Multiplexed target enrichment and high-resolution NGS analysis of on- and off-target CRISPR-Cas9 editing events via rhAmpSeq technology

Mollie Schubert[1]*, Garrett R Rettig[1], Gavin Kurgan[1], Matthew McNeill[1], Junzhou Wang[1], Stephanie Fiedler[1], Michael A Collingwood[1], Christopher Vakulskas[1], and Mark A Behlke[1]

[1] Research & Development, Integrated DNA Technologies, Coralville, IA

* Corresponding author: **mschubert@idtdna.com**

Introduction

CRISPR-Cas9 facilitates genome editing using a targeting guide RNA (gRNA) that directs the Cas9-RNA complex to a genomic DNA locus. However, this prokaryotic nuclease system is not always well adapted to maintain high specificity when targeting much larger eukaryotic genomes. Therefore, a streamlined pathway to identify, quantify, and characterize the full array of on- and off-target genome editing events is needed. Here, we describe such a pathway. Our solution first uses established techniques to identify double-stranded breaks (DSBs) for a given genome editing experiment. This is followed by accurate quantification of editing efficiency at each edited locus, using a multiplexed, targeted enrichment approach called rhAmpSeq technology (Figure 1).

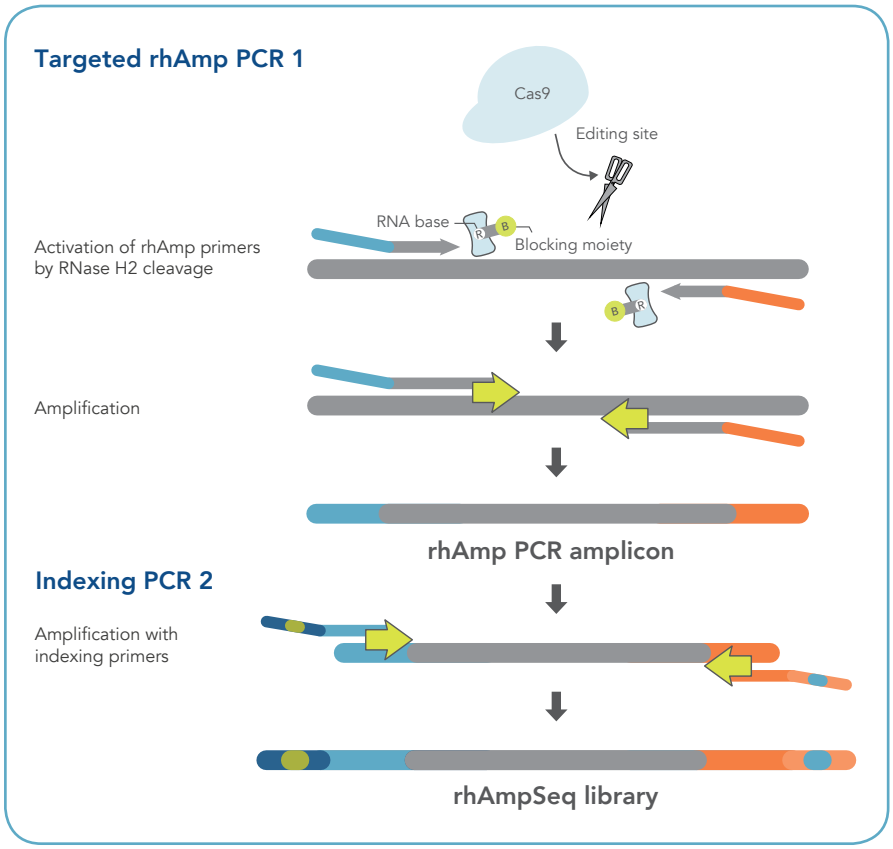


Figure 1. rhAmpSeq amplicon sequencing methodology.

In Step 1, RNase H2 recognizes and cleaves the rhAmp primers via substrate recognition of the RNA base within the DNA:RNA duplex, removing a 3'-blocking group from those primers that are perfectly hybridized [3]. The resulting "activated primers" are extended to generate locus-specific amplicons, enriching validated and/or predicted Cas9 target sites, thus eliminating primer-dimers or mis-amplification. In the second amplification step, Illumina sample indexes and P5/P7 sequences are incorporated using a high throughput PCR and purification protocol.

