How to monitor and minimize off-target events during genome editing

The CRISPR-Cas9 genome editing system has seen exponential growth in adoption with broad applications from basic research to therapeutics. The system is composed of Cas9 nuclease and a target-specific guide RNA (gRNA) that as a complex induces a double-stranded DNA break at a desired location. CRISPR-mediated genome editing is an extremely powerful tool that enables researchers to create different cellular models by removing, adding, or altering sections of a DNA sequence in the genome in a wide range of different cell types and gene loci. While it can achieve high editing efficiencies, the CRISPR-Cas9 system can also cleave the target DNA at unintended locations (known as off-target events) that can result in undesired phenotypes or loss of functional gene activity, which is especially detrimental for therapeutic applications.

To minimize the occurrence of off-target events in applications of the CRISPR-Cas9 system, several factors must be evaluated including: (1) delivery format of the CRISPR-Cas9 system—the purified protein format of Cas9 nuclease offers the fastest clearance time upon delivery, decreasing the time for off-target edits to occur; (2) optimization of gRNA design—leveraging an *in silico* predictive tool to design and select gRNA with a high score and less predicted off-target events further decreases the potential for unintended edits; (3) the specificity of Cas9 enzyme—using a high-fidelity Cas9 enzyme with improved specificity further minimizes off-target events. While optimizing any one of these factors will help decrease the chance of off-target events, the best possible outcome is achieved when all three factors are considered.

Even with the optimization of each of these factors, off-target events cannot be completely eliminated, and thus still pose risks for the genome editing project. As a result, accurate detection and monitoring of off-target events is an important step in any genome editing project, especially if the end goal is a therapeutic application.

Here we discuss strategies to optimize the various factors that cause off-target events and describe an unbiased analysis system, TEG-seq, to help detect off-target events. We also describe the use of this system to screen for a high-fidelity Cas9 mutant with improved specificity.

Method to monitor and detect off-target events

To optimize the various factors that result in unintended edits, a robust analytical method is required to reproducibly measure off-target events. Over the years, several off-target analysis systems have been developed, most as *in vitro* methods [1-4] where the genome editing reaction is performed in tubes without cellular context (such as histones and other DNA-binding proteins), which may influence the off-target potential of a system. Several *in cellulo* detection methods have also been developed including GUIDE-seq (genome-wide, unbiased identification of double-stranded breaks (DSBs) enabled by sequencing), which is currently the most widely used method for off-target detection [5-6]. Unlike *in vitro* methods, the editing reaction in *in cellulo* methods is done directly in the cells where specific cellular context is represented. However, most of the currently available *in cellulo* detection methods are not sensitive enough to detect low-frequency off-target events.

To address the sensitivity limitation of GUIDE-seq, we developed a more targeted genome-wide off-target screening method: TEG-seq (tag-enriched GUIDE-seq) [7]. In this method, specific 5′ phosphorylated primers are used for PCR amplification and differential marking of amplicons containing a double-stranded DNA tag (dsTag) inserted in DSB sites (Figure 1). With this alternative *in cellulo* method, only amplicons with a dsTag are phosphorylated at the 5′ end and can be ligated to a barcoded adaptor (BC-A) for further amplification and enrichment. These improvements significantly reduce nonspecific amplification and improve sensitivity of DSB detection. TEG-seq was also applied in experiments to detect and predict off-target events in engineered rats and mice embryos, which showed it is better than other methods [8].

