Chapter 3

Triple A syndrome

3.1 Knockout of *AAAS* exon 2

To knockout the function of *AAAS*, the biallelic targeting strategy was to replace one copy of exon 2 with a puromycin (puro) selection cassette by homologous recombination (HR) and simultaneously damage the other allele by non-homologous end-joining (NHEJ; Figure 3.1.1).

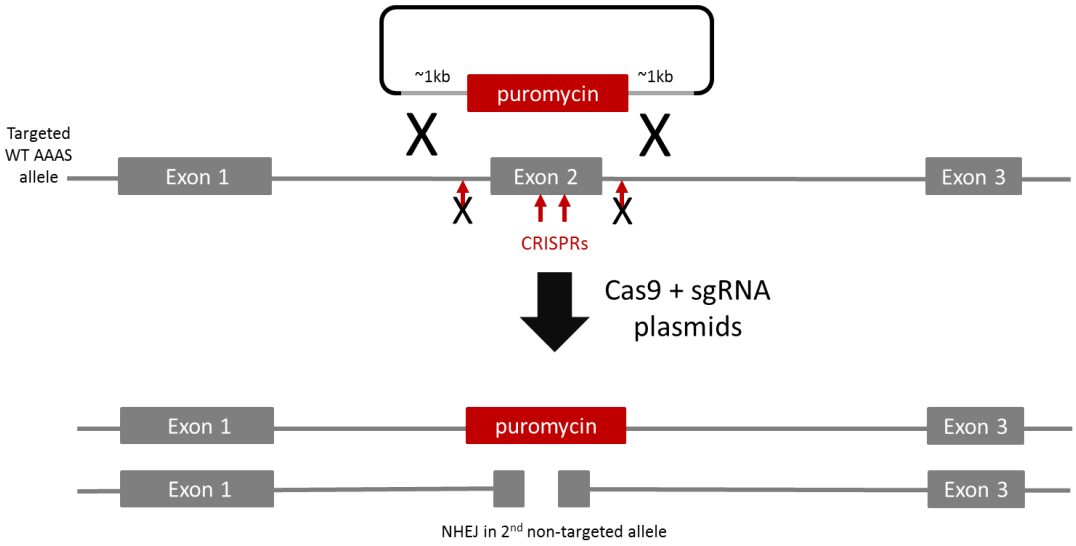


Figure 3.1.1: **CRISPR-Cas9 biallelic gene targeting strategy for AAAS**. One copy of exon 2 of *AAAS* is deleted by CRISPR-Cas9 assisted homologous recombination with a puro-resistant plasmid donor. NHEJ damage to the second non-targeted allele was identified in puro-resistant clones by PCR amplification and sequencing of *AAAS* exon 2. Arrows indicate functional CRISPR positions; crossed arrows indicate flanking CRISPRs which did not yield viable cells. "X" refers to homologous recombination.

The knockout of *AAAS* exon 2 (AAAS2) was a plasmid-based nuclease-assisted gene knockout in KOLF2-C1 hiPS cells. Cells were co-transfected with CRISPR gRNA expression plasmids, Cas9 nuclease, and targeting vector plasmids. Two experiments were conducted: one using two CRISPRs flanking exon 2 and the second using two CRISPRs lying within the exon. The targeting vector contained the EF1α-puro-polyA cassette for introduction into the alleles by homologous recombination, thus knocking out the gene, enabling for the selection of the puro resistant knockout clones. On the 3rd day after nucleofection, culture media supplemented with 0.5µg/ml puromycin was used to select for the puro-resistant colonies. The EF1a promoter, highly active in most human cells, was found to be optimal for gene editing experiments in human iPS cells (Koutsourakis & Skarnes, unpublished data). The three-day delay in supplementing the culture media with puromycin is to allow the cells enough time for expression of Cas9 protein and gRNAs and eventually for homologous recombination to occur

between genomic DNA and targeting vector, allowing expression of puro-resistance in the cells.

The targeting donor vector used for biallelic knockout of *AAAS* exon 2 was constructed in a step-wise process (Figure 3.1.2). A region of genomic DNA containing *AAAS* exon 2, with approximately 1kb homology arms on either sides of the exon, was first cloned into pUC19_RV plasmid by InFusion cloning forming pAAAS_exon 2 clone. The next step was replacing *AAAS* exon 2 of the pAAAS_exon 2 clone with Zeo/PheS (Z/P) Gateway cassette (Skarnes et al. 2011), by recombineering followed by the introduction of a drug selection cassette, pL1L2 EF1α-puro-polyA (Koutsourakis & Skarnes, unpublished), by a Gateway exchange reaction.

