CRISPR/CAS9-MEDIATED GENE EDITING IN HERDA EQUINE

by

Joseph R. Hawkes

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

of

MASTER OF SCIENCE

in

Animal, Dairy, and Veterinary Sciences

Approved:

Zhongde Wang, Ph.D.
Co-Major Professor

Dirk K. Vanderwall, D.V.M., Ph.D.
Co-Major Professor

Mirella Meyer-Ficca, Ph.D.
Committee Member

Richard S. Inouye, Ph.D.
Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2019
ABSTRACT

CRISPR/Cas9-mediated Gene Editing in HERDA Equine

by

Joseph R. Hawkes, Master of Science

Utah State University, 2019

Co-Major Professors: Drs. Zhongde Wang and Dirk Vanderwall
Department: Animal, Dairy, and Veterinary Sciences

To date, no research group has published gene editing in Equine. Heritable Equine Regional Dermal Asthenia (HERDA) is an autosomal recessive genetic skin disease found mainly in elite American Quarter Horse breeding lines. HERDA is caused by a single nucleotide mutation (c.115 G>A) in exon 1 of the peptidyl-prolyl Isomerase B (PPIB) gene. HERDA-affected horses are not suitable for performing and are oftentimes euthanized. Any prospect of “breeding out” HERDA from American Quarter Horses is unlikely, given the wide spread HERDA-causing mutation among elite American Quarter Horses.

Since no effective treatment exists for HERDA, it representing an opportunity to use gene editing to correct this genetic disease. In this thesis project, CRISPR/Cas9 and adenine base editor 7.10-max (ABEmax) were employed to edit the HERDA-causing mutation in equine fibroblasts. I designed a single guide RNA (sgRNA) to direct CRISPR/Cas9 to induce double strand breaks (DSBs) 15 nucleotides (nt) downstream of the mutation which achieved very efficient insertions or deletions (indels) (~90%). Assisted by a small 100nt. reverse compliment single-stranded donor oligonucleotide as
template, CRISPR/Cas9 edited the HERDA-causing mutation in equine fibroblasts with ~10% efficiency. Four single-cell derived colonies, out of 41, were established from transfected cells and genetically characterized, and showed both monoallelic and biallelic gene editing frequencies of 7.3% and 2.4% respectively. The colonies were cryopreserved for future animal cloning experiments to produce HERDA-free animals. Performance evaluation of the cloned animals may also answer the question on whether the HERDA mutation is linked to the extraordinary flexibility of joints and super performance of elite horses.

Lengthening the highly efficient sgRNA that was used with CRISPR/Cas9 by 2 nucleotides broadened the base editing window, giving ABEmax access to the HERDA-causing mutation at protospacer position 19. Horse fibroblasts were transfected with ABEmax, where single-cell colony selection and sequencing revealed seamless A to G conversion in 3 out of 20 single-cell colonies. Another 40% of isolated colonies may have been edited but carry an unspecific PCR band. Further investigation is needed to determine these genetic alterations.

(74 pages)
PUBLIC ABSTRACT

CRISPR/Cas9-mediated Gene Editing in HERDA Equine

Joseph R. Hawkes

HERDA (Heritable Equine Regional Dermal Asthenia) is a genetic skin disease mainly found in Quarter Horses, but also in Appaloosa and American Paint breeds. HERDA is similar to Ehlers-Danlos syndrome in humans, with symptoms including stretchy skin, hyperflexible joints, and, unique to HERDA equine, spontaneous skin sloughing. Horses affected by HERDA are not suitable for performing and are oftentimes euthanized. Some carriers for the HERDA-mutation are very competitive in the American Quarter Horse industry, especially in cutting events where it is believed, yet unproven, to give them an advantage with increased flexibility. It is also possible that the genomic locus (or loci) that links to the competitive performance traits is located close to the HERDA-causing mutation, which could lead to the co-segregation of this performance trait with the HERDA-causing mutation.

Direct-line breeding strategies in the last 30 years have increased the number of HERDA-affected equine causing this disease to increase in frequency among the Quarter Horse breed. Since no treatment exists for HERDA, owners often heavily invest in HERDA horses before the symptoms arise at around two years of age. These horses are often euthanized to alleviate pain and stress on the horse and to mitigate the costly upkeep by the owner.
HERDA-affected horses carry a homozygous single nucleotide mutation (c.115 G>A) in exon 1 of peptidyl-prolyl Isomerase B (PPIB). Gene editing approaches would be preferable for correcting this genetic disease, since it can precisely correct the mutation without altering any other genetic traits in the elite horse breeds that have been heavily selected for. By employing the CRISPR/Cas9 system, we have sought to correct the HERDA-causing mutation in the PPIB gene. The CRISPR/Cas9 system is comprised of a bacterial endonuclease protein called Cas9 and a guide RNA sequence to direct Cas9 to target the genome in a sequence-specific manner by introducing DNA double-strand breaks (DSBs). The introduction of DNA DSBs promotes the activation and recruitment of homologous recombination (HR)-mediated DNA repair machineries to repair the broken DNA; if oligonucleotides with the desired DNA sequence are co-delivered with the CRISPR/Cas9 system into cells, the HR-mediated DNA repair mechanism can replace the targeted sequence in the genome with the oligonucleotide’s sequence, therefore, achieving gene correction or editing. We designed sgRNAs to target genomic sequences in close vicinity of the HERDA-causing mutation and a single-stranded DNA oligonucleotide containing the normal (wild type) PPIB genotype. Co-delivery of the CRISPR/Cas9/sgRNA complex with the donor oligonucleotide has successfully led to the production of gene edited cells. We established single-cell derived colonies from the edited cells and achieved 7.3% monoallelic and 2.4% biallelic editing frequencies. The gene edited fibroblasts were cryopreserved as an initial step for future HERDA-free equine cloning projects to develop the first gene edited horses.
We also investigated novel approaches to correct the HERDA-causing mutation by means of Adenine Base Editors (ABEs). Optimized for mammalian expression, ABEmax is composed of the Cas9 nickase fused to TadA, a tRNA adenosine deaminase lab-evolved to catalyze A to G base conversions in DNA. Seamless base editing was achieved in 15% of single-cell derived colonies, correcting the HERDA-causing mutation, while unusual genotypes were detected in 40% of isolated colonies.
ACKNOWLEDGMENTS

I would like to thank my Master’s Degree advisors Dr. Zhongde Wang and Dirk Vanderwall for mentoring me during my research and studies at Utah State and providing me with this opportunity to use cutting-edge technology. Zhongde provided specific project counseling and troubleshooting in regards to gene editing approaches in HERDA equine. Dirk provided unique, hands-on experiences with horse sample collection and project support. They served diligently with Mirella Meyer-Ficca as my supervisory committee members; enthusiastically aiding my thesis development and encouraging me to derive more single-cell colonies which contributed significantly to the results in this thesis.

I’d like to thank Nick Robl and Kristy Boyce for developing the Max HERDA fibroblast cell line, finding published primers for PCR, initiating RFLP analysis procedures for diagnosing HERDA, and performing the sgRNA1 transfection and RFLP. I’d like to thank the summer undergraduate intern Jaman Al-Haddad who aided in xABE and REPAIR plasmid isolations and performed PCR reactions and enzyme digests.

I would like to thank the Wang lab postdoctoral researchers Rong Li and Yanan Liu for their contributions to this project and my laboratory technique training. Rong taught me how to do RFLP, transfections, cell culture, etc. Both he and Yanan offered invaluable comments, assistance, and feedback during critical times. They’re great friends that I hope to remain in contact with.

I’d like to thank my parents for encouraging me pursue a Master’s Degree at Utah State University. They helped me endure the research ups and downs; I owe so much love
and appreciation for their sustaining help and support to complete my work on this project.

The Utah State University ADVS department has helped me feel part of the team of graduate students. I appreciate their friendly environment and emphasis on preparing students for presentations and research in the science field.

I’d like to thank my wife, Annika who helped me practice presentations, brought me food, and provided tons of feedback and companionship.

I’d like to thank Feng Zhang, Omar Abudayyeh, Jonathan Gootenberg, Linyi Gao, David Liu, Johnny Hu, Nicole Gaudelli, and Dominik Paquet, who offered invaluable suggestions on alternative DNA editing methods found in this thesis, including REPAIR, constructing xABE and guide plasmids, ABE 7.10 and ABEmax, broadened window editing, and estimating CRISPR/Cas9 sgRNA3 editing efficiencies while offering ssODN guide design suggestions.

Joseph R. Hawkes
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>PUBLIC ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xiv</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>6</td>
</tr>
<tr>
<td>Heritable Equine Regional Dermal Asthenia</td>
<td>6</td>
</tr>
<tr>
<td>Histopathological Signs and Identification</td>
<td>8</td>
</tr>
<tr>
<td>Potential Environmental Influences on HERDA Equine</td>
<td>10</td>
</tr>
<tr>
<td>Palliative Treatments</td>
<td>11</td>
</tr>
<tr>
<td>HERDA-causing Mutation Origin</td>
<td>11</td>
</tr>
<tr>
<td>Genetically Testing Equine for HERDA</td>
<td>13</td>
</tr>
<tr>
<td>Molecular Basis for HERDA</td>
<td>14</td>
</tr>
<tr>
<td>Ehlers-Danlos Syndrome</td>
<td>16</td>
</tr>
<tr>
<td>CRISPR/Cas9</td>
<td>19</td>
</tr>
<tr>
<td>Engineering the Cas9 System</td>
<td>21</td>
</tr>
<tr>
<td>CRISPR/Cas9 DNA Breaks – DNA Repair</td>
<td>24</td>
</tr>
<tr>
<td>Adenine Base Editors</td>
<td>29</td>
</tr>
<tr>
<td>Cas13 mRNA Editor (REPAIR)</td>
<td>31</td>
</tr>
<tr>
<td>III. MATERIALS AND METHODS</td>
<td>33</td>
</tr>
</tbody>
</table>
IV. RESULTS ........................................................................................................... 41

CRISPR/Cas9-mediated Gene Editing in Fibroblasts ................................. 41

sgRNA Designs, PCR-RFLP, Donor Oligonucleotides .................41

Testing sgRNA Efficiency ................................................................. 42

Donor Oligo Design ................................................................. 42

Gene Editing Efficiency ................................................................. 43

Screening Single-cell Derived Colonies ................................. 44

ABEmax ....................................................................................... 45

