Supplementary Information for

Expansion and Precise CRISPR-Cas9 Gene Repair of Autologous T Memory Stem Cells from Patients with T Cell Immunodeficiencies

Xun Li [1,#], Van Trung Chu [1,2], Christine Kocks [1,2], and Klaus Rajewsky [1,*]

[1] Immune Regulation and Cancer, Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

[2] Genome Engineering and Disease Models, Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

This PDF file includes:

Figure S1
Figure S2
Figure S3
Figure S4
Figure S5
Figure S6
Figure S7
Figure S8

Figure S1. Two-part guide RNA (gRNA) hybridized to form a crRNA:tracr RNA complex. Schematic drawing of the RNA components of the Alt-R CRISPR-Cas9 System for directing Cas9 endonuclease to genomic targets (Integrated DNA Technologies, Inc. [IDT]). Alt-R CRISPR-Cas9 crRNAs are 36 nucleotide (nt) RNA oligos containing a (variable) 20 nt target-specific spacer region, along with a non-variable 16 nt domain complementary to tracrRNA. The 67 nt Alt-R CRISPR-Cas9 tracrRNA is shorter than the natural *S. pyogenes* tracrRNA and contains proprietary chemical modifications that confer increased nuclease resistance. The crRNA and tracrRNA molecules are annealed to each other and form a complex with Cas9 endonuclease enzymes.

