Generating Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) deletions in Bladder Cancer Cell Lines

The School of Biosciences

Adam Dooner

2019

A thesis submitted to the University of Kent for the degree of MRes in Cancer Genetics in the Faculty of Science, Technology and Medical Studies.

I Declaration

No part of this thesis has been submitted in support of an application for any degree or qualification of the University of Kent or any other University or institute of learning

Adam Dooner

II Acknowledgements

My thanks go to all in the Fenton lab, especially Tim Fenton, for all of his help and supervision throughout the year. This has been a fantastic opportunity and I have enjoyed every moment of it. I would also like to thank Nerissa Kirkwood for her help with sequencing RT-qPCR and TOPO cloning! Nikki Smith, for help with all the formulas, and Maxmilan Jeya with general lab organisation. Many thanks go to Martin Michaelis for providing the cells I worked with, and Jo Bird for help setting up cell culture with them.

I would also like to thank my family and friends, who have always believed in me and supported me throughout my studies.

Lastly, I would like to dedicate this to my late grandmother, none of this could be possible without her support throughout the years which will continue to drive me for many more to come. You will be sorely missed.

III Contents

IV List of Figures

V Abbreviations

A3	APOBEC3
A3A	APOBEC3A
A3B	APOBEC3B
AID	Activation-Induced cytidine Deaminase
APOBEC	Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like
BER	Break Excision Repair
BFP	Blue Fluorescent Protein
BFTC	BFTC-905 Cell Line
bp	Base Pairs
bpA	Bovine growth hormone PolyAdenylation signal
cDNA	Complementary DNA
CNA	Copy-Number Alteration
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSR	Class-Switch Recombination
DDR	DNA Damage Response
DMSO	Dimethyl Sulfoxide
DNA	DeoxyriboNucleic Acid
DSB	Double-Strand Break
dsDNA	Double-Stranded DNA
FGFR3	Fibroblast Growth Factor Receptor 3
FRT	Flippase Recognition Target
GAPDH	GlycerAldehyde 3-Phosphate DeHydrogenase
GFP	Green Fluorescent Protein
gRNA	Guide RNA

HDR	Homology Directed Repair
HIV	Human Immunodeficiency Virus
HPV	Human PapillomaVirus
HSV-TK	Herpes Simplex Virus Thymidine Kinase Gene
IFN	Interferon
IMDM	Iscove's Modified Dulbecco's Medium
MMR	MisMatch Repair
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
PGK	PhosphoGlycerate Kinase
RNA	RiboNucleic Acid
RPA	Replication Protein-A
RPKM	Reads Per Kilobase of transcript per Million mapped reads
RT-qPCR	Reverse Transcription Quantitative PCR
SHM	Somatic Hyper-Mutation
SNP	Single-Nucleotide Polymorphism
SCC	Single-Cell Cloning
ssDNA	Single-Stranded DNA
TBP	TATA-Binding Protein
TLS	TransLesion Synthesis
UBC	Urothelial Bladder Cancer
UNG	Uracil-DNA Glycosylase

VI Abstract

Bladder cancer was responsible for 200,000 deaths worldwide in 2018, it has a well-documented ability to evade and resist chemotherapy treatment. Apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) enzymes are a family of enzymes and form part of the innate immune system in human cells where they mutate cDNA from infecting viruses. 2 APOBECs: APOBEC3A (A3A) and APOBEC3B (A3B) have been linked not only to bladder cancer, but to half of all cancers. These APOBECs leave a distinct mutational signature which has led to many computational studies into APOBEC involvement in cancers.

APOBECs seem to have a direct role in evading chemotheraputic agents; through upregulation in cancers and increased mutational signature after rounds of chemotherapy. Because of this it has been postulated that development of techniques to downregulate APOBEC expression in cancers, alongside traditional chemotherapy drugs, will stop recurrence.

To test this theory cancer-cell lines with homozygous knockouts for A3A and/or A3B must be created to assess the link between these genes and resistance to traditional chemotherapy drugs. This thesis attempts to create these knockouts in BFTC-905, T24 and 5637 bladder cancer cell lines.

A3A was successfully knocked out in one line: BFTC-A2-F, and A3B was potentially knocked out in BFTC-Both3-F and BFTC-Both4-F. This experiment has laid the foundation for more A3A and A3B knockouts to be cloned from pools of knock-out targeted cell lines: 5637, T24 and BFTC-905 or through BFTC-A2-F which can be used as a platform to attempt to knock out A3B too. It would appear that trying to knock-out both A3A and A3B at the same time creates some unexpected results, possibly due to how close these genes are together (~30kb) in the APOBEC3 locus. Deletion of A3B appears hard to achieve, as only a heterozygous deletion in BFTC-B5-F was achieved, these results echo a previous experiment in a normal cell line.

1 Introduction

1.1 Bladder cancer

Urothelial Bladder Cancer (UBC) is a high mortality rate cancer with 3.2 and 0.9 deaths per 100,000 in men and women respectively (worldwide), with an estimated 550,000 new cases and 200,000 deaths worldwide in 2018 (1). Around 25% of patients present with highly aggressive Muscle-Invasive Bladder Cancer (MIBCs), this form requires radical surgery and radiotherapy quickly, often still leading to a poor outcome. The other 75%, Non-Muscle-Invasive Bladder Cancers (NMIBCs) require local therapy but still have a high chance of recurrence or progression afterwards. Bladder cancer is caused by environmental carcinogens often coming from occupational exposure which accumulate in the bladder, smoking is the main attributing factor accounting for nearly 50% of UBCs. It is believed that the cancer-causing agents are aromatic amides, this is reflected by the genetic susceptibility of some patients who have a lower ability to reduce the presence of aromatic amides in their bladder (2).

A commonly activated oncogene in UBCs is Fibroblast Growth Factor Receptor 3 (FGFR3) (2, 3), most often via the recurrent S249C mutation (T<u>C</u>C → T<u>G</u>C), found in 65% NMIBCs and 15% of MIBCs (20). On top of this, a sequence variant 4p16.3 (just upstream of FGFR3), is often found in UBCs with an activating FGFR3 mutation (2).

1.2 Sub-clones

The UBC high rates of relapse and progression (2, 4), appear to be due to its ability to often avoid and adapt to treatment by anti-cancer drugs (4, 5, 6). Many cancers, specifically in esophagus and lung cancer, appear to be more deadly when there is a more diverse population within a tumour. Sub-clones within tumours may develop new mutations that are beneficial to the cancer and would allow for a drug resistant line to develop. These resistant sub-clones are positively selected for when the patient is treated with chemotherapy drugs, increasing the disease's chance of recurrence (7). Indeed, UBCs tend to be a highly heterogeneous cancer with many sub-clones harbouring different mutations (4, 5, 6). A mechanism for such heterogeneity would require mass mutation and rearrangement of genomic DNA, a group of enzymes thought to be able to contribute to this is the APOBEC family of cytosine deaminases.

1.3 APOBEC

The APOBEC family of enzymes are common throughout mammals and span 3 chromosomes. The family consisting of: Activation-Induced cytidine Deaminase (AID), APOBEC1, APOBEC2 which seems to have no catalytic ability, differing copy-numbers of APOBEC3 (A3) and APOBEC4 which is uncategorised. Humans have 7

versions of the A3 gene all residing on chromosome 11: A3A, A3B, A3C, A3D, A3F, A3G and A3H leading to 11 APOBECs in total. All APOBECs (bar APOBEC2 and 4) target single-stranded DNA (ssDNA), some with the ability to also target RNA. AID is the most studied APOBEC mechanistically, it is part of Somatic Hyper-Mutation (SHM) and Class Switch Recombination (CSR) driving the diversification of antibodies in B-lymphocytes. A3A and A3B have an affinity for TCW sites (where W represents Thymidine or Adenine), where they deaminate Cytosine to Uridine (TCW → TUW). A3A and A3B are part of the cell's innate immune system, induced by Interferons (IFN) α/β and γ respectively. Usually they attack viral cDNA in the cytoplasm and nucleus preventing infection and replication (7, 8, 9, 10, 11).

1.4 APOBEC Mutational Mechanism

AIDs mechanism of action has previously been explained using yeast models (9), with a similar mechanism thought to apply in A3A and A3B, despite no full model has been described for these enzymes yet. Nor has the total crystal structures been resolved, which would lead to better understanding of the full function of the enzyme, though the active site has been resolved and is well understood (7).

During immunoglobulin SHM of B cells, AID uses a non-canonical Break Excision Repair (BER) pathway to attack ssDNA either side of a double stranded break after mis-match repair (MMR) (8, 9, 10, 11, 12). AID has been found to deaminate G-rich DNA supporting a theory that it attacks R-loops during transcription, where ssDNA is generated on the lagging strand (9, 13). An example of this is the common FGFR3

S248C mutation mentioned earlier, this is thought to be a gene replicated primarily as a lagging strand, which Shi et al. propose could provide hairpins for A3s to attack (3). On top of this, it was found in CSR that AID is required for Double-Strand Breaks (DSBs), where it marks these sites and recruits γ-H2AX (12).

AID produces these DSBs by deaminating dC → dU, dU is then removed by Uracil-DNA Glycosylase (UNG) leaving an abasic site. An endonuclease then removes the DNA backbone from the abasic site creating a nick, if left unrepaired another abasic site may form on the opposite strand leading to formation of a DSB (8, 9, 12). CSR has been found to be dependent on UNG accompanied by the proper BER proteins, without them defects occur (12). When these sites are not repaired, they lead to a C>T transition (8, 9, 12), Translesion Synthesis (TLS) polymerases can insert bases into the abasic site left by UNG during replication. TLS polymerases commonly insert Adenine opposite the abasic site during replication, later leading to a to C>T transition on the original strand, when replicated. Other TLS polymerases insert other bases, such as Cytosine, leading to C>G another common APOBEC signature mutation (8, 9).

It is thought that other APOBECs work in a similar fashion, though induction differs; A3A is induced when cytoplasmic DNA has been recognised by the cell. This DNA is transcribed by RNA polymerase III leading to double-stranded RNA intermediates which are captured by RIG-I, inducing type-I IFN production via a cascade, in turn inducing A3A. It is known that IFN-α increases levels of A3A in a cell but not A3B. Suspene et al. find these results by introducing ssDNA to the cytoplasm of

monocytes; IFN-α and IFN-β were released, increasing the expression of A3A. The induction of A3A by IFN-α/β leads to destabilisation of cytoplasmic DNA which is then digested by exonuclease (13). In theory, constant inflammation of cells (constant expression of IFNs) could lead to extended periods of A3A overexpression, genomic DNA destabilisation and then cancer (5, 8, 9, 13). Suspene et al. wanted to find which IFN induced A3A directly, they designed an experiment which included antibodies for IFNs; IFN-α, IFN-β and IFN-γ. These antibodies were added separately while monocytes were transfected with mtDNA. Suspene et al. find that when IFN-β antibodies were present there was a reduction in A3A expression, while other antibodies had no effect, suggesting that IFN-β has a paracrine effect on A3A expression (13).

1.5 APOBEC Mutation in Cancer

As previously mentioned, A3s leave behind mutational markers based on their target TC sites and the usual C>T or C>G mutation at specific sequence sites. These sites allow for computational studies of known cancer genomes, it has been found that APOBEC based mutations C>T/G at TC sites is a large signature in many cancers. Though it is not in all cancers, the APOBEC signature is common in cervical, breast and bladder cancers, being found in at least 17 others (8, 9, 10, 14). A3B shows very high expression in around 50% of cancers (Figure 1), of note are Bladder, Breast, Lung cancers and Cervical, where APOBEC mutation has been well studied (7). In 2013, a computational study by Alexandrov et al. found 21 mutational signatures

across all cancers, they state that 2 such signatures describe APOBEC mutagenesis; signatures 2 and 13 (Figure 2) (15).

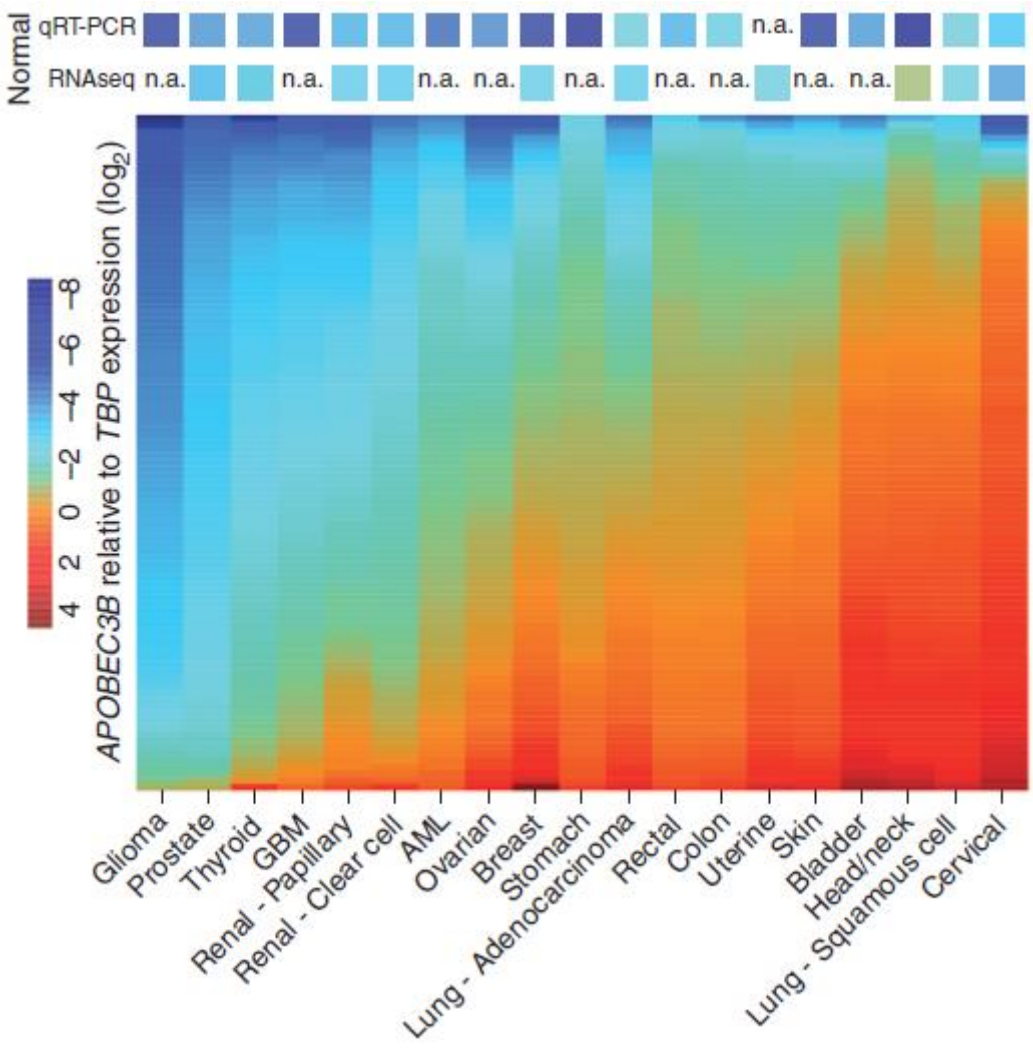

Figure 1. Heatmap of APOBEC3B Expression in Various Cancer Types.
Expression of A3B in various cancers, normalised to expression of the housekeeping gene: TBP. With high expression of A3B (orange/red) in a gradient with lower expression of A3B (blue), where $4\log_2$ TBP expression is red/high A3B. Normal tissue data is shown above each column in boxes and colour only describes expression (8). Expression data for cancerous tissues is shown below as a gradient depicting all data from the original source (14). Raw data used to make this graph is from (14) where it was taken from the Cancer Genome Atlas Data Matrix. Expression data was mined from RNAseqV2 and Reads Per Kilobase of transcript per Million mapped reads (RPKM) values from RNAseq.