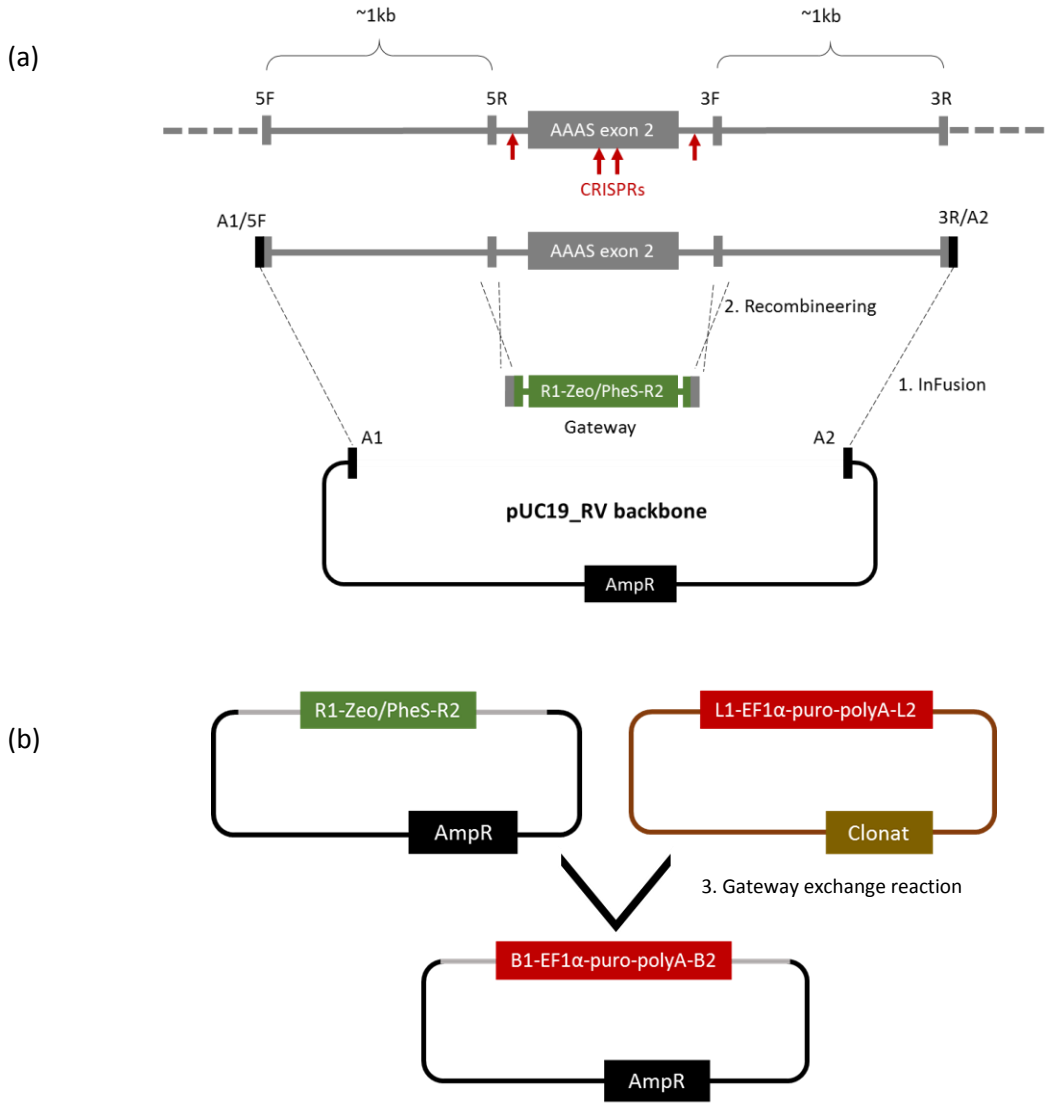


Figure 3.1.2: **Construction of targeting donor vector for knockout of *AAAS* exon 2.**
(a) Introduction of Z/P Gateway cassette into pAAAS_exon 2 by recombineering. Arrows indicate CRISPR positions. 5F and 5R are forward and reverse primers, respectively, on the

5’end of the targeted exon; 3F and 3R are on the 3’end of the exon. 5F and 3R are InFusion PCR primers for amplification of AAAS exon 2. (b) Introduction of drug selection cassette, pL1L2 EF1α-puro-polyA, into Z/P recombineered clone by LRII clonase Gateway exchange reaction.

Many puro-resistant colonies were recovered using only the CRISPRs within the exon. In contrast, only a few colonies were recovered using CRISPRs flanking exon 2. The reason for this discrepancy could be due to poor Cas9 cleavage at the intronic CRISPR sites which could not introduce the puro-cassette at the target site leading to cell death in puro-supplemented media. It is, however, also possible that this experiment failed due to other technical problems.

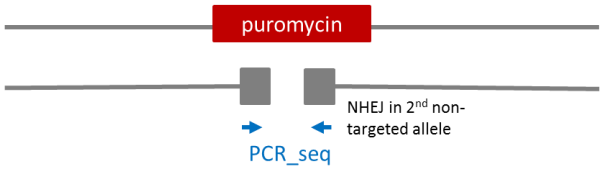
Sequencing of non-targeted allele

Analysis of 96 colonies by PCR amplification and sequencing of the non-targeted allele demonstrated that different indels were introduced by NHEJ in this allele in a small number of colonies (Table 3.1.1). Three of the clones contained frameshift mutations of which two contained deletions of 16bp and 11bp, and one had an insertion of 1bp (Figure 3.1.3). Therefore, the biallelic targeting efficiency defined as the fraction of clones with two inactive alleles was 3%. In addition, four other clones contained in-frame NHEJ indels (Figure 3.1.4). In 22 clones, overlapping sequence traces downstream of DSB site were also observed indicating they were mixed or mosaic clones. In theory, mosaic clones can be explained by cleavage of the non-targeted allele after cell division, resulting in the formation of a colony with a mixture of alleles. Almost half of the clones (52%) were wild-type containing no mutation. These clones could have suffered a DSB in their DNA, but were repaired correctly. Finally, the results suggest that the CRISPR1 exonic site was cleaved more efficiently than the CRISPR2 site as all of the indels overlapped the CRISPR1 site (Figure 3.1.3 (b)).

Table 3.1.1: Summary of the genotypes different clones obtained upon biallelic targeting of *AAAS* exon 2

No of colonies screened	**NHEJ Indel**		**Mosaic**	**Wild-type**	**No/ poor quality sequence**
	frameshift	in-frame			
96	3 (3%)	4 (4%)	22 (23%)	50 (52%)	17 (18%)

(a)



(b)

	CRISPR 1	CRISPR 2	
TCCTACAACTGACAAAGG	**ATCCCCTAAAGACCCC.TGGAAGG**	CTGGA**CCATGGCACAAGAACTGCCTTCA**	TCCATCACCGGGAGCAAGTGTGGAAGA	**WT**
TCCTACAACTGACAAAGGATCCCCTAAAGA----.	----------	CCATGGCACAAGAACTGCCTTCATCCATCACCGGGAGCAAGTGTGGAAGA	-16bp
TCCTACAACTGACAAAGGATCCCCTAAA------.	-----GGCTGGACCATGGCACAAGAACTGCCTTCATCCATCACCGGGAGCAAGTGTGGAAGA	-11bp
TCCTACAACTGACAAAGGATCCCCTAAAGACCCC**T**	TGGAAGGCTGGACCATGGCACAAGAACTGCCTTCATCCATCACCGGGAGCAAGTGTGGAAGA	+1bp