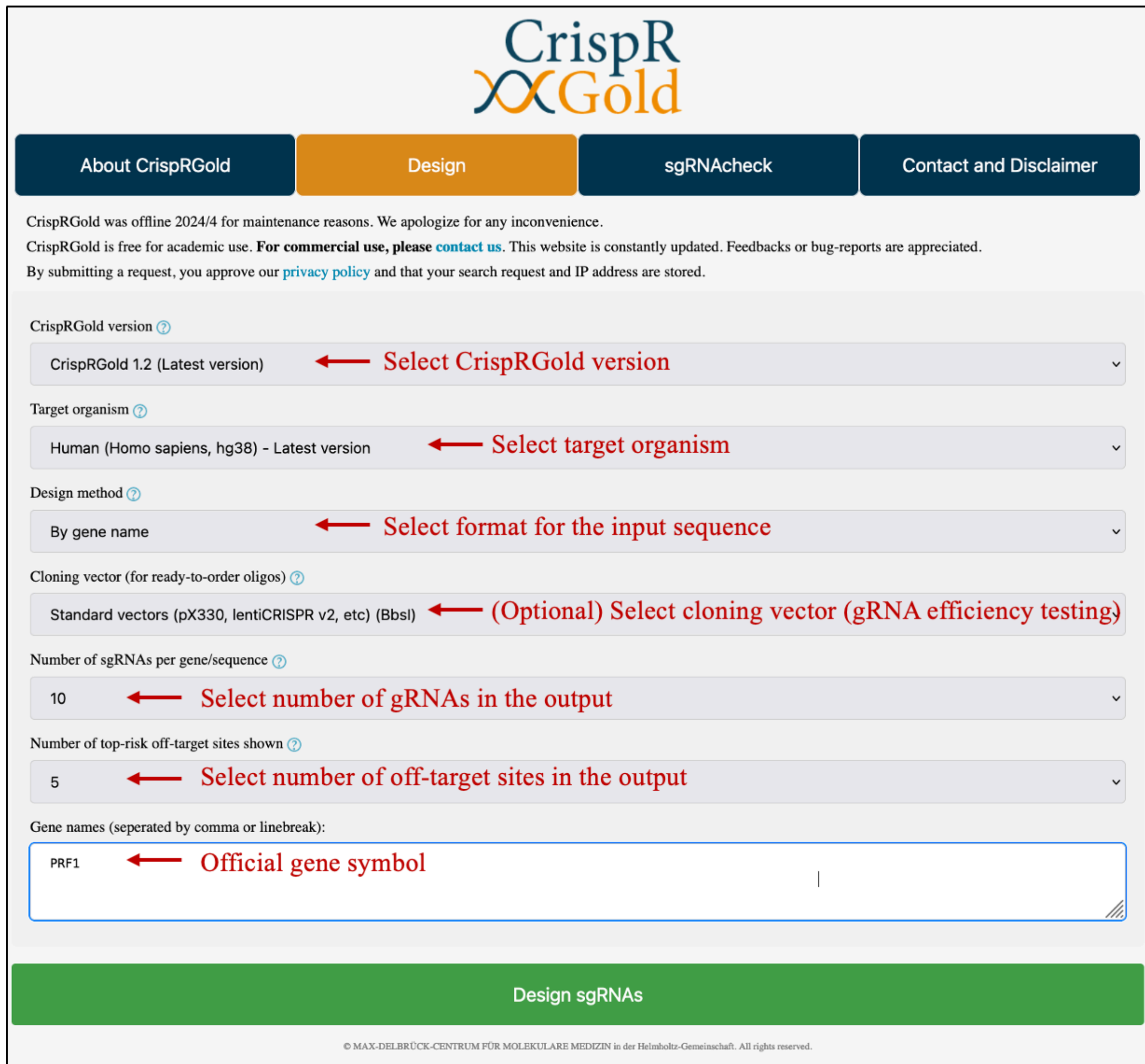
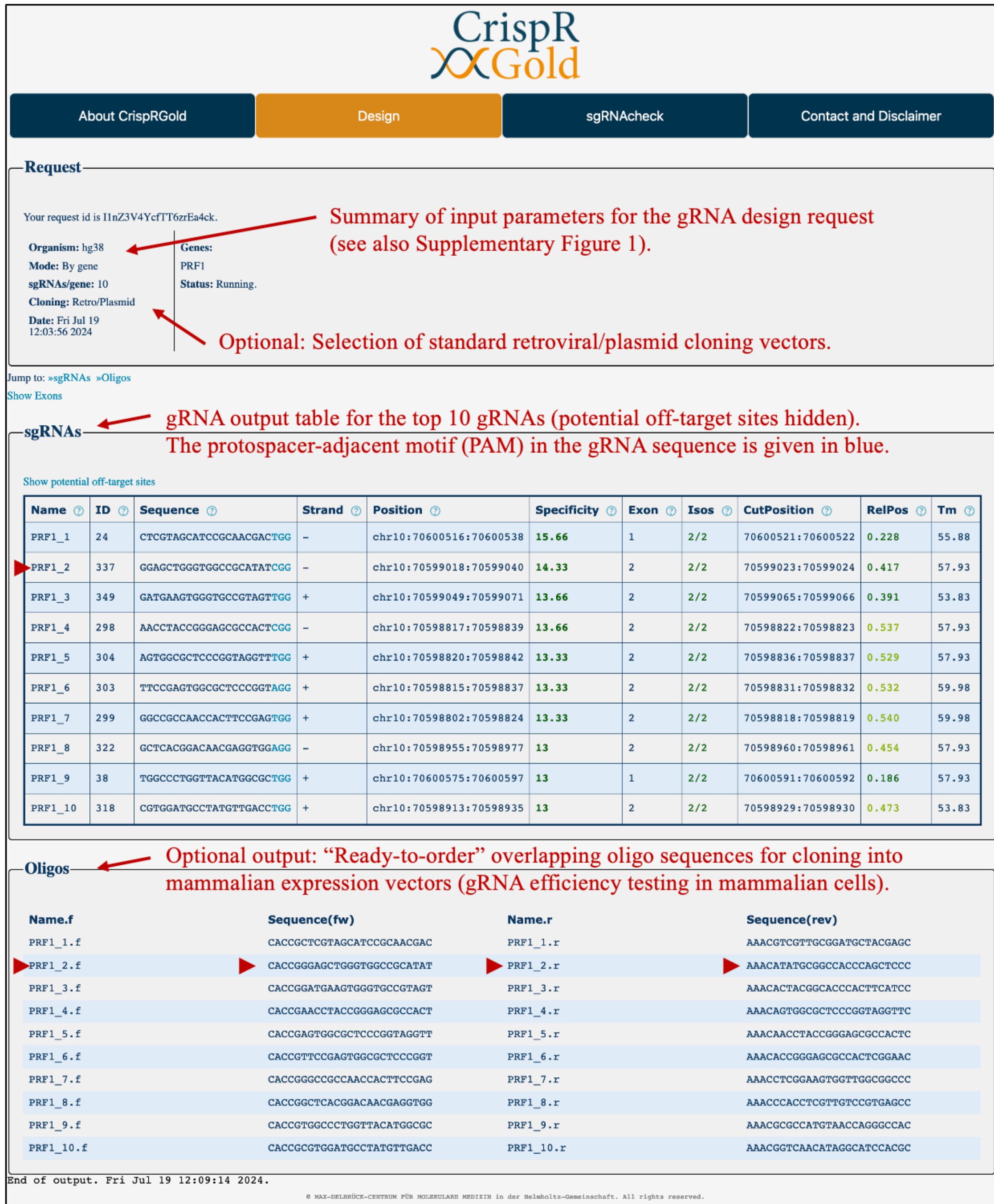