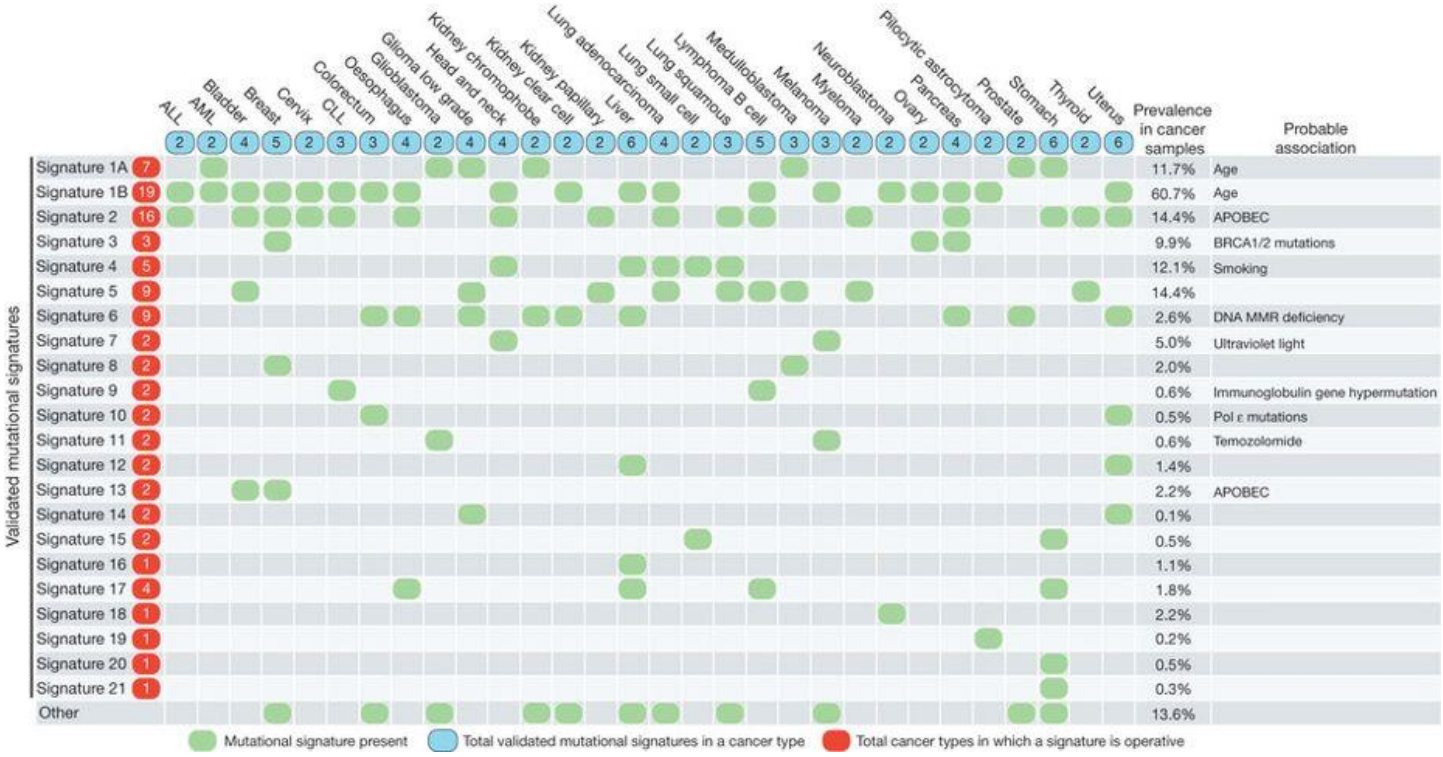

Figure 2. Table of Mutational Signatures (Rows) and the Cancer Types they are found in (Columns).

Alexandrov et al. cover data computational sequence data from (NUMBER) of cancer samples, they find the recurrent signatures shown in the graph. Each signature has a probably association, this is surmised from the signature mutation. Signatures 2 and 13 describe mutations by APOBEC, these signatures appear in 16 of the cancer types studied by Alexandrov et al., with both signatures prevalent in 14.6% of cancers studied. (15)

Due to some disagreement about which A3 was causing most of the APOBEC signature seen in cancers (9, 7, 10, 16, 17) Silvas et al. investigated sequence specificity of A3A, as it has the highest catalytic activity of the A3s (16). They found that A3A seemed to have a strong affinity to (T/C)TC(A/G) sites, on top of this they also found that A3A has a preference for hairpin-loops. The affinity between the bases at -2 and +1 (T/C and A/G) led Silvas et al. to suggest that intra-DNA interactions of ssDNA have a defining role in the specificity of A3A. They find that substrates folded into a U-shape have a much stronger affinity for A3As active site (16, 17), leading to a higher catalytic activity when the interaction between -2 and +1 was present. In addition, it was discovered that A3A bound to RNA hairpins similarly to that of DNA ones (though, with a much higher affinity for DNA) (16), which could

mean that A3A overexpression could do more damage than seen only in genomic sequencing. In fact, the hairpin-loop affinity fits in with the active site of A3A and A3B containing conserved U-shape binding for TC in the active site (17). Chan et al. compared signatures of APOBEC mutagenesis in multiple cancers finding that, of the TCW motif, TCA was a more commonly mutated target. When working in yeast, they find that A3A prefers YTCA (where Y is T/C) and A3B preferred RTCA (where R is A/G). Chan et al. next determined which APOBEC was mutating DNA more often, finding that A3A was contributing up to an 11-fold increase of mutations across bladder, breast, head and neck squamous cell, lung adenocarcinoma/squamous cell cancers. This is despite A3B mRNA being expressed at much higher levels than A3A, which could be due to A3As much higher enzyme activity, or possibly that A3A mRNA is more often translated into protein (10).

Cells overexpressing A3A were found to activate the DNA Damage Response (DDR) pathway, as Replication Protein-A (RPA) was found to be phosphorylated in these cells' DNA. In fact, cells expressing A3A also express γ-H2AX (18), suggesting A3A may mark DSBs in a similar way to AID (8, 9, 12). UNG was found to be required for this response to activate as no DSBs were found when an inactive form of the enzyme was transfected into the cells. When A3A is overexpressed cells accumulate in S-phase but do not become apoptotic, so A3A can induce DSBs which do not lead to cell death. When the catalytic domain of A3A is removed these responses are not activated, suggesting that it is DNA damage which leads to these responses (18).

APOBECs can use DSBs to gain access to long strands of ssDNA, especially during repair of said breaks. Roberts et al. found that APOBEC hijacks the DDR pathway

leading to hyper-mutations of TpC sites in the chromosomal DNA of yeast, results which have parallels in cancer sequencing (19). This event has subsequently been named kataegis (8, 9, 7). Signatures show that APOBEC may switch strands, with passages on the sense strand having stretches of C mutations followed by G mutations, this is suggestive of bidirectional ssDNA exposed at DSBs (7, 8, 9, 18, 19). Burns et al. (14) found that kataegic foci (two or more gDNA mutations within 10,000 bp of each other) were often associated with genomic rearrangement when A3B was over expressed in yeast. Whereas, Alexandrov et al. found that most rearrangement sites in human samples did not have kataegic foci, suggesting that there must be other factors involved (15). As well as causing DSBs A3s are thought to hijack stalled replication forks that are being repaired by BER, exposing single-stranded genomic DNA. This mechanism is predicted as DDR proteins are required for kataegis to take place. Not only can kataegis be induced by an IFN response to replication stress (i.e. stalled replication forks and DSBs) but it also causes replication stress through mass mutation, inducing an IFN response, increasing A3 expression thus creating a positive feedback loop (7, 8, 9) this has been linked to A3B overexpression (9).

1.6 APOBECs and Anti-Cancer Drugs

UBCs, exhibiting A3 mutational signature, often become resistant to anti-cancer drugs. DNA damage from anti-cancer drugs causes replication stress which, in turn, may induce A3B expression in bladder cancer cells starting the positive feedback

loop of A3 expression (20). Middlebrooks et al. observed a large increase in A3B expression in bladder cancer cells that were exposed to bleomycin. They believe this overexpression may be due to a Single-Nucleotide Polymorphism (SNP) rs1014971 acting as an A3B enhancer found 20kb upstream of A3A. SNP rs1014971 is strongly associated with bladder cancer and smoking. Middlebrooks et al. found that bladder cancer cells heterozygous for rs1014971 show binding of promoter probes to the T-risk allele, in contrast breast cancer lines showed association to both the T-risk allele and the S-non-risk allele. They also find that an increased mutation rate by APOBECs in bladder cancer is linked to longer survival, with the double S-risk rs1014971 SNP showing better survival rates to that of heterozygous cancers (20). As an overly increased mutational burden leads to cell death due to excessive genomic DNA damage, it could be possible that an APOBEC driven cancer can kill itself through overexpression. To get around this, it could be that mutation happens in several bursts as and when needed, with over and under expressing subclones dying off. These bursts could possibly be in response to viral or anti-cancer drug presence, with subclones that express a median amount of APOBEC being selected for (7, 8, 9).

High heterogeneity in UBC tumours can allow for selection of driver genes in some clones, for example; Roper et al. find that a patient had metastases with only activating HRAS mutations, whereas the primary tumour did not have this gene mutated at all (5). When comparing treated UBC to that of untreated there is a clear difference in evolution and which subclones were selected for, the strong selection

pressure of an anti-cancer drug seems to lead towards an increase of APOBEC related mutations, this is without mutational load increasing (2, 5, 6, 4). Oddly enough it was found that even driver mutations may change after chemotherapy; with some samples even having different mutations in the same driver genes post-therapy. These changes would suggest that a sub-clone resistant to chemotherapy survived and grew to later represent the majority of the cancer. This shows heterogeneity within tumour populations, as there must already be subclones with differing driver mutations before selection via therapy (2).

It could be that the introduction of replication stress induced by anti-cancer drugs leads to an increase of APOBEC expression, with stalled replication forks and DSBs formed by the drugs. Treatment by chemotherapeutic agents could provide ample ssDNA substrate for APOBEC to cause kataegis, leading to novel mutations in many subclones during chemotherapy (2, 10).

The wide variety of subclones in a heterogeneous population may harbour cells with the necessary mutations leading to APOBEC overexpression, producing resistance via rapid evolution during chemotherapy or selection of a subclone already resistant to the drug that has come about by previous APOBEC mutation. If it is the previous, then many clones may be wiped out due to overexpression or underexpression of APOBEC, leaving only those with just the right amount of APOBEC expression (the 'Goldilocks zone'), with survivors that have adapted to the treatment through rapid evolution (5). Though it seems likely that there is probably a mixture of the right amount of mutation beforehand alongside ability to express APOBECs at the right level during chemotherapy.

Indeed this 'Goldilocks zone' appears to require the right timing too, as there is evidence of APOBEC mutagenesis being episodic in nature (2, 4, 6), possibly occurring early in tumourogenesis and then reoccurring at other times leading to large scale mutation and cancer evolution, though there is no evidence as to what causes the increase in APOBEC expression as of yet (21).

Despite there being no increase of mutational burden in post chemotherapy UBC cells, post treatment heterogeneity seems to predict a worse overall survival (6).
More recently a mixture of chemotherapy and immunotherapy has been tested, it was thought that a two-pronged attack would help to stop tumours reoccurring at a later date. Though it appears, in cancers expressing the APOBEC mutational signature (such as UBC), tumours are able to avoid both therapies at the same time (4). Immunotherapy targets neoantigens, these are altered antigens expressed only in cancers, they are associated with an increased mutational burden and can be targeted by immune surveillance (18). It is thought that the mechanism which causes A3s to mutate retroviruses like; Human ImmunoDeficiency Virus (HIV) and Human PapillomaVirus (HPV) allowing them to avoid the immune system, may occur in a similar fashion for cancers. The increased replication stress caused by A3s lead to chromosomal instability as well as Copy-Number Alterations (CNAs) which could allow for increased tolerance of genomic instability. Replication stress brought about by anti-cancer drugs seems to enhance stress brought on by A3s, leading to genomic instability. This instability allows for CNAs and rearrangements, not

increasing mutational burden, and this creates new neoantigens, rendering immunotherapy ineffective (4, 6).

Subclones resistant to chemotherapy may also have differing neoantigens, leading to this subclone to become the main population of cancer cells post-treatment, mutational burden may not increase as no cancer evolution is needed to create new neoantigens. It has been observed that there is a loss of neoantigens after dual therapy has been carried out (4), suggesting that dual therapy of APOBEC cancers is not an effective treatment strategy.

It has been suggested that A3 based cancers might need to be treated with chemotherapy alongside an agent which reduces expression of A3s. This strategy could reduce the cancers ability to evolve and possibly survive chemotherapy. A mechanism suggested by Law et al. was to use shRNAs to knockdown A3B expression in tandem with chemotherapy. After testing this, as they found that the knockdown of A3B in breast cancer cells slowed down resistance gain to tamoxifen in vitro (22).

1.7 Aims of this Project

It seems that A3s play a large part in UBCs ability to avoid and resist chemotherapy, though the question remains, which A3 is it? The consensus seems split, with evidence for A3B being the main culprit (6, 9, 15) as well as some for A3A (2, 13, 16, 17). Both sides of the argument have strong points: A3B being upregulated in so

many cancers (7) and showing a correlation between its expression and mutational signature (7). A3A has the higher catalytic ability but has a much lower mRNA expression, although its mutational signature appears to be higher in sequenced cancers (9, 13, 16).

The question becomes: Would the downregulation of A3A and/or A3B in tandem with traditional chemotherapy drugs kill cancer and stop recurrence? The knockdown of A3A and/or A3B would theoretically stop cancer cells' ability to hypermutate, evolve and evade anti-cancer drugs. To test this theory, we first need to create cancer cell lines with these genes knocked out and afterwards, try to make them resistant to traditional anti-cancer drugs. The high APOBEC signature of UBC (Figures 1 and 2), alongside high recurrence and resistance, makes it an ideal model. The aim of this study is to create A3A, A3B and A3A/A3B double knockouts for future study of UBC drug resistance. The creation of these tools will require use the of CRISPR/Cas9 to insert a targeting vector capable of knocking out gene expression (Figures 3 and 4) into the first exon of both genes. After integration of the targeting vector Single-Cell Cloning (SCC) will be used to isolate cells that are homozygous or heterozygous for the deletion of A3A and/or A3B. Polymerase Chain Reaction (PCR) will be used throughout to validate integration and subsequent removal of the targeting vector. Finally, Reverse Transcriptase Quantitative PCR (RT-qPCR) will confirm knockout of both A3’s mRNA expression.

2 Materials and methods

2.1 Cell Culture

BFTC-905, T24 and 5637 Urothelial Bladder Cancer cells (Provided by Martin Michaelis) were maintained in Iscove's Modified Dulbecco's Medium (IMDM) (PanBiotech) with 10% Fetal Bovine Serum (FBS) Supreme (PanBiotech) and 1% Penicillin/Streptomycin (ThermoFisher). Cells were cultured at 37°C in a 5% CO_2 static incubator. Cells were grown to >80% confluency before passaging, freezing down or harvesting.

Passaging protocol: 1ml of 0.05% Trypsin-EDTA (PanBiotech) was added to a T25, 3-5ml to T75s, cells were replaced in the incubator for 5 minutes. After incubation cells were agitated until resuspended, or placed back into the incubator, this process was repeated until most cells were detached. Once cells were in suspension, IMDM was added making it up to 5ml for T25s and 10ml for T75s. Cells were then spun down at 300xG for 5 minutes, supernatant was removed and cells resuspended in 10ml of IMDM. 1ml of this cell suspension was added to the new flask, with IMDM added to make the flasks up to 5ml (T25) and 10ml (T75).

Cell storage: Cells were stored in freezing medium: antibiotic-free IMDM containing 10% Dimethyl Sulfoxide (DMSO) (Sigma), in 2ml Cryotubes. Cells were first trypsinised and spun down and resuspended in the described freezing medium, where each cryotube contained 1ml of cell suspension at $1x10^6$ cells/ml.

2.2 DNA and RNA Extraction

DNA and RNA were extracted from cultured cells at various stages after being tyrpsinised and frozen immediately at -80°C in dry ice, then transferred to a -80°C freezer. Extractions were performed using; AllPrep DNA/RNA/Protein extraction kit (Qiagen), DNA mini kit (Qiagen) and RNeasy mini kit (Qiagen). The included protocols were used in tandem with QiaShredder (Qiagen) for cell homogenisation. DNA and RNA were then quantified using a NanoDrop (ThermoFisher).