(c)

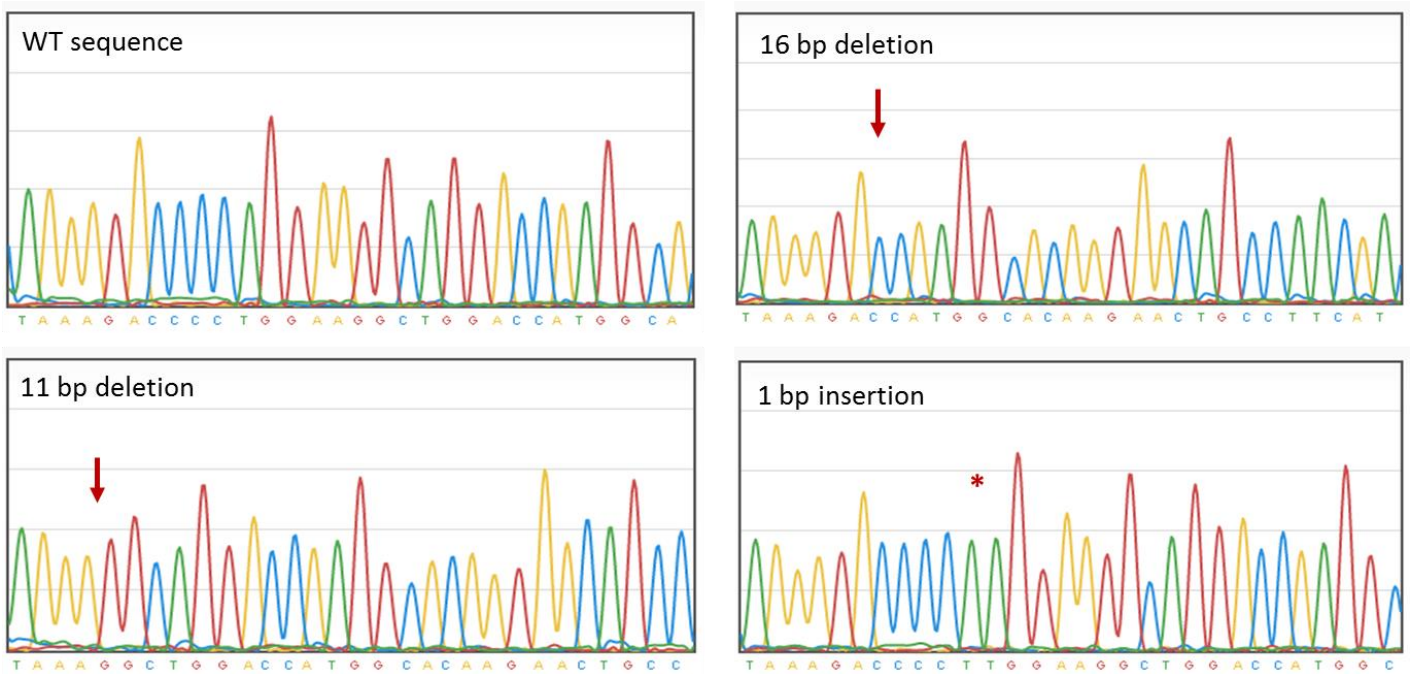


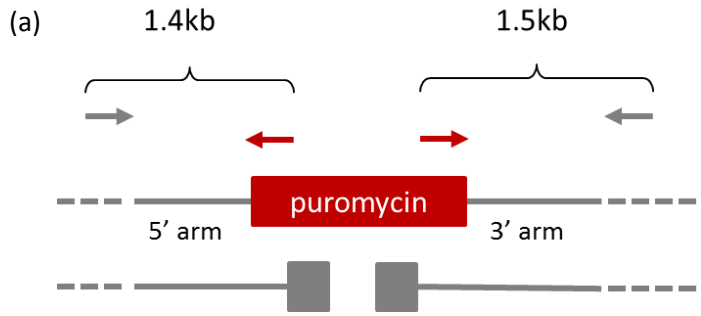
Figure 3.1.3: **Screening for NHEJ indels in non-targeted allele of *AAAS* gene.** (a) *AAAS* exon 2 showing PCR sequencing across the damaged region of non-targeted allele (sequencing primers indicated with blue arrows) (b) DNA sequences of wild-type *AAAS* exon 2 and three frameshift mutations recovered in the screen. CRISPR sites are indicated in red and PAM sites are underlined. Deleted sequences (16bp & 11bp) are indicated with dashes and inserted sequences (single base pair) in green. (c) Sequence traces of the wild-type and mutated alleles. Arrows indicate breakpoints at which deletion occurred and asterisk (*) indicates insertion of a base (T).

Figure 3.1.4: **DNA sequence alignment of in-frame NHEJ indels in non-targeted allele of *AAAS* exon 2.** CRISPR sites are indicated in red and PAM sites are underlined. Deleted sequences are indicated with dashes.

Genotyping of puro-targeted allele

The three clones containing frameshift indels in the non-targeted allele were genotyped to confirm if the puro-cassette was correctly targeted by homologous recombination. PCR amplification of the homology arms was done with gene-specific primers and puro-cassette-specific primers (Figure 3.1.5 (a); refer to Appendix, Table A8 for sequences). The gene-specific primers were designed outside of the homology arms to avoid the detection of random insertions of the plasmid donor. All three clones showed the correct-sized PCR products for both the 5' and 3' homology arms (Figure 3.1.5 (b) and (c)), suggesting that the puro-cassette was correctly inserted in each of these clones.

