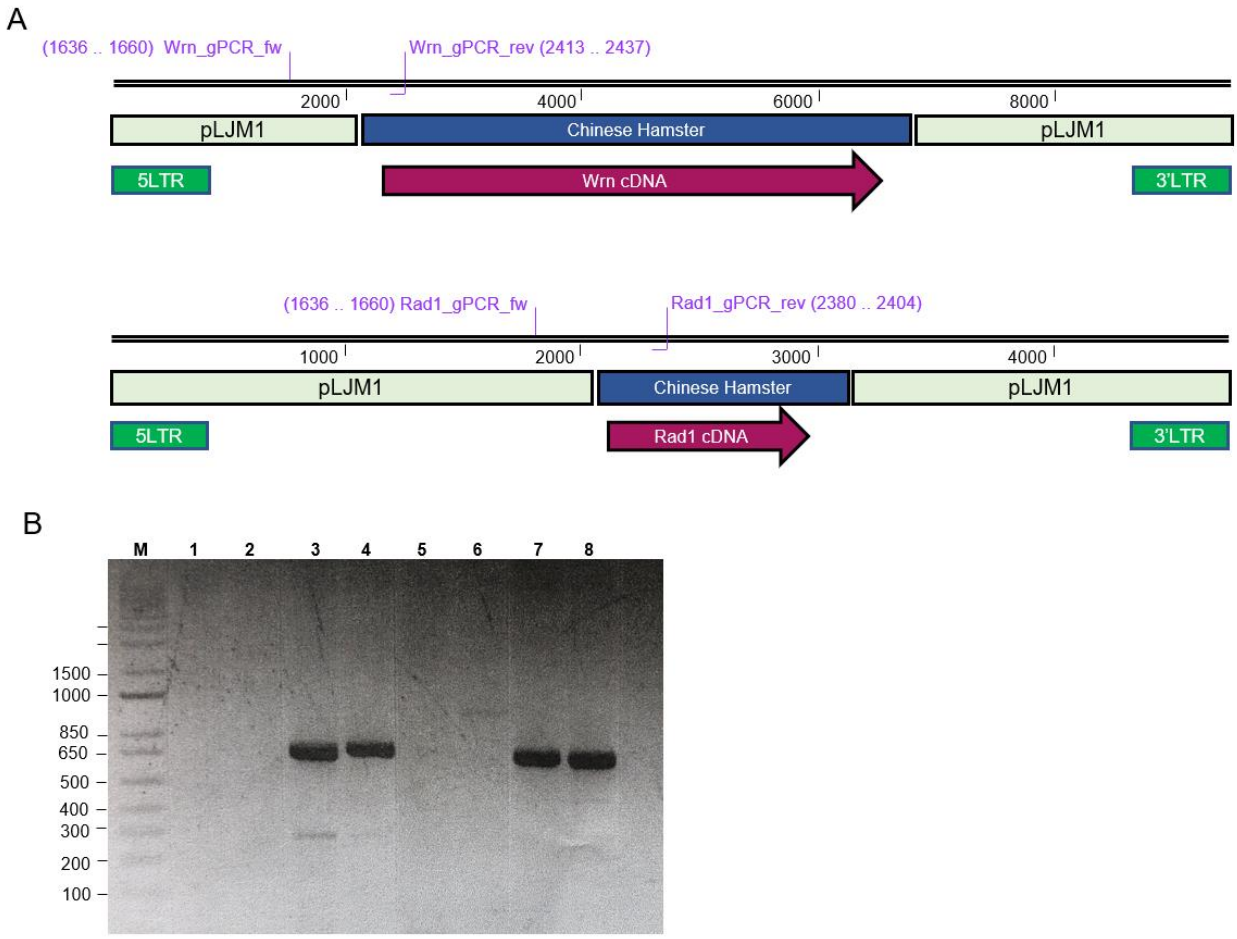


Figure 5: Integration of plasmid into CHO-SEAP. A) Top- schematic of primer placement for Wrn; Bottom- schematic of primer placement for Rad1; The forward primers Wrn_gPCR_fw and Rad1_gPCR_fw lay on the vector backbone of pLJM1. Blast and pLJM1.Bleo, respectively, while the reverse primers Wrn_gPCR_rev and Rad1_gPCR_rev specifically target Rad1 and Wrn; **B)** PCR samples run on an agarose gel. (M) 1 kb plus ladder; (1) Wrn negative control; (2) WT CHO-SEAP; (3) CHO-SEAP Wrn; (4) CHO-SEAP Wrn + Xrcc6; (5) Rad1 negative control; (6) WT CHO-SEAP; (7) CHO-SEAP Rad1 (8) CHO-SEAP +Xrcc6.