2.3 Cell Genotyping

PCR was performed using 5µl of HiFi HotStart ReadyMix Polymerase (KAPA Biosystems) 1µl of each primer (10µM) and 100ng of DNA to a total volume of 10µl inside a 0.2ml PCR tube.
Primers for A3A were: Forward; TGAGCTCACACCAGAACCAC
Reverse; TAGAGCCCAGAGAAGGTCCC
A3B: Forward; AAACAACCACTGGGAGTCCG
Reverse; GTCCAACTCTGGCCTTTCCA
PCR was performed with conditions: 95°C for 3 minutes → 98°C for 20 seconds → 68°C for 15 seconds → 72°C for 3 minutes returning to the second step for 30 cycles finishing at 72°C for 5 minutes.
Samples were then run on Ethidium Bromide Agarose gels at 100 Volts for 45 minutes alongside a 1KB Plus DNA Ladder (Invitrogen). Expected product lengths are described in Figures 3 and 4.

2.4 Targeting Vectors

Targeting vectors (Figures 3 and 4) for the knock-out of both A3A and A3B were provided and designed by Tim Fenton. They contain 1kb 5' and 3' homology arms to direct them to the appropriate site and allow for Homology Directed Repair (HDR) to incorporate the targeting vector at the target site (1st Exon). The gene knockout is achieved through a translation and transcription termination site added at the end of Exon 1, in both A3A and A3B, made up of stop codons placed in all 3 reading frames and bovine growth hormone polyAdenylation signal (bpA), a transcription terminator. Flippase Recognition Targets (FRTs) flank a PhosphoGlycerate Kinase (PGK) promoter-driven Herpes Simplex Virus Thymidine Kinase Gene (HSV-TK) dual selection cassette which encodes for puromycin resistance and ganciclovir sensitivity.

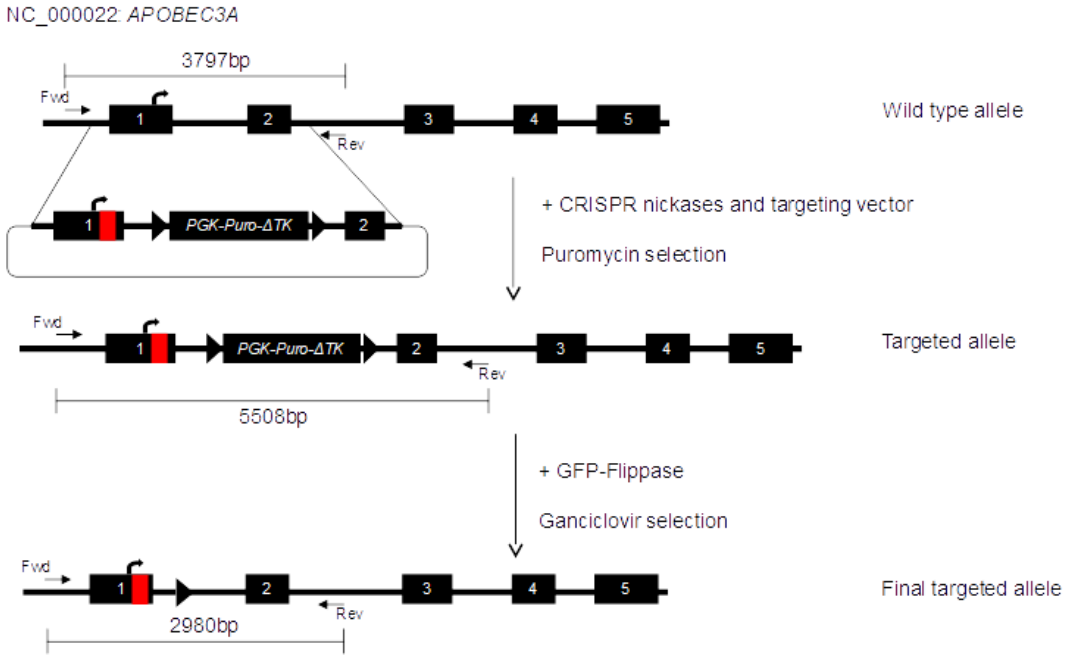

Figure 3. A3A Targeting Vector.
This figure describes the targeting vector (1711bp) for A3A, as well as its incorporation into the locus. This vector inserts a bpA site towards then end of the first exon, stopping translation, a PGK driven puro-HSV-TK dual selection cassette inserted between exons 1 and 2. This cassette is flanked by FRT sites. Insertion of the vector increases PCR product for A3A primers (section 2.3) from 3793bp to 5508bp, when targeted by flippase the PCR product is then reduced to 2980bp. All bp lengths are predicted with SnapGene. (Unpublished, Tim Fenton)

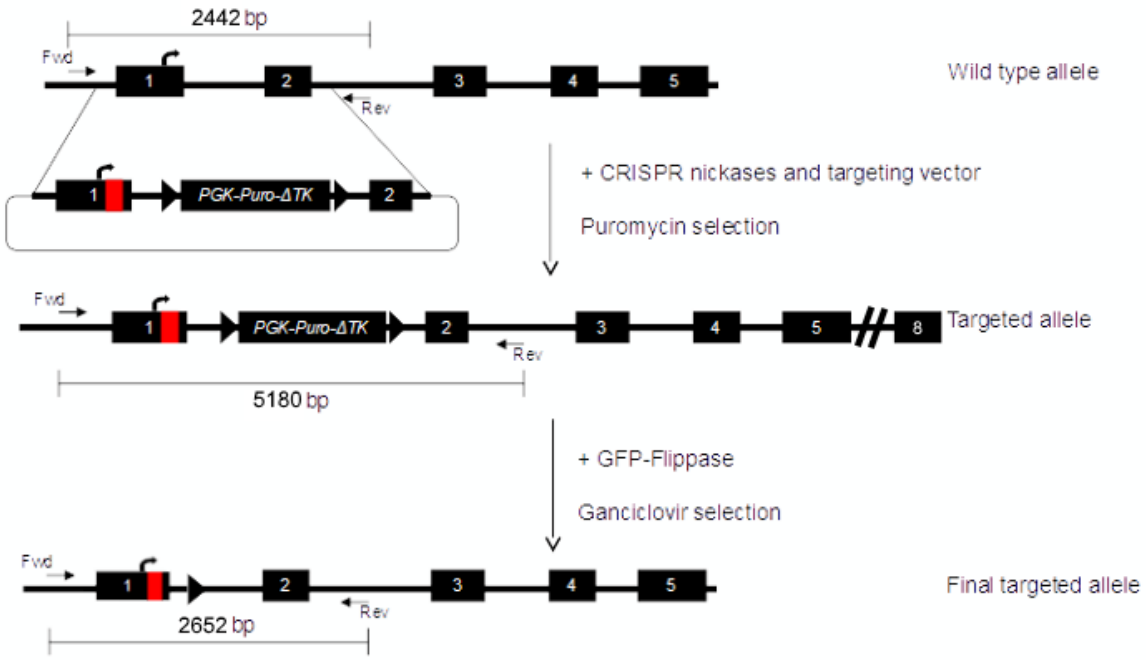

Figure 4. A3B Targeting Vector.
This figure describes the targeting vector (2738bp) for A3A, as well as its incorporation into the locus. This vector inserts a bpA site towards then end of the first exon, stopping translation, a PGK driven puro-HSV-TK dual selection cassette inserted between exons 1 and 2. This cassette is flanked by FRT sites. Insertion of the vector increases PCR product for A3A primers (section 2.3) from 2442bp to 5180bp, when targeted by flippase the PCR product is then reduced to 2652bp. All bp lengths are predicted with SnapGene. (Unpublished, Tim Fenton)

2.5 Plasmids

This experiment uses 7 plasmids; 2 guide RNA (gRNA) plasmids (Figure 5) for each targeting vector, 4 in total. These plasmids contain code for transient expression of mutant Cas9 enzyme which cuts only one strand of DNA, the plasmids also encode a gRNA, which is recruited by the Cas9 enzyme, guiding it to the target site. These gRNAs target either upstream (left) or downstream (right) of the target site (Figure NewC). Combination of both gRNAs creates a nickase pair in the transfected cell and thus cuts either side of the target site (around 10bp apart from each other) allowing for HDR.

gRNA Sequences:

A3A N-terminus left sense RNA: CACCGCTTGCGACTTGCTCAAGGCG

A3A N-terminus right sense RNA: AAACCGCCTTGAGCAAGTCGCAAGC

A3B N-terminus left sense RNA: CACCGCACAGACCAGGAACCGAGAA

A3B N-terminus right sense RNA: AAACTTCTCGGTTCCTGGTCTGTGC

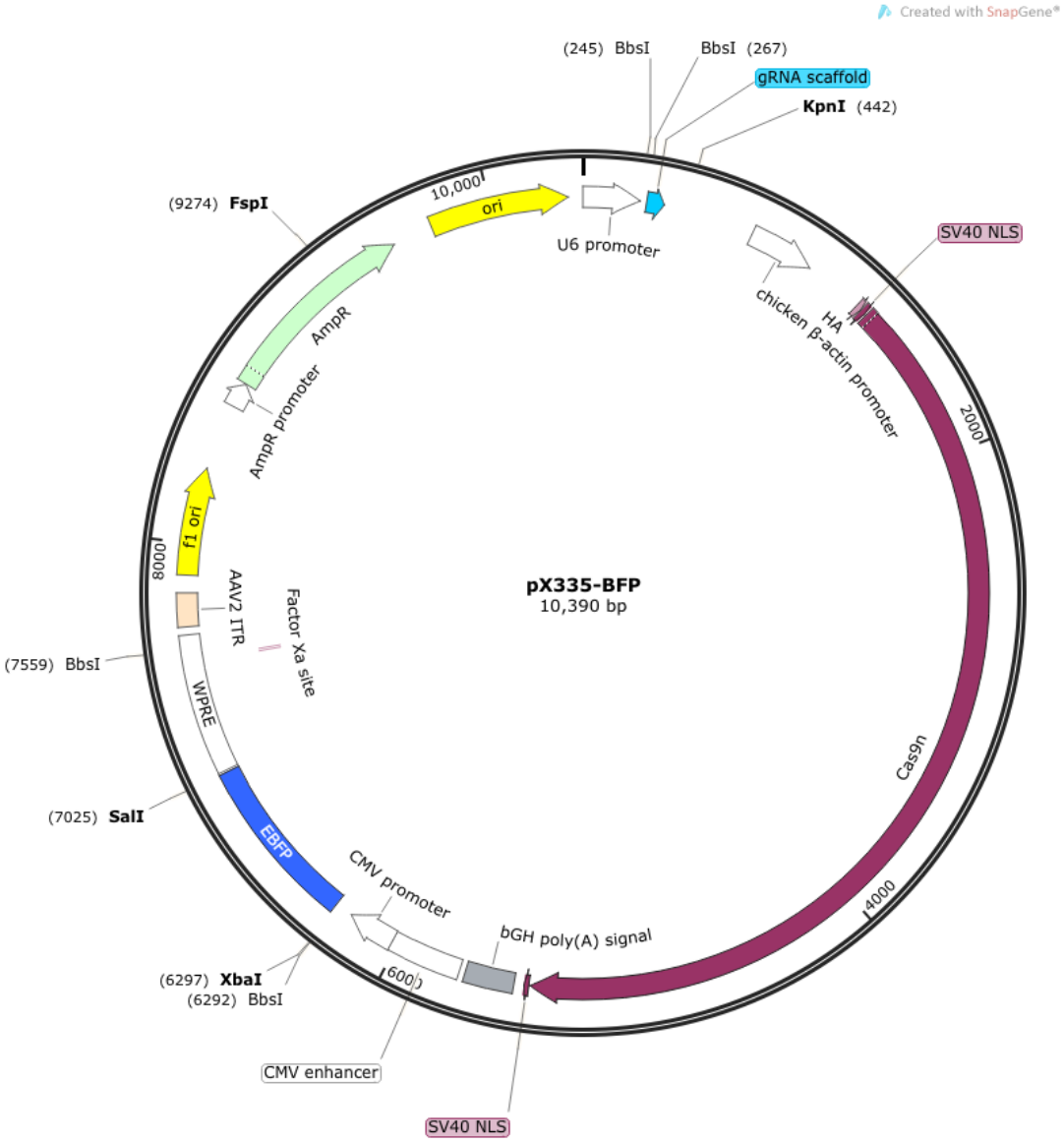

Figure 5. Guide RNA Plasmid.
This figure describes the plasmid containing the 'left' gRNA for either transfection alongside A3A or A3B targeting vectors (Figures 3 and 4). The gRNA is represented with a blue arrow and the fluorescent protein (reporter) is BFP (blue block), in the 'right' gRNA this would be GFP. An Ampicillin resistance gene is also incorporated into this plasmid to allow for selection when grown up in bacteria. DNA is then extracted from a resistant colony of these bacteria for transfection alongside targeting vectors. (Tim Fenton, Unpublished)

Each targeting vector (Figures 3 and 4) was contained in a plasmid as shown for A3B in Figure 6.

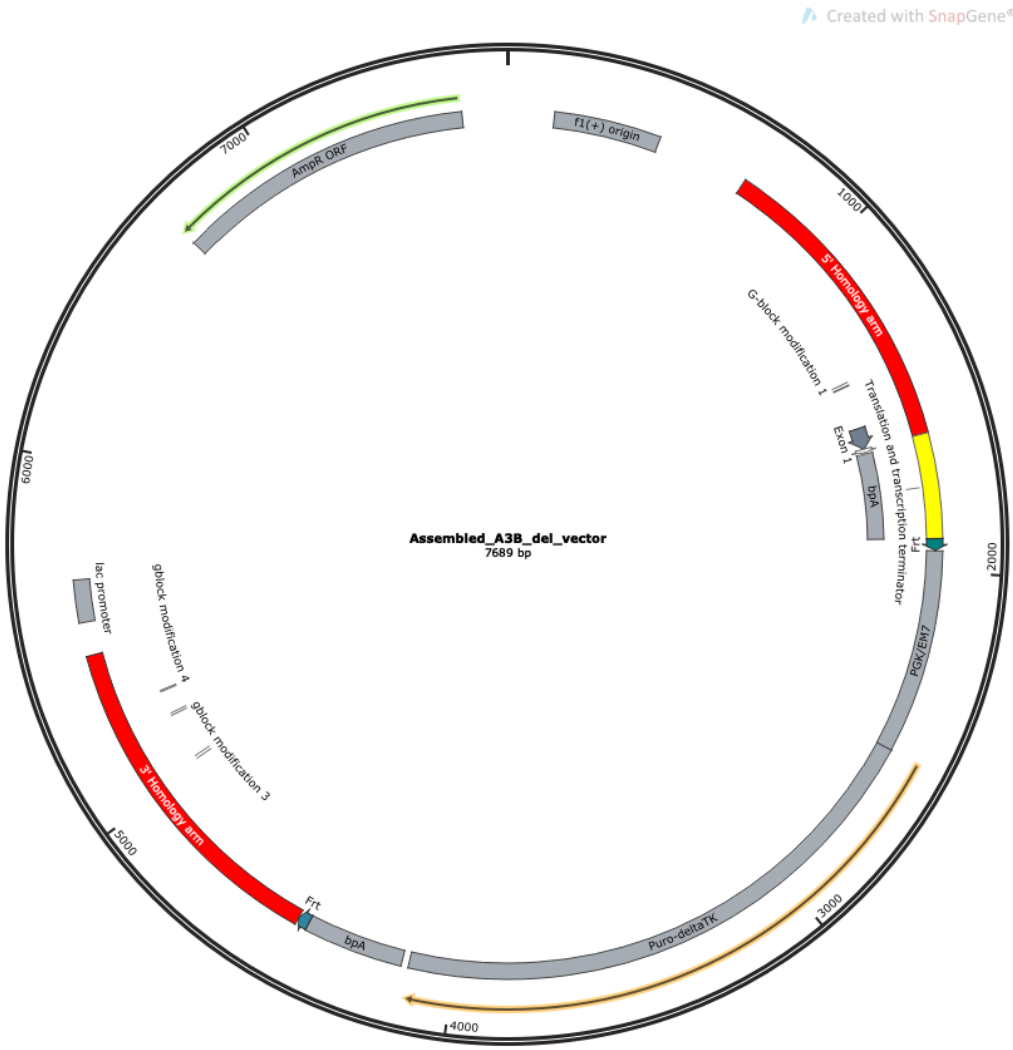

Figure 6 A3B Targeting Vector Plasmid
This figure describes the plasmid containing the targeting vector for A3B deletion, housing 5' and 3' homology arms, bpA and a PGK driven puro-delta-TK dual selection cassette flanked by FRT sites. In addition to the vector, this plasmid contains an ampicillin resistance gene for amplification in competent bacteria. This plasmid was provided by Tim Fenton.