Two putative heterozygous clones were also chosen from the archived clones as controls. They were puro-resistant clones in which the non-targeted allele showed a wild-type sequence. However, only one of the heterozygous clones (Het1) had a correctly targeted puro cassette giving bands of expected sizes for both the 5' and 3' arms. The other clone (Het2) did not produce the expected bands and instead showed non-specific bands or smears similar to the wild-type negative control (Figure 3.1.5 (b) and (c)). This indicates that the donor plasmid was not correctly targeted but rather integrated elsewhere in the genome of this clone.



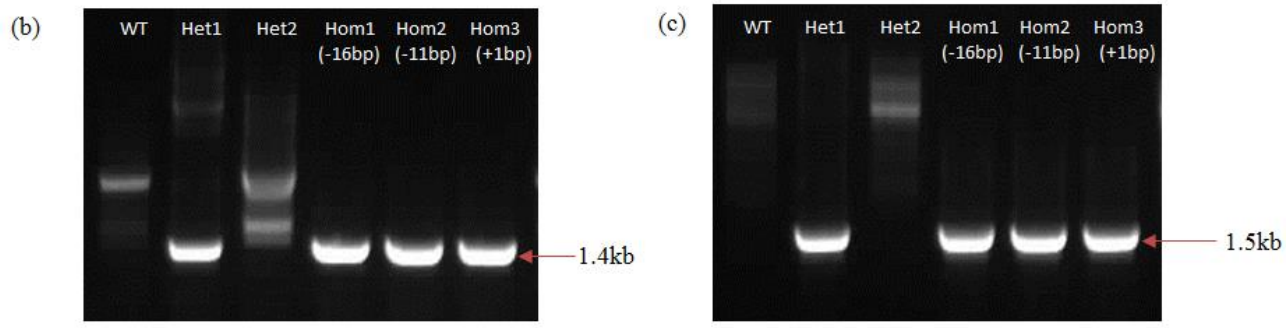


Figure 3.1.5: **Genotyping of puro-targeted allele.** (a) Schematic representation of the targeted allele showing primers for amplification of the homology arms. Red arrows indicate puro-cassette specific PCR primers, grey arrows indicate gene-specific primers outside the homology regions. Each of the homology arms are indicated along with their sizes. (b) and (c) Gel electrophoresis of PCR amplicons of (b) 5' homology arm [1.4kb] and (c) 3' homology arm [1.5kb] of the puro cassette.

3.2 *AAAS* exon 1 point mutation

This project of generating hiPS disease models for the point mutation in *AAAS* exon 1 was extended from the study by Krumbholz and colleagues (2006). Their findings (Figure 3.2.1) suggest that a single base substitution of 43C>A in exon 1 of *AAAS* gene led to the introduction of a novel splice donor site. Aberrant splicing at this site results in an exon frameshift in the gene, leading to a premature stop codon downstream. This frameshift mutation eventually resulted in a truncated ALADIN protein causing the disease manifestations as observed in patients.

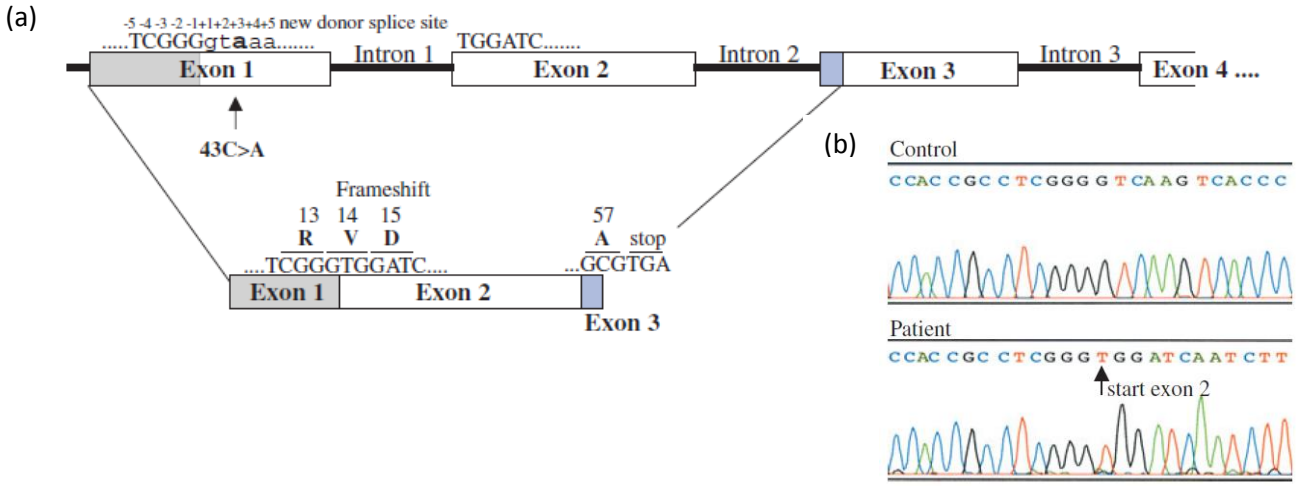


Figure 3.2.1: **Homozygous point mutation of 43C>A in *AAAS* exon 1 results in a novel splice donor site.** (a) Splicing due to the 43C>A mutation in exon 1 leads to a frameshift mutation, hence a truncated protein. (b) Electropherogram traces showing the sequences across *AAAS* exon 1 of a healthy control and that of a patient with aberrant splicing and start of exon 2 as indicated by arrow (adapted from Figure 2; Krumbholz et al. 2006)

For this part of the study, KOLF2_C1 cells were co-transfected with preassembled Cas9-RNP and ssODN (101 bases; refer to Appendix, Table A6) containing the C>A point mutation for its introduction into genomic DNA by homology-directed repair (HDR). The crRNA, tracrRNA, and Cas9 used for the Cas9-RNP preassembly were all purchased commercially. The ssODN, also bought commercially, was symmetric containing 50-base homology arms on either side of the C>A point mutation. The ssODN was non-complementary to the CRISPR gRNA oligo to prevent it from annealing to the gRNA. Two experiments were conducted, each on $8x10^5$ cells, with Cas9 protein from two commercial sources. The first experiment using 12µg of recombinant Cas9 protein (Feldan) showed decreased activity in the pool of transfected cells when compared to the second experiment done with 18µg of Cas9 (Thermofisher) based on the T7E1 assay (Figure 3.2.2).

