

https://dspace.org/

An In Ovo expression system for therapeutic proteins

Lemmer, Caitlin

2023
https://hdl.handle.net/10539/35720

Downloaded from a DSpace repository. https://dspace.org/



An *In Ovo* expression system for therapeutic proteins

by

Caitlin Lemmer

1601518

Dissertation

Submitted in fulfilment of the requirements for the degree

Master of Science

in

Molecular and Cell Biology

in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Prof. Natalya Nikitina

January 2023

Plagiarism declaration

I, **Caitlin Lemmer** (**1601518**), am a student registered for the degree of Masters of Science by Dissertation in the academic year 2022.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the research proposal submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature_  __ 15 January 2023

Acknowledgements

I would like to first acknowledge my supervisor, Professor Natalya Nikitina, for all the guidance and support she provided throughout my MSc, while still encouraging me to gain independence in the lab. In addition to this, she provided me with a project, which may not have been known at the time, which fulfilled everything that I desired to gain from this experience. I have had the privilege to gain exposure to multiple techniques that are relevant across many fields as well as knowledge, that although seems to contribute to only a small branch of genetics on the surface level, has its roots in many areas of biological science. Additionally, I would like to thank Zane Oberholzer for providing the red fluorescence protein plasmid used to test injection and trace editing in the cells.

I would also like to thank the University of the Witwatersrand for providing me with the funds to complete my degree.

Lastly, I would like to thank my friends and family, especially my mom, for their patience and support throughout this journey.

Abstract

With the global therapeutic protein market estimated to reach 316 billion USD by 2025, it is to no surprise that there is a peaked interest in research into alternative protein bioreactors. Currently, Chinese Hamster Ovarian cells are the most widely used vector for therapeutic expression, however, chickens present a potentially higher yielding and more cost-effective vehicle compared to this mammalian system. To fully utilize this avian bioreactor, gene drive technology in combination with CRISPR Cas9 needs to be employed. This will allow for the targeted editing of the Ovalbumin and Ovomucoid loci allowing for the maximum expression of a therapeutic protein in the albumin. To accomplish this, we first constructed a gene drive, with restriction digest, that contained CRISPR Cas9, the therapeutic protein human clotting Factor 9 (FIX), homology arms, and a meiosis-specific promoter. Primordial Germ Cells (PGCs) were then isolated from the blood of HH 14-17 embryos and cultured for 4 months to be the vehicle for gene transfer into the chicken. After this, kill curves were performed and the best concentration for selection with puromycin and neomycin of PGCs, was determined. The best concentration of busulfan for depleting native PGCs in gonads while conserving viability as much as possible, was also determined through IHC staining. The gene drive constructs were then electroporated and lipofected into PGCs and integration was assessed with Homology Directed Repair and Non-Homologous end joining specific primers. Separately, PGCs edited with an mRFP-Tol2 tracer were injected into HH 13-15 embryos and repopulation was assessed on these busulfan ablated embryos. The use of 75µg of busulfan in 50µl of sesame seed oil and DMSO was determined to be most effective at ablating native PGCs and conserving viability while 1.5µg/ml and 400µg/ml was determined to be the ideal selection concentration for puromycin and neomycin, respectively. While integration into the PGC genome was unsuccessful, transfection conditions were established that successfully transferred the constructs into PGCs. Future work will need to include treatment with HDR promoting factors in cell culture or gRNA redesign to ensure successful integration of the gene drive into PGCs. To address unsuccessful repopulation, different methods of injecting, such as using a microinjector, could be attempted or the technique could be repeated with more cells. Nonetheless, this study provides important preliminary data to create a transgenic chicken bioreactor for the expression of FIX at a lower cost than current methods.

Contents

Plagiarism declaration....2

Acknowledgements....3

Abstract....4

List of figures....8

List of tables....10

List of abbreviations....11

1.Introduction....13

1.1 Background: Therapeutic proteins as a multibillion-dollar market and common bioreactors used for production....13

1.2 The chicken as a bioreactor and major routes for genome modification....18

1.3 Transgenesis in chickens for protein production....23

1.4 Methods of PGC extraction and cell culture for creating transgenics....28

1.5 Transfection methods for PGCs with naked constructs....31

1.6 Sterilisation of chickens to remove native germ cells....32

1.7 A word on Hemophilia B and Human FIX as a therapeutic....35

1.8 Aim and objectives....37

Objective 1....37

Objective 2....37

Objective 3....37

Objective 4....37

Objective 5....37

2. Methods....38

2.1 Bioinformatics....38

2.2 Construction of the gene drive....38

2.2.1 Bacterial culture and plasmid extraction....38

2.2.2 Digestion and ligation of gene drive fragments....40

2.2.3 NEBuilder construction of the antibiotic resistance genes 42
2.2.4 Gel extraction and PCR cleanup for the purification of gene drive fragments 45
2.2.5 Order of cloning 46
2.3 Animal ethics and general egg care 59
2.4 Cell culture and establishing primary cell lines 59
2.4.1 Establishing a CEF primary culture and cryopreservation of CEFs 59
2.4.2 Preparation of the CEF feeder layer 60
2.4.3 Establishing PGC culture from germinal crescent extraction 60
2.4.4 Establishing PGC culture from blood extraction 61
2.4.5 Passaging of PGCs 62
2.4.6 Staining to identify PGCs 62
2.5 Editing of primordial germ cells *in vitro* 63
2.5.1 Kill curves for neomycin and puromycin 63
2.5.2 Transfection of PGCs 63
2.5.3 Assessment of gene drive integration in PGCs 64
2.6 Gonadal repopulation of partially sterilized embryos with edited PGCs 66
2.6.1 Busulfan ablation of native PGCs 66
2.6.2 Repopulation with edited PGCs 68
2.7 Statistical analysis 68
3. Results 69
3.1 Restriction mapping demonstrates the successful construction of the gene drives 69
3.2 Microscopy and ICC staining to identifying PGCs in culture 75
3.2.1 Germinal crescent PGCs successfully identified in culture by EMA-1 staining 75
3.2.2 Blood PGCs successfully identified in culture with SSEA-1 staining and morphological characteristics 76
3.3 IHC analysis reveals the best concentration of busulfan for ablation of native PGCs and viability conservation 78
3.4 Optimal concentrations for the selection of PGCs with neomycin and puromycin determined using kill curve plots 82

3.5 PCR analysis reveals the successful transfection of PGCs and the absence of genome modification at the OVA and OVM loci 84
3.6 mRFP-Tol2 tracer analysis reveals that repopulation of embryo gonads with edited PGCs is largely unsuccessful 89
4. Discussion 91
4.1 The successfully altered gene drive is ready for use in downstream modification experiments 91
4.2 A line of PGCs is successfully established allowing for editing of the chicken germline 93
4.3 Busulfan is effective at eliminating native PGCs and preparing embryos for repopulation 94
4.4 Neomycin and puromycin successfully select for edited PGCs allowing for a means to remove unedited cells from culture 95
4.5 PGCs are successfully transfected for the OVA and OVM gene drives but do not undergo genome modification 96
4.6 Re-population of embryo gonads requires optimisation 96
4.7 Improvements, future work, and conclusion 97
5. References 100
6. Supplementary 110
6.1 Sequencing for gene drives 110
6.2 Sexing gels 117
6.3 CEF culture 118
6.4 Defects found at various concentrations of busulfan treatment 118
7. Recipes 119
Recipe 1:Terrific Broth 119
Recipe 2: Crude extraction lysis buffer 119
Recipe 3: Ringer's solution 119

List of figures

Figure 1: Homology directed repair within the chicken genome in the presence of a gene drive expressing CRISPR Cas9. 25
Figure 2: Mendelian inheritance as compared with Super-Mendelian inheritance brought about by a CRISPR gene drive. 27
Figure 3: Overview of the process of PGC culture and transgenic creation. 31
Figure 4: The original plasmids (A, B and C) used to amplify the necessary fragments to create the final Neomycin and puromycin resistance genes that were inserted into the OVA and OVM gene drives, respectively (D and E). 45
Figure 5: Brief overview of the cloning strategy used for the OVM, and OVA derived gene drives. 47
Figure 6: Cloning step for insertion of FIX into the OVA-HA gene drive. 49
Figure 7: Cloning step for insertion of Sycp1 into the OVA-HA-FIX gene drive. 50
Figure 8: Cloning step for insertion of FIX into the OVM-Syc gene drive. 51
Figure 9: Cloning step for insertion of the ssOVM gRNA cassette into the OVM-FIX-Syc gene drive. 53
Figure 10: Cloning step for insertion of the puromycin resistance gene into the OVA-HA-FIX-Syc gene drive. 55
Figure 11: Cloning step for insertion of the neomycin resistance gene into the OVM-FIX-Syc-ss gene drive. 56
Figure 12: Cloning step for the insertion of the homology arms into the OVM-FIX-Syc-ss-Neo gene drive. 58
Figure 13: Stage 15 White Leghorn chicken embryo indicating where blood was drawn from for PGC extraction. 62
Figure 14: Image demonstrating where the needle is positioned when piercing the egg to inject busulfan for ablation purposes. 67
Figure 15: Digestions confirming the removal of FIX from p Mini T, the removal of Tenecteplase from the gene drive and the successful insertion of FIX and presence of Sycp1 in the OVM-Syc gene drive after cloning. 70
Figure 16: Digestions confirming the removal of CMV and Tenecteplase and the successful insertion of FIX and Sycp1 in the OVA-HA-Syc-FIX gene drive after cloning. 71
Figure 17: Digests of the OVM-Syc-FIX gene drive to confirm the removal of the old cassette and insertion of the new ssOVM gRNA cassette. 72
Figure 18: Successful amplification of the products needed for NEBuilder ligation of the antibiotic resistance genes for neomycin and puromycin.. 73

Figure 19: Restriction digests for both OVA-HA-Syc-FIX-Puro and OVM-Syc-FIX-ss-Neo gene drives to confirm the removal of zeocin resistance and the insertion of puromycin and neomycin resistance genes, respectively. 73
Figure 20: Digestion of the OVM gene drive in order to insert the 5` and 3` homology arms and confirmation thereof. 74
Figure 21: Staining of germinal crescent PGCs to confirm their presence in culture. 75
Figure 22: Blood PGCs on a CEF feeder layer stained with DAPI and SSEA-1 as well as images monitoring morphology and colony formation over 3 weeks of culture. 77
Figure 23: Right gonads of male embryos treated with various busulfan concentrations, extracted at day 7.5 (HH stage 31-33). 79
Figure 24: IHC analysis on the ablation efficiency of busulfan at various concentrations and assessment of viability. 82
Figure 25: Kill curve demonstrating the effect of puromycin on PGC and CEF viability after 5 days of treatment. 83
Figure 26: Kill curve demonstrating neomycin treatment effect on viability of cells after 7 days of treatment. 84
Figure 27: mRFP-Tol2 edited PGCs observed under a fluorescent microscope. 85
Figure 28: mRFP-Tol2 and gene drive edited PGCs on a CEF feeder layer. 86
Figure 29: PCR analysis of circular transfected gene drives for genome integration into PGCs. 87
Figure 30: PCR analysis of linearized transfected gene drives for genome integration in PGCs. 89
Figure 31: Gonadal repopulation with the RFP tracer construct (mRFP-Tol2) of HH 13-15 male chick embryos. 90
Figure 32: FIX sequencing results for the OVA-HA-FIX gene drive. 111
Figure 33: Sycp1 sequencing results for the OVA-HA-Syc-FIX gene drive. 111
Figure 34: Puromycin resistance gene sequencing results for the OVA-HA-Syc-FIX-Puro gene drive. 112
Figure 35: Sycp1 and FIX sequencing results for the OVM-Syc-FIX gene drive 115
Figure 36: Neomycin resistance gene sequencing results for the OVM-Syc-FIX-ss-Neo gene drive. 116
Figure 37: ssOVM g RNA cassette sequencing results for the OVM-Syc-FIX-ss gene drive 117
Figure 38: Sexing gel for PGC extraction. 118
Figure 39: CEFs plated in a T75cm^2 flask one day after harvest. 118

List of tables

Table 1: Comparison between various non-mammalian and mammalian bioreactor systems for human therapeutic protein production. 18
Table 2: Reaction set up for all restriction digestions. 41
Table 3: Reaction set up for ligations peformed for the insertion of various fragments into the gene drives. 41
Table 4: Primers used to amplify the various fragments for NEBuilder assembly of the antibiotic resistance genes for neomycin and puromycin. 42
Table 5: PCR conditions for the various NEBuilder antibiotic resistance gene fragments amplified with Q5 DNA polymerase. 43
Table 6: PCR conditons for overlap extension PCR for joing the individual puromycin fragments and individual neomycin fragments. 44
Table 7: Nomenclature used to represent the gene drive at each cloning step and what change was implemented in the gene drive 47
Table 8: PCR reaction set-up for sexing of embryos. 61
Table 9: PCR reaction conditions for sexing PCR 61
Table 10: Primers and PCR conditions used to screen edited PGCs for gene drive integration at the 5` and 3` homology arms as well as to screen the different gene drives for NHEJ insertion 65
Table 11: Primers and PCR conditions used to screen edited PGCs for gendrive presence as well as eSpCas9-T2A-Puro presence 66
Table 12: Defects recorded across embryos exposed to different concentrations of busulfan. 119

List of abbreviations

bFGF	Basic Fibroblast Growth Factor
BHK21	Baby Hamster Kidney Fibroblasts
BRL	Buffalo Rat Liver
Cas9	CRISPR associated 9
CEFs	Chicken Embryonic Fibroblasts
CHO	Chinese Hamster Ovarian
cPGCs	Circulating PGCs
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Dazl	Deleted in azoospermia gene
DDX04	Dead Box Protein 4
DMF	NN Dimethyl Formamide
DMSO	Dimethyl Sulfoxide
E. coli	*Escherichia coli*
EGFP	Enhanced Green Fluorescent Protein
EMA-1	Embryonic Mouse Antigen 1
FACs	Fluorescent Activated Cell sorting
Fkbp	fk506-binding protein
FIX	Factor 9
FX	Factor 10
FVI	Factor 6
FVIII	Factor 8
GD	Gene drive
gRNA	Guide RNA
HDR	Homology Directed Repair
hFIX	Human Factor 9
iFNβ	Interferon Beta
HH	Hamburger Hamilton
HIV	Human Immunodeficiency Virus
hSCF	Human stem cell factor
kb	Kilo base
Lac	Lactose
LIF	Leukemia Inhibitory Factor

MACs	Microfluids Assisted Cell screening
MEGA	Molecular Evolutionary Genetics Analysis
mgstII	Microsomal glutathione s-transferase 2
MUSCLE	Multiple sequence alignment
NHEJ	Non homologous end joining
NS0	Non-secreting 0
OVA	Ovalbumin
OVM	Ovomucoid
OVT	Ovotransferrin
pDNA	Plasmid DNA
PGCs	Primordial germ cells
RFP	Red Fluorescent Protein
Sycp1	Synaptonemal complex 1
SSCs	Spermatogonial stem cells
SSEA-1	Stage specific mouse embryonic antigen
ssOVM	Spider silk Ovomucoid
STO	Sandos inbred mice (SIM) 6-thioguanine-resistant ouabain-resistant
TALEN	Transcription activator-like effector and a nuclease domain
TF	Tissue factor
tPA	Tenecteplase
VASA	VASA

1.Introduction

1.1 Background: Therapeutic proteins as a multibillion-dollar market and common bioreactors used for production

The market for biopharmaceuticals, such as therapeutic proteins, monoclonal antibodies and vaccines, is rapidly expanding and is predicted to reach $316 billion by 2025 [1]. With therapeutic proteins being the fastest growing sector within the biopharmaceuticals market [2–4], it has garnered a lot of attention from the transgenic bioreactor research industry. This is especially true since the cost of protein therapeutics is largely unaffordable due to production line bottlenecks such as maintenance and running costs of expression facilities[2,3]. These therapeutics are used to treat a range of diseases and ailments, such as post trauma and post operative bleeding, and Hemophilia [5–7]. Unfortunately, due to the unaffordability of these therapeutics, majority of the population can be deprived of treatment. A transgenic bioreactor can offer a reprieve to the healthcare sector that must meet the demands of an ever-expanding population, only if it can produce a product at a cost that suites the majority.

Originally, therapeutic proteins used in the treatment of blood disorders were obtained directly from human blood, however, there have been issues with purification, and the presence of blood borne viruses left in the sample pose considerable risk. Chemical synthesis has, therefore, been considered as a non-human means to producing products that safely eliminates the risk of viral infection, however, it is complicated as reagent handling and disposal need to be considered. In addition to the cost that these requirements entail, the use of reagents makes it significantly more expensive, thus, an animal bioreactor would provide a promising, cheaper alternative[8,9]. To solve the issues around affordability and disease risk, mammalian production hosts, where protein is expressed under the dominant milk promoters in the mammary glands, have been thoroughly investigated. For all mammalian bioreactors, therapeutic protein expression is controlled under the αs1 and βcasein promoters, which are responsible for the expression of 73% of protein in milk[10,11].

Goats are one of the most recently reviewed mammalian expression systems that have the potential to increase protein yield. On average these mammals have a gestation period of 5 months and reach reproductive age at 8 months after which they produce 2 offspring. Once fertilized, goats produce roughly 600 liters of milk during their 300-day lactation period. The highest recorded production of therapeutic protein in goats' milk was 20 mg/ml for α1-proteinase inhibitor over the entire lactation period. While large expression volumes in this system have proven to be successful, the gestation and sexual maturation timeline,

unfortunately, provide significant delay for therapeutic protein production, making them infeasible[11,10,12].

A smaller mammalian alternative to goats, that possesses a 0.75-month gestation time and takes 1 month to reach sexual maturity, are mice. While the gestation time is significantly less limiting in this system, they only produce around 0.0015 liters of milk per lactation period of 3 weeks. The maximum therapeutic protein yield recorded for this period was human alpha 1-proteinase inhibitor at 35mg/ml[11,10,12]. While the yield in this system has the potential to reduce therapeutic costs, there are behavioral and physical limits mice, and goats, impose and it is laborious and expensive to induce transgenesis in these animals. The most commonly used method for creating transgenics, is microinjection of foreign DNA into harvested oocyte nuclei. This has a low success rate (on average 3%) in offspring and the technique is invasive to the cell. Microinjection is also time consuming and requires special equipment which is costly and limiting[12] further dampening the prospects for both these expression systems.

Other potentially higher yielding mammalian bioreactors that have been explored are rabbits and cows, however, high husbandry costs for maintaining rabbits and the 3-year long gestation to sexual maturity period for cattle often exclude them from further consideration. Rabbits have a 1-month gestation time and take 5 months to reach sexual maturity and produce approximately 1.25 liters per 21-day lactation period. While yield prospects for therapeutic protein production in rabbit milk is promising, with the maximum yield recorded at 8 mg/ml human α-glucosidase [11–14], production of antibodies is still unaccomplished due to the need to inactivate endogenous immunoglobulin loci. Despite this hurdle, researchers managed to demonstrate the production of 200μg/ml monoclonal antibody B as a concept, which is insufficient for the effort and time it takes to create this system. In addition to this, the high husbandry costs associated with rabbits unfortunately make them less appealing for protein expression as a whole[11].

Perhaps offering a larger potential yield but still limited by rearing and maintenance costs, are cows. In order to produce therapeutics in bovine, fertilized oocytes removed from slaughtered animals are obtained and pronuclei are injected with DNA. After 7 days *in vitro*, these edited oocytes are transferred to a host. This method is incredibly labor intensive and time consuming from obtaining the oocyte to implantation [13]. Even though the method of transgenesis is limiting, cows could theoretically yield 10mg/ml over a 305 day lactation period in which they produce 1120l [14]. Despite this large expression potential, cows have

only been recorded to yield 5 mg/ml of alpha-lactalbumin[10], meaning the time and effort to create this transgenic far outweigh the yield this system could provide and, thus it has not been further investigated.

Non-mammalian production systems, such as bacteria, have also been investigated as they offer a significantly cheaper alternative to the abovementioned mammalian systems. *E. coli* are commonly considered the cheapest and fastest way to produce proteins as they require minimal maintenance, have a doubling time of 20 minutes, they can be grown in huge bioreactors with simple media, and can be edited easily and cheaply. Bacteria also have the advantage of containing well characterized promoters for the expression of exogenous protein that can easily be targeted with restriction digest or CRISPR. While non-mammalian glycosylation patterns are the major issue in these expression systems, and can affect the function of the therapeutic negatively, correct glycosylation patterns can be engineered into *E. coli* to potentially evade a non-functional end product. This, however, can only be performed on already folded proteins, which *E. coli* frequently have difficulty with, especially with proteins greater than 30KDa [15–17]. Therapeutic products can also be chemically glycosylated through pegylation; however, this would be incredibly time consuming and expensive, especially to perform at a scale that meets the market demands. Nevertheless, *E. coli* have always been seen as the cheapest and easiest way to scale up protein production. The highest yield for protein produced in *E. coli* was 2 mg/ml of human antibody fragment [18]. While scale-up is theoretically possible in most cases, there are often issues that arise when products form insoluble build-up and lead to bacterial death, thus, limiting scale-up in some instances [19,20] .

The other easily manipulated and cost-effective non-mammalian alternative are yeast. They have a doubling time of approximately 2 hours and can be grown on a large scale in a bioreactor tank, similar to *E. coli*. Yeast are more capable of glycosylating similar to mammalian systems, however, they lack the ability to add terminal sugar residues to the final product which in turn reduces the half-life of the therapeutic. While half-life can be improved through the addition of functional groups, this would again be time consuming and expensive[19,21]. Another downfall to this production system is that the media composition for yeast causes the degradation of proteins and there is often insoluble protein formation and incorrect folding in the final product. Yeast have been recorded to produce up to 12 mg/ml of tetanus toxin fragment C in strains of *P. pastoris*[22], which would make it an attractive expression system if mammalian folding and post translational modifications were not of importance to the therapeutic.

Other non-mammalian alternatives, such as insects, are significantly more stress resistant compared to most systems and can, thus be used for large scale expression. Unfortunately the main issue is that insects add glycans not present in humans, thus, triggering an immune response, making them largely infeasible for human therapeutic expression[23]. In addition to this, while potential yield may be large, the highest recorded yield was only 0.3mg/ml for human collagenase IV, which would in turn drive up the cost of the therapeutic[19].

Plants have recently been investigated in order to find a host that produces a large amount of protein at a low cost while maintaining the necessary post translational modifications. They offer the potential of easy scale-up and cheap maintenance while allowing for more human compatible glycosylation patterns[24]. They are capable of producing approximately 100 tons of material hectare^{-1} per year^{-1}. The most common method of modifying plants includes using a gene gun or *Agrobacterium tumefaciens* transformation to reach the nuclear genome for nuclear expression. While editing is made simple through these technologies, gene silencing, and low expression levels are a common side effect of these random integration techniques and the nucleus alone is not sufficient to provide a worthy increase in protein yield when compared to mammalian systems [23,25]. Another benefit, however, to using plant expression systems is that other organelles can be used for gene modification and thus potentially increase yield. Chloroplasts are one of these organelles and the main appeal is that they are numerous and could, therefore, offer high expression if harnessed correctly. Transformation of exogenous DNA into the chloroplast is, unfortunately, difficult because of the double membrane and having no known viruses that penetrate it. To circumvent this, modification is usually done with tungsten coated particles in a gene gun, however, random integration is possible[20,26,27].

Another alternative to expression in plants which removes issues associated with random integration, is transient expression. While this has a limit on scalability, as expression plasmids will be diluted out over time, it has the potential to provide a high yield for short term applications. Another method used to overcome gene silencing from random integration is the inclusion of flanking homologous sequences on the inserted gene to allow for HDR (homology directed repair) insertion and, thus targeted expression. Even though this would circumvent silencing the chloroplast is the one organelle in which no glycosylation occurs, thus, creating an alternate issue for human therapeutic expression. The other issue, not centred around human therapeutic expression, is that the viruses and bacteria (Tabacco Mosaic Virus or *Agrobacterium tumefaciens*) that are used to carry the transgene into the plant can enter the environment and kill native plants creating substantial ecological risk[23,28].

As opposed to protein production in the whole plant, production in plant cells can be used to express transgenes which eliminates the potential ecological risk while still maintaining easy scale-up. This also offers a cheaper alternative to mammalian cell culture as media composition is simpler[29]. Glycosylation patterns, however, are unfortunately still an issue with products containing immunogenic sugars[23]. While the optimal production in plants was recorded at 1 g/kg of leaf mass, plant cells, on the other hand, were shown to produce approximately 3 mg/ml, however, expensive infrastructure is required to reach these yeilds[30]. While these amounts may be attractive for a sizeable plot of plants or large-scale bioreactor, this would only increase the effort and cost of maintenance, leaving this method only slightly better than mammalian cell culture, with the added disadvantage of incorrect glycosylation.

Currently, mammalian cells or specifically, Chinese hamster ovarian cells, are the main means of production for human therapeutic proteins. They are often modified with a vector containing the protein of interest under the dihydrofolate reductase (dhfr) promoter, which relies on random integration. After this, cells are screened in the hundreds to isolate a cell line that successfully expresses the protein [31]. These bioreactors offer the correct glycosylation patterns for human therapeutic proteins, which circumvent the disadvantage posed by non-mammalian production systems. They are also easily edited to incorporate the transgene that expresses the desired recombinant, which, again, is the appeal of some of the cheaper non-mammalian systems. Other examples of mammalian cells used for human therapeutic production are Baby Hamster Kidney cells (BHK21) and Murine Myeloma cells (NS0). These lines are successful in producing a recognizable and safe therapeutic protein suitable for human use with a yield of approximately 1-5 mg/ml [31,32]. While the yield produced by mammalian cell lines is sufficient, cell culture facilities remain expensive to maintain, in turn increasing the cost of the product.

Depending on the purpose of the protein that is required as well as the size and necessary post translational modifications, the systems discussed may be sufficient for producing large amounts of therapeutic. In terms of human therapeutics, however, these systems all suffer from the lack of correct glycosylation patterns or high running costs. Even though research has provided an alternative to harvesting therapeutic proteins from human blood, we are still left with the need for a bioreactor that has a short generation time, correct glycosylation patterns, and is cheap to maintain while having a large output of product to reduce unnecessary costs of treatment. To this end, the chicken as a protein expression system has recently come under the spotlight as it could provide the ideal host that would remedy all these issues. The table below summarizes all the bioreactors discussed as well as the chicken,

which will be discussed later, and compares various important aspects necessary for therapeutic protein production centered around cost and yield. As one can see, the chicken offers the most well-rounded system when yield, post translational modifications, scalability, gestation time, maintenance cost and ease of extraction are compared against other mammalian and non-mammalian systems.

Table 1: Comparison between non-mammalian and mammalian bioreactor systems for human therapeutic protein production. The higher the score for each aspect compared, the better the bioreactor is at performing the task or the closer it is to mimicking human protein production. Yield refers to the maximum potential yield possible in the system across its lactation or expression period, glycosylation is the common post translational modification necessary for human use, and scalability refers to how easily these systems can be used on a large scale i.e., land required or growth in a large bioreactor. Gestation time refers to the time from when the animal conceives to when the offspring are born or when the therapeutic protein can be obtained, cost of maintenance refers to growing and care costs over the lifetime of the vector, and ease of extraction refers to where the protein can be extracted from and, thus how easily it is accessed e.g., from the milk or egg. The information in this table was collated from sources 10-32.

System	*E. coli*	Yeast	Plants	CHO	Rabbits	Mice	Goats	Cows	Chicken
Yield	+++++	+++++	++++	+++	++	+	++++	+++++	++++
Glycosylation	+	++	++	+++++	+++++	+++++	+++++	+++++	++++
Scalability	+++++	+++++	++++	++	+++	++++	++	+	++++
Gestation time	+++++	+++++	++++	++++	+++	++++	++	+	+++
Cost of maintenance	+++++	+++++	+++	+	++	+++	++	++	++++
Ease of extraction	+++++	+++++	++	+++	++	++	+++	+++	++++

1.2 The chicken as a bioreactor and major routes for genome modification

What makes the chicken such an attractive alternative for therapeutic production is that it brings improvement to all the lacking areas that currently face protein production in mammalian or cell expression systems. Most importantly is the ability of the chicken to offer correct glycosylation across many human therapeutics. In addition to this, gestation time in chickens is 21 days and at 5 months the birds reach sexual maturity, which is significantly shorter than other larger mammalian systems. Egg laying occurs once a day per bird, which is extremely beneficial if the albumin (which comprises 66% of the egg) [33] were to be repurposed for expression, such as in the case of all mammalian expression systems where the milk promoters are utilized. This would be most ideal for the purposes of increasing yield,

however, targeting this area requires the manipulation of the chicken's genome. Two aspects need to be considered for inducing modification in chickens and that is where the modification will be administered, i.e., what cell type will be targeted, what promoter will be used promoter, or what physical area of the embryo will be modified, and how exactly these changes will be induced at these locations.

Targeting the gametes is the only way to gain access to the chicken's genome and elicit a change that could potentially be passed down to offspring. For this reason, all routes to obtaining gametes have been explored. Firstly, an understanding of gamete development in chickens is required to fully understand the different routes for gene modification and transfer available in avians as well as the advantages and disadvantages of manipulation at each possible stage. To begin, fertilization takes place after ovulation in the oviduct after which the ovum is encased in albumen and shell as it passes down the oviduct to be laid. Before oviposition, syngamy occurs followed by the first cleavage division. After incubation begins, the hypoblast is formed from bilayering of the area pellucida and the epiblast undergoes gastrulation to form the primitive streak, beginning the stages 2-4 under Hamburger Hamilton (1951) classification[33,34].

Ova in adult hens have been investigated as a potential area to manipulate the chicken's genome, however, the ovum would need to be harvested before oviposition and, therefore, from the oviduct, to be genetically manipulated. This would involve invasive surgery and the egg would then need to be placed back in the oviduct to ensure development as normal. While there have been experiments that successfully culture embryos in an artificial environment, the survival rate was only 10-15%[35–37]. Other disadvantages to this technique include the pronuclei not being visible which makes editing challenging and not having access to more than one ovum per hen brings in a major limiting factor for a technique that is already cumbersome[33].

Another point of potential genome modification are chicken primordial germ cells (PGCs), which are the stem cell like precursors to mature gametes in avians. PGCs are induced by the localization of maternal germplasm around fertilization, or epigenesis from cell interactions around gastrulation. The expression of transcriptional repressor b lymphocyte induced maturation proteins (prdm1 and prdm14) reduces somatic programming and signals pluripotency in PGCs, respectively[38,39]. In addition to this, VASA, which is found in the basal part of the cleavage furrow at first cleavage (Stage I), is incorporated into the blastodisc at stage V to aid in the formation of PGCs. Approximately 130 PGCs are localized at the area

pellucida at stage X[40,41]. At this stage it is also worth noting that blastodermal cells can be dissected out, edited and transferred to a recipient stage X embryo, however, this technique is challenging, not entirely specific to germ cells, and has a low success rate at accomplishing germline transgenesis [42]. After incubation begins, PGCs then migrate into the blood vessels that form in the germinal crescent at stage 5-11. Circulation throughout the embryo then begins at stage 13-17 and PGCs peak in the blood at stage 14[40,43]. PGCs exit the vessels and migrate along the dorsal mesentery and settle in the germinal ridges during stage 28-30. It is suspected that these cells are attracted to the germinal ridges by chemokine stromal cell derived factor 1 (sdf-1). PGCs only differentiate into mature male PGCs at stage 39 and germ cells in the gonads peak at stage 43[44–46]. Mature PGCs then come to rest in the gonads until spermatogonia restart division 10 weeks after hatching and differentiate into spermatogonial stem cells (SSCs). At sexual maturity these cells are responsible for forming the mature sperm. On the other hand, female PGCs start differentiation into primary oocytes at stage 34 and the 1st meiosis division begins at stage 39, halting at diplotene. Oocyte maturation and growth occurs at sexual maturity with primary oocytes resuming their first meiotic division before ovulation and then the second division follows and halts at metaphase until ovulation begins[47–49].

Given the migratory pathway of PGCs during embryo development, these cells and, thus the chickens genome, can be easily accessed from the germinal crescent, which contains other germinal crescent cells[50], the blastoderm which also contains other cell types[51], the blood at stage 13-17, the gonads at stage 28-30, or as SSCs present 10 weeks after hatching. The decision of which cell stage to utilize largely depends on whether changes need to be made to the adult or embryo chicken. Both stages have disadvantages, however, one is far more suitable in the case of limited funding and resources[46,52–54].

Using embryos as opposed to adult chickens brings with it many benefits. Embryos do not require as much upkeep as adult chickens, they require less space, ethics clearance is easier to obtain, suffering is minimized if experiments are performed before half development and a smaller number of PGCs are required to edit the chicken (~1000 cells) which is particularly beneficial due to the difficult nature and time needed to culture PGCs to a high enough density[46]. The successful transfer of PGCs from blood stage embryos at stage 13-17 to embryos at the same stage has been demonstrated several times [46,52–54] and, thus the technique is somewhat well defined. In addition to using PGCs edited in culture, modification of an organisms genome can also be performed via transfer of edited cells into the blastoderm or

germinal crescent, however, these techniques have a low success rate and are not specific to the cells in target[33].

Further on in chicken development, SSCs provide a similar vehicle for gene transfer to PGCs, however, these experiments of modification are often performed in adult roosters. It has been demonstrated that SSCs and/or PGCs transplanted into sterile roosters can successfully complete spermatogenesis, however, adult chickens require 2×10^6 cells for sufficient modification to take place within the gonads as opposed to embryos requiring only 1000. After obtaining the necessary number of cells needed for transgenesis, implanting edited cells into adult chickens requires invasive surgery which requires strict planning as to when the cells will be ready in sufficient amounts, how many times the surgery will be performed and, thus there is very little room for trial and error. In addition to this, acquiring adult animals takes more time and is a much more rigorous process than obtaining embryos[42,55,53]. Lastly, it inflicts unnecessary suffering on an animal with a fully developed nervous system where an alternative is available. As a result of all these disadvantages that follow using adult chickens, PGCs and embryos are considered the ideal candidates for gene transfer.

Once the preferred point of entry to the chicken`s genome is determined, the region of the genome to be modified needs to be decided upon. As previously mentioned, the chicken is an ideal bioreactor if expression of the therapeutic can be targeted to the albumin. Understanding the basic composition of the albumin is of importance to understanding which promoters are of use for modification and providing high expression of a therapeutic. Albumin includes lipids and proteins at 24%, carbohydrates, vitamins, and minerals at 13%, major egg white proteins such as ovalbumin (OVA) at 54 %, ovotransferrin (OVT) at 12%, ovomucoid (OVM) at 12%, lysozyme at 3,4% and ovomucin at 3.5%. Due to their dominant occurrence, OVA and OVM are often the subject of transgenic studies in chickens[57] .

The ovalbumin portion of egg white is synthesized in the oviduct with the 40kb gene expressing this protein forming a cluster of OVAX and OVAY paralogs on chromosome 2. The gene contains 8 exons of which the second contains the start codon and the first is untranslated. The region encodes a 386 amino acid (a.a) long protein [58–60]. The OVA promoter is well characterized and strong promoter found upstream of the coding sequence and is usually accessed via exon 2 as this does not disrupt splicing of the gene [61]. The function of OVA is not entirely understood; however, the protein was found to be unaltered at different stages of growth and is suspected to be an amino acid source for the embryo. It is also suspected to have antibacterial, antihypertensive and antioxidant properties. The paralog

of OVA, specifically OVAY, is a protease inhibitor and, thus protects the sperm and early oocytes or fertilized embryos from degradation. The other paralog, OVAX, is suspected to be an antimicrobial inhibitor [60,62–64].

While the function of ovalbumin appears important, modifications at the locus have been accomplished without effect on the viability of the embryo, at least in the case where one allele is conserved. Researchers have successfully expressed hINFβ under the OVA locus by knocking in the sequence under the 2nd exon using a CRISPR construct fitted with homology regions complementary to the region on either side of the exon, to allow HDR. Up to 4.42mg/ml of therapeutic protein expression under just the ovalbumin promoter, was recorded [65]. In another study, EGFP was successfully expressed at the OVA locus. This was introduced with CRISPR and homology arms designed around the tata box, to promote high expression of the protein[66].