Figure S2. Guide RNA (gRNA) design with free CrispRGold software (Graf et al 2019). As an example we show the selection of input parameters for the human *PRF1* gene (National Center for Biotechnology Information (NCBI) Gene ID 5551). Explanatory labels are given in red next to the red arrows.



Name	ID	Sequence	Strand	Position	Specificity	Exon	Isos	CutPosition	RelPos	Tm
PRF1_1	24	CTCGTAGCATCCGCAACGACTGG	–	chr10:70600516:70600538	15.66	1	2/2	70600521:70600522	0.228	55.88
PRF1_2	337	GGAGCTGGGTGGCCGCATATCGG	–	chr10:70599018:70599040	14.33	2	2/2	70599023:70599024	0.417	57.93
PRF1_3	349	GATGAAGTGGGTGCCGTAGTTGG	+	chr10:70599049:70599071	13.66	2	2/2	70599065:70599066	0.391	53.83
PRF1_4	298	AACCTACCGGGAGCGCCACTCGG	–	chr10:70598817:70598839	13.66	2	2/2	70598822:70598823	0.537	57.93
PRF1_5	304	AGTGGCGCTCCCGGTAGGTTTGG	+	chr10:70598820:70598842	13.33	2	2/2	70598836:70598837	0.529	57.93
PRF1_6	303	TTCCGAGTGGCGCTCCCGGTAGG	+	chr10:70598815:70598837	13.33	2	2/2	70598831:70598832	0.532	59.98
PRF1_7	299	GGCCGCCAACCACTTCCGAGTGG	+	chr10:70598802:70598824	13.33	2	2/2	70598818:70598819	0.540	59.98
PRF1_8	322	GCTCACGGACAACGAGGTGGAGG	–	chr10:70598955:70598977	13	2	2/2	70598960:70598961	0.454	57.93
PRF1_9	38	TGGCCCTGGTTACATGGCGCTGG	+	chr10:70600575:70600597	13	1	2/2	70600591:70600592	0.186	57.93
PRF1_10	318	CGTGGATGCCTATGTTGACCTGG	+	chr10:70598913:70598935	13	2	2/2	70598929:70598930	0.473	53.83