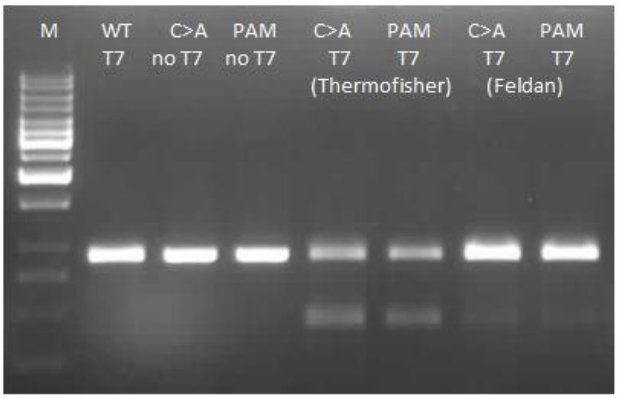
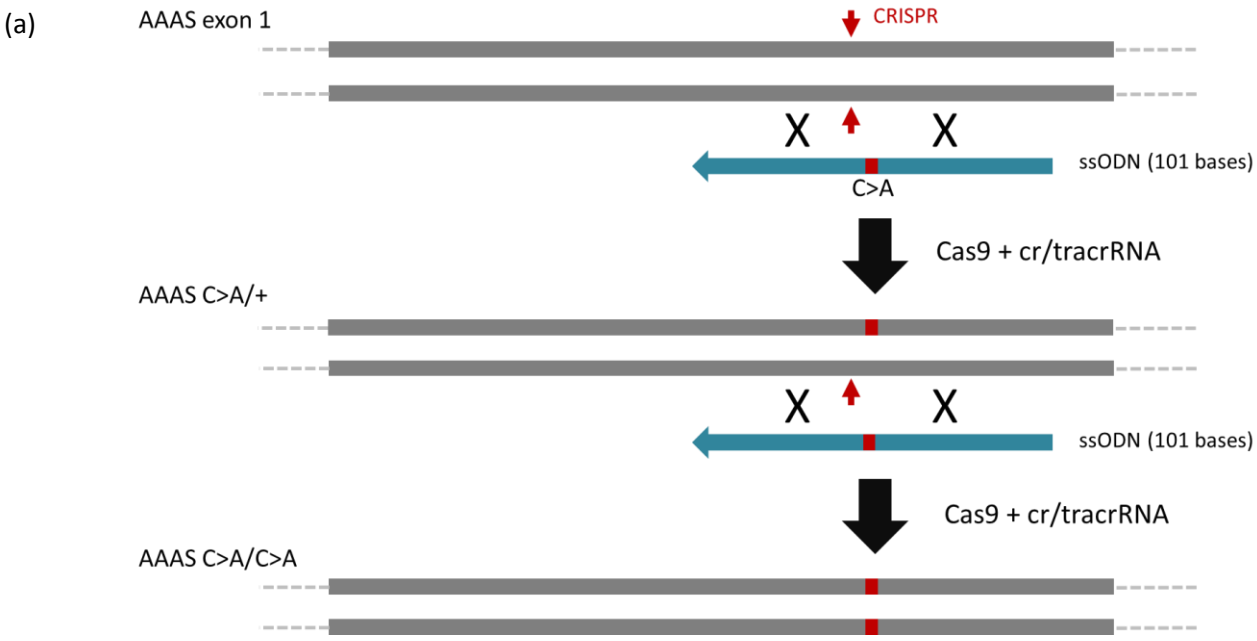


Figure 3.2.2: **T7E1 assay demonstrating Cas9 activity in a pool of transfected cells**. WT-T7: wild-type cells treated with T7E1 indicate no Cas9 activity. C>A: cells transfected with ssODN containing only the C>A point mutation; PAM: cells transfected with the control ssODN containing both C>A and PAM mutations. C>A-no T7 and PAM-no T7 are negative controls with no T7E1 treatment. C>A-T7 and PAM-T7 labelled as Thermofisher or Feldan are cells treated with T7E1 that had been transfected with Cas9 from respective manufacturers.

However, PCR genotyping of clones suggested that lesser amount of Cas9 was more efficient in producing the desired specific point mutations in the genome by HDR. The initial transfection of wild-type cells with Cas9-RNP and ssODN resulted in heterozygotes only, but no homozygotes. Thus, in order to obtain homozygotes, the experiment was repeated by re-transfecting the heterozygotes (Figure 3.2.3).



(b)

```
----CCGGCAAGATGTGCTCTCTGGGGTTGTTCCCTCCTCCACCGCCTCGGGGTCAAGTCACCCTATATGAGCACAATAACGAGCTGGTGACGGGCAGTAGCTAT----  WT
    :::::::::::::::::::::::::::::::::::::::::::::::::: :::::::::::::::::::::::::::::::::::::::::::::::::::
3'- GGCCGTTCTACACGAGAGACCCCAACAAGGGAGGAGGTGGCGGAGCCCCATTTCAGTGGGATATACTCGTGTTATTGCTCGACCACTGCCCGTCATCGATA -5' ssODN

----CCGGCAAGATGTGCTCTCTGGGGTTGTTCCCTCCTCCACCGCCTCGGGGTAAAGTCACCCTATATGAGCACAATAACGAGCTGGTGACGGGCAGTAGCTAT---- C>A
```

Figure 3.2.3: **Generation of the C>A point mutation in exon 1 of *AAAS* gene.** (a) CRISPR-Cas9 assisted gene targeting strategy. Cells were co-transfected with Cas9, crRNA, tracrRNA, and a ssODN containing the C>A mutation to recover clones heterozygous for the C>A mutation (C>A/+). The experiment was repeated in heterozygous (C>A/+) cells to recover the mutation in both alleles (C>A/C>A). "X" refers to homologous recombination; C>A mutation marked red; CRISPR positions indicated with red arrows (b) Sequence of wild-type *AAAS* exon 1 showing the CRISPR site in red (PAM sequence is underlined), the ssODN containing the C>A mutation (antisense strand; G>T change in green), and the C>A mutation (sense strand; C>A change in green) in mutated allele.