A 7th plasmid encodes for transient expression of flippase, an enzyme which cuts at FRT, sites which are incorporated by the A3A and A3B targeting vectors (Figures 3 and 4).

All plasmids contain an Ampicillin resistance gene for amplification via bacterial vectors. All plasmids were provided by Tim Fenton.

2.6 Picking colonies for Mini and Midi Preps

TOPO cloning was performed as per the protocol described in section 2.13. After transformation competent bacteria was then streaked on agar plates containing Ampicillin. Targeting vector plasmids for A3A and A3B were retransformed and streaked on Ampicillin containing plates by Tim Fenton.

After overnight growth (at 37°C, static incubation) a bacterial colony was picked using a pipette tip and deposited in 5ml and 50ml Ampicillin containing LB for TOPO and targeting vectors respectively. After growth in a shaking incubator at 37°C at 250rpm for 24 hours. Samples were then spun down at 13,000 rpm for 10 minutes and extracted following Qiagens Mini and Midi plasmid prepping protocols using these kits.

2.7 Transfection

0.08×10^6 cells were plated in a 6-well plate 24 hours pre-transfection, transfection was commenced once confluency was 80% or higher (but not 100%). BFTC-905, T24 and 5637 cells were each transfected using Lipofectamine 3000 with 2 gRNA plasmids each plus the targeting vector: A3A and A3B (Figures 3 and 4). The manufacturer's protocol was followed with amount of targeting plasmid and gRNA plasmids added (µg) described in Table 1.

	A3A Targetting Plasmid	A3A Forward gRNA Plasmid	A3A Reverse gRNA Plasmid	A3B Targetting Plasmid	A3B Forward gRNA Plasmid	A3B Reverse gRNA Plasmid
A3A Knockout Line	1.25	2.5	2.5			
A3B Knockout Line				1.25	2.5	2.5
A3A/A3B Double Knockout Line	0.75	1.25	1.25	0.75	1.25	1.25

Table 1. Visual representation of amount of plasmid added per transfection (µg)
The top row describes the plasmid added, with the left-most column describing which A3 was targeted. Total amount of plasmid DNA added per transfection is unchanged, even when using 2 targeting vector plasmids and accompanying gRNA plasmids.

A second round of transfections were performed with a flippase plasmid using Lipofectamine 3000 for the majority of samples, 2 samples; BFTC-B3 and BFTC-Both3 were transfected with FuGene HD (Promega). FuGene HD led to a longer recovery time with a lower observed ratio of green cells.

Transfections were performed in a 6-well plate in 2.5ml antibiotic-free IMDM, Lipofectamine 3000 reagents were mixed with DNA in Opti-MEM and incubated for 15 minutes before being added to wells. Green Fluorescence Protein (GFP) fluorescence was observed under a Lumascope after 48 hours and the media was changed. Samples were placed under selection (puromycin or ganciclovir) after 72 hours to allow for some recovery after transfection.

All transfections had no targeting vector controls (still transfected with gRNA plasmids) as well as no reagent controls.

2.8 Positive Selection

All transfections with the A3A and/or A3B knockout vector cells, alongside their controls, were originally treated with 750 ng/ml of puromycin in T24 and 5637 cell lines. This was later increased to 1500 ng/ml for BFTC-905, T24 and 5637 cell lines as control cells were not dying within 2 weeks of the start of selection. Cells were kept in selection media until all control cells were dead, the criteria for this being: no cells seen attached to the surface of the plate for 2 days.

2.9 cDNA Synthesis

Genomic RNA was extracted from cell lines as described in section 2.2. cDNA was made using the following reagents: GoScript Reverse transfection system (Promega) for parental lines, and LunaScript RT SuperMix Kit (New England Biolabs) for single-cell clones, manufacturer's protocols were followed for each reagent set. 1µg of RNA was used alongside random and oligo(dT) primers. The reaction volume was made up to 20µl (topped up with nuclease free water) in 0.2ml PCR tubes. GoScript samples first had RNA and primers mixed (5µl) and placed in a 70°C heat block for 5 minutes, on ice for 5 minutes after, this mix was then combined with GoScript reagents and run in a thermocycler at 25°C for 5 minutes 42°C for 1 hour and then 70°C for 15 minutes. Lunascript samples were run in a thermocycler at 25°C for 2 minutes, 55°C for 10 minutes then 95°C for one minute.

2.10 RT-qPCR

RT-qPCR was performed using SYBR green reagent (ThermoFisher), according to the manufacturers protocol, using cDNA produced as per section 2.9. Samples were run in an unskirted 96-well PCR plate, 0.5µl of forward and reverse primers (0.5µM,

diluted in nuclease free water) were added with 26.25µl reaction mix, to which 100ng of cDNA was added, each reaction was made up to 50µl with nuclease free water. The qPCR primers targeted A3A and A3B (exon 5 and 6???????) mRNA, with TATA-Binding Protein (TBP) and GlycerAldehyde 3-Phosphate DeHydrogenase (GAPDH) HouseKeeping Gene primers running in tandem per sample for use as references for relative expression. Non-template controls were run in duplicate per primer set per plate. Samples were run using QuantStudio 3 (ThermoFisher) using QuantStudio Design & Analysis Software.
Calibration curves for each primer were provided by Nerissa Kirkwood.
CT values for each repeat (2 per primer per sample)
Analysis of raw data was performed in Excel (Microsoft) using the equation:
10^((CT – Primer Calibration Curve Intercept) / Calibration Curve Gradient) produced the copy-number of each replicate including housekeeping genes. For a comparison of expression in relation to housekeeping genes this value was then divided by that for the housekeeping gene per replicate per sample.

2.11 Single-Cell Cloning (SCC)

After genotyping transfected cells, confirming some had either the A3A and/or A3B knockout insert, they were serially diluted in a 96-well cell culture plate. Confluent cells were first trypsinised (Section 2.1), once in suspension they were then diluted to a total of 10ml of trypsin/IMDM mixture, 100µl was placed into well A1 of a 96-well cell culture plate. This well was then made up to 200µl by adding more IMDM, 100µl was then removed and added in a serial dilution down the plate to well H1. Wells were then filled to 200µl before further dilution across the plate.

Cells were checked daily, with wells containing only one visual cell/colony being marked and kept track of.

Once cells had grown to confluency in the 96-well plate they were moved up to a T25 cell-culture flask and then a T75 cell-culture flask before being frozen down and harvested for DNA and RNA extraction before being genotyped.

2.12 Cell Viability Assay for Ganciclovir Based Negative Selection

Death curves for transfected cells compared to parental cells in the presence of ganciclovir dissolved in DMSO were made using a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay by incubating cells grown in ganciclovir for 5 days with 20µl of Cell Titer 96 AQueous One Solution Cell Proliferation Assay (Promega) for 4 hours before reading absorbance (490nm). Cell viability curves were made using the formula:

(Absorbance of Well – Average Negative Control Absorbance) / (Average Positive Control Absorbance – Average Negative Control Absorbance) was done for all 3 replicates at each concentration of ganciclovir these results were then averaged and multiplied by 100 for % cell viability.

Cells were plated in a 96-well cell culture plate at 10,000 cells/well 2 hours before (allowing for cells to attach to the plate) normal media was removed and replaced with ganciclovir containing media. 5 concentrations of ganciclovir were made starting at 50µM, this was then serially diluted 1:3 7 times.

SCCs were transfected with flippase, as described in section 2.7, after which they were grown in 20µM ganciclovir until control transfections had no viable cells visible under a microscope for 2 days.

2.13 TOPO Cloning

Zero Blunt™ TOPO™ PCR Cloning Kit, with PCR™-Blunt II-TOPO™ Vector (ThermoFisher) and One Shot™ TOP10 Chemically Competent *E. coli* (ThermoFisher) were used as per the recommended protocol to clone the PCR products from Section 2.3. DNA was then purified using QIAprep Spin Miniprep kit (Qiagen) as per suggested protocol. A small amount (1µl) of cleaned PCR product was run on a 1% agarose gel to confirm correct product was recovered, 4µl of cleaned PCR product was combined with 1µl salt solution and 1µl pCR II-Blunt-TOPO, leading to a final reaction volume of 6µl, this reaction mix was incubated at room temperature for 15 minutes and transformed into competent cells as per section 2.6. Transformed and amplified competent cells then had the plasmids purified as per 2.6 and were prepped for sequencing. Sequencing was performed by Eurofins Genomics using SupremeRun Tube with primers: SP6 and T7.

Analysis of the sequence was done using SnapGene Viewer with reference sequences of expected targeting vector incorporation provided by Tim Fenton.

3 Results

3.1 Insertion of Deletion Vector

3.1.1 Knockout Strategy and Targeting Vector

The knockout strategy is described by Figure 7 and is further outlined below:
Both deletion targeting vectors contain homology arms of around 1kb 5' and 3' of the gRNA target site around exon 1 of A3A or A3B respectively (Figure 5). gRNA plasmids transfected into cells transiently express Cas9 D10A mutant, as are the gRNAs, which bind to the mutant Cas9 and direct them to the target site (Figure 7a/b). These Cas9 mutants can only cut one strand of DNA and so offset nicks are created around the target site when both gRNA plasmids are transfected. The gRNA plasmids also transiently express either Blue Fluorescent Protein (BFP) or GFP, depending on which gRNA was present, either binding upstream (left) of the nick site with BFP (Figure 5), or downstream (right) of the nick site with GFP (Provided by Tim Fenton). HDR is then exploited by the added transfection of the targeting vector plasmid, as the DNA break is repaired with reference to the targeting vector (Figure 7a to b). The long 1kb homology arms direct the targeting vector to the transfection site allowing for DNA repair machinery to incorporate the targeting vector to the locus, adding stop codons in all 3 reading frames at the end of exon 1, followed by a bpA site to terminate transcription (Figure 7b). The A3A targeting vector shortens the A3A gene further by removal of a repeat region at the end of first intron (Figure 3). Through incorporation of the targeting vector via HDR the dual selection cassette, driven by a PGK promoter-driven dual selection cassette, allows for selection of successfully transfected cells (Figure 7c to d). Selection is conferred; positively through resistance to puromycin and negatively through sensitivity to ganciclovir. Cells are first selected with puromycin (Figure 7d) then go through a single cell cloning process (Figure 7e), this allows for picking a clone that is either homozygously or heterozygously targeted (Figure 7f). The dual-selection cassette is flanked by FRT sites (Figure 6), transfection of positively selected cells with flippase will remove the cassette but, leave behind a single FRT site and the gene

transcription and translation terminators keeping knock-out genotype (Figure 7f). Cells untransfected with flippase will still contain the dual-selection cassette and so, will continue to be sensitive to ganciclovir exposure. Sensitivity is conferred by as the HSV-TK gene as it phosphorylates ganciclovir, creating a toxic metabolite which is incorporated into elongating DNA causing replication termination and cell death (Figure 7g).

This strategy leads to complete deletion of the target genes, as the stop codons halt transcription, and removal of the selection cassette leads to minimal disruption of the gene. Loss of gene expression should be measurable with RT-qPCR, this method is used instead of measuring protein expression due to a lack of specific antibodies for A3A and A3B.

With the cassette inserted the gene length of each APOBEC should increase by 1711bp and 2738bp for A3A and A3B respectively. This insertion should be easily validated by the genotyping assay described in section 2.3. After transient transfection with flippase, the gene length with the same primers should be 2980bp and 2652 for A3A and A3B respectively. This difference in gene size between insertion of gene, selection of pools and flippase activity allows for analysis of single-cell clones by genotyping with the same primers, with A3As expected final size being 817bp shorter than the WT (due to removal of a repeat region) it can be validated on an agarose gel.

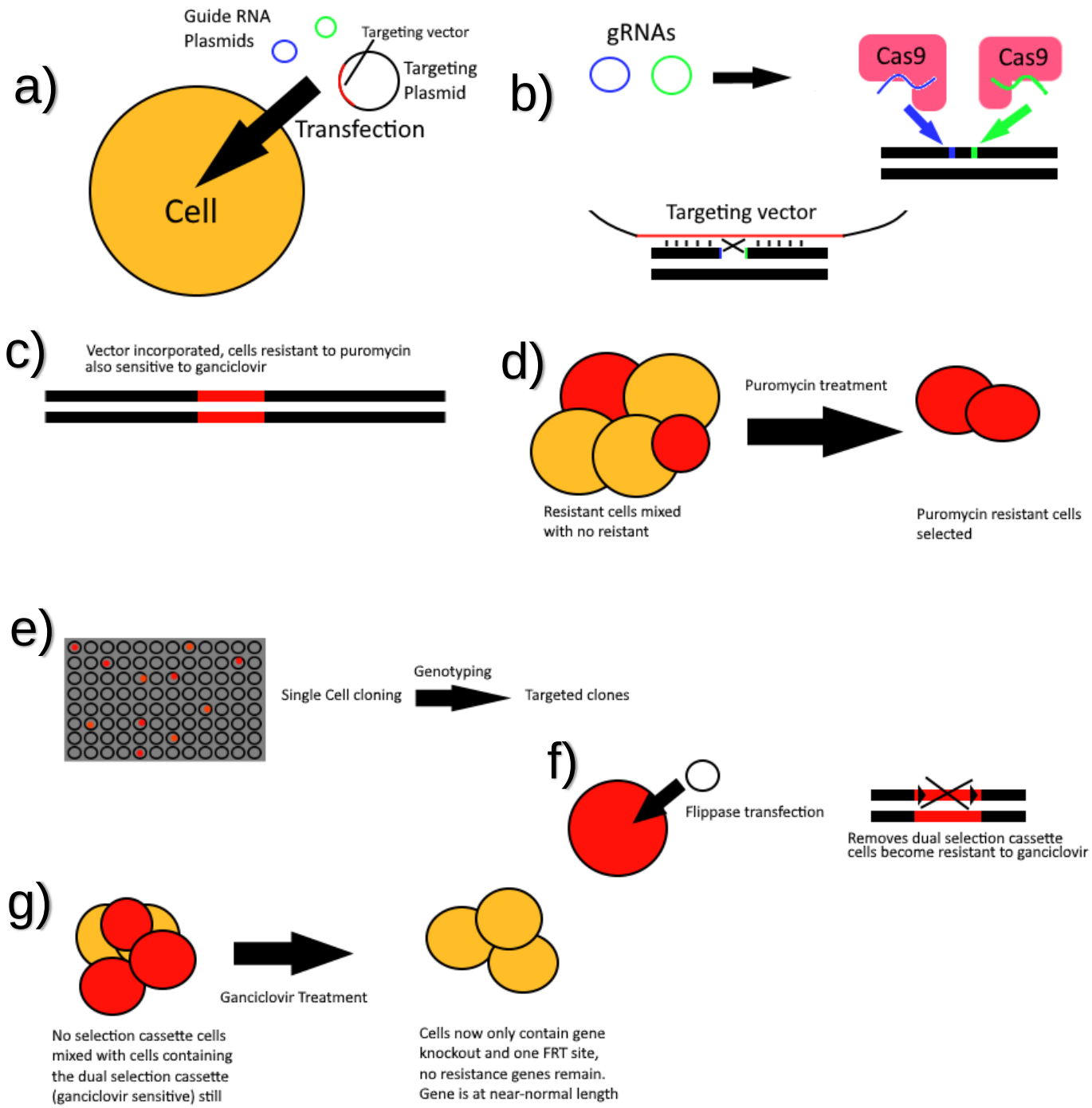

Figure 7 Knockout Strategy

a) UBC cells are initially transfected with A3A, A3B or both sets of targeting plasmids and accompanying gRNA plasmids. b) gRNAs encode for transient expression of a mutant Cas9 enzyme, as long as a single gRNA. The mutant Cas9 only cuts a single strand, requiring a nickase pair to properly cut the target sequence. This allows the targeting vector plasmid to be used as a template in HDR of the resulting cut. c) Cells that incorporate the targeting vector now contain a dual selection cassette, conferring puromycin resistance and ganciclovir susceptibility. d) Puromycin treatment removes non-targeted cells. e) Cells go through single-cell cloning to capture a possible homozygous knockout. f) Cells homozygous or heterozygous for incorporation of the targeting vector are transfected with flippase, this removes the dual-selection cassette. g) Treatment of cells with ganciclovir removes any still containing the dual-selection cassette, the remaining population has a near-normal gene size.