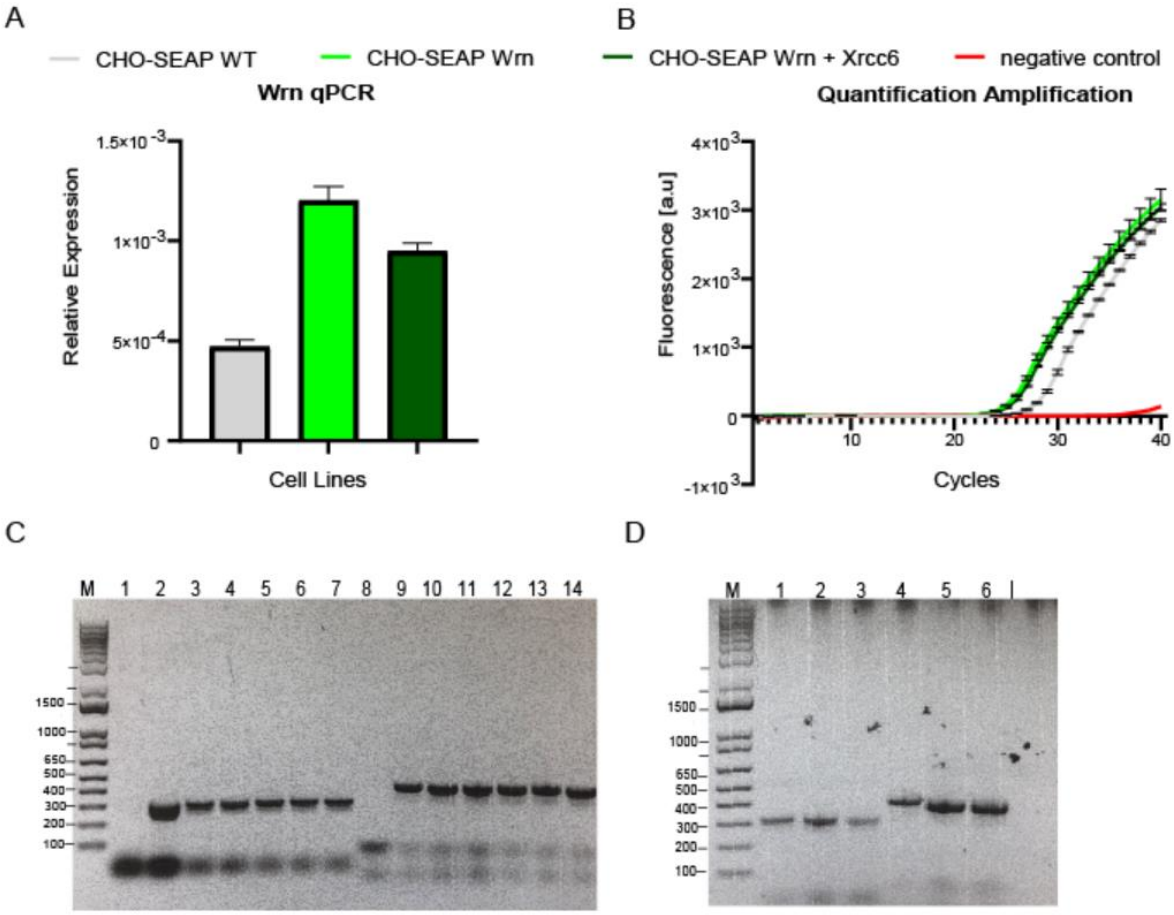


Figure 6: Quantitative PCR (qPCR) comparison of Wrn cell lines. A) Relative expression levels of CHO-SEAP Wrn, CHO-SEAP Wrn + Xrcc6, and CHO-SEAP WT normalized to GADPH housekeeping gene. Expression levels of CHO-SEAP WT is the lowest compared to CHO-SEAP Wrn and CHO-SEAP Wrn + Xrcc6 with CHO-SEAP Wrn expressing the Wrn the greatest. **B)** Amplification plot of Wrn cell lines. **C)** Agarose gel containing WT and Wrn qPCR samples. (M) 1 kb plus ladder; (1) Wrn_qPCR_fw/rev negative control; (2-4) Wrn_qPCR_fw/rev WT; (5-7) Wrn_qPCR_fw/rev Wrn; (8) GADPH_qPCR_fw/rev negative control; (9-11) GADPH_qPCR_fw/rev WT; (12-14) GADPH_qPCR_fw/rev Wrn. **D)** Agarose gel containing Wrn + Xrcc6 samples. (M) 1 kb plus ladder; (1-3) Wrn_qPCR_fw/rev Wrn + Xrcc6; (4-6) GADPH_qPCR_fw/rev Wrn + Xrcc6.

