OMTN, Volume *24*

Supplemental information

Base editing strategy for insertion of the A673T mutation in the APP gene to prevent the development of AD *in vitro*

Antoine Guyon, Joël Rousseau, Francis-Gabriel Bégin, Tom Bertin, Gabriel Lamothe, and Jacques P. Tremblay

SUPPLEMENTAL RESULTS:

Table S1: List of sgRNA used in the study.

SpCas9 VQR/EQR sgRNAs targeting sequences of 17 to 22 nucleotides tested:	
sgRNA 22bp	ATTCTGCATCCATCTTCACTTC
sgRNA 21bp	TTCTGCATCCATCTTCACTTC
sgRNA 20bp	TCTGCATCCATCTTCACTTC
sgRNA 19bp	CTGCATCCATCTTCACTTC
sgRNA 18bp	TGCATCCATCTTCACTTC
sgRNA 17bp	GCATCCATCTTCACTTC
SaCas9 sgRNAs tested:	
sgRNA 22bp	ATTCTGCATCCATCTTCACTTC
sgRNA 21bp	TTCTGCATCCATCTTCACTTC
sgRNA 20bp	TCTGCATCCATCTTCACTTC
sgRNA 19bp	CTGCATCCATCTTCACTTC
sgRNA 18bp	TGCATCCATCTTCACTTC
sgRNA 17bp	GCATCCATCTTCACTTC

Table S2: Example of Deep-Sequencing analysis. Indels calculated with the script2 of Figure S7.

	Target-AID-SpCas9nVQR 19	BE3_SpCas9nVQR 19
Total reads	100%	100%
Wild-Type	39,3	66,0
C1	4,8	1,0
C2	3,8	0,3
C3	0,1	3,6
C4	0,4	0,0
C5	0,8	0,3
C1+C2	36,3	0,0
C1+C3	0,3	2,1
C1+C4	0,1	0,0
C1+C5	0,2	0,0
C2+C3	0,1	0,2
C2+C4	0,1	0,0
C2+C5	0,3	0,0
C3+C4	0,0	0,2
C3+C5	0,0	0,2
C4+C5	0,4	0,0
C1+C2+C3	2,3	1,0
C1+C2+C4	0,8	0,0
C1+C2+C5	5,5	0,0
C2+C4+C5	1,8	0,0
C2+C3+C5	0,0	0,0
C1+C2+C3+C4	0,6	0,1
C2+C3+C4+C5	0,3	0,2
C1+C3+C4+C5	0,0	0,0
C1+C2+C4+C5	0,5	0,1
C1+C2+C3+C5	0,6	0,0
C1+C2+C3+C4+C5	0,4	0,2
Global C2	53,3	2,5
Global C2 (excluding C3 and C4))	42,6	0,4
Indels	1,4	2,6
Total	75,60%	75,70%
Mis-sequencing	24,40%	24,30%

Table S3: Percentage of reduction of Aβ40 and Aβ42 peptides secretion induced by the addition of the A673T mutation to wild type APP gene or to an APP gene containing the London mutation. Last column shows the percentage of reduction induced by the addition of E674K and A673T (C1+C2) mutations to wild type APP gene. The addition of the A673T mutation reduced the production of Aβ40 and Aβ42 peptides in all 3 situations. Experiment performed in SH-SY5Y transfected with plasmids containing an APP695 cDNA. Results from Guyon et al. 2020 (DOI: 10.1371/journal.pone.0237122)

FAD mutation	APPWT	APPV717I	APPC1+C2
Abeta42 Decrease (%)	-46	-65	-53
Abeta40 Decrease (%)	-63	-81	-44

Figure S1.