Name.f	Sequence(fw)	Name.r	Sequence(rev)
PRF1_1.f	CACCGCTCGTAGCATCCGCAACGAC	PRF1_1.r	AAACGTCGTTGCGGATGCTACGAGC
PRF1_2.f	CACCGGGAGCTGGGTGGCCGCATAT	PRF1_2.r	AAACATATGCGGCCACCCAGCTCCC
PRF1_3.f	CACCGGATGAAGTGGGTGCCGTAGT	PRF1_3.r	AAACACTACGGCACCCACTTCATCC
PRF1_4.f	CACCGAACCTACCGGGAGCGCCACT	PRF1_4.r	AAACAGTGGCGCTCCCGGTAGGTTC
PRF1_5.f	CACCGAGTGGCGCTCCCGGTAGGTT	PRF1_5.r	AAACAACCTACCGGGAGCGCCACTC
PRF1_6.f	CACCGTTCCGAGTGGCGCTCCCGGT	PRF1_6.r	AAACACCGGGAGCGCCACTCGGAAC
PRF1_7.f	CACCGGGCCGCCAACCACTTCCGAG	PRF1_7.r	AAACCTCGGAAGTGGTTGGCGGCCC
PRF1_8.f	CACCGGCTCACGGACAACGAGGTGG	PRF1_8.r	AAACCCACCTCGTTGTCCGTGAGCC
PRF1_9.f	CACCGTGGCCCTGGTTACATGGCGC	PRF1_9.r	AAACGCGCCATGTAACCAGGGCCAC
PRF1_10.f	CACCGCGTGGATGCCTATGTTGACC	PRF1_10.r	AAACGGTCAACATAGGCATCCACGC

Figure S3. Guide RNA (gRNA) design with free CrispRGold software (Graf et al 2019). Upper panel: Summary of input parameters (see also Figure S1.) Middle panel: Output of gRNA design request. Potential off-target sites are hidden. gRNAs are ranked by the CrispRGold Specificity score. For targeting exon 3 of the *PRF1* gene, the gRNAs must be located within the coding part of exon 3 sequence (this corresponds to Coding Exon 2 ("Exon 2") in the CrispRGold software output; genomic coordinates 70,598,053 to 70,599,181). The Specificity score ideally should be ≥ 12. Thus, the top 8 gRNAs positioned in Exon 2 in the output shown above are candidates to be chosen for gRNA efficiency testing. Lower panel: Optional output for "ready-to-order" oligos for cloning into mammalian sgRNA expression vectors (see also Figure S4). Explanatory labels are given in red next to the red arrows. Red arrow heads point to the gRNA PRF1.2 (see Table 1) and corresponding cloning oligos used in our study (Li et al 2024).

sgRNAs

gRNA output table for the top 5 gRNAs. In this view, the top 5 potential off-target sites are expanded. We recommend to examine 10 to 15 potential off-target sites per gRNA and to choose gRNAs with a predicted low risk for off-target activity.