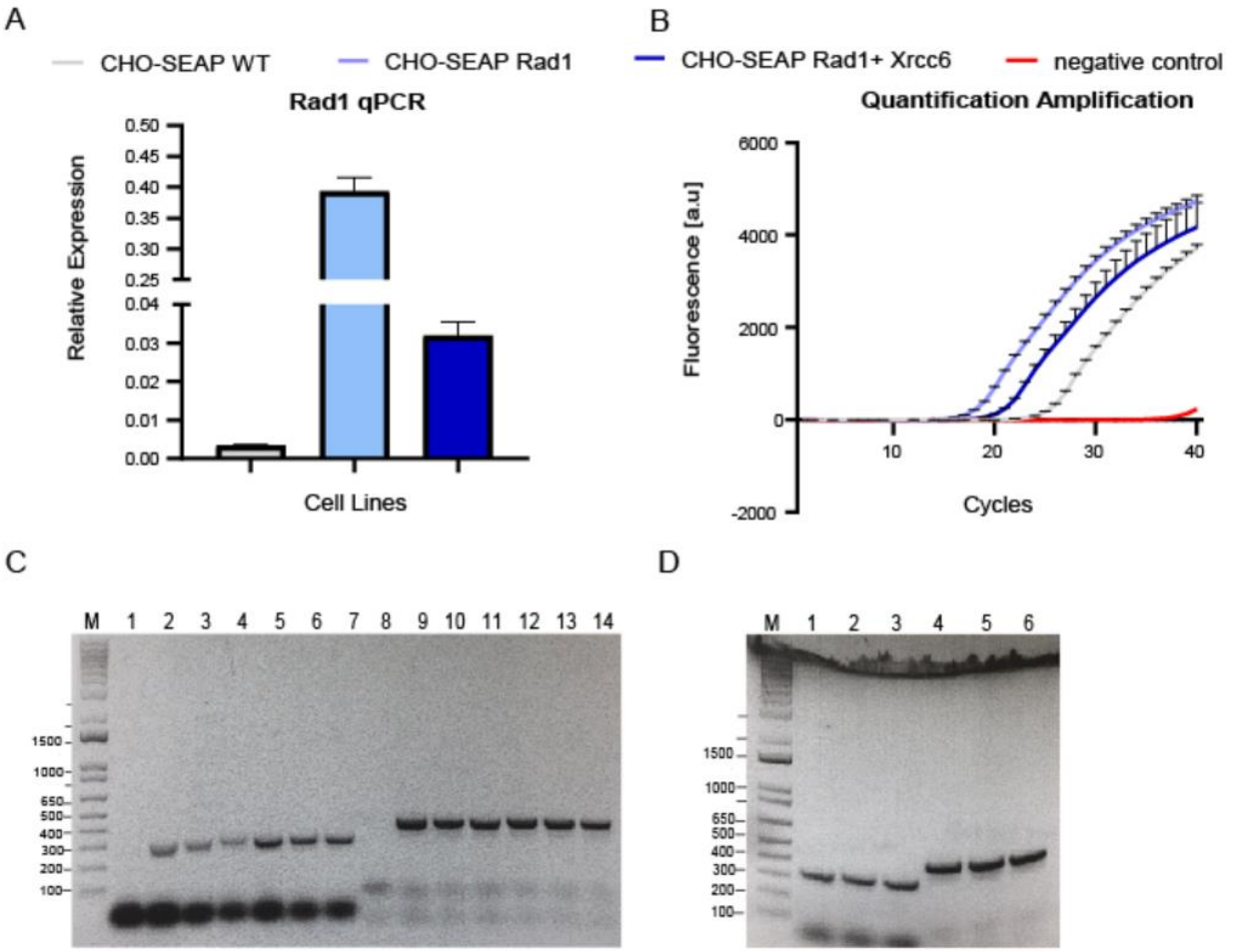


Figure 7: Quantitative PCR (qPCR) comparison of Rad1 cell lines. A) Relative expression levels of CHO-SEAP Rad1, CHO-SEAP Rad1 + Xrcc6, and CHO-SEAP WT normalized to GADPH housekeeping gene. CHO-SEAP WT expression levels remained the lowest, with CHO-SEAP Rad1 being 6.8 folds greater than WT and 3.6 folds greater than CHO-SEAP Rad1 +Xrcc6. **B)** Amplification plot of Rad1 cell lines. **C)** Agarose gel containing WT and Rad1 qPCR samples. (M) 1 kb plus ladder; (1) Rad1_qPCR_fw/rev negative control; (2-4) Rad1_qPCR_fw/rev WT; (5-7) Rad1_qPCR_fw/rev Rad1; (8) GADPH_qPCR_fw/rev negative control; (9-11) GADPH_qPCR_fw/rev WT; (12-14) GADPH_qPCR_fw/rev Rad1. **D)** Agarose gel containing Rad1 + Xrcc6 samples. (M) 1 kb plus ladder; (1-3) Wrn_qPCR_fw/rev Rad1 + Xrcc6; (4-6) GADPH_qPCR_fw/rev Rad1 + Xrcc6.

4.3 DSB repair efficiency and SEAP titer was compared to CHO-SEAP WT

By transfecting several CHO-SEAP cell lines with different unmutated DSB repair genes, the overexpression of that particular gene was expected to supplement for the mutated repair gene and aid in the DSB repair pathway or increase the rate at which the breaks are fixed leading to a decrease in genomic and production instability. A modified EJ5-GFP assay modified from Bennardo and Stark, 2010 (Fig. 3) was run for each cell line. This assay reported the number of cells that were slow or faulty in repairing the DSB induced by the

two sgRNAs by fluorescing green. These green signals were identified through FACS and plotted in *Figure 8A & B*. CHO-SEAP Wrn and CHO-SEAP Wrn + Xrcc6 were compared to the WT (Fig. 8A) and revealed surprising results. CHO-SEAP Wrn failed to repair DSBs at a ratio of 48.5%, while CHO-SEAP WT had a similar failed DSB reparation ratio of 48.7%. However, what was more uncharacteristic was the ratio of unrepaired DSBs in the double repair gene cell line. Given that the CHO-SEAP Wrn + Xrcc6 cell line had overexpression of two different unmutated DSB repair genes. It was believed that the cell line would rescue deficient repair in the DSB repair pathway and increase the DSB repair rate. The ratio of unrepaired DSB (64.3) exceeded the WT ratio by 15.6. Most notably, the double repair gene cell line CHO-SEAP Rad1 + Xrcc6 showed more of an improvement in DSB repair than just the single repair gene Rad1 (59.0), obtaining a ratio of 48.7 for unrepaired DSBs. These results were contrary to the results for CHO-SEAP Wrn + Xrcc6, which begged the question of whether Xrcc6 is aiding or hindering DSB repair based on its interaction with Wrn or Rad1.