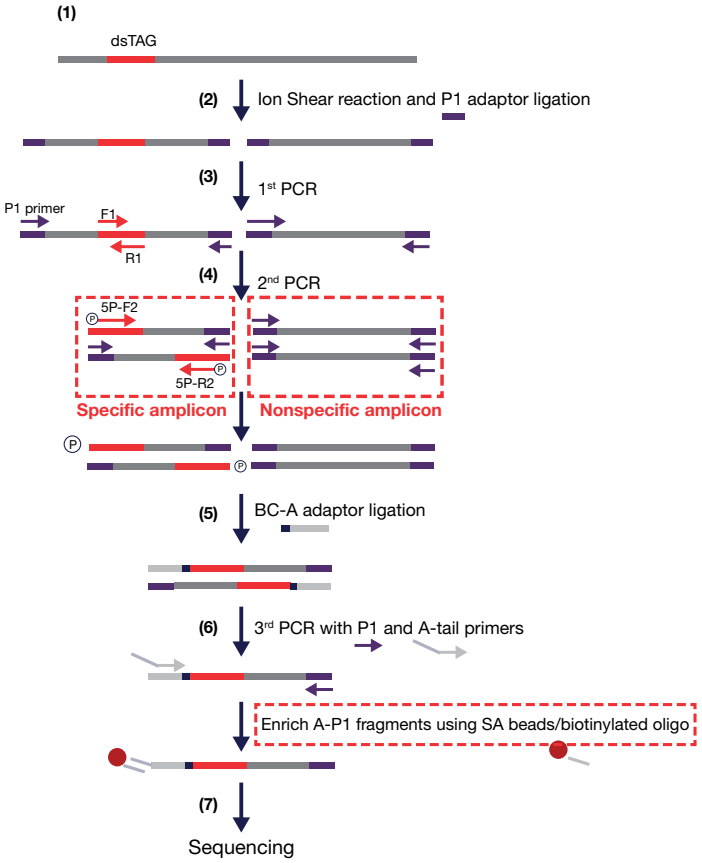

Figure 1. The TEG-seq workflow. (1) A double-stranded DNA tag (dsTag) is cotransfected with Cas9–gRNA ribonucleoprotein (RNP) or plasmid that expresses Cas9 and gRNA. The dsTag integrates at any site containing a DSB. (2) The genomic DNA is extracted and fragmented to 200–600 bp using enzyme-based Ion Shear™ chemistry. The P1 adaptor is ligated to the fragmented genomic DNA. (3) The first PCR is performed in separate tubes using P1/F1 for the forward primers and P1/R1 for the reverse primers. (4) In the second (nested) PCR, 5′ phosphorylated primers (5P-F2 and 5P-R2) are used that generate PCR products containing a 5′ phosphate only in the 5P-F2/P1 and 5P-R2/P1 products, but not the P1/P1 product. (5) A non-phosphorylated barcode adaptor (BC-A) is specifically ligated to the 5P-F2/P1 and 5P-R2/P1 products, but not to the P1/P1 product. (6) A third PCR is performed using P1/A-tail primers followed by a bead enrichment via a biotinylated capture oligo that is complementary to the A-tail primer. (7) The enriched amplicons are then subjected to next-generation sequencing.

In comparison to GUIDE-seq, TEG-seq detected more total off-target events under similar depth of next-generation sequencing (NGS) (Figure 2). Additionally, the read number for an individual target from TEG-seq is on average 10-fold higher than GUIDE-seq. This clearly indicates that TEG-seq is more sensitive and specific than GUIDE-seq. To further verify the off-target events detected by TEG-seq, we used the targeted amplicon-seq validation (TAV-seq) method for the quantification of off-target editing (Figure 3). The results showed that TEG-seq can detect low-frequency off-target events at 0.001% probability level as detected by TAV-seq.

As discussed later, TEG-seq also works efficiently in different cell types including primary T cells (Figure 6) and iPSCs (Figure 7) for the screening of off-target events that occur with clinically relevant gRNA targets, using Cas9–gRNA RNP format. Thus, TEG-seq is an unbiased genome-wide analysis method that effectively detects off-target events at low frequency in a wide range of cell types.