AAVS1 gRNA survey

Literature precedence for gRNA AAVS1-T2 led to its common use as a gRNA to target AAVS1, a "safe-harbor" locus for transgene integration in human cells. However, the off-target homology and predictive score suggest better guides could be used. Three alternative gRNAs targeting the AAVS1 locus were predicted to have improved specificity by the IDT CRISPR-Cas9 gRNA design checker (www.idtdna.com/Cas9checker) (Table 1). Guides are scored on a scale of 1–100 where a higher value indicates a lower off-target risk.

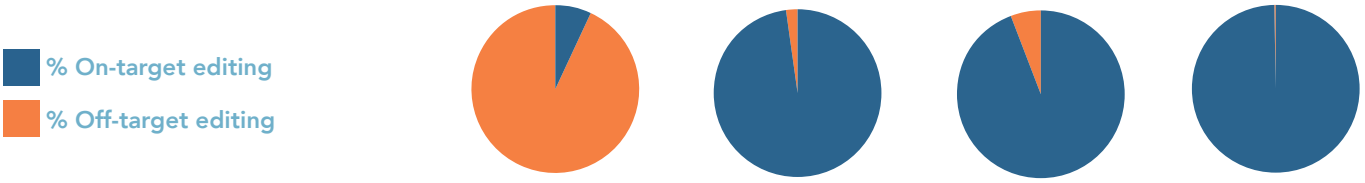
Table 1. Prediction of improved Cas9 gRNA specificity at the AAVS1 locus.

Guide name	20nt guide sequence (5′→3′)	Off-target score
AAVS1-T2	GGGGCCACTAGGGACAGGAT	22
AAVS1-292	GGACGCACCATTCTCACAAA	71
AAVS1-692	TAAGCAAACCTTAGAGGTTC	53
AAVS1-670	CCTCTAAGGTTTGCTTACGA	87

The 4 gRNAs were compared in GUIDEseq [1] experiments to validate on- and off-target cleavage sites (Table 2). This was done in an unbiased fashion via nucleofection of 10 µM Alt-R crRNA:tracrRNA complexes (IDT) into HEK-293 cells with stable expression of Cas9. The results demonstrate that guide RNAs derived from the IDT gRNA design tool offer superior on- to off-target ratios for the human AAVS1 locus when compared to the more commonly used AAVS1-T2. In particular, AAVS1-670 exhibits a perfect balance between editing efficiency and accuracy. Data analysis was as published with improved local alignment capability [2].

Table 2. Confirmation of gRNA specificity using GUIDEseq.

	Guide			
	AAVS1-T2	AAVS1-292	AAVS1-692	AAVS1-670
On-target editing (% total reads)	7.4	97.9	94.3	99.8
Off-target sites (number detected)	297	8	8	4



rhAmpSeq panel design

Using the rhAmpSeq Custom Design Tool (IDT), rhAmp primers were computationally designed to flank each Cas9 cleavage site identified by GUIDEseq (see Table 2). rhAmp primers were synthesized, analyzed by ESI-MS, pooled, and used in target enrichment via multiplex amplification in a single reaction.