V. DISCUSSION ......................................................................................... 48

VI. CONCLUSIONS .................................................................................. 50

VII. REFERENCES ...................................................................................... 52
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>List of Reagents ................................................................. 38</td>
</tr>
<tr>
<td>2</td>
<td>PCR Primers ........................................................................ 39</td>
</tr>
<tr>
<td>3</td>
<td>ssODN Designs ........................................................................ 40</td>
</tr>
<tr>
<td>4</td>
<td>Guide RNA Designs .................................................................. 42</td>
</tr>
<tr>
<td>5</td>
<td>Gene Edited HERDA-equine Single-Cell Colonies ......................... 45</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clinical signs of HERDA in Affected Equine</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Joint Hypermobility in HERDA-affected Equine</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Dermal Lesions Observable in HERDA-affected Equine</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>HERDA Collagen Fiber Organization</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Genotyping Horses for the HERDA-causing mutation</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>HERDA <em>PPIB</em> Protein Alignment</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td><em>PPIB</em> (CypB) Crystal Structure</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>Phenotypic Similarities between HERDA and Ehlers-Danlos Syndrome</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>Non-HERDA Clinical Signs of Ehlers-Danlos in Equine</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>Targeting Genomic DNA with CRISPR/Cas9</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>Indels from CRISPR/Cas9 Double-Strand Breaks</td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>Double Strand Break Repair Pathways</td>
<td>25</td>
</tr>
<tr>
<td>13</td>
<td>Homology directed repair after double-strand break formation</td>
<td>27</td>
</tr>
<tr>
<td>14</td>
<td>Estimated CRISPR/Cas9 Editing Frequencies</td>
<td>28</td>
</tr>
<tr>
<td>15</td>
<td>Adenine Base Editor Schematic Representation</td>
<td>30</td>
</tr>
<tr>
<td>16</td>
<td>Base Editing with Broadened Protospacers</td>
<td>31</td>
</tr>
<tr>
<td>17</td>
<td>Targeting the HERDA-causing Mutation with CRISPR/Cas9</td>
<td>41</td>
</tr>
<tr>
<td>18</td>
<td>CRISPR/Cas9 pooled DNA RFLP Analysis for Indels and Editing</td>
<td>43</td>
</tr>
<tr>
<td>19</td>
<td>CRISPR/Cas9 Single-cell Colony RFLP and Sequencing Analysis</td>
<td>44</td>
</tr>
<tr>
<td>20</td>
<td>Targeting the HERDA-causing Mutation with ABEmax</td>
<td>45</td>
</tr>
<tr>
<td>21</td>
<td>ABEmax Single-cell Colony RFLP and Sequencing Analysis</td>
<td>46</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

AQH = American Quarter Horse

AQHA = American Quarter Horse Association

CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats

CypB = Cyclophilin B

DSB = Double-Stranded Break

EDS = Ehlers-Danlos Syndrome

HDR = Homology-Directed Repair

HERDA = Heritable Equine Regional Dermal Asthenia

Indels = insertions or deletions

KI = Knock in

KO = Knock out

NHEJ = Non-Homologous End Joining

PAM = Protospacer Adjacent Motif

PPIB = peptidyl-prolyl isomerase B

REPAIR = RNA Editing for Programmable A to I Replacement

RNP = Ribonucleoprotein complex

sgRNA = single-guide RNA

xABE = x Adenine Base Editor

ABEmax = Adenine Base Editor max
CHAPTER I
INTRODUCTION

Hereditary Equine Regional Dermal Asthenia (HERDA) is an autosomal recessive disease of American Quarter horses, with the highest prevalence in reining and cutting horses (A. Rashmir-Raven, 2013). HERDA is caused by a single-nucleotide missense mutation (c.115G>A) in exon 1 of the peptidyl-prolyl Isomerase B (PPIB) gene which codes for cyclophilin B (CypB) protein (Tryon, White, & Bannasch, 2007). The HERDA-causing mutation alters CypB’s interaction with other collagen-folding enzymes, delaying collagen folding, decreasing collagen tensile strength, and propagating disorganized fiber packing (Bowser et al., 2014; Brounts, Rashmir-Raven, & Black, 2001; Grady, Elder, Ryan, Swiderski, & Rashmir-Raven, 2009; Ishikawa et al., 2012). Symptoms include fragile skin, prone to ulcerative lesions, spontaneous skin sloughing, hypermobility, subdermal hematomas, and decreased cornea thickness (Badial et al., 2015; A. M. Rashmir-Raven & Spier, 2015).

In Rashmir-Raven’s study (2013), HERDA-affected equines reportedly begin showing pathological lesions by age two, when the horse is saddled to begin riding, while some do not develop the disease until 4-5 years old, or occasionally, never develop typical skin lesions, though hyperelastic (i.e., “stretchy”) skin is present. She found that occasionally, mildly affected horses are sound enough to compete, while some, aided by great care, have lived into their late 20’s. Due to increased DNA testing availability and HERDA awareness, most horses with HERDA are humanely euthanized by age 3, but Rashmir-Raven notes that often, HERDA is diagnosed after considerable financial and
emotional investments have been made. She also reports that the number of horses affected by HERDA is increasing due to the common practice of line breeding carrier horses, to the point that today, HERDA is one of the most commonly reported inherited diseases in the equine industry. HERDA-affected Quarter Horses, Appaloosas, and American Paints are found throughout the world, reported in at least the United States, Canada, Mexico, Brazil, England, France and the Netherlands (Badial et al., 2015; Rendle, Durham, & Smith, 2010; Tryon et al., 2009; White & Bourdeau, 2011). Truly, HERDA is a world-wide issue without a cure.

Gene-editing technology provides an opportunity to correct the HERDA-causing mutation, as a solution for alleviating animal suffering without euthanasia and propagating these elite genetic carrier lines without HERDA. Gene editing of performance horses, either heterozygous or homozygous for the HERDA-causing mutation, could offer horse breeders more flexibility in selecting mating pairs as edited clones would be restored to wild-type, free of the mutation.

 Originally derived from microbial adaptive immunity, clustered regularly interspaced short palindromic repeats (CRISPR) has been harnessed to mediate targeted and efficient modification of virtually any location within the eukaryotic genome by a short RNA guide (Cong et al., 2013a). In complex with the guide RNA, CRISPR-associated endonuclease Cas9 can base pair with the sequence of interest, thereafter inducing double-strand breaks (DSBs) (Jinek et al., 2012). This DNA cleavage may result in insertions or deletions (indels) which disrupt the gene in an uncontrollable manner, or through the Homology Directed Repair Pathway, available with a sister chromatid or provided DNA template which edits the gene (Hsu, Lander,
Precise short gene editing (< 50bp) may be achieved by providing the cells with a short single-stranded oligodeoxynucleotide (ssODN), designed to have ~50-80bp homology arms of varying symmetry around the DSB and sense in relation to the PAM-containing strand (Paquet et al., 2016; Richardson, Ray, DeWitt, Curie, & Corn, 2016; Wang et al., 2016). For large insertions/deletions (>100bp), plasmid donors are commonly used with 800bp homology arms. Co-delivered with CRISPR/Cas9, the donor ssODN or plasmid instructs the cell repair machinery which mutations to integrate or retract from the cellular DNA (Yang et al., 2013).

In order to increase editing efficiency and reduce indels, base editor technology was developed by David Liu’s group at the Broad Institute of MIT and Harvard (Gaudelli et al., 2017). They describe the Adenine Base Editor (ABE) as a Cas9 nickase fused to Escherichia coli TadA, a tRNA adenosine deaminase, which acts on DNA instead of its usual RNA substrate. In the process of A to G conversion, they describe an intermediate called inosine, which is converted to guanosine along with the complimentary thymine-base converted to cytosine, after either DNA replication or repair. In order to target a specific adenine base for editing, ABE uses an NGG protospacer adjacent motif (PAM) and 20nt guide RNA to edit adenine bases within an editing window of positions 13-17 upstream of the PAM sequence.

ABE7.10 was the first ABE version developed by Gaudelli (2017) which catalyzed an A (adenine) to G (guanine) conversion in ~50% of human cells with very low INDEL rates (typically ≤0.1%). Although much more efficient than Cas9 nuclease and ssODN editing, Koblan (2018), from the same group, reported that ABE7.10 editing bottlenecks in the process of plasmid expression, which reduces base editing efficiency.
His group optimized ABE7.10 by modification of nuclear localization signals and codon usage, and ancestral reconstruction of the deaminase component. The resulting ABEmax editor corrects mutations with ~5 to 7-fold greater efficiency than ABE7.10 and generates indels rarely ($\leq 1.6\%$) although elevated from the virtually undetectable indel levels of ABE7.10.

Another base editing tool was developed by Feng Zhang’s group at the Broad Institute of MIT and Harvard to convert A to I (inosine) in RNA instead of DNA (Abudayyeh et al., 2017; Cox et al., 2017). They fused Cas13b RNA nuclease to ADAR, a human deaminase that catalyzes A to I conversion in RNA, naming it REPAIR (RNA Editing for Programmable A to I Replacement). They reported that REPAIR can be used to selectively edit adenosine to inosine in RNA when used in conjunction with a guide RNA where inosine is read as guanosine by translational machinery. Additionally, REPAIR does not require PAMs or operate within a base editing window, but rather requires modification and delivery of a 50bp guide RNA, where a base flip (C) at the target A directs base conversion optimally at protospacer position 37 (Matthews et al., 2016). Zhang’s group (2017), developed several versions of REAPIR where a mutation in the ADAR deaminase domain confers increased specificity, and Cas13b is truncated for adeno-associated virus (AAV) expression. With these modifications, Cas13b has powerful potential as a gene therapy tool.

This thesis project is the first to implement CRISPR/Cas9 to correct the HERDA-causing mutation in equine fibroblasts, in anticipation of cloning the first gene-edited horses by using the gene-edited fibroblasts as nuclear donors. In parallel,
research was also carried out to establish DNA and RNA editing techniques in equine for developing therapeutics for affected horses.
Hereditary Equine Regional Dermal Asthenia

Hereditary equine regional dermal asthenia (HERDA) is an autosomal recessive skin disorder that affects primarily American Quarter Horses (A. Rashmir-Raven, 2013). In a clinical study by Borges and associates (2005) in Brazil, they noted that HERDA is also known as hyperelastoid cutis, as it belongs to a group of inherited congenital connective tissue dysplasias. They reported clinical sign of the disease from three horses with HERDA, where they found that tender bilateral lesions of their trunk and lumbar regions cause the horse pain upon handling. They also demonstrated that these sensitive regions contain hyperelastic skin due to thinner and smaller collagen fibrils which are loosely organized. In White’s study (2014), it was reported that horses begin showing additional HERDA disease signs such as skin lesions, seromas, and hematomas by one and a half years of age, after saddling which commonly causes lesions along the dorsal aspect in HERDA-affected equine. Rashmir-Raven (2013) also found that HERDA-affected horses have an increased range of motion in their joints and greater ligament elasticity than other horses.
Figure 1. Clinical signs of HERDA in Affected Equine. A) Spontaneous skin sloughing in a 2-year-old Quarter Horse filly (Hrd/Hrd), typical of a horse severely affected with HERDA. B) Disfiguring scars on a 6-year-old Quarter Horse stallion with HERDA (Hrd/Hrd) C) An example of ‘doughy’ skin on a 3-year-old Quarter Horse gelding with HERDA (Hrd/Hrd). This horse was gelded at approximately 18 months of age without ill effects. D) An example of loose, hyperextensible skin on a 3-year-old Quarter Horse gelding with HERDA (Hrd/Hrd). This horse was also gelded at approximately 18 months of age without ill effects. (Rashmir-Raven and Spier, 2015).