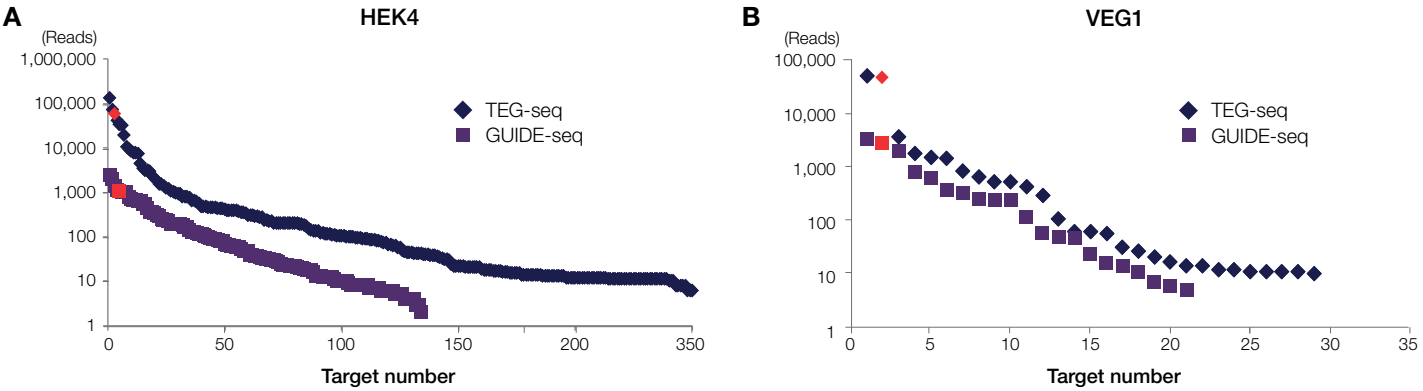

Figure 2. Comparison of TEG-seq and GUIDE-seq. Read numbers for **(A)** *HEK4* and **(B)** *VEG1* loci were plotted from individual on-target events (red) and off-target events from TEG-seq (blue) and GUIDE-seq (purple). The total off-target events detected by TEG-seq is 252 for HEK4 and 27 for VEG1, and the total off-target events detected by GUIDE-seq is 132 for HEK4 and 21 for VEG1. The read number for an individual target is also higher in TEG-seq than GUIDE-seq with a similar level of NGS sequencing depth. Cas9 and gRNA were delivered using plasmid format.

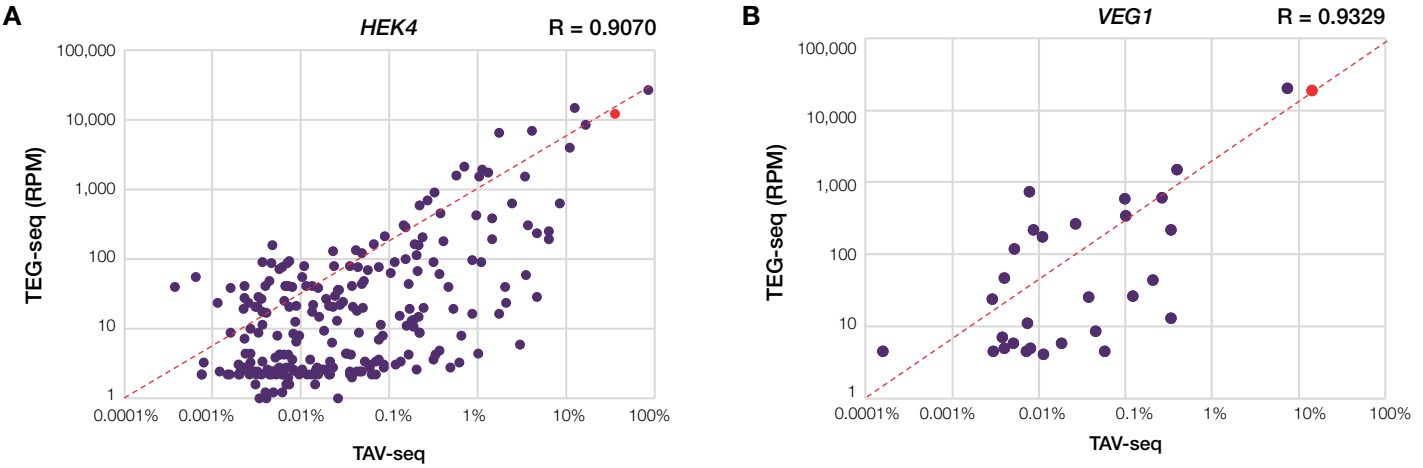

Figure 3. Comparison of off-target detection level between TEG-seq and TAV-seq. Off-target events for **(A)** *HEK4* and **(B)** *VEG1* loci were detected by TEG-seq and TAV-seq. Events detected by TEG-seq are plotted in reads per million (RPM) against the percentage of cleavage detected by TAV-seq. The correlation factor (R value) is indicated on the upper-right side of each graph. The on-target activity is indicated by red color.

Strategies to minimize off-target events through design

The CRISPR-Cas9 system is a powerful genome editing tool that only requires the presence of the Cas9 nuclease and gRNA. The Cas9–gRNA complex searches for NGG protospacer-adjacent motifs (PAMs) in the genome. When a sufficient match between the gRNA and the dsDNA target is detected, the Cas9 nuclease cleaves the DNA and produces a double-stranded break (DSB). While the CRISPR-Cas9 system typically cleaves the genome at the target site with high efficiency, cleavage at undesired sites with mismatches of one to several bases can occur. These undesired cleavage events are known as off-target effects and should be minimized to help prevent undesired side effects. Several factors can be leveraged to strategically generate a CRISPR-Cas9 system with minimal off-target events.