Guide	Guide coordinates				Sequence alignments																		GS reads
	Chr	Start	Stop	Strand	G	G	G	G	C	C	A	C	T	A	G	G	G	A	C	A	G	G	A	T	7064
AAVS1-T2	chr12	107092486	107092506	+	A	T	G	G	C	C	A	C	T	A	A	G	G	A	C	A	G	G	A	A	9315
	chr6	36797686	36797706	+	G	G	G	A	C	C	A	T	C	A	G	G	G	A	C	A	G	G	A	T	8527
	chr19	16064179	16064199	-	G	G	G	G	C	T	T	C	T	A	A	G	G	A	C	A	G	G	A	T	6100
	chr21	41521017	41521037	-	G	G	G	G	C	C	C	C	T	G	G	G	G	A	C	A	G	A	A	T	5766
	chr15	89933458	89933478	-	G	C	T	G	C	C	A	C	T	T	G	G	G	A	C	A	G	G	A	G	5511
	chr18	48222163	48222183	-	A	G	G	G	A	C	A	C	C	A	G	G	G	A	C	A	G	G	A	T	5368
	chr2	204531403	204531423	+	G	G	G	G	A	T	G	C	T	A	G	G	G	A	C	A	G	G	A	T	5041
	chr13	105960562	105960582	+	G	G	G	G	C	C	A	A	T	T	A	G	G	A	C	A	G	G	A	T	4988
	chr11	61343764	61343784	+	C	A	G	G	C	C	A	C	C	A	G	G	G	T	C	A	G	G	A	T	4519
	chr8	22778070	22778090	-	G	G	G	G	C	A	A	C	T	A	G	A	G	A	C	A	G	G	A	A	4181
	chr10	130782551	130782571	+	G	A	G	G	C	C	A	T	A	A	G	G	G	A	C	A	G	G	G	T	3372
	chr22	44303220	44303240	-	G	G	T	G	C	C	A	C	C	A	G	G	G	A	G	A	G	G	A	T	2649
	chr11	118846778	118846798	-	A	T	G	G	C	C	A	C	A	A	G	A	G	A	C	A	G	G	A	G	2518
	chr7	2104507	2104527	+	T	C	T	G	C	C	A	C	C	A	G	A	G	A	C	A	G	G	C	T	2497
	chr8	143802949	143802969	+	G	G	T	G	C	C	A	C	T	A	G	G	C	A	C	A	G	G	A	G	2480
	chr11	27064833	27064853	+	G	G	G	G	C	C	A	C	-	A	G	G	G	G	C	A	G	G	A	T	2246
	chr17	75895457	75895477	-	G	G	G	G	T	C	A	G	C	A	G	G	G	G	C	A	G	G	A	T	2158
	chr7	74106811	74106831	+	G	C	T	G	C	C	A	C	T	A	G	A	G	C	C	A	G	G	A	T	1935
	chr7	51607198	51607218	-	C	T	G	G	C	C	A	C	C	A	A	G	G	A	C	A	G	G	C	T	1543
Guide	**Guide coordinates**				**Sequence alignments**																		**GS reads**
	Chr	**Start**	**Stop**	**Strand**	**C**	**C**	**T**	**C**	**T**	**A**	**A**	**G**	**G**	**T**	**T**	**T**	**G**	**C**	**T**	**T**	**A**	**C**	**G**	**A**	9034
AAVS1 670-AS	chr3	97914653	97914673	+	C	C	T	C	A	A	A	G	G	T	T	T	-	C	T	T	A	C	A	A	12
	chr8	22778056	22778076	+	C	C	A	T	C	A	C	G	G	T	T	C	C	C	T	T	C	C	T	G	4
	chr9	97300276	97300296	+	C	C	A	C	A	A	G	G	T	A	C	A	G	G	T	T	G	G	G	A	3
	chr14	90773070	90773090	-	C	C	T	G	A	A	G	G	C	C	C	T	G	C	C	T	G	C	C	A	2

Evaluation of genome editing

rhAmpSeq panels for AAVS1-T2 (111-plex) and AAVS1-670 (9-plex) enable detection of all targets of interest in a single reaction. This results in valuable time- and cost-savings and the ability to simultaneously test multiple variables (Cas9 source, gRNA modifications, cell lines, RNP dose, HiFi vs. WT Cas9, etc.) (Figure 2). Each experiment is sequenced by Illumina sequencing at an exceedingly high read map rate and with excellent target uniformity (see Figure 3).