The other dominant egg white protein, OVM, is located on chromosome 13 and spans 5.6kb. It contains 9 exons and the 0.63kb coding region expresses a 210 a.a product [66,67]. OVM is a trypsin inhibitor and is a suspected antimicrobial agent, however, its function isn't fully understood. It is also theorized that it protects the proteins in the absorbed egg white at day 11, as it inhibits proteases [68–70]. Similarly, to OVA, knockout at the OVM locus, at the 3rd exon, was successfully achieved and again, CRISPR with homology arm complementary sequences were employed to insert the fragment through HDR. This was accomplished without any effect on the embryo's viability suggesting that the function is not crucial to survival if one allele is conserved [70].

Another benefit to modifying these major egg white proteins, and perhaps disrupting and removing their presence, is that 2.5% of children are victim to egg allergies which could result in death in serious cases. OVA together with OVM are the two major allergenic proteins in egg white. Denaturation of OVA through heating and proteolytic digestion has been used to reduce the allergenicity of eggs, however, OVM is incredibly stable, thus, requires knockout through genetic means such as CRISPR. Knockout of OVA has also been accomplished for allergenicity purposes using TALENs (DNA-recognition transcription activator-like effector and a nuclease domain) [70–72].

Another major egg white protein that has not been modified due to suspicion that it is crucial for life, is OVT. It has a 2.4kb coding sequence and is located on chromosome 9. The protein product consists of 705a.a and can exist in a form bound or unbound to Fe^{3+} [73 74]. OVT has been found to lyse bacterial membranes, scavenge free radicals, and inhibit fungus and

viruses. It induces cytotoxicity in cancer cells, enhances immune responses during infection though the activation of macrophages, and enhances secretion of immunoglobulin A[75], all of which are important for the developing embryo contained in a porous shell.

1.3 Transgenesis in chickens for protein production

Genetic manipulation of chickens was first accomplished in the late 1980s[76]. Since then, there have been enormous strides in research into the best possible ways to induce transgenesis in these birds. Knowing that insertion of exogenous DNA under the OVA and OVM loci is the most feasible way to accomplish a higher yield of protein is only half the protocol required to create a transgenic bioreactor. While the chicken as a bioreactor would be a huge step forward in therapeutic protein manufacturing, avian genomes are unfortunately difficult to modify because of the zygotes yolky structure and silencing of the transgene often occurs due to random integration. While we have already circumvented the inaccessibility of the yolk structure by determining that PGCs are the best means of accessing the chicken's genome, here we need to consider the different methods which can be employed to induce transgenesis at the desired region in the genome while avoiding the common issues mentioned above. Although CRISPR is the main means of genome modification nowadays, multiple other methods of avian transgenesis exist that have brought the field of biotechnology to the point it is at today.

Viral based vectors were the original method of gene editing used in chickens. Viruses such as the Lentivirus and retrovirus that contain the target sequence were transfected into chickens allowing the exogenous DNA to insert at random into the genome. These viruses were commonly injected into PGCs, into blastodermal cells at oviposition [77], into the heart of 55 hour old embryos, or into the blood of stage 13-17 embryos[78,79]. In any of these methods of transgenesis, cells can either be modified by direct injection of the vector into the embryo or cells can be edited and then transferred to an unedited embryo to bring about modification. Lentiviral vectors were successfully used to create chickens that express miR24 and hIFNβ in the egg white by directing expression to the albumin by using the 5` end of the ovalbumin gene upstream from the coding sequence in front of the target DNA [78]. Similarly with Harvey *et al*, 2002 who demonstrated that β galactosidase and β lactamase could be expressed in the albumin if the viral vectors were injected into the sub germinal cavity of freshly laid eggs. In this case, a 1.4kb region upstream of the Ovalbumin coding sequence was used to drive specific expression [79]. While viral vectors brought about the necessary modifications, they unfortunately posed the issue of long-term dilution and silencing from possible random integration, making them suboptimal.

The use of lipofection and electroporation methods opened doors for avoiding the use of viral vectors for gene transfer. One such example is testis mediated gene transfer in which genes are transferred to spermatozoa and spermatogonia via lipofection or electroporation procedures and these sperm are then used for artificial insemination. Other means of inducing transgenesis in sperm include electroporation, restriction enzyme mediated integration, receptor mediated gene transfer and intra cytoplasmic sperm injection [80]. While these avoid the need for encapsulating a sequence that is to be introduced, the sequence itself needs to be modified to allow for genomic integration.

Recently, there has been large insight into avian transgenics with new methods being explored that would allow for naked DNA transfection of PGCs. Methods such as TALENs have been used to create DDXO4 knockout hens which were genetically infertile. The construct containing the coding region of interest was fitted with homology regions to the DDX04 gene and PGCs were edited and re-implanted to later breed sterile offspring [81]. Unfortunately, TALENs are expensive to make, and the efficiency of repair is low (8%), thus, there is a need for a cheaper system that performs modification in a similar targeted manner. Another study that demonstrated the use of an alternative to TALENs, which used a dimeric ZNF, failed to show editing at the OVA locus in chickens. Since then, this method has not been explored further [82] leaving us with the need for a precise editing tool.

With the discovery of CRISPR, accessing the chicken`s genome has become more straightforward and affordable. The efficiency of CRISPR was shown to be 90% when compared to TALENs, which was 33% [70], demonstrating its ability to be a more efficient and precise tool. Various studies have shown the successful expression of proteins such as hIFNβ and β-galactosidase in the albumin[65], with this editing technology. One of the main areas of genome modification that has seen improvement is the longevity of the system as these CRISPR based constructs can integrate and be passed down instead of existing transiently.

In order to create a self-propagating editing system, CRISPR would need to be used in combination with an editing technology that allows for permanent and precise change that could be passed down through generations in a non-Mendelian manner. Gene drives, which are constructs that bias the inheritance of a particular trait, in combination with CRISPR create this ideal editing system. This would work by employing a construct like that delivered by viral vectors, however, it requires the addition of gRNA sequences which are complementary to the target region in the genome, a pam sequence to allow for interaction of Cas9 with the genome at these regions, as well as homology arms complementary to the area

externally flanking where the gene drive will be inserted. To make cutting possible, CRISPR will also need to be included in the vector/s used in such a way that it inserts into the genome with the gene drive. The gene drive will then be inserted through HDR and expression of the desired therapeutic will be controlled under the native promoter in the target region. HDR occurs predominantly during meiosis due to the presence of sister chromatids and is a mechanism of conserving important regions in the genome upon DNA damage. Commonly, the sister chromatid, which contains a region complementary to the damaged region, is used for repair. The flanking homology arms will allow for the use of the desired insert in repair instead of the sister chromatid. Outside of meiosis, double stranded breaks are repaired through NHEJ (non-homologous end joining) and would, thus preclude gene drive insertion from taking place. For this reason, it is important for Cas9 to be under the control of a meiosis-specific promoter[61]. Below (Figure 1) is a diagram depicting a simplified version of HDR repair with a gene drive and strand conversion of the second allele once the gene drive has integrated into the first allele.

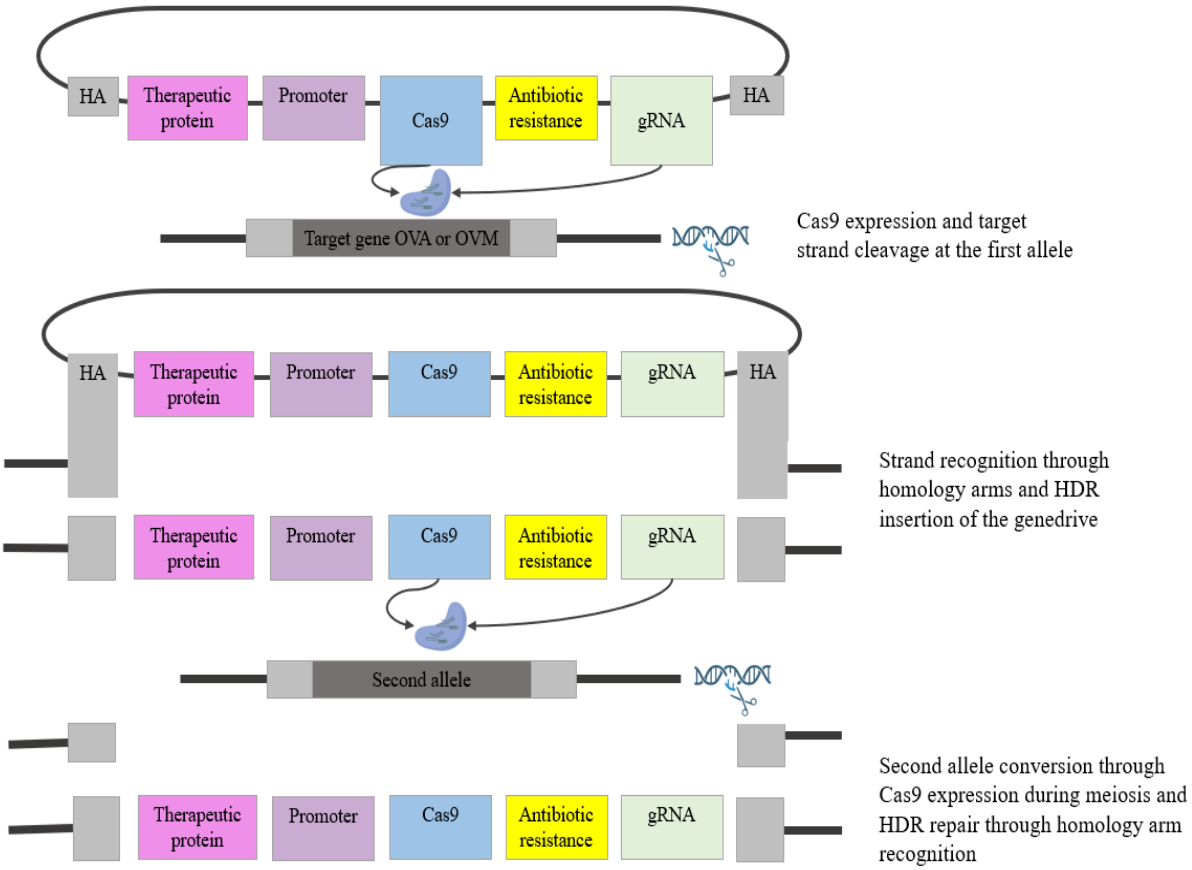


Figure 1: Homology directed repair within the chicken genome in the presence of a gene drive expressing Cas9. The genome is first cut by Cas9 which is expressed from the gene drive (Blue). This Cas9 can also be expressed from a secondary plasmid (not shown) if the Cas9 in the gene drive is controlled for meiosis-specific expression and editing is occurring *in*

vitro. Homology arms contained in the construct assist with complementary recognition after which strand invasion occurs and, thus HDR, inserting the gene drive into the target locus. Cas9 is then expressed during meiosis, *in vivo*, and the second allele can be converted similarly to the first, except this time by the first allele and not the transient gene drive. This image was created with PowerPoint (2022).

One of the most recent uses of this gene drive technology was demonstrated by Grunwald *et al* in mice. Here, Super-Mendelian inheritance was demonstrated in mammals for the first time, where triple homozygosity was accomplished at the tyrosinase locus, a feat which only has a 1.5% chance of occurring in nature[83]. This was, however, not the first instance of gene drive technology being used for genome modification. A study in 2013 was the first to allude to this technology when insertion of CRISPR Cas9 from one locus into another, by homology directed repair, creating a homozygote, was accomplished in *Drosophilia melanogaster*. Further to this, the power of these genetic constructs was demonstrated when producing infertility by mutation at the *asdgx* locus in the malaria carrying *Anopheles* mosquito. This targeted integration eliminated the downfalls associated with viral vectors and allowed for a more controlled and trackable expression[83–85]. These powerful gene drive tools allow for the much-needed advancement in avian transgenesis and coupled with the ideal expression system, provide the potential for a groundbreaking therapeutic protein revelation. In the case of the chicken bioreactor specifically, Super-Mendelian inheritance at the OVA and OVM loci within the albumin, which is demonstrated in Figure 2, could bring about maximum therapeutic expression and a rapid method for creating these “therapeutic eggs”.

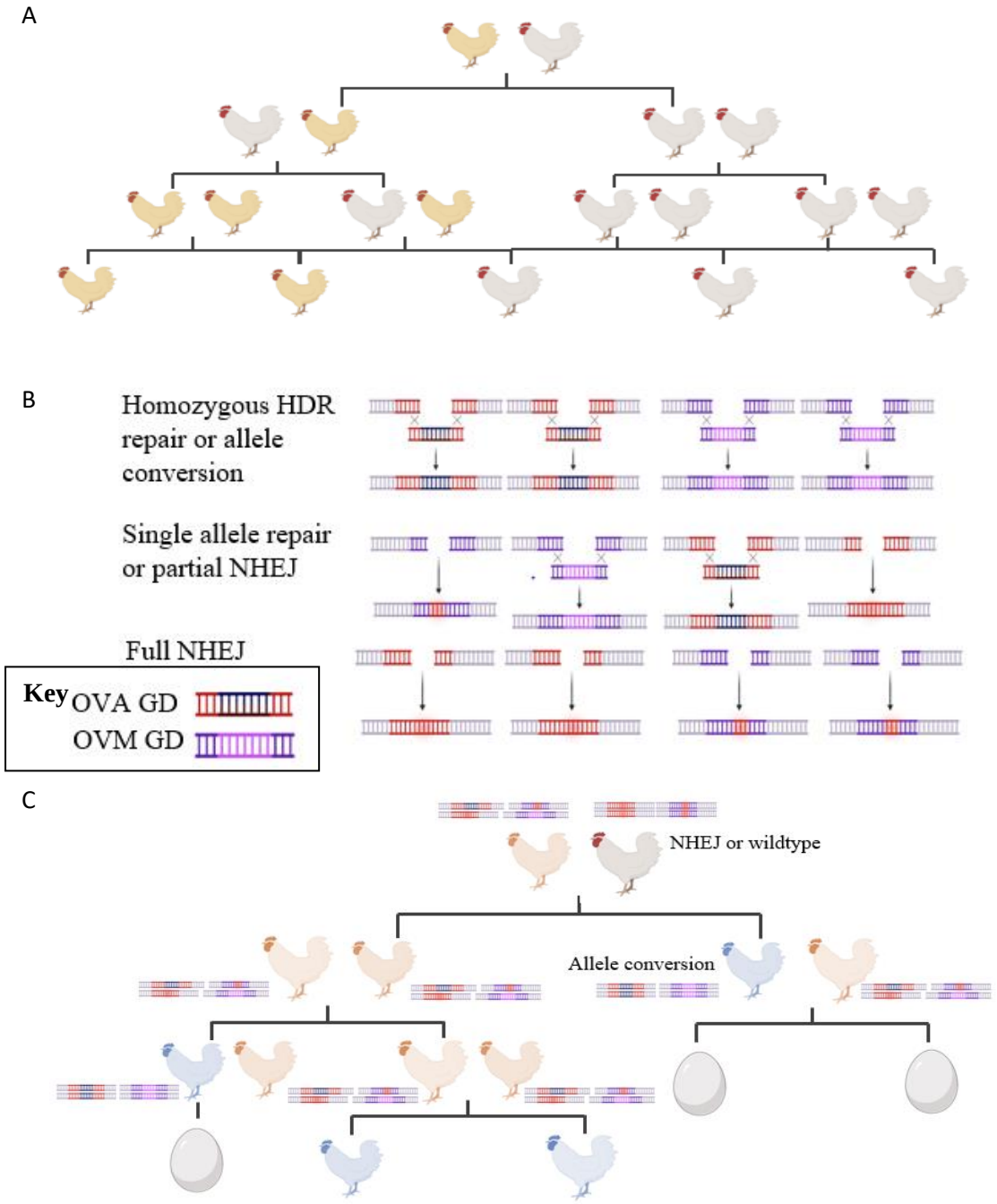


Figure 2: Mendelian inheritance as compared with Super-Mendelian inheritance brought about by a CRISPR gene drive. (A) represents normal mendelian inheritance where yellow chickens are homozygous recessive and white chickens are heterozygotes. (B) demonstrates the ways in which alleles can be repaired in the presence of gene drive editing during meiosis. (C) demonstrates Super-Mendelian inheritance and the creation of chickens that produce "therapeutic eggs". For each case in the Super-Mendelian pedigree, the alleles are demonstrated for both OVA and OVM loci as an example. Allele conversion is demonstrated, whereby the individual should have been heterozygote or unedited in normal instances but has undergone allele conversion to become a homozygous edited individual. Blue chickens

are sterile and lay eggs that contain therapeutic protein inserted under the two loci (OVA and OVM). These will be layer hens no matter if fertilization takes place or not. For the sake of this diagram, all offspring depicted are female crossed with an outsider male rooster. This image was created with BioRender (2022).

In the above diagram, Super-Mendelian inheritance is demonstrated to be a useful tool at accomplishing target modification within, potentially, the first generation through strand conversion. This editing tool makes it a lot quicker to induce change to the genome as opposed to normal constructs that are confined to the pattern of normal mendelian inheritance. In the case of chickens used for the production of therapeutics, the goal is to breed a homozygote layer hen, that produces eggs which will be sterile but will still contain the therapeutic of choice. Gene drive technology makes editing more convenient in this case as the breeding program can be started with either an unedited wildtype or a heterozygous, edited rooster (Figure 2 – white and orange bird, respectively).

1.4 Methods of PGC extraction and cell culture for creating transgenics

As discussed previously, PGCs are the chosen cell target of genome editing in chickens as they allow access to the germline. PGCs, thus need to be maintained in culture for a period sufficient to allow for editing, recovery, and transplantation into a host. Chicken primordial germ cells can be maintained in culture for long periods without losing their properties making them useful for conservation efforts for avians which require storage of PGCs as well as for the field of biotechnology and transgenics. As PGCs are the most ideal vector for creating transgenic chickens, different ways of culturing PGCs have been researched extensively over the years.

The dominant points of extraction for primordial germ cells are the blood of stage 13-17 embryos (380 PGCs[86]) and the gonads of stage 28-30 embryos (500- 2500 PGCs [87]). PGCs can also be extracted from the germinal crescent, however, there are far fewer cells present at this stage (250 PGCs[88]). While this method allows for fast extraction due to a low incubation time, the gonads and blood are preferred as they offer enough cells to encourage substantial proliferation in early culture. There is also persistent tissue in the culture after germinal crescent extraction which could hinder downstream processes. After extraction, PGCs can be purified using Nycodenz or Percoll purification, immunomagnetic cell sorting and fluorescence activated cell sorting (FACS), however, this is unnecessary for blood PGCs as the only contaminating cells are blood cells which cannot be maintained in culture. While gonadal PGCs are most ideal as they have the highest concentration in comparison to other

collection points, there is usually contamination from other cell types and, thus needs FACS, MACS or Percoll purification which can be an issue as these methods ar not always accessible. To circumvent the possible inaccessibility to these methods of purification, it was demonstrated that gonads extracted from embryos can be incubated at 37˚C in PBS which will discharge the PGCs from the tissue at 50% purity. For this method of purification, however, germline transmission has not been demonstrated [46].

After extraction and purification, PGCs need to be cultured in an environment that allows them to be maintained in their primordial cell state and not undergo differentiation into mature PGCs or germ cells. Across all protocols, there are 3 hormones that appear crucial for PGC development and maintenance. As stem cells, PGCs require human stem cell factor (hSCF) to ensure normal proliferation, differentiation, and survival. In addition to this, leukemia inhibitory factor (LIF) is added to prevent differentiation and basic fibroblast growth factor (bFGF) is added to encourage proliferation of cells in culture[88–91]. Another important factor for PGC culture is a feeder layer, of typically fibroblast cells, to maintain PGC colonies. PGCs are suspected to respond well to feeder layers as they secrete continuous SCF and LIF, which mimics their natural environment. Why these natural hormones perform better is suspected to be due to the fact that manufactured hormones only have a half-life of approximately 8 hours in standard culture conditions (37˚C, 5% CO_2)[92]. Commonly, chicken embryonic fibroblast cells are used as a feeder layer as it is easy to obtain material from left over eggs after PGC extraction. Gonadal stroma cells have been shown to be more effective than CEFs as a feeder layer, however, this is only for two weeks of initial culture and then the PGCs need to be transferred onto CEFs, STO (Sandoz inbred mouse-derived thioguanine-resistant and ouabain-resistant) or BRL cells [90,91,93,94], which are also commonly used feeder layer cells that can be purchased. To culture PGCs in the absence of a feeder layer, researchers recently demonstrated the concept of 3D culture. Using polymer fp003 in media increased the viscosity enough so that PGCs did not sediment, thus, allowing for better nutrient reach and expansion [95] .

Another approach to obtaining PGCs from a source other than the blood or gonads of chicken embryos, is to reprogram CEFs into PGCs. This has the advantage of producing high numbers of PGCs as CEFs are easy to culture and have a doubling time of under 24 hours as opposed to PGCs that have a doubling time of 2-3 days or more. To reprogram CEFs into germ cells, researchers expressed *Oct4*, *Sox2*, *Nanog*, *Lin28A* (*OSNL*) to induce iPSCs. These stem cells were then cultured in BMP4, BMP8b and EGF to further induce PGC formation [96]. This could potentially be a highly efficient means of obtaining high yields of PGCs,

however, the introduction of new genes will require additional work, the expression abilities of these reprogrammed cells are unknown, and there will be an increase in the cost of obtaining PGCs due to the additional hormones required.

Lastly, identifying PGCs is an important aspect of PGC culture. Certain markers are expressed on the surface of PGCs depending on the stage of their development. Staining of embryonic mouse antigen 1 (EMA-1) and stage specific embryonic antigen 1 (SSEA-1) help identify PGCs. These markers, however, are found to reduce by half at 2.5 days of development. Dazl and VASA are germline specific markers and bind to the vasa gene mRNA and RNA binding proteins within the cell. These markers were shown to decrease by 25% after day 4.5 of development. In addition to these germ cell specific markers, PGCs can also be identified by their large spherical shape (9-20μm), nuclei and refractive cytoplasmic lipid granules [86,87] and RT PCR. RT PCR is, however, limited for the identification of PGCs as often the markers used are not expressed at all stages and there are very few that do not overlap with CEFs, which are commonly used as a feeder layer. Nonetheless, there are two markers that are specifically detected in PGCs only, at most stages of development, and these are integrin α6 and integrin β1[87].

In general, the method for establishing a line of PGCs to create a transgenic follows a standard protocol. This includes extraction from either the blood or gonads of chicken embryos, after which cells are cultured for a month. Cells can then be identified through staining or RT PCR. PGCs then undergo some form of genome editing with a construct containing an antibiotic resistance gene. After this, the cells are subject to selection and are then re-implanted into a recipient embryo, usually into the gonads via the blood stream. The diagram below (Figure 3) summarizes the general culture and downstream procedure for PGCs, in the process of creating a transgenic protein bioreactor.

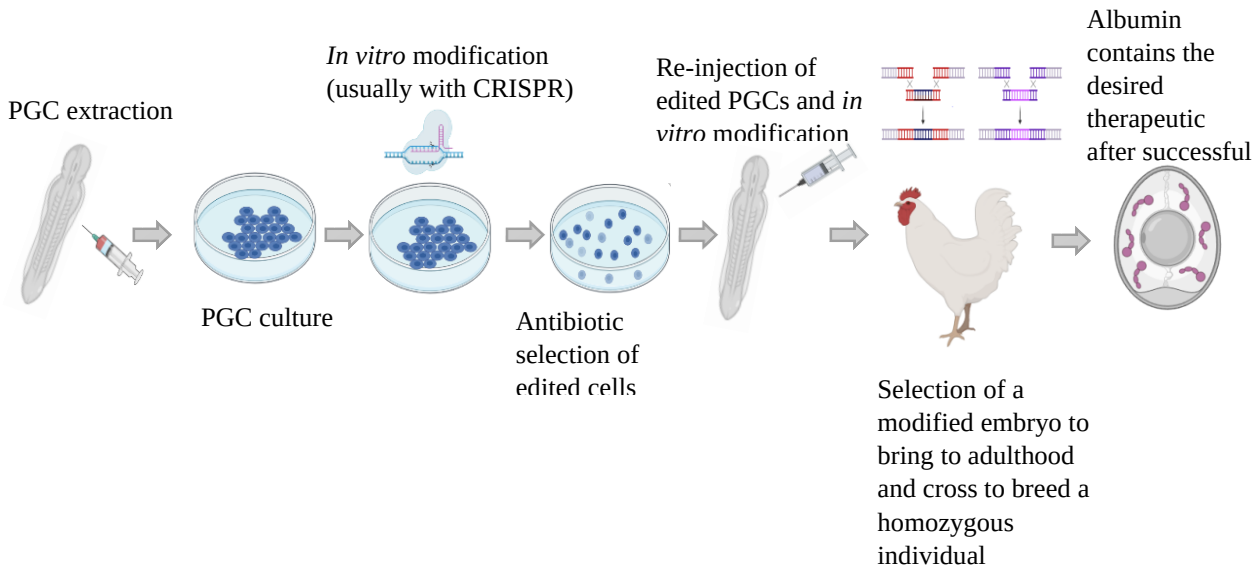


Figure 3: Overview of the process of PGC culture and transgenic creation. PGCs are extracted from embryos after which they are subjected to genome modification and antibiotic selection. After this, PGCs are injected into the blood stream of embryos and edited chickens are screened, bred, and crossed to yield expression of a therapeutic protein in the albumin. This image was created with BioRender (2022). Based on sources 65,66,70,90-95.

1.5 Transfection methods for PGCs with naked constructs

Once PGCs are cultured to an appropriate density and a method of editing has been chosen, the desired construct needs to be administered. Transfection of PGCs has been studied extensively to achieve the optimal rates of transgenesis. Multiple methods have been compared and improved upon to achieve this, but the two that stand out above the rest are electroporation and lipofection.

One way in which transfection of PGCs can be accomplished is via adding lipofectamine to the cell culture media. This is a lipid-based system that works by carrying the construct into the cell when the polymer merges with the membrane. This has the advantage of being less damaging to the cell than other methods, such as electroporation, however, it has been demonstrated that lipofection only has a 17 % transfection efficiency[97] which was significantly less than that demonstrated for other methods such as electroporation[89]. Transfection with transposons and lipofection was attempted to improve efficiency and was shown to give a 60 % transfection efficiency of the construct[98], however, the use of transposons may not always be possible depending on the research being done.

Transfection with electroporation provides a more efficient alternative to lipofection and can be accomplished with low voltage (approximately 240-250 V) pulses applied to cells to open

pores in their membrane. This then allows constructs in the media to pass in and integrate into the genome. There is, however, the downside that electroporation kills a larger number of cells than lipofection, but this is insignificant when compared to the number of transfected cells that can survive downstream processing, such as selection with antibiotics. In addition to this, it was demonstrated that electroporation of PGCs had a much higher transfection efficiency, of 80%, compared to the 17% demonstrated for lipofection[97]. A similar method to electroporation, which allows for a more reliable transfection of cells, is nucleofection. This is performed like electroporation except the media composition allows for construct passage into the nucleus. The disadvantage to this method, however, is that it requires special equipment, such as a nucleofector, which may not always be available [98].

The other major advantage to using electroporation is that extensive research has been done to improve the success of transfection with this method. DMSO (dimethyl sulfoxide) was found to increase electroporation efficiency of PGCs by 2.2-fold when adding 1.25 % DMSO to the electroporation mix. The exact mechanism of this is unknown but it is suspected that the compound increases cell membrane permeability or aids in stabilizing the membrane after electroporation, allowing for better recovery [97]. Other researchers have also demonstrated that PGCs subjected to electroporation and Percoll purification have the best transfection efficiency compared to Percoll and lipofectamine [89]. This is suspected to be due to the Percoll removing unwanted cell debris which may interfere with electroporation pulses [99].

While PGC manipulation has been thoroughly researched, it is still multistep and suffers from low efficiency. Combining injection, electroporation and lipofection it was demonstrated that transposons could be injected into the sub germinal cavity of HH stage 1 embryos followed by electroporation to promote insertion into PGCs. Additionally, DNA combined with lipoplexes and highly concentrated carboxy methyl cellulose was injected into HH stage 1 and stage X embryos with an efficiency of over 20 % integration in PGCs [100]. Although transfection efficiency with these methods is low, actual repopulation of the gonads with edited PGCs from culture is comparable to that with traditional electroporation and lipofection in whole embryos. These methods would eliminate the need for PGC culture, and the multistep process involved in creating transgenics.

1.6 Sterilisation of chickens to remove native germ cells

Once edited PGCs are obtained, they need to be transplanted into a host's gonads that are sterile or partially sterile. Generating sterile recipients for edited PGCs is of high importance as this will determine the success of repopulation efficiency as well as the number of edited

offspring and edited gametes being passed down to the next generation. Multiple methods exist for the sterilization of chickens such as exposure to γ radiation, UV radiation, manual removal of progenitor cells, busulfan treatment, and genetic sterilization. The choice of method depends largely on the amount of research done, the cost it entails, and accessibility to any necessary equipment.

Some of the less popular techniques have been phased out over the years due to cost and inaccessibility. One of these techniques is sterilization with γ irradiation which works by exposing unincubated eggs to a co-source and destroying stem cell like cells[101]. UV exposure specifically blocks PGC migration to the germinal anlagen and, thus population of the gonads[102]. While this technique is relatively simple, as it is accomplished with a 5W lamp, it produces considerable teratogenic effects. The use of a UV laser as opposed to a lamp source was found to reduce the number of teratogenic effects as it is more precise, however, equipment for this method is very expensive and accessibility is more limited [102]. Other types of sterilization include X ray exposure. Although this can have detrimental effects similar to UV exposure, shielding embryos with lead can significantly reduce these. This additional step, however, requires thin sheets of lead as well as specialized equipment like a soft X ray apparatus (B-4:softex Tokyo Japan) [103]. The last of the less popular techniques to acquire sterilization is the manual removal of cells that form PGCs. This is done with a micropipettor in the sub germinal cavity. While this in combination with UV exposure produces the best viability in treated embryos (80%), compared to γ radiation exposure (70 %) and has the best depletion rate of native PGCs at 56 % compared to UV and γ irradiation (38 and 27 %, respectively) [101], the equipment needed is highly specialized and the method is time consuming.

Perhaps a more accessible, and cost-effective approach to sterilization is treatment with busulfan. Busulfan is an alkylating agent commonly used to prepare the body for stem cell transplant in leukemia patients. It stops the growth of cells and is nonspecific, however, PGCs are shown to be particularly susceptible to its effects until 48 hours after egg incubation [104]. Recent studies have shown that concentrations of 100μg busulfan in 50μl sesame seed oil and DMF (N-N Dimethyl formamide) have the best ablation efficiency. Survivability at day 7.5 at this concentration was significantly less (68%) than at 75μg (74%) with significant decrease as incubation increased. In addition to this, hatchability was at 36% and 41 % at 100 and 75μg, respectively. In addition to this, 100μg or higher was shown to be highly teratogenic. For these reasons, 75μg is favored for sterilization[105]. While busulfan

treatment was shown to better germline transmission rates, chimeras were still present after ablation, thus, demonstrating that it is not a means of complete sterilization [52].

Research has also been done on the best way to administer busulfan to maximize PGC ablation. The current standard is to use sesame seed oil as a carrier for busulfan dissolved in DMF or DMSO. Sesame seed oil mimics the viscosity of egg yolk and DMF/DMSO functions to dissolve busulfan which is otherwise insoluble. DMF and DMSO dissolve busulfan in a similar way, however, DMSO has less detrimental effects than DMF. The speed at which busulfan is released into the yolk can also be controlled. Sustained release busulfan treatment differs from immediate release by the time it takes for the busulfan to dissolve into the yolk. For sustained release, busulfan is dissolved in DMF which is further diluted in PBS and emulsified in oil. As busulfan is not soluble in PBS it takes longer for it to penetrate through the yolk. Immediate release contains busulfan dissolved in DMF and emulsified in oil. It was found that after injection into the yolk, busulfan short release emulsion reached the embryo immediately as it rose to the top of the subgerminal cavity as it was more lipophilic, and the sustained release emulsion was delayed. While sustained release removed more PGCs, and had the same survivability as short release, there were more teratogenic effects than in the short release treated embryos. Neither of the methods had a better re-population efficiency than the other[104],[53].

Time of busulfan administration was also demonstrated to have an effect on the ablation efficiency and morphology of the embryo. Treatment at 0 hours of incubation opposed to at 24 hours of incubation was found to reduce the number of PGCs more even though survivability decreased. Survivability was 10 % less but hatchability was similar between the two injection times at 3 %. There were also only 1 % less teratogenic effects compared to 24 hours of incubation. The number of circulating PGCs (cPGCs) in the blood of 0 hour treated embryos was significantly less than 24 hour treated embryos and donor PGCs were significantly more in 0 hour treated groups due to the fact that busulfan is suspected to have been degraded by the 48 hour mark in 0 hour treated embryos where as in the 24 hour treated embryos, residual busulfan may kill off any re-injected PGCs at stage 14[106]. To avoid these effects of residual busulfan traces, researchers demonstrated that busulfan resistance can be induced in PGCs *in vitro*. They expressed human glutathione s transferase (mgstII) in PGCs which conferred resistance to 1μm - 4μm concentrations of busulfan. This also aids in repopulation efficiency as cells can be selected for in culture to reduce the number of unedited PGCs that may have escaped selection with an antibiotic marker [107].

The most recent and effective means of sterilization demonstrated, was genetic sterilization induced in hens and roosters. ICaspase 9 male chickens were created by placing ICaspase9 or human caspase 9 fused to the FK506-binding protein (FKBP) in the 3`-end of the Dazl gene through HDR. Fusion of the truncated human form of Caspase9 to FKBP allows for the dimerization and activation of ICaspase 9. Dazl is a germ cell dependent marker that is highly expressed in migratory and gonadal PGCs. The action of ICaspase 9 is induced by exposing embryos to ap20187 which leads to apoptotic cell death and removal of PGCs. In females the dead box protein 4 (DDX04) gene is knocked out to produce genetically sterile hens that can lay eggs from donor PGCs [108]. It was also recently demonstrated how this can be used to biobank and aid in conservation efforts for avians if gonads are harvested and cryopreserved to serve as the donor cells for these sterile carriers [109,110].

1.7 A word on Hemophilia B and Human FIX as a therapeutic

Overall, the methods described above highlight a means to creating a more affordable expression system for therapeutic proteins. In this project in particular, human FIX is the target of expression. Human FIX is a serine protease as it functions as a clotting factor in the clotting cascade. It is used to treat Hemophilia B (HB) by replacing missing or non-functional FIX. The cost of treatment for severe – moderate HB is approximately 21-22 million USD for a lifetime in one adult [111]. HB is usually treated with prophylaxis with FIX (Benefix/nonacog alfa) where it is administered weekly in severe phenotypes or when necessary, in mild phenotypes. In addition to being relevant for HB treatment, FIX is also used as a clotting factor treatment for post operative and post trauma bleeding, which are some of the leading causes of death worldwide[5]. FIX was traditionally harvested from human plasma, however, due to the risk of contaminating viruses it is now produced in CHO cells from where it undergoes chromatography for purification and viral inactivation[112–115].

HB is determined by mutations on the X chromosome in which the FIX protein is rendered nonfunctional or completely missing. It occurs in 1 in 40000 people and over 1000 causative mutations have been characterized which can influence the severity of the phenotype of the disease. More specifically, point mutations result in HB Leiden, which is characterized by a severe phenotype, as well as missense, nonsense and, silent mutations which affect the availability, stability, activation, and function of FIX and result in moderate to severe phenotypes[112–114,116]. The main characteristics of a mild phenotype include significant bleeding after operation or injury, moderate phenotypes include bleeding after minor trauma and severe phenotypes include sporadic and heavy bleeding. Symptoms range from hemarthrosis, hematoma, cerebral hemorrhage, arthropathy to loss of limbs or life [117].