1. **Delivery format of the CRISPR-Cas9 system:** The delivery format of the Cas9–gRNA complex influences the system's clearance time and duration of nuclease expression. The use of Cas9–gRNA RNP complex containing purified Cas9 protein results in an initially high level of complex followed by rapid decay or clearance. As such, the Cas9–gRNA RNP has less time to cause undesired off-target effects. Therefore, the use of the Cas9–gRNA RNP format is recommended because it offers a high level of editing efficiency combined with faster clearance, resulting in minimal off-target effects.
2. **Optimization of gRNA design:** The use of high-scoring gRNA can help reduce the off-target events associated with the CRISPR-Cas9 system. Many genome editing design tools are available, including the Invitrogen™ TrueDesign™ Genome Editor, that enable researchers of all experience levels to easily design, select, and order reagents for accurate and successful gene editing experiments. Based on several criteria, including the probability of off-target activity, the design tool assigns a score to each gRNA. The higher the score, the less potential for off-target events. However, limitations such as the availability of PAM sites, proximity to target loci, and overall efficiency could preclude researchers from identifying gRNAs with low off-target events.
3. **High fidelity of the Cas9 enzyme:** The use of high-fidelity Cas9, an enzyme that is engineered to demonstrate improved specificity, can reduce the occurrence of off-target events. The CRISPR-Cas9 system is an extremely powerful tool that has completely transformed cell engineering as we know it. While the wild-type Cas9 nuclease can achieve high editing efficiency in a wide variety of cell types, the high editing efficiency of the wild-type Cas9 nuclease comes at an expense of increased off-target effects. The same properties that make the wild-type Cas9 nuclease so effective in cutting the genome at the desired locus inherently make it an effective tool at cutting the genome at undesired locations.

To improve the specificity of the wild-type Cas9 protein, we set out to engineer a high-fidelity Cas9 nuclease variant that would retain as much of the original on-target editing efficiency as the wild-type Cas9 nuclease, but demonstrate improved specificity at the same time. The next section outlines the steps we took to build a high-fidelity Cas9 variant that strikes the right balance between on-target editing and increased specificity.

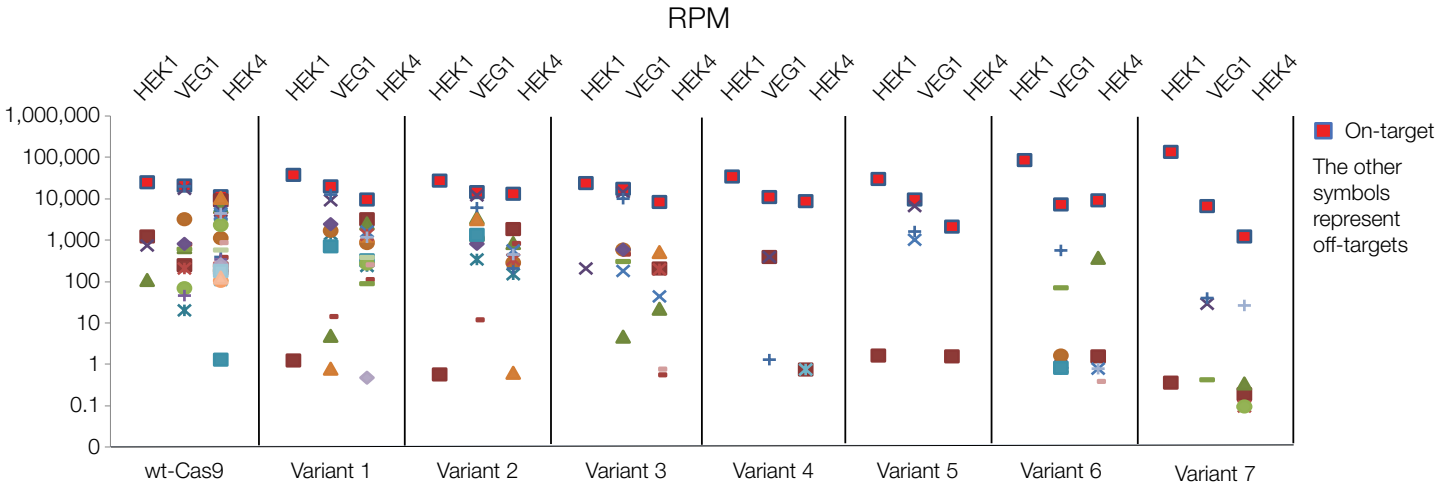

Figure 4. Genome-wide off-target screening for Cas9 variants using TEG-seq. Three commonly studied gRNAs (targeting *HEK1*, *VEGFA1*, and *HEK4*) that represent low, medium, and high potential off-target events were mixed with Cas9 protein and cotransfected in HEK293 cells. Invitrogen™ TrueCut™ Cas9 Protein v2 (wt-Cas9) was used as a control in a parallel screen with 7 high-fidelity Cas9 variants. Samples were barcoded using the Ion Xpress™ Barcode kit and sequenced using the Ion S5™ XL System. The in-house–developed Motif_Search tool was used for off-target analysis. RPM from each barcoded sample was calculated and plotted in log scale (y-axis). Red squares are on-target events, and all other markers are off-target events.