Hide potential off-target sites

Name	ID	Sequence	Strand	Position	Specificity	Exon	Isos	CutPosition	RelPos	Tm
PRF1_1	24	CTCGTAGCATCCGCAACGACTGG **Top off-target sites** .T.........T...G...A.. C...G.........G...G.... ....t.c.....g.... C...G.........G...G.... AG.C...........A......	- **Strand** + - - - +	chr10:70600516:70600538 **Position** chr20:41619215:41619236 chr1:16700714:16700735 chr8:126448732:126448753 chr1_KI270714v1_random:26545:26566 chr7:132509128:132509149	15.66 **MutDist** 15 16 16 16 17	1 **Nmut** 3 4 3 4 4	2/2 **Risk** low low low low lim→0	70600521:70600522 **Annotation (if relevant)** Intergenic ESPNP.Intron Intergenic Intergenic	0.228	55.88
▶ PRF1_2	337	GGAGCTGGGTGGCCGCATATCGG **Top off-target sites** .T..........T....G.A.. ......A.A.........G.G.. ..A..-..............A. .T..-...G............. ........G...T.....CA..	- **Strand** - + + + -	chr10:70599018:70599040 **Position** chr5:5838464:5838485 chr15:62578684:62578705 chr15:21006768:21006788 chr14:102776927:102776947 chr16:28938707:28938728	14.33 **MutDist** 14 14 15 15 15	2 **Nmut** 3 3 3 3 3	2/2 **Risk** low low low low low	70599023:70599024 **Annotation (if relevant)** Intergenic TLN2.Intron;TLN2.RNA.Intron Intergenic Intergenic CD19.CDS	0.417	57.93
PRF1_3	349	GATGAAGTGGGTGCCGTAGTTGG **Top off-target sites** .....c.....a...c...... ....a..tc...... ....a..tc...... ....a..tc...... ....a..tc......	+ **Strand** + - - + +	chr10:70599049:70599071 **Position** chr11:20708369:20708390 chrX:85622236:85622257 chr10:98737387:98737408 chr6:130174516:130174537 chr19:53210659:53210680	13.66 **MutDist** 13 14 14 14 14	2 **Nmut** 3 3 3 3 3	2/2 **Risk** low low low low low	70599065:70599066 **Annotation (if relevant)** NELL1.Intron Intergenic HPSE2.Intron;HPSE2.RNA.Intron SAMD3.Intron;SAMD3.RNA.Intron Intergenic	0.391	53.83
PRF1_4	298	AACCTACCGGGAGCGCCACTCGG **Top off-target sites** .......A.....A.....G.. ....G..A.....T.....A.. g...................t...t.. T...-.A............... ...CT........T.....T..	- **Strand** + - + - +	chr10:70598817:70598839 **Position** chr4:72590398:72590420 chr18:23227929:23227950 chr12:72824153:72824175 chr9:2161948:2161968 chrX:153507108:153507129	13.66 **MutDist** 13 14 14 14 14	2 **Nmut** 3 3 3 3 3	2/2 **Risk** low low low low low	70598822:70598823 **Annotation (if relevant)** Intergenic CABLES1.Intron;CABLES1.RNA.Intron Intergenic SMARCA2.Intron BGN.CDS	0.537	57.93
PRF1_5	304	AGTGGCGCTCCCGGTAGGTTTGG **Top off-target sites** ....G..............G.A ....G..............-.. .G.........C.......C... .................C....GG.. ....T-....T...........	+ **Strand** - - + + -	chr10:70598820:70598842 **Position** chr5:180133138:180133159 chr5:180133139:180133159 chr18:26549393:26549414 chr9:127405050:127405072 chr22:33236766:33236786	13.33 **MutDist** 12 14 14 15 15	2 **Nmut** 2 1 3 3 3	2/2 **Risk** low low low low low	70598836:70598837 **Annotation (if relevant)** RASGEF1C.Intron RASGEF1C.Intron KCTD1.Intron;KCTD1.alt5UTR SLC2A8.Intron;SLC2A8.RNA.Intron LARGE.RNA.Intron	0.529	57.93

Figure S4. Output of CrispRGold (Graf et al. 2019) gRNA design request with off-target sites shown in expanded mode. gRNAs are ranked by the CrispRGold Specificity score. It is desirable to choose gRNAs with a predicted low risk for off-target activity such as the gRNAs shown here. Off-targets in intergenic regions may be less critical than off-targets in gene loci. Explanatory labels are given in red next to the red arrows. A red arrow head points to the gRNA PRF1.2 (Table 1) used in our study Li et al 2024.

5′-*CACCG*GGAGCTGGGTGGCCGCATAT-3′

|||||||||||||||||||||

3′-*C*CCTCGACCCACCGGCGTATA*CAAA*-5′

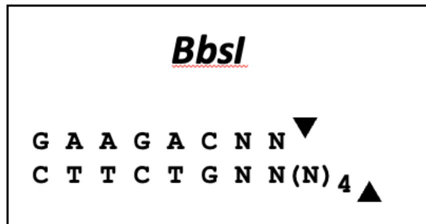


Figure S5. Cloning oligo nucleotides with *BbsI* sticky ends for gRNA standard vectors (pX330, MSCV retroviral vectors, etc.). As example we show annealed forward and reverse cloning oligos for the gRNA PRF1.2 (Table 1 and CrispRGold output shown in Figure S2, lower panel). The oligos must be phosphorylated at the 5'end before annealing. Bases colored in blue and set in italics correspond to cloning-compatible overhangs as configured by CrispRGold (Figure S2, lower panel). (Note that sticky ends from different *BbsI* sites may not be compatible.)