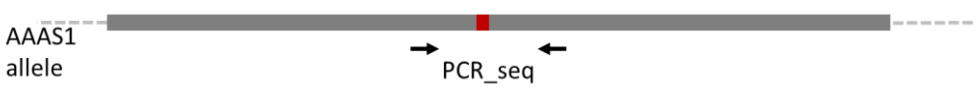
Generation of heterozygotes:

Analysis of 48 colonies by PCR amplification and sequencing revealed that 3 clones contained heterozygous C>A point mutation at the target site (Figure 3.2.4). NHEJ indels but with no C>A mutation was observed in about 14 clones, and almost half of the clones (46%) were wild-type (Table 3.2.1) which could have suffered DNA damage but were repaired correctly. These results support that NHEJ indels are the more common events in cellular DNA repair, while HDR is rare.

Table 3.2.1: Summary of the genotypes of different clones obtained in the first round of transfection of wild-type hiPS cells with the two different amounts of Cas9

No. of colonies screened after 1st transfection	**Amount of Cas9 protein; manufacturer**	**Heterozygote (C>A/+)**	**Homozygote (C>A/C>A)**	**NHEJ indels (no C>A mutation)**	**Wild-type (+/+)**	**No/ poor quality sequence**
48	12µg; Feldan	3 (6%)	0 (0%)	14 (29%)	22 (46%)	9 (19%)
96	18µg; Thermofisher	0 (0%)	0 (0%)	59 (62%)	30 (31%)	7 (7%)

(a)



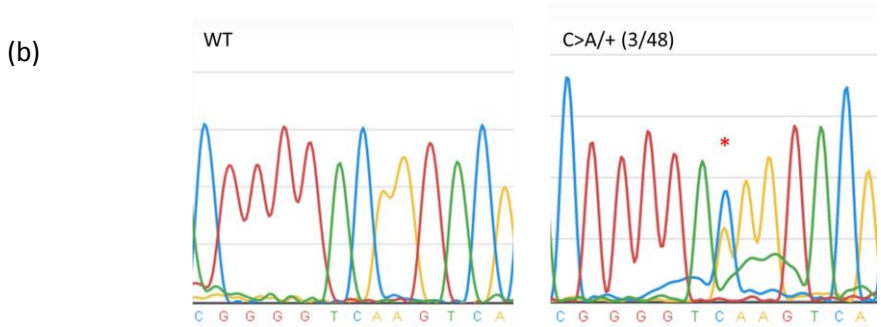


Figure 3.2.4: **Screening for the C>A point mutation in *AAAS* exon 1 after the first round of transfection.** (a) *AAAS* exon 1 allele showing PCR sequencing across the target C>A point mutation (sequencing primers indicated with black arrows). (b) Sequence traces of PCR products from representative wild-type (WT) and heterozygous (C>A/+) cell lines. The position of the C>A mutation is marked with a red asterisk.

On the other hand, clones from the other experiment in which cells were transfected with greater amount (18µg) of Cas9 nuclease were also analyzed. Screening 96 colonies revealed that none harbored the desired C>A point mutation in *AAAS* exon 1 either in the homozygous or heterozygous state (Table 3.2.1). Many of the clones (62%) had multiple sequence traces downstream of the DSB site indicative of NHEJ indels in either or both alleles. (Figure 3.2.5). Besides several colonies were wild-type containing no mutations. These results suggested that higher amounts of Cas9 led to more frequent NHEJ indels and less frequent HDR.

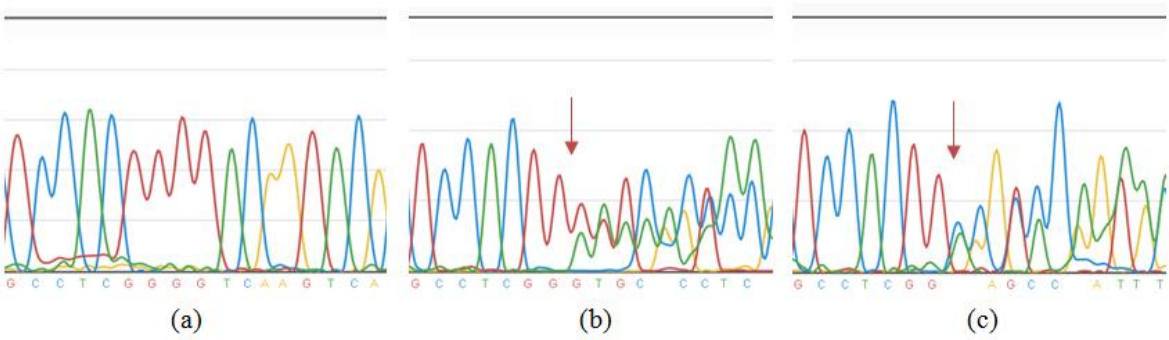


Figure 3.2.5: **Screening of *AAAS* exon 1 around DSB site showing indels.** (a) wild-type sequence, (b) and (c) show indels in one and both alleles respectively, with DSB sites indicated by arrows.