Screening of Cas9 mutants for improved specificity

Using TEG-seq, we set out to screen for a high-fidelity Cas9 variant with improved off-target profiles. Seven high-fidelity Cas9 candidates were identified and included in the screen from in-house engineered variants and published Cas9 candidates. Three commonly studied gRNAs targeting different loci (*HEK1*, *VEGFA1*, and *HEK4*) were selected and screened for off-target events with the Cas9–gRNA RNP delivery format in HEK293 cells. Shown in Figure 4 is an example of the Cas9 screening results where variant 4 outperformed other candidates. Variant 4 generated the least number of off-target events and lower read number (or number of actual cuts) at each individual off-target site compared to other Cas9 candidates in the panel.

Variant 4 was selected for further verification and compared to Sniper-Cas9 (a recently published high-fidelity Cas9 [9]) and a high-fidelity Cas9 from another supplier (Supplier I). Table 1 shows one example of the TEG-seq data on *HEK4*. Although Sniper-Cas9 and Supplier I generated less off-target events compared to TrueCut Cas9 Protein v2 (wt-Cas9), they both generated much higher off-target events compared to variant 4.* Data from Table 1 suggested that our high-fidelity Cas9 candidate, variant 4, generated 80% less off-target cleavage sites compared to TrueCut Cas9 Protein v2. Variant 4 became the new Invitrogen™ TrueCut™ HiFi Cas9 Protein.

* Total number of off-target sites: TrueCut Cas9 Protein v2 (wt-Cas9): 34, Sniper-Cas9: 18, Supplier I: 13, TrueCut HiFi Cas9 Protein (variant 4): 7.

Table 1. Reads per million (RPM) for off-target events detected by TEG-seq using *HEK4* gRNA in HEK293 cells.

Target	MM	Align sequence	PAM	wt-Cas9	Sniper-Cas9	Supplier I	TrueCut HiFi Cas9
On	0	GGCACTGCGGCTGGAGGTGG	GGG	25,950	112,147	57,977	41,848
Off-1	2	. . . . . GA.	GGG	23,050	26,225	6,608	691
Off-2	2	. . . . A . . . . . C . .	AGG	20,196	37,895	21,393	497
Off-3	2	. . . G. . . . . G.	AGG	18,843	8,898	1,074	7
Off-4	3	A. . T. . . . . A. .	GGG	16,942	3,890	24	0
Off-5	3	A. . . G. . . . . A.	TGG	10,310	5,654	629	0
Off-6	3	T. . . . . C. . . . . A. .	TGG	9,697	13,852	12,438	10
Off-7	3	A .G. . . . . G.	TGG	8,763	4,072	881	5
Off-8	4	. A. . . CA. . . . . A . . .	TGG	6,934	619	0	0
Off-9	3	. . . . TCA.	AGG	5,215	0	0	2
Off-10	2	. . . . . T. . . . . C.	AGG	3,113	976	0	2
Off-11	2	. . . . . G . . T	GGG	2,988	0	2,180	0
Off-12	2	. . . . . T . . . . . G	TGG	1,984	172	0	0
Off-13	2	. . . T . . . . . G . .	TGG	1,386	1,987	208	0
Off-14	2	. . . . – . . . . . g	AGG	1,272	0	0	0
Off-15	3	A . A . . . . . T	TGG	1,182	0	0	0
Off-16	3	CC . . . . . G	GGG	1,128	0	0	0
Off-17	3	T . . . . . T . . . . . A . . .	GGG	1,014	2	2	0
Off-18	3	. . T . . . CT	TGG	908	0	0	0
Off-19	3	. C . . . . . A . . . A	AGG	869	0	0	0
Off-20	3	. . . . . g . . A . . C	TGG	800	344	718	0
Off-21	3	. . . T . C . . A	GGG	744	0	0	0
Off-22	3	. . . . . A . A . . . . G	GGG	628	0	0	0
Off-23	4	A . . . . . A . . . . . GA	AGG	609	676	0	0
Off-24	3	. A . . . . . A . . . A	GGG	550	0	0	0
Off-25	3	T . G . . . . . a . . .	AGG	511	271	114	0
Off-26	2	. . . . . G . . . . . C . .	GGG	498	2,145	0	0
Off-27	4	. A . . . C.T. A	AGG	414	182	0	0
Off-28	3	. . . . . G . . . . . G . . A . .	GGG	320	0	353	0
Off-29	2	. . . . . A . . G	GGG	216	0	0	0
Off-30	4	. . . TG . . . . . CA	AGG	211	0	0	0
Off-31	2	. . A . . . . T	CAG	194	287	0	0
Off-32	3	. . . . . G . A . . . . –	TGG	135	0	0	0
Off-33	3	. C . . . . . G . . . . . G	GGG	80	0	0	0
Off-34	3	. . A . . . . G . . . . . G	GGG	42	0	0	0