Rashmir-Raven and Spier (2015) later demonstrated that HERDA-affected horses exhibit loose, stretchy, hyperextensible skin that is easily pulled away from the body (Figure 1D). They reported that affected areas of the abnormal skin may range from a small area of 2-3 cm in diameter to 80% of the horse’s body; some skin becomes so fragile that it sloughs off, leaving disfiguring scars, common among HERDA equine (Figure 1B). Also, they noted that oddly, some horses do not develop lesions until 4yrs old or later while other horses appear to never develop lesions, although they exhibit
hyperextensible skin and signs of discomfort when saddled. Their research indicates that males appear to be most severely affected. Although HERDA-affected equine show poor recovery from some wounds, successful healing may occur after routine or traumatic repair surgeries (Castori et al., 2012).

Rashmir-Raven (2013) reports that other skin conditions may mimic HERDA phenotype, especially Alopecia (horse baldness), often worrying the horse owners whether their horses are affected by HERDA. She and Brown (2015) list other skin conditions sometimes confused with HERDA, which include “Rain Rot” (bacterial skin infection), mud fever, sweet itch, hives, sunburn, warts, ringworm, mange, or lice.

![Figure 2. Joint Hypermobility in HERDA-affected (Hrd/Hrd) Equine.](image)

**Figure 2. Joint Hypermobility in HERDA-affected (Hrd/Hrd) Equine.** A yearling American Paint Colt undergoing two types of leg stretches, revealing an abnormal range of motion (Rashmir-Raven, 2013).

**Histopathological Signs and Identification**

In separate clinical studies, groups led by Badial (2015), Bowser (2014), Brinkman (2017), (Gardner, Arnoczky, & Lavagnino, 2011) and Grady (2009) reported that HERDA-affected equine may show signs of the disease not only in skin and joints but also the cornea, great vessels, tendinoligamentous structures, and heart valves. They
also noted that these horses have higher amounts of total soluble collagen than unaffected horses, osteoarthritic lesions, increased joint inflammation, greater incidence of corneal opacities, decreased corneal thickness, abnormal collagen fibril arrangement, and weaker (lower tensile strength) tendinoligamentous tissues and great vessels.

**Figure 3. Dermal Lesions Observable in HERDA-affected Equine.** A) Hematoma on the withers of an affected horse, caused by blood and serum accumulation in the upper-deep dermis. B) Dorsal skin ulcerations complicated by UV radiation and saddle riding (Tyron et al., 2007).

A group led by Brounts (2001) analyzed tissues from HERDA-affected equine, and found that thinning and fragmentation of the skin may result in separated skin layers where the space between layers fills with fluid, forming large blisters (bulla) and hematomas (Figure 3A). In some areas, the upper half of the skin is completely separated from the lower half of the deep dermis by a cleft-like space, known as zonal dermal separations (Figure 4C). Also present in the deep dermis are interstitial microhemorrhages and increased vascularity (Brounts et al., 2001).
Figure 4. HERDA Collagen Fiber Organization. Dermal skin biopsy from a 1yr. old filly affected with HERDA, fixed on a slide and imaged by electron microscopy (TEM; x54,000). A) Variations in collagen fiber diameter, a trait unique to HERDA. B) collagen fibers packed in a loose and unorganized fashion, C) cross-section revealing the separation of the superficial and deep dermis (Brounts et al., 2001).

Potential Environmental Influences on HERDA Equine

The reason for variation in affected horses’ clinical signs is partially unknown, while potential environmental influences may play a role in severity of the disease. UV light appears to increase severity of HERDA, by the disease’s manifestation on the dorsal midline where UV light hits more directly and heat is also more intense (Rashmir-Raven, 2013). UV light exposure disrupts the dermal extracellular matrix by activating matrix metalloproteinases (MMPs) including collagenases (Rittié & Fisher, 2002). Due to delayed collagen folding and loosely packed/disorganized collagen fibers, the MMPs may degrade collagen more rapidly than in humans (A. M. Rashmir-Raven & Spier, 2015). Rashmir-Raven (2013) reports that horses with dermal lesions often show signs of
recovery after being moved indoors for 2-4 weeks. She also notes that colder weather, UV light protective sheets, and northern climates also decrease the effects of HERDA (A. Rashmir-Raven, 2013).

**Palliative Treatments**

No effective treatment for HERDA exists (White, 2014). Rashmir-Raven and Spier (2015) suggest that palliative therapy, which focuses on preventing or relieving symptoms, for HERDA-affected equine, includes administering phenylbutazone and chondroprotective nutraceuticals. They mention that some horse owners strongly recommend supplementing HERDA horses with lysine, but its benefits are not medically confirmed. Rashmir-Raven (2013) suggests that supplementing horses with dietary copper and vitamin C may help manage HERDA, but concludes that minimizing trauma, careful wound management, and restriction from heat and sunlight are some of the best ways to care for HERDA-affected equine.

**HERDA Mutation Origin**

Hedgepeth (1990) explains that the Quarter Horse breed originated from crosses between English and Spanish horses as early as 1611 that were originally bred for their use in short quarter-mile races in colonial America. These horses were later interbred with the Mustang, where he adds that during the American Westward migration, these “Quarter Horses”, as they became known, were invaluable for their versatility, strength, and agility, as they were employed in ranches and cattle management.

Although HERDA was first reported in young Quarter horses in 1978, this mutation has been traced back to a Quarter Horse foundation sire, Poco Bueno who was
foaled in 1944 (A. Rashmir-Raven, 2013; Tilstra & Byers, 2002). Rashmir-Raven (2013) recounts that Poco Bueno was immensely popular, gentle, smart, and easy-to-handle. She and the American Quarter Horse Association (AQHA) (2010) report that he remained an influential sire during the 1940s, ‘50s and ‘60s—siring 405 registered foals, 50% of which were carriers for HERDA. The AQHA (2010) also records that 36 of his offspring were AQHA champions, 3 of which, as confirmed carriers, were inducted into the National Cutting Horse Association Hall of Fame, and later, in 1990, Poco Bueno was inducted posthumously.

A group led by Tryon (2007) concluded that robust breeding of HERDA carrier stallions has caused inbreeding loops, where many affected horses’ pedigrees trace back to the popular sires bred in the late 1990s, resulting in higher incidence HERDA-affected equine. Out of 5000 horses’ DNAs tested, 55 were found homozygous for the HERDA-causing mutation (Hrd/Hrd), whose lineages trace back to Poco Bueno. In a separate evaluation of 75 HERDA-affected horses, 100% exhibited consanguinity to Poco Bueno within seven generations (Rashmir-Raven et al., 2013). Today, carrier frequency of the HERDA-causing mutation is estimated at 3.5% in the overall American Quarter Horse population but estimated at 28.3% in elite cutting horses (Tryon et al., 2009). HERDA is now reported as one of the most common diseases in the equine industry, reported in Quarter Horses, Appaloosas, and American Paint Horses. In a study of Brazilian Quarter Horses, Borges’s group (2005) estimated HERDA allele and carrier frequencies at 2.9% and 5.8%, respectively (Badial et al., 2015; Rendle et al., 2010; Tryon et al., 2009; White & Bourdeau, 2011). HERDA is a world-wide issue without a cure that many horse owners encounter.
Genetically Testing Equine for HERDA

In order to prevent the spread of HERDA, the AQHA (2018) requires breeders to complete the DNA (parentage) and panel tests on stallions and report which mares were bred. They offer a five-panel genetic test which covers HERDA, glycogen branching enzyme deficiency, hyperkalemic periodic paralysis, malignant hyperthermia, and polysaccharide storage myopathy. Breeders may order a kit online through the AHQA website and mail in a DNA sample of about 50 tail hairs for genotyping, where samples are processed by the University of California-Davis for HERDA. Current panel tests cost $100 for AQHA members or $155 for nonmembers, and results may be obtained in about two weeks. Some universities offer HERDA genotyping services, diagnosing through Restriction Fragment Length Polymorphism (RFLP) analysis (University of California-Davis) where the difference in homologous DNA sequences can be detected by the presence of differing PCR fragments after restriction enzyme digest (Figure 5A), or sanger sequencing (Figure 5B) (Texas A&M University) (Heather & Chain, 2016) NCBI, 2017).

Although several groups refer to horse zygosity for the HERDA-causing mutation differently, this thesis project refers to the three genotypes as wild type (WT/WT); carriers (WT/Hrd), which are heterozygous for the HERDA-causing mutation, and (Hrd/Hrd) which are homozygous for the HERDA-causing mutation.
Figure 5. Genotyping Horses for the HERDA-causing mutation. A) Restriction Fragment Length Polymorphism (RFLP) analysis where DNA fragments are sorted by length in gel electrophoresis and B) Sanger Sequencing to identify the HERDA-causing mutation in equine. The horses Coralee, Eunice, and Eileen are wild type (WT/WT); Nike and Solo carry the mutation (WT/Hrd) and Max is homozygous for the mutation (Hrd/Hrd). The HERDA mutation is circled in red, the mutant glycine to arginine change (p.39G>R) is boxed in yellow. The blue arrow indicates the double peak for HERDA carriers.

Molecular Basis for HERDA

The HERDA-causing mutation, located on equine chromosome 1 (ECA1), was identified by Tryon’s group (2007), with fine-structure single nucleotide polymorphisms (SNPs) mapping. They compared the HERDA PPIB gene sequence to other species such as human, chimpanzee, dog, cow, mouse, and Quarter Horses (Figure 6). They identified the correlation between the HERDA-causing mutation (c.115G>A) and HERDA-affected individuals in results from assaying samples of 64 HERDA-affected horses with a control group of 1079 horses. Additionally, extracellular cyclophilins, such as cyclophilin B (cypB) of the PPIB gene, have been reported to be associated with procollagen processing and inflammatory response (Arora et al., 2005; Bachinger, Morris, & Davis,
1993; Smith, Ferreira, Hebert, Norris, & Sauk, 1995; Steinmann, Bruckner, & Superti-Furga, 1991).