3.1.2 Validation of Cell Transfection

5637, BFTC-905 and T24 cells were transfected with the A3A and/or A3B deletion insertion and were subsequently given the suffix A, B or Both to describe which transfection pools they were originally in (BFTC-905 was also shortened to 'BFTC') (Figure 4). The transfection was carried out using Lipofectamine 3000 as described in section 2.7. The gRNA plasmid allowed for identification of cells expressing gRNAs visually via the encoded GFP and BFP genes, though GFP fluorescence was only observed under a microscope to confirm successful transfection 48 hours post-transfection. All lines were treated with 1500 ng/ml puromycin until only cells which had been treated with gRNA plasmids, targeting vector plasmid and transfection reagent remained.

After puromycin selection, transfected cell pools were genotyped (Figures 8, 9 and 10) using the PCR method described in section 2.3. Genomic DNA from pools of puromycin-selected NIKS (a human keratinocyte cell line in which A3A and A3B had previously been targeted using the same vector and gRNA plasmids) are labelled as A3A Control and A3B Control, (provided by Tim Fenton) and display bands of the correct height for both the targeted and untargeted loci.

After transfection of 5637 (Figure) there were populations of cells that had the insertion for A3A in 5637-Both but there were no populations with bands for the A3B insertion in either 5637-B or 5637-Both. The T24 transfections (Figure 9) had populations that gained the A3A insert in T24-A only. Whereas the A3B insert appears in populations of both T24-B and T24-Both. Transfection of BFTC-905 lines was the most successful (Figure 10), with populations of BFTC-A and BFTC-Both showing bands for the insertion of the A3A knockout. Populations of BFTC-B and BFTC-Both contain the A3B insert also.

After transfection lines that show 2 or more populations of cells with either one insertion or more were selected for SCC (as described in Section 2.11). T24-A, T24-B, BFTC-A, BFTC-B and BFTC-Both were cloned due to the possibility of obtaining a homozygous clone of either (or both) deletions.

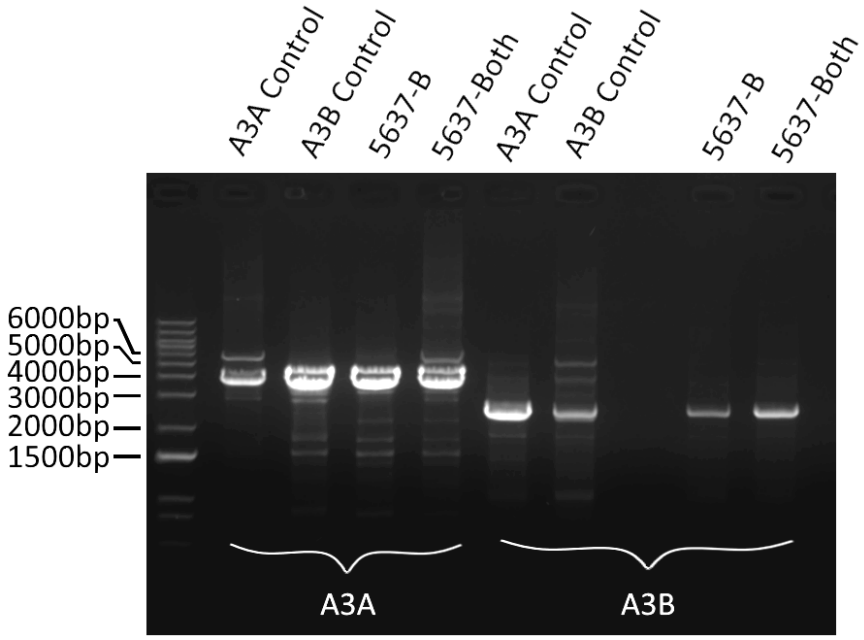

Figure 8. 5637 Pooled Transfection Populations.
1% agarose gel displaying PCR products from A3A and A3B (left, right) genotyping PCR reactions (Section 2.3). Bands appear at the expected height (Figures 3 and 4) with a few, less-bright, off-target bands also appearing. When comparing to controls, 5637-Both has a band at the height of the insert for the A3A knockout, but no samples show the A3B insertion.

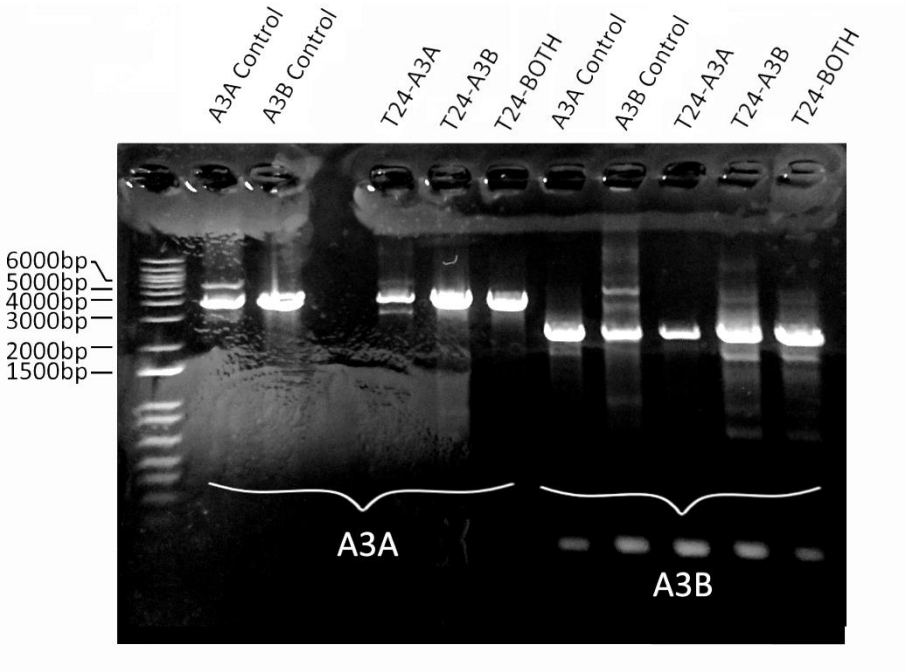

Figure 9. T24 Pooled Transfection Populations.
1% agarose gel displaying PCR products from A3A and A3B (left, right) genotyping PCR reactions (Section 2.3). Bands appear at the expected height (Figures 3 and 4) with a few, less-bright, off-target bands also appearing. When comparing to controls, T24-A has some incorporation of the A3A knockout vector but T24-Both does not. T24-B and T24-Both have some incorporation of the A3B knockout.

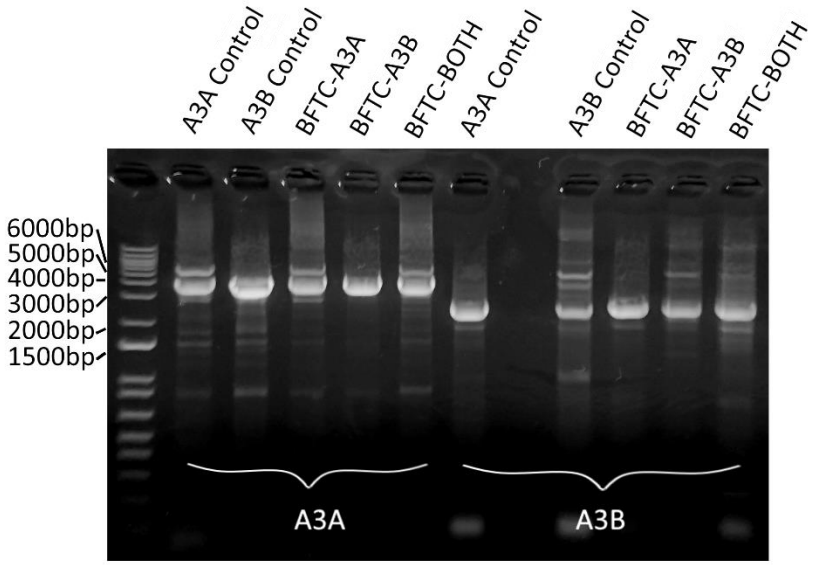

Figure 10. BFTC-905 Pooled Transfection Populations.
1% agarose gel displaying PCR products from A3A and A3B (left, right) genotyping PCR reactions (Section 2.3). Bands appear at the expected height (Figures 3 and 4) with a few, less-bright, off-target bands also appearing. When comparing to controls, samples BFTC-A and BFTC-Both show bands for the A3A insertion. Samples BFTC-B and BFTC-Both show bands for the A3B insertion.

3.2 Single-Cell Cloning

3.2.1 The Survivors

DNA from surviving SCCs was genotyped (Figure 11 and 12) and compared to NIKS insertion controls. This revealed 9 clones isolated from the original pools which were at least heterozygous for the insertion deletion. As all of the homo/heterozygous clones were originally from the BFTC-905 line they were grown up for flippase transfection, to excise the dual-selection cassette from the target vector locus, leaving behind a single FRT site and the transcription termination site at the end of the first exon.

Of the remaining SCCs; 3 clones appeared to be homozygous for the A3A insertion: BFTC-A2, BFTC-Both4 and BFTC-Both2.

2 more clones appeared heterozygous for the A3A insertion: BFTC-Both5 and BFTC-Both3.

And 2 clones appeared heterozygous for the A3B insertion: BFTC-B3 and BFTC-B5.

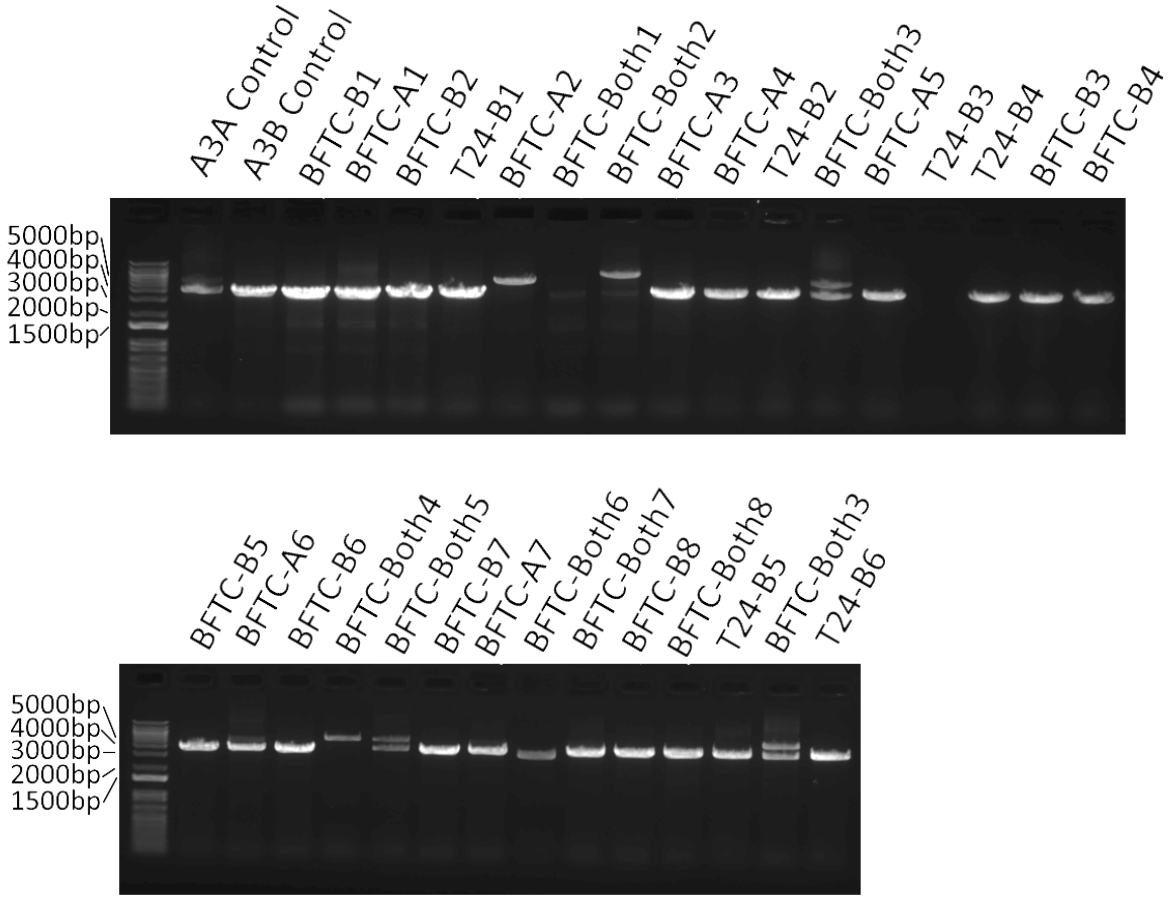

Figure 11. Single-Cell Cloning; A3A Genotype.
1% agarose gel displaying PCR products from A3A genotyping PCR reactions with DNA from SCCs genotyped. The higher band of 5508bp describes the deletion insertion allele, the lower 3797bp band describes the wild-type allele (Figure 3). After SCC 3 samples; BFTC-A2, BFTC-Both2 and BFTC-Both4 appear to be homozygous for the deletion insertion. 2 samples are heterozygous; BFTC-Both3 (repeated) and BFTC-Both5, with a faint, but unclear, band in BFTC-A6. Missing samples are covered in Supplementary data Figures S1 and S2.

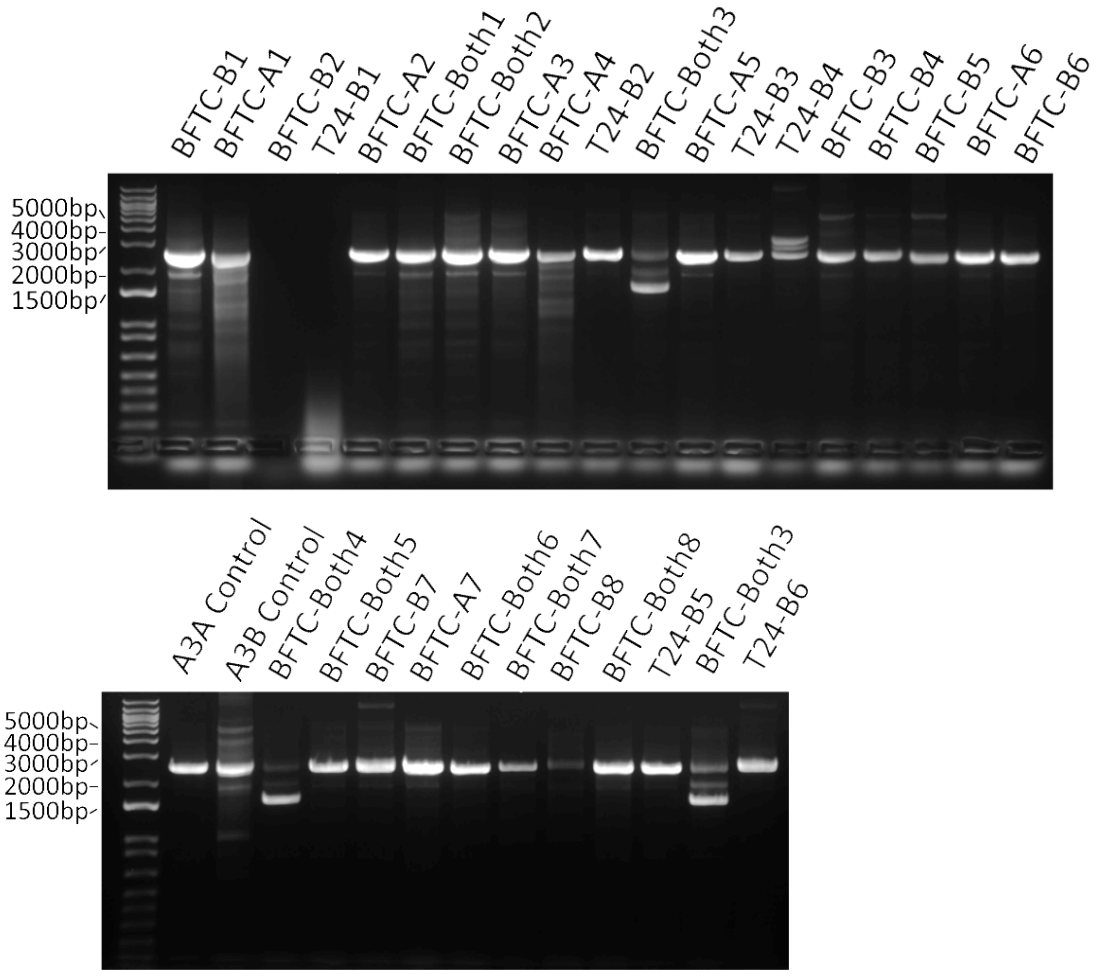

Figure 12. Single-Cell Cloning; A3B Genotype.
1% agarose gel displaying PCR products from A3B genotyping PCR reactions with DNA from SCCs genotyped. The high band of 5180bp describes the predicted insertion allele size (Figure 4) as compared to the A3B control. 2 samples show heterozygous bands: BFTC-B3 and BFTC-B5. There are several unpredicted bands such as BFTC-B7s large band around 9500bp, and the much lower and stronger bands of BFTC-Both3 and BFTC-Both4 at 1579bp. Missing samples are covered in Supplementary data Figures S1 and S2.