The ability of these CHO-SEAP cell lines expressing an intact copy of a chosen DNA repair gene or genes, to retain production titer of secreted alkaline phosphatase—a reporter protein that is secreted in cell culture (Yang et al. 1997), was tested by measuring SEAP titer (fluorescence) every 1 to 2 weeks over a period of about 8 weeks. All cell lines were initially supplemented with MTX—an inhibitor used select for transfected cells and thus drive cells to adapt, for the first week before measuring SEAP titer. Observation of each cell line and condition (WT +MTX and WT + Ku), it was shown that at initial time point (P0) the cells all fluoresced at roughly the same level of $1\text{-}2x10^4$ a.u, with the exception of Rad1 and Wrn + Xrcc6 which were notably higher. Throughout passaging, the SEAP titer stayed relatively constant until the 13^{th} passage (P13), where the single repair gene cell lines Wrn and Rad1 had increased cell titer relative to the other cell lines and conditions. This observation

contradicted the expected outcome for Rad1, as it was shown to not as effectively repair DSBs compared to the other cell lines (Fig.8B). Nonetheless, the WT SEAP titer was not decreasing as much, and was equally comparable to WT + MTX SEAP titer, however WT + Ku 60019 (Ku) inhibitor was much lower, as expected since the Ku inhibitor blocks ATM activation, an important regulator of signaling in the DSB repair pathway (Álvarez-Quilón et al. 2014).

After passaging a few more times the cell lines were irradiated to induce DSBs two weeks prior to measuring SEAP titer at passage 21 (P21). As a result, SEAP titer was reduced in all cell lines. As the cells were continuously cultured and measured, SEAP titer measurements demonstrated an increase in cell lines expressing Xrcc6 that surpassed WT. However, although the SEAP titer of Xrcc6 expressing cell lines stayed higher compared to cell lines overexpressing single DNA repair genes (Wrn + Rad1), SEAP titer dropped lower than the WT levels. This observation could be in part to not waiting the necessary incubation period before measuring luminescence of each sample CHO-SEAP cell line.