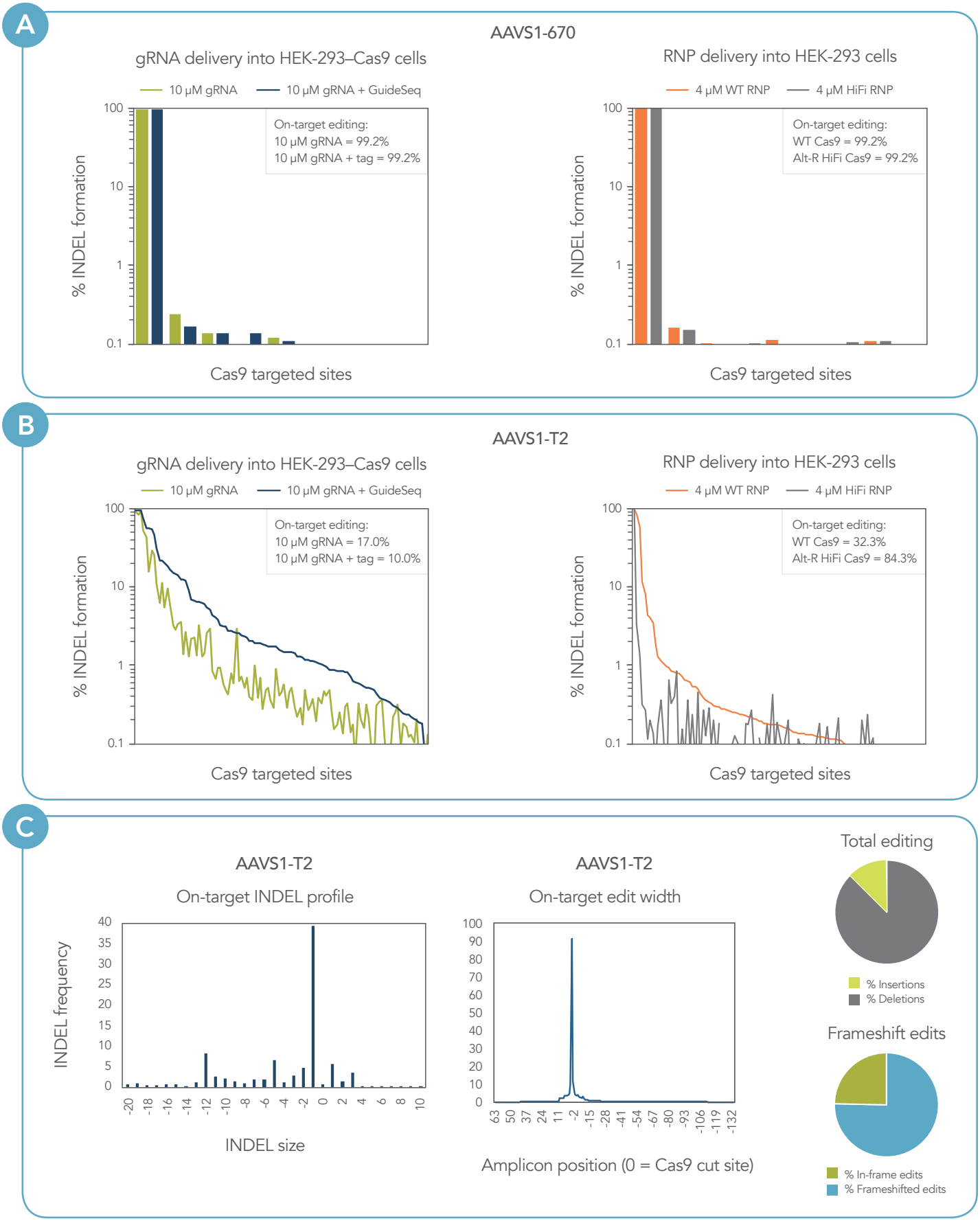


Figure 2. Quantification and characterization of genome editing using a custom rhAmpSeq panel. (**A**) rhAmpSeq technology confirms specificity of editing at only the AAVS1 locus is >99% for Alt-R HiFi Cas9 RNP delivery of AAVS1-670. (**B**) Editing by the low-specificity gRNA is likewise dramatically improved to >84% on-target accuracy using Alt-R HiFi Cas9 RNP electroporation into HEK-293 cells. (**C**) The complete characterization of editing events available for all targets for the AAVS1-T2 gRNA.

Panel uniformity assessment

A 282-plex, custom rhAmpSeq panel for target enrichment was used to generate sequencing libraries from 2 sources of Coriell gDNA at 10, 15, 25, and 50 ng amounts. On-target mapping rates and target uniformity were assessed for each condition (Figure 3).