3.2.2 A3B Gene Truncation

An interesting note is that 2 samples: BFTC-Both4 and BFTC-Both3 presented a band far shorter in length for A3B PCR (around 800bp) than the original gene length of 2442bp (Figures 4 and 12). This would suggest that when transfecting with both the A3A and A3B targeting vectors and accompanying gRNAs, an event occurs leading to a possible loss of exon 1 from A3B in one allele. In comparison, transfection with one targeting vector and its gRNAs at a time this event is not seen.

The event in question could be a rearrangement, if both genes are essential (but able to cover for loss of the other), with A3B possibly being saved as it may be able to cover for loss of A3A. But it would appear this is not the case as the homozygous insertion of the A3A vector does not impact this genotype; BFTC-Both3 is heterozygous for the A3A insertion but still has the deletion in A3B. Heterozygosity would not suggest enough of a loss of A3A expression for a knock-out of this gene.

These PCR products were cloned into TOPO vectors as described in section 2.13 and sent off for sequencing. When comparing the sequence to that of the predicted A3B targeted locus and the WT A3B from the BFTC-905 parental line, a large 800bp deletion is seen, with 200bp upstream and 600bp downstream of the insertion site. This deletion covers the whole of Exon 1 continuing into 5' and 3' homology arms (Figures 12 and 13) with a conserved point deletion several bases downstream. It looks as though both gRNAs for A3B were transfected, expressed and able to cut at the target site, possibly no A3B targeting vector was present and a recombination event occurred. BFTC-Both3 survived puromycin treatment, most likely, because it has one targeting vector incorporated at the A3A locus, BFTC-Both4 may have off-target integration of the targeting vector.

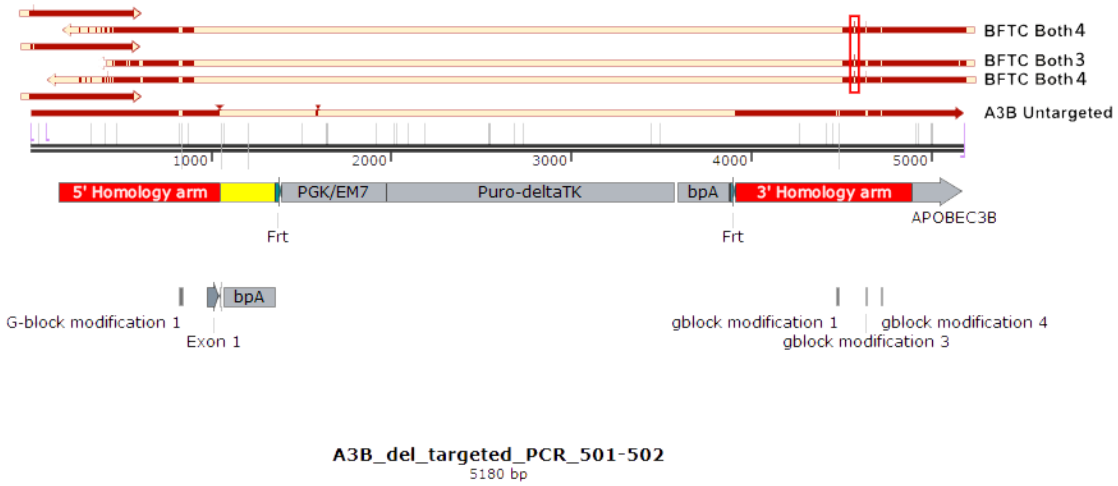

Figure 13. Comparison of A3B Truncated Genes to Targeting Vector Incorporation and BFTC-905 Wild Type.
The bottom-most line (arrow headed) represents the targeting vector (Figure 4).Cream lines depict missing bases, red depicts matched bases to A3B with the predicted targeting vector incorporation sequence (Provided by Tim Fenton). Truncated A3Bs (BFTC-Both3 and BFTC-Both4) lose ~200bp upstream and ~600bp downstream of the insertion site, interestingly these samples also lose 3 bases just downstream of the large deletion (marked in red box) this deletion is not found in the untargeted A3B of BFTC-905.

3.3 Removal of Selection Cassette

3.3.1 Ganciclovir Treatment

Selected cell lines were transfected with flippase, to remove the dual-selection cassette, they were then treated with ganciclovir. Any cells which did not express the flippase plasmid after transfection would still be sensitive to ganciclovir due to the sensitivity conferred by the HSV-TK gene in the remaining cassette.

The correct concentration of ganciclovir was determined using a cell viability assay which was run via the method described in section 2.12. This assay compared ganciclovir sensitivity of pre-flippase BFTC SCCs (containing the selection cassette) to that of parental lines (Figures 11 and 12). Results show that after 5 days only 25% of cells were viable at both 16.667μM and 50μM (Figure 14), in comparison to no

loss of cell viability in the parental line (BFTC-905) at 16.667μM and a very low percentage of 2% at 50μM (Figure 15). A control was also run with cells in the same volume of DMSO (used to dissolve ganciclovir in) data is shown in supplementary figures 3 and 4, the DMSO concentration was deemed not to effect results below the comparative ganciclovir concentration of 16.667μM. From these data a concentration of 20μM was used, as a slight increase in concentration to ensure cell death quickly with no expected toxicity from DMSO. Cells were in selection media for 2 weeks, until control lines (SCC lines subjected to transfection reagent but no flippase) were dead.

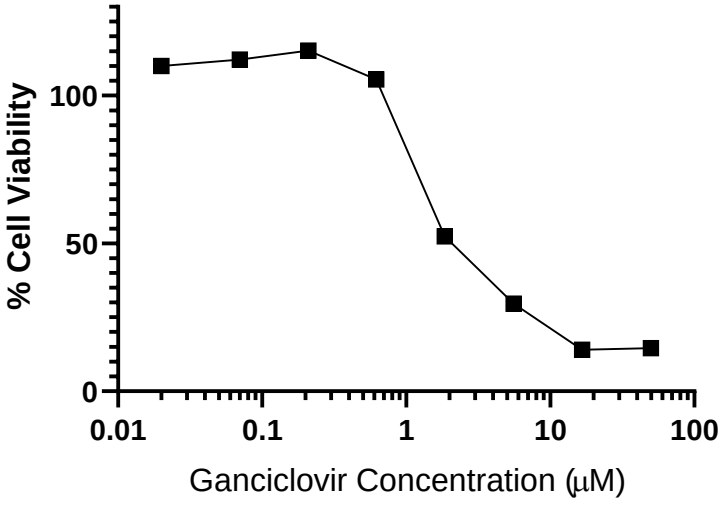

Figure 14. Cell Viability of BFTC-905 cells Containing the Targeting Vector.
Viability of cells containing the dual-selection cassette grown in differing concentrations of ganciclovir from: 0.02μM to 50μM (in a repeated 1:3 serial dilution) for 5 days. Cells become sensitive to ganciclovir at concentrations between 0.62μM and 1.85μM. After 5 days in concentrations of 16.67μM there is viability of 15% which does not increase as concentration of ganciclovir increases. Cell viability assay was normalised to DMSO only controls as ganciclovir was dissolved in DMSO.

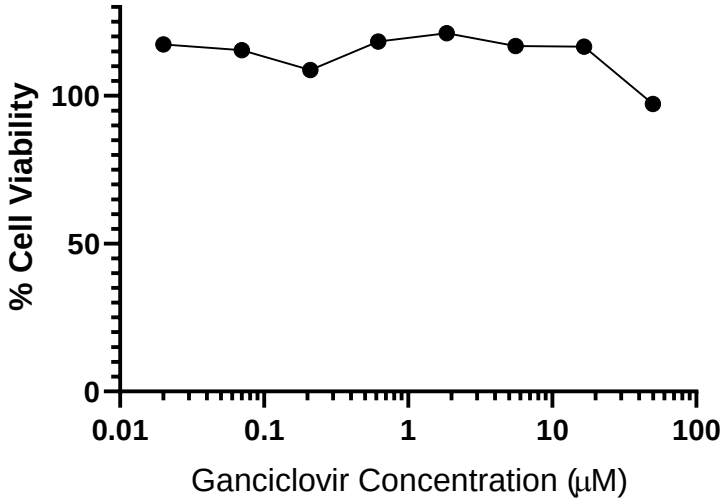

Figure 15. Cell Viability of BFTC-905 Parental Cell Line.
Viability of BFTC-905 parental cells grown in differing concentrations of ganciclovir from: 0.02μM to 50μM (in a repeated 1:3 serial dilution) for 5 days. Parental line cells show some sensitivity to ganciclovir at a concentration of 50μM after 5 days. With effects being relatively similar throughout all other samples. Cell viability assay was compared to DMSO only controls, which showed no change in viability, this was because ganciclovir was dissolved in DMSO.

During this time cells that had been affected by flippase continued to grow whilst in ganciclovir containing media, this effect is described in Figure 15, where parental BFTC-905 cells continued to grow at all concentrations of ganciclovir. Interestingly enough, one control cell line BFTC-Both5-C showed colonies that had spontaneously become resistant to ganciclovir and had started to grow similarly to lines where flippase had been added. These spontaneously resistant cells were frozen down with no further work performed on them.

3.3.2 Post-Flippase Genotyping

As shown in Figure 16 sample BFTC-A2-F (where “F” denotes flippase treated) is homozygous for the knockout of A3A (2980bp), as is BFTC-Both2-F10-F. Sample BFTC-Both2-F has a larger band than that of the insertion height (~8000bp), post-flippase this band reduces by around 2000bp which is similar to the expected

reduction in band size, despite still being around 6000bp. Another unexpected result is that BFTC-Both3-F, which before flippase treatment appeared heterozygous for the A3A insertion (Figure 11), now presents a single band at the height of a homozygous deletion (Figure 16). BFTC-Both5-F remains homozygous after flippase treatment but now has an extra band at 10,000+bp. BFTC-B5-F shows one band after treatment (Figure 16), when compared to the pre-flippase genotype it is predicted that the sample is still heterozygous, where the targeted allele containing the insertion is now of similar height to the WT band. The post-flippase A3B targeted gene is predicted to be ~200bp longer than the WT allele and would be hard to distinguish on an agarose gel. Sample BFTC-Both4-F shows no difference in length of the truncated A3B allele post-flippase, further confirming that the targeting vector was never incorporated. BFTC-Both3 has a 3rd band in the pre-flippase sample, this is missing, alongside a less intense wild-type band, in the post-flippase sample.

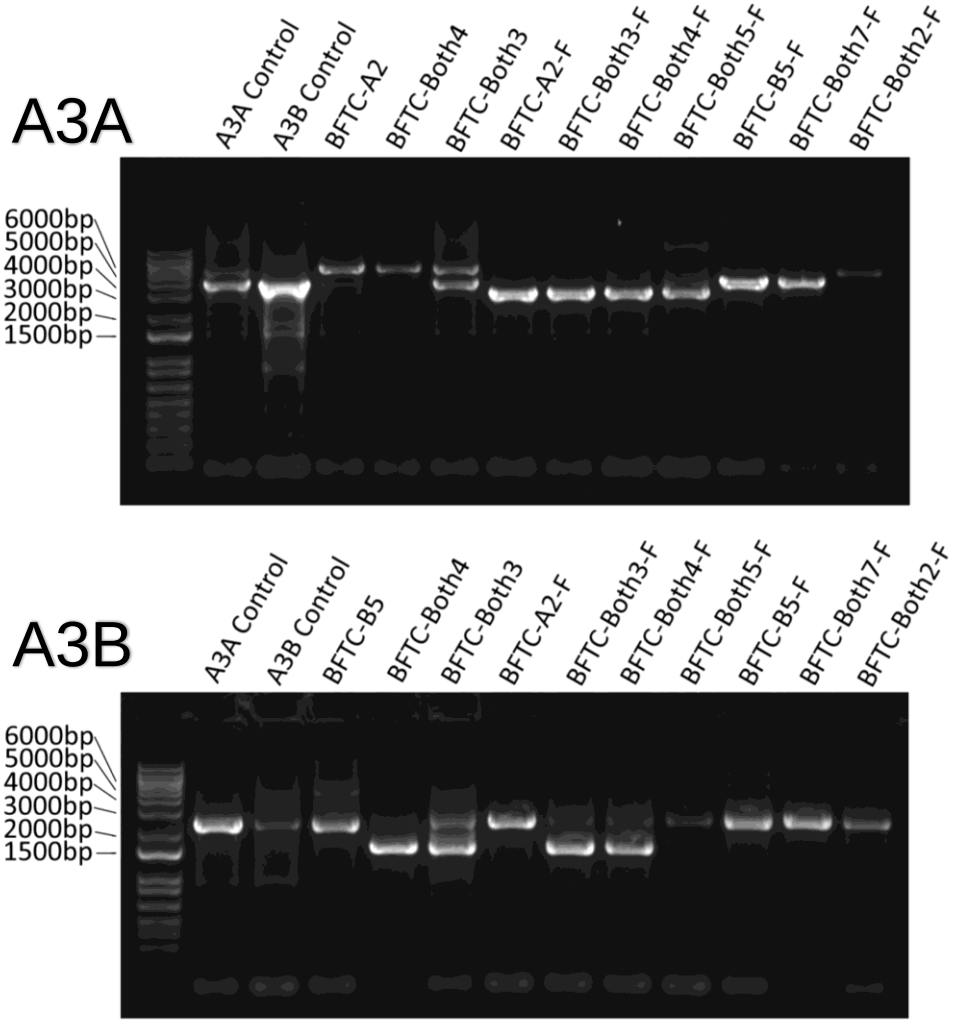

Figure 16. Genotyping of Post-Ganciclovir, Flippase-Treated BFTC-905 SCCs.
1% agarose gel displaying PCR products from A3A and A3B genotyping PCR reactions. Samples are compared against 3 samples pre-flippase. **A3A:** BFTC-Both1-G10-F still has a band for the deletion insertion whereas other A3A knockout samples show bands lower than the wild-type at the expected height of 2980bp (Figure 3). **A3B:** BFTC-B1-F8-F shows the loss of the band at insertion height, BFTC-Both2-E11-F shows loss of intensity of the wild-type gene (Figure 4).