Figure S6. Analysis of the genome editing efficiency of the guide RNA PRF1.2 with the free web-based tool 'Inference of CRISPR edits' (ICE) (Conant et al 2022). Screenshots of the output of the ICE analysis tool (Syntego). Upper panel: Contributions of molecular species to the Sanger sequence in %. Lower panel: Sanger sequence traces with annotations. For a good guide RNA, the % of sequences with Indels should be above 60%. Explanatory labels are given in red next to the red arrows.

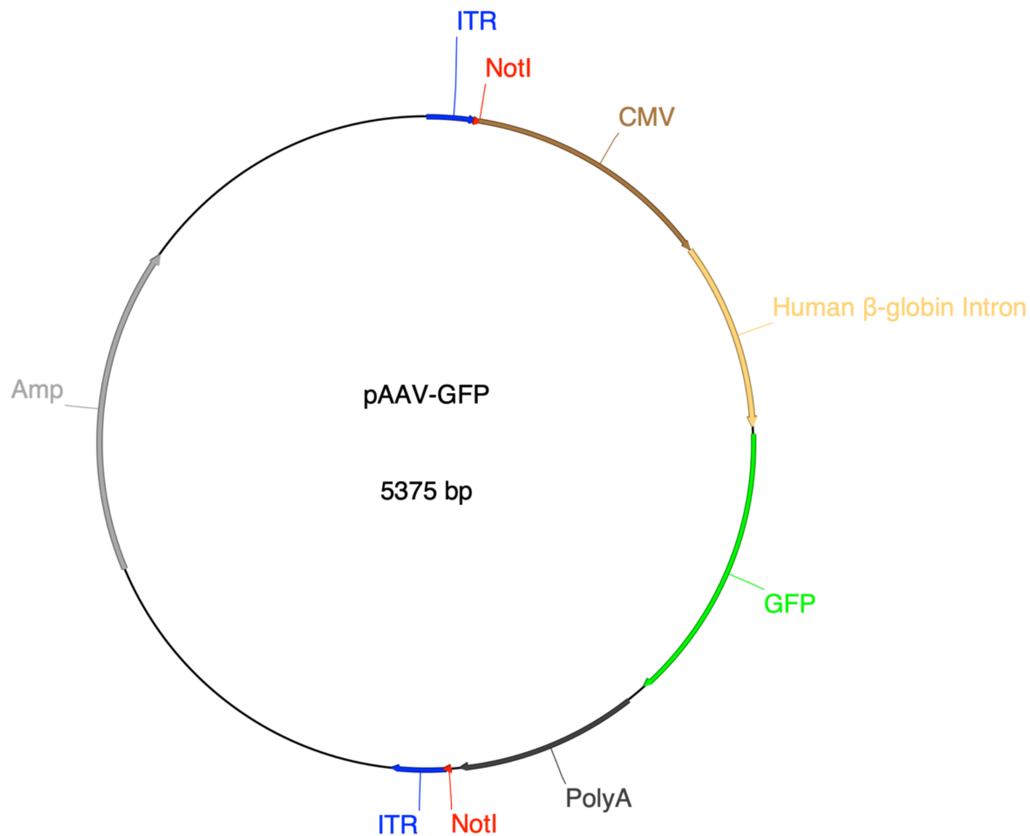


Figure S7. Plasmid map of the pAAV-GFP control vector. DNA donor templates for homology-directed repair (HDR) are cloned into this vector by 'one-step cloning' (Procedure part B.4). The pAAV-GFP vector is cut with NotI and the insert between the two NotI sites is replaced by the HDR DNA donor template (consisting of 5' and 3' homology arms flanking the codon-modified *PRF1* exon 3 or T2A-*PRF1*-cDNA). The intact pAAV-GFP control vector also serves as a positive control for AAV production (Procedure Part E). The packaging capacity of this vector is 4 to 4.6 kb (corresponding to the maximum fragment length in between the NotI sites).

pAAV-GFP Control Vector (Cell Biolabs Inc., cat. No. VPK-400)
https://www.cellbiolabs.com/paav-gfp-control-plasmid
Vector features: 1-130 Left inverted terminal repeat (ITR), 139-798 Cytomegalovirus (CMV) promoter, 806-1298 Human beta-globin intron, 1321-2061 Green Fluorescent Protein (GFP), 2120-2598 PolyA, 2638-2778 Right ITR, 3695-4555 Ampicillin Resistance.