Control experiments with PAM mutations in ssODN:

Control experiments were to be conducted with ssODN containing a silent mutation in the PAM site and the C>A mutation. Mutations in the PAM site are expected to abolish re-cleavage

of the locus following incorporation of the ssODN by HDR. Thus, this experiment was conducted to test if PAM site mutations increased the rate of HDR. In exon 1 of AAAS gene, there were frequent consensus splice donor sites along the sequence, for which a silent mutation, for instance, a substitution mutation from a pyrimidine to a purine base in the PAM was likely to give rise to a splice donor site which would result in unexpected mutations and splicing in the gene. Thus, a missense mutation in the PAM sequence was done instead. This was, however, not going to affect the cell models as it was a control experiment to only test if PAM site mutations increased the rate of HDR. For this part of the experiment, it was observed that, for the reduced amount of Cas9 (12µg), there were less clones containing NHEJ indels (25/96) and more were wild-type (52/96) clones (Table 3.2.2). However, two of the clones contained heterozygous C>A and PAM mutations which agreed with the previous observations of increased HDR frequency using reduced amounts of Cas9. Contrarily, when greater amount of Cas9 was used, there were more NHEJ indel events (41/96), like earlier, however none of the clones contained the control (C>A and PAM) mutations.

Table 3.2.2: Summary of the genotypes of clones obtained upon targeting cells with two different amounts of Cas9 protein using control ssODN containing an additional PAM mutation besides the target C>A point mutation.

No. of colonies screened	**Amount of Cas9 protein; manufacturer**	**Heterozygote (C>A_PAM/+)**	**Homozygote (C>A_PAM/ C>A_PAM)**	**NHEJ indels**	**Wild-type (+/+)**	**No/ poor quality sequence**
96	12µg; Feldan	2	0	25	52	17
96	18µg; Thermofisher	0	0	41	24	31

The results indicate that more NHEJ damage, but less HDR are observed in the presence of higher amounts of Cas9 protein. This can be explained by re-cleavage of the modified locus and secondary damage to the locus. This higher frequency of NHEJ damage was probably demonstrated in the T7E1 assay as a higher activity. On the contrary, less amount of Cas9 protein exhibited decreased activity and frequency of NHEJ indels. This can be explained in that lesser amounts of Cas9 protein and sgRNA are likely to be degraded early in cells, lowering the effective concentration of the Cas9-RNP complex and its duration in cells (Wu et al. 2014). Adding mutations in the PAM unexpectedly did not improve rates of HDR. This could be explained by poorer incorporation of the ssODN oligo when more mismatches to the target

locus are present. The different manufacturers could also explain the difference in the Cas9 activities which can, however, be determined by performing experiments keeping Cas9 amounts and all the other factors, along with amount of sgRNA constant. Besides, the high frequency of NHEJ events can be prevented by inhibiting the DNA ligase IV which plays the major role in ligating Cas9-cleaved DNA strands in the NHEJ repair pathway. Scr7, an inhibitor of the DNA ligase IV can be used to block the NHEJ repair pathway forcing the damaged DNA to use ssODN to repair itself by HDR (Maruyama et al. 2015). This way, by inhibiting NHEJ, efficiency of HDR events can be enhanced eventually increasing the rates of HDR mediated precise genome editing with desired mutations.

Generation of homozygotes:

To obtain homozygotes, a second round of transfection was performed on two of the selected heterozygotes using same conditions as in the first round, that is 12µg of Cas9 (Feldan), 8µg of cr:tracrRNA, and 500pmoles of ssODN. 192 colonies were analyzed, with 96 colonies being screened by PCR and sequencing for each of the two heterozygotes transfected.

A total of 3 homozygotes were obtained out of the 192 colonies screened, which contained the C>A mutation in both the alleles following the second transfection (Figure 3.2.6). NHEJ indels were observed in 49 clones and almost half of the colonies were not modified, that is, they were heterozygotes, as were initially started with. Interestingly, two clones were observed to have reverted back from heterozygote to wild-type with both the alleles having the wild-type base-C in the position of interest. This suggests that gene conversion took place whereby the mutant DNA strand was cleaved again by the Cas9-RNP and repaired by the wild-type allele as the template, thus converting heterozygous mutant cells back to wild-type. However, it is possible that these two clones represent low level contamination by wild-type cells. A considerable number of colonies (47/192) had no or poor quality sequence which also means that the results likely underestimate the efficiency of HDR. All results are tabulated in Table 3.2.3.

Table 3.2.3: Summary of the genotypes of different clones obtained after the second round of transfection of heterozygote clones

No of colonies screened after 2nd transfection of heterozygotes	**Homozygote (C>A/C>A)**	**NHEJ indels**	**No mutation (C>A/+)**	**WT (+/+)**	**No/ poor quality sequence**
192	3 (2%)	49 (25.5%)	91 (47%)	2 (1%)	47 (24.5%)

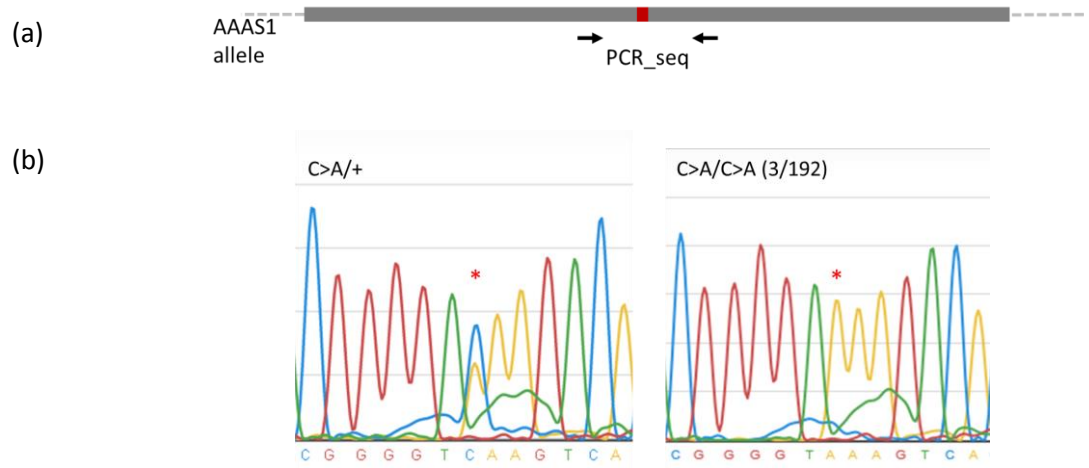


Figure 3.2.6: **Screening for the C>A point mutation in *AAAS* exon 1 after the second round of transfection.** (a) *AAAS* exon 1 allele showing PCR sequencing across the target C>A point mutation (sequencing primers indicated with black arrows). (b) Sequence traces of PCR products from representative heterozygous (C>A/+) and homozygous (C>A/C>A) cell lines. The position of the C>A mutation is marked with a red asterisk.