To further evaluate the effectiveness of TrueCut HiFi Cas9 Protein in a more diverse set of cell types, particularly in therapeutically relevant primary T cells, we conducted additional off-target screening to compare TrueCut HiFi Cas9 Protein against TrueCut Cas9 Protein v2 (wt-Cas9) and enzyme from Supplier I. Twenty-one gRNAs were selected targeting four therapeutically relevant genes (*CD52*, *TRAC*, *TRBC*, and *PD1*) in T cells (Figure 5). Some of these gRNAs have been evaluated for CAR T cell gene therapy [10-12]. To demonstrate the difference in fidelity between the three Cas9 proteins, we intentionally included three gRNAs (TRBC-4, PD1-4, and PD1-5) that had low score from *in silico* analysis to represent gRNAs with high predicted off-target potential. In general, TrueCut HiFi Cas9 Protein generated much fewer off-target events and lower off/on ratio at individual off-target sites compared to TrueCut Cas9 Protein v2 and enzyme from Supplier I across different probability scales (Figure 5B).

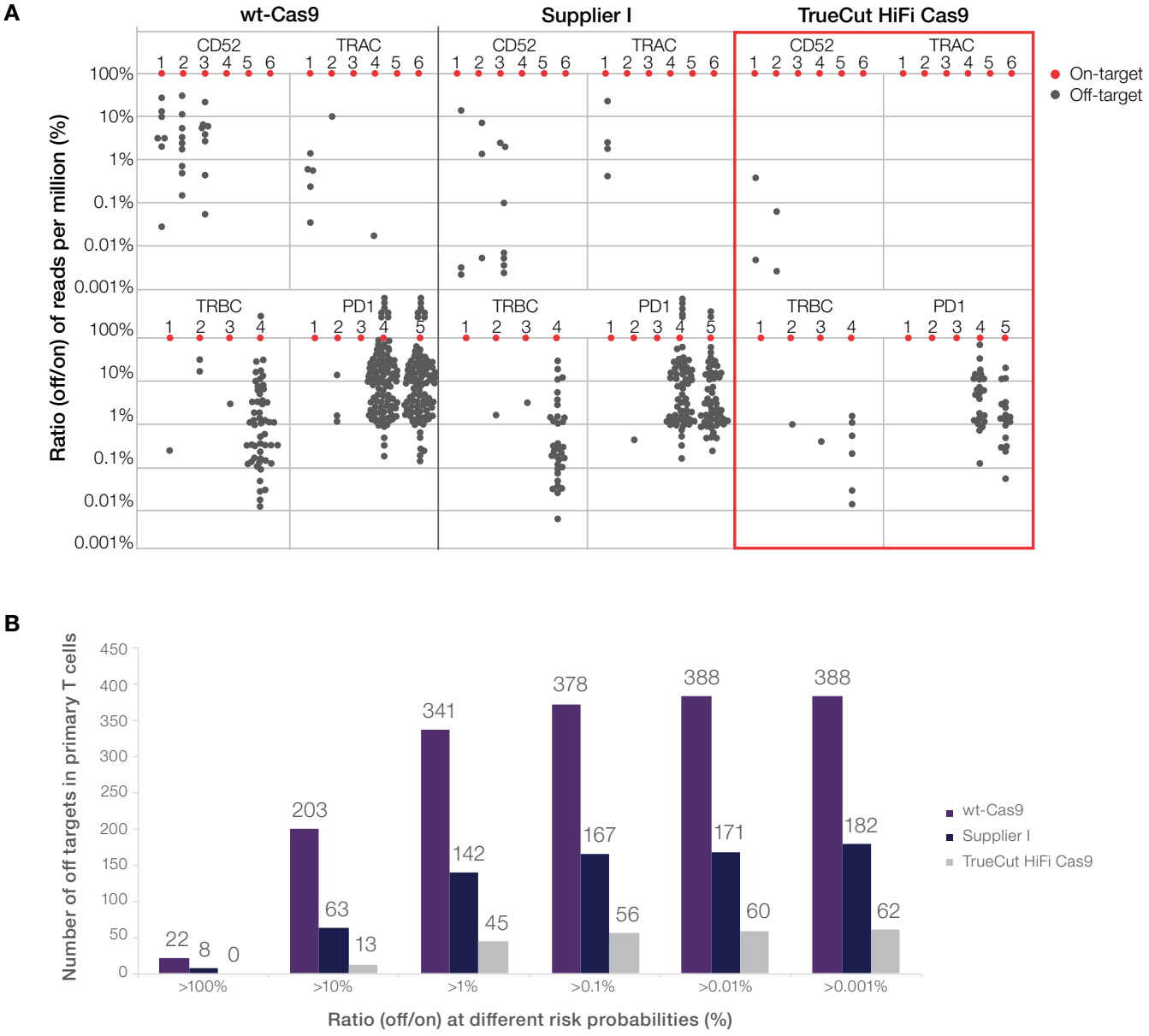

Figure 5. Genome-wide off-target events detected in T cells by TEG-seq and their off/on ratios at different percentile scales. (A) The 21 gRNAs targeting four genes *(CD52*, *TRAC*, *TRBC*, and *PD1*) were individually cotransfected with TrueCut Cas9 Protein v2 (wt-Cas9), enzyme from Supplier I, or TrueCut HiFi Cas9 Protein. The number underneath each gene name is the gRNA ID number. The off/on ratio was calculated based on RPM from individual off-target events divided by the RPM from the corresponding on-target events. Red dots are on-target events normalized to 100%. Gray dots are the off/on ratio. The gray dots above the red dots are off-target events with reads higher than on-target events. **(B)** The total number of off-target events in different percentiles based on risk probability.