The structure of FIX lends itself to the protein`s function and, thus its mechanism in HB. The FIX protein is encoded by the *F9* (*factor 9*) gene which contains 8 exons. F9 spans 34kb on the X chromosome. The final 55KDa protein encoded is 461 a.a long with an N terminal signal peptide, propeptide, GLA domain, two EGF like domains and a linker sequence, an activation peptide, and a serine protease domain. The Signal peptide and propeptide are cleaved to form the mature FIX protein and the activation peptide is cleaved to form the light chain and heavy chain (a serine protease linked by disulfide bond). FIX is synthesized in the liver and from here it undergoes post translational modifications such as glycosylation, sulfation, carboxylation, and phosphorylation. The propeptide region is responsible for carboxylation as it binds FIX to the endoplasmic reticulum (ER). The propeptide has a hydrophobic region that stops translation in the cytosol, thus, needs to be removed for mature FIX to form. The GLA domain helps membrane binding occur during clotting and tissue factor (TF) binding which prompts conversion to active FIX (FIXa) after the activation peptide is cleaved off. The EGF domains interact with TF which is of importance in the final steps of the clotting cascade[118],[116].

Due to the causative mutations of HB, FIX`s function in the clotting cascade is affected. Normally, TF proteolysis of FIX converts it to FIXa after which it functions as a serine protease in which it cleaves FX to eventually signal fibrin clot formation and activate platelets at the site of injury. The EGF2 domain of FIX binds to the platelet membrane and fVIIa and FX to assemble the FIXa-FVIIa complex which activates platelets. FIXa also cleaves out the activation peptide of FIX through interaction with its EGF1 domain. The signal peptide on FIX allows for a FX recognition site, heparin binding site and binding to FVIIIa to gain full activity. FIX is usually activated by TF or FXI, however, due to a change in amino acid residues caused by Hemophilia B mutations, FIX can no longer function to sever FX to form FXa which binds to fVIIIa as a co-factor and which eventually converts prothrombin to thrombin forming crosslinked fibrin [118,119].

1.8 Aim and objectives

This project aims to transfect chicken PGCs with a human clotting factor 9 gene drive, ablate native PGCs in chicken embryos, and assess repopulation of chicken gonads with edited PGCs.

Objective 1: Construct the gene drive necessary to insert FIX under the albumen genes in domestic chickens. The baseline gene drive was already synthesized by GenScript, however, Factor 9, the Sycp1 meiosis-specific promoter (synaptonemal complex protein 1), homology arms, and antibiotic resistance genes, needed to be added.

Objective 2: Establish the concentration necessary for antibiotic resistance selection of PGCs. This will be accomplished using neomycin and puromycin kill curves.

Objective 3: Determine the necessary ablation concentration for eliminating native PGCs in embryos. This will be determined through immunohistochemical staining (IHC).

Objective 4: Establish a line of primordial germ cells and transfect them with an OVA and OVM targeting FIX gene drive. This will be accomplished using electroporation and lipofection.

Objective 5: Assess repopulation of embryo gonads with edited PGCs. Ablated embryo gonads will be repopulated with PGCs containing an Red Fluorescent Protein (RFP) expressing tracer construct.

2. Methods

2.1 Bioinformatics

All primers designed were run through Primer-BLAST (NCBI) against the necessary gene drive to identify any possible off targets. After it was clear that the primers designed had no off-targets, the sequences were run through the Integrated DNA Technologies (IDT) OligoAnalyzer tool (https://eu.idtdna.com/calc/analyzer) to identify any hairpin structures or self/hetero dimer structures that may present an issue. A delta G greater than -9 kcal/mole would deem the primer unacceptable and would, thus need to be redesigned.

Sequencing was done by Inqaba Biotechnology or GenScript and results were analysed using MUSCLE (Multiple Sequence Comparison by Log-Expectation) alignment in MEGA X [120] and represented in SnapGene (2022). All plasmids were provided at a final concentration of 50 ng/µl for Inqaba Biotechnology and a minimum of 4 µg was blotted onto sterile filter paper for GenScript.

2.2 Construction of the gene drive

2.2.1 Bacterial culture and plasmid extraction

2.2.1.1 Practices for intermediate cloning steps

Firstly, the two constructs needed to target the OVA and OVM loci, to induce expression of therapeutic FIX in the albumin, had to be grown in *Stbl E. coli* and plasmid extracted. These gene drives already contained some modifications that were made to the original construct, synthesized in the pUC19 backbone by GenScript, however, additional changes were required to ready the constructs for transfection. After each change was induced in the various gene drives, through restriction digest and ligation, bacterial culture and plasmid extraction was used to prepare the construct for the next round of modification and cloning.

For the transformation and liquid culture of gene drives, 100 ng of the gene drive was transfected into *Stbl E. coli* by adding the plasmid to 20 µl of cells and placing the cells on ice for 30 minutes. The cells were then heat shocked at 42 ˚C for 30 seconds in a preheated water bath. The total 20 µl of transformed cells was then plated on an ampicillin plate (100µg/ml) and allowed to grow overnight at 30˚C. Colonies were selected the next day and placed into 5 ml of liquid broth containing 100 µg/ml ampicillin. These were left to shake at 180 rpm, overnight, at 30 ˚C.

Next, plasmids from liquid cultures were extracted using an alkaline-lysis method. Liquid culture, at 6 mls, was spun down at 17000xg for 30 seconds. Ice cold resuspension buffer (composition: 50 mM glucose, 25 mM Tris-Cl (pH 8.0), 10 mM EDTA (pH 8.0)) was then added at a volume of 250 µl and the pellet was resuspended by pipetting up and down. Lysis solution (composition: 0.4 N NaOH, 1% w/v SDS) was added next at 300 µl and the tube was inverted 5 times and allowed to sit on ice for 2 minutes. Next, 400µl of ice-cold neutralization buffer (5 M potassium acetate and 11.5 % glacial acetic acid from 100 % stock) was added and the tube was inverted until white precipitate formed. This was spun down for 5 minutes at 17000 xg. The supernatant was then transferred to a new tube and an equal volume of 100 % isopropanol was mixed in by inverting 5 times. The samples were placed at -80 ˚C for 30 minutes and after incubation, the samples were spun down for 5 minutes at 17000 xg. After this, the supernatant was completely removed, and the pellet was resuspended in 250 µl TE buffer (composition: 1 M Tris, 0.5 M EDTA pH 8.0) containing 20 µg/ml RNase A. RNase A treatment was performed over 1 hour at 37 ˚C in a heating block. After treatment, 600 µl of 0.2 M potassium acetate isopropanol (Composition: 100 % isopropanol, 5 M potassium acetate) was added to the TE RNAse mixture and the sample was vortexed briefly. Tubes were then spun for 5 minutes at 17000 xg. The supernatant was removed, and the pellet was washed with 70 % ethanol. Once the remaining ethanol was removed, the pellet was allowed to dry at 37 ˚C in a heating block for 5 minutes. Pellets were then resuspended in TE buffer and the concentration of the extracted plasmid DNA was measured using a Nanodrop One (Thermo Fisher Scientific).

2.2.1.2 Practices for prepping final constructs for transfection

In order to prepare endotoxin free samples of the constructed gene drives necessary for PGC transfection, the PureYeild Plasmid Midiprep kit from Promega was used (A2495). In addition to the two final gene drive constructs, it was also necessary to prepare samples of the mRFP-Tol2 plasmid designed by Zane Oberholzer and the Esp-Cas9 general expression construct. The RFP tracer construct was needed to visualize repopulation and transfection in PGCs. The Esp-Cas9 general expression construct was needed to allow for Cas9 expression upon transfection of PGCs with the gene drives as the final gene drives contain the Sycp1 promoter, preventing expression of Cas9 until the cells enter meiosis *in vivo*.

Briefly, fresh transformants for each of the constructs were plated and transferred to a 250 ml liquid culture in Terrific Broth (See recipes 1). This was allowed to shake at 250 rpm for 24 hours at 30 ˚C. The OD (optical density) was measured until it reached 2, using a spectrophotometer. Next, 125 µl of liquid culture was spun down in bottles at 5000 xg for 10

minutes. The pellet was resuspended in 6 ml of cell resuspension buffer and 3 ml of the cell suspension was mixed with an equal volume of cell lysis solution. This was kept at room temperature for 5 minutes with constant inverting. Next, 5 ml of neutralisation solution was added, and the tube was inverted until the solution was completely neutralised. This was confirmed when the lysate turned clear.

The lysate was then centrifuged at 15000 × g for 15 minutes and transferred to the clearing column and a vacuum was applied to pull the lysate through the clearing and binding columns. Endotoxin removal wash was added at a volume of 5 ml to the binding column and removed with a vacuum. Column wash was then added at 20 ml and similarly removed from the column. The column was allowed to vacuum dry for 2 minutes until no ethanol residue was left. The sample was then vacuum eluted into a 1.5 ml tube using 600 µl of DNA elution water.

In order to concentrate the plasmids to over 1 µg/µl, which is necessary for optimal transfection, 1/10th volume 3 M sodium acetate and 2.5 volumes 95 % ethanol were added to the eluted DNA and mixed by inverting. The tube was then placed at -80 ˚C for 30 minutes after which the sample was spun at 14000 xg for 10 minutes. Seventy percent ethanol was then added to the pellet and spun at 14000 xg for 20 minutes. After this, ethanol was removed, and the pellet was dried at 37 ˚C for 5 minutes. Finally, the sample was resuspended in 100 µl elution water, and the concentration was measured with a Nanodrop One (Thermo Fisher Scientific).

2.2.2 Digestion and ligation of gene drive fragments

2.2.2.1 Cloning and check digests

After plasmids were extracted, they would undergo restriction digest depending on what needed to be inserted. Digested products were then either gel extracted, or PCR purified and after undergoing ligation, the extracted plasmids were always cheaper digested with various enzymes, usually the cloning enzymes, to confirm the presence of the desired insert. All downstream incubations were done in the T100 thermocycler (Bio-Rad), and all gels were viewed with the GelDoc XR imager (Bio-Rad). All digests, whether it be for gel extraction or visualisation, were mixed with 50 % purple loading dye and loaded onto a 1 % agarose gel, stained with ethidium bromide. In the case of each digest performed, the enzymes used were supplied by New England Biolabs and the buffer used was the one that came with the relevant enzyme. The reaction set-up and conditions for both digestions for cloning and checking are described in the table below (Table 2). Once all the necessary changes had been made in the gene drive, the final OVA and OVM gene drives were digested with *Pac*I,

separately, in order to linearize the constructs for easier integration into the PGCs. Digests were performed as previously described and two 50 µl reactions of 10 µg each were pooled and purified according to the Monarch PCR clean-up kit (NEB T1030S) described below in section 2.2.4.2.

Table 2: Reaction set up for all restriction digestions. The reaction conditions for products needed for further cloning as well as check digests required to confirm the successful insertion of fragments post-cloning, are indicated under each relevant heading.

Digests for further cloning		
Component	Amount	Incubation time at 37˚C
Buffer (Cutsmart Buffer supplied with the enzyme)	5 µl	Overnight
Enzyme	1 µl (each)	
Sample	5 µg	
Water	Up to 50 µl	
Check digest		
Buffer (Cutsmart Buffer supplied with the enzyme)	2.5 µl	1 hour
Enzyme	0.2 µl (each)	
Sample	1 µg	
Water	Up to 25 µl	

2.2.2.2 Ligations

In order to set up the necessary ligations required for cloning in the various fragments into the gene drive constructs, the concentration of gel extracted, or PCR purified product, was determined using a nanodrop. The reaction conditions for all ligations, except the NEBuilder ligation, are set out in the table below (Table 3) and all reactions were mixed thoroughly by pipetting up and down after set-up. The NEBio calculator (version 1.15.0) was used to determine the amount of insert required for 100 ng of vector in a 3:1 ratio.

Table 3: Reaction set up for ligations peformed for the insertion of various fragments into the gene drives. This protocol is the same for all insertion steps except the antibiotic resistance genes, which are accomplished with NEBuilder.

Component	Amount	Incubation time at room temperature
Buffer (NEB M0202)	2 µl	1 hour
Ligase (NEB M0202)	1 µl	
Vector	100 ng	
Vector: insert ratio	3:1	
Water	Up to 20 µl	

2.2.3 NEBuilder construction of the antibiotic resistance genes

2.2.3.1 Primer design for NEBuilder

One of the changes to be made to the gene drive constructs was to insert a different antibiotic resistance gene in each. For this, neomycin resistance was chosen for the OVM gene drive and puromycin resistance for the OVA gene drive. Two different resistances would allow for the separate selection of each gene drive within a single cell. The primers for the assembly of the neomycin and puromycin resistance genes (Table 4) were designed using the NEBuilder Assembly Tool (version 2.6.1). Annealing temperatures were approximately 60˚C and the overlaps between various fragments were the maximum allowed for the NEBuilder HiFi DNA Assembly Cloning kit (E5520S), at 20bps for each overlap. Primers were synthesized and cartridge purified through Inqaba biotechnology. Primer sequences were as follows:

Table 4: Primers used to amplify the various fragments for NEBuilder assembly of the antibiotic resistance genes for neomycin and puromycin. The primer sequence as well as the primer direction and designated symbol are given in the first 3 columns. The fragments that are joined are described in the "function" column and the size of the fragments generated is described under "target size". GD is short for gene drive, Puro is short for Puromycin, and Neo is short for Neomycin. Fwd indicates the forward primer and Rev indicates the reverse primer. The amplified fragment that is produced from primer pairs is given under the full fragment amplified column. * Primer B1 and D1 pair to produce the full neomycin fragment of 1349 bps. 3` UTR is the 3` UTR for the anitbiotic resistance gene and SV40 is the promoter.

Symbol	Direction	Primer sequence 5`-3`	Function	Target size	Full fragment amplified
A2	Fwd	TGTCTTCTTGACGAGCATTCGAC TCTGTCACCGGACTG	3` UTR+ GD	642	3` UTR
A2	Rev	CTGGTTCTTTCCGCCTCAGAGCC TCTATTGTGTGCACTT	3` UTR + sv40		
B1*	Fwd	AAGTGCACACAATAGAGGCTCT GAGGCGGAAAGAACCAG	3` UTR + SV40	397/13	Sv40/Neo*
B2	Rev	GTGGGCTTGTACTCGGTCATTTT GCAAAAGCCTAGGCCTC	sv40 + puro	97*	
C1	Fwd	GAGGCCTAGGCTTTTGCAAAATG ACCGAGTACAAGCCCAC	sv40 + puro	912	
C2	Rev	TCGGCAGTGCCTGAAGGAGCGG CCTCCCCAGCATGCCTGCTATTC	GD + puro		Puro
D1*	Rev	TCGGCAGTGCCTGAAGGAGCGG CCCACACAAAAAACCAACACAC AGA	Neo + GD	1357	Neo*

2.2.3.2 PCR amplification of the various fragments needed

Next, the important fragments needed to assemble the antibiotic resistance genes were PCR amplified. The sv40 promoter was obtained from a neomycin cloning vector (Figure 4B). Additionally, puromycin, neomycin as well as their respective poly adenylation signals needed to be amplified and isolated from the relevant plasmids (Figure 4A and B, respectively). The 3` UTR from the CRISPR gene drive stock also had to be amplified as it was removed during the excision of zeocin resistance from the stock gene drive (Figure 4C). All these vectors were already available in the lab. The 3` UTR of Cas9 then needed to be placed upstream of the neomycin and puromycin resistances as previously it was behind IRES and the zeocin resistance, which was removed when the gene was excised. All PCR for the various fragments was done using Q5 DNA polymerase (NEB M0491) in a T100 thermocycler (Bio-Rad) (Table 5). PCR reactions were set up as follows: 10 µl 5X reaction buffer was added to a PCR tube with 1 µl dNTPs, 2.5 µl 10 µM forward and reverse primer, 10 ng DNA, 0.5 µl enzyme and water to 50 µl. The PCR conditions for amplifying the NEBuilder fragments were as follows:

Table 5: PCR conditions for the various NEBuilder antibiotic resistance gene fragments amplified with Q5 DNA polymerase. The relevant fragment being amplified under certain conditions is indicated under the "fragment" column. For all fragments, denaturation, annealing and cycle number are the same. SV40 is the resistance gene promoter, 3` UTR is the untranslated region of the gene and Neo and Puro are short for neomycin and puromycin, respectively.

Fragment	Denaturation	Denaturation	Annealing	Extension time (s)	Cycles
Sv40	98˚C 30s	98˚C 10s	72˚C 30s	11.91	30
3` UTR				19.26	
Neo				41.91	
Puro				27.36	

These fragments for neomycin and puromycin resistance were then joined with overlap extension PCR using the same Q5 protocol described above and modified for an extension cycle (Table 6). Antibiotic resistance genes had to be joined before insertion into the gene drive due to the fact that the NEBuilder kit had a limit on the number of pmol of inserts that could be included, thus, joining the entire fragment was necessary as it helped the insert fall within the pmol limits given within the kit. Exceeding or falling short of the recommended limits would have resulted in a ligation with low efficiency.

All reactions were initially allowed to run for 7 cycles without the addition of primers or the final extension at 72 ˚C and the cool down step at 14 ˚C. These 7 cycles allowed for the annealing of the different fragments at the overlapping sections created by the NEBuilder primers (Table 4). After this initial annealing step, primers were added, and the reaction ran for 15 cycles, according to the conditions below, in order to amplify the necessary genes:

Table 6: PCR conditons for overlap extension PCR for joing the individual puromycin resistance fragments and individual neomycin resistance fragments. This step is necessary to make a complete antibitoic resistance fragment for each before inserting it into the gene drive. For each procedure, the denaturation, anealing and cycle number are the same.

Product	Primers	Denaturation	Denaturation	Annealing	Extension	Cycles
Neomycin	A2 + D1 (Table 4)	98˚C 10s	98˚C 30s	72˚C 30s	58.5	15
Puromycin	A2 + C2 (Table 4)				56.3	

These fragments that had been joined to create the full length puromycin and neomycin resistance genes were then mixed with 50 % purple loading dye, run on a 1 % agarose gel stained with ethidium bromide and gel extracted as described below in section 2.2.4.1.

2.2.3.3 Nebuilder ligation and transformation

Neomycin and puromycin fragments were then combined with the gene drives digested for *Avr*II and *Not*I, in the NEBuilder ligation reaction. The reaction was set up as follows; 2 µl enzyme, 5 µl buffer, 50 ng vector and insert in a 1:2 vector:insert ratio. The NEBio Calculator (version 1.15.0) was used to determine amount of insert. The reaction was made up to a final volume of 20 µl with water and allowed to incubate at 50˚C for 1 hr. This was transformed into *Stbl E. coli* cells. The transformation reaction took place as follows; 2 µl of ligation was added to 50 µl of *Stbl* competent cells, this was then swirled gently with a pipette tip. Cells were placed on ice for 30 minutes before heat shocking at 42 ˚C for 30 seconds. After this, the cells were placed back on ice for 5 minutes and were mixed with 900µl NEB *stbl/10 beta* outgrowth media (NEB 9035S) and put in the shaking incubator for 1hr at 30˚C and 250rpm. 300µl of the transformed cells were then plated and left to grow at 30 ˚C overnight. The colonies were then grown in liquid culture and the plasmid DNA extracted as previously described in section 2.2.1.1.

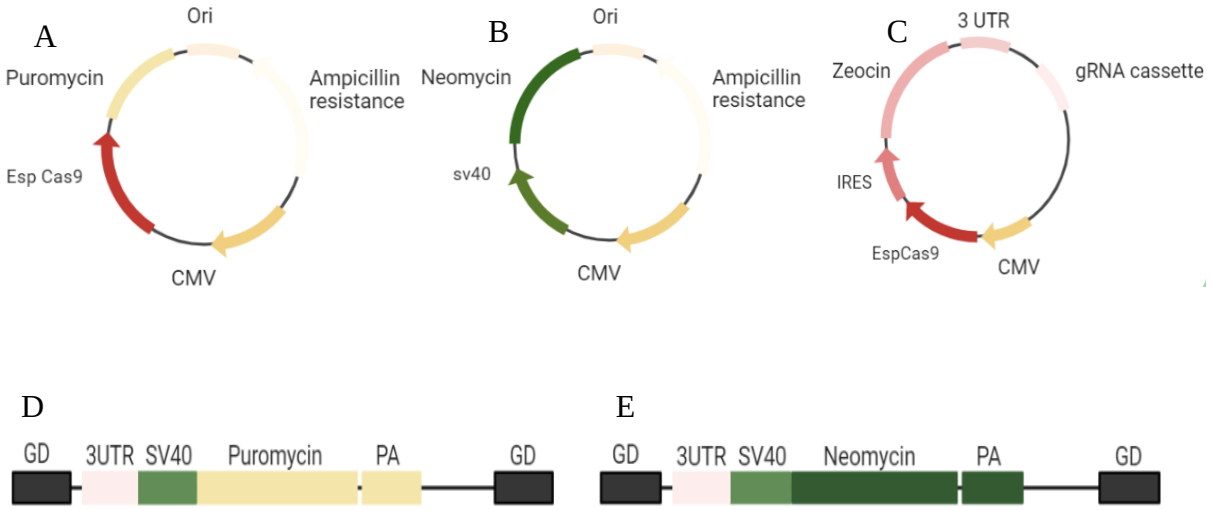


Figure 4: The original plasmids (A, B and C) used to amplify the necessary fragments to create the final Neomycin and puromycin resistance genes that were inserted into the OVA and OVM gene drives, respectively (D and E). The 3` untranslated region is indicated by the 3` UTR (C), the promoter for the antibiotic resistance by SV40 (B) and the poly A tail by PA (In A or B the PA is joined to the antibiotic resistance gene). Plasmid A contains the puromycin resistance gene and B contains the neomycin resistance gene. CMV indicates the general cytomegalovirus promoter, Ori is the origin of replication and IRES is the internal ribosome entry site.

2.2.4 Gel extraction and PCR cleanup for the purification of gene drive fragments

2.2.4.1 Gel extraction

In order to isolate the necessary fragments needed for downstream ligation reactions, gel extraction and purification was performed using the Zymo clean gel DNA recovery kit (Zymo research D4007/D4008). Gel purification was the chosen method of extraction in instances where the excised product size was above 20 bp, which is the limit for the alternate extraction protocol, PCR purification. Briefly, a blade was sprayed with ethanol and flamed. The gene drive fragment was then excised from the gel with the aid of a Dark reader, transilluminator. The gel fragment was placed in a 1.5 ml tube and weighed. Three volumes of binding buffer were added to the tube and the fragment was allowed to melt at 50 ˚C for 30 minutes or until the melted agarose had no solid chunks or strands visible. 1 volume of water was then added if the excised fragment was greater than 8kb. The dissolved agarose was loaded into a binding column and spun at 16000 xg for 1 minute. Wash buffer was added at a volume of 200 µl, and the sample was spun at 16000 xg for 30 seconds. This step was repeated after which the binding column was placed in a clean 1.5 ml tube. Finally, elution buffer, preheated to 70 ˚C, was added to the column in two steps of 6 µl with a one-minute

wait before spinning for 1 minute at 16000 xg. Concentration was determined with a nanodrop.

2.2.4.2 PCR clean-up

In the case where any small fragments were removed from the gene drive (~ 20 bps), such as in the case of the homology arm digestion, the gene drive was PCR purified using the Monarch PCR clean-up kit (NEB T1030S) instead of being gel extracted. When PCR purification could be used instead of gel extraction it was always ideal as this method reduced the loss of DNA after purification and, thus made ligations more efficient. Briefly, binding buffer was added in a 2:1 volume ratio to the digested sample which was then spun through a collection tube at 16000 xg for 1 minute. The column was then washed twice with 200 µl wash buffer and spun for 1 minute to remove the buffer. Two steps of 6µl of elution buffer were passed through the column into a clean 1.5 ml tube. Elution in a two-step process was found to better the yield of DNA from the column.

2.2.5 Order of cloning

A brief overview of the various cloning steps performed to construct the final gene drives necessary for PGC transfection and future protein expression, is outlined below (Figure 5). In each case of the two gene drives that target different loci, different changes had to be made depending on what was already contained within the gene drive. The OVM targeting construct already contained Sycp1 cloned in by a PhD student and the OVA targeting gene drive already contained the OVA homology arms as these had to be included before antibiotic resistance insertion.

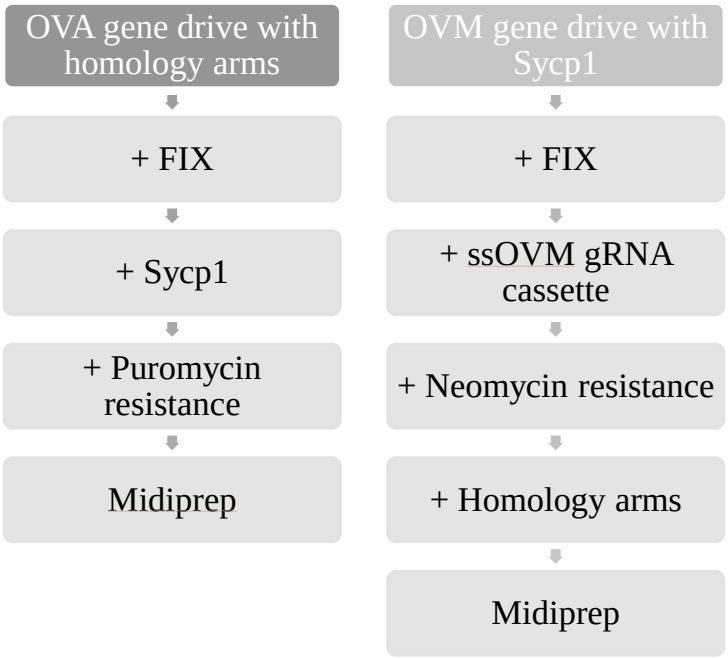


Figure 5: Brief overview of the cloning strategy used for the OVM, and OVA derived gene drives. Each step with a + indicates a new addition to the gene drive and the final step is the point at which the gene drive is midiprepped for transfection. The dark grey boxes at the top are the original OVA and OVM gene drives that already contain the homology arms and OVA gRNA cassette, and the Sycp1 promoter, respectively.

For the purpose of this report, a series of nomenclature was designed to keep track of the gene drives (Table 7) as the different changes were made. All changes were made to a gene drive that was originally synthesised by GenScript, in the pUC19 backbone and provided as a stab culture.

Table 7: Nomenclature used to represent the gene drive at each cloning step and what change was implemented in the gene drive. Changes are described under the "addition description" column and are incorporated into the nomenclature as a shorthand form of the word indicated under the "symbol column". The changes to the gene drive names can be followed under the OVM and OVA derived gene drive columns. Changes are represented from earliest at the top of the table to latest at the bottom of the table. In each case of the OVM and OVA gene drive column, the starting point for this project is indicated by * and the final gene drives are indicated by **.

OVM derived gene drive	OVA derived gene drive	Addition description	Symbol
OVM-Syc	OVA-HA	Starting point* Already contains Sycp1 and OVA HA, respectively.	Syc/HA
OVM-Syc-FIX	OVA-HA-FIX	FIX coding sequence	FIX
	OVA-HA-Syc-FIX	Sycp1 promoter	Syc
OVM-Syc-FIX-ss		ssOVM gRNA cassette	ss
OVM-Syc-FIX-ss-Neo	OVA-HA-Syc-FIX-Puro**	Neomycin and puromycin resistance genes for OVM and OVA, respectively.	Neo/Puro
OVM-HA-Syc-FIX-ss-Neo**		The 5` and 3` OVM homology arms.	HA

2.2.5.1 Insertion of FIX and Sycp1

The original OVM-Syc and OVA-HA gene drives were digested with *Mlu*I (R0198S) and *Sal*I (R0138S) to remove the Tenecteplase coding sequence, which the gene drives were synthesised with, and replace it with the FIX coding sequence that was obtained in my Honours year. When the sequence was amplified from HepG2 cell RNA in 2020, a His-Tag was included on the 3`end of the FIX coding sequence. This was different from the traditional appendage to the 5` end as the end of the protein is removed during its functional pathway in

the clotting cascade. The human FIX sequence is required for future protein expression within PGC cells implanted in chickens.

In addition to the removal of Tenecteplase from the gene drive, the FIX sequence contained in the cloning plasmid, pMini T 2.0, was also digested with *Mlu*I (R0198S) and *SalI* (R0138S) and gel extracted as explained above. This together with the two gene drives devoid of the Tenecteplase sequence, were placed in separate ligation reactions. After incubation, the ligation was transformed into *Stbl E. coli*, grown in liquid culture and plasmid extracted. After extraction, the desired insert in both OVA-HA-FIX and OVM-Syc-FIX was confirmed through restriction digest using the enzymes *Mlu*I and *Sal*I.

Next, Sycp1, which is the meiosis-specific promoter needed for meiosis-specific expression of Cas9 in PGCs *in vivo*, was inserted into the OVA-HA-FIX gene drive using enzymes *Spe*I (NEB R3133) and *XhoI* (NEB R0146S). The promoter only had to be inserted in the case of the OVA derived gene drive, as the OVM derived gene drive already had the promoter cloned in. This step was necessary to remove and replace the CMV promoter already contained in the gene drive as this promotes general expression of Cas9. A check digestion was then done on OVA-HA-Syc-FIX after insertion, using *Spe*I (NEB R3133) and *Xho*I (NEB R0146S), to confirm the presence of Sycp1. The Sycp1 fragment did not need to be extracted from a cloning plasmid as there was already gel extract available in the lab.

A visual aid demonstrating the cloning procedure (Figure 6, 7 and 8), which includes restriction digest and ligation, to create the OVA-HA-Syc-FIX and OVM-Syc-FIX gene drives, is provided below. These figures detail the insertion of the FIX sequence and Sycp1 into the OVA-HA gene drive and the insertion of FIX into the OVM-Syc gene drive.

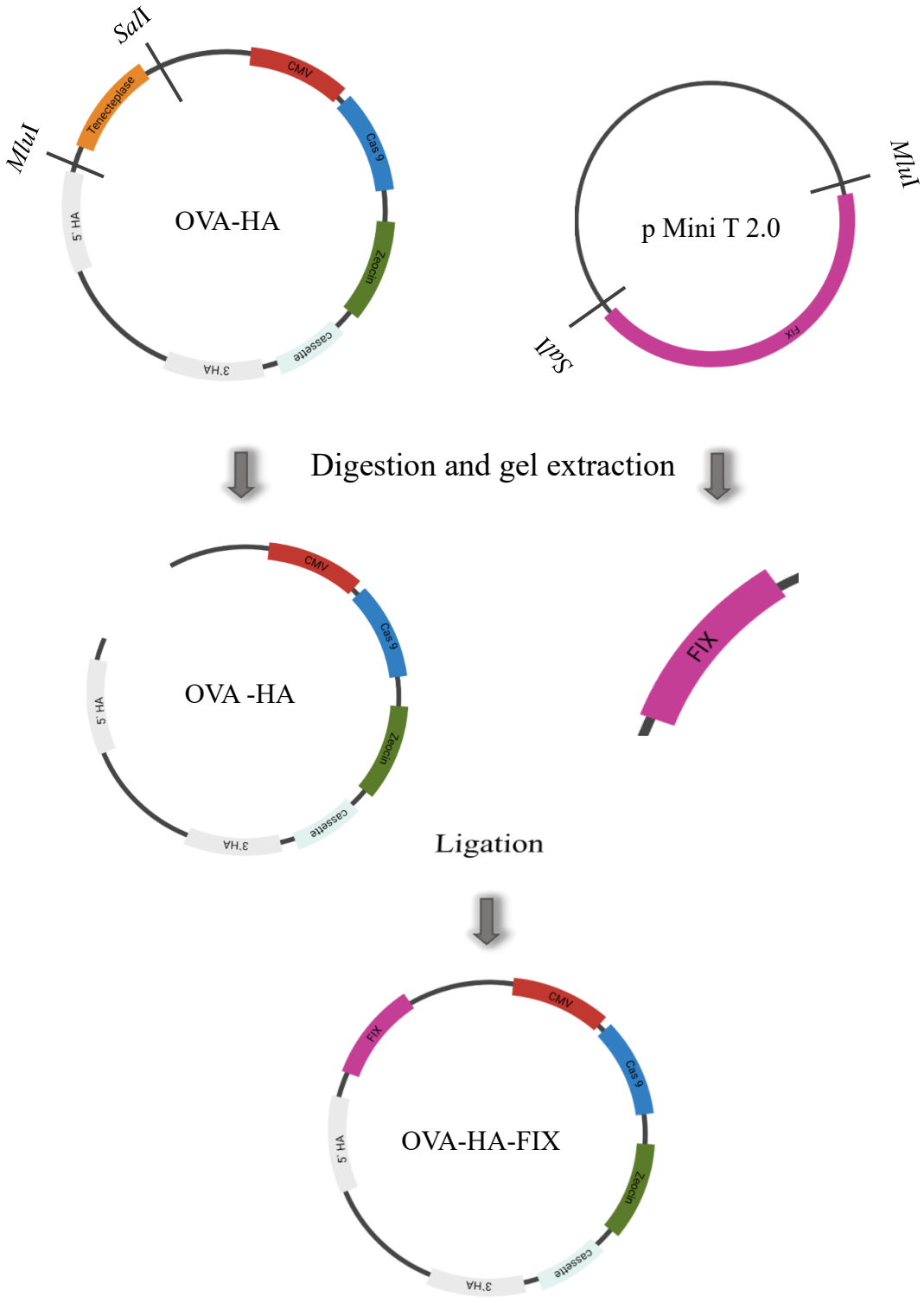


Figure 6: Cloning step for insertion of FIX into the OVA-HA gene drive. This figure demonstrates the digestion and ligation steps required to insert FIX into the relevant gene drive.

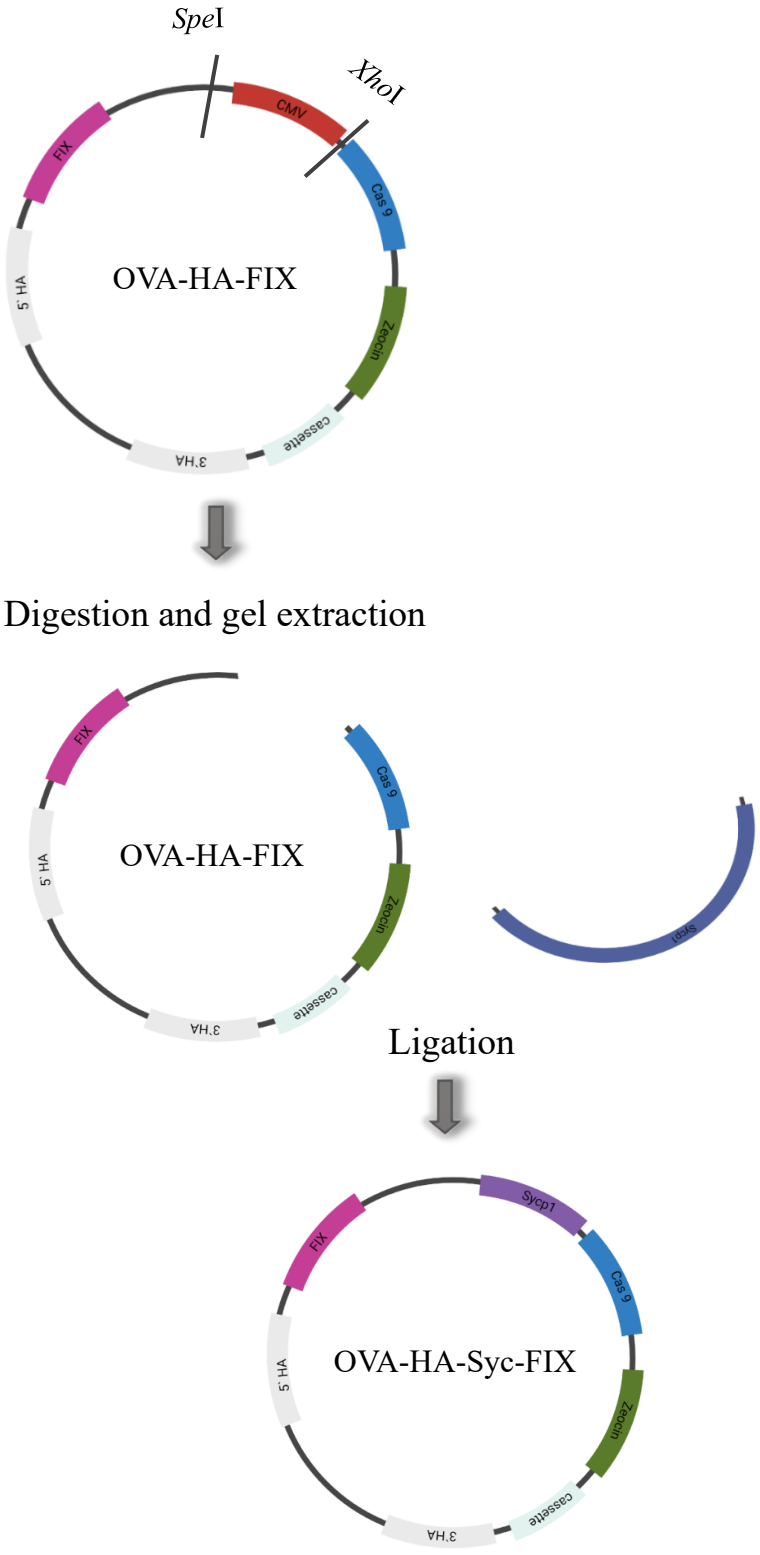


Figure 7: Cloning step for insertion of Sycp1 into the OVA-HA-FIX gene drive. This figure demonstrates the digestion and ligation steps required to insert Sycp1 into the relevant gene drive. In this case, Sycp1 was already provided as gel extract and did not need to be removed from a cloning plasmid.