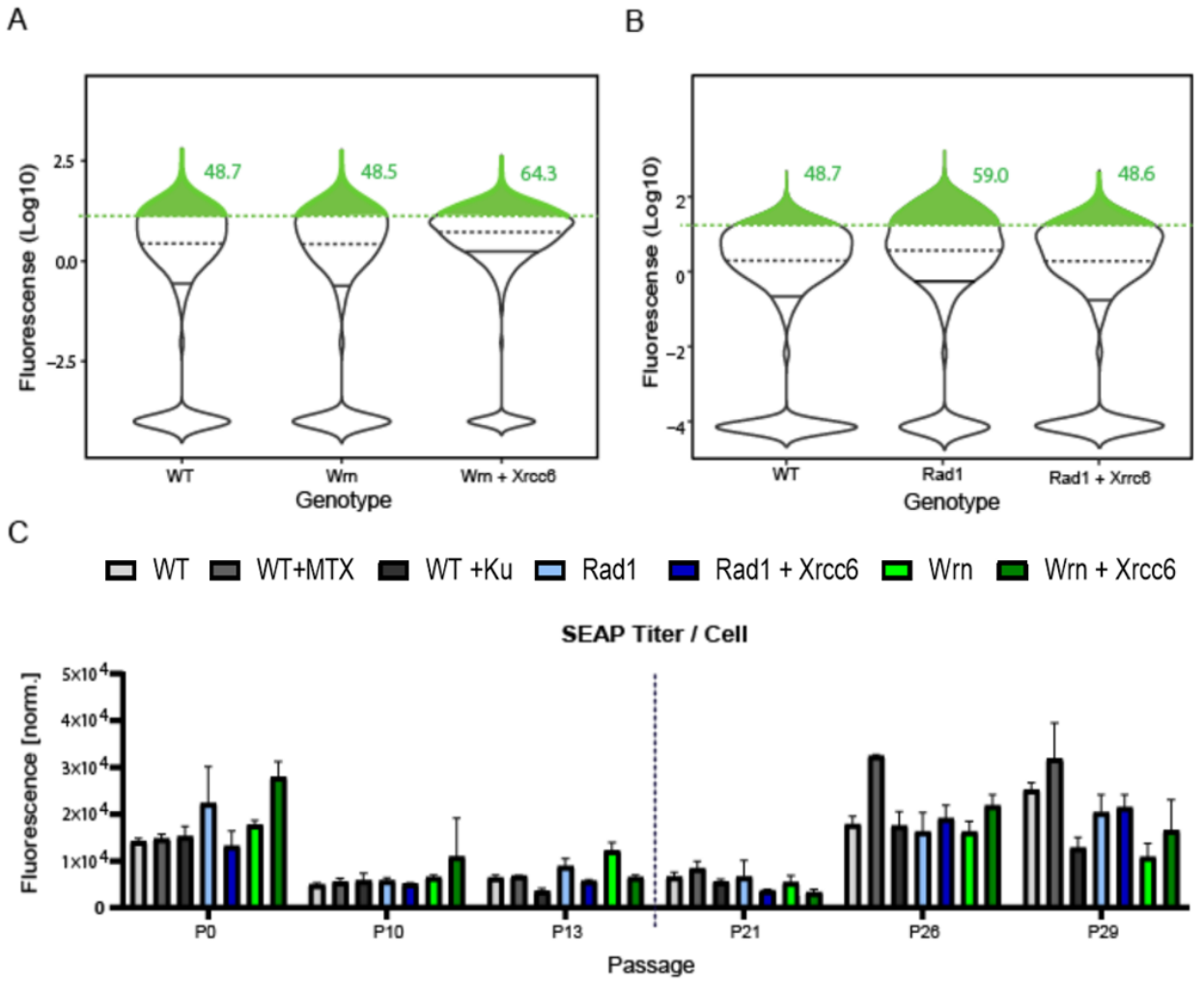


Figure 8: DSB repair efficiency and SEAP titer measurements of engineered CHO cell lines. A) Violin plot of unrepaired DSB of CHO-SEAP Wrn and Wrn + Xrcc6 compared to WT. The green represents the ratio of unrepaired DSBs to total cells, measured through FACS following the DSB repair assay (Fig.3). WT + Xrcc6 contains highest ration of unrepaired DSBs (64.3) compared to WT (48.7) and Wrn (48.5). **B)** Violin plot of unrepaired DSB of CHO-SEAP Rad1 and Rad1+ Xrcc6 compared to WT. Rad1 contains highest ration of unrepaired DSBs (59.0) compared to WT (48.7) and Wrn (48.6). **C)** SEAP titer measurements on the different CHO cell lines following supplementation of 5 µM MTX to each cell line and condition for a week. Following passage 13 (P13) cells underwent Irradiation to induce DSBs, denoted by the dotted black line. Note that the last two passages were done at 1:10 dilutions of the CHO-SEAP cell culture medium .

5 Conclusion

CHO cells inability to retain production titer after long-term culture has often been linked to its ability to repair DSBs. We hypothesized that the inability to efficiently fix these DSBs are due to deficient DSB related genes. Whole sequencing analysis of CHO cells across several cell lines showed SNPs in key genes related to DSB repair. In this study, deficient DNA repair genes were handpicked, based on the sequencing analysis done previously by colleagues in the lab and restored by knock-in of the unmutated DNA repair gene. A comparison of DSB repair and SEAP titer was done for each cell line expressing a single or double repair gene (Wrn, Wrn + Xrcc6, Rad1, or Rad1 + Xrcc6) to the CHO-SEAP WT cell line.

Overexpression of Wrn and Rad1 + Xrcc6 had roughly equal or slightly positive effects on the repair of DSBs when compared to WT. Meanwhile, Rad1 and Wrn + Xrcc6 showed a decrease in DSB repair compared to WT. Although, the cell lines did not show promise for DSB repair it was essential to record the cell lines ability to retain or increase SEAP titer as cells are cultured long term. Contrary to how the cell lines repaired DSBs, the cell lines expressing single DSB repair gene (Wrn or Rad1) increased in SEAP titer compared to the WT and the WT supplemented with MTX. After inducing DSBs by irradiation the SEAP titer of Wrn and Rad1 maintained similar pattern compared to the double expressed cell lines, although reduced significantly. After several passages, the pattern then began to shift, and the SEAP titer was increased in cell lines expressing Xrcc6. This pattern switch in SEAP titer suggests that Xrcc6 plays a key role in CHO-SEAP repair system.

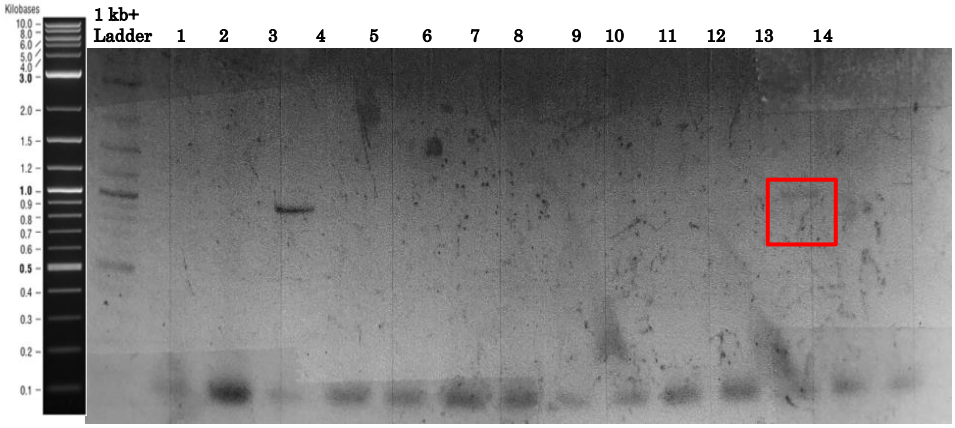
Undoubtedly, based on the observations made, it is crucial to note that reparation of mutated DSB repair genes to combat titer loss is not a straightforward solution. More research would be needed to conduct a high-throughput screening of targets that affect multiple repair pathways, not just NHEJ and HR, and thus find better suited gene targets