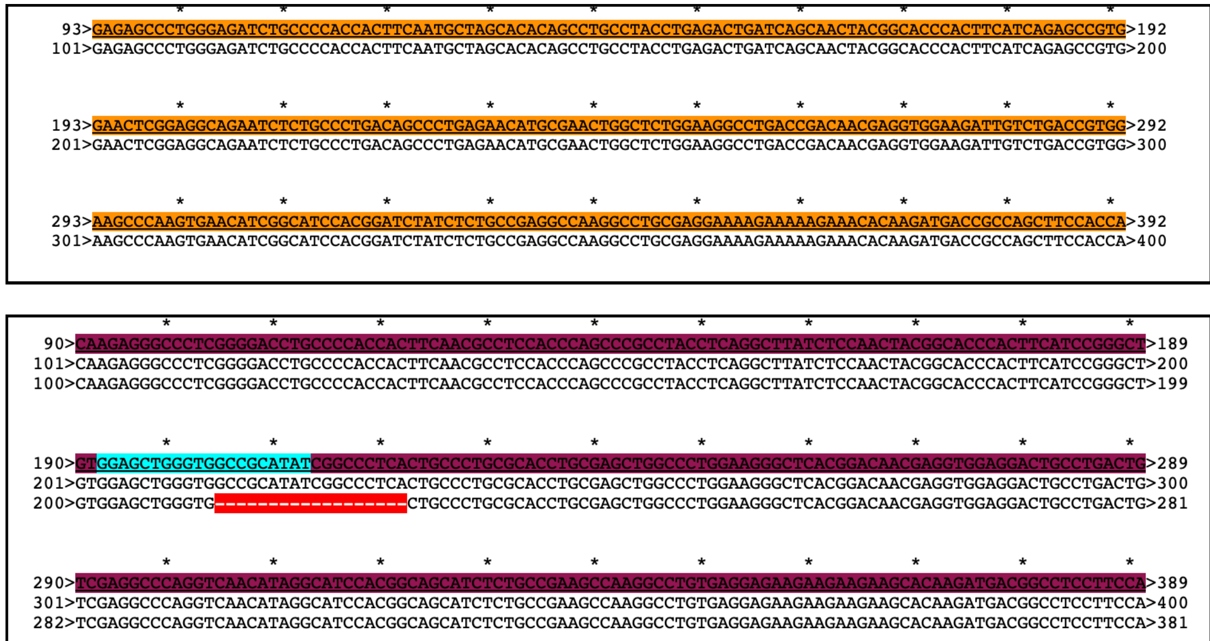


Figure S8. Example of HDR, NHEJ and WT events after gene repair as determined by Sanger sequencing. Sanger sequences were aligned to the respective reference sequences (top strands). Upper panel: The top sequence corresponds to the predicted 'codon-modified' exon 3 sequence of the *PRF1* gene (pos. 42 to 342; highlighted in orange). The bottom sequence corresponds to an HDR event (knock-in corresponding to a perfectly repaired allele): It perfectly matches the reference sequence. Lower panel: The top sequence corresponds to the 5' end of the native exon 3 of the *PRF1* gene (Gene ID 5551) (pos. 40 to 339; highlighted in dark red). The target site of the guide RNA (gPRF1.2) (Table 1) is highlighted in turquoise blue. The middle sequence corresponds to a non-targeted allele and matches the reference sequence. The bottom sequence corresponds to an NHEJ Indel event. The sequence carries a deletion (highlighted in bright red) overlapping the gRNA target site. In order to calculate the knock-in efficiencies shown in Table 2, events in each category (HDR, NHEJ, Non-targeted) are counted and expressed as % of total events.