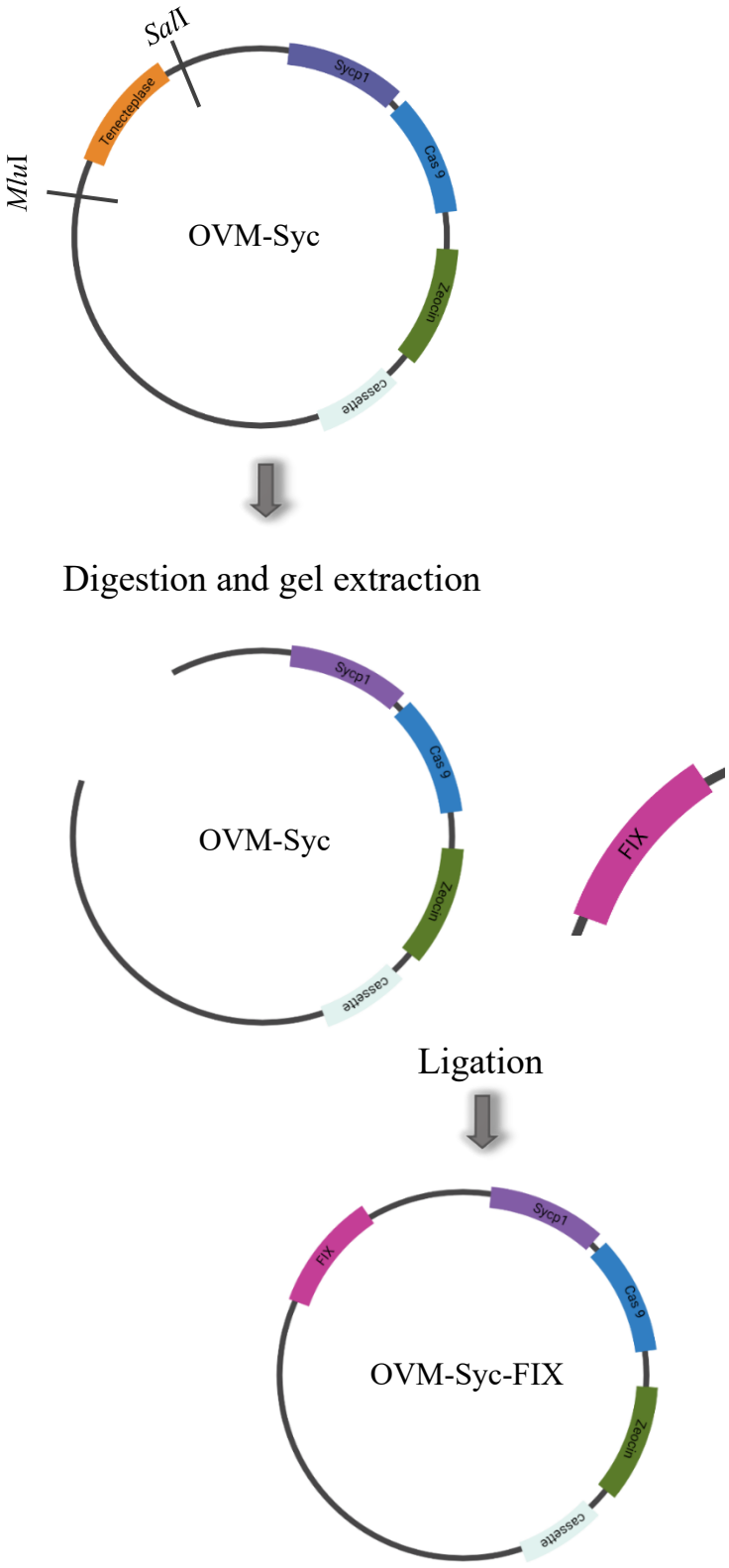


Figure 8: Cloning step for insertion of FIX into the OVM-Syc gene drive. This figure demonstrates the digestion and ligation steps required to insert FIX into the relevant gene drive. In this case, FIX was already provided as gel extract as it was extracted previously as seen in Figure 6.

2.2.5.2 Insertion of the ssOVM gRNA cassette

In order to replace the OVA, OVM and OVT gRNA cassette present in the original OVM gene drive, the OVM derived gene drive was digested with *BsteII* (NEB R3162S) and *NotI* (NEB R0189S). Replacing the old gRNA cassette in the case of the OVM-Syc-FIX gene drive was necessary as originally both gene drives contained a cassette that had gRNAs to cut at OVA, OVM and OVT, as this formed part of the original project plan. If all three gRNAs were kept in the gene drive, they would have caused unnecessary cutting at loci that were no longer the target. The OVA-HA-Syc-FIX gene drive already contained the OVA gRNA cassette, however, the OVM-Syc-FIX gene drive needed to be fitted with a new ssOVM (spidersilk OVM) gRNA cassette. The ssOVM cassette differs from that of the OVA gRNA cassette as it contains additional tRNA sequences that are important for a different project that is running in the lab that focuses on producing spider silk *In Ovo*.

The insertion of the gRNA cassette had to occur before the insertion of the Neomycin resistance gene as the restriction site of *Not*I is not conserved in the NEBuilder primers and would, thus be unusable after resistance cloning had taken place. Digests were also set up on the pUC19 cloning plasmid, which contained the ssOVM cassette, using enzymes *Bste*II (NEB R3162S) and *Not*I (NEB R0189S). All the fragments were then gel extracted and ligated as explained for previous cloning experiments. The ligation was then transformed into *Stbl E. coli* and liquid cultures were plasmid extracted. The desired insert was then confirmed using *Bste*II and *Not*I to digest the final plasmid (OVM-Syc-FIX-ss) for the ssOVM cassette.

A visual aid demonstrating the cloning procedure (Figure 9), which includes restriction digest and ligation, to create the OVM-Syc-FIX-ss gene drive, is provided below. This figure details the insertion of the ssOVM gRNA cassette into the OVM-Syc-FIX gene drive.

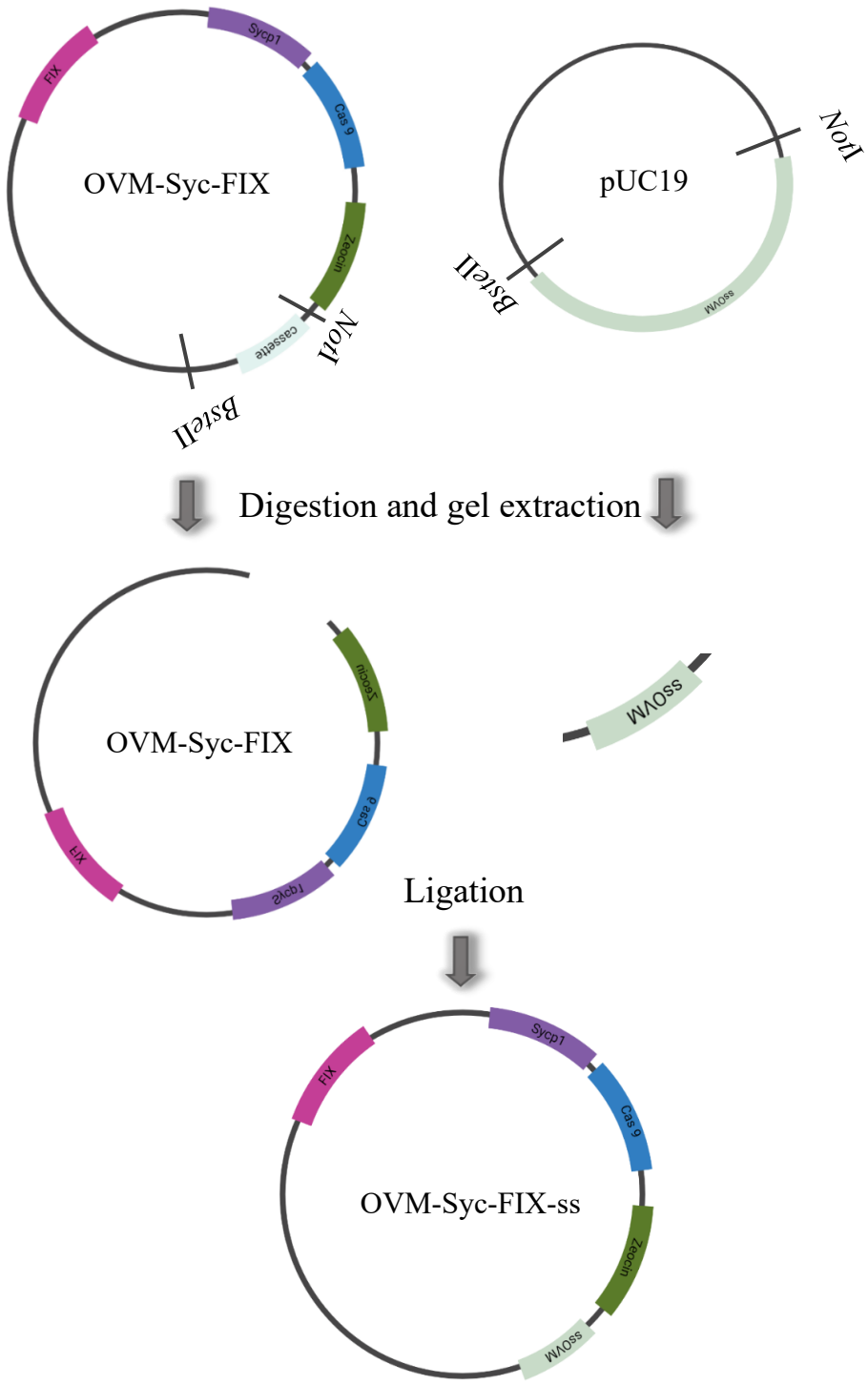


Figure 9: Cloning step for insertion of the ssOVM gRNA cassette into the OVM-FIX-Syc gene drive. This figure demonstrates the digestion and ligation steps required to insert the cassette into the relevant gene drive.

2.2.5.3 Puromycin and neomycin resistance insertion

In order to select for PGCs that had successfully taken up the gene drives upon transfection, the construct required a different antibiotic resistance gene for the OVA-HA-Syc-FIX gene drive and the OVM-Syc-FIX-ss gene drive, especially if the activity of these constructs was

to be assessed separately. In addition to this, the original antibiotic resistance gene (zeocin) placed in the gene drive needed to be replaced as its mechanism of action might have been detrimental to future embryos[75].

Using the NEBuilder HiFi DNA Assembly Cloning kit (New England biolabs E2621S) and restriction digestion, neomycin and puromycin resistance genes were inserted into the two separate gene drive constructs. Firstly, constructs were digested with *Avr*II and *Not*I overnight (R0174S and R0189S) to remove the zeocin resistance gene fragment. Gel extraction was then performed followed by ligation of the gene drives with the fully amplified resistance genes in section 2.2.3.2. Check digests were then run on the edited gene drives to confirm insertion of the desired fragment using the enzymes of *Xho*I and *Bste*II.

A visual aid demonstrating the cloning procedure (Figure 10 and 11), which includes restriction digest and ligation using the NEBuilder HiFi enzyme, to create the OVM-Syc-FIX-ss-Neo and OVA-HA-Syc-FIX-Puro gene drive, is provided below. This figure details the insertion of the the neomycin and puromycin resistance genes into the OVM-Syc-FIX-ss and OVA-HA-Syc-FIX gene drives, respectively.

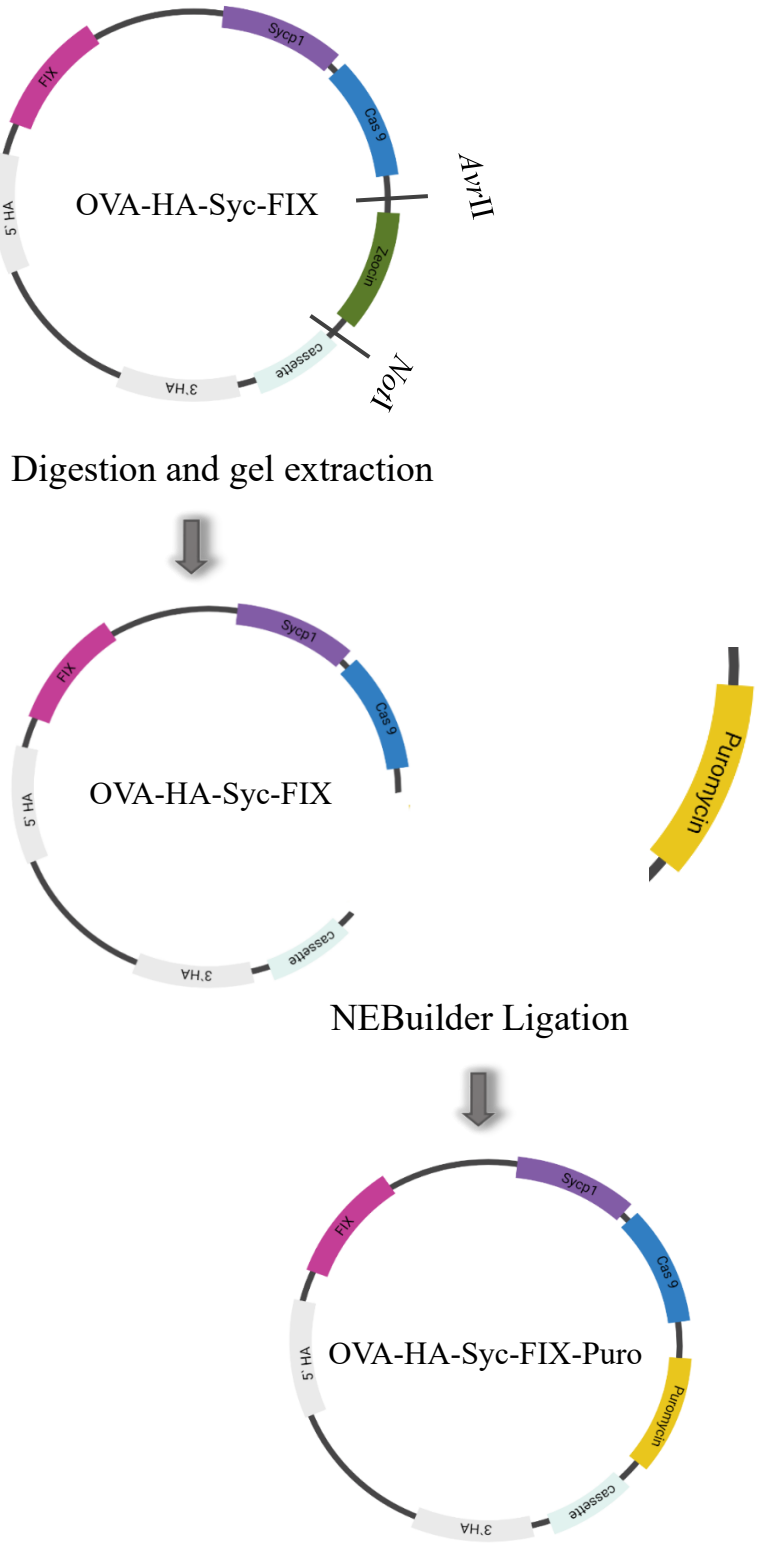


Figure 10: Cloning step for insertion of the puromycin resistance gene into the OVA-HA-FIX-Syc gene drive. This figure demonstrates the digestion and ligation steps required to insert the resistance gene into the relevant gene drive. The puromycin resistance gene was obtained from PCR amplification in section 2.2.3 and the overlapping regions complementary to the gene drive on the 5` and 3` end allow for ligation.

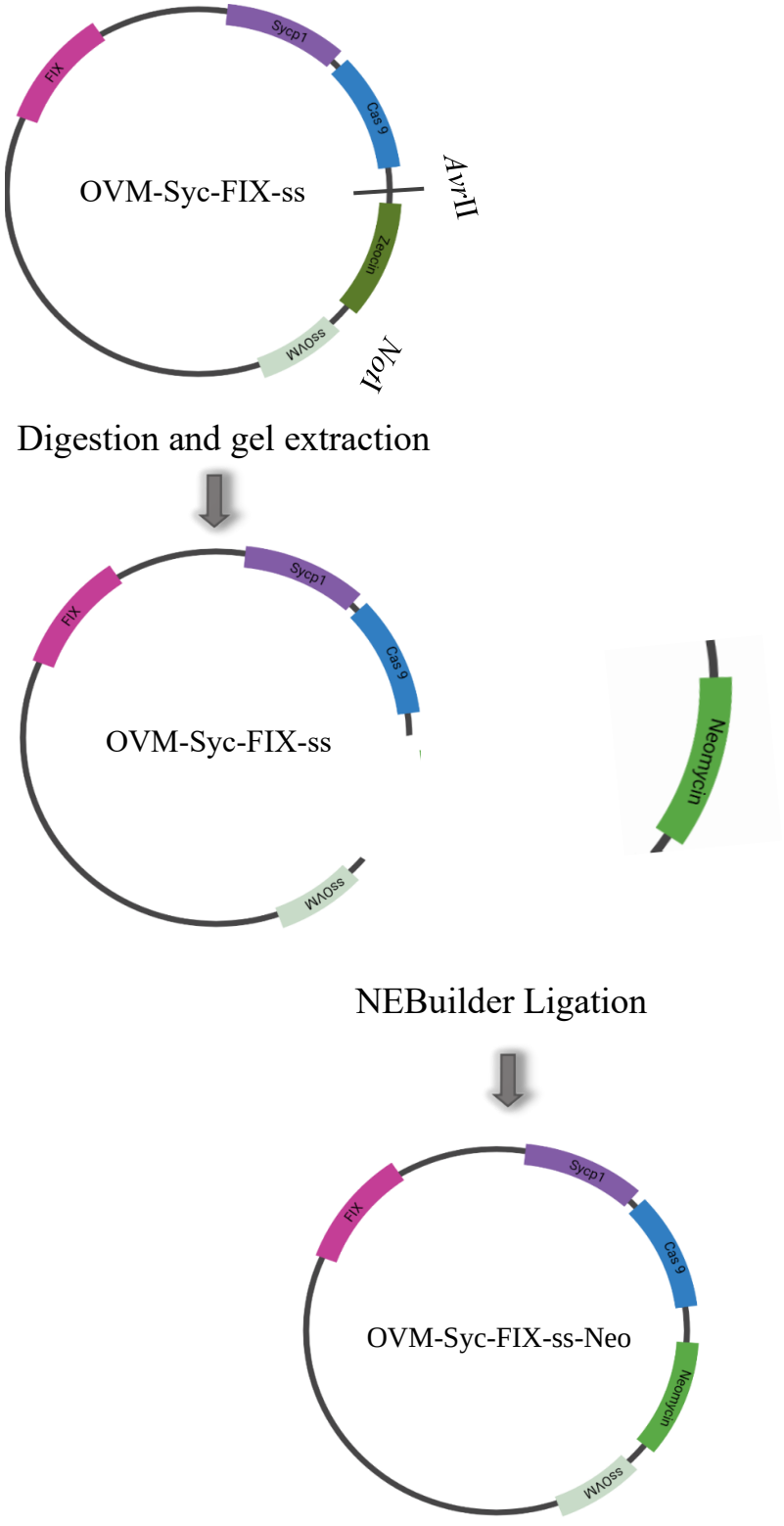


Figure 11: Cloning step for insertion of the neomycin resistance gene into the OVM-FIX-Syc-ss gene drive. This figure demonstrates the digestion and ligation steps required to insert the resistance gene into the relevant gene drive. The neomycin resistance gene was obtained from PCR amplification in section 2.2.3 and the overlapping regions complementary to the gene drive on the 5` and 3` end allow for ligation.

2.2.5.4 Cloning of the homology arms onto the OVM GD

In order to attach the 5` and 3` homology arm sequences to the OVM-Syc-FIX-ss-neo gene drive, that are necessary for allowing HDR integration into PGCs in future transfection experiments, the gene drive was digested using *Nhe*I (R3131S) and *Asis*I (R0630S). This fragment was then PCR purified instead of gel extracted. The 5`-homology arm was then digested out of the pUC 19 plasmid using *Nhe*I and *Asis*I and was ligated into the gene drive. Check digests were performed using the cloning enzymes before digesting the gene drive for *PacI* (R0547S) and *Bste*II to insert the 3`-homology arm. The gene drive was again purified using PCR cleanup and the 3`-homology arm was similarly digested and gel extracted from the pUC19 cloning vector. The ligation and transformation procedures were the same as for all other cloning. The successful insertion was confirmed using the cloning enzymes *PacI* and *BsteII*. Homology arms only needed to be cloned onto the OVM-HA-Syc-FIX-ss-neo gene drive as the OVA-HA-Syc-FIX-puro gene drive was provided with homology arms. It was crucial that the OVA homology arms were cloned on before insertion of puromycin as the puromycin resistance gene contains a duplicated site for *Bste*II. Similarly, for the OVM homology arms which had to be cloned on after insertion of the Neomycin fragment as the gene contained the *Avr*II site.

A visual aid demonstrating the cloning procedure (Figure 12), which includes restriction digest and ligation, to create the OVM-HA-Syc-FIX-ss-Neo gene drive, is provided below. This figure details the insertion of the 5` and then 3` homology arms into the OVM-Syc-FIX-ss-Neo gene drive.

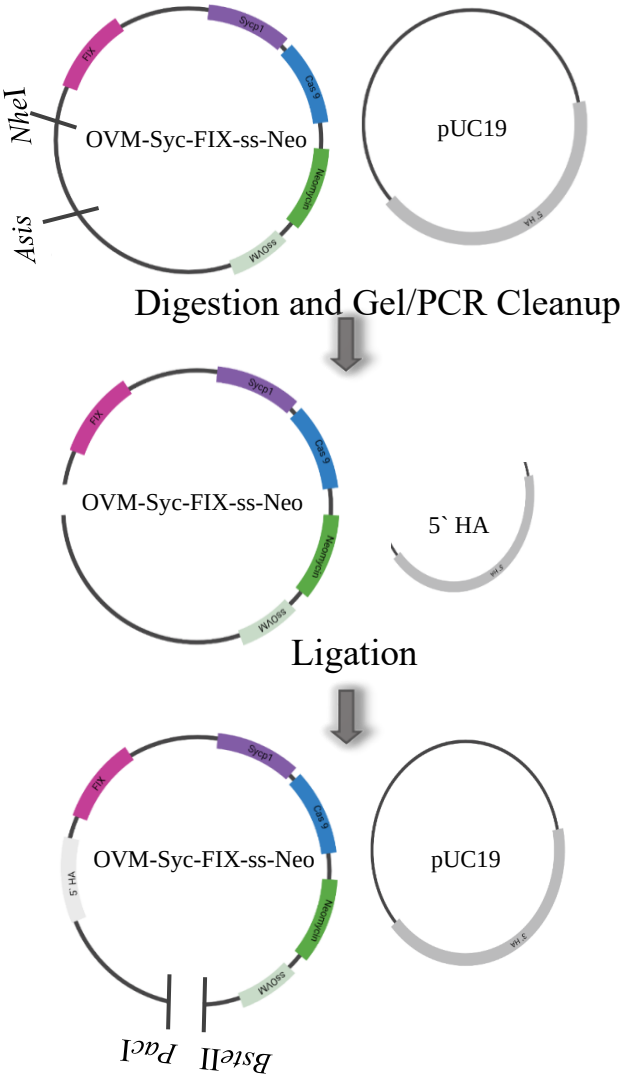

NheI
Asis
OVM-Syc-FIX-ss-Neo
pUC19
Digestion and Gel/PCR Cleanup
OVM-Syc-FIX-ss-Neo
5` HA
Ligation
OVM-Syc-FIX-ss-Neo
pUC19
BsteII
PacI
Digestion and Gel/PCR Cleanup
OVM-Syc-FIX-ss-Neo
3` HA
Ligation
OVM-HA-Syc-FIX-ss-Neo

Figure 12: Cloning step for the insertion of the homology arms into the OVM-FIX-Syc-ss-Neo gene drive. This figure demonstrates the digestion and ligation steps required to insert the 5` and 3` homology arms into the relevant gene drive. The gene drives were PCR purified and the homology arms were gel purified.

2.3 Animal ethics and general egg care

Ethics for the use of White leghorn chicken embryos was obtained to allow for PGC extraction, PGC genome modification and injection of modified PGCs into embryos no older than 10 days (HH stage 36). Ethics was obtained from the WITS Animal Ethics Research Committee (2021/09/02/O). Ethics to run in a BSL laboratory 1 was obtained from the Wits Institutional Biosafety Committee (20190801Lab). In addition to this, the laboratory (GH708) was granted clearance by the Department of Agriculture, Forestry and Fisheries to perform genetic modification (in bacteria, cells, and chickens) under the Genetically Modified Organisms Act, 1997, with registration number 39.2/Witwatersrand – 20/165 (03/06/2020-03/06/2023).

All specific pathogen free eggs were obtained from AviFarms through the Wits Medical School, CAS animal unit. Eggs were incubated in a sterile commercial egg incubator (Pleysier incubators) at 37 ˚C with 50 % relative humidity. Eggs were not allowed to develop past day 10 (HH stage 36) and embryos were harvested and euthanised through decapitation before this time. All equipment for dissection was sterilised by autoclaving and eggs were always incubated horizontally.

2.4 Cell culture and establishing primary cell lines

All cell culture and dissection work necessary for establishing primary cells lines, unless otherwise stated, was done in a biosafety cabinet (BSC), model Esco AC2-4S8. All equipment used was autoclaved and the working area was thoroughly sterilised with 70 % ethanol before and after use. Cells were maintained in a NuAire DH AutoFlow 5500 Air Jacketed CO2 Incubator at 37 ˚C and 5 % CO2. All cell counts were done on a TC20 automated cell counter (Bio-Rad) or Neubauer haemocytometer.

2.4.1 Establishing a CEF primary culture and cryopreservation of CEFs

To establish the line of chicken embryonic fibroblast cells to be used as a feeder layer for PGCS, White Leghorn chicken embryos were incubated at 37 ˚C and 60 % humidity for 9-10 days (HH stage 36). For dissection, eggs were sprayed with 70% ethanol and wiped down before embryos were removed (approximately 3) with sterile scissors and tweezers and the viscera, limbs and head were cut off. The torso of the embryo was then macerated in a petri dish and added to 10ml of 1 x trypsin/PBS. This was allowed to shake at 150 rpm for 15 min

at 37 ˚C. The larger pieces of tissue were then allowed to settle for 5 minutes, and the supernatant was removed and split equally into two 15ml tubes. The sample was then spun down at 400 xg for 4 minutes. The supernatant was removed, and the pellet was resuspended in CEF media, and the cells were plated in two t75 cm^2 flasks. CEF media was composed of 10 % FBS, 2 % chicken serum, 1 % L-glutamine, 1 % non-essential amino acids, 0.2 % primocin and 85.8 % DMEM.

The next day, cells were lifted using 1 x trypsin and placed at 37 ˚C for 5 minutes. Cells were then split in half to be passaged and seeded into a t75cm^2 flask or they were spun down at 400 xg for 4 minutes and the pellet was resuspended in CEF freezing media and placed in a Corning Costar cell cryopreservation tube. CEF freezing media was composed of 10 % DMSO, 10 % FBS, 2 % chicken serum and 78 % DMEM. Cells that were to be frozen were then placed in an isopropanol container and left at -80 ˚C overnight. Cells were then placed in a liquid nitrogen reservoir the next day for indefinite storage.

2.4.2 Preparation of the CEF feeder layer

To establish the feeder layer, CEFs were lifted with 1x trypsin PBS at 100 % confluency at 37 ˚C for 5 minutes. 1 x FBS was added to inactivate trypsin treatment and cells were spun down at 400 xg for 4 minutes. The pellet was then resuspended in fresh media and a cell count was taken using 10 µl cell suspension and an equal volume of trypan blue loaded onto a Dual-Chamber cell counter slide (Bio-Rad, 1450019). Cells were seeded at 1×10^6 into 6 well plates and cultured in CEF media. The cells were then inhibited with 10 µg/ml Mitomycin C (Merck, M4287-2MG) diluted in 2 ml culture media for 4 hours at 37 ˚C when they reached 80 % confluency. After inhibition, media was discarded, and cells were washed twice with PBS. Finally, fresh media was supplemented into the well and the feeder layer was ready for use.

2.4.3 Establishing PGC culture from germinal crescent extraction

PGCs were originally extracted from the germinal crescent of HH stage 5 embryos as this allowed for the shortest incubation time to obtaining PGCs. For dissection, eggs were wiped with 70 % ethanol before a 4 cm hole was created on the side of the eggshell. PGCs were the extracted by cutting out the germinal crescent with sterile scissors. Once the germinal crescent was excised, it was placed into a media filled well in a 24 well plate until each sample could be sexed before combining males and females. Sexing was done by removing the tail section and processing with crude DNA extraction lysis buffer (Recipe 2). Tail tissue was then combined with proteinase K and Crude lysis buffer and incubated at 55 ˚C for 5

minutes. The sample was then vortexed and further heat denatured at 55 ˚C for 15 minutes and 95 ˚C for 5 minutes. This was then vortexed for 1 minute in a Minifuge and the supernatant was added as sample to a PCR tube. The tables (Table 8 and 9) below show the PCR reaction setup and PCR conditions for sexing PCR.

Table 8: PCR reaction set-up for sexing of embryos. This protocol is based on Clinton et al[121].

Component	Volume
Sample	1 µl
Master mix	5 µl
18s primer (0.5 µM)	0.5 µl
W primer (1 µM)	1 µl
dH2O to	10 µl

Table 9: PCR reaction conditions for sexing PCR. At each step the temperature and time is given. This protocol is according to Clinton et al[121].

Initial denaturation	Denaturation	Annealing	Elongation	Final elongation	Cycles
94 ˚C 2 min	94˚C 5 sec	56 ˚C 5 sec	72 ˚C 5 sec	72 ˚C 5 sec	25

Cells were cultured for 3 weeks in PGC media before immunocytochemical (ICC) analysis was performed (section 2.4.6). PGC media consisted of the following: DMEM (Gibco) to 250 ml , 2.5 % chicken serum, 10 % FBS, 0.1 mM β-Mercaptoethanol, 1 X Primocin, 1 X NEAA, 1 X Glutamax, 10 ng/ ml BFGF and 10 ng/ ml LiF. Media was replaced the day after plating and then every 2nd day from there onwards.

2.4.4 Establishing PGC culture from blood extraction

PGCs were then extracted from the blood of HH stage 14-17 embryos using pulled glass needles created with a Narishige puller (27-9 Minami- Karasuyama 4-chome setagaya-ku), attached to a mouth pipette. This method was attempted in order to increase the number of PGCs extracted initially, to encourage faster growth in culture. After incubation at 37˚C and 50 % humidity for approximately 3 days, eggs were wiped down with 70 % ethanol and a 4 cm hole was made at the top of the horizontal egg on the side of the embryo, on a sterilised benchtop. Eggs were then placed under a light microscope and pulled glass needles were used to pierce various vessels. Blood was obtained from the dorsal aorta, heart, and vitelline arteries of embryos (see Figure 13). Approximately 30µl of blood was obtained from 100 males in batches of 20 across the course of the project. The blood was then added to 80µl of PGC culture media composed of: DMEM (Gibco) to 250 ml , 2.5 % chicken serum, 10 % FBS, 0.1 mM β-Mercaptoethanol, 1 X Primocin, 1 X NEAA, 1X Glutamax, 10 ng/ml BFGF,

10 ng/ml LiF, 5 ng/ml hSCF. The same sexing protocol was carried out for each sample as described above for germinal crescent extraction (Table 8 and 9) except this time the entire head was excised as sample tissue for PCR. Once the sex was known, all females and all males were pooled together in a single well in a 6 well plate on a CEF feeder layer and were maintained for approximately 3 more weeks in PGC media before staining took place.

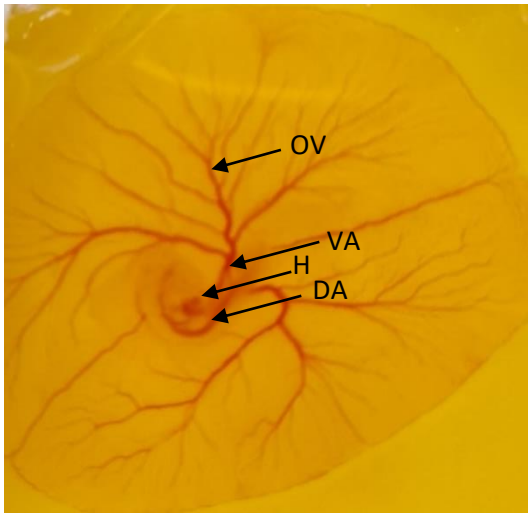


Figure 13: Stage 15 White Leghorn chicken embryo indicating where blood was drawn from for PGC extraction. DA indicates the dorsal aorta, H indicates the heart and VA indicates the vitelline arteries. OV indicates the outer vitelline vessels. Embryos were windowed and then imaged under a light microscope.

2.4.5 Passaging of PGCs

PGCs were passaged every week in order to encourage colony formation. This was done using 1% trypsin in PBS. Cell media was spun down and the pellet was transferred to a fresh feeder layer. In addition to this, the remaining cells were washed with 1 x PBS and then covered in 2 ml of 1x trypsin. Cells were allowed to incubate for 10 minutes at 37 ˚C or until all cells had lifted. Trypsin was then inactivated with equal volumes of culture media and the cells were pelleted at 400 xg for 4 minutes. Cells were then resuspended in fresh media and split in half onto freshly inhibited CEF feeder layers. By following this protocol, cells in suspension and those attached to the feeder layer were preserved.

2.4.6 Staining to identify PGCs

After three weeks of culture, cells were stained using Anti-EMA-1 (for germinal crescent cells) (Developmental studies Hybridoma Bank) and Anti-SSEA-1/MC-480 (for blood PGCs) (Developmental studies Hybridoma Bank) to identify the presence of PGCs in culture. These are germ cell specific surface antigens expressed during the primordial stage of germ cell development. Media was discarded and cells were washed gently with 2 ml of 1 x PBS, for 5

minutes (x3). Cells were then fixed with 4 % PFA for 15 minutes at room temperature. After fixing, washes were repeated and 4 % BSA was added to the cells to block for 1 hr at 37˚C. Primary antibody (Anti-SSEA-1 or EMA-1) was added in a concentration of 5µg/ml diluted in 4 % BSA. After 1 hour of incubation at 37˚C, the antibody was removed, and 3 more 5-minute washes were done. A 1/50 dilution of secondary antibody (Rabbit Anti Mouse IgG FITC, abcam; ab97045) was made in 4 % BSA and allowed to incubate at 37 ˚C for 1 hour. Secondary antibody was then discarded and 0.5µg/ml of DAPI (4′,6-diamidino-2-phenylindole) (Sigma-Aldrich) diluted in 4 % BSA was added to stain the nuclei of all cell types. This was allowed to incubate at room temperature for 15 minutes. DAPI was then removed and fresh PGC media was added to cover the cells. Staining was observed under white, blue, and green light with the FLoid Cell Imaging Station (Thermo Fisher Scientific). A negative primary antibody control stain was set up on CEFs using only secondary antibody. The procedure for the control was the same as outlined for the experiment.