3.4 Single-Cell Clone A3A/A3B Gene Expression via RT-qPCR

To validate gene deletion RT-qPCR was performed on SCCs in comparison to the parental, BFTC-905, line. Unfortunately, RNA from samples BFTC-Both7-F and BFTC-Both4-F was lost, therefor RNA extracted from the same lines pre-flippase was used instead. The use of pre-flippase RNA should still provide comparable data

as the A3A and A3B genes should still be disrupted as the qPCR primers for A3A and A3B are downstream of the transcriptional termination sites (Figures 3 and 4). Of the A3A homozygous knockouts; BFTC-A2-F shows no measurable expression of A3A at all and BFTC-Both2-F shows little expression of A3A (Figure 17). BFTC-Both5-F, genotyped as an A3A homozygous knock-out (Figure 16), also shows a large loss in expression of A3A. While BFTC-Both4 and BFTC-Both3-F, which present as A3A homozygous knockouts in genotyping (Figure 16) both have the largest expression of A3A, with a 5-fold increase in expression when compared to the parental line (Figure 17). Expression of A3B (Figure 18) did not differ much between the heterozygous knock-out BFTC-B5-F and the parental line, but expression drops further in BFTC-Both5-F and BFTC-Both2-F which showed no incorporation of the A3B knock-out vector at any stage of genotyping. More interestingly samples BFTC-Both4 and BFTC-Both3-F, samples which have complete loss of Exon 1, both show a ~10-fold loss in expression of A3B in comparison to the parental line.

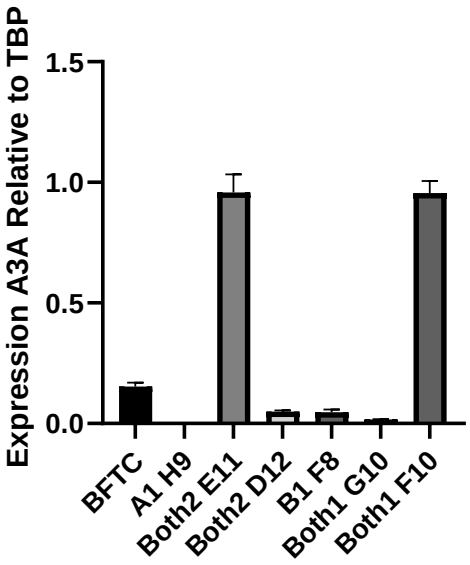

Figure 17. RT-qPCR Expression of A3A in Cell Lines Treated with Flippase.
Expression is shown relative to TBP expression per repeat. Normal expression of A3A for this line is shown by BFTC-905. A3A expression is undetectable in the homozygous knock-out BFTC-A2-F, and is reduced in other lines, especially BFTC-Both2-F. Though lines BFTC-Both5-F and BFTC-Both4 show 5-fold increases in expression.

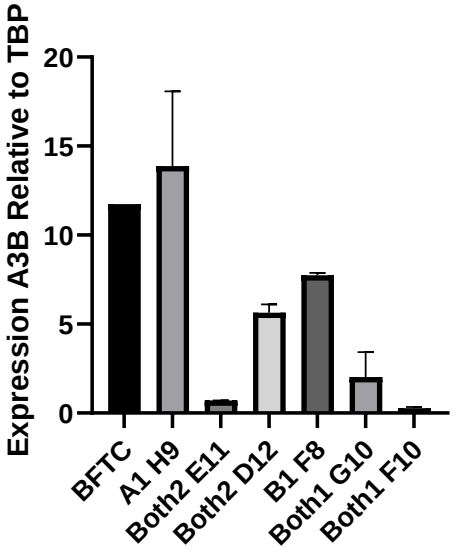

Figure 18. RT-qPCR Expression of A3B in Cell Lines Treated with Flippase.
Expression is shown relative to TBP expression per repeat. Normal A3B expression for this line is shown by BFTC-905. BFTC-Both3-F and BFTC-Both3 show a more than 10-fold loss of expression, BFTC-B5-F around a 3rd is lost, BFTC-Both5-F's expression is halved and BFTC-Both2-F is quartered.

3.5 Summary

Through use of gRNA plasmids the target site was successfully targeted in the first exons of A3A and A3B, transient expression of GFP confirmed initial transfection. Targeting vector plasmids for A3A and A3B were successfully incorporated into the target site of BFTC-905, T24 and 5637 cell lines, confirmed through use of PCR (FIGURE). The incorporated dual-selection cassette allowed for initial selection of transfected cells in transfected pools, the following single-cell cloning stage was fairly unsuccessful, with only 29 clonal lines produced. Of these 29 pools only one homozygous knockout for A3A was confirmed, after ganciclovir treatment and RT-qPCR analysis of mRNA expression.

4 Discussion

The deletion of A3 genes in UBC cell lines should be an effective tool in understanding the ability of cells to avoid and evolve against anti-cancer drugs. The targeting vectors and gRNAs used in this experiment have produced a homozygous knock-out of A3A (BFTC-A2-F) successfully, which can now be used as a tool to explore A3A's role in UBC. The knock-out of A3B was less successful, where BFTC-B5-F was the only surviving singly targeted sample, appearing to be a heterozygous knock-out, this will need to be verified via targeted Next Generation Sequencing (NGS) of the A3 region on chromosome 11. Though, there is a faint band at the insertion size (Figure 16) suggesting that some of the population still contains the selection cassette in one allele after flippase treatment and ganciclovir selection. Another round of cloning would ensure all cells have the same post-flippase genotype.

It would seem that the limiting factor of this experiment was the SCC stage, where only 29 clones were produced from 10 96-well plates. This could be due to the technique used, an in-plate dilution strategy leads to a large percentage of wells in the plate being seeded with more than one cell, effectively reducing the total plate

number to 2.5. To improve the SCC stage efficiency an out-of-plate serial dilution to a concentration of 0.5 cells per well (2.5 cells per ml) is a better option, increasing the chance of there being one cell per well. Another possibility is that clones which have incorporated the targeting vector may have a harder time surviving in single-cell conditions, this was likely the case in the T24 line where only 4 clones survived, which were all were wild type. Adding feeder cells during the cloning process may help colonies survive, as it was observed that some colonies, of both T24 and BFTC-905, would form only to die after a few divisions. T24 in general appeared to struggle more than BFTC-905 derived cells in a single cell environment, with few wells even forming small colonies.

Another stage which possibly reduced chance of homozygous deletions being produced, was during initial transfection. Due to problems with the flow cytometry machine, cells were initially only selected through exposure to puromycin, utilising the selectivity cassette incorporated via the targeting vector. Selectivity could have been enhanced via an additional cell sorting step. gRNA plasmids also encoded for Fluorescent proteins to be transiently expressed: 'left' gRNA expressed BFP, while the 'right' gRNA GFP, this allows for sorting of cells which only show expression of both gRNAs. Cells that express both BFP and GFP would be more likely to form a nickase pair either side of the DNA strand at the desired site and therefore lead to incorporation of the targeting vector. Puromycin selection is still needed to remove cells which did not incorporate the targeting vector.

By selecting cells only via puromycin treatment there could be cells which have an off-target incorporation of the targeting vector and therefore gain resistance without loss of A3. Though SCC is a method of easily removing these populations, the combination of cell sorting and Puromycin selection would help to enrich for cells with correct targeting and may reduce the number of SCCs which had off-target incorporation of the targeting vector.

During puromycin treatment, puromycin concentration had to be increased to achieve the death of the control cells in all lines (T24, 5637 and BFTC-905). This could allow for untransfected cells to become resistant and survive treatment and confound SCC. It was also observed during treatment, that cells seemed to stagnate rather than grow as resistant cells should. Nevertheless, BFTC-905 transfected pools showed strong bands for the insertion of these vectors in genotyping (Figure 10). One way to fix this issue would be to run viability assays on the parental lines pre-transfection and use the results to predict a working concentration of puromycin.

Flippase treatment could have also used cell sorting as the plasmid it was expressed from also transiently expresses GFP. But, through the use of viability assays, ganciclovir treatment was, in general, effective for A3A knockouts, though seemingly not so for A3B. For one sample BFTC-Both5, the control line became spontaneously resistant to ganciclovir. This line was heterozygous for A3A deletion pre-flippase, which could suggest that heterozygous lines are less sensitive to ganciclovir as they may have less expression of HSV-TK. This could possibly lead to these cells having a higher chance of spontaneously deleting the sensitivity cassette through longer natural survival, whilst only requiring deletion of the gene on one chromosome. Or it could be that this line, which after flippase activity has a very large band >10kb (Figure 16), went through a round of rearrangement whilst passaging which could allow for easier deletion of the cassette. Again, NGS via Targeted Sequencing of the A3 region could reveal what has happened in both the control and flippased variants of this line.

Samples transfected with both A3A and A3B targeting vectors made up one third of all SCCs and most of these samples took up at least one targeting vector (Figures 8 and 9). 2 of these samples do knock down A3A expression by a large amount in comparison to the wild type (Figure 17). While 2 other samples have a full deletion of exon 1 in A3B (Figures 12 and 13) whilst also appearing to have homozygous knock-outs of A3A (Figure 16), but RT-qPCR shows a 5-fold increase in A3A expression (Figure 17) and a >10-fold knock-down of A3B expression (Figure 18), in

these clones. Though sample BFTC-Both4 has not been treated with flippase, it should still have the transcription termination site and therefore not express A3A.

The deletion of 2 genes which are very close together, using two nickase pairs may lead to unexpected deletions, rearrangements or copy-number changes. There are multiple combinations of what may have happened to BFTC-Both3-F and BFTC-Both4 via expression of 1 or more gRNAs. Concentration of each gRNA and targeting vector was halved in dual targeting vector transfections, in comparison to lines were only one A3 was targeted (Section 2.5). Large sections of the A3 locus could be removed (~30kb) as A3A and A3B lie along chromosome 11 in the same order, complete repair of such damage could be quite hard. It was originally thought that A3B deletion could be due to the cell trying to save either gene, though through RT-qPCR and sequencing it appears it was instead a deletion (Figure 18). The small amount of mRNA expressed, as seen in RT-qPCR, is nonsense mRNA that is being removed via nonsense mediated decay. Indeed, in the sequencing of this deletion an ATG was left just before the deletion took place, although this is unlikely to be able to be targeted by promoters and lead to expression, promoters upstream on the A3 locus may be driving expression in a different reading frame. Melt curves were also at the same temperatures for all samples and all primers, suggesting that everything expressed was the same size as expected. The only way to see what has happened is to use targeted sequencing of the A3 locus surrounding both A3A and A3B, especially for the apparent A3A homozygous deletion.

With A3A deleted in one line, more SCCs should be grown up from the BFTC-905 A3A deletion pool to create biological replicates for experimentation. It also confirms the ability to create homozygous deletions using the A3A targeting vector and accompanying gRNAs. The next step in this line of experimentation is to observe ability of cells without A3A and/or A3B to become resistant to anti-cancer drugs, this can be done by growing the lines in the drugs until a population becomes resistant. A comparison could be drawn between BFTC-905 lines which are already resistant

to these drugs, BFTC-A2-F and the BFTC-905 parental line, a useful control for this experiment would be cells that have gone through the same process as BFTC-A5-F and survived without A3 loss, BFTC-Both7-F would be a good line for this. A comparison of RT-qPCR data for expression of A3A and A3B between resistant lines, deletion and parental may shed some light on A3s involvement in resistance. But it could be that due to the episodic nature of APOBEC mutation (2, 4, 6) that no real difference can be observed.

Another use for the A3A homozygous deletion is that these lines can now be targeted for A3B deletion and a possible double-knock-out homozygous line could be created. The deletion of either A3 seemed quite rare in deletion pools (Figures 8, 9 and 10), though the A3B insert band is a lot less bright than that of the A3A insert band, suggesting that incorporation of the A3B targeting vector is rarer. It could be that A3A may affect transfection efficiency, so the knock-out line may be a much better target for A3B transfection than targeting both A3A and A3B at the same time. Though this 2-step transfection may not be the most ideal way of knocking out both genes, as cells will be passaged a lot, requiring SCCs of both the A3A knock-out line and the parental BFTC-905 line again for comparisons. This idea originally came about when the loss of exon 1 in A3B was discovered in BFTC-Both4 and BFTC-Both3 with the suggestion that the deletion of one or more allele of A3A (Figures 11, 12 and 16) though RT-qPCR has shown A3A expression was not lost (Figure 17).

It could be that transfection efficiency of hard-to-transfect cell lines could benefit from A3A inhibitors. Huerfano et al found that INF-β is upregulated and DDR proteins are expressed in mouse fibroblasts transfected via nucleofection with plasmid DNA (23). In human cells cytoplasmic DNA has been found to upregulate INF-α, which in turn massively upregulates A3A but not A3B. Suspène et al. find that exogenous dsDNA causes activation of the INF-I pathway inducing A3A which then attacks ssDNA made through an exonuclease deaminating cytosines which are then removed by UNG creating abasic sites and leading to catabolism of cytoplasmic ssDNA (Figure

19) (13). Huerfano et al. did not look at IFN-α expression and used only mouse cells, which have only one A3 gene, Suspene et al. found dsDNA is required for IFN-I upregulation (23). Because of this, a novel experiment comparing BFTC-A3A homozygous knockouts between the parental line in a transfection efficiency assay with differing amounts of gRNA plasmids may reveal a negative correlation between A3A expression and transfection efficiency. This could be repeated, with results compared to lines where A3A was massively upregulated, after verification by targeted sequencing, it could confirm a negative correlation.

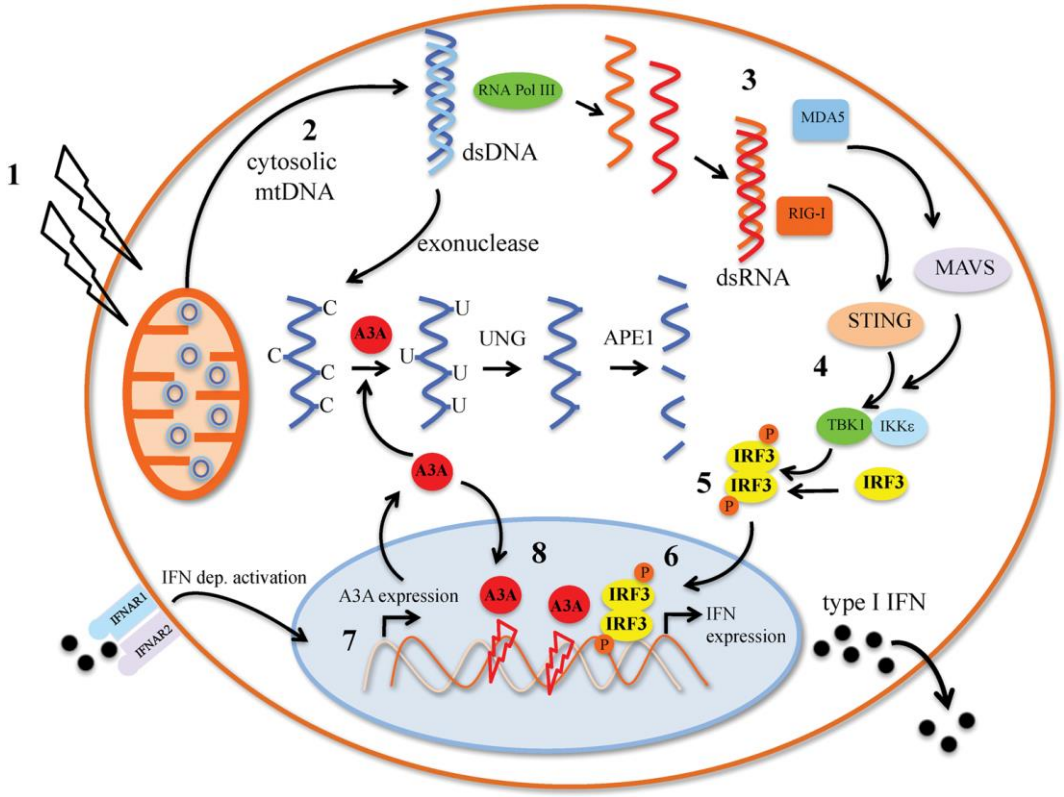

Figure 19. Mechanism of A3A Upregulation by Mitochondrial and Exogenous DNA.
1. mtDNA or exogenous DNA enters the cell. 2. RNA Polymerase III transcribes dsDNA into RNA. 3. dsRNA forms from the transcription of dsDNA and is sensed by RIG-I 4-5. which signals for IRF-3 to be phosphorylated by TBK1. 6. Interferons are induced and 7. upregulate A3A. A3A catalbolises ssDNA in the cytoplasm, where UNG removes Urasil leaving abasic sites. 8. Upregulated A3A can translocate to the nucleus and hypermutate genomic DNA. (Figure is from Figure 6 in Suspène et al.'s 2017 paper (13))

Results seen here mirror what has previously been achieved in the same lab with the same tools on a NIKS cell line (Tim Fenton, Unpublished), where the control DNA for A3A and A3B incorporation in cell pools came from. This experiment, though in normal cells, managed to produce an A3A homozygous deletion and an A3B heterozygous deletion. This suggests that A3B, using these targeting vectors, is very hard to produce a homozygous knock-out in. It may be that the complete loss of Exon 1, as seen in BFTC-Both3 and BFTC-Both4, does knock-out A3B, but this has only been survivable for the cell by some form of rearrangement and compensation of A3A expression which might be facilitated by both sets of gRNAs and targeting vectors being present.