for DSB repair. Future directions would include testing CHO-SEAP cell line expressing repaired Xrcc6 as it has shown a promising ability to improve DSB repair and production titer. Moreover, Xrcc6 coupled with the correct target gene that is directly involved in DSB repair could have the potential to improve the CHO-SEAP repair pathway.

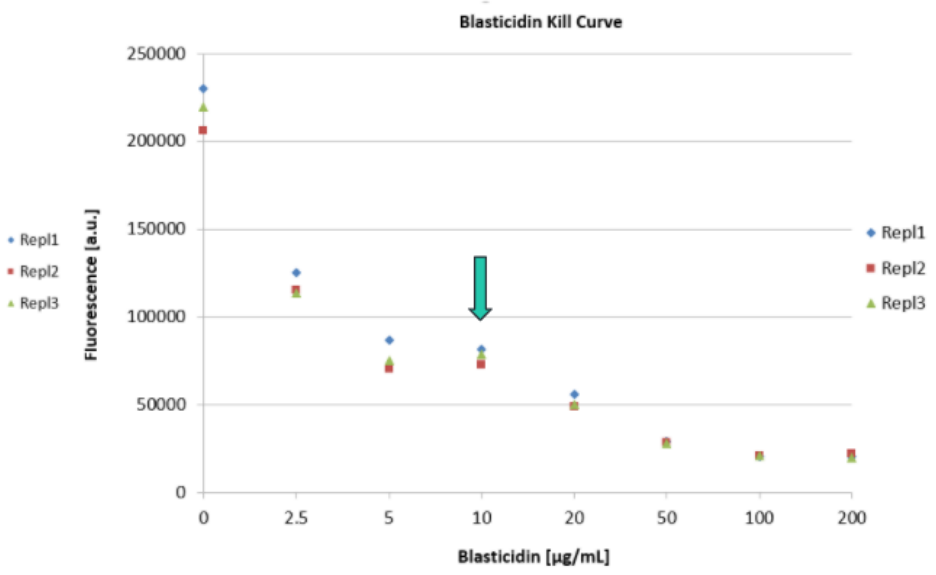
APPENDIX

Supplementary Table 1: Primer list

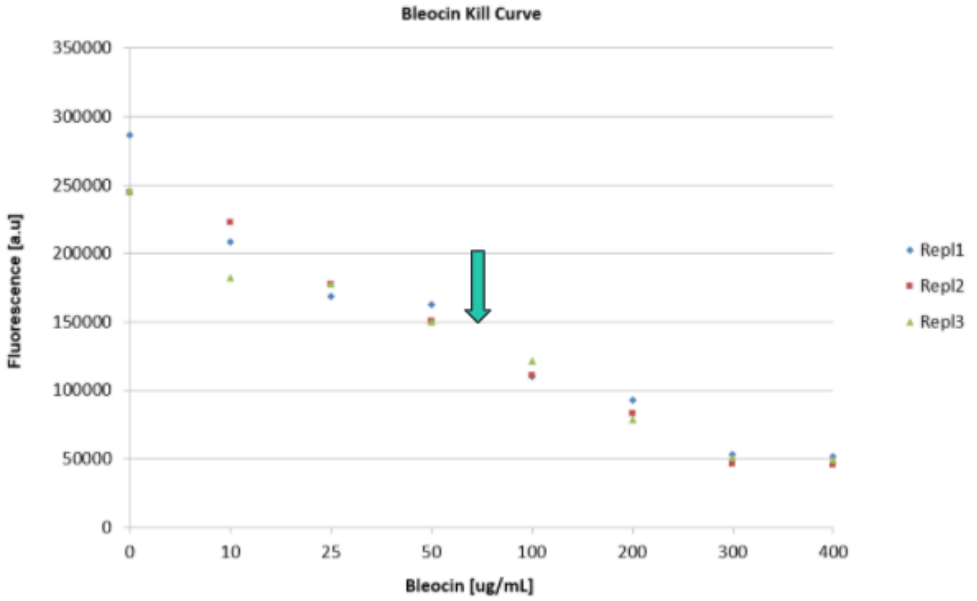
Primer Pair	Sequence
Wrn_Nhe_F2/Wrn_Bst_R2	TATTGCTAGC gcctgagttgtgggctttgg/ TATTTTCGAA catgtctcctgatgcgggaag
Rad1_Nhe_F/Rad1_Bstb1_R	TATTGCTAGCtgaccgtgcagtggaccatcc/ TATTTTCGAA gagaattcgccaacactgctc
Rad1_gPCR_Fw/rev	TTGCATATACGATACAAGGCTGTTA/AGATTCTTCCCGAATGATA AACTCT
Wrn_gPCR_Fw/rev	TTGCATATACGATACAAGGCTGTTA/AGCTTCATAACTGTAAACG ATGGAC
Rad1_qPCR_Fw/rev	TTCGGATGTGTTACCAAGGTTATGG/CAGTGAAATGCTTCCATCA AGTCAG
Wrn_qPCR_fw/rev	ataggcttctgcctctcacagcaat/gctcacacagctctcttcagtgcta



Supplementary Figure 1: Rad1 colony PCR.