2.5 Editing of primordial germ cells *in vitro*

2.5.1 Kill curves for neomycin and puromycin

To determine the best concentrations required for neomycin and puromycin selection on PGCs, kill curves were set up in triplicate wells. For this, 0.5×10^5 PGCs were seeded into 24 well plates on a feeder layer of CEFs at a density of 0.3×10^5. Experiments were prepared for concentrations ranging 0, 0.5, 1, 1.5, 2, 2.5, 3 µg/ml for puromycin and 0, 100, 200, 300, 400, 500, 600 µg/ml for neomycin. A control of CEFs treated at each of these concentrations for each antibiotic was prepared to account for the feeder layer in the kill curve. Cells were treated for 5 days in the case of puromycin and 7 days in the case of neomycin and media was changed every two days and supplemented with fresh antibiotic. A cell count was taken at the end of the treatment week and the CEF feeder layer was subtracted from the final cell count in each of the experiments. The PGC-only counts were then averaged as well as the triplicates for each CEFs-alone and PGCs-on-CEFs and plotted in a kill curve. All cell counts were done on a Neubauer hemocytometer under a light microscope.

2.5.2 Transfection of PGCs

In order to transfect the final gene drives as well as the mRFP-Tol2 tracer and eSpCas9 constructs into PGCs, electro gradients of 240 v 10 ms pulse length, 1s pulse interval and varied pulses: 1, 2 and 3, were tested on PGCs using only the mRFP-Tol2 tracer construct. All electroporation was done with the BTX Harvard ECM 830 Square Wave electroporator in electroporation cuvettes (BTX). Transfected PGCs that had been growing for a week, were gently pipetted off from the CEFs and any chunks were trypsinised for 5 minutes at 37˚C.

Trypsin was then inactivated with equal volumes of FBS after which the cells were spun down and resuspended in media. In the case of electroporation 1×10^5 cells were resuspended in Advanced MEM (Thermo Fisher 124901) in an electroporator cuvette and placed within the electroporator. Once electroporated, cells were added directly to prewarmed PGC culture media in a 6 well plate containing a fresh feeder layer. Cells were then observed under the FLoid Cell Imaging Station (Thermo Scientific) after 3 days in culture and the cells that looked the healthiest and had the most fluorescence were determined to have undergone the best electroporation conditions and were used for repopulation.

Once the ideal settings were determined, the gene drive, eSpCas9 vector and RFP tracer construct were electroporated into the cell at a total of 5 μg in a 3:1:1 ratio (Gene drive: mRFP-Tol2:eSpCas9 vector). In addition to this, the constructs were also lipofected into the cells following the lipofectamine 3000 protocol (Thermo Fisher Scientific). Briefly, cells were pipetted off the feeder layer and trypsinised in 1 x trypsin, if needed as stated above. Cells were then resuspended in PGC culture media and added to a fresh feeder layer in a 6 well plate. Briefly 3.75 μl and 7.5 μl lipofectamine 3000 reagent were made in separate tubes to a final volume of 125 μl in Advanced MEM. DNA mixture (5 μg) was then diluted in 10 μl p3000 reagent and made to a final volume of 250 μl Advanced MEM. The separate 125μl diluted lipofectamine 3000 reagent was then added to 125 μl diluted DNA in p3000 reagent. This was allowed to incubate for 15 minutes at room temperature. The reaction was then added to the suspended cells in the 6 well plate and left at 37 ˚C overnight. The next day media was removed for both electroporated and lipofected cells and cells were treated with 7.5 μM RS-1 and 0.1 μM SCR-7 for 24 hours. Neomycin and puromycin at a concentration of 500 and 400 μg/ml, and 1.75 and 1.5 μg/ml, respectively, were added to the media and selection took place for 2 days with one concentration then a further 2 with the next. Cell media was changed at the end of selection, and cells were allowed to recover for 4 days. Finally, cells were counted per microscope field of view (N=3) to quantify the best method of transfection. The lipofection and electroporation conditions above were used for the transfection of circular gene drives, however, the transfection of the linearized gene drives was done with only the electroporation conditions.

2.5.3 Assessment of gene drive integration in PGCs

After 4 days in culture, PGCs were pipetted off the feeder layer and spun down at 1000 xg for 1 minute. The DNA of the PGCs was then extracted using the Monarch genomic DNA purification kit by New England Biolabs (T3010s). The cell pellet was resuspended in 100 μl cold PBS and 1ul Proteinase K and 3 μl RNase A were added to the suspended cells and

vortexed briefly. Cell lysis buffer was then added and was mixed in by vortexing thoroughly. The sample was incubated at 56 ˚C in a heating block for 5 minutes with vortexing every minute. Binding buffer was then added at a volume of 400 µl, and the mix was then passed through a binding column at 1000 xg for 3 minutes followed by 13000 xg for 1 minute. 500 µl gDNA wash buffer was added to the column and spun through at 13000 xg for 1 minute. After the wash step was repeated a second time, the column was placed in a 1,5 ml tube and 40 µl of elution buffer, preheated to 60 ˚C, was added to the column, allowed to incubate for 1 minute, and spun through at 14000 xg for 1 minute.

Integration of the gene drives was assessed in PGCs using primers as laid out in Table 10, designed by Michael Edwards, a past MSc student. These primers overlap the region in the genome outside the cut site of the gRNAs as well as the 5` and 3` end of the gene drive. For PCR assessment, genomic DNA was set up in a reaction with Platinum SuperfiII DNA polymerase (Invitrogen, cat no. 12361010), according to the manufacturer's instructions. In a 10 µl reaction, 2 µl buffer was combined with 0.25 µl forward primer, 0.25 µl reverse primer, 0.2 µl dNTPs, 20 ng DNA, 0.2 µl enzyme, 5 % DMSO and made to a final volume of 10 µl. PCR conditions for these reactions were as follows:

Target	Primers 5`-3`	Denaturation	Denaturation	Annealing	Extension	Size	Cycles
5` OVM HA	`TCCCCTCGAAGGAAACCCATA`					2915	
	`CCTTGGGATTCCATTGACT`						
3` OVM HA	`ACGTTGTCACTGAAGCGG`					2691	
	`ACTGCTTCATCCATCTGCCC`						
5` OVA HA	`ACCTGCGTGATACCCCCTAT`					3202	
	`CCTTGGGATTCCATTGACT`	98 ˚C 30s	98˚C 10s	60˚C 10s	72˚C 30s/kb		25
3` OVA HA	`GCAAGAACTCTTCCTCACGC`					2508	
	`CCTCTGTGTGCTGCTCCTAC`						
NHEJ OVM	`GTGAAACCGCCAAGAAACCC`					1793	
	`CAGAGTGCGAGATGAACCGA`						
NHEJ OVA	`AAGGGTGAAGTACCTGCGTG`					623	
	`CAAGGTCTGTCCCATCACCC`						

Table 10: Primers and PCR conditions used to screen edited PGCs for gene drive integration at the 5` and 3` homology arms as well as to screen the different gene drives for NHEJ insertion. In each case of the primers given, the forward primer is given first then the reverse primer. The reactions differ by the size of the fragment and extension time.

Genomic DNA then underwent further PCR to confirm the presence of the desired plasmids in the cells after electroporation and lipofection (Table 11). The cells were screened for puromycin resistance and the gene drive backbone to confirm the OVA-HA-Syc-FIX-Puro gene drive and neomycin resistance and the backbone to confirm the OVM-HA-Syc-FIX-Neo gene drive. Cells were also screened for the 5`end of the eSpCas9-T2A-Puro plasmid.

Table 11: Primers and PCR conditions used to screen edited PGCs for gene drive presence as well as eSpCas9-T2A-Puro presence. In each case of the primers given, the forward primer is given first then the reverse primer.

Target	Primers 5`-3`	Denaturation	Denaturation	Annealing	Extension	Size (bps)	Cycles
Neomycin resistance	`AATAGAGGCTCTGA `TCGGCAGTGCCTGA					1316	
Puromycin resistance	`AAGTGCACACAAT `TCGGCAGTGCCTGA	98 ˚C 30s	98˚C 10s	60˚C 10s	72˚C 30s/kb	1249	25
Backbone	`ACCTCTGACACATG `TTCGAGCTCGGTAC					385	
eSpCas9-T2A-Puro	`ACCTCTGACACATG `ACGCAGCGACTCCC					1800	

2.6 Gonadal repopulation of partially sterilized embryos with edited PGCs

2.6.1 Busulfan ablation of native PGCs

Fertilized, unincubated White Leghorn eggs were obtained from AviFarms for injection with various busulfan concentrations to determine the best concentration for removal of native PGCs from embryos to prepare them for repopulation. First eggs were allowed to acclimatize at room temperature overnight (N=68). Eggs were then wiped down with 70 % ethanol and busulfan/control mixtures were prepared as follows; oil/DMSO controls consisted of 50 % sesame seed oil and 50 % DMSO (v/v). For the experiment concentrations such as 65 µg/50µl: busulfan was dissolved in 1ml DMSO then emulsified with 1ml sesame seed oil (Endomex, wholesale) to give a concentration of 65 µg in 50 µl. The same was done to prepare the 75 µg/50µl and 85 µg/50µl concentrations of busulfan in sesame seed oil and DMSO. Emulsification was done by passing the mixture through a 0.45 µm syringe filter 50 times and then vortexing in between use.

For busulfan administration, a hole was made on the sharp end of the egg approximately one centimeter from the top of the eggshell to avoid the air pocket (see Figure 14) and 50 µl busulfan/DMSO/oil mixture was injected into the yolk with a 29-gauge insulin needle. Each time a new egg was injected the syringe was washed out with PBS and a fresh 50 µl of busulfan/oil mixture was aliquoted into the tube. Syringes and aliquot tubes were changed between each concentration. Eggs were then placed in the incubator for 7.5 days (HH stage

31-33) at 37 ˚C and 50 % humidity. All Embryos were observed at day 7.5 as this provided a middle point where it would be late enough to see PGCs that should have settled and started repopulating but not too late for results to be skewed by native PGCs that had proliferated again.

To view, embryos were dissected with scissors by removal from the yolk and immediately decapitated. Gonads were removed while still attached to the back tissue and the side flaps, head, viscera, and limbs of the embryo were cut out with dissection scissors. The forelimbs of the embryos were kept for molecular sexing as described above for sexing of PGCs. Embryos were then fixed in 4 % PFA overnight at 4 ˚C then moved to 5 % sucrose for 5 hrs, 15 % sucrose overnight and 30 % sucrose for 6 hrs until the embryo sunk, all at 4 ˚C. Embryos were placed in cryomatrix embedding fluid (Mercedes Scientific: FIS 6769006) overnight and flash frozen in liquid nitrogen the next day. These samples were stored at -20 ˚C until sectioning.



Figure 14: Image demonstrating where the needle is positioned when piercing the egg to inject busulfan for ablation purposes. The circled area indicates the blastodisc.

Embedded embryos were then sectioned at 10 µm and placed on a positively charged slide to dry overnight. Slides were then washed in 1x TBST-D (Tris Buffered Saline TWEEN and DMSO) thrice, for 10 minutes to rehydrate the sections. Slides were blocked in 4 % BSA for 1 hour at room temperature and then incubated with primary antibody (SSEA-1 at 1:50) overnight at 4 ˚C. The next day, slides were washed 3 times for 10 minutes and incubated with secondary antibody, diluted 1/100 in 4 % BSA (Rabbit Anti Mouse IgG FITC), for 2 hours at 37 ˚C. After incubation, slides were washed twice for 10 minutes at a time and incubated with 5 ng/µl DAPI for 15 minutes at room temperature. All incubations were done in a humidified container containing 1x TBS. Slides were then viewed under the FLoid Cell

Imaging Station and PGCs were counted in 20 sections (10 for each gonad) for each experiment and control group, across 5 samples for each group (N=400). To determine viability, living and dead embryos were recorded for each group as well as the natural duds.

2.6.2 Repopulation with edited PGCs

To repopulate embryo gonads with PGCs modified for the mRFP-Tol2 construct, embryos were ablated as described above using the 75 μg/50μl busulfan concentration. Eggs were wiped with 70% ethanol and a small hole was made in the top of the eggshell where busulfan was administered to the yolk (N=21). After 3 days of incubation, eggs were sexed with blood according to Clinton *et al*[121], and female embryos were discarded. mRFP-Tol2 transfected PGCs were pipetted off the feeder layer of CEFs and were injected into the vitelline arteries of day HH 13-15 embryos (Figure 13). Approximately 1000 PGCs suspended in 50μl of 1xPBS were injected into the embryos with a mouth pipetted and pulled glass needle. Eggs were then rehydrated with Ringer's solution (Recipe 3) and sealed with parafilm and placed into the incubator at 37˚C 50 % relative humidity. For viewing, embryos were dissected out of the egg at day 6 (HH stage 28-29) and the limbs, head and viscera were removed. Whole mount observations of mRFP-Tol2 edited PGCs in the gonads were done using the Zeiss AXIO Zoom.V16 with the AXIOCam MRm (ZEISS). PGCs were then counted across 5 embryos and the average number of PGCs in the embryo gonads was computed. Survival of the embryos was also noted.

2.7 Statistical analysis

Statistical analysis to determine the mean and the standard error was performed on excel (2021). Standard deviation/error is represented as SD±. Statistical significance was determined using a ONE-Way ANOVA and Tukey's Post Hoc test for ablation efficiency and the average number of PGCs in embryos. A Chi-square test was used to assess significance for viability (Excel, 2021). $P<0.01$ and 0.05 were considered significant.

3. Results

3.1 Restriction mapping demonstrates the successful construction of the gene drives

In order to accomplish one of the main aims of this project, which was to transfect chicken PGCs with a human FIX gene drive, the editing constructs that were available in the lab had to be modified. These changes included the addition of the human FIX coding sequence for expression *in ovo*, an Sycp1 meiosis-specific promoter for the timed expression of Cas9, neomycin and puromycin resistance genes in order to allow for selection of transfected cells, an OVM specific gRNA cassette, and the addition of the OVM homology arms for HDR integration into cells. Some changes had already been made prior to this project such as the addition of the OVA homology arms as well as an OVA specific gRNA cassette. In addition to making the necessary changes, final constructs then had to be prepared in a midiprep to meet concentrations necessary for transfection as well as remove traces of endotoxins.

Gene drive construction was accomplished through restriction digest and cloning in *Stbl E. coli* cells and the final products were subject to check digestions to confirm the desired insert was present. For each step performed, the cloning digests are shown as well as the final check digestion. After clones were confirmed to have the desired insert, sequencing was performed through either Inqaba Biotechnology or GenScript to confirm the sequence integrity. The sequencing results are available in the supplementary section (Figure 32-37).

Firstly, the FIX coding sequence was inserted into the OVM-Syc gene drive. This was done by digesting the FIX sequence out of the p MiniT plasmid (the cloning of which had been completed in my Honours year) demonstrated in Figure 15A (Lane 1), at 1.4 kb. In conjunction with this, Tenecteplase was removed from the OVM-syc gene drive (Figure 15A at 1.5 kb) and both p MiniT and the digested gene drive were gel purified, and the products were joined in a ligation reaction. Sycp1 digested out of a cloning plasmid is not shown as gel extract was obtained from a PhD student in the lab.

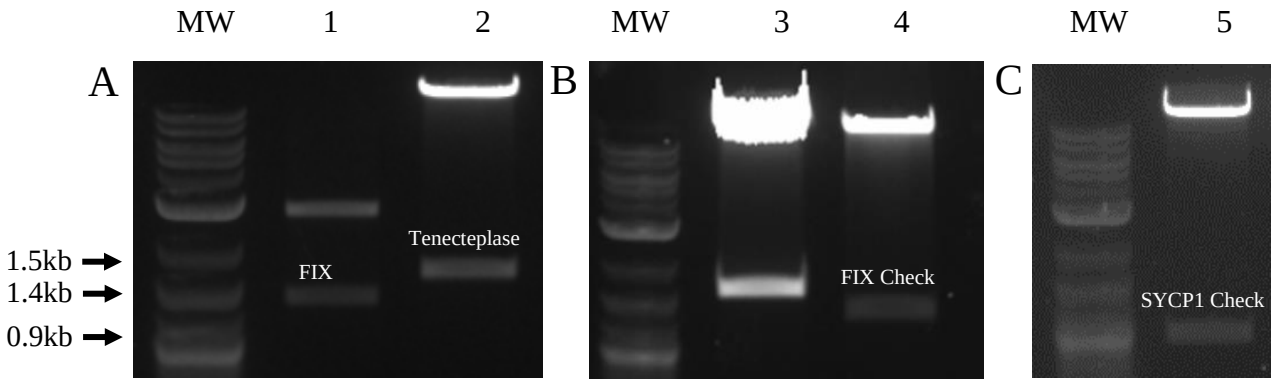


Figure 15: Digestions confirming the removal of FIX from p Mini T, the removal of Tenecteplase from the gene drive and the successful insertion of FIX and presence of Sycp1 in the OVM-Syc gene drive after cloning. Lanes are labeled at the top of the figure and lanes containing ladder are labeled as MW (NEB 1kb plus). (A) Lane 1 is the p-Mini T cloning plasmid (upper band) digested for FIX (indicated at 1.4kb), lane 2 is the gene drive (upper band) digested for Tenecteplase (indicated at 1.5kb), (B) lane 3 is a control digest of Tenecteplase (lower band) out of the gene drive (upper band) in order to differentiate the size of FIX in the next lane, lane 4 is a check digest on a clone confirming the insertion of FIX (indicated by 1.4 kb) into the genedrive, (C) lane 5 is a check digest of Sycp1 (indicated at 0.98 kb), on that same clone, which had already been inserted by a PhD student. Samples are run on a 1 % agarose gel stained with ethidium bromide.

After bacterial culture and plasmid extraction, a check digestion was performed to confirm the presence of FIX, using the same enzymes for cloning (*Mlu*I and *Sal*I). Figure 15B demonstrates the successful insertion of the FIX sequence as indicated at 1.4kb in lane 4. Figure 15C was a check digest performed on the same gene drive for the already inserted Sycp1 fragment, using enzymes *Spe*I and *Xho*I. The presence of the promoter was confirmed as shown in lane 5 at 0.98 kb. These gels confirm the creation of the OVM-Syc-FIX gene drive. In addition to this, sequencing results for the FIX and Sycp1 sequences in this gene drive can be found in the supplementary section, Figure 35.

Next, FIX and Sycp1 needed to be inserted into the OVA-HA gene drive in order to replace Tenecteplase and CMV which were included in the gene drive upon original synthesis. First, the gene drive was digested with *Mlu*I and *Sal*I to remove Tenecteplase as shown in Figure 16A at 1,5 kb and once insertion had been confirmed, *Spe*I and *Xho*I were used to remove CMV (shown in Figure 15A at 0.6 kb) to replace it with Sycp1. FIX obtained from the p Mini T cloning plasmid shown in Figure 15A above was used for the ligation reaction below.

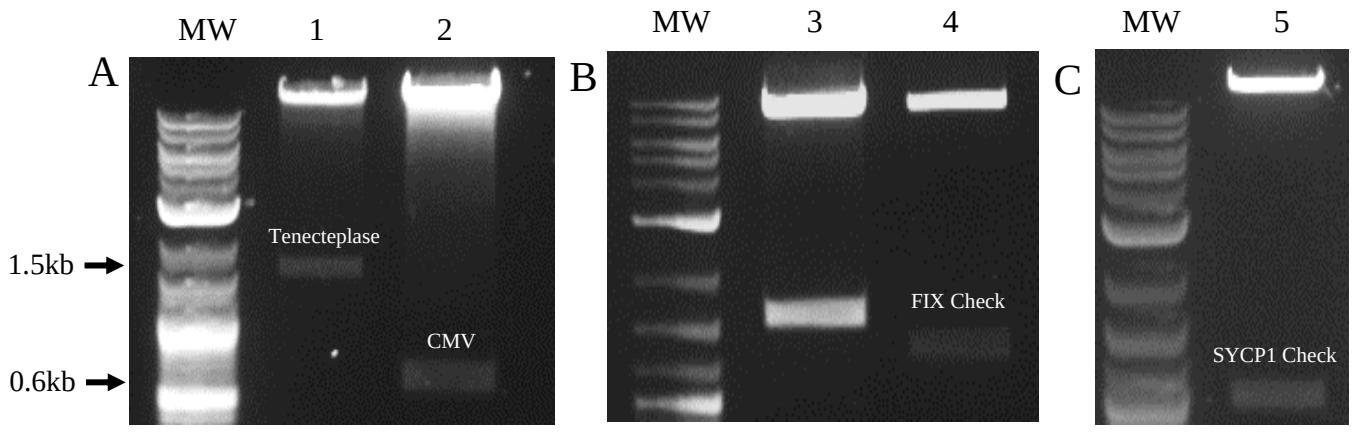


Figure 16: Digestions confirming the removal of CMV and Tenecteplase and the successful insertion of FIX and Sycp1 in the OVA-HA-Syc-FIX gene drive after cloning. Lanes are labeled at the top of the figure and lanes containing ladder are labeled as MW (NEB 1kb plus).(A) Lane 1 is the gene drive containing the OVA homology arms, lane 2 is the same gene drive digested at a later stage for CMV, (B) lane 3 is a control digest in order to differentiate the size of FIX in the next lane, lane 4 is a check, (C) lane 5 is a check digest of Sycp1 in the same gene drive. Samples are run on a 1% agarose gel stained with ethidium bromide.

Once the selected colonies had undergone plasmid extraction, successful insertion of FIX was confirmed in Figure 16B at 1.4 kb using enzymes *Mlu*I and *Sal*I. This was similarly done for Sycp1 in Figure 16C at 0.98 kb using enzymes *Spe*I and *Xho*I. These digests demonstrate that the OVA derived gene drive contained the necessary coding sequence and meiosis-specific promoter after modification. Sequencing was performed (Figure 32-33, supplementary) which further confirmed the desired changes.

Next it was necessary to insert the ssOVM gRNA cassette into the OVM-Sycp1-FIX gene drive. In Figure 17A below, successful removal of the OVA-OVM-OVT gRNA cassette is demonstrated at 1.3 kb (lane 1) as well as the removal of the new ssOVM cassette from the pUC19 cloning plasmid at 0.9 kb (lane 2).

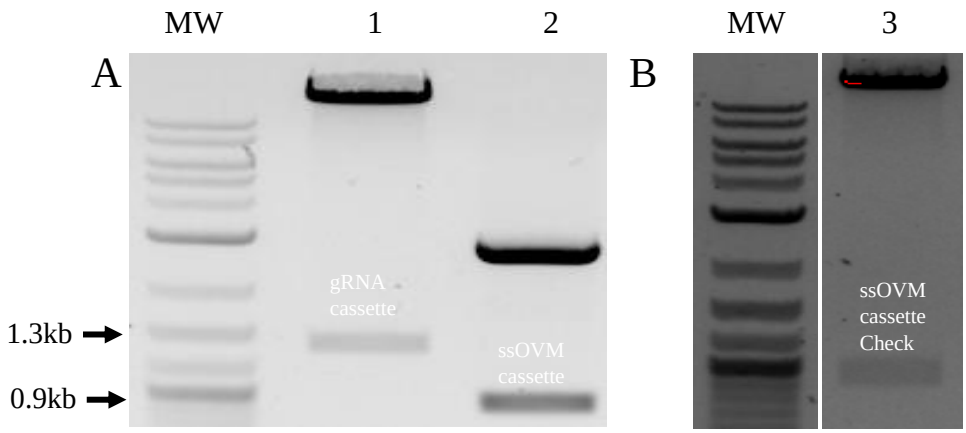


Figure 17: Digests of the OVM-Syc-FIX gene drive to confirm the removal of the old cassette and insertion of the new ssOVM gRNA cassette. Lanes are labeled at the top of the figure and MW represents lanes with ladder (NEB 1kb plus). Lane 1 is a digest of the OVM-Syc-FIX gene drive for the old gRNA cassette, lane 2 is a digest of the new ssOVM cassette out of the pUC 19 cloning plasmid (A), (B) lane 3 shows the ssOVM fragment digested out of the OVM-Syc-FIX-ss gene drive.

The successful insertion of the new ssOVM cassette is indicated in Figure 17B lane 3 at 0.9kb, which is the ssOVM gRNA cassette digested out of the gene drive. All check digests were done using *Not*I and *Bste*II. In addition to this, sequencing was done to confirm the presence and integrity of the new cassette (Figure 37, supplementary).

Following modification of the cassette, it was necessary to insert the puromycin and neomycin resistance sequences, which are important for antibiotic selection of cells modified with these gene drives, into the OVA-HA-Syc-FIX and OVM-Syc-FX-ss constructs. While this step renders the OVA derived gene drive complete, it was necessary to insert the neomycin resistance sequence into the OVM-Sycp1-FIX-ss at this particular step as the 3`-homology arm contains the *Avr*II restriction site, thus, all cloning involving this site needed to be completed before the homology arms could be inserted. Previously, zeocin resistance was included in the gene drive but the mechanism of the antibiotic causes DNA damage in cells, therefore, it had to be replaced with neomycin or puromycin resistance genes. In addition to this, two separate resistances were chosen to make selection easier of the two different gene drive constructs in separate cell lines. Figure 18 (lane 3-6) shows the successful PCR amplification of the fragments necessary to assemble the full neomycin resistance fragment (lane 1) and full puromycin resistance fragment (lane 2), through overlap extension PCR.

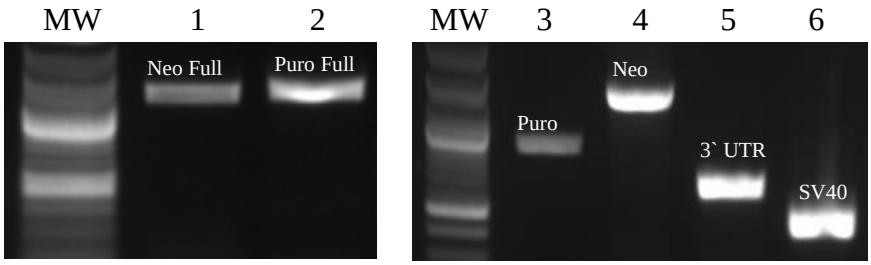


Figure 18: Successful amplification of the products needed for NEBuilder ligation of the antibiotic resistance genes, neomycin and puromycin. Lane 1 is the overlap extension PCR product of the full neomycin resistance gene and lane 2 is the same product for puromycin resistance. Lane 3-6 are the fragments used in overlap extension PCR to assemble the full-length neomycin and puromycin resistance genes in lanes 1 and 2. Lane 3 is the individual puromycin fragment, lane 4 is the individual neomycin fragment, lane 5 is the individual 3` UTR fragment and lane 6 is the individual SV40 fragment.

Removal of the zeocin resistance gene is demonstrated in Figure 19A at 1.6 kb for both the OVM-Syc-FIX-ss (lane 1) and OVA-HA-Syc-FIX (lane 2) gene drives. This step was performed using the *Avr*II and *Not*I enzymes. After ligation through the NEBuilder protocol, the gene drives underwent restriction mapping in order to confirm the successful insertion of neomycin and puromycin resistances. This had to be done with enzymes different from those used in cloning as the *Avr*II and *Not*I sites were not conserved in the NEBuilder primers.

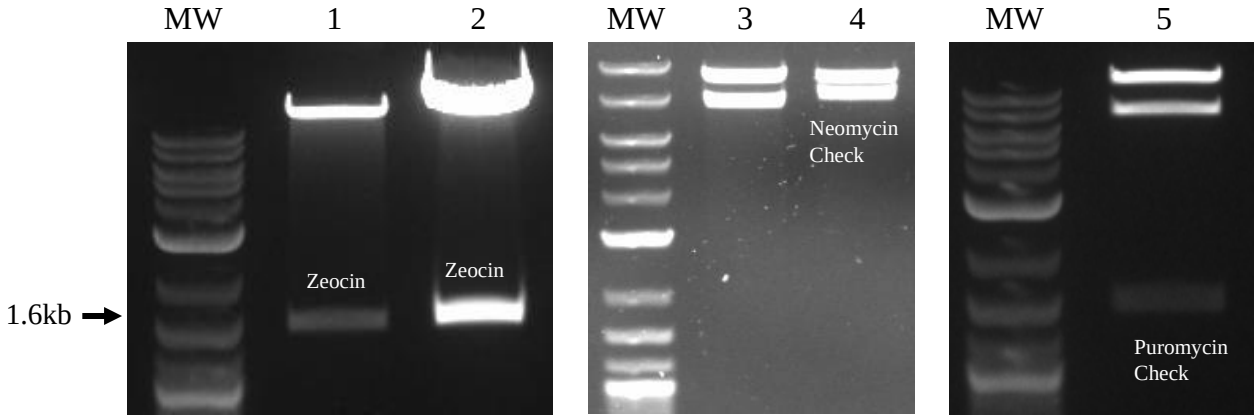


Figure 19: Restriction digests for both OVA-HA-Syc-FIX-Puro and OVM-Syc-FIX-ss-Neo gene drives to confirm the removal of zeocin resistance and the insertion of puromycin and neomycin resistance, respectively. Lanes are labeled at the top of the figure and MW represents lanes containing ladder (NEB 1kb plus). Lane 1 is the OVM gene drive digested for zeocin resistance, lane 2 is the OVA gene drive digested for zeocin resistance, lane 3 is a gene drive containing zeocin resistance digested for *Bste*II and *Xho*I. Similarly in lane 4,

however, this gene drive contains the newly inserted neomycin resistance fragment and lane 5 contains the newly inserted puromycin resistance fragment.

Successful insertion of the neomycin fragment into the OVM-Syc-FIX-ss-Neo gene drive is demonstrated in Figure 19, lane 4. Lane 3 demonstrates what the same digest would look like had neomycin resistance not been inserted and zeocin resistance remained. Successful insertion of puromycin resistance into the OVA-FIX-Sycp1-HA gene drive is also demonstrated in lane 5. Lane 3 can be used for reference of how the gene drive would appear if the puromycin fragment was not taken up and the gene drive still contained zeocin resistance. Sequencing results for both puromycin and neomycin resistance genes within the gene drives can be found in the supplementary section (Figure 34 and 36, respectively).

Lastly, the OVM homology arms needed to be cloned onto the OVM-Syc-FIX-ssNeo gene drive to render it complete and ready for transfection. Figure 20, lanes 1 and 2, show the successful digestion of the OVM 5` and OVM 3` HA out of the pUC19 cloning plasmid and lane 3 and 4 demonstrates the successful digestion of the gene drive for the 5` and 3` homology arms. Gene drives appear linearized as the product between the homology arm restriction sites is less than 20 bps and, thus the digested gene drive product was PCR purified and run on a gel purely for record.

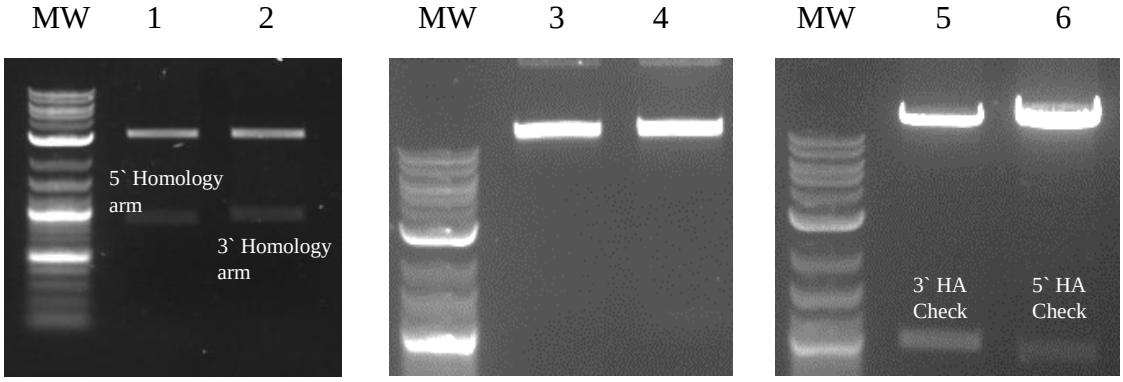


Figure 20: Digestion of the OVM gene drive in order to insert the 5` and 3` homology arms and confirmation thereof. Lanes 1 and 2 show the 5` and 3` homology arms digested out of the pUC19 cloning plasmid, respectively. Lane 3 shows the gene drive digested for the 5`-homology arm and lane 4 shows the gene drive digested for the 3`-homology arm. Lanes 5 and 6 demonstrate a check digestion of the 3` and 5` homology arms, respectively, in the same gene drive, after it had undergone ligation.

Successful insertion of the OVM 3` and OVM 5` homology arms is demonstrated in Figure 20, lane 5 and 6, respectively. This concludes the final cloning step needed to use the gene

drives in transfection. Gene drives were then cleared of endotoxins and concentrated to greater than 1 µg/µl with the Promega Pure Yield Plasmid Midiprep kit. After being transfected in their circular form, gene drives were linearized with *Pac*I and transfected again, following the same protocol.

3.2 Microscopy and ICC staining to identifying PGCs in culture

3.2.1 Germinal crescent PGCs successfully identified in culture by EMA-1 staining

Establishing a culture of PGCs was crucial to this project as they would be the target of transfection as well as the main means of assessing repopulation in ablated embryo gonads. The first attempt at PGC extraction was done on the germinal crescent as this method allowed for the least incubation time for eggs and, thus the quickest wait time to obtain PGCs from embryos. This was done by excising the germinal crescent region from stage 5 embryos and placing it in a culture dish. Figure 21A and B represent staining of germinal crescent PGCs with EMA-1 and Rabbit Anti-Mouse conjugated FITC, after 3 weeks of culture.

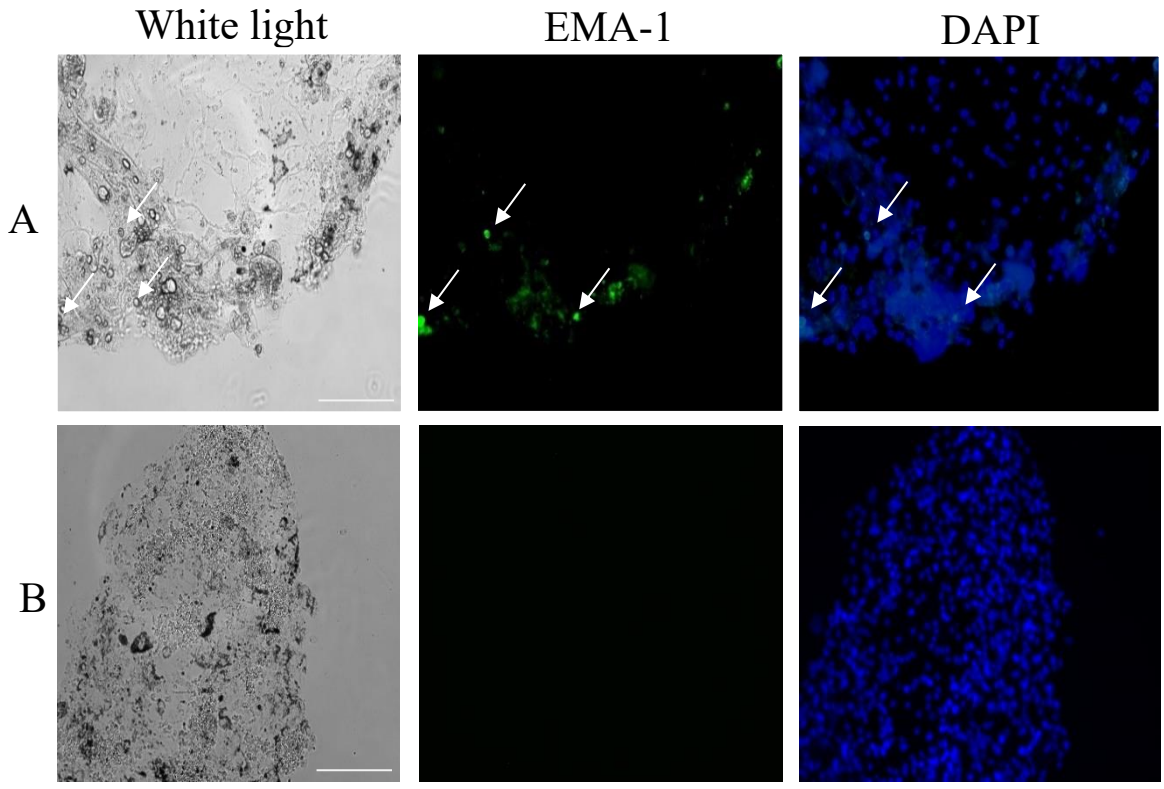


Figure 21: Staining of germinal crescent PGCs to confirm their presence in culture. Lane A (experiment) and B (negative control) are mixed-sex PGCs extracted from the germinal crescent of HH 5 embryos stained with EMA-1 and Rabbit Anti-Mouse FITC as well as DAPI. Scale bar = 100 *µm.* Arrows indicate PGCs.