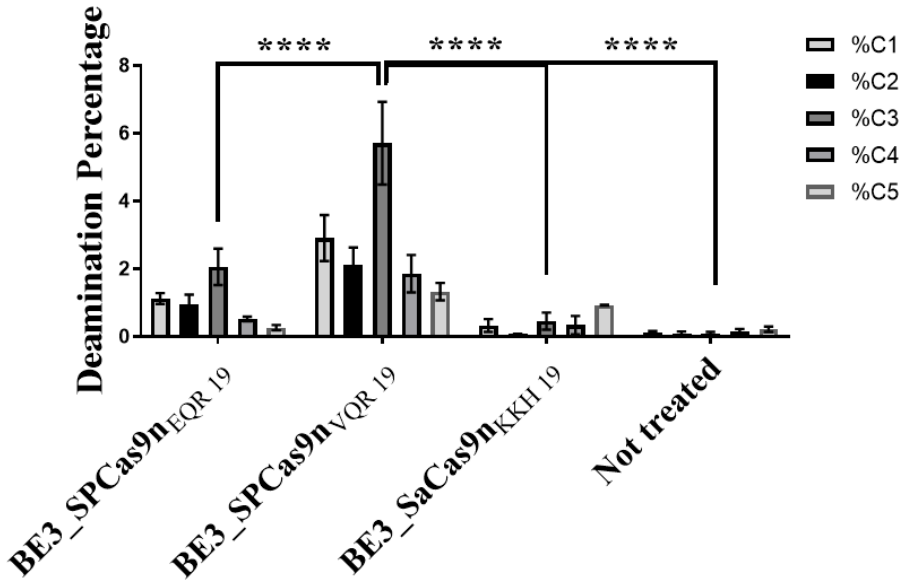

Figure S1: **Percentages of cytidine deamination produced by various enzymes and sgRNAs.** BE3_SpCas9nEQR, BE3_SpCas9nVQR, BE3_SaCas9nKKH enzymes tested in SH-SY5Y cells. The figure illustrates the means +/- SEM (n=4). Statistical test: two-way ANOVA Tukey's multiple comparisons test (n=4). P value style **** p<0.0001.

Figure S2.

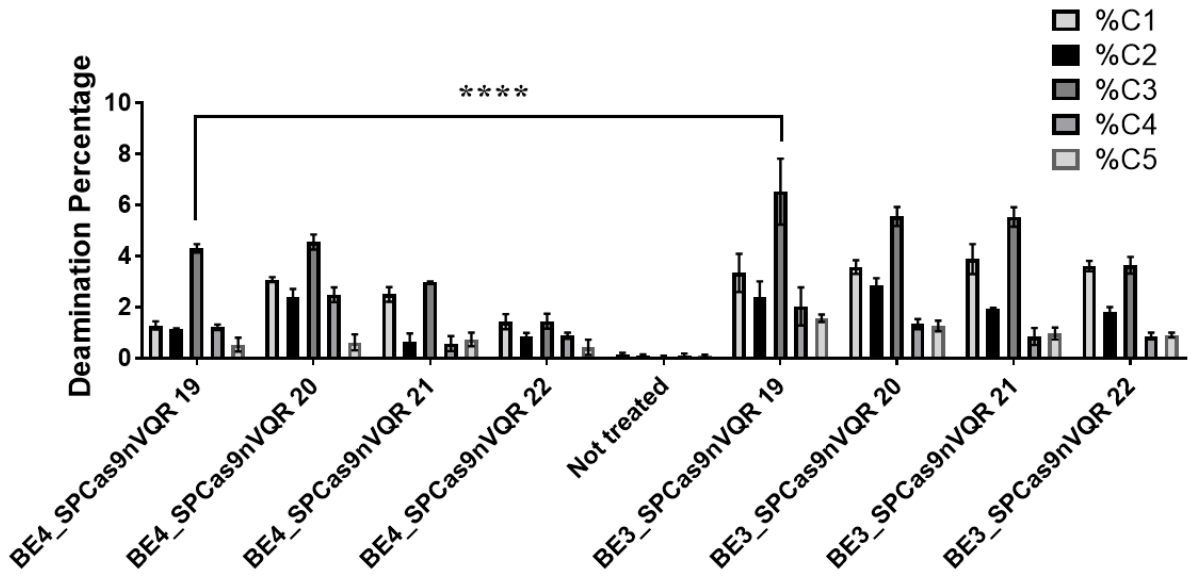

Figure S2: **Percentages of cytidine deamination produced by various enzymes and sgRNAs.** BE4_SpCas9nVQR and BE3_SpCas9nVQR enzymes test in SH-SY5Y cells. The figure illustrates the means +/- SEM (n=3). Statistical test: two-way ANOVA Tukey's multiple comparisons test (n=3). P value style **** p<0.0001.

Figure S3.