Supplementary Figure 2: Blasticidin kill curve used to determine ideal concentration for selection



Supplementary Figure 3: Bleocin kill curve used to determine ideal concentration for selection

REFERENCES

Altamirano, Claudia, Julio Berrios, Mauricio Vergara, and Silvana Becerra. 2013. "Advances in Improving Mammalian Cells Metabolism for Recombinant Protein Production." *Electronic Journal of Biotechnology: EJB* 16 (3). https://doi.org/10.2225/vol16-issue3-fulltext-2.

Álvarez-Quilón, Alejandro, Almudena Serrano-Benítez, Jenna Ariel Lieberman, Cristina Quintero, Daniel Sánchez-Gutiérrez, Luis M. Escudero, and Felipe Cortés-Ledesma. 2014. "ATM Specifically Mediates Repair of Double-Strand Breaks with Blocked DNA Ends." *Nature Communications* 5 (February): 3347.

Bennardo, Nicole, Anita Cheng, Nick Huang, and Jeremy M. Stark. 2008. "Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair." *PLoS Genetics* 4 (6): e1000110.

Bennardo,N. and Stark,J.M. (2010) ATM limits incorrect end utilization during non-homologous end joining of multiple chromosome breaks. *PLoS Genet.*, **6**, 16–18.

Clark, David P., and Nanette J. Pazdernik. 2016. "Recombinant Proteins." *Biotechnology*. https://doi.org/10.1016/b978-0-12-385015-7.00010-7.

Dahodwala, Hussain, and Kelvin Lee. 2019. "The Fickle CHO: A Review of the Causes, Implications, and Potential Alleviation of the CHO Cell Line Instability Problem." *Current Opinion in Biotechnology* 60 (December): 128–37.

Jackson, Stephen P. 2002. "Sensing and Repairing DNA Double-Strand Breaks." *Carcinogenesis* 23 (5): 687–96.

Jayakpal, K, Katie F. Wlaschin, Wei-Shou Hu Miranda G. S. Yap. 2007."Recombinant Protein Therapeutics from CHO Cells - 20 Years and Counting" *Chemical Engineering Progress* 103 (10): 40-47.

Jeggo, P. A., and M. Löbrich. 2007. "DNA Double-Strand Breaks: Their Cellular and Clinical Impact?" *Oncogene*. https://doi.org/10.1038/sj.onc.1210868.

Jia, Yan-Long, Xiao Guo, Jiang-Tao Lu, Xiao-Yin Wang, Le-Le Qiu, and Tian-Yun Wang. 2018. "CRISPR/Cas9-Mediated Gene Knockout for DNA Methyltransferase Dnmt3a in CHO Cells Displays Enhanced Transgenic Expression and Long-Term Stability." *Journal of Cellular and Molecular Medicine* 22 (9): 4106–16.

Kim, Minsoo, Peter M. O'Callaghan, Kurt A. Droms, and David C. James. 2011. "A Mechanistic Understanding of Production Instability in CHO Cell Lines Expressing Recombinant Monoclonal Antibodies." *Biotechnology and Bioengineering* 108 (10): 2434–46.

Kuo, Chih-Chung, Austin W. T. Chiang, Isaac Shamie, Mojtaba Samoudi, Jahir M. Gutierrez, and Nathan E. Lewis. 2018. "The Emerging Role of Systems Biology for Engineering Protein Production in CHO Cells." *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2017.11.015.

Lee, Kyoung Ho, Masayoshi Onitsuka, Kohsuke Honda, Hisao Ohtake, and Takeshi Omasa. 2013. "Rapid Construction of Transgene-Amplified CHO Cell Lines by Cell Cycle Checkpoint Engineering." *Applied Microbiology and Biotechnology* 97 (13): 5731–41.

Li, Shengwei, Xiaoping Gao, Rui Peng, Sheng Zhang, Wei Fu, and Fangdong Zou. 2016. "FISH-Based Analysis of Clonally Derived CHO Cell Populations Reveals High Probability for Transgene Integration in a Terminal Region of Chromosome 1 (1q13)." *PLOS ONE*. https://doi.org/10.1371/journal.pone.0163893.

López-Camarillo, César, Mavil Lopez-Casamichana, Christian Weber, Nancy Guillen, Esther Orozco, and Laurence A. Marchat. 2009. "DNA Repair Mechanisms in Eukaryotes: Special Focus in Entamoeba Histolytica and Related Protozoan Parasites." *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 9 (6): 1051–56.

Mao, Zhiyong, Michael Bozzella, Andrei Seluanov, and Vera Gorbunova. 2008. "Comparison of Nonhomologous End Joining and Homologous Recombination in Human Cells." *DNA Repair* 7 (10): 1765–71.