Staining with anti EMA-1, a stage specific surface antigen marker, revealed that very few PGCs were obtained from the germinal crescent, even after 3 weeks of culture (21A and B). In addition to this, germinal crescent staining and white light observation revealed persistent tissue in culture after 3 weeks. There was also an issue to obtain sufficient material for sexing (data not shown). As sexing was unsuccessful due to the lack of tissue, the samples were combined to make a mixed culture of male and female PGCs to assess how they would perform in culture after this extraction method. A successful sexing gel can be seen in Figure 38 (Supplementary) for blood PGC extraction.

3.2.2 Blood PGCs successfully identified in culture with SSEA-1 staining and morphological characteristics

To improve PGC yield by 3 weeks of culture, which proved to be a limiting factor for germinal crescent extraction, PGCs were re-extracted from the blood of HH 14-17 embryos and placed on a CEF feeder layer (successful establishment of the CEF primary culture used for the feeder layer can be seen in Figure 39, supplementary). The morphology of PGCs was monitored over 3 weeks and the presence of colonies was used to indicate the presence of PGCs in combination with staining. Staining with Anti SSEA-1 was performed in order to confirm that the suspected ‘colonies` were in fact PGCs (Figure 22).

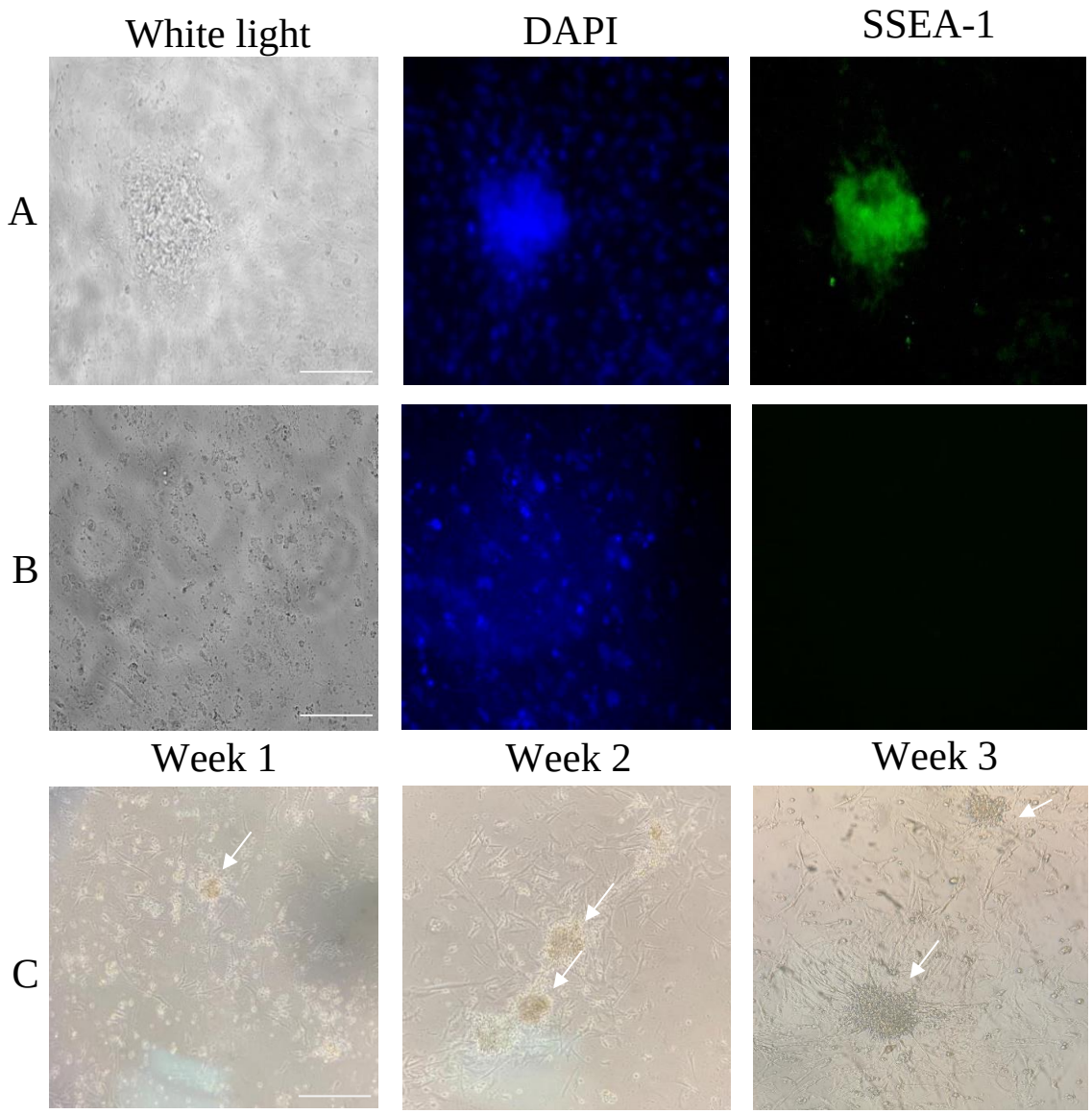


Figure 22: Blood PGCs on a CEF feeder layer stained with DAPI and SSEA-1 as well as images monitoring morphology and colony formation over 3 weeks of culture. Lane A is the experiment, B is the negative control using CEFs and C is the 3-week progression of PGCs in terms of morphology. Scale bar = 100 μm for row A and B, and 120 μm for C. Arrows indicate PGC clusters and individual PGCs.

Figure 22A demonstrates the classic colony forming characteristic of PGCs after 3 weeks of culture on a CEF feeder layer. These were also confirmed to be PGCs with SSEA-1 staining, a stage specific PGC marker. In addition to this, images were captured under the light microscope at 1-week intervals until PGCs formed noticeable colonies. Figure 22C indicates single PGCs or clusters of a few cells at 1 week of culture. Week 2 of culture, PGCs are forming larger clusters and at week 3 of culture, PGCs have formed the largest colonies.

3.3 IHC analysis reveals the best concentration of busulfan for ablation of native PGCs and viability conservation

Once a line of PGCs had been established, gonads had to be prepared for repopulation with PGCs containing the red fluorescent protein expressing tracer, mRFP-Tol2. The sterilisation and, thus removal of native PGCs was accomplished by injecting embryos with busulfan. In order to accomplish sterilisation most efficiently, while still conserving the survival of embryos, the optimal concentration of busulfan needed to be determined. For this, embryos were injected at stage X of development with 65 μg, 75 μg and 85 μg busulfan in 50 μl sesame oil and DMSO mixture. Embryos were then harvested at day 7.5 and the surrounding tissue was removed leaving the back tissue and gonads attached to kidneys. Samples were then preserved in cryomatrix media and sectioned and stained with SSEA-1 and DAPI (Figure 23).

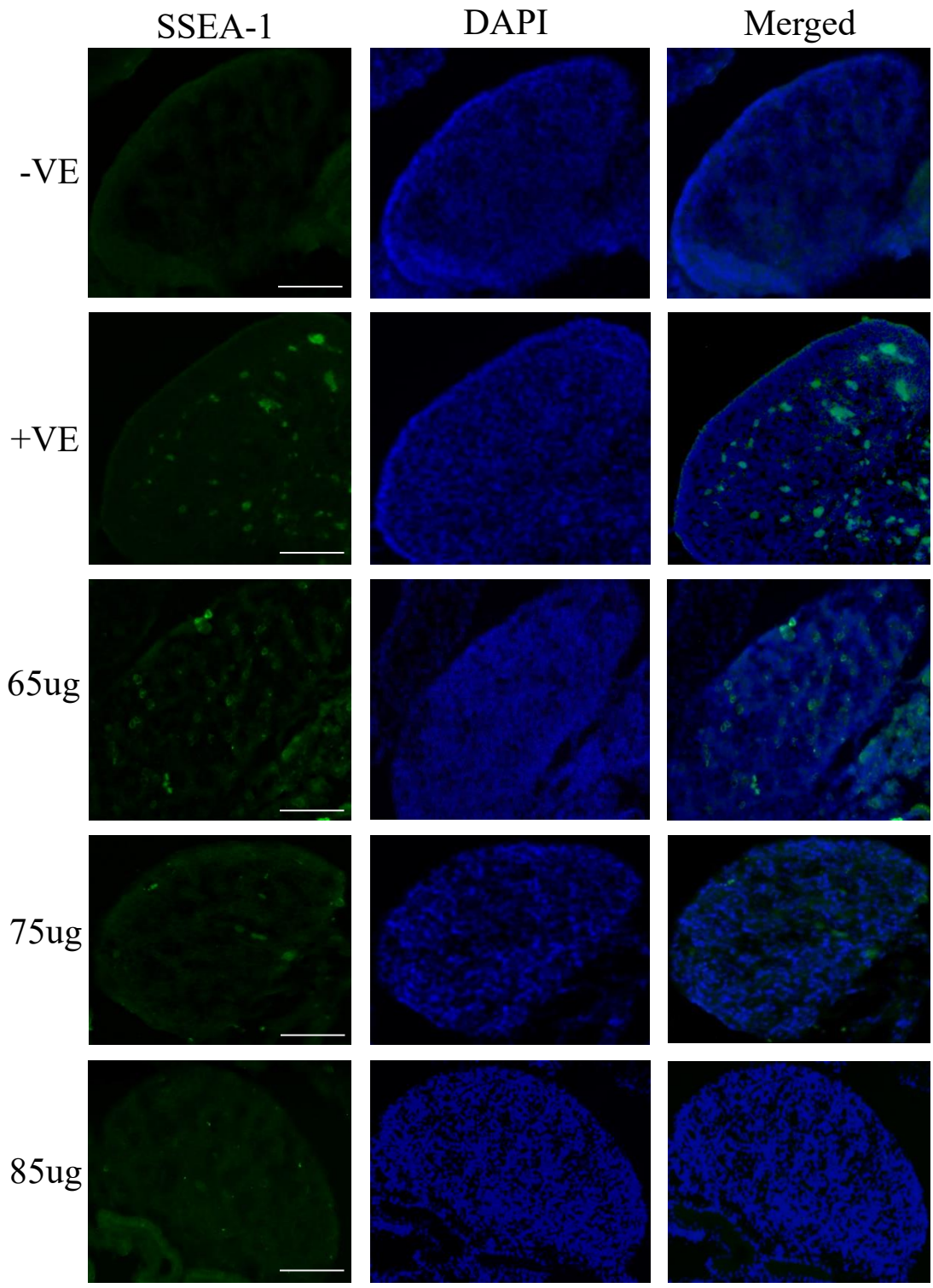
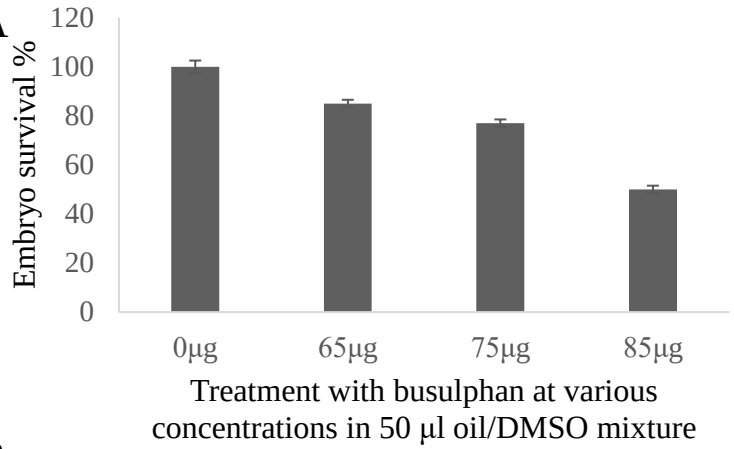


Figure 23: Right gonads of male embryos treated with various busulfan concentrations, extracted at day 7.5 (HH stage 31-33). Gonads were sectioned (10 μm) and stained with Anti-SSEA-1 and DAPI. Each row represents a different concentration of busulfan, and the first two lanes represent a no primary antibody control (-ve) and an untreated control, respectively. N=500 and the scale bar = 135 μm. In each image the outline of the gonad is visible and green flourescent cells withing this region are PGCs.

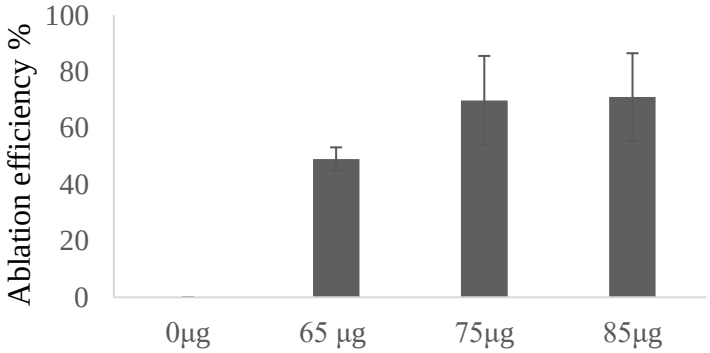
Visually, it appears that 85 μg had the best effect on removing native PGCs, however, viability at this concentration needed to be determined against actual ablation efficiency which involved counting the PGCs across sections and samples. In order to do this, stained cells were counted across 20 sections (10 sections for the left and right gonad separately) for 5 samples in each concentration group. Figure 24A demonstrates viability after busulfan treatment and 24B shows PGC ablation as determined by counting the number of PGCs left in the gonads after busulfan treatment. Only embryos that had developed to at least stage 13, were counted as viable for graph A and embryos considered to be affected by busulfan were any that developed to stage 10 and died. Eggs with no development were considered natural duds and were counted towards natural death in graph D. Figure 24C indicates the number of PGCs left in the gonads after ablation compared to non-ablated embryos. Finally figure 24D represents the viability after treatment corrected for the natural viability. All averages were computed on excel with the standard deviation (SD±) representing standard error. A one-way ANOVA and Tukey's Post Hoc test was then performed to determine the significance of the concentrations and their effect on ablation and PGC number and a Chi-square test was done to determine the significance of viability at those concentrations against the controls. For this, $p<0.01$** and $p<0.05$* was considered significant.

A



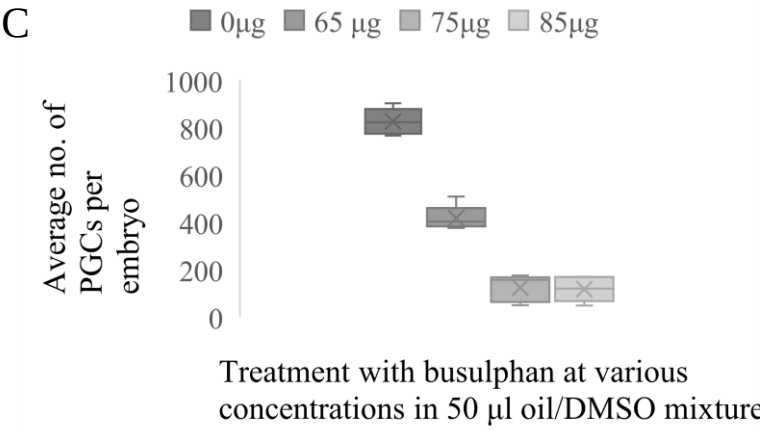
Treatment with busulphan at various concentrations in 50 µl oil/DMSO mixture

B



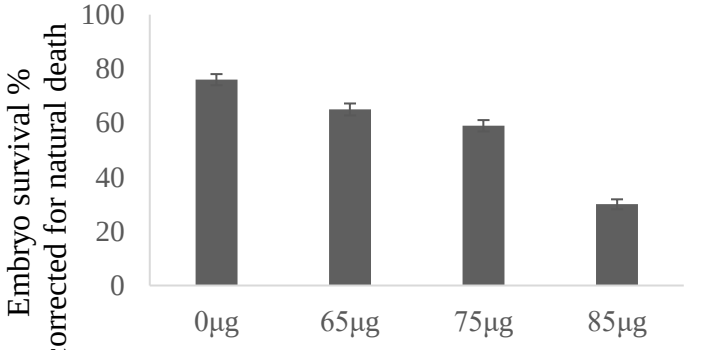
Treatment with busulphan at various concentrations in 50 µl oil/DMSO mixture

C



Treatment with busulphan at various concentrations in 50 µl oil/DMSO mixture

D



Treatment with busulphan at concentrations in 50 µl oil/DMSO mixture

Figure 24: IHC analysis on the ablation efficiency of busulfan at various concentrations and assessment of viability. Graph A indicates the survivability of embryos at varying concentrations and graph B indicates the ablation efficiency of various concentrations of busulfan treatment. The average number of PGCs in each gonad, either treated or untreated, is noted in graph C. Graph D demonstrates the viability of embryos after busulfan treatment corrected for natural duds. For ablation efficiency and the number of PGCs, P** for all concentrations except 65 µg/50µl vs 0 µg/50µl. P** for 75 µg/50µl vs 85 µg/50µl. P** for the number of PGCs in an embryo for 65 µg vs 0 µg. P was insignificant for viability. N= 400 for ablation and average PGC assessment and N=68 for viability assessment.

Figure 24A and B indicate that the concentration that ablated the most PGCs while having the best viability was 75 µg/50µl busulfan in DMSO/Sesame oil (SD± 2). The 85 µg/50µl (SD±2) concentration only had a 1 % higher ablation efficiency than 75 µg however it had a survival rate of 50 % (SD± 1.5) which when corrected for natural death fell to 30 % (SD± 2), which was 27 % lower than 75 µg (SD± 2). The 65 µg/50µl (SD±1.7) concentration had an ablation efficiency of 50 % with a survival rate of 85 % (SD± 1.5) which fell to 65 % (SD± 2) after correction for natural death. The control was considered to have 0 % ablation efficiency and a survival rate of 100 %. Corrected viability is available in 24D at 76 %. This ablation efficiency was further confirmed by the number of PGCs left in the embryos in figure 24C. 65 µg of busulfan had a minimum of 376 and maximum of 509 (mean 411, SD±43). 75 µg had a minimum of 50 and maximum of 176 (mean 123, SD±56) and 85 µg had a minimum of 48 and maximum of 170 PGCs (mean 118, SD±52) left in the gonads after ablation. Finally, the control had a minimum of 767 and a maximum of 904 (mean 823, SD±33).

3.4 Optimal concentrations for the selection of PGCs with neomycin and puromycin determined using kill curve plots

Once the optimal sterilisation concentration of busulfan was decided on, the best selection concentration for PGCs with neomycin and puromycin had to be determined. In order to do this, PGCs were seeded on CEFs and selected for 5 or 7 days. In order to account for the CEF feeder layer, the same kill curve was set up for CEFS alone. Media was replaced every two days, and supplemented with fresh antibiotic, and cells were lifted and counted at the end of the selection period. The CEF feeder layer was then subtracted from the PGCs containing the CEF feeder layer and the average of triplicate wells was computed in excel. Standard error was calculated as standard deviation (SD± 2). Averages for the number of cells left after 7 days for puromycin or 5 days for neomycin treatment were plotted in kill curves below (Figure 25 and 26).

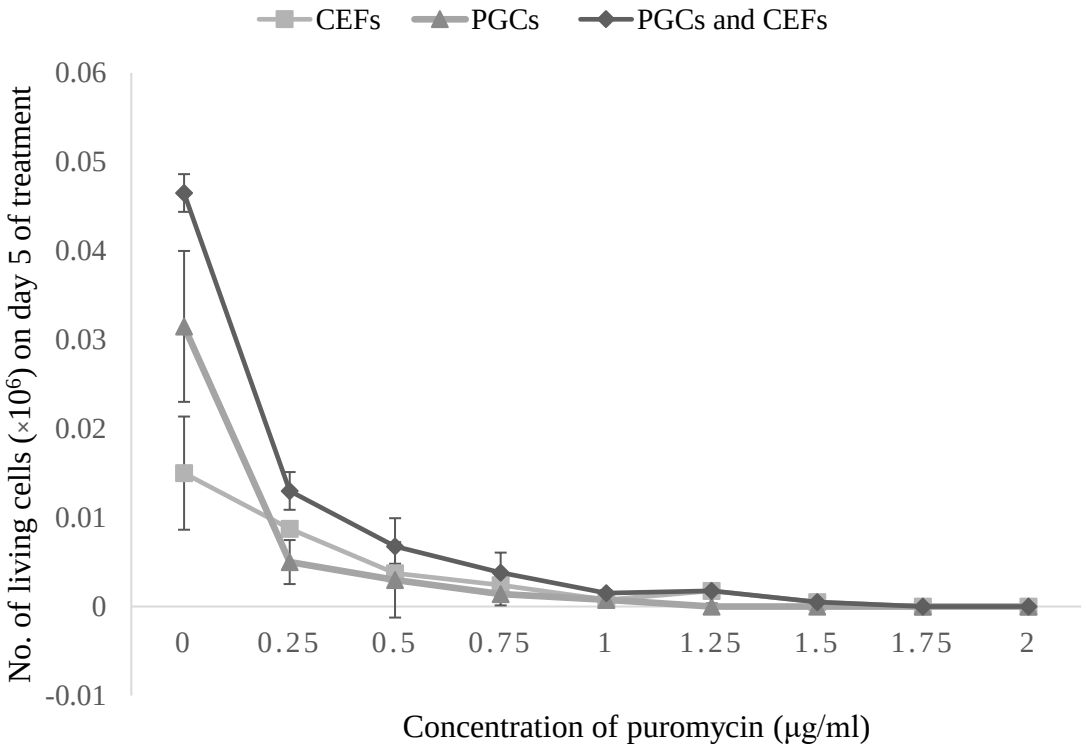


Figure 25: Kill curve demonstrating the effect of puromycin on PGC and CEF viability after 5 days of treatment.**P<0.01 for 1.5 μg/ml against 0 μg and all other concentrations.

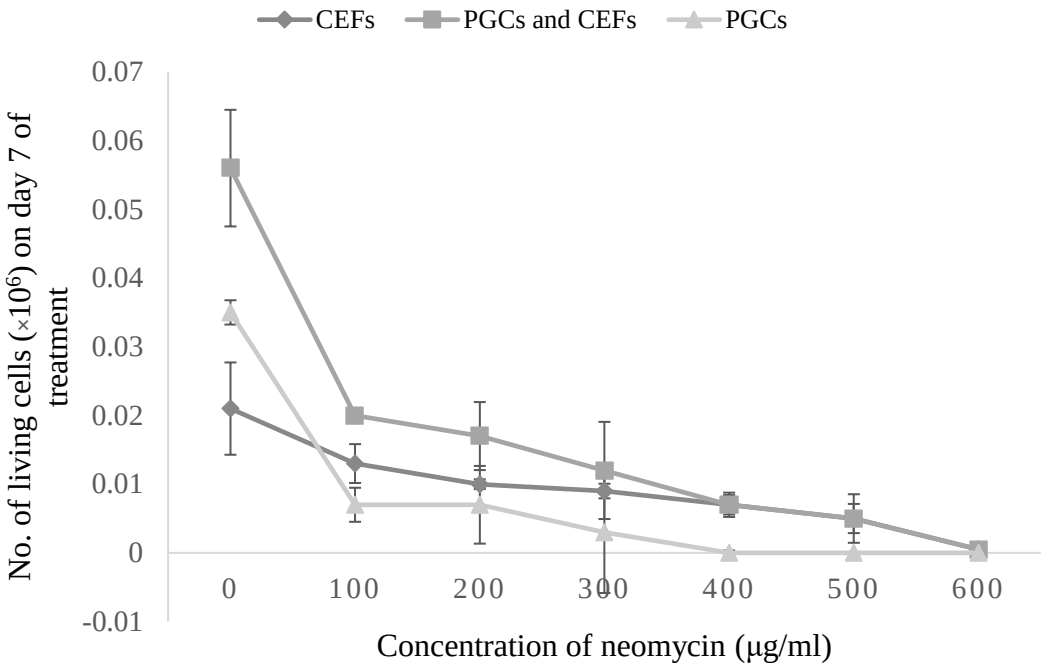


Figure 26: Kill curve demonstrating the effect of neomycin treatment on the viability of cells after 7 days of treatment. P<0.05* for 400 µg/ml against 0 µg and all other concentrations.

Figures 25 and 26 above indicate that the concentration that kills off PGCs after 5 and 7 days is 1.5 µg/ml for puromycin and 400 µg/ml for neomycin, respectively, which means that the these are the concentrations required to successfully select for PGCs with these antibiotics. The CEFs required a higher concentration for both puromycin (1.75 µg/ml) and neomycin (600 µg/ml), respectively. The lethal dose was not present for either antibiotic but the higher concentrations of 1.75 µg/ml and 500 µg/ml for puromycin and neomycin, respectively, were noted to kill cells after approximately 2-3 days.

3.5 PCR analysis reveals the successful transfection of PGCs and the absence of genome modification at the OVA and OVM loci

Once the best selection concentration was determined for each form of resistance across the two gene drives, the PGCs had to be edited with the OVA and OVM targeting constructs. In order for this to happen, two methods of transfection were attempted, namely electroporation and lipofection with lipofectamine 3000 (Invitrogen L3000015). Firstly, PGCs had to be transfected with the mRFP-Tol2 construct (Figure 27). This provided the opportunity to test electroporation conditions with a smaller construct and find the optimal conditions for transfection with the larger gene drive constructs. Electroporation gradients of 240 v, 10 ms pulse length, 1s pulse interval and varied pulses : 1, 2 and 3, were tested.

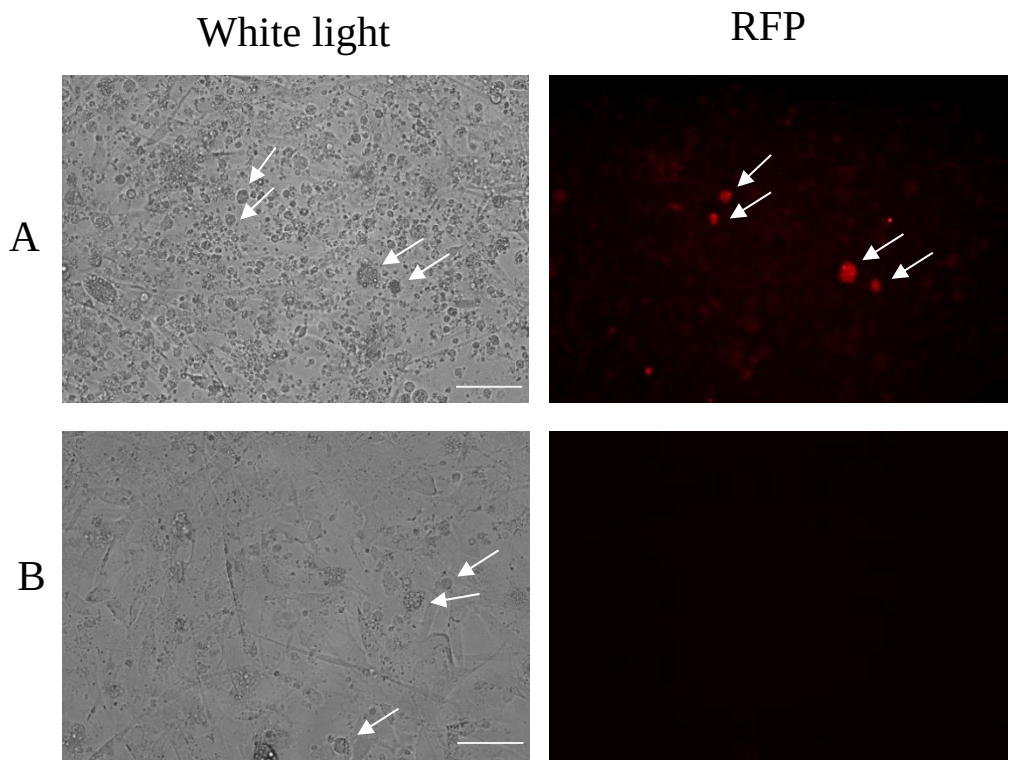


Figure 27: mRFP-Tol2 transfected PGCs observed under a fluorescence microscope. A demonstrates cells electroporated at 2 pulses, B is a negative transfection control. Scale bar = 125 μm and arrows indicate PGCs.

It was demonstrated that 2 pulses, with the abovementioned conditions, appeared to be the best electroporation condition as this had the highest viability of cells and the only fluorescence. Electroporation at 3 pulses appeared to be lethal to cells as all cells were dead (not shown) and electroporation at 1 pulse had little to no fluorescence (not shown). These results also confirmed the presence of the mRFP-Tol2 tracer construct in PGCs and that it was functional.

Knowing the best electroporation conditions for PGCs, gene drives were then electroporated and lipofected into the cells along with the tracer mRFP-Tol2 construct and the eSpCas9-T2A expression construct. Lipofection with lipofectamine 3000 had to be done at a volume of 0.15 and 0.75 μl as recommend in the manufacturer's instructions. After transfection, cells were placed on a fresh feeder layer of inhibited CEFs. Transfected cells were then selected for with puromycin and neomycin for 2 days and allowed to recover for 4 days before viewing under the microscope and 5 days before genomic DNA extraction. To view, cells were visualised under the fluorescent microscope and morphology, viability and fluorescence were noted to determine the best means of transfection (Figure 28).

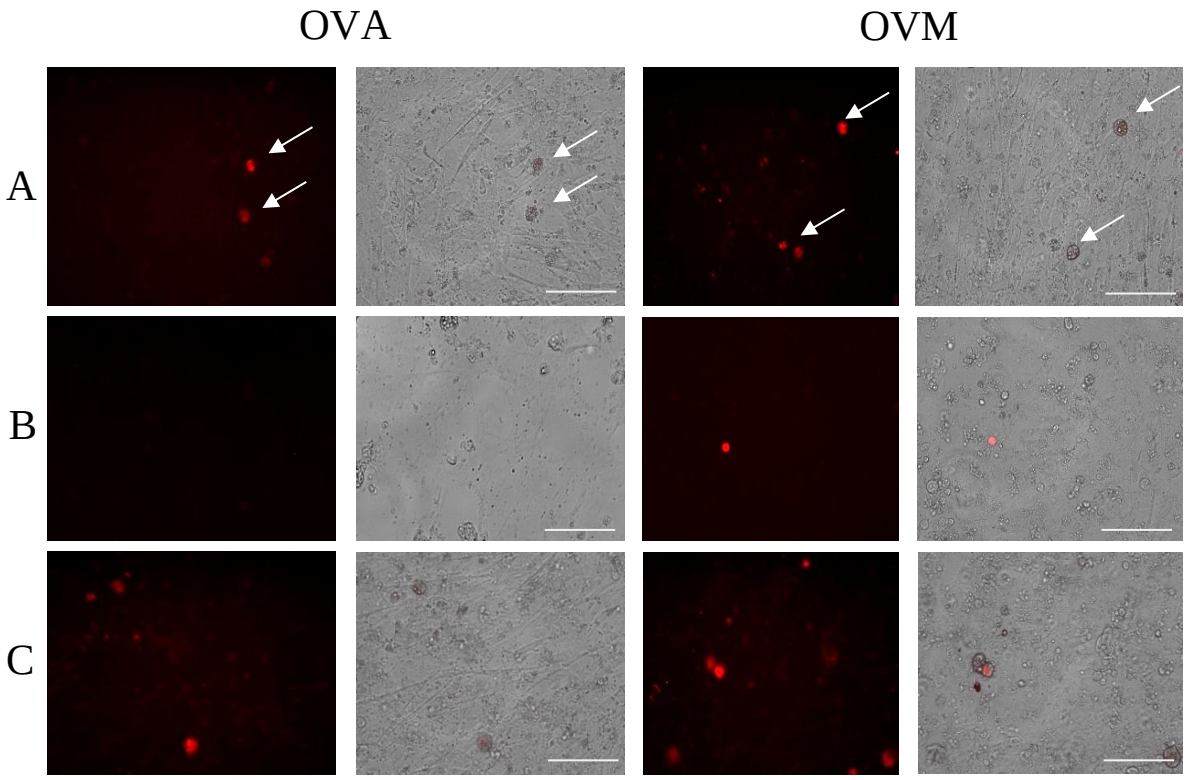


Figure 28: mRFP-Tol2 and gene drive edited PGCs on a CEF feeder layer. A is OVA-HA-Syc-FIX-Puro and OVM-HA-Syc-FIX-ss-Neo transfected with the higher concentration of lipofectamine, B is the lower concentration of lipofectamine, and C are cells transfected with electroporation. N=18 and the scale bar = 100 μm. PGCs are indicated by arrows and are shown as an example in the first row only.

In the case of transfection with multiple constructs, electroporation appeared to have the most cells transfected with the mRFP-Tol2 construct and thus assumed the other constructs. The higher concentration of lipofectamine 3000 also appeared to yield the best fluorescence, when lipofection was compared alone. Fluorescent cell clusters were counted across 3 microscope fields of view and the average was computed. The average across the electroporated group was the highest at 6 (SD±1) and 10 (SD±2.26) clusters per field of view for OVA and OVM, respectively. The second best was the highest concentration of lipofectamine at 6 (SD±2.6) and 7 (SD±2.6) (for OVA and OVM, respectively), and the least efficient was the lowest concentration of lipofectamine at 3 (SD±1.7) and 2 (SD±1) clusters (for OVA and OVM, respectively) per field of view. $P<0.01^{**}$ for all samples except the low concentration lipofectamine samples.

Next cells for each transfection method were harvested and processed with a genomic DNA extraction kit. Primers designed by a previous MSc student were used to screen for HDR integration, NHEJ integration, gene drive presence in the cell, and the presence of the eSpCas9 expression plasmid (Figure 29).

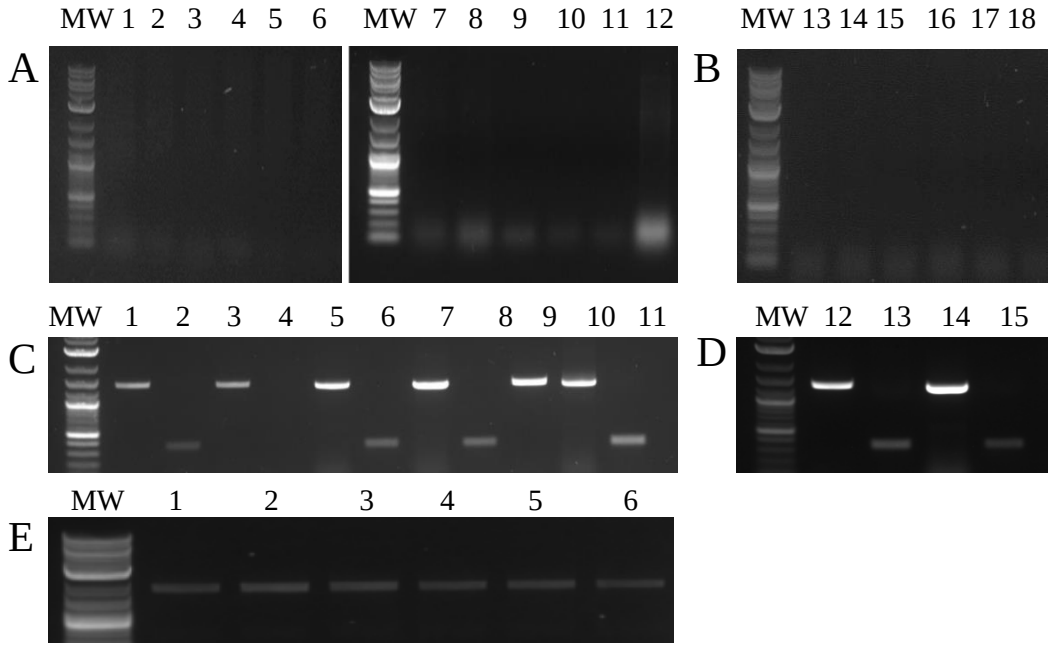


Figure 29: PCR analysis of circular transfected gene drives for genome integration into PGCs. A and B show the assessment of HDR intgeration of the OVA and OVM GDs at both the 3` and 5` homology arm, and NHEJ into the genome. (A) Lane 1 and 2 are the 3` and 5` homology arm of the OVM electroporated GD, respectively. Lane 3 and 4 are the 3` and 5` homology arm of the OVM GD lipofected at the higher concentration, respectively. Lane 5 and 6 are the 3` and 5` homology arm of the OVM GD lipofected at the lower concentration, respectively. Lane 7 and 8 are the 3` and 5` homology arm of the OVA GD electroporated, respectively. Lane 9 and 10 are the 3` and 5` homology arm of the OVA GD lipofected at the higher concentration, respectively. Lane 11 and 12 are the 3` and 5` homology arm of the OVA GD lipofected at the lower concentration, respectively. (B) Lane 13, 14 and 15 are NHEJ assessment for the OVM GD for electroporation, lipofection (high) and lipofection (low), respectively. Lane 16,17 and 18 are NHEJ assessment for the electroporated and lipofected (high and low), respectively. C and D show the assessment of gene drive transfection into PGCs. Lane 1 is the OVM GD amplified for neomycin resistance (electroporation), lane 2 is the OVM GD amplified for the backbone (electroporation), 3 is the OVM GD amplified for neomycin resistance (lipofection high), 4 is the OVM GD amplified for the backbone (lipofection high), 5 is the OVA GD amplified for puromycin

resistance (electroporation), 6 is the OVA GD amplified for the backbone (electroporation), 7 is the OVA GD amplified for puromycin resistance (lipofection high) 8 is the OVA GD amplified for the backbone (lipofection high). Lane 12 and 13 are PCR on the neomycin fragment and backbone of the lipofected (low concentration) OVM GD and lane 14 and 15 are PCR on the puromycin fragment and backbone of the lipofected (low concentration) OVA GD. (E) Circular gene drive transfection and screening for the eSpCas9 plasmid of the OVA and OVM electroporated samples (lane 1 and 2 respectively.), the high concentration lipofected OVA and OVM samples (lane 3 and 4 respectively.) and the low concentration lipofected OVA and OVM samples (lane 5 and 6, respectively).