**Figure 6. HERDA PPIB Protein Alignment.** Equine PPIB shares 97% identity with canine PPIB, while it is 88% identical with the six-mammalian species in this figure. *HERDA-causing mutation (c. 115G>A), results in a glycine to arginine change (p. 39G>R) in the putative N-terminal domain of the cyclophilin B protein, otherwise conserved in all other sequenced vertebrates. (Tryon et al., 2007).

CypB has been shown to play a major role in the biosynthesis of procollagens, as it interacts with the many collagen-modifying enzymes, such as P3H1-CRTAP-CypB complex, calnexin, calreticulin, calmegin, HSP47, and protein-disulfide isomerase (Chang, Barnes, Cabral, Bodurtha, & Marini, 2009; Ishikawa, Vranka, Wirz, Nagata, & Bächinger, 2008; Kozlov, Määtänen, Thomas, & Gehring, 2010; Smith et al., 1995; J. Zhang & Herscovitz, 2003). CypB participates in proline residue 3-hydroxylation, and after triple helix formation, CypB and other cyclophilins catalyze the rate-limiting step of peptide bond cis-trans isomerization (Bächinger, 1987; Bachinger et al., 1993; Steinmann et al., 1991). Procollagen then folds with the assistance of special chaperones including HSP47 and FKBP65 as likely candidates (Ishikawa et al., 2012, 2008; Makareeva & Leikin, 2007). Ishikawa’s group (2012) discovered that the HERDA-causing mut-CypB protein has a disordered side chain where Arg-6 predominantly occupies the position of normal lys-5 (Figure 7), displacing the main chain and preventing normal interaction with...
the P-domain of procollagen modifying enzymes and chaperones. They suggest that mutCypB’s interactions with collagen-folding chaperones results in delayed collagen folding in this rate-limiting step, affecting its secretion into the extracellular matrix, which was confirmed when they detected decreased hydroxylysine residues in collagen fibrils from HERDA-affected equine.

**Figure 7.** HERDA *PPIB* (CypB) Crystal Structure. Three superimposed crystal structures of horse WT CypB (green), horse HERDA mut-CypB (red), and human WT CypB (white) in a duplex with the proline-rich P-domain of calmodulin (yellow). Note the displacement of Lys-5 (green arrow) and occupation of normal Lys-5 position by Arg-6 (shown as sticks; indicated by orange arrow) in HERDA CypB, leading to significant displacement of the main chain and preventing normal interaction with calmodin’s P-domain. Although only calmodin P-domain is shown here in complex with CypB, the P-domains of several procollagen modifying enzymes and chaperones interact with CypB, such as calreticulin, lysyl hydroxylase 1 the P3H1-CRTAP-CypB complex, and others (Ishikawa et al., 2012)

**Ehlers-Danlos Syndrome**

HERDA is clinically similar to Ehlers-Danlos syndrome type VIA (EDSVI) in humans, which is categorized within a larger group of connective tissue fragility syndromes (A. Rashmir-Raven, 2013). Ehlers-Danlos Syndrome (EDS) is caused by mutations in genes encoding various collagen types, collagen-modifying enzymes, and
other critical components of the dermal extracellular matrix (Mao & Bristow, 2001a; Myllyharju & Kivirikko, 2004a; Seidler et al., 2006). EDSVI is caused by a deficiency of lysyl hydroxylase 1 due to \textit{PLOD1} gene mutations and is inherited autosomal recessively (Hautala, Heikkinen, Kivirikko, & Myllyla, 1993; Wilcox, 2003). EDSVI is characterized by hypotonia (muscle weakness) at birth; progressive kyphoscoliosis (abnormal posterior and sideways curvature of the spine); skin hyperelasticity and fragility; joint hypermobility and subluxations; microcornea; rupture of arteries and the eye globe; and/or osteopenia (bone weakness) (Beighton, De Paepe, Steinmann, Tsipouras, & Wenstrup, 1998). Occurrence of EDSVI in humans is estimated at 1:100,000 (Yeowell & Steinmann, 2014). Classical EDS of hypermobile and classical forms are most common, affecting between 1 in 5,000 to 20,000; other forms of EDS are rare, often linked to affected families (Tilstra & Byers, 2002).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8}
\caption{\textbf{Figure 8. Phenotypic Similarities between HERDA and Ehlers-Danlos Syndrome.} A) Hyperextensible skin in a human patient affected by Ehlers-Danlos syndrome (Malfait, Wenstrup, & De Paepe, 2010) and B) hyperelastic skin characteristics in HERDA-affected equine (White, 2014).}
\end{figure}

Mutations affecting enzymes involved in post-translational modification of collagens have been associated with autosomal recessive forms of EDS in humans and
with similar recessive syndromes in animals (Colige et al., 2002; Mao & Bristow, 2001b; Myllyharju & Kivirikko, 2004b). The most common mutations for Ehlers Danlos affected humans occur in collagen genes, with 75% occurring in Col5A1 gene, while null mutations in LEPRE1 and CRTAP cause severe recessive osteogenesis imperfecta (Marini, Cabral, & Barnes, 2010; Ritelli et al., 2013; Symoens et al., 2012).

Two novel mutations in the PPIB gene were recently reported by Jiang’s group (2017) in a rare pedigree from a Chinese cohort with autosomal recessive osteogenesis imperfecta type IX. With next-generation sequencing (NGS), they sequenced the parents’ whole exomes, revealing that the father carried a PPIB (c.25A>G) mutation in exon 1 and the mother carried a PPIB (c.509G>A) mutation in exon 4. These human cases of EDS hypermobility type (HEDS) represent the most similarity to the HERDA-causing mutation (c.115G>A) in location and type, where gene editing the HERDA-causing mutation in equine could potentially serve as a model for curing EDS in humans.

EDS has been observed in Quarter Horses, not affected by HERDA. In one study by Steelman’s group (2014), a 7-year-old black Quarter Horse gelding, in good body condition, had a 12-month history of skin lesions and scarring but did not exhibit joint hypermobility or ocular abnormalities normally associated with the HERDA, nor did it have any relation to HERDA-affected horses. The group performed a necropsy, which revealed fibrous lesions of the withers and tenting lesions on dorsal skin (Figure 9). Additionally, they found that collagen fibers were slightly fragmented and tissue layers were separated throughout the skin. Although they did not discover a genetic mutation as a cause for this case report, the group hypothesized that a novel and likely spontaneous mutation in the many collagen synthesizing, propeptide cleaving, or folding associated
proteins is responsible for the observed Ehlers-Danlos symptoms, and suggests that horses with similar clinical signs without the *PPIB* mutation should be classified as having EDS.

Figure 9. Non-HERDA Clinical Signs of Ehlers-Danlos in Equine. Gross lesions revealed by autopsy of a 7-year-old gelding with Ehlers Danlos syndrome, which tested negative for HERDA. A) Skin tenting resulting from lesions. B) Dermal separation characteristic of HERDA but observed in this equine with Ehlers Danlos (Steelman et al., 2014).

**CRISPR/Cas9**

Since HERDA is caused by the single-nucleotide mutation (c.115G>A) in exon 1 of *PPIB* gene, several gene editing tools may be employed to genetically reverse it (Paquet et al., 2016; Tryon et al., 2007). Type II CRISPR (clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR associated protein 9) is derived from the adaptive prokaryotic immune system and targets genomic DNA sites, guided by short single-stranded guide RNA (sgRNA). After Cas9 binds to target loci, it nicks both strands of DNA, generating double strand brakes (DSBs) which generally result in indels but with a donor oligonucleotide template present, can stimulate cellular repair machinery to incorporate desired edits to genomic DNA (Cong et al., 2013a). CRISPR/Cas9 has improved the efficiency of genetic engineering which replaces previous gene editing
tools like designer zinc fingers (ZFs), transcription activator-like effectors (TALENs) which required homodimer formation (McCarty, Kleiger, Eisenberg, & Smale, 2003).

Figure 10. Targeting Genomic DNA with CRISPR/Cas9. Schematic representation of CRISPR/Cas9 targeting genetic loci and inducing DSBs. After complexing with the guide RNA, Cas9 is directed to a gene of interest, cuts both DNA strands, and triggers DSB repair, which may result in indels or HDR which may introduce desired edits if an oligonucleotide donor is present (Ding, Li, Chen, & Xie, 2016).

CRISPR loci in the bacterial genome are composed of DNA repeat sequences (21-48bp) that are flanked with spacers (27-72bp), as observed in Thermophilus bacterial response to phage infection during milk fermentation (Labrie, Samson, & Moineau, 2010). Many phage-resistant bacteria carry CRISPR-Cas systems with multiple repeat-
spacers identical to the infectious phage genomic sequences; for example, Methanococcus jannaschii contains up to 20 CRISPR loci in its genome, and CRISPR loci can be transferred horizontally via conjugation (Deveau, Garneau, & Moineau, 2010). Defense against phages occurs in two phases: adaptation and interference (Agari et al., 2010; Horvath & Barrangou, 2010). Adaptation refers to the process of bacteria acquiring short fragments of invading phage DNA and integrating them into their genomes (Brouns et al., 1993; Browne, Li, Chong, & Littman, 2005; Bult et al., 1996). Interference refers to the Cas complex binding CRISPR RNA (crRNA) from the CRISPR locus and cleaving foreign DNA which matches the previously integrated spacers which threaten to integrate into the host’s DNA (Agari et al., 2010).

**Engineering the Cas9 System**

Cas9 is composed of six domains, REC I, REC II, Bridge Helix, PAM Interacting, HNH and RuvC. REC I binds guide RNA; the role of REC II is not yet well understood; the Bridge Helix domain is crucial for initiating cleavage activity upon binding of target DNA; the PAM Interacting domain confers PAM specificity for discriminating between self and nonself sequences, and is responsible for initiating binding to target DNA; and HNH and RuvC are nuclease domains that cut single-stranded DNA (Anders, Niewoehner, Duerst, & Jinek, 2014; Jinek et al., 2014; Nishimasu et al., 2014). The sgRNA is composed of two parts; tracrRNA which binds Cas9, paired to the crRNA, which base pairs with the target DNA, forming a heteroduplex (Cong et al., 2013; Jiang & Doudna, 2017). These two RNAs are oftentimes delivered as a single-guide RNA (sgRNA), where after complexing with Cas9, lead it to cut DNA at endogenous genomic loci adjacent to the PAM, as shown in human, mouse, and other mammalian species,
including the golden Syrian hamster as demonstrated by our laboratory (Cong et al., 2013a; Fan et al., 2014).