Similar performance analysis was also conducted in iPSCs to demonstrate the difference in fidelity between the three Cas9 proteins. Genome-wide off-target screening was performed in iPSCs on 4 gRNAs: one gRNA targeting a commonly studied *HEK4* target, two gRNAs targeting two SNPs in the hemoglobin β subunit (*HBB*) gene that cause sickle cell disease, and one gRNA to knock out *BLC11A*, as a potential cure for sickle cell disease. As shown in Figure 6, off-target events were detected in 3 gRNAs (*HBB1*, *HBB2*, and *HEK4*) while no off-target events were detected from *BCL11A* gRNA (data not shown). Similar to its efficiency in other cell types, TrueCut HiFi Cas9 Protein also generated fewer off-target events and lower off/on ratio for individual off-target sites compared to TrueCut Cas9 Protein v2 (wt-Cas9) and protein from Supplier I.

A

HBB1			**Reads per million (RPM)**		
Target	Align sequence	PAM	wt-Cas9	Supplier I	TrueCut HiFi Cas9
On	CTTGCCCCCACAGGGGCAGTAA	CGG	141,922	258,580	284,917
Off1	TCA.....	GGG	126,583	970	132
Off2	T.....T.G.	CAG	13,100	15,871	1,229

HBB2			**Reads per million (RPM)**		
Target	Align sequence	PAM	wt-Cas9	Supplier I	TrueCut HiFi Cas9
On	CTTGCCCCCACAGGGGCAGTAA	AGG	431,212	356,556	452,904
Off1	...AA.....	TGG	246,927	3,153	319
Off2	..A..a.....C.....C.....	GGG	929	228	7
Off3	G.....A.	AGG	118	2,197	0

HEK4			**Reads per million (RPM)**		
Target	Align sequence	PAM	wt-Cas9	Supplier I	TrueCut HiFi Cas9
On	GGCACTGCGGCTGGAGGTGG	GGG	82,924	147,851	159,782
Off1	...G.....G.....	AGG	149,096	848,833	8,695
Off2	.....GA.....	GGG	151,950	390	105
Off3	A.G.....G.....	TGG	246,887	927	124
Off4	.....T.....C.....	AGG	118,633	119	65
Off5	.....A.....C...	AGG	12,949	8	0
Off6	.....G.....C...	GGG	3,005	21	0
Off7	T.....C.....A...	TGG	1,734	0	0
Off8	.....–G.....g.....	AGG	99	0	0
Off9	.....T.....G...	TGG	52	0	0