Figure 29A demonstrates failed PCR for the amplification of the gene drive and genome region that would be present if HDR had occurred in OVA and OVM gene drives. Figure 29B shows failed PCR for the amplification of gene drive region if NHEJ had occurred in any of the samples. Figures 29C and D show the presence of both gene drives in the separate cell lines for each method of transfection. Figure 29E demonstrates the presence of the eSpCas9 construct in the cells that were transfected.

The two OVA and OVM derived gene drives were then digested at the *PacI* site after the 3`-homology arm. Another attempt at transfection was made with the linearised gene drives using only electroporation, however, HDR screening showed that there was no insertion at the target locus (Figure 30A lanes 1-4) and NHEJ screening demonstrated the same issue as with the circularised plasmids (Figure 30A Lanes 5-6). In this case the gene drives (Figure 30B lanes 1-4) and the eSpCas9 plasmid (Figure 30C Lanes 1-2) were still present, similar to the above scenario. Figure 30A, Lanes 5-6, demonstrate that even after gene drive linearisation, there appears to be no integration into the genome or cutting at the target site, similarly to the circular gene drive in Figure 29.

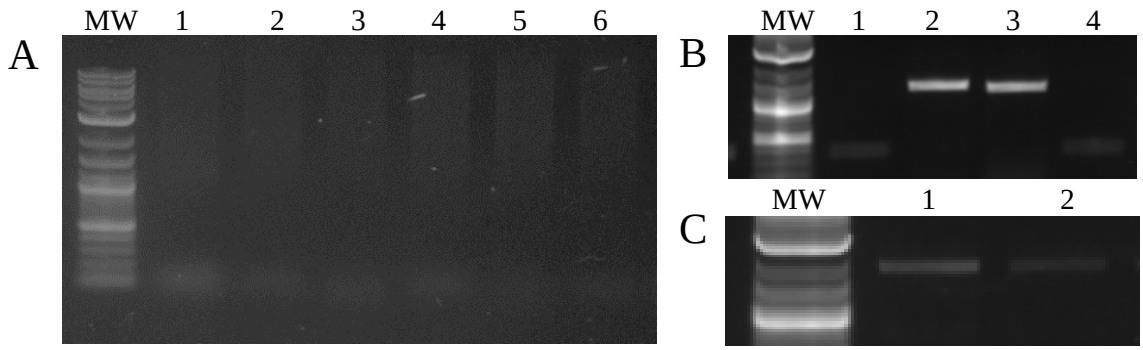


Figure 30: PCR analysis of linearized transfected gene drives for genome integration in PGCs. A: Assessment of NHEJ and HDR for linearised GDS. Lane 1 and 2 are HDR assessment for the OVM GD (5` and 3`) and lanes 3 and 4 are HDR assessment for the OVA GD (5` and 3`). Lanes 5 and 6 are NHEJ assessment for the OVM and OVA GDs, respectively. B: Assessment of integration of linear GD electroporated into cells. Lane 1 is the backbone of the OVA gene drive, lane 2 is the puromycin fragment, lane 3 is the neomycin fragment in the OVM gene drive and lane 4 is its backbone. C: Linear gene drive transfection screened for the eSpCas9 vector. Lane 1 is the PCR on the OVA gene drive containing gDNA and lane 2 is PCR on the OVM gene drive containing gDNA.

3.6 mRFP-Tol2 tracer analysis reveals that repopulation of embryo gonads with edited PGCs is largely unsuccessful

Repopulation was then attempted with the RFP edited PGCs. Cells were resuspended in Advanced MEM and injected into stage 13-15 embryos with a glass needle attached to a mouth pipette. Cells were injected into the outer vitelline arteries as injection into the dorsal aorta resulted in death. Embryos were then harvested at day 6 (HH stage 27-28) and observed under a fluorescent microscope (Figure 31). Approximately 1000 cells were injected into the embryos.

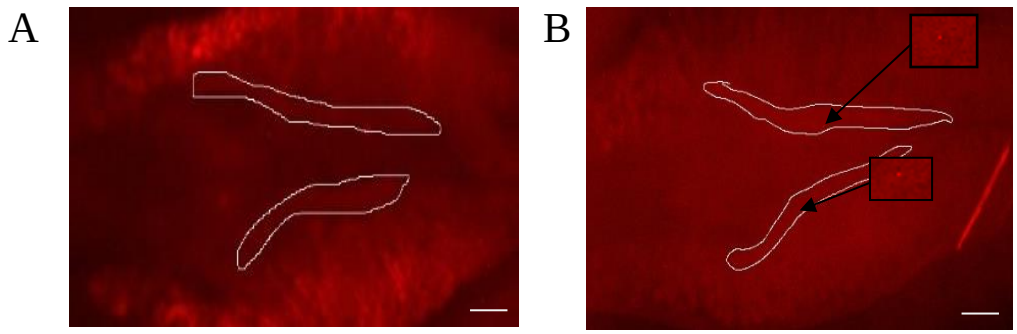


Figure 31: Gonadal repopulation the RFP tracer construct (mRFP-Tol2) of HH 13-15 male chick embryos. A is a control embryo that did not receive the tracer edited PGCs and B is the experiment embryo. N=5. Scale bar = 250 μm

Single fluorescent specks are visible in each gonad of the experiment. These fluorescent cells within the gonads were counted across 5 experimental embryos. The average number of PGCs was then determined to be 1.2 (SD±0.81) PGCs per embryo and the viability was determined to be 23 % (N=21).

4. Discussion

In order to alleviate the high cost that comes with producing therapeutic proteins, an expression system is required that is both cost effective and easy to maintain. For these reasons, the domestic chicken is an ideal candidate. Chickens are capable of producing many eggs (one per day per bird) where the albumin can be used as a vehicle for therapeutic protein expression. While the yolky structure of the zygote has limited avian transgenics in the past, thanks to recent advancements with CRISPR, the genome is now more easily accessible for transgenic production and, thus bioreactor creation. In this project we sought to produce a genetic construct that would in time be responsible for editing the Ovalbumin and Ovomucoid genes in chickens and replacing them with a human FIX expressing gene drive. This would in turn yield more protein than has ever been produced by other systems, such as CHO cells, at a much cheaper cost. In order to accomplish this, a gene drive construct first had to be modified through a combination of cloning and restriction digests, a line of PGCs had to be established as these cells were the most straightforward way to gain access to the chicken's germline, and PGCs had to be transfected with the aforementioned gene drive so that its activity could be assessed.

4.1 The successfully altered gene drive is ready for use in downstream modification experiments

First, the human FIX coding sequence, which was obtained in my honour's year, had to be inserted in place of the Tenecteplase coding sequence in the editing construct. This sequence was already present in the gene drive as it was synthesised with it. Tenecteplase also forms part of a similar project that is being undertaken in Prof. Natalya Nikitina`s lab. Figures 15 and 16B demonstrate the successful insertion of the FIX sequence into the OVM and OVA gene drives through restriction digest. Similarly, the native CMV promoter had to be replaced with a meiosis-specific promoter, Sycp1. CMV was originally included in the construct in order to evaluate the expression and function of Cas9 and thus, the gRNAs, as well as gene drive activity within CEFs, outside of meiosis. This proved useful as the cell line used currently, PGCs, is not as proliferative as CEFs but they could still provide a chicken analogue for acquiring information on gene drive activity *in vitro*.

Having a meiosis-specific promoter, such as Sycp1, is important for controlling the expression of Cas9 to the meiosis stage in PGCs. This will in turn negate any NHEJ repair that would occur outside of meiosis or at least lessen the chances and in turn reduced the possibility of the construct failing to insert into the genome. Figure 16C demonstrates successful insertion of the Sycp1 promoter into the OVA gene drive as well as confirmation

of the previously cloned in promoter in the OVM gene drive (Figure 15C). These steps completed the addition of two important components to the gene drive. With the addition of these, the construct will be able to maintain expression of the therapeutic FIX protein under the Ovalbumin or Ovomucoid promoters and have timed expression of Cas9 so as to avoid any off-target effects.

Next, the ssOVM gRNA cassette had to be cloned into the OVM-Syc-FIX gene drive. As original work on this project assumed that the OVA, OVM and OVT genes would be targeted for transgenesis, the gRNA cassette contained within the gene drive had gRNAs for each of these three loci. The original plan to target OVT had to be abandoned as research indicated that the OVT protein is crucial for life and thus it is suspected to render embryos edited for it inviable, hence why no prior studies have made modifications to it[75]. In the case of this project that planned to only target the OVA and OVM loci separately, this old cassette would have resulted in unnecessary cutting within the genome which could have had unwarranted effects. In the instance of OVT specifically, this could prove lethal. Successful insertion of the new cassette was demonstrated in Figure 17, where the band at 0.9 kb represents the cassette. Further to the listed reasons for modifying the cassette, the new OVM cassette contained additional tRNA sequences as it is necessary for another project in the lab that is focusing on expressing spidersilk *in Ovo*. Expression of the spider silk protein requires more tRNA residues than what is naturally available for protein synthesis in chickens. Using the same cassette across gene drives that are tailored for the expression of different proteins helps standardise the gene drives somewhat, making the various intermediates more manageable.

Next, two separate antibiotic resistance genes needed to be added to the two separate gene drives. This would offer the advantage of being able to select for the OVA-HA-Syc-FIX-Puro and OVM-HA-Syc-FIX-Neo gene drives separately. This would be especially useful in future experiments where the gene drives are both transfected into a single cell line and cells would need to undergo sequential selection with two antibiotics to eliminate any cells that contained only one gene drive. In addition to needing two separate resistances, the original resistance to zeocin had to be removed as the antibiotics mechanism of action works to damage DNA which could cause unnecessary issues in future adult chickens. The successful cloning of neomycin and puromycin resistance genes are demonstrated in Figure 19 (lanes 4 and 5) while the successful PCR of the fragments required to assemble these products that were inserted into the gene drives, is shown in Figure 18.

Lastly, the OVM homology arms had to be appended to the OVM-syc-FIX-Neo gene drive in order to allow for HDR integration into the chicken genome upon transfection into PGCs. Figure 20 (lanes 5 and 6) demonstrates the successful cloning of the homology arms onto the gene drive, as indicated by the lower bands at 0.9 kb for the 5` and 3` homology arms. These homology arms had to be added as the final step, as opposed to the OVA gene drive, as they contained a restriction site also present in the resistance gene. Thus, if the steps were performed the other way around, the homology arm would have been cut and upon ligation could have been excluded from the gene drive together with a sizeable chunk of the construct. In addition to this, the gene drive could have been re-arranged. Similarly with the OVA gene drive, which contained a restriction site used for homology arm cloning in the puromycin resistance fragment. These changes could have negatively impacted gene drive function in the future. Upon successful completion of these gene drives, it is the idea that these constructs, once transfected into cells, will integrate under their relevant loci (OVA and OVM). This would eventually translate into FIX expression under both of these promoters and allow for the constructs to be passed down from generation to generation.

4.2 A line of PGCs is successfully established allowing for editing of the chicken germline

After the constructs were completed, PGCs had to be prepared as they would be the vehicle for transfection. PGCs were initially obtained from the germinal crescent of embryos, however, they were too few and the surrounding tissue persisted in culture for too long (Figure 21A). The cells did not appear to multiply to a sufficient number at the end of three weeks as evident by minimal cells present after staining and, thus a different method of extraction was attempted, that according to literature, provides a better yield. Too few PGCs at this stage is in line with literature as at stage 4-11 only 130 PGCs are predicted to be present[40,41].

PGCs were, thus harvested from the blood of HH 14-17 embryos and plated on a dish containing a CEF feeder layer. While some literature demonstrates that PGCs can be cultured successfully on nothing but the plastic cell culture material, it was found that majority of the other publications demonstrate that PGCs grow better and proliferate faster when placed on a feeder layer[54,90,92,106]. This method was thought to be most ideal especially in the case of limited time. Once extracted from the blood of embryos and plated, PGCs were identified through IHC staining (Figure 22A), after 3 weeks in culture. PGCs reached colony formation within 2-3 weeks (Figure 22C) as is demonstrated in literature[15,52]. It is suspected that this is due to the proliferative response elicited when they come in contact with other cells, namely

the CEFs. As PGCs are so few upon extraction, they would need to reach far greater numbers to proliferate faster and thus placing them on the feeder layer provided the contact stimulation which would otherwise be lacking. The CEF feeder layer also has the benefit of secreting hormones necessary for PGC maintenance, which in combination with added hormones provides the ideal culture environment. PGCs would reach confluency of approximately 5×10^5 cells after three weeks with passaging every week. PGCs also responded well to freshly isolated cells being added to the culture. Cells were maintained for extended periods (up to 2-3 months) before being replenished with fresh cells. While this method provided PGCs at a sufficient number to accomplish the objectives in this project, the number of PGCs that could be reached in culture was still a limiting factor in cases where replicates and repetition was required.

4.3 Busulfan is effective at eliminating native PGCs and preparing embryos for repopulation

While PGCs were kept in culture to reach sufficient numbers for downstream editing and transfection experiments, native PGCs were treated with busulfan for ablation in the gonads of HH 14-17 embryos. There is a general trend shown after staining with SSEA-1, in Figure 23, that PGCs decreased overall from the control to higher concentrations of busulfan treatment. According to these results, 85μg of busulfan had the best effect on PGC depletion, however, this only gave a qualitative result and the viability needed to be compared and the number of stained cells in each treated and untreated sample needed to be calculated in order to confirm the ablation efficiency quantitatively.

Overall, PGC proportion in controls was significantly higher than in most busulfan groups, as demonstrated by Figure 24C. Teratogenic effects were more prevalent with increased busulfan concentrations (Table 12 supplementary), and survival was shown to decrease in a dose dependent manner from 65 to 85μg/50μl (Figure 24A and D). All these findings were in line with literature and with expectations as busulfan is toxic to cells other than PGCs, even though they are most susceptible. Although statistical analysis revealed that there were no significant differences between survival rates and different concentrations of busulfan treatment, most likely due to a small sample size, the survival values were still considered as it has been shown in other papers that viability is affected[105,106]. In Figure 24A we can see that viability is higher at lower concentrations and drops at higher concentrations of busulfan treatment. This is the case for when natural viability is considered (24D) as well as viability only in terms of busulfan treatment (24A). This drop in viability at 85 μg/50μl is not justifiable for the number of PGCs ablated compared to the 75 μg/50μl (Figure 24B), which

has a better survival rate. In addition to this, there was found to be no statistical significance between these two concentrations in terms of ablation efficiency. Viability needs to be conserved as much as possible as external factors contribute further to future hatchability number and for this reason 75 μg/50μl was chosen for future ablation experiments. This was in line with literature where this concentration was determined to be the best compromise between ablation efficiency and viability[52],[105]. In the case of the lowest concentration, 65 μg/50μl, the number of cells remaining was still too high in comparison to the 75 μg/50μl concentration which had a better ablation efficiency (Figure 24B) and significantly fewer PGCs after treatment (Figure 24C).

Even in the case of the maximum and minimum PGCs counted in the samples, treatment with 75 μg/50μl gives a similar maximum and minimum to 85 μg/50μl (Figure 24C) with no statistical significance between the two. Thus, even in the worst-case scenario, the number of PGCs removed between the two experiments is similar. In the case of the 65 μg/50μl, the worst-case scenario presents a possibility of having samples with not many more PGCs removed than is present in an untreated control and for this reason treatment with 65μg of busulfan was disregarded as being sufficient. The number of PGCs counted in this study as well as the viability and ablation efficiency at day 7.5 are in line with other studies[87], however, it must be noted that there are variable results across many papers which is suspected to be due to the antibody used in staining being differentially expressed at different points during PGC development.

4.4 Neomycin and puromycin successfully select for edited PGCs allowing for a means to remove unedited cells from culture

Next, treatment concentrations for both neomycin and puromycin selection of PGCs were determined (Figures 25-26). This was of importance to future experiments that would involve PGCs transfected with the different gene drives across two separate cell lines. The optimal concentration for puromycin was determined to be 1.5 μg/μl and neomycin was determined to be 400 μg/μl, as this is what killed off all cells at the end of the treatment period. The number of PGCs decreased in a dose dependent manner and overall, there were more PGCs in the controls at the end of treatment than in the experiments. Concentrations were considered optimal after 7-5 days of treatment as this meant that these concentrations were not too strong so as to kill cells off rapidly, thus, killing cells that may contain the gene drive, or not too weak where it would still leave cells that did not contain the gene drive. These findings are in line with literature where similar concentrations are used to select for PGCs in culture [90,98,99].

4.5 PGCs are successfully transfected for the OVA and OVM gene drives but do not undergo genome modification

Once the best concentration for selection was known, PGCs were edited with the gene drives as well as the mRFP- Tol2 and eSpCas9 constructs. In the ideal scenario, PGCs would have undergone HDR and incorporated the gene drive under the Ovomucoid and Ovalbumin promoters. A separate cell line was transfected with the mRFP-Tol2 (Figure 27) construct, in order to assess repopulation down the line. Gene drive transfected cells were not used for repopulation assessment as the activity of the construct still needed to be assessed within PGCs. The mRFP-Tol2 construct alone would provide valuable insight into whether the method of injecting PGCs into an embryo with a pulled glass needle and mouth pipette was viable before wasting edited PGCs that are difficult to culture. Figure 27 demonstrates that the mRFP construct successfully passed into PGCs upon electroporation. It was noted that the condition that had the least cell death and most fluorescence was the 2-pulse condition. A single pulse was not efficient at allowing the construct to pass through the membrane and three pulses were too strong and killed PGCs. Figure 28 shows that the best condition for transfection is electroporation, compared to lipofection, as this had the highest average fluorescent cell/cell cluster count per microscope field of view.

Figures 29-30 demonstrate that PGCs transfected with the gene drive never successfully experienced gene drive integration (Figure 29A and 30A lane 1-4) even though they took up the various constructs (Figure 29 C and D and 30B and C). While the gene drive was present in the cells, amplification of the NHEJ product was performed which confirms that the gene was in fact never removed (Figure 29B and 30A lanes 5-6). This is suspected to be due to the gRNAs not promoting cutting at the target site or having reduced efficiency and thus preventing integration of the gene drive. This may be because the cassette is non-functional or that the target site is partly inaccessible to the gRNAs and Cas9 because the chromatin is in heterochromatin form. In addition to this, editing efficiency could naturally be low and repeating the experiment in duplicates or triplicates may be useful, however, this would require more PGCs, which is a major limiting factor in this method.

4.6 Re-population of embryo gonads requires optimisation

The repopulation of gonads was then attempted using the edited tracer PGCs and ablated embryos (Figure 31). These results demonstrate that PGC repopulation is largely unsuccessful as indicated by the mean of 1.2 PGCs per embryo over 5 samples. While this indicates the presence of PGCs it is important to note that the results are inconclusive as counter staining would need to be performed to confirm that the fluorescent cells are in fact

PGCs. This could not be performed due to time constraints of ordering the necessary antibody that would bind to RFP after exposure to formaldehyde.

Low or lack of PGC repopulation could be as a result of the method as a whole being technically challenging or as a result of repetition and practice needed which could not be assessed due to limited PGCs. The first issue with this method is that after ablation there are few embryos to work with due to the low survivability as well as sex ratio. Viability further decreases as embryos are disturbed a second time for sexing and then PGC injection. Embryos must then be grown to at least day 5,5 to visualise repopulation. The viability of embryos has been shown to decrease drastically as incubation time increases after modification and sterilisation[52],[105]. Additionally, the method on its own is incredibly difficult to perform gently as well as consistently penetrate the vessels. Commonly, a microinjector is used which allows for the control of pressure and volume as well as allows for dexterity when performing the process under a microscope. These machines also come with needles; thus, needle size can be controlled which eliminates possible avenues of error. Unfortunately, these pieces of equipment are expensive to purchase and thus accessibility is limited.

4.7 Improvements, future work, and conclusion

Going forward, in order to be able to make any improvements to the work done in this project, a larger number of PGCs is required. In order to improve on PGC culture and obtain what would be necessary to troubleshoot downstream experiments, an alternative method of PGC extraction could be attempted- gonadal extraction. PGCs in the gonads are estimated to be anywhere between 500-2500[87] cells which is significantly more than what is present in the blood. While gonadal PGCs proliferate slower initially, they eventually overtake their blood counterparts and exceed them in number in later culture. Another method of increasing the number of PGCs, is to reprogramme CEFs through the introduction of stem cell specific genes and exposure to stem cell hormones in culture. While this was demonstrated to be feasible it must be noted that expression ability of these reprogrammed cells has not been fully investigated[96]. In addition to different extraction methods, culture conditions can also be optimised in order to encourage proliferation of PGCs. It was demonstrated that 3D culture offers a significant improvement in growth, however, additional reagents would need to be acquired to attempt this and they might prove costly[95]. For this reason, the first change that should be made to improve the number of PGCs is varying the extraction method.

Once sufficient numbers are obtained, PGC editing can undergo troubleshooting. Improving the likelihood of HDR could be done by exposing PGCs to different RS-1 and SCR-7

concentrations and then testing editing efficiency to identify which amount best promotes integration. In the event that these meiosis inducing chemicals fail, one could employ deheterochromatizing agents, such as epigenetic modification[122], in order to uncoil heterochromatic areas that could be preventing the gene drive from integrating and the gRNAs from cutting. Another strategy would be to re-design gRNAs to cut at a different area so as to avoid the heterochromatin or problem area completely. gRNAs could be designed in exon 2 and 3 of the OVA and OVM genes, respectively, as these regions were demonstrated to be modifiable by Oishi *et al* [65,99]. The gRNAs were originally designed to remove the entire OVA and OVM gene which may be difficult if there is super-coiling in the DNA in these regions. gRNA cutting efficiency should also be determined using ENGEN to ensure that the cassette is functional, before proceeding with the aforementioned measures. In addition to all these changes, PGCs can be Percoll purified before transfection in order to increase the number of cells transfectable as this was shown to improve transfection efficiency greatly[89]. This will be especially helpful if the lack of integration is purely due to low efficiency.

In order to improve repopulation efficiency, a larger number of PGCs would need to be injected or electroporation of constructs directly into the embryo could be explored. Additionally, cells could be allowed more time to recover after transfection, in order to maximise the number of recovered PGCs available for repopulation. A tracer construct containing an antibiotic resistance gene could also be employed to select for cells and, thus isolate only RFP containing cells, rather than using manual isolation. These suggestions should solve the issue around repopulation, however, to enact these would require a significant extension to the timeline.

Once integration and repopulation are accomplished, embryos would need to be hatched and assessed for integration using PCR, similarly to what was done for integration assessment in transfected PGCs. Alongside this, protein expression experiments will be performed in oviduct cells to determine whether the FIX protein is expressed. The presence of FIX can be easily identified in culture with a dip test for the His tag included in the FIX sequence. Further to this, a protein function assay will need to be done, known as a coagulation assay, which will determine the serine protease efficiency of the FIX protein.

In conclusion, this study provides valuable insight into the activity of the gene drive constructs in PGCs and where areas need improvement in order to accomplish integration in the chicken germline. It also provides a method of ablating native PGCs in order to prepare

embryos for repopulation with edited PGCs, it highlights areas where repopulation needs to be improved, and it provides a method of how to culture PGCs as well as select for them for transfection purposes. Additionally, it has proved valuable transfection conditions that can be used in the future as well as created constructs that can be used for editing PGCs and expressing human FIX in a transgenic chicken bioreactor. These preliminary findings accomplish the important first major steps towards creating cheaper therapeutic protein products and if successful in the future, could provide a construct capable of modifying other animal bioreactors for therapeutic expression.

5. References

1. Markets, R. and. Protein Therapeutics Market Analysis and Trends Report 2016: Therapeutic Proteins, Application & Function - Forecast to 2025 for the $315.9 Billion Market. *GlobeNewswire News Room* https://www.globenewswire.com/en/news-release/2017/01/26/911299/28124/en/Protein-Therapeutics-Market-Analysis-and-Trends-Report-2016-Therapeutic-Proteins-Application-Function-Forecast-to-2025-for-the-315-9-Billion-Market.html (2017).

2. Herron, L. R. *et al.* A chicken bioreactor for efficient production of functional cytokines. *BMC Biotechnology* **18**, 82 (2018).

3. Bertolini, L. R. *et al.* The transgenic animal platform for biopharmaceutical production. *Transgenic Res.* **25**, 329–343 (2016).

4. Pavlou, A. K. & Reichert, J. M. Recombinant protein therapeutics—success rates, market trends and values to 2010. *Nat Biotechnol* **22**, 1513–1519 (2004).

5. Herbert, H. K., Hyder, A. A., Butchart, A. & Norton, R. Global Health: Injuries and Violence. *Infectious Disease Clinics* **25**, 653–668 (2011).

6. Rogaev, E. I., Grigorenko, A. P., Faskhutdinova, G., Kittler, E. L. W. & Moliaka, Y. K. Genotype analysis identifies the cause of the 'royal disease'. *Science* **326**, 817 (2009).

7. Saini, S. & Dunn, A. L. Chapter 111 - Hemophilia B. in *Transfusion Medicine and Hemostasis (Third Edition)* (eds. Shaz, B. H., Hillyer, C. D. & Reyes Gil, M.) 685–690 (Elsevier, 2019). doi:10.1016/B978-0-12-813726-0.00111-2.

8. Cai, K. *et al.* Ensuring the Biologic Safety of Plasma-Derived Therapeutic Proteins. *BioDrugs* **19**, 79–96 (2005).

9. Tan, Y., Wu, H., Wei, T. & Li, X. Chemical Protein Synthesis: Advances, Challenges, and Outlooks. *J. Am. Chem. Soc.* **142**, 20288–20298 (2020).

10. Ziomek, C. A. Commercialization of proteins produced in the mammary gland. *Theriogenology* **49**, 139–144 (1998).

11. Fan, J. & Watanabe, T. Transgenic rabbits as therapeutic protein bioreactors and human disease models. *Pharmacology & Therapeutics* **99**, 261–282 (2003).

12. Pollock, D. P. *et al.* Transgenic milk as a method for the production of recombinant antibodies. *J Immunol Methods* **231**, 147–157 (1999).

13. Colman, A. Production of proteins in the milk of transgenic livestock: problems, solutions, and successes. *The American Journal of Clinical Nutrition* **63**, 639S-645S (1996).

14. Jänne, J. *et al.* Transgenic Animals as Bioproducers of Therapeutic Proteins. *Annals of Medicine* **24**, 273–280 (1992).

15. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* **27**, 297–306 (2009).

16. Kowarik, M. *et al.* N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* **314**, 1148–1150 (2006).

17. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol Cell Biochem* **307**, 249–264 (2008).

18. Carter, P. *et al.* High Level Escherichia coli Expression and Production of a Bivalent Humanized Antibody Fragment. *Nat Biotechnol* **10**, 163–167 (1992).

19. Schmidt, F. R. Recombinant expression systems in the pharmaceutical industry. *Appl Microbiol Biotechnol* **65**, 363–372 (2004).

20. Verma, D., Samson, N. P., Koya, V. & Daniell, H. A protocol for expression of foreign genes in chloroplasts. *Nat Protoc* **3**, 739–758 (2008).

21. Marsian, J. & Lomonossoff, G. P. Molecular pharming - VLPs made in plants. *Curr Opin Biotechnol* **37**, 201–206 (2016).

22. Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. & Romanos, M. A. High-Level Expression of Tetanus Toxin Fragment C in Pichia Pastoris Strains Containing Multiple Tandem Integrations of the Gene. *Nat Biotechnol* **9**, 455–460 (1991).

23. Burnett, M. J. B. & Burnett, A. C. Therapeutic recombinant protein production in plants: Challenges and opportunities. *PLANTS, PEOPLE, PLANET* **2**, 121–132 (2020).

24. Moustafa, K., Makhzoum, A. & Trémouillaux-Guiller, J. Molecular farming on rescue of pharma industry for next generations. *Crit Rev Biotechnol* **36**, 840–850 (2016).

25. Shahid, N. & Daniell, H. Plant-based oral vaccines against zoonotic and non-zoonotic diseases. *Plant Biotechnol J* **14**, 2079–2099 (2016).

26. Daniell, H., Lee, S. B., Panchal, T. & Wiebe, P. O. Expression of the native cholera toxin B subunit gene and assembly as functional oligomers in transgenic tobacco chloroplasts. *J Mol Biol* **311**, 1001–1009 (2001).

27. Daniell, H., Lin, C.-S., Yu, M. & Chang, W.-J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol* **17**, 134 (2016).

28. Shih, S. M.-H. & Doran, P. M. Foreign protein production using plant cell and organ cultures: Advantages and limitations. *Biotechnol Adv* **27**, 1036–1042 (2009).

29. Twyman, R. M., Stoger, E., Schillberg, S., Christou, P. & Fischer, R. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* **21**, 570–578 (2003).

30. Vézina, L.-P. *et al.* Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol J* **7**, 442–455 (2009).

31. Jayapal, K. P., Wlaschin, K. F., Hu, W.-S. & Yap, M. G. S. Recombinant Protein Therapeutics from CHO Cells - 20 Years and Counting. 8.

32. Kim, J. Y., Kim, Y.-G. & Lee, G. M. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl Microbiol Biotechnol* **93**, 917–930 (2012).

33. Perry, M. M. & Sang, H. M. Transgenesis in chickens. *Transgenic Research* **2**, 125–133 (1993).

34. Nakanishi, A., Utsumi, K. & Iritani, A. Early nuclear events of in vitro fertilization in the domestic fowl (Gallus domesticus). *Mol Reprod Dev* **26**, 217–221 (1990).

35. Naito, M., Nirasawa, K. & Oishi, T. Development in culture of the chick embryo from fertilized ovum to hatching. *J Exp Zool* **254**, 322–326 (1990).

36. Naito, M. *et al.* Embryonic expression of beta-actin-lacZ hybrid gene injected into the fertilized ovum of the domestic fowl. *Int J Dev Biol* **35**, 69–75 (1991).

37. Perry, M. M. A complete culture system for the chick embryo. *Nature* **331**, 70–72 (1988).

38. Yamaji, M. *et al.* Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* **40**, 1016–1022 (2008).

39. Ohinata, Y. *et al.* Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**, 207–213 (2005).

40. Nakamura, Y. *et al.* Migration and proliferation of primordial germ cells in the early chicken embryo. *Poult Sci* **86**, 2182–2193 (2007).

41. Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T. & Noce, T. Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. *Development* **127**, 2741–2750 (2000).

42. Petitte, J. N., Clark, M. E., Liu, G., Verrinder Gibbins, A. M. & Etches, R. J. Production of somatic and germline chimeras in the chicken by transfer of early blastodermal cells. *Development* **108**, 185–189 (1990).

43. Tajima, A. *et al.* Study on the concentration of circulating primordial germ cells (cPGCs) in early chick embryos. *J Exp Zool* **284**, 759–764 (1999).

44. Stebler, J. *et al.* Primordial germ cell migration in the chick and mouse embryo: the role of the chemokine SDF-1/CXCL12. *Dev Biol* **272**, 351–361 (2004).

45. Fujimoto, T., Ukeshima, A. & Kiyofuji, R. The origin, migration and morphology of the primordial germ cells in the chick embryo. *The Anatomical Record* **185**, 139–153 (1976).

46. Nakamura, Y., Kagami, H. & Tagami, T. Development, differentiation and manipulation of chicken germ cells. *Development, Growth & Differentiation* **55**, 20–40 (2013).

47. Swift, C. H. Origin of the sex-cords and definitive spermatogonia in the male chick. *American Journal of Anatomy* **20**, 375–410 (1916).

48. Hughes, G. C. THE POPULATION OF GERM CELLS IN THE DEVELOPING FEMALE CHICK. *J Embryol Exp Morphol* **11**, 513–536 (1963).

49. Perry, M. M. Nuclear events from fertilisation to the early cleavage stages in the domestic fowl (Gallus domesticus). *J Anat* **150**, 99–109 (1987).

50. Simkiss, K., Luke, G. & Behnam, J. Female chromosomes in cockerel ejaculates. *Proc Biol Sci* **263**, 1245–1249 (1996).

51. Marzullo, G. Production of chick chimaeras. *Nature* **225**, 72–73 (1970).

52. Nakamura, Y. *et al.* Germline Replacement by Transfer of Primordial Germ Cells into Partially Sterilized Embryos in the Chicken1. *Biology of Reproduction* **83**, 130–137 (2010).

53. Nakamura, Y. *et al.* Increased proportion of donor primordial germ cells in chimeric gonads by sterilisation of recipient embryos using busulfan sustained-release emulsion in chickens. *Reprod. Fertil. Dev.* **20**, 900 (2008).

54. Yamamoto, Y. *et al.* A Novel Method to Isolate Primordial Germ Cells and Its Use for the Generation of Germline Chimeras in Chicken1. *Biology of Reproduction* **77**, 115–119 (2007).

55. Trefil, P. *et al.* Male fertility restored by transplanting primordial germ cells into testes: a new way towards efficient transgenesis in chicken. *Sci Rep* **7**, 14246 (2017).

56. Li, B. *et al.* Efficient generation of transgenic chickens using the spermatogonial stem cells in vivo and ex vivo transfection. *SCI CHINA SER C* **51**, 734–742 (2008).

57. Li-Chan, E. C. Y., Powrie, W. D. & Nakai, S. The Chemistry of Eggs and Egg Products. in *Egg Science and Technology* (CRC Press, 1995).

58. McReynolds, L. *et al.* Sequence of chicken ovalbumin mRNA. *Nature* **273**, 723–728 (1978).

59. Heilig, R., Muraskowsky, R., Kloepfer, C. & Mandel, J. L. The ovalbumin gene family: complete sequence and structure of the Y gene. *Nucleic Acids Res* **10**, 4363–4382 (1982).

60. Da Silva, M. *et al.* The Family Secrets of Avian Egg-Specific Ovalbumin and Its Related Proteins Y and X. *Biol Reprod* **93**, 71 (2015).

61. Bahrami, S. *et al.* Designing A Transgenic Chicken: Applying New Approaches toward A Promising Bioreactor. *Cell J* **22**, 133–139 (2020).

62. Dávalos, A., Miguel, M., Bartolomé, B. & López-Fandiño, R. Antioxidant activity of peptides derived from egg white proteins by enzymatic hydrolysis. *J Food Prot* **67**, 1939–1944 (2004).