A critical step in genome editing is designing an efficient guide RNA to target the genomic site of interest while accounting for possible off-targets. Guide RNA scoring software is available for SpCas9, and some Cas9 variants, through many online tools which use metrics from large-scale, empirical data to optimize sgRNA design, to enable precise and efficient genome engineering (Doench et al., 2016). Recently, new rules for guide RNAs were discovered, where TT- and GCC-motifs located in the 4 PAM-proximal bases of the targeting sequence block CRISPR/Cas9-mediated gene editing (Graf, Li, Chu, & Rajewsky, 2019). Additional factors such as decreased chromatin accessibility, secondary structure formation, or potential off-target DNA sites may also affect the choice and efficiency of sgRNA design (Cong et al., 2013a).

Figure 11. Indels from CRISPR/Cas9 Double Strand Breaks. Insertions or deletions (indels) detected via sequencing, in EMX1 locus of human 293FT cells edited by human codon-optimized SpCas9 plasmid (Cong et al., 2013). The red dashes represent deletions, and red letter is an insertion. Note the PAM sequence overlined in red, Cas9 cleavage site indicated by the purple arrow, and the protospacer in blue.

The PAM is crucial for Cas9 binding genomic DNA, a sequence specific to the species of Cas9, but in most commonly used, SpCas9’s PAM is composed of three
nucleotides in the order NGG after the 3’ end of the protospacer sequence which occurs every 16bp on average within the human genome (Figure 11) (Cong et al., 2013; Hu et al., 2018). The frequency of the PAMs within a gene of interest is vital to the success of homology directed repair (HDR)-mediated gene editing, which require precise positioning and short cut-to-mutation distances (~10-20bp away), respectively, for efficient editing (Paquet et al., 2016). Due to the infrequency, or to overcome a great distance of NGG PAMs from genetic locus to be edited, Cass proteins from more than 20 bacterial species, requiring PAMs such as NGRRT, NGRRN, NNNNGATT, NNNNRYAC, NNAGAAW, or NAAAAC have mediated gene editing in a variety of cell types, while thousands more Cas proteins remain to be discovered (Amrani et al., 2018; Kim et al., 2017; Kleinstiver et al., 2015; Koonin & Makarova, 2013). These orthologues’ sequences are limited in their similarity, except for conserved HNH and RuvC nuclease domains, and vary highly in length (~900-1600 amino acid residues) (Jiang & Doudna, 2017). SpCas9 is the most commonly used Cas9 version with a length of 1,369 amino acids, which limit its application for therapeutics, where shorter orthologues such as Staphylococcus aureus Cas9 (SaCas9) which is 1kb shorter than Cas9 or Cas-X which was isolated from a non-pathogenic bacterium with the intent of reducing the human pre-existing adaptive immune response. These Cas traits are desirable therapeutic applications such as packaging into viral vectors like AAV (~4.7kb DNA storage) and preventing an immune response in those receiving treatment (Charlesworth et al., 2019; Kleinstiver et al., 2015; Liu et al., 2019; Wang et al., 2016).

Although many new Cas9 variants are discovered frequently, none are more efficient than SpCas9 in editing mammalian cells with a PAM less restrictive than NGG
(Hu et al., 2018). Engineered variants of SpCas9 such as Sniper-Cas9, xCas9-3.7, Sp-Cas9-NG and others have been developed by directed evolution, resulting in higher specificity to reduce off-targets, and/or greater PAM recognition flexibility, for increasing the number of genomic sites where genome editing may occur (Hu et al., 2018; Lee et al., 2018; Mazhar, Hill, & McAuliffe, 2018).

Another RNA-guided DNA-cutting nuclease known as Cas12a (Cpf-1) has been recently adapted for genome editing, which processes its own guide RNA and recognizes TTTV PAMs, generating staggered or “sticky” ends after DNA cleavage, 18-23nt distal to the target site (Yamano et al., 2016; Zetsche et al., 2015). Recently, Gao’s group (2017) has engineered Cas12a variants through structure-guided mutagenesis, which increased its targeting range three-fold, to one cleavage site per ~11bp in human coding sequences. These Cpf-1 TYCV and TATV PAM-recognizing variants show enhanced performance in human cells and high DNA-targeting specificity, qualities which are invaluable for editing NGG PAM-scarce loci (Gao et al., 2017).

CRISPR/Cas9 DNA Breaks – DNA Repair

One study in humans reported that DNA damage occurs naturally at an estimated 20,000-100,000 times a day, with DSBs being the most deleterious (Khanna et al., 2001). Originally observed by gamma irradiation, DSB repair was theorized to use broken ends with homologs and later proven to use sister chromatin as templates for repair (Orr-Weaver, Szostak, & Rothstein, 1981; Resnick, 1976). If not repaired, DSBs can induce apoptosis, loss of heterozygosity through copying sequences from sister chromatids, cell senescence, or cancer (Burma, Chen, & Chen, 2006).
The non-homologous end joining (NHEJ) pathway directs the religation of segregated DNA, while not requiring a homologous template. Due to the presence of repeats or microhomology, miscellaneous nucleotides at the DSB site are deleted to allow for rejoining of the strands (Figure 12) (Jasin & Rothstein, 2013; Weterings & Chen, 2008). The steps of NHEJ include (I) DNA end recognition, (II) DNA end bridging, (III) DNA end processing, and (IV) ligation of broken ends, which result in insertion or deletion of DNA bases at the DSB site, making NHEJ pathway a generally unfavorable pathway for genome editing (Davis & Chen, 2013).

![Figure 12. Double Strand Break Repair Pathways.](Jasin & Rothstein, 2013)

The Homology Directed Repair (HDR) pathway uses an undamaged DNA strand as a template for repairing the damaged strand, such as a sister chromatid or
synthetically-produced DNA (J. P. Zhang et al., 2017). The first step of HDR is 5’ to 3’ end resection, which leaves a 3’ ended single-stranded DNA (ssDNA) region (Kakarougkas & Jeggo, 2014). Replication protein A (RPA) quickly binds ssDNA tails to prevent secondary structure formation, followed by RAD51 tail binding, which displaces RPA, and promotes invasion onto the undamaged template and strand displacement, forming a D-loop structure which provides a 3’ end for DNA synthesis or gap filling by polymerase, resulting in a heteroduplex molecule called a double Holliday Junction (dHJ) (Figure 13) (Jasin & Rothstein, 2013; Kakarougkas & Jeggo, 2014). Many HDR avenues may follow, but to avoid cross-over events, endonucleases induce symmetrical incisions across the junctions, resulting in nicked duplex products (Shah Punatar, Martin, Wyatt, Chan, & West, 2017). The newly synthesized DNA strand disassociates from the D loop, dissolved by branch migration and topoisomerase activity of the BLM (Sgs1)/TOP3α/RMI1 complex, and anneals to the other DNA end, followed by end- ligation which finishes the DNA repair event (Jasin & Rothstein, 2013).
Figure 13. Homology Directed Repair after Double-strand Break Formation. (Shah Punatar et al., 2017).

The HDR pathway may be used to insert exogenous DNA sequences into the genome, enabling generation of site-specific knock-ins (Gutschner, Haemmerle, Genovese, Draetta, & Chin, 2016). HDR-mediated knock-in efficiencies may reach 10% in mammalian cells when used in conjunction with donor molecules which are necessary for providing a template for the cell’s repair machinery to read and insert altered gene sequences (Jasin & Rothstein, 2013). These homology donors may range from short single-strand oligodeoxynucleotides (ssODNs) with 50-100bp homology arms, to larger external linear sequences with circular or linearized donor template with 1kb or longer flanking arms to achieve knock-in of larger reporter genes, which has been achieved in many cell types, including human cells via CRISPR/Cas9-induced HDR (Chen et al.,
According to findings from three groups, led by Wang (2016), Richardson (2016), and Prykhozhij (2018), ssODN optimization involves several factors including length, sense, symmetry, PAM mutations and chemical modifications. These groups found that an optimal ssODN design included a length of 120nt., antisense orientation, asymmetric alignment to the DSB (77bp upstream homology and 43bp downstream), a 1nt. PAM mutation and phosphorothioate modifications to prevent endonuclease cleavage, increasing HDR 3 to 10-fold in Zebrafish gene editing experiments.

![Diagram of Estimated CRISPR/Cas9 Editing Frequencies](image)

**Figure 14. Estimated CRISPR/Cas9 Editing Frequencies.** As Cas9 cuts further away from the mutation of interest, the lower the probability of isolating biallelically edited zygotic clones. Heterozygous (blue) frequency climaxes at 10-15 bp away. Frequency of homozygous wild type (green) increases with distance as frequency of homozygous edited (red) decreases (Paquet et al., 2016).

In many instances, HDR is desired for knocking in a genetic sequence but it’s difficult to achieve due to lower efficiency when compared to NHEJ. In the case of single-nucleotide changes, HDR efficiencies depend largely on the distance-to-cut mutation, preferably <15bp (Figure 14) (Paquet et al., 2016). Other ways to increase
HDR events in cells include small molecule inhibitors of NHEJ, which are shown to increase point mutation generation by up to 4-fold in some cell lines (Riesenberg & Maricic, 2018). Alternatively, HDR booster chemicals SCR7, L755507, and Resveratrol increased HDR by nearly 2-3 fold, and in some cases as high as 50% in screened colonies (Li et al., 2017). Additionally, Integrated DNA Technologies (IDT) has developed a proprietary Alt-R HDR chemical enhancer (Schubert, Thommandru, and Wang, 2018). Although HDR enhancers and NHEJ inhibitors have increased HDR in many cases, these chemicals do not have universal effects on all cell types and targeted genes and may not increase efficiency or may alter cell growth (Riesenberg & Maricic, 2018).

**Adenine Base Editors**

Gaudelli (2017) from David Liu’s group at MIT developed an Adenine Base Editor which converts A·T to G·C pairs in genomic DNA without DSBs. Her 7th generation Adenine Base Editor (ABE7.10) contains a Cas9 nickase fused to Escherichia coli TadA, a tRNA adenosine deaminase, which was adapted through extensive directed evolution and protein engineering to act on DNA instead of its usual RNA substrate. She reported that ABE7.10 requires an NGG PAM and a 20nt guide RNA to target genomic loci to convert A to G within a base editing window which spans 5 nucleotides in protospacer positions 13-17 from the PAM and is sometimes referred to as window positions 4-8 on the PAM-containing strand (Figure 15A). She also described the mechanism for base conversion where Cas9n binds and nicks genomic DNA, and TadA hydrolytically deaminates adenosines to inosine within the editing window. In order to complete the base conversion, Gaudelli reported that during subsequent DNA repair or
replication, inosine is permanently replaced by guanosine and the complimentary thymine base is changed to cytosine, and virtually no indels are generated (typically ≤0.1%).