B

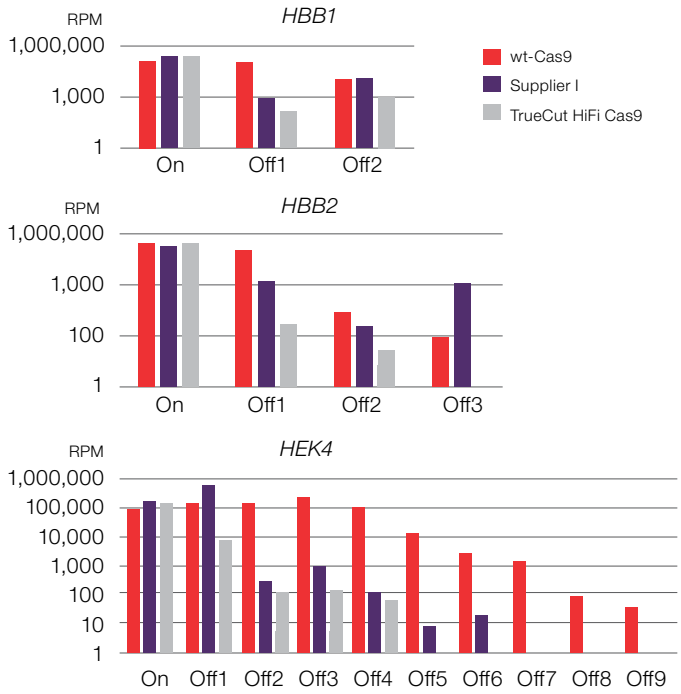

Figure 6. Off-target events detected in iPSCs with two gRNAs targeting two SNPs in *HBB* and one gRNA targeting *HEK4*. (A) TEG-seq data table containing the sequence and RPM for all on- and off-target events from each gRNA. **(B)** Bar graph representation of the RPM results for on- and off-target events from the table. TrueCut HiFi Cas9 Protein showed higher fidelity compared to TrueCut Cas9 Protein v2 and protein from Supplier I.

The high-fidelity Cas9 variant that we identified retains sufficient on-target editing efficiency for standard use in genome editing experiments while significantly reducing off-target events commonly observed when using the wild-type protein. The use of the TrueCut HiFi Cas9 Protein is especially beneficial when it is necessary to use a suboptimal gRNA option due to the limited availability of PAM sites near the cut site. TrueCut HiFi Cas9 Protein is also beneficial in applications where off-target events can result in undesired phenotypes or loss of functional gene activity.

invitrogen

Conclusion

The CRISPR-Cas9 system is a widely adopted genome editing tool with broad applications from basic research to therapeutics. While it can achieve high editing efficiencies, off-target events must be minimized to prevent undesired phenotypes or loss of functional gene activity, which is especially detrimental for therapeutic applications. As a result, accurate detection of off-target events is essential, and appropriate design choices must be made to minimize off-target events. Here we demonstrated the effectiveness of TEG-seq as an *in cellulo* analysis method, with 10-fold more sensitivity and specificity compared to GUIDE-seq. We later leveraged TEG-seq for the identification of a high-fidelity Cas9 (TrueCut HiFi Cas9 Protein) that exhibited superior off-target profiles compared to TrueCut Cas9 Protein v2 and another supplier's high-fidelity Cas9 enzyme in a wide range of cell types, including primary T cells and iPSCs.

Genome editing products and services

For more information on TrueCut HiFi Cas9 Protein, go to **thermofisher.com/cas9**

For gRNA design and ordering, go to **thermofisher.com/trueguide**

For TEG-seq services, go to **thermofisher.com/engineeringservices**

References

1. Crosetto N, Mitra A, Silva MJ et al. (2013) Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods* 10:361–368.
2. Frock RL, Hu J, Meyers RM et al. (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* 33:179–187.
3. Tsai SQ, Nguyen NT, Malagon-Lopez J et al. (2017) CIRCLE-seq: a highly sensitive *in vitro* screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* 14:607–614.
4. Kim D, Bae S, Park J et al. (2015) Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat Methods* 12:237–243.
5. Wang X, Wang Y, Wu X et al. (2015) Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. *Nat Biotechnol* 33:175–179.
6. Tsai SQ, Zheng Z, Nguyen NT et al. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* 33:187–197.
7. Tang P-T, Ding B, Peng L et al. (2018) TEG-seq: an Ion Torrent–adapted NGS workflow for *in cellulo* mapping of CRISPR specificity. *Biotechniques* 65:259–267.
8. Anderson KR, Haeussler M, Watanabe C (2018) CRISPR off-target analysis in genetically engineered rats and mice. *Nat Methods* 15:512–514.
9. Lee JK, Jeong E, Lee J (2018) Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat Commun* 9(1):3048.
10. Stadtmauer EA, Fraietta JA, Davis MM et al. (2020) CRISPR-engineered T cells in patients with refractory cancer. *Science* 367(6481):eaba7365.
11. Graham C, Jozwik A, Pepper A et al. (2018) Allogeneic CAR-T cells: more than ease of access? *Cells* 7(10):155.
12. Schumann K, Lin S, Boyer E et al. (2015) Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *PNAS* 112(33):10437–10442.

Find out more at **thermofisher.com/crispr**

ThermoFisher SCIENTIFIC