3.3 Expression of ALADIN protein by *AAAS* mutant clones

For the *AAAS* exon 2 knockout and *AAAS* exon 1 point mutation clones, the expression of ALADIN protein was determined by Western blot. Using rabbit polyclonal anti-AAAS antibodies, ALADIN expression was observed in wild-type and heterozygote clones, but not in the homozygotes. This observation validates that ALADIN expression was completely knocked out in these CRISPR-Cas9 gene-edited homozygotes (Figure 3.3). The anti-AAAS antibody used in this experiment recognizes the epitope present in three of the isoforms of ALADIN – Q9NRG9-V1 (60kDa), Q9NRG9-V2 (56kDa), and F8VZ44 (46kDa). Presumably, the band in the assay is not the isoform F8VZ44 because initiation of translation of this isoform happens in exon 3 and thus it should be expressed regardless of any mutations in exons 1 and 2. However, the protein observed in the assay was knocked out only in the homozygotes; it is possible that the F8VZ44 isoform is not expressed at all in undifferentiated cells, as in this case. Therefore, it can be assumed that the band could correspond to either V1 or V2 of the Q9NRG9 isoform. It can be further characterized as to which of these isoforms is being expressed by performing RT-PCR of the mRNA to its cDNA and validating which alternative transcript is being produced in these iPS cells.

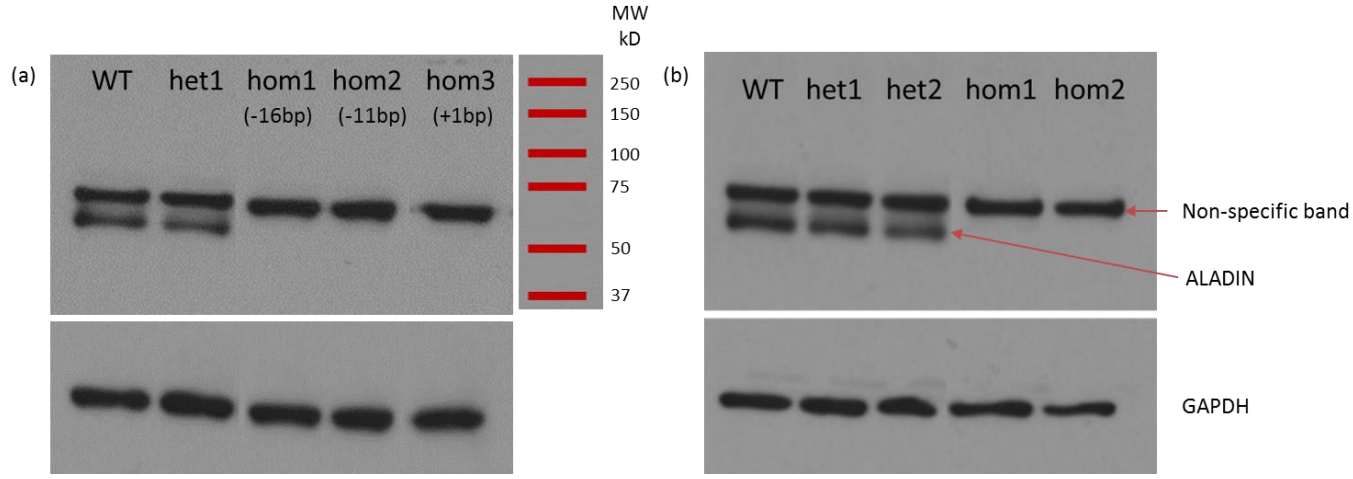


Figure 3.3: Western blot of (a) *AAAS* exon 2 knockout clones with the respective indels of the non-targeted allele indicated in brackets and (b) *AAAS* exon 1 43C>A point mutation clones. Wild-type (WT) and heterozygotes (het) express the ALADIN protein, which is however knocked out in homozygotes (hom). GAPDH was used as a loading control.

An additional slower-migrating (upper) band was observed in all the clones, which is probably due to cross-reaction to an AAAS paralog or to an unrelated protein. The multiple isoforms of ALADIN are formed by skipping of exon 6 to form 56kDa isoform, or translation

initiation in exon 3 to form the shorter 46kDa isoform. Mutations in exons 1 or exon 2, generated in this study, will result in a premature stop codon within exon 3. The mutated protein for *AAAS* exon 1 point mutation is 57 amino acids long weighing approximately 6kDa, whereas the lengths for the *AAAS* exon 2 knockout clones were 87 (for 16bp deletion), 81 (11bp deletion) and 85 (1bp insertion) amino acids weighing around 10kDa, 9kDa, and 10kDa respectively. Thus, the upper band cannot be explained by any isoform of ALADIN. BLAST results of the Q9NRG9 isoform against the human genome indicated minor identities to a processed pseudogene on chromosome 6 (refer Appendix, Figure A15). This pseudogene, however, does not encode any functional protein, and thus, rules out the possibility of a AAAS paralog. Therefore, the upper band could represent non-specific binding to an unrelated protein in the human genome, an epitope of which might cross-react with the polyclonal antibody. To resolve the origin of the upper band on the Western blot, monoclonal antibodies could be used that recognize a unique epitope of ALADIN, different to that recognized by the polyclonal antibody (Proteintech) used in this experiment (refer Appendix A14).

The observations made in this study suggest that *AAAS* gene expression and protein function was completely knocked out. The truncated protein was not detectable in the assay, either because it was too small to be detected, or due to the probable decay of the mutated transcript or the degradation of the truncated protein within the cells. In any case, it is truncated and therefore not functional.