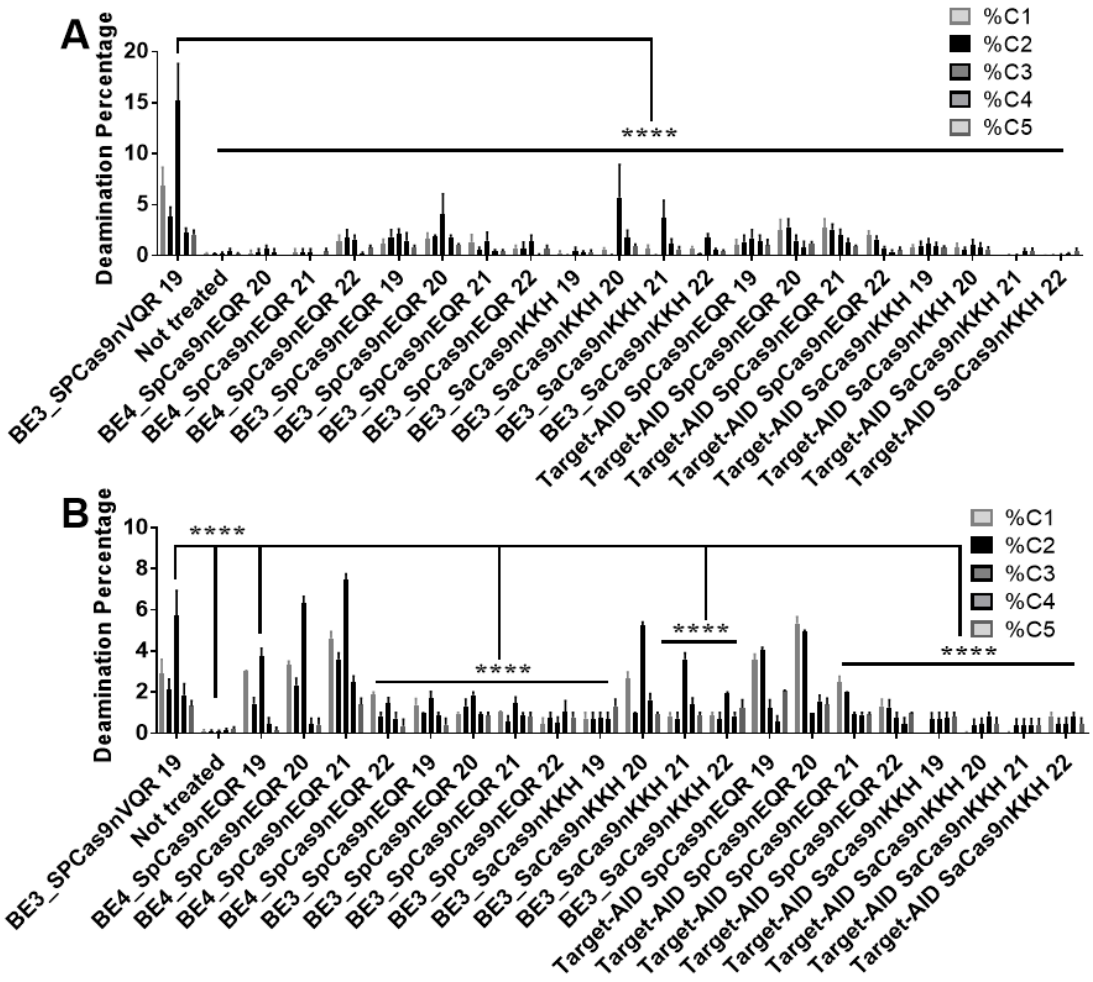

Figure S3: **Percentages of cytosine deamination produced by BE3_SpCas9nVQR, BE4_SpCas9nEQR, BE3_SpCas9nEQR, BE3_SaCas9nKKH, Target-AID_SpCas9nEQR, Target-AID_SaCas9nKKH enzymes.** In **A**, test in HEK293T cells. In **B**, test in SH-SY5Y. The figure illustrates the means +/- SEM (n=3). Statistical test: two-way ANOVA Tukey's multiple comparisons test (n=3). P value style **** $p<0.0001$.

Figure S4.