The applications of an A3A homozygous knockout, as described above, will allow better understanding on how the targeted knockdown of A3As expression may affect the recurrence of UBC in patients. The experiment described above, for adaptation of BFTC-A5-F cells to various platinum-based anti-cancer drugs will help to explain A3As role in rapid resistance gain to these various drugs. BFTC-A5-F will also act as a platform to create an A3A/A3B double knockout line which can further explain the two enzymes role in resistance gain and recurrence in UBC, with a repeat of the above experiment. For BFTC-905 A3B only knockout lines, as with T24 and 5637 cells, the transfected pools from this experiment can be utilised to obtain new SCCs, as would more biological repeats for the A3A double-knockout line.

Conclusion

One homozygous A3A deletion was generated, BFTC-A5-F, it can be used for investigation into A3As role in UBC becoming resistant to anti-cancer drugs, as a platform to create an A3A/A3B homozygous double deletion and has raised a question about transfection efficiency. This study was able to provide insight and improvement into how further knock-out UBC cells could be generated, as well as pools of cells from T24 and 5637 lines ready to undergo SCC and possibly repeated SCC on BFTC-905 pools to get more clones with homozygous deletions for comparison. Previous work has been built on top of from Tim Fenton's experiments with the same vectors in normal cells producing an A3A homozygous knock-out, there is now a cancer comparison.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel R, Torre L, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. 2018;68(6):394-424.
2. Faltas B, Prandi D, Tagawa S, Molina A, Nanus D, Sternberg C et al. Clonal evolution of chemotherapy-resistant urothelial carcinoma. Nature Genetics. 2016;48(12):1490-1499.
3. Shi M, Meng X, Lamy P, Banday A, Yang J, Moreno-Vega A et al. APOBEC-mediated Mutagenesis as a Likely Cause of FGFR3 S249C Mutation Over-representation in Bladder Cancer. European Urology. 2019;76(1):9-13.
4. Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada S et al. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. Annals of Oncology. 2018;29(3):563-572.
5. Roper N, Gao S, Maity T, Banday A, Zhang X, Venugopalan A et al. APOBEC Mutagenesis and Copy Number Alterations are Drivers of Proteogenomic Tumor Evolution and Heterogeneity in Metastatic Thoracic Tumors. SSRN Electronic Journal. 2018;.
6. Liu D, Abbosh P, Keliher D, Reardon B, Miao D, Mouw K et al. Mutational patterns in chemotherapy resistant muscle-invasive bladder cancer. Nature Communications. 2017;8(1).
7. Henderson S, Fenton T. APOBEC3 genes: retroviral restriction factors to cancer drivers. Trends in Molecular Medicine. 2015;21(5):274-284.
8. Swanton C, McGranahan N, Starrett G, Harris R. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. Cancer Discovery. 2015;5(7):704-712.
9. Rebhandl S, Huemer M, Greil R, Geisberger R. AID/APOBEC deaminases and cancer. Oncoscience. 2015;2:320.

10. Chan K, Roberts S, Klimczak L, Sterling J, Saini N, Malc E et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nature Genetics. 2015;47(9):1067-1072.
11. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O et al. Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). Cell. 2000;102(5):565-575.
12. Chaudhuri J, Alt F. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nature Reviews Immunology. 2004;4(7):541-552.
13. Suspène R, Mussil B, Laude H, Caval V, Berry N, Bouzidi M et al. Self-cytoplasmic DNA upregulates the mutator enzyme APOBEC3A leading to chromosomal DNA damage. Nucleic Acids Research. 2017;:gkx001.
14. Burns M, Temiz N, Harris R. Evidence for APOBEC3B mutagenesis in multiple human cancers. Nature Genetics. 2013;45(9):977-983.
15. Alexandrov L, Nik-Zainal S, Wedge D, Campbell P, Stratton M. Deciphering Signatures of Mutational Processes Operative in Human Cancer. Cell Reports. 2013;3(1):246-259.
16. Silvas T, Hou S, Myint W, Nalivaika E, Somasundaran M, Kelch B et al. Substrate sequence selectivity of APOBEC3A implicates intra-DNA interactions. Scientific Reports. 2018;8(1).
17. Shi K, Carpenter M, Banerjee S, Shaban N, Kurahashi K, Salamango D et al. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. Nature Structural & Molecular Biology. 2016;24(2):131-139.
18. Landry S, Narvaiza I, Linfesty D, Weitzman M. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. EMBO reports. 2011;12(5):444-450.

19. Roberts S, Sterling J, Thompson C, Harris S, Mav D, Shah R et al. Clustered Mutations in Yeast and in Human Cancers Can Arise from Damaged Long Single-Strand DNA Regions. Molecular Cell. 2012;46(4):424-435.
20. Middlebrooks C, Banday A, Matsuda K, Udquim K, Onabajo O, Paquin A et al. Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. Nature Genetics. 2016;48(11):1330-1338.
21. Nordentoft I, Lamy P, Birkenkamp-Demtröder K, Shumansky K, Vang S, Hornshøj H et al. Mutational Context and Diverse Clonal Development in Early and Late Bladder Cancer. Cell Reports. 2014;7(5):1649-1663.
22. Law E, Sieuwerts A, LaPara K, Leonard B, Starrett G, Molan A et al. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. Science Advances. 2016;2(10):e1601737.
23. Huerfano S, Ryabchenko B, Forstová J. Nucleofection of Expression Vectors Induces a Robust Interferon Response and Inhibition of Cell Proliferation. DNA and Cell Biology. 2013;32(8):467-479.

Supplementary Data

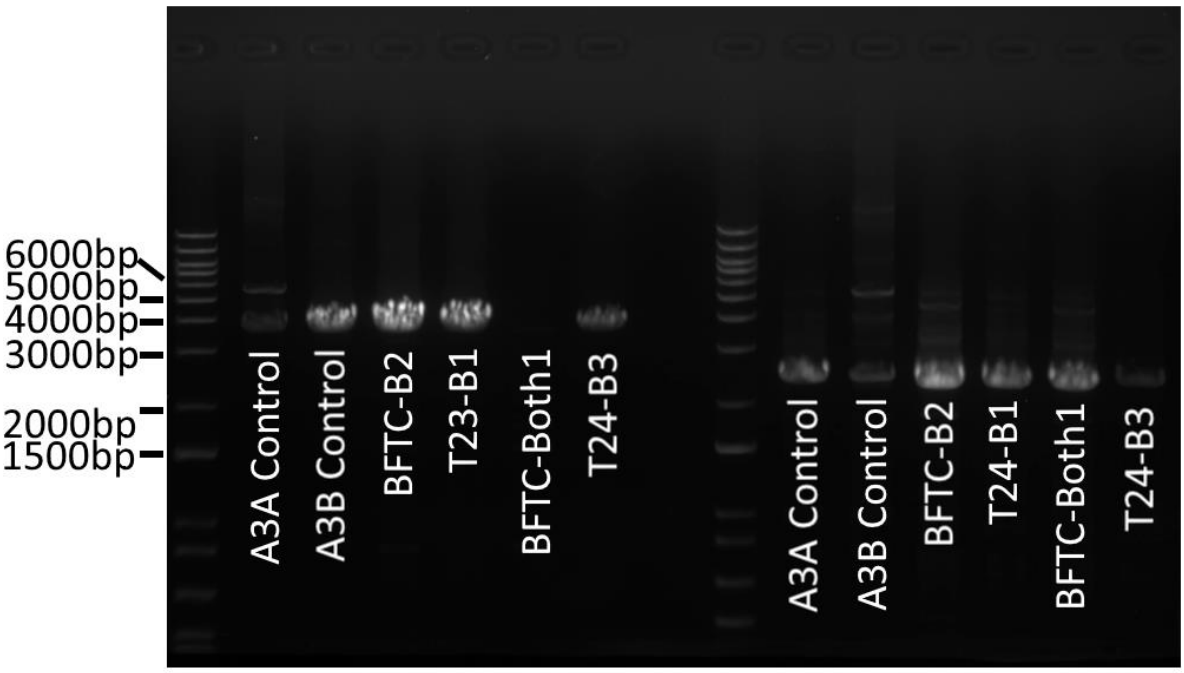

Figure S1. A3A and A3B SCC Repeats. 1% Agarose gel of products from PCR with A3A and A3B genotyping primers. Samples on the left were genotyped for A3A and A3B on the right. Genotyping was repeated as these samples had one or more missing band in Figures 11 and 12.

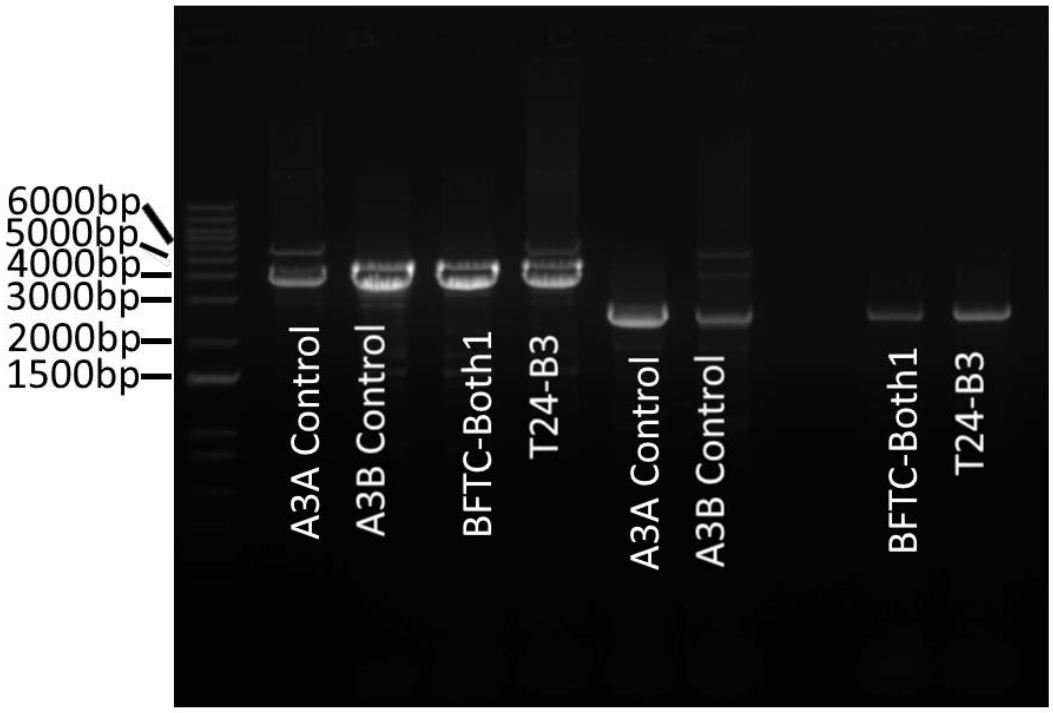

Figure S2. A3A and A3B SCC Repeats 2. 1% Agarose gel of products from PCR with A3A and A3B genotyping primers. Samples on the left were genotyped for A3A and A3B on the right. Genotyping was repeated as these samples had one or more missing band in Figure S1.

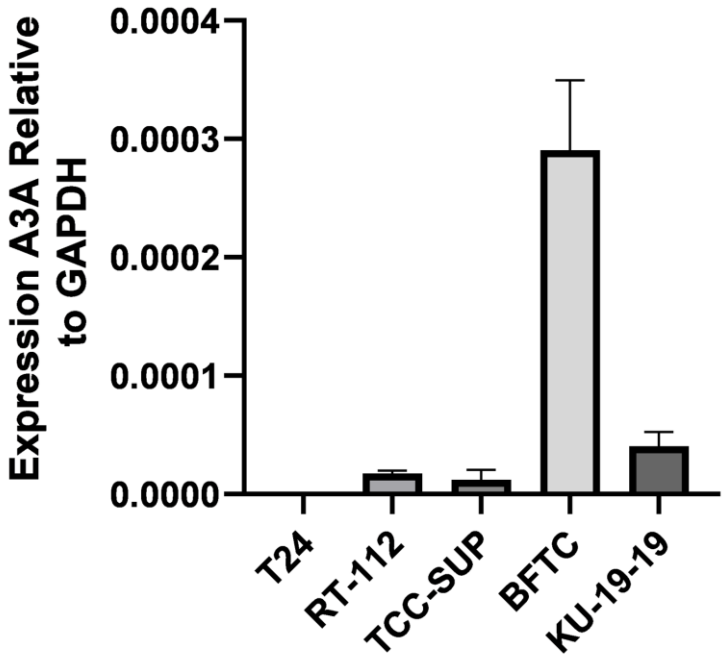

Figure S3. A3A Expression Levels across 5 Cancer Lines relative to GAPDH. RT-qPCR data comparing expression of A3A in 5 cancer cell lines show BFTC-905 has the largest expression of A3A in comparison to other lines. 5637 was missed out due to inconsistent data.

Data shown in in Figures S3 and S4 describe expressions of A3s in untransfected UBC cancer lines. Lines were chosen for transfection, from parental lines, on ease of upkeep, with T24 and 5637 proving difficult to get the targeting vector incorporated BFTC-905 was chosen to be transfected 3rd. BFTC-905's high A3A expression and moderate A3B expression would allow for accurate comparisons between original expression levels and those after transfection. A3A is particularly difficult in RT-qPCR as it is often expressed at such low levels it cannot be properly detected.

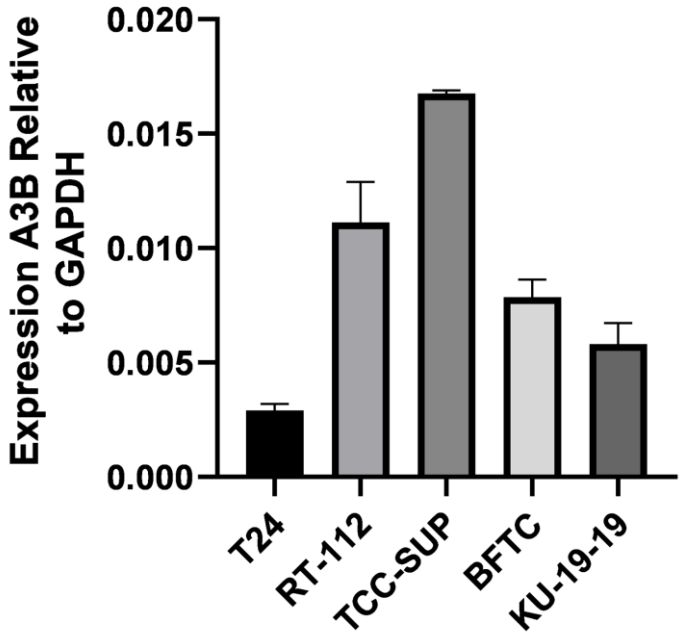

Figure S4. A3B Expression Levels across 5 Cancer Lines relative to GAPDH. RT-qPCR data comparing expression of A3B in 5 cancer cell lines show BFTC-905 has a median expression of A3B in comparison to other lines, whereas TCC-SUP has the highest A3B expression. 5637 was missed out due to inconsistent data.