If no NGG PAMs are available to adjust the target adenine within the editing window, ABE 7.10 cannot be used to correct the mutation. To increase PAM flexibility, Hu (2018) from David Liu’s group engineered the SpCas9 so that it recognizes NG, GAA, and GAT PAMs, and fused it to TadA, resulting in xCas9(3.7)ABE7.10. He notes in the study that editing efficiencies with these candidate PAMs vary by selection and additionally by guide sequence design.

Additionally, to improve the already reliable ABE7.10, it was optimized by modification of nuclear localization signals and codon usage, and ancestral reconstruction of the deaminase component to increase plasmid expression Koblan (2018) from David Liu’s group. The resulting ABEmax editor corrects mutations ~5 to 7-fold more efficiency than ABE7.10 and generates indels rarely (≤1.6%) although elevated from the virtually undetectable indel levels of ABE7.10 (Koblan et al., 2018).

Figure 15. Adenine Base Editor 7.10 Editing Window (A) Adenine DNA Base Editor and (B) Adenine RNA Base Editor (REPAIR). DNA ABEs contain one WT TadA deaminase and one evolved TadA deaminase to act on DNA. Base editing window is shown for ABE 7.10 (Molla & Yang, 2019).
It was also discovered that the editing window may be widened to accommodate mutations which fall outside the editing window by Ryu’s group (2018). By adjusting the guide length 1-2nt for a desired position (Figure 16), ABE editing windows are broadened to protospacer positions 12-20. He showed that adenine base editing at the distant 18 and 19 protospacer positions reached 20% in HEK293T cells, demonstrating that the editing window may be adjusted to accommodate adenine base conversion at distant positions which are unattainable with a 20nt. guide RNA.

![Figure 16. Base Editing with Broadened Protospacers.](image)

**Figure 16. Base Editing with Broadened Protospacers.** A→G editing in HEK293T cells outside of conventional protospacer positions 12-17 by delivering lengthened sgRNAs. About 10-12% normal efficiency at protospacer position 19 (blue arrow), compared to that of the ideal target position 16. Since the HERDA mutation lies at position 19, delivery of 22nt. sgRNA could allow for base editing as high as 10% (Ryu et al., 2018).

**Cas13 mRNA Editor (REPAIR)**

In 2017, Gootenberg, Abudayyah, and Cox (2017) from Feng Zhang’s lab at MIT, developed an RNA base editing system composed of type VI CRISPR-Cas which...
contains the programmable single-effector RNA-guided RNase, Cas13 which is capable of robust knockdown and RNA editing by dCas13 to direct A to I conversion by Adenosine Deaminase that acts on RNA (ADAR) in mammalian cells, referred to as RNA Editing for Programmable A to I Replacement (REPAIR). They reported that Cas13 does not require a PAM to target RNA, but is differentially regulated by accessory proteins Csx27 and Csx28. They designed 51 nucleotide sgRNA to target mRNA transcripts with REPAIR, containing a C mismatch at the target A in the guide. This “base flip” directs the ADAR2 enzyme to deaminating the target A, exposed in a bulge or loop (Matthews et al., 2016). They reported that efficiency of REPAIR is dependent on the position of the target A in the guide sequence (as designated by the C base flip), where the 34th position from the direct repeat provides the best editing, and a 5’ G may be appended to cap the guide which increases efficiency.
CHAPTER III
MATERIALS AND METHODS

Cell Culture

Equine skin sample collections for genetic analysis and genome editing were taken from a privately owned HERDA-affected horse named Max, and a Utah State University owned horse named Badger (IACUC-2771). The biopsies were sanitized with ethanol, washed in PBS, submerged in DMEM while resectioned into tiny pieces, and placed in at 100mm cell culture dishes with 10% or 20% FBS DMEM for several days at 37°C, 5% CO2 to allow for outgrowth of fibroblasts from the tissue fragments (Table 1). The cells were then trypsinized, expanded in T-75 flasks, and cryopreserved.

Gene Targeting in HERDA Equine Fibroblasts

To target the PPIB genomic loci in HERDA equine fibroblasts, RNA guides were designed with Benchling life science data management/collaboration platform tool (https://www.benchling.com). Primers for sgRNA synthesis and oligonucleotide repair templates for the targeting sites were synthesized by Integrated DNA Technologies (IDT, Iowa, USA). sgRNAs were synthesized “in house” via GeneArt™ Precision gRNA Synthesis Kit (Thermo Fisher Scientific) according to the manufacturer’s protocol. 2 μg sgRNAs were complexed with 2 μg Cas9 protein (Trucut©, Thermo Fisher Scientific) at room temperature for 15-30 minutes, forming the ribonucleoprotein (RNP) (Table 1). 5x10⁵ HERDA fibroblasts were transfected using AMAXA 4D nucleofection (Program EN-150) according to the manufacturer’s protocol with sgRNA3 RNP and ssODN RC7 (PAGE purified single stranded DNA donor oligonucleotide (IDT) 100uM final concentration) (Table 3). The CRISPR/Cas9 RNPs and ssODN RC7 were transfected into
fibroblasts derived from a HERDA-affected horse named Max (Hrd/Hrd). Negative controls were derived from untransfected Max cells and positive controls from non-affected and untransfected horse (WT/WT) fibroblasts.

Cells were grown for 2 days, after which half of cells were harvested for genotyping. Genomic DNA was extracted with the DNeasy blood and Tissue Kit (Qiagen, Cat No./ID: 69506) according to the manufacturer’s protocol (Table 1).

The *PPIB* gene containing the HERDA mutation was amplified using PPIBex1F forward primer and PPIBex1R reverse primer (Table 2). Each 50 μl PCR reaction was assembled with 0.25 Ex Taq Polymerase (Clonetech) according to the manufacturer’s instructions using 1 μM forward and reverse Pub1 primers. PCR reactions were carried out under the following conditions: 98°C for 2 min, then 32 cycles of 98°C for 10 sec, 60°C for 30 sec, 72 for 30sec, then final extension 72°C for 5 minutes. PCR products were verified by comparison to DNA standard (1-kb Plus DNA Ladder) and to WT or HERDA controls on a 2% agarose gel with SYBR Safe DNA stain (Life Technologies) (Table 1). 2-3 μl of PCR product were digested with Stu1 (Eco 1417) (Thermo Fisher Scientific) at 37°C for 15 minutes to check for editing, or the with Taa1 (Thermo Fisher Scientific) with the same parameters to check for indels. PCR fragments were loaded onto a 2% agarose gel and compared to WT and HERDA control digests along with 1-kb Plus DNA ladder (Table 1). Since Stu1 cuts PCR product containing HERDA, bands at 250bp were classified as WT, and those lower at 120bp were considered HERDA. Gel image analysis was performed using Image J software (1.47p, NIH) to quantify band relative intensity for knockout and editing efficiencies. SYNTHEGO ICE, a CRISPR
poly peaks parser analysis tool, was used to found in mixed PCR sequences for estimating INDEL species in single colonies or pooled PCR.

**HERDA Single-Colony Selection:**

One-third of transfected cells were detached 2 days after transfection and pipetted into 2mL fresh media for counting and establishing single-cell derived colonies by limiting dilution, and the other two-thirds were cryopreserved. The average cell count was performed with a hemocytometer, where the appropriate volume was pipetted into 50mL fresh 20% FBS DMEM (Thermo Fisher Scientific) and distributed to 5 96-well plates (Thermo Fisher Scientific) (Table 1). Cells were incubated for 10-15 days and screened for single-cell colonies within cell-positive wells. Upon reaching at least 80% confluency, cells were transferred to 24-well plates, and then to 6-well plates where one-half of the cells were harvested for genotyping. If cells had monoallelic or biallelic editing, 5x10^5 cells were cryopreserved per cyrovial by addition of 10% DMSO to 90% DMEM (20%FBS), and placed in a -80°C freezer, then transferred to liquid nitrogen the next day.

**Gene Targeting with Base Editors xABE and ABEmax**

xCas9(3.7)-ABE(7.10) plasmid (Addgene #108382), pCMV ABEmax plasmid (Addgene #112095), guide vector PFYF 1320 EGFP Site#1 (Addgene plasmid #47511), and pX330-SpCas9-NG (Addgene plasmid #117919) were purchased from Addgene (https://www.addgene.org) (Table 1). These plasmids were received as bacterial stabs. To amplify these plasmids, bacterial stabs were spread on agarose dishes (+amp) and incubated at 37°C for 12 hours. Colonies were selected and inoculated in 2ml amp+ LB broth and cultured at 37°C with shaking at 225rpm for 12 hours. xABE and pFYF1320
plasmid-carrying colonies were isolated by Gene JET Plasmid Miniprep kit (Thermo Scientific) according to manufacturer’s instructions, verified by enzyme digest and gel electrophoresis (Table 1). Plasmids for sgRNA expression were constructed using one-piece blunt-end ligation of a PCR product with the KLD kit (New England Biolabs) according to the manufacture’s protocol (Table 1). PCR was performed using Ex Taq Polymerase (Clonetech) with guide sequence-containing primers and the guide expression plasmid pFYF1320 as a template according to the manufacturer’s instructions. DNA vector amplification was carried out using single-use JM109 competent cells (Promega). Bacteria were transformed, cultured, and verified according to the manufacturer’s instructions and as stated above.

**Plasmid Transfection:** 2 μg plasmid (xABE, ABEmax, or pX330-SpCas9-NG, and 2.5 μg of sgRNA expression plasmid were mixed with 100 μl P3 solution (LONZA) and transfected into 5x10⁵ HERDA cells according to LONZA protocol; cells were cultured and genotyped as stated according to methods stated above. The same procedure was carried out with 2 μl GFP plasmid (LONZA) as a positive control, where although not quantified, nearly every fibroblast was fluorescent under blue/UV light.