63. Matoba, N., Usui, H., Fujita, H. & Yoshikawa, M. A novel anti-hypertensive peptide derived from ovalbumin induces nitric oxide-mediated vasorelaxation in an isolated SHR mesenteric artery. *FEBS Lett* **452**, 181–184 (1999).

64. Pellegrini, A., Hülsmeier, A. J., Hunziker, P. & Thomas, U. Proteolytic fragments of ovalbumin display antimicrobial activity. *Biochim Biophys Acta* **1672**, 76–85 (2004).

65. Oishi, I., Yoshii, K., Miyahara, D. & Tagami, T. Efficient production of human interferon beta in the white of eggs from ovalbumin gene–targeted hens. *Scientific Reports* **8**, 1–12 (2018).

66. Shi, M., Kawabe, Y., Ito, A. & Kamihira, M. Targeted knock-in into the OVA locus of chicken cells using CRISPR/Cas9 system with homology-independent targeted integration. *J Biosci Bioeng* **129**, 363–370 (2020).

67. Lindenmaier, W. *et al.* Isolation and characterization of the chicken ovomucoid gene. *Nucleic Acids Res* **7**, 1221–1232 (1979).

68. Agarwal, V., Reddy, I. K. & Khan, M. A. Oral Delivery of Proteins: Effect of Chicken and Duck Ovomucoid on the Stability of Insulin in the Presence of α-Chymotrypsin and Trypsin. *Pharmacy and Pharmacology Communications* **6**, 223–227 (2000).

69. Holen, E., Bolann, B. & Elsayed, S. Novel B and T cell epitopes of chicken ovomucoid (Gal d 1) induce T cell secretion of IL-6, IL-13, and IFN-gamma. *Clin Exp Allergy* **31**, 952–964 (2001).

70. Oishi, I., Yoshii, K., Miyahara, D., Kagami, H. & Tagami, T. Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Scientific reports* (2016) doi:10.1038/srep23980.

71. Bernhisel-Broadbent, J., Dintzis, H. M., Dintzis, R. Z. & Sampson, H. A. Allergenicity and antigenicity of chicken egg ovomucoid (Gal d III) compared with ovalbumin (Gal d I) in children with egg allergy and in mice. *J. Allergy Clin. Immunol.* **93**, 1047–1059 (1994).

72. HONMA, K. *et al.* Allergenic epitopes of ovalbumin (OVA) in patients with hen's egg allergy: inhibition of basophil histamine release by haptenic ovalbumin peptide. *Clin Exp Immunol* **103**, 446–453 (1996).

73. Abeyrathne, E. D. N. S., Lee, H. Y. & Ahn, D. U. Egg white proteins and their potential use in food processing or as nutraceutical and pharmaceutical agents—A review. *Poultry Science* **92**, 3292–3299 (2013).

74. Jeltsch, J.-M. & Chambon, P. The Complete Nucleotide Sequence of the Chicken Ovotransferrin mRNA. *European Journal of Biochemistry* **122**, 291–295 (1982).

75. Rathnapala, E. C. N., Ahn, D. U. & Abeyrathne, S. Functional properties of ovotransferrin from chicken egg white and its derived peptides: a review. *Food Sci Biotechnol* **30**, 619–630 (2021).

76. Salter, D. W. & Crittenden, L. B. Artificial insertion of a dominant gene for resistance to avian leukosis virus into the germ line of the chicken. *Theoret. Appl. Genetics* **77**, 457–461 (1989).

77. Okon, B., Ibom, L. A. & Njume, G. N. Transgenesis techniques and its application in poultry production. *Global Journal of Agricultural Sciences* **15**, 11–15 (2016).

78. Lillico, S. G. *et al.* Oviduct-specific expression of two therapeutic proteins in transgenic hens. *Proc Natl Acad Sci U S A* **104**, 1771–1776 (2007).

79. Harvey, A. J., Speksnijder, G., Baugh, L. R., Morris, J. A. & Ivarie, R. Consistent production of transgenic chickens using replication-deficient retroviral vectors and high-throughput screening procedures. *Poult Sci* **81**, 202–212 (2002).

80. Cyriac, S., George, C. & Churchil, R. TRANSGENIC CHICKEN: METHODS AND APPLICATIONS. **10**, 67–71 (2012).

81. Taylor, L. *et al.* Efficient TALEN-mediated gene targeting of chicken primordial germ cells. *Development* **144**, 928–934 (2017).

82. Fan, B. *et al.* Assembly and in vitro Functional Analysis of Zinc Finger Nuclease Specific to the 3′ Untranslated Region of Chicken Ovalbumin Gene. *Animal Biotechnology* **22**, 211–222 (2011).

83. Grunwald, H. A. *et al.* Super-Mendelian inheritance mediated by CRISPR–Cas9 in the female mouse germline. *Nature* **566**, 105–109 (2019).

84. Gantz, V. M. & Bier, E. The mutagenic chain reaction: A method for converting heterozygous to homozygous mutations. *Science* **348**, 442–444 (2015).

85. Kyrou, K. *et al.* A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. *Nature Biotechnology* **36**, 1062–1066 (2018).

86. Szczerba, A., Kuwana, T. & Bednarczyk, M. Concentration and total number of circulating primordial germ cells in Green-legged Partridgelike chicken embryos. *Poult Sci* **100**, 319–324 (2020).

87. Motono, M., Ohashi, T., Nishijima, K. & Iijima, S. Analysis of chicken primordial germ cells. *Cytotechnology* **57**, 199–205 (2008).

88. Chojnacka-Puchta, L., Kasperczyk, K., Płucienniczak, G., Sawicka, D. & Bednarczyk, M. Primordial germ cells (PGCs) as a tool for creating transgenic chickens. *Polish journal of veterinary sciences* **15**, 181–8 (2012).

89. Chojnacka-Puchta, L. *et al.* Obtaining chicken primordial germ cells used for gene transfer: in vitro and in vivo results. *J Appl Genet* **56**, 493–504 (2015).

90. Kong, L. *et al.* Long-term in vitro culture and preliminary establishment of chicken primordial germ cell lines. *PLOS ONE* **13**, e0196459 (2018).

91. MIYAHARA, D. *et al.* Chicken stem cell factor enhances primordial germ cell proliferation cooperatively with fibroblast growth factor 2. *J Reprod Dev* **62**, 143–149 (2016).

92. Makoolati, Z., Movahedin, M. & Forouzandeh-Moghadam, M. Proliferation in culture of primordial germ cells derived from embryonic stem cell: induction by retinoic acid. *Biosci Rep* **36**, e00428 (2016).

93. Macdonald, J., Glover, J. D., Taylor, L., Sang, H. M. & McGrew, M. J. Characterisation and Germline Transmission of Cultured Avian Primordial Germ Cells. *PLoS One* **5**, e15518 (2010).

94. Park, T. S. & Han, J. Y. Derivation and characterization of pluripotent embryonic germ cells in chicken. *Mol Reprod Dev* **56**, 475–482 (2000).

95. Chen, Y.-C. *et al.* Three-dimensional culture of chicken primordial germ cells (cPGCs) in defined media containing the functional polymer FP003. *PLOS ONE* **13**, e0200515 (2018).

96. Zhao, R. *et al.* Production of viable chicken by allogeneic transplantation of primordial germ cells induced from somatic cells. *Nat Commun* **12**, 2989 (2021).

97. Hong, Y. H., Moon, Y. K., Jeong, D. K. & Han, J. Y. Improved Transfection Efficiency of Chicken Gonadal Primordial Germ Cells for the Production of Transgenic Poultry. *Transgenic Res* **7**, 247–252 (1998).

98. Sawicka, D. & Chojnacka-Puchta, L. Effective Transfection of Chicken Primordial Germ Cells (PGCs) using Transposon Vectors and Lipofection. *folia biol (krakow)* **67**, 45–52 (2019).

99. Oishi, I. Improvement of Transfection Efficiency in Cultured Chicken Primordial Germ Cells by Percoll Density Gradient Centrifugation. *Bioscience, Biotechnology, and Biochemistry* **74**, 2426–2430 (2010).

100. Eghbalsaied, S. & Kues, W. A. Development of a transposon-based technology for transfection of day 0 chicken embryos. *Gene* **730**, 144318 (2020).

101. Motono, M., Komatsu, H., Kawabe, Y., Nishijima, K. & Iijima, S. Sterilization of Chicken Primordial Germ Cells. (2009) doi:10.1007/978-1-4020-9646-4_26.

102. Wentworth, B. C., Tsai, H., Hallett, J. H., Gonzales, D. S. & Rajcic-spasojevic, G. Manipulation of Avian Primordial Germ Cells and Gonadal Differentiation1. *Poultry Science* **68**, 999–1010 (1989).

103. Atsumi, Y. *et al.* Depletion of Primordial Germ Cells (PGCs) by X-irradiation to Extraembryonic Region of Chicken Embryos and Expression of Xenotransplanted Quail PGCs. *Journal of Poultry Science - J POULT SCI* **46**, 136–143 (2009).

104. Aige-Gil, V. & Simkiss, K. Sterilisation of avian embryos with busulphan. *Research in Veterinary Science* **50**, 139–144 (1991).

105. Song, Y., D'Costa, S., Pardue, S. L. & Petitte, J. N. Production of germline chimeric chickens following the administration of a busulfan emulsion. *Molecular Reproduction and Development* **70**, 438–444 (2005).

106. Nakamura, Y. *et al.* Effects of Busulfan Sustained-release Emulsion on Depletion and Repopulation of Primordial Germ Cells in Early Chicken Embryos. *Journal of Poultry Science - J POULT SCI* **46**, 127–135 (2009).

107. Kim, Y. M., Park, K. J., Park, J. S., Jung, K. M. & Han, J. Y. In vivo enrichment of busulfan-resistant germ cells for efficient production of transgenic avian models. *Sci Rep* **11**, 9127 (2021).

108. Ballantyne, M. *et al.* Direct allele introgression into pure chicken breeds using Sire Dam Surrogate (SDS) mating. *Nat Commun* **12**, 659 (2021).

109. Hu, T. *et al.* A low-tech, cost-effective and efficient method for safeguarding genetic diversity by direct cryopreservation of poultry embryonic reproductive cells. *eLife* **11**, e74036.

110. NAKAMURA, Y. Poultry genetic resource conservation using primordial germ cells. *J Reprod Dev* **62**, 431–437 (2016).

111. Li, N. *et al.* Adult lifetime cost of hemophilia B management in the US: payer and societal perspectives from a decision analytic model. *Journal of Medical Economics* **24**, 363–372 (2021).

112. Bolton-Maggs, P. H. B. & Pasi, K. J. Haemophilias A and B. *Lancet* **361**, 1801–1809 (2003).

113. Morfini, M. *et al.* Evaluation of prophylactic replacement therapy in haemophilia B. *Scand J Haematol* **16**, 41–47 (1976).

114. Srivastava, A. *et al.* Guidelines for the management of hemophilia. *Haemophilia* **19**, e1–e47 (2013).

115. Franchini, M., Frattini, F., Crestani, S., Sissa, C. & Bonfanti, C. Treatment of hemophilia B: focus on recombinant factor IX. *Biologics* **7**, 33–38 (2013).

116. Shen, G., Gao, M., Cao, Q. & Li, W. The Molecular Basis of FIX Deficiency in Hemophilia B. *Int J Mol Sci* **23**, 2762 (2022).

117. Franchini, M. & Mannucci, P. M. Haemophilia B is clinically less severe than haemophilia A: further evidence. *Blood Transfus* **16**, 121–122 (2018).

118. Liu, J., Jonebring, A., Hagström, J., Nyström, A.-C. & Lövgren, A. Improved Expression of Recombinant Human Factor IX by Co-expression of GGCX, VKOR and Furin. *Protein J* **33**, 174–183 (2014).

119. Smith, S. A., Travers, R. J. & Morrissey, J. H. How it all starts: initiation of the clotting cascade. *Crit Rev Biochem Mol Biol* **50**, 326–336 (2015).

120. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **35**, 1547–1549 (2018).

121. Clinton, M., Haines, L., Belloir, B. & McBride, D. Sexing chick embryos: A rapid and simple protocol. *British Poultry Science* **42**, 134–138 (2001).

122. Janssen, J. M., Chen, X., Liu, J. & Gonçalves, M. A. F. V. The Chromatin Structure of CRISPR-Cas9 Target DNA Controls the Balance between Mutagenic and Homology-Directed Gene-Editing Events. *Mol Ther Nucleic Acids* **16**, 141–154 (2019).

6. Supplementary

6.1 Sequencing for gene drives

Sequencing was performed for each modification in the gene drive, except the homology arms, in order to reconfirm that insertion did take place as well as confirm the sequence was free of any major mutations that could hinder downstream and future applications. Homology arms were not sequenced as sequence variability in this region is forgiving.

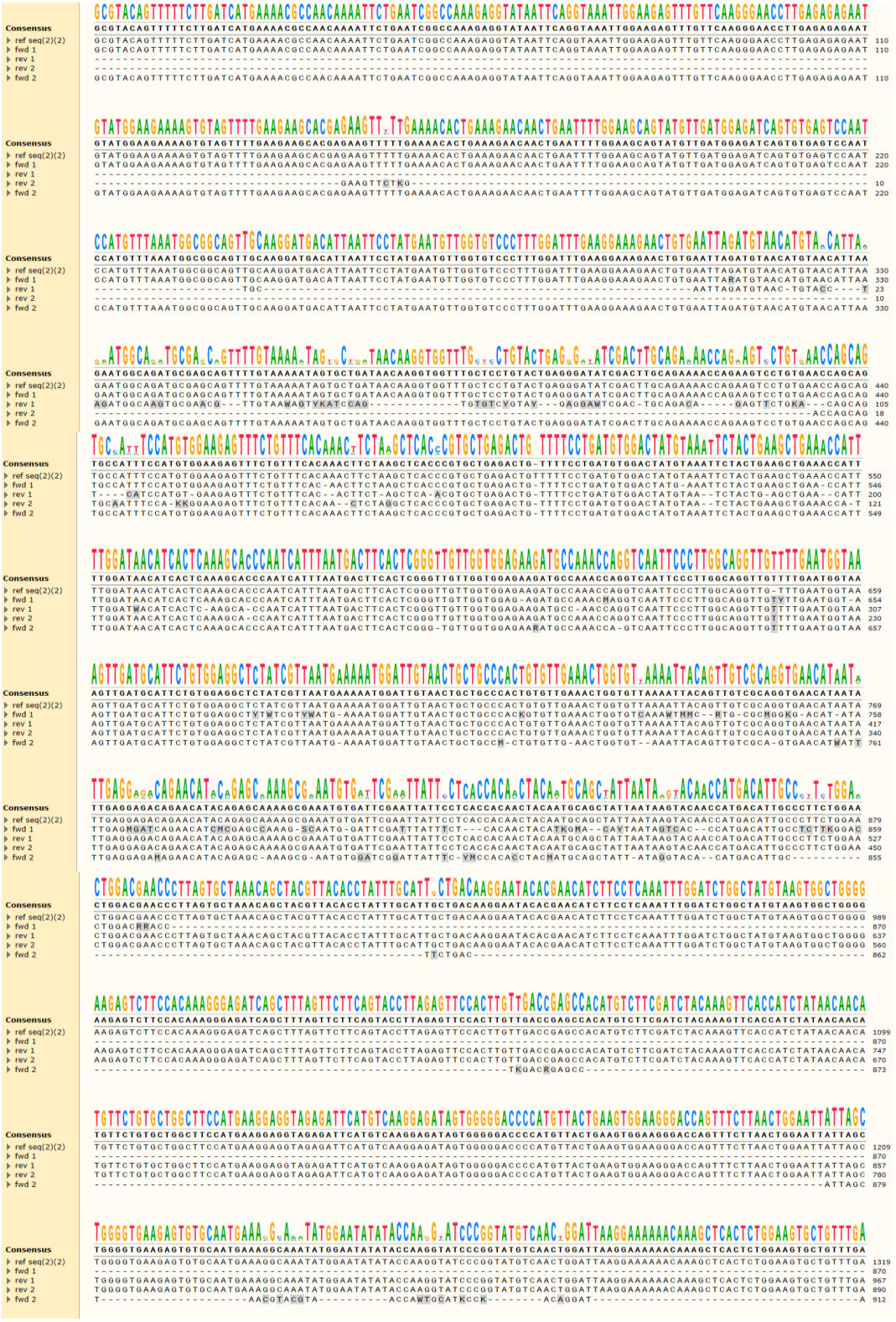


Figure 32: FIX sequencing results for the OVA-HA-FIX gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing results to this sequence are highlighted in grey. Multiple sequence results for a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.

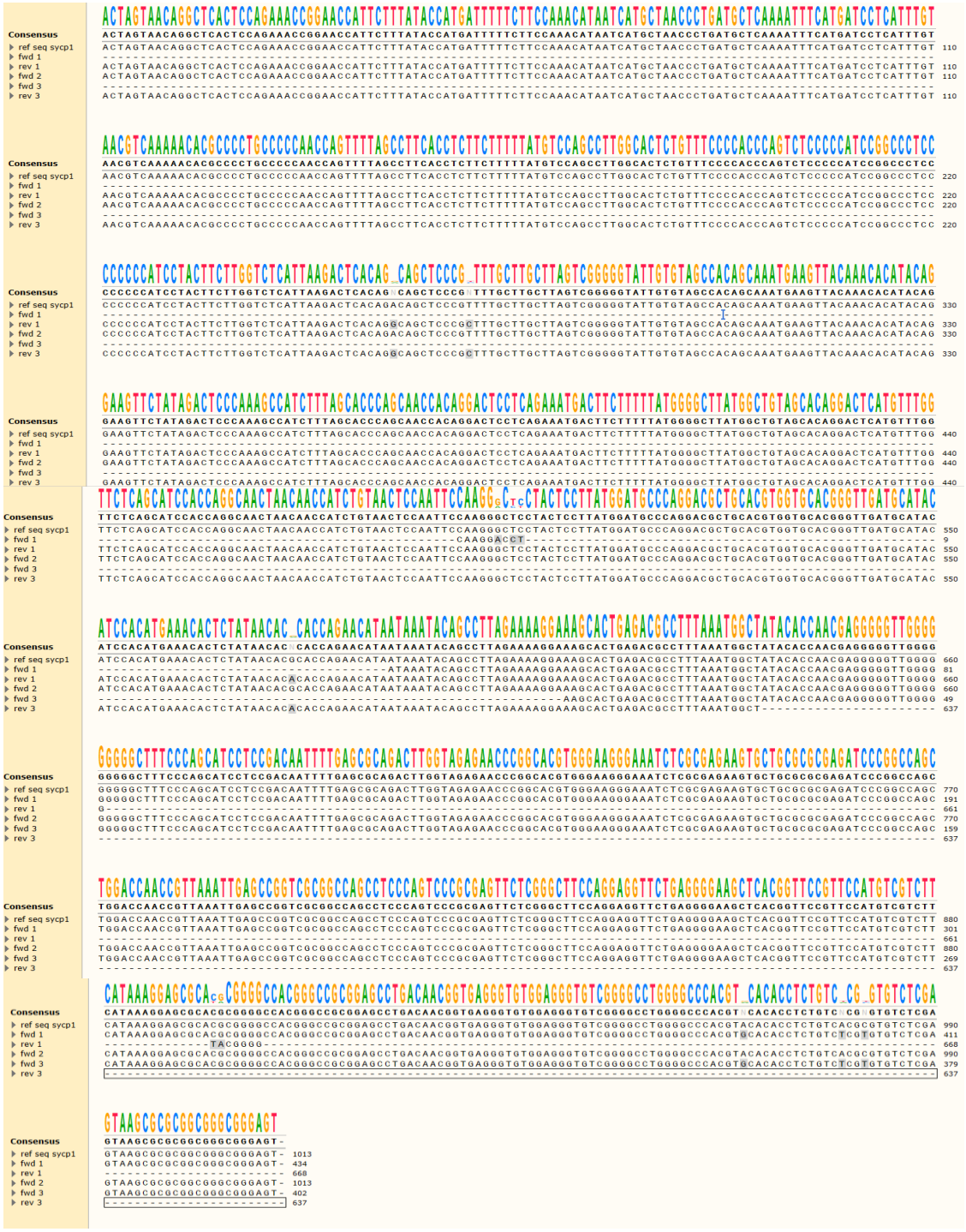


Figure 33: Sycp1 sequencing results for the OVA-HA-Syc-FIX gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing

results to this sequence are highlighted in grey. Multiple sequence results for a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.

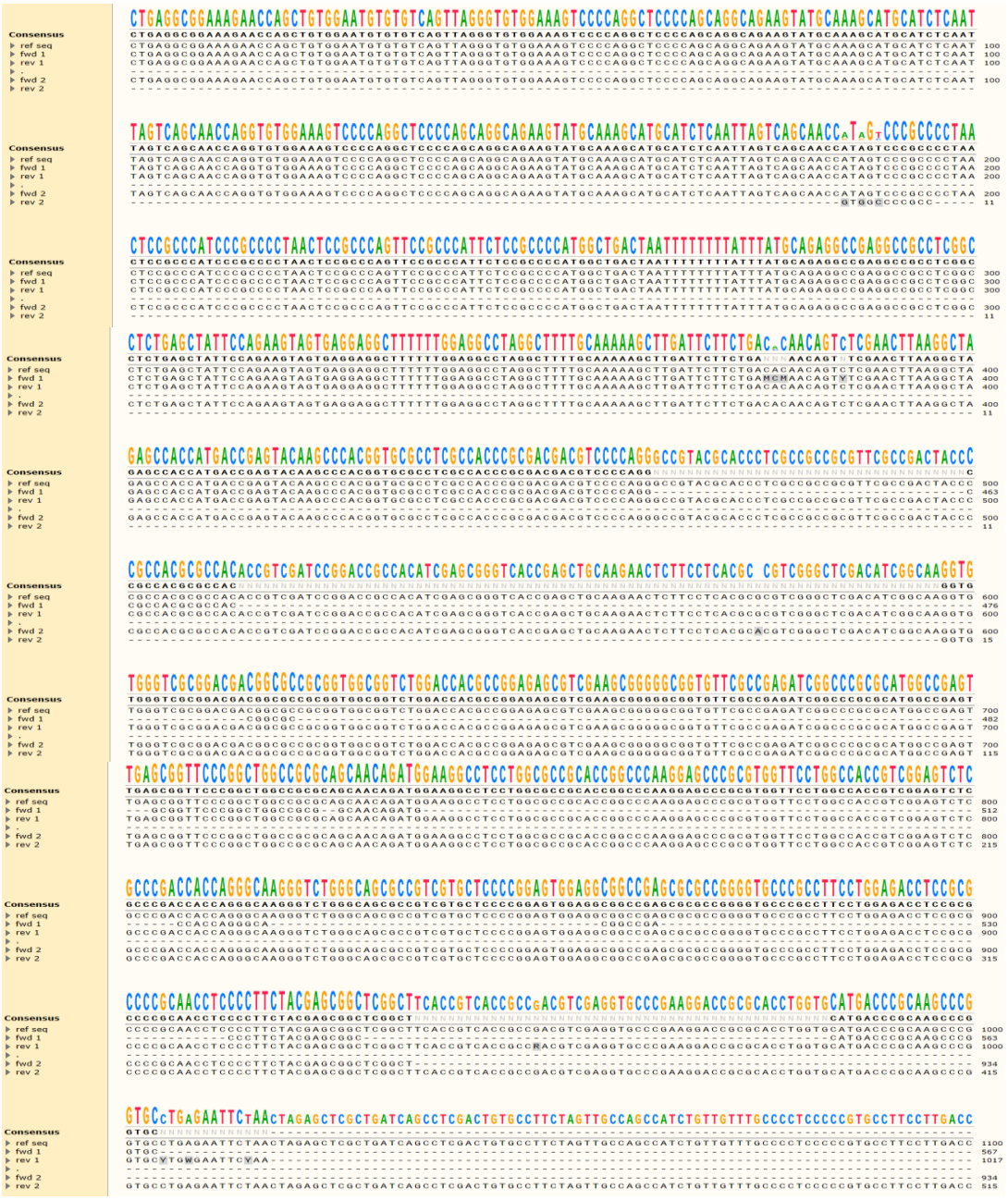


Figure 34: Puromycin resistance gene sequencing results for the OVA-HA-Syc-FIX-Puro gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing results to this sequence are highlighted in grey. Multiple sequence results for a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.

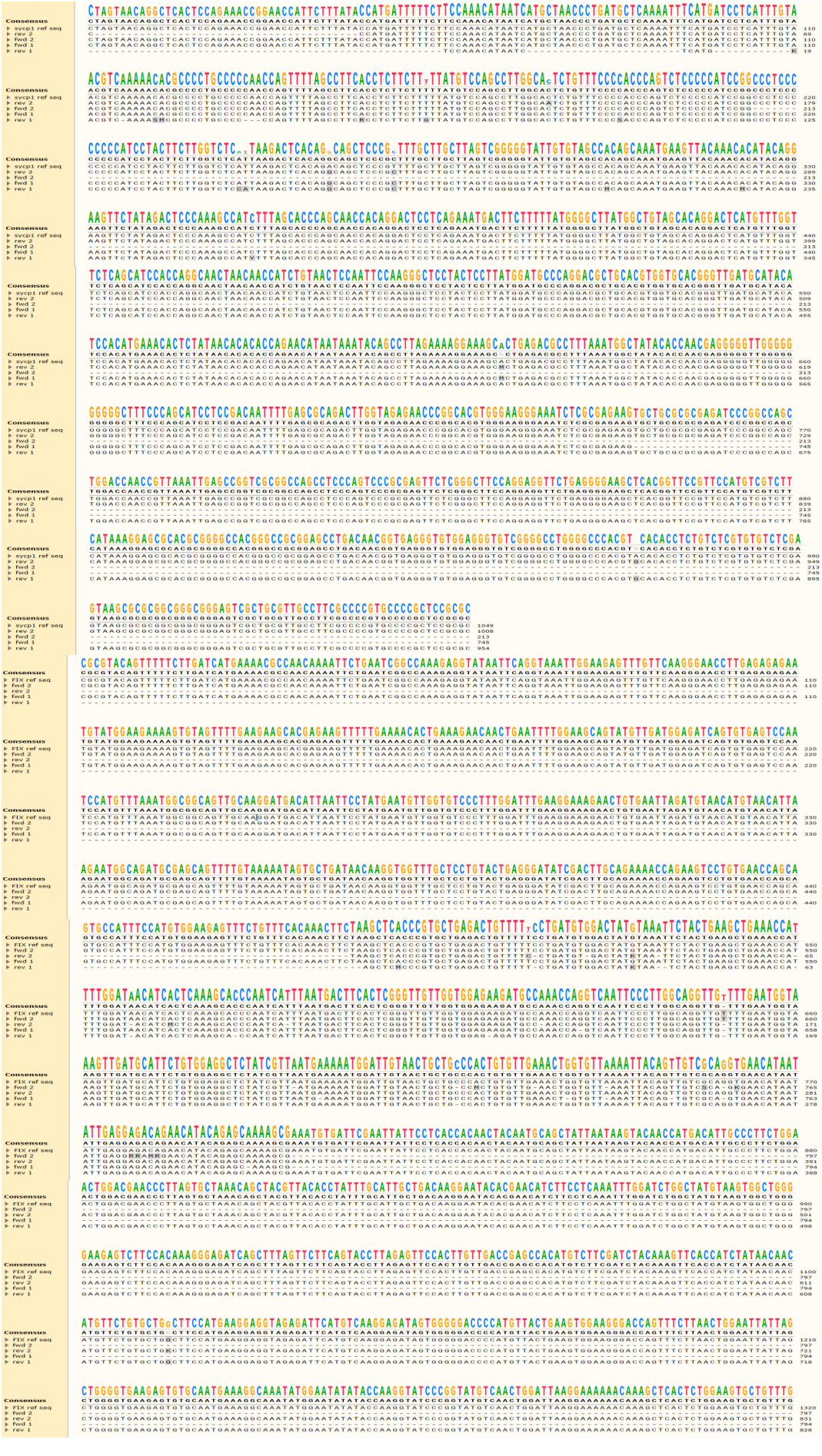


Figure 35: Sycp1 and FIX sequencing results for the OVM-Syc-FIX gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing results to this sequence are highlighted in grey. Multiple sequence results for a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.

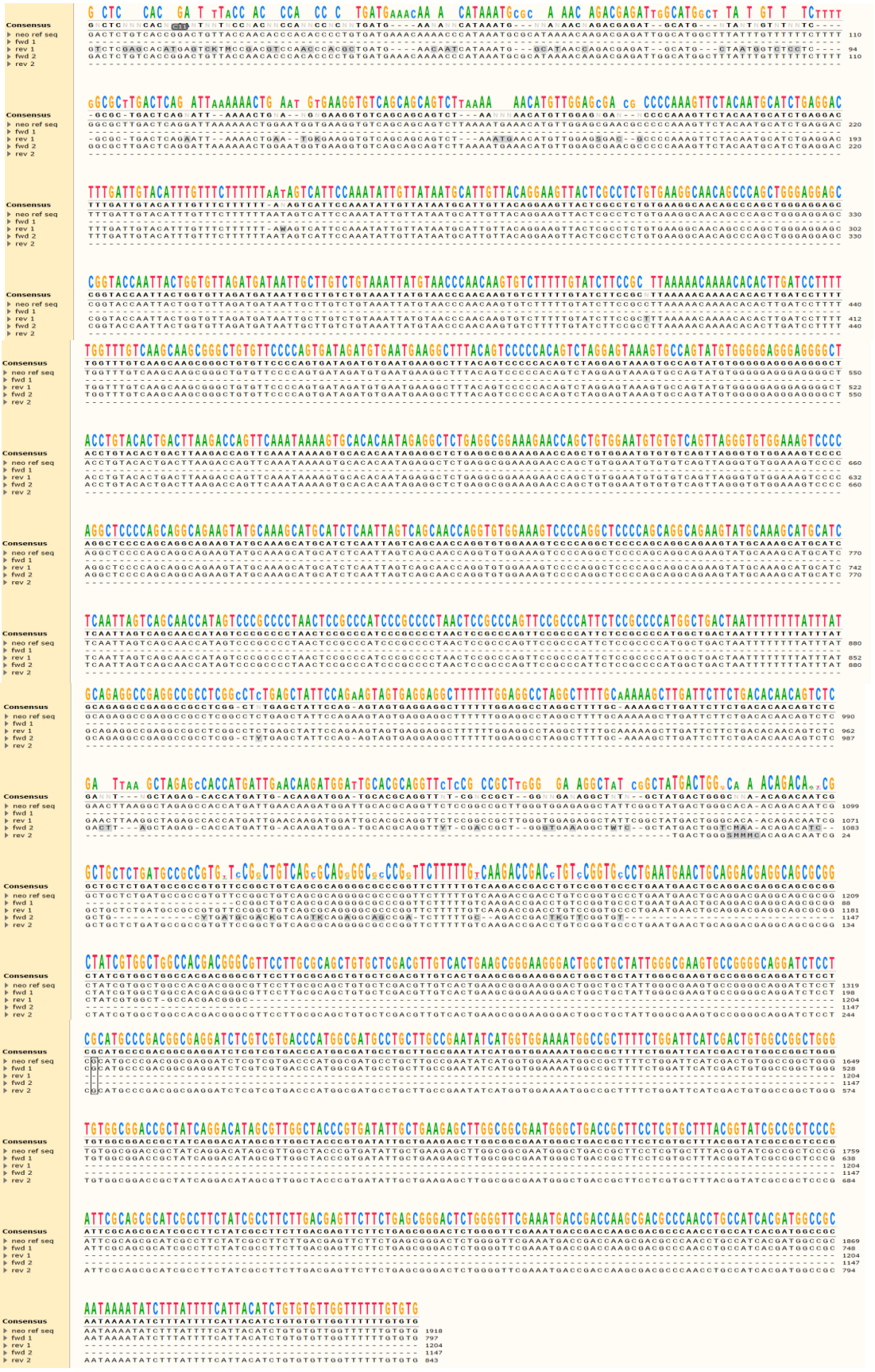


Figure 36: Neomycin resistance gene sequencing results for the OVM-Syc-FIX-ss-Neo gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing results to this sequence are highlighted in grey. Multiple sequence results for

a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.



Figure 37: ssOVM g RNA cassette sequencing results for the OVM-Syc-FIX-ss gene drive. The reference sequence is presented first as “ref seq” and all bases that do not match in the sequencing results to this sequence are highlighted in grey. Multiple sequence results for a single strand are shown as “fwd” or “rev” depending on which direction they were sequenced in.

All sequencing results indicate that the necessary insert was in fact present as well as being the correct sequence, free of major mutations. Base changes were disregarded if at least one read had a base that matched the reference sequence.

6.2 Sexing gels

Sexing was performed for PGC extraction with a representational gel below as well as in the case of busulfan ablation and repopulation assessment.

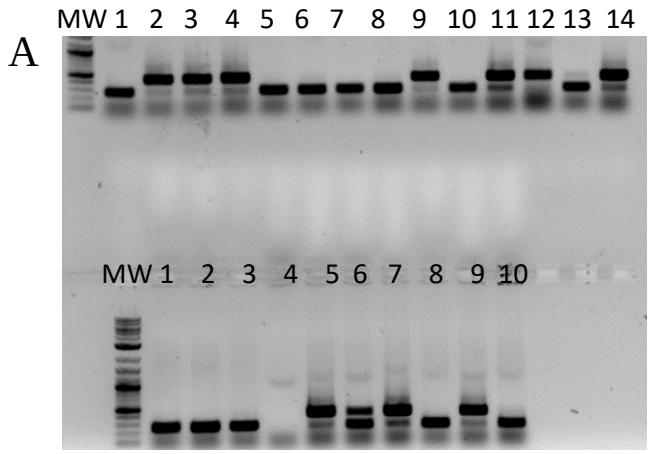


Figure 38: Sexing gel for PGC extraction. In each case the presence of one band indicates a male and in the case of two bands a female. The lower band is a 18s control, and the upper band amplifies a region on the W chromosome, present only in females. For each of these, only males were selected, and females were discarded for further experiments.

6.3 CEF culture

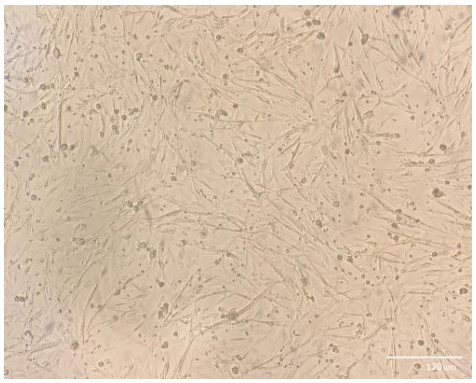


Figure 39: CEFs plated in a T75cm^2 flask one day after harvest. CEFs were harvested from day 9-10 embryo torsos. CEFs are present at 80% and confirm the successful isolation of CEFs from day 9-10 embryos for use as a feeder layer. Scale bar = 120 μm.

6.4 Defects found at various concentrations of busulfan treatment

Table 12: Defects recorded across embryos exposed to different concentrations of busulfan. Embryos were observed at day 7.5 of development.

Concentration of busulfan µg/50µl	Defect	No. of embryos out of a total of 17 per group that had the defect	Natural dud embryos in each group
0	None	0	4
65	None	0	6
75	Development delay, small eyes	1	6
85	Development delay, small eyes, overproduction of blood/enlarged vessels, underdeveloped iris	3	6

7. Recipes

Recipe 1:Terrific Broth

Component	Amount/action
Buffer A	
Yeast extract	24 g
Tryptone	12 g
Glycerol	4 g
dH20	900 ml
Autoclave	
Buffer B	
KH_2PO_4	2.3 g
K_2HPO_4	16.4 g
dH2O	100 ml
Filter sterilize	
Add A to B	

Recipe 2: Crude extraction lysis buffer

Component	Amount/action
For Blood	
Citric acid	0.48 % m/v
Sodium citrate	1.32 %
For tissue	
Tris-HCL/Base	1 M
EDTA	0.5 M
NaCl	1 mol
dH2O	To 20 ml

Recipe 3: Ringer's solution

Component	Amount/action
NaCl	7.2 g
$CaCl_2.2H_2O$	0.23 g
KCl	0.37 g
dH_2O	1 l
pH 7.4	