Matsuyama, Rima, Noriko Yamano, Namiko Kawamura, and Takeshi Omasa. 2017. "Lengthening of High-Yield Production Levels of Monoclonal Antibody-Producing Chinese Hamster Ovary Cells by Downregulation of Breast Cancer 1." *Journal of Bioscience and Bioengineering* 123 (3): 382–89.

Moritz, Benjamin, Peter B. Becker, and Ulrich Göpfert. 2015. "CMV Promoter Mutants with a Reduced Propensity to Productivity Loss in CHO Cells." *Scientific Reports* 5 (November): 16952.

Nematpour, Fatemeh, Fereidoun Mahboudi, Behrouz Vaziri, Vahid Khalaj, Samira Ahmadi, Maryam Ahmadi, Saedeh Ebadat, and Fatemeh Davami. 2017. "Evaluating the Expression Profile and Stability of Different UCOE Containing Vector Combinations in mAb-Producing CHO Cells." *BMC Biotechnology* 17 (1): 18.

Pham, Phuc V. 2018. "Medical Biotechnology." *Omics Technologies and Bio-Engineering*. https://doi.org/10.1016/b978-0-12-804659-3.00019-1.

Reinhart, David, Wolfgang Sommeregger, Monika Debreczeny, Elisabeth Gludovacz, and Renate Kunert. 2014. "In Search of Expression Bottlenecks in Recombinant CHO Cell Lines—a Case Study." *Applied Microbiology and Biotechnology*. https://doi.org/10.1007/s00253-014-5584-z.

Rupp, Oliver, Madolyn L. MacDonald, Shangzhong Li, Heena Dhiman, Shawn Polson, Sven Griep, Kelley Heffner, Inmaculada Hernandez, Karina Brinkrolf , Vaibhav Jadhav , Mojtaba Samoudi, Haiping Hao, Brewster Kingham, Alexander Goesmann, Michael J Betenbaugh, Nathan E Lewis, Nicole Borth, Kelvin H Lee .2018. "A Reference Genome of the Chinese Hamster Based on a Hybrid Assembly Strategy."*Biotechnology and Bioengineering* 115 (8): 2087–2100.

Scully, Ralph, Arvind Panday, Rajula Elango, and Nicholas A. Willis. 2019. "DNA Double-Strand Break Repair-Pathway Choice in Somatic Mammalian Cells." *Nature Reviews. Molecular Cell Biology* 20 (11): 698–714.

Sharma, S., and S. C. Raghavan. 2016. "Nonhomologous DNA End Joining." *Encyclopedia of Cell Biology*. https://doi.org/10.1016/b978-0-12-394447-4.10047-1.

Varga, Tamas, and Peter Aplan. 2005. "Chromosomal Aberrations Induced by Double Strand DNA Breaks." *DNA Repair* 4 (9): 1038–46.

Vcelar, Sabine, Vaibhav Jadhav, Michael Melcher, Norbert Auer, Astrid Hrdina, Rebecca Sagmeister, Kelley Heffner, Anja Puklowski, Michael Betenbaugh, Till Wenger, Friedrich Leisch, Martina Baumann, and Nicole Borth. 2018. "Karyotype Variation of CHO Host Cell Lines over Time in Culture Characterized by Chromosome Counting and Chromosome Painting." *Biotechnology and Bioengineering* 115 (1): 165–73.Wang, Wensheng, Patrick Brandt, Marie L. Rossi, Laura Lindsey-Boltz, Vladimir Podust, Ellen Fanning, Aziz Sancar, and Robert A. Bambara. 2004. "The Human Rad9-Rad1-Hus1 Checkpoint Complex Stimulates Flap Endonuclease 1." *Proceedings of the National Academy of Sciences of the United States of America* 101 (48): 16762–67.

Wang, Ziyan, Heng Lin, Fang Hua, and Zhuo-Wei Hu. 2013. "Repairing DNA Damage by XRCC6/KU70 Reverses TLR4-Deficiency-Worsened HCC Development via Restoring Senescence and Autophagic Flux." *Autophagy* 9 (6): 925–27.

Wurm, Florian M. 2004. "Production of Recombinant Protein Therapeutics in Cultivated Mammalian Cells." *Nature Biotechnology* 22 (11): 1393–98.

Xu, Jianlin, Peifeng Tang, Andrew Yongky, Barry Drew, Michael C. Borys, Shijie Liu, and Zheng Jian Li. 2019. "Systematic Development of Temperature Shift Strategies for Chinese Hamster Ovary Cells Based on Short Duration Cultures and Kinetic Modeling." *mAbs* 11 (1): 191–204.

Yang, T. T., P. Sinai, P. A. Kitts, and S. R. Kain. 1997. "Quantification of Gene Expression with a Secreted Alkaline Phosphatase Reporter System." *BioTechniques* 23 (6): 1110–14.

Yannone, S. M., S. Roy, D. W. Chan, M. B. Murphy, S. Huang, J. Campisi, and D. J. Chen. 2001. "Werner Syndrome Protein Is Regulated and Phosphorylated by DNA-Dependent Protein Kinase." *The Journal of Biological Chemistry* 276 (41): 38242–48.