**REPAIR Plasmid Preparations:** pC0054-CMV-dPspCas13b-longlinker-ADAR2DD(E488Q/T375G) (Addgene plasmid # 103870), pC0043-PspCas13b crRNA backbone (Addgene plasmid # 103854), and pC0039-CMV-dPspCas13b-GS-ADAR2DD(E488Q) (Addgene plasmid # 103849) were purchased from Addgene (https://www.addgene.org) (Table 1). pC0043 guide expression vector was linearized in the following 100 μl reaction: 1500 ng BbsI-linearized vector DNA (pc0043), 10 μl cutsmart buffer, 2 μl BbsI-HF, 85 μl H2O. The reaction was incubated at 37°C for 1 hour
and inactivated at 65°C for 20 minutes. Digest with BbsI was purified using QIAquick® PCR purification kit (Ref# 28106). For ligation, primers were ordered from IDT, melted at 95°C for 2 minutes, then annealed by cooling to 25°C on bench top. Ligation reaction included 2 µl T4 DNA ligase buffer, 50 ng vector DNA (pc0043 crRNA backbone), 2.5 ng annealed primers, 1 µl T4 DNA ligase (M0202), and 14 µl water (New England Biolabs). Concentration of insert oligos was calculated using the NEB calculator. Ligation proceeded for 10 minutes on benchtop at room temperature and was inactivated by incubation for 10 minutes at 65°C. 5 µl ligation product was transformed into JM109 competent cells (Promega) according to manufacturer’s instructions and incubated overnight at 37°C with shaking (225rpm) on amp+ agarose. Colonies were selected and inoculated in 2ml LB broth and cultured at 37° with shaking at 225rpm for 12 hours. Plasmids were isolated with the Gene JET Plasmid Miniprep kit (Thermo Scientific) according to manufacturer’s instructions and verified by enzyme digest, gel electrophoresis, and sequencing (ACGT Inc.; m13R primer 5’-GGAAACAGCTATGACCATG- 3’). Colony 3.3 (positive for insert into pc0043) was added to 50ml LB broth and incubated overnight at 37°C with shaking (225rpm). The plasmid was purified from the bacteria using Plasmid Plus Midi kit (Qiagen, Ref# 12943).

**REPAIR Transfection:** 1.5 µg pC0054 or pC039 and 4 µg of sgRNA expression plasmid (pC0043+insert) were mixed with 100 µl P3 solution (LONZA) and 5x10⁵ HERDA equine cells and transfected according to LONZA protocol and repeated with half the amount of plasmid. Cells were harvested after 48 hours in order to harvest optimal amounts expected edited RNA transcripts. Medium was removed and cells were
lysed by addition of 300 μl lysis buffer with 2-mercaptoethanol. Lysate was homogenized by passing through an 18-gauge syringe needle. RNA was purified from lysate using the PureLink® RNA Mini Kit (Life technologies, Cat# 12183018A) according to manufacturer’s instructions. RNA was treated with DNase I after purification to assure highly pure RNA without genomic DNA contamination (5 μl RNA, 2 μl DNaseI+buffer, 3 μl water, room temperature incubation for 15min, 65°C for 10 min). cDNA was synthesized using the SuperScript™ III First-Strand Synthesis Super Mix according to the manufacturer’s instructions (Invitrogen, Thermo Fisher Scientific). Each 20 μl PCR reaction was assembled with 0.2 μl Ex Taq Polymerase (Clonetech) according to the manufacturer’s instructions (2 μl buffer, 1.6 μl dNTP) using primers PPIBF2-RT, PPIBR5-RT (Table 2) and 2 μl cDNA from previous reaction (unquantified). PCR reactions were carried out under the same conditions as described above and genotyped in the same manner.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Company</th>
<th>Catalogue #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cas9 protein (Truecut©)</td>
<td>Thermo Fisher Scientific</td>
<td>A36498</td>
</tr>
<tr>
<td>GeneArt™ Precision gRNA Synthesis Kit</td>
<td>Thermo Fisher Scientific</td>
<td>A29377</td>
</tr>
<tr>
<td>KLD kit (two tubes: enzyme and buffer)</td>
<td>New England Biolabs</td>
<td>M0554S</td>
</tr>
<tr>
<td>SuperScript™ III First-Strand Synthesis Super Mix</td>
<td>Invitrogen, Thermo Fisher Scientific</td>
<td>18080400</td>
</tr>
<tr>
<td>pFYF1320 EGFP Site#1</td>
<td>Addgene</td>
<td>47511</td>
</tr>
<tr>
<td>pAsCpf1(TYCV)(BB) (pY211)</td>
<td>Addgene</td>
<td>89352</td>
</tr>
<tr>
<td>pC0043-PspCas13b crRNA backbone</td>
<td>Addgene</td>
<td>103854</td>
</tr>
<tr>
<td>pC0039-CMV-dPspCas13b-GS-ADAR2DD(E488Q)</td>
<td>Addgene</td>
<td>103849</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>pC0054-CMV-dPspCas13b-longlinker-ADAR2DD(E488Q/T375G)</td>
<td>Addgene</td>
<td></td>
</tr>
<tr>
<td>xCas9(3.7)-ABE(7.10)</td>
<td>Addgene</td>
<td></td>
</tr>
<tr>
<td>pCMV_ABEmax</td>
<td>Addgene</td>
<td></td>
</tr>
<tr>
<td>pX330-SpCas9-NG</td>
<td>Addgene</td>
<td></td>
</tr>
<tr>
<td>StuI (Eco 1471)</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Taa1</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>BbsI</td>
<td>New England Biolabs</td>
<td></td>
</tr>
<tr>
<td>PureLink™ RNA Mini Kit</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Agarose A</td>
<td>Bio Basic</td>
<td></td>
</tr>
<tr>
<td>DNA ladder</td>
<td>Clonetech</td>
<td></td>
</tr>
<tr>
<td>Nunc™ Biobanking and Cell Culture Cryogenic Tubes</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Trypsin-EDTA (0.05%), phenol red</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>DMEM</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>DPBS</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>100mm Cell culture dish</td>
<td>Falcon, Life Sciences</td>
<td></td>
</tr>
<tr>
<td>FBS</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>DMSO</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Gene JET Plasmid Miniprep kit</td>
<td>Thermo Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>JM109 competent cells</td>
<td>Promega</td>
<td></td>
</tr>
<tr>
<td>T4 DNA ligase</td>
<td>New England Biolabs</td>
<td></td>
</tr>
<tr>
<td>96-well Clear Flat Bottom TC-treated Culture Microplate</td>
<td>Falcon, Life Sciences</td>
<td></td>
</tr>
<tr>
<td>LB Agar, Miller (Powder)</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Amplicillin</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. List of Reagents.** Reagents used in experiments with respective companies and catalogue numbers.

**Table 2. PCR primers**
<table>
<thead>
<tr>
<th>ssODN Name</th>
<th>Sequence</th>
<th>length (nt.)</th>
<th>symmetry</th>
<th>sense</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC7</td>
<td>5’−CAGGCAGGTCTGGAGAGGGAGGCC AGATCTTGACAGTGACTTTAGGCCCTT CTTCTTCTCATCGGCCGTGGAGGGTCCTG GCAACAGCAGGAAGAAGA−3’</td>
<td>100</td>
<td>asymmetric</td>
<td>Anti-sense</td>
</tr>
<tr>
<td>*120nt</td>
<td>5’−AGGACCCTCCACGGCCGATGAGAAG AAGAAGGGCCCTAAAGTCACCTGTCAAA GTCTGGCCTCCCCCTCTCCAGACCTGCGGG TCGAGGCCCGCCCTGGGGTGCCGGCCTCG GTGCCGCTGGT−3’</td>
<td>120</td>
<td>asymmetric</td>
<td>sense</td>
</tr>
</tbody>
</table>

Table 3. ssODN Designs. Sequences of ssODNs used in fibroblast transfections for editing the HERDA-causing mutation in fibroblasts. *this ssODN was not shown to mediate editing in pooled RFLP analysis, but may yet be useful for increasing HDR efficiency since it was a late-stage ssODN design (Wang et al., 2016) for sgRNA3 that never had single-cell colonies derived from a transfection.
RESULTS

CRISPR/Cas9-mediated Gene Editing in Fibroblasts

sgRNA Designs, PCR-RFLP, Donor Oligonucleotides

The first approach used to edit the HERDA (c.115G>A) mutation in fibroblasts was with CRISPR/Cas9 and ssODN templates. Since HDR efficiency in gene editing decreases dramatically when DSBs are generated >15bp from the mutation, we designed sgRNA1 (Table 4), which recognizes an AGG PAM resultant from the HERDA-causing mutation, to generate indels 3bp upstream from the mutation. Reported as a unit less on/off target score where values closer to 100/100 represent the best guide, sgRNA 3 had the best predicted on/off target score, designed to generate indels 15bp downstream of the HERDA mutation (Benchling guide scorer) (Table 4, Figure 17).

Figure 17. Targeting the HERDA-causing Mutation with CRISPR/Cas9. A segment of genomic DNA from the horse PPIB gene that contains the HERDA-causing mutation (red). Annotations include sgRNA sequences (highlighted in corresponding colors), their corresponding PAM locations (underlined in brown).
<table>
<thead>
<tr>
<th>Guide name</th>
<th>Sequence (5’ to 3’)</th>
<th>Mutation-to-cut distance (bp)</th>
<th>Predicted on/off target scores</th>
<th>Observed indel efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>sgRNA1</td>
<td>CGGCCGATGAGAAGAAGAAG</td>
<td>3</td>
<td>59/26</td>
<td>none</td>
</tr>
<tr>
<td>sgRNA3</td>
<td>GAGGCCTAAAGTCACTGTCA</td>
<td>15</td>
<td>69/51</td>
<td>90%</td>
</tr>
<tr>
<td>sgRNA4</td>
<td>AGGCCTCCTTCTTCTTCTCAT</td>
<td>11</td>
<td>53/34</td>
<td>none</td>
</tr>
<tr>
<td>sgRNA7</td>
<td>GCTGTGGCCAGGACCCTCCA</td>
<td>23</td>
<td>62/32</td>
<td>15%</td>
</tr>
<tr>
<td>PX330-NG-4</td>
<td>GAGAAGAAGAAGAGGCCCTAA</td>
<td>4</td>
<td>n/a</td>
<td>55%</td>
</tr>
<tr>
<td>sgRNA3x22</td>
<td>AAGAGGCCTAAAGTCACTGTCA</td>
<td>15</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 4. Guide RNA Designs. Details of sgRNAs tested in equine fibroblasts with the HERDA-causing mutation.

Testing sgRNA Efficiency

A major challenge to designing sgRNAs for editing the HERDA-causing mutation was that three GAA repeats are found upstream of the mutation, resembling a figurative NGG PAM “desert”. We also suppose that these GAA repeats formed secondary structures when incorporated into sgRNAs, preventing sgRNA1 and sgRNA4, which spanned these GAA repeats, from generating indels (analyzed via RFLP). SgRNA 3 was selected for gene editing because it generated indels 15bp downstream of the mutation, with high efficiency (Figure 18C), while sgRNA 7 was the only other guide to generate indels but at much further distance of 23 nucleotides from the HERDA-causing mutation.