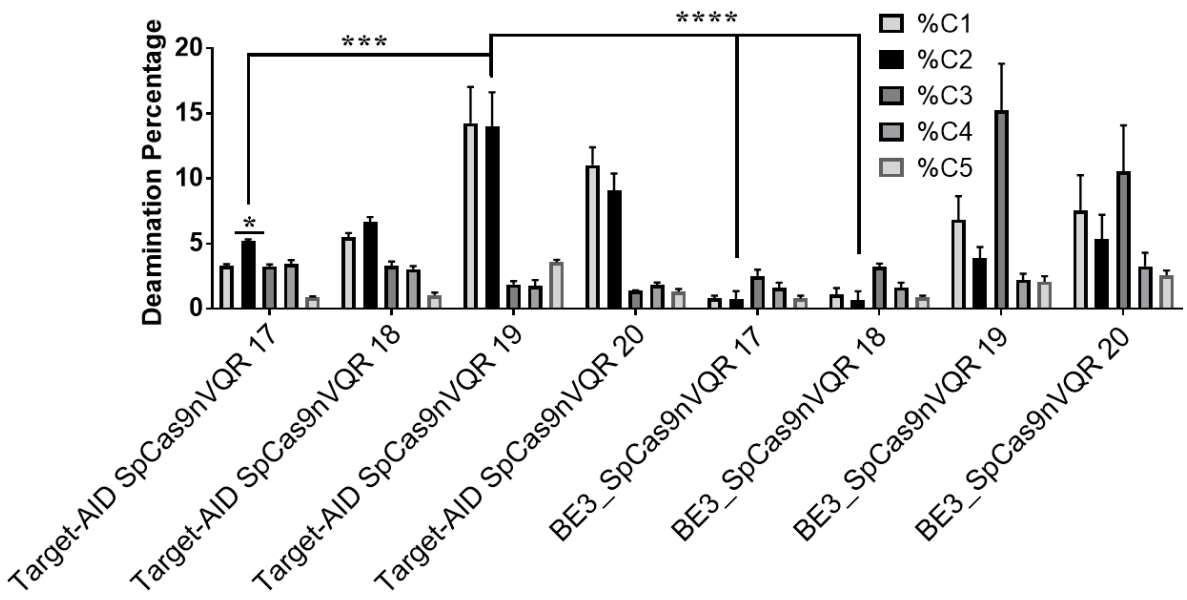

Figure S4: **Deamination efficiencies using various Cas9n-deaminases and sgRNAs targeting various numbers of nucleotides**. The difference of deamination in HEK293T cells of cytidines C1 to C5 produced by the Target-AID-SpCas9nVQR and BE3_SpCas9nVQR enzymes and two copies of a sgRNA targeting 17 to 20 nucleotides. The figure illustrates the means +/- SEM (n=3). Statistical test: two-way ANOVA Tukey's multiple comparisons test (n=3). P value style *** p<0.0002, **** p<0.0001.

Figure S5

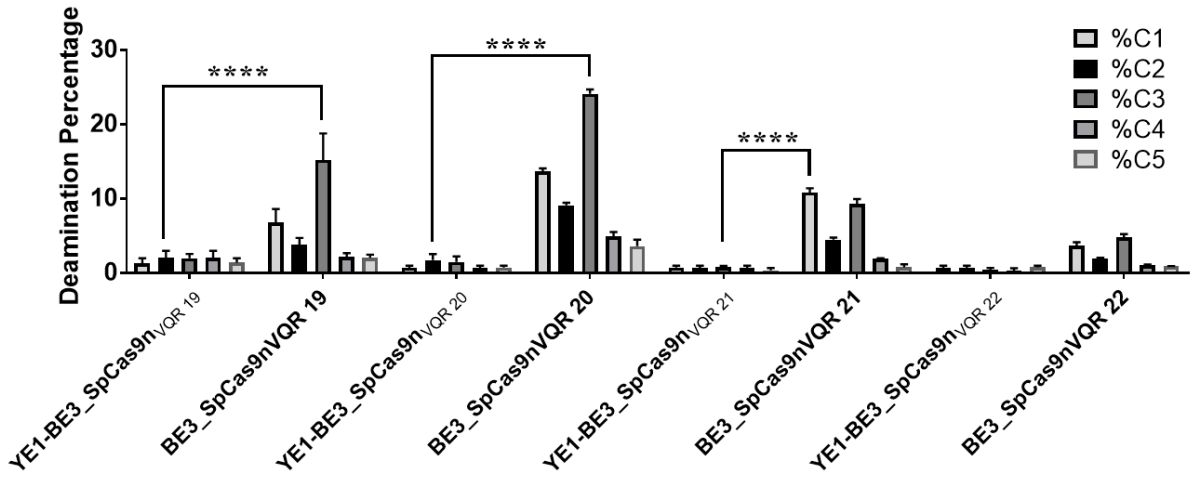

Figure S5: **Deamination efficiencies using various Cas9n-deaminases and sgRNAs targeting various numbers of nucleotides**. Difference between YE1-BE3_SpCas9nVQR and BE3_SpCas9nVQR in HEK293T cells. The figure illustrates the means +/- SEM (n=4). Statistical test: two-way ANOVA Tukey's multiple comparisons test (n=4). P value style **** p<0.0001.