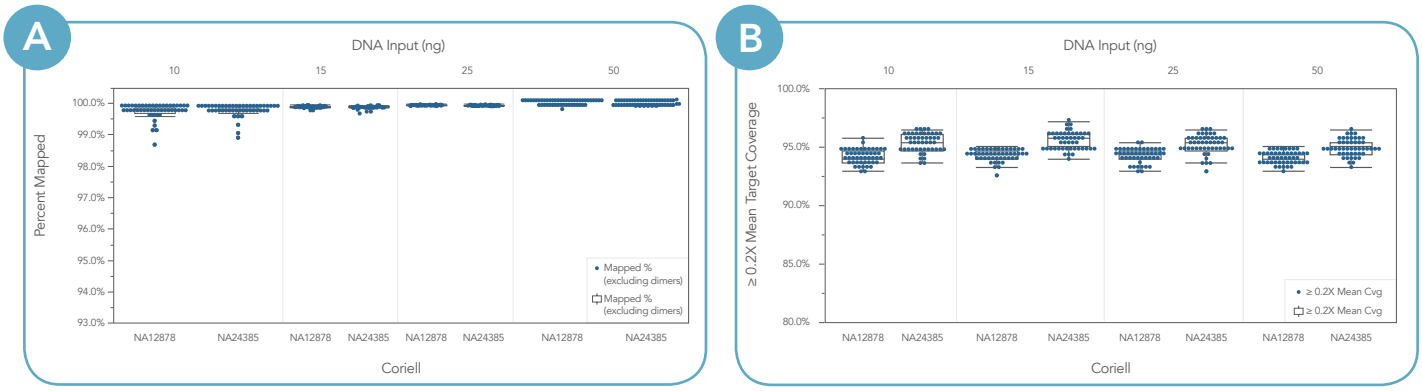


Figure 3. Custom 282-plex rhAmpSeq pool gives very high mapping rate and uniformity. (**A**) "Percent-mapped" is defined as the number of merged reads mapping to the human genome over the total number of reads. For this custom panel of 282 primer pairs, the mean percent-mapped is >99% for all conditions tested (n = 48 for each gDNA source and amount). (**B**) The same libraries were used to assess uniformity of target amplification. Uniformity is measured by the fraction of targets that have a read depth that is >0.2X of the mean read depth. The requirement of the custom panel is that >85% of all targets meet this metric. For both percent mapped and target uniformity measurements, the resultant read counts are from Illumina sequencing (MiSeq® system, 2 x 150 paired-end reads, v2 chemistry)

Conclusions

- Novel rhAmpSeq technology facilitated an accurate evaluation and determination of the most specific Cas9 gRNA option that targets the AAVS1 locus.
- The percent mapped and target uniformity displayed (Figure 3) for the custom 282-plex rhAmpSeq panel demonstrates reliability of the rhAmpSeq technology, which is built upon the primer design tool and the RNaseH2-containing rhAmpSeq master mix.
- Absent *in silico* gRNA screening to alleviate off-target-editing, rhAmpSeq analysis of genome-wide editing by Alt-R HiFi Cas9 indicates that this enzyme dramatically reduces off-target editing and maintains excellent on-target performance [4].
- The multiplex enrichment and data analysis provide an efficient workflow from primer design through sequencing, enabling accurate and comprehensive characterization of editing events.

References

1. Tsai SQ, Zheng Z, et al. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol. 33(2):187–197.
2. Tsai SQ, Topkar VV, et al. (2016) Open-source guideseq software for analysis of GUIDE-seq data. Nat Biotechnol. 34(5):483.
3. Dobosy JR, Rose SD, et al. (2011) RNase H-dependent PCR (rhPCR): improved specificity and single nucleotide polymorphism detection using blocked cleavable primers. BMC Biotechnol. 11:80.
4. Vakulskas CA, Dever DP, et al. (2018) A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. Nat Medicine. 24(8):1216–1224.

Find more information on rhAmpSeq technology at www.idtdna.com/rhampseq.

NGS-10139-FL 04/19

IDT
INTEGRATED DNA TECHNOLOGIES