Donor Oligo Design

Several oligonucleotides were designed for editing the HERDA-causing mutation according to the findings of the Richardson (2016), Kankan (2016), and Paquet (2016)
groups, for gene editing (Table 3). These oligonucleotides differed by length, symmetry around the DSB, and sense with respect to the PAM-containing strand. An antisense asymmetric ssODN donor template with 69 and 31 nucleotide homology arms was designed, named RC7 (Reverse compliment 7). This ssODN carried two mutations: R39G to edit HERDA and the sgRNA3 PAM-blocking mutation (AGG→AGA) (Figure 18A).

**Gene Editing Efficiency**

CRISPR/Cas9 and ssODN (sgRNA3 and RC7) were transfected in fibroblasts established from a HERDA-affected horse, named Max, which is homozygous for the HERDA-causing mutation. RFLP analysis revealed that sgRNA 3 indel formation was 90% efficient (Figure 18C) while fragments resistant to Stu1 reached 20% (Figure 18D), indicating that editing occurred.
Figure 18. CRISPR/Cas9 pooled DNA RFLP Analysis for Indels and Editing. A) Genomic sequence and donor oligonucleotide RC7 (antisense, asymmetric) shown for site PPIB G39R editing. B) sgRNA3 and ssODN RC7-transfected HERDA pooled PCR product with WT and HERDA (Max) controls C) Taal enzyme digest which is unaffected by HERDA point mutation editing or the PAM-blocking mutation, revealing indels only, as uncut upper band and D) modified (diluted to 30 μl) Stu1 digest, revealing editing and large indels as uncut upper band (blue arrow).

Screening Single-cell Derived Colonies

In order to determine point mutation editing efficiency in fibroblasts, 41 single-cell colonies were isolated and analyzed with RFLP and confirmed by sequencing. Colonies 3D6, 1B6, and 3D5 had monoallelic editing and one colony, 3B2, had biallelic editing. These colonies were cryopreserved for future experiments (Table 5). Overall editing efficiency observed in single colony selection was 9.8% (4/41), while monoallelic and biallelic frequencies were 7.3% and 2.4% respectively.
Figure 19. CRISPR/Cas9 Single-Cell Colony RFLP Analysis and Sequencing. A) colony 3D6 is monoallelically edited, B) 3B2 is biallelically edited, C) 1B6 is monoallelically edited, and D) 3D5 (although small peaks) likely monoallelically edited. RFLP fibroblast Controls: WT = WT/WT alleles, Hondo = WT/Hrd alleles, Max = Hrd/Hrd alleles. Unlabeled lanes indicate unedited or indel-containing colonies. Blue arrows and indicate point mutation editing (circled in red); green arrows indicate the PAM-blocking mutation editing (circled in purple).

<table>
<thead>
<tr>
<th>Colony Name</th>
<th>Genotype</th>
<th>Vials Cryopreserved</th>
<th>Cells per cryovial</th>
<th>Confluency at cryopreservation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D6</td>
<td>WT/4-6bp Deletion</td>
<td>5</td>
<td>5x10⁴</td>
<td>90%</td>
<td>Very fast growth and good morphology</td>
</tr>
<tr>
<td>3B2</td>
<td>WT/WT</td>
<td>2</td>
<td>5x10³</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>3D5</td>
<td>WT/HRD</td>
<td>2</td>
<td>5x10⁴</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>1B6</td>
<td>WT/1bp Deletion</td>
<td>1</td>
<td>5x10⁵</td>
<td>60%</td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Gene Edited HERDA-equine Single-Cell Colonies. Colonies’ names and genotype derived from sgRNA 3 and RC7 ssODN transfection and single-colony selection.

ABEmax

Since sgRNA3 was highly efficient (90%), it was also used to base edit the HERDA-causing mutation as the guide for ABEmax. sgRNA3 was lengthened by two nucleotides for broadening the editing window to positions 12-19. This guide RNA modification gave ABEmax access to the HERDA-causing mutation at position 19 (Figure 20).

Figure 20. ABEmax Editing Diagram. The equine PPIB genomic sequence to be edited with ABEmax contains the sgRNA3x22 sequence (annotated in green) and its
corresponding PAM location (underlined in brown). The curly bracket indicates the expanded base editing window location (positions 12-19) with the lengthened 22 nucleotide sgRNA. Complimentary strands are separated to demonstrate base editing window.

After transfecting fibroblasts carrying homozygous HERDA mutation, gene editing was not detected in pooled cells with the PCR-RFLP assay (Figure 21A) but after deriving 20 single-cell colonies, monoallelic base editing was observed in three colonies and confirmed by Sanger sequencing (Figures 21C and D). Seven other colonies were also resistant to Stu1 digestion but unfortunately showed an unspecific PCR band just below 250bp also apparent after Stu1 digestion (Figures 20B and C). When sequenced, mixed peaks made the chromatogram difficult to interpret, indicating that the PCR products contain an unrelated sequence and are unspecific bands, which can be further investigated through TA cloning.
Figure 21. ABEmax Single-cell Colony RFLP and Sequencing Analysis. A) Pooled RFLP for transfection with ABEmax and sgRNA3x22 showing no HERDA-edited cells. B) PCR product from single and mixed colonies isolated from a 96-well plate. C) Digest with Stu1 for editing detection. Colonies 1, 17, 19 show monoallelic editing. Colonies 3-6, 12-14 show editing and another possible indels (extra band). D) Sequence results from single-cell colonies 1, 17, and 19 show double-peaks at the HERDA mutation, revealing monoallelic adenine base editing, without indels. HERDA is circled in red and the blue arrow indicates the double-peaks characteristic of heterozygotes.
DISCUSSION

HERDA is an increasingly problematic genetic disease that remains difficult to resolve in elite Quarter Horses throughout the world. “Breeding out” the disease by carefully genotyping registered Quarter Horses for the HERDA-causing mutation may reduce the spread of the genetic disease, but given the HERDA-causing mutation frequency among elite American Quarter Horses, it is unlikely. Also, since there are no effective treatments for HERDA-affected horses, they are oftentimes euthanized. Gene-editing technology provides an opportunity to correct the HERDA-causing mutation.

In my thesis research, two gene editing strategies were successful in correcting the HERDA-causing mutation in horse fibroblasts cultured from a HERDA-affected horse. First, CRISPR/Cas9 nuclease with an ssODN template corrected either one or both alleles in ~10% of transfected cells. Sanger sequencing confirmed that three single-cell colonies derived from the transfection were edited monoallelically, and one biallelically. These four colonies were cryopreserved with the goal of cloning the first gene edited horses in this breeding line without HERDA. The second strategy used was adenine base editing where ABEmax monoallelically corrected HERDA in 3 single-cell colonies derived from a transfection. Seven other colonies showed likely editing, in which case total ABEmax editing efficiency would be 55%, but an unspecific PCR band prevented their genotyping by sequencing. Further investigation is needed to characterize the genetic alterations in these colonies.

Transfection via nucleofection maximized the efficiency of the guide RNA used for CRISPR/Cas9, generating 90% indels. Detecting editing in pooled cells with the
PCR-RFLP assay was challenging, and determining editing efficiency from single-cell colonies proved more reliable, but also time and work intensive. It is possible that gene edited loci are selected against during PCR and/or digested in concentrated Fast Enzyme Digest reactions. When enzyme digests were diluted to 30ul reactions, digests from pooled DNA showed editing. Nevertheless, increased editing efficiency is paramount to correcting HERDA in affected equine.

Novel CRISPR tools emerge at an accelerated rate, but their applicability to many species, including equine, is unknown. We employed emerging gene editing techniques to genetically edit equine fibroblasts homozygous for the HERDA-causing mutation and demonstrated that both CRISPR/Cas9 combined with an ssODN and ABEmax are efficient in editing the equine genome.

Although development may be costly to develop for actual treatment, we foresee their use in gene editing fibroblasts established from champion Quarter Horses which carry the HERDA-causing mutation followed by cloning animals from using the mutation-corrected fibroblasts as nuclear donors. Cloning these performance horses could offer horse breeders more flexibility in selecting mating pairs as edited clones could propagate these elite genetic lines without instances of HERDA.
CONCLUSIONS

The *PPIB* HERDA (c.115 G>A) mutation was efficiently targeted by CRISPR/Cas9, generating 90% indels with sgRNA3 with a 15bp mutation-to-cut distance. Cas9 nuclease-mediated editing reached 10% in HERDA-affected equine fibroblasts. ABEmax was also used to target the HERDA mutation with sgRNA3, lengthened 2nt. (sgRNA3x22). This allowed for using the highly efficient sgRNA3 with ABEmax to edit HERDA at protospacer position 19, observed at 15% editing in single-cell colonies, but may be as high as 55% upon further analysis of unspecific PCR bands.

Other new technologies that emerged during this project were tested in HERDA-affected equine fibroblasts, including other Adenine Base Editors—xABE and REPAIR, Nureki’s SpCas9-NG, and the Cpf-1 TYCV variant. Editing was not achieved with these tools, except SpCas9-NG, which generated 50% indels only 4bp from the HERDA-causing mutation. Single-cell colonies from SpCas9-NG and ssODN transfections are needed to determine editing efficiency, which were not analyzed due to time constraints.

Out of the 41 single-cell colonies derived from the CRISPR/Cas9 ssODN transfection, four were edited—three monoallelically and one biallelically. These four colonies were cryopreserved for later use in somatic cell nuclear transfer. Editing the HERDA-causing mutation in horses may demonstrate whether the mutation is responsible for high performance or if other genetic components independent of the mutation contribute to their high performance. This precise gene editing could also provide an alternative to “breeding out” the HERDA-causing mutation by retaining elite performance horse genetics in edited clones, used for breeding.
Additionally, HERDA gene therapies based on CRISPR/Cas9 and Adenine Base Editor technology, which corrected the mutation in horse fibroblasts, could improve affected equine health and performance. Restoring normal collagen folding, even at low rates, could help affected horses perform in the American Quarter Horse industry, thus reducing animal suffering and investment loss by owners.
REFERENCES


Smith, T., Ferreira, L. R., Hebert, C., Norris, K., & Sauk, J. J. (1995). Hsp47 and cyclophilin B traverse the endoplasmic reticulum with procollagen into pre-Golgi intermediate vesicles: A role for Hsp47 and cyclophilin B in the export of procollagen from the endoplasmic reticulum. *Journal of Biological Chemistry, 270*(31), 18323–18328. https://doi.org/10.1074/jbc.270.31.18323