Figure S6:

Sequence	PAM	Score	Gene	Locus
CTGCATCCATCTTCACTTC	AGAG	100.0	APP (ENSG00000142192)	chr21:+25897633
CTGAAGCCATCTTCACTTC	GGAG	1.4		chr5:-74133250
CTGCCTCCATCTTCACATG	TGAG	0.5		chr11:+70758797
CTGCTTCCAACTTCACTTT	GGAG	0.5	SPOCK2 (ENSG00000107742)	chr10:-72059638
CAGGATCCATCTTAACTTC	TGAG	0.4		chr12:-47654519
CTGCTTCCATCTTCTGTTC	AGAG	0.4		chr3:-36729648
CTGCATCCTTCTCCACTTG	GGAG	0.4		chr8:-67603274
CTGAATCAATCTCCACTTC	AGAG	0.4		chr12:-56572869
ATCCATGCATCTTCACTTC	AGAG	0.3		chr11:-20359838
CTGCCCCCACCTTCACTTC	TGAG	0.3		chr9:+85398673
CTGCATCCATCTCTCCTTC	AGAG	0.3		chr3:+176747607
CTATTTCCATCTTCACTTC	AGAG	0.3		chr9:-21442550
ATGTATCCATCTTCACTGT	TGAG	0.1		chr14:-35505959
TTTCATCCATCTCCACTTT	AGAG	0.1		chr5:-44542656
TTTCATCCATCTTAACTAC	AGAG	0.1		chr10:-120808163
ATCCATCCACCTTCACTTG	TGAG	0.1		chr4:+10179124

Figure S6: **Off target analysis performed with Crispr.mit.edu algorithm (Hsu et al. 2014)**

Figure S7:

Script1.sh:

```
(echo -n "Sample#n#"; for i in $(cat list);do echo -n "$i#" ;done; echo; for j in $(ls *.bam);do echo -n "$j#"; samtools view $j |wc -l| awk '{printf $1 "\t"}' ;for i in $(cat list);do samtools view $j | grep -c $i | awk '{printf $1 "\t"}'; done; echo; done)|sed s/#/\\t/g
```

List associated with Script1.sh:

ATTCTGCATCCATCT
ATTTTGCATCCATCT
ATTCTGTATCCATCT
ATTCTGCATTCATCT
ATTCTGCATCTATCT
ATTCTGCATCCATTT
ATTTTGTATCCATCT
ATTTTGCATTCATCT
ATTTTGCATCTATCT
ATTTTGCATCCATTT
ATTCTGTATTCATCT
ATTCTGTATCTATCT
ATTCTGTATCCATTT
ATTCTGCATTTATCT
ATTCTGCATTCATTT
ATTCTGCATCTATCT
ATTCTGCATCTATTT
ATTCTGCATCCATTT
ATTTTGTATTCATCT
ATTTTGTATCTATCT
ATTTTGTATCCATTT
ATTCTGTATTTATCT
ATTCTGTATCTATTT
ATTCTGTATTCATTT
ATTTTGTATTTATCT
ATTCTGTATTTATTT
ATTTTGCATTTATTT
ATTTTGTATCTATTT
ATTTTGTATTCATTT
ATTTTGTATTTATCT
ATTTTGTATTTATTT

Script2.sh:

Part1:

```
for i in $(ls *.bam);do samtools view $i |grep -o TGAGTCAT.*CGTCTTGA|sort|uniq -c| awk '{print $2 "\t" $1}' > $i.count;done
```

Part2:

```
ls *.count|perl Parse2Table.pl 1 > 20170629_Counts.tsv
```

Figure S7: **Deep sequencing file analysis script.** Script1 allows you to find a specific sequence in a NGS file. The sequences are inserted into an attached "list" file. Script 2 allows to see all sequences between two specific areas of the genome by modifying the highlighted bases. Commonly used to calculate indels. The same results can be achieved with CRISPResso2 but with less details and research versatility.