2019 ISSCR

F-3083

TEG-seq: An Ion Torrent-adapted NGS workflow for *in cellulo* mapping of CRISPR specificity

Pei-zhong Tang, Bo Ding, Lansha Peng, Vadim Mozhayskiy, Chris Reyes, Jason Carte, Jarrod Clark, Jason Potter and Jonathan D. Chesnut.

ThermoFisher Scientific. 5781 Van Allen Way, Carlsbad, CA 92008

pei-zhong.tang@thermofisher.com

760-4766058

ABSTRACT

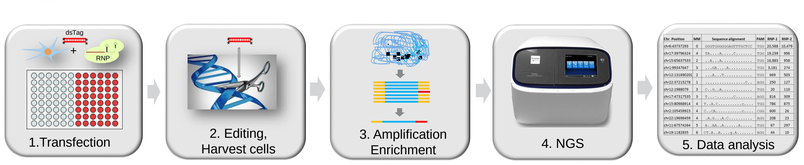
Engineered nucleases, including the CRISPR/Cas9 system, have been widely used for genome editing, and is now being developed to create gene and cell therapies to treat human disease. However, lack of specificity leading to off-target cleavage is still a concern. To measure this, an *in cellulo* method, genome-wide unbiased identification of double stranded breaks enabled by sequencing (GUIDE-seq) was developed and has been widely used (1). However, this method as originally reported was associated with a significant level of non-specific target amplification which reduced sensitivity and increased the cost to detect low-frequency off-target events. In an attempt to improve robustness and sensitivity, we developed a modified method termed Target-Enriched GUIDE-seq (TEG-seq) (2). The modification improves the sensitivity approximately 10 fold compared to GUIDE-seq. In addition to the increased specificity, we developed high-throughput workflow and data analysis tool that led TEG-seq became more cost-effective. Using TEG-seq, we evaluated a panel of Cas9 mutants to identify potential high-fidelity Cas9 protein that will be a critical for genome editing, especially for gene and cell therapy. We also used TEG-seq to map on- and off-target cleavage events on 22 gRNAs targeting a set of therapeutically relevant SNPs. Finally, TEG-seq was used to evaluate CRISPR off-target profiling for therapeutic applications in different cells including iPSC and CAR-T cells and an animal model (3). TEG-seq off-target detection with the use of high-fidelity Cas9 proteins will be one of the crucial steps in genome-editing and gene therapy.

INTRODUCTION

CRISPR-Cas9 system promises the powerful concept of directly correcting mutations or disrupting abnormal genes in order to cure and prevent diseases, particularly inherited genetic disorders. However, CRISPR-Cas9 is known to induce off-target mutations at sites with homology to the target sites. Gene and cell therapeutic applications of CRISPR-Cas9 require a comprehensive knowledge of their off-target effects to minimize the risk of deleterious outcomes. GUIDE-seq, the only *In cellulo* approach that could mapping the specificity of CRISPR retains high non-specific background that leads to low sensitivity of detection and cost-effectiveness. We developed a modified method and successfully adapted it to Ion-Torrent NGS in a high throughput format that is now more sensitive and cost effective. A study of CRISPR off-target analysis in genetically engineered rats and mice demonstrated that TEG-seq was a good predictor of in vivo activity (3)

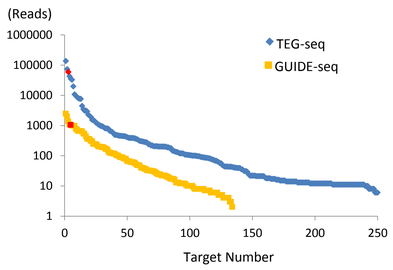
RESULTS

Figure 1: The workflow of TEG-seq



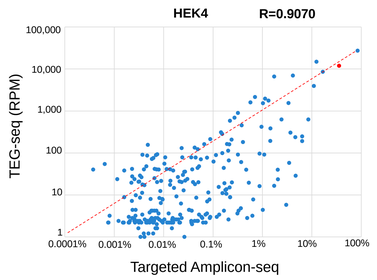
1) Transfection of cells with a double strand DNA Tag (dsTag) (red), gRNA and Cas9 in different formats (plasmid or RNP); 2) dsTag integrated to double strand DNA break (DSB) sites cleaved by CRISPR-Cas9; 3) Genome DNA is sheared and ligated with a universal Ion-P1 adaptor (yellow), and amplified using P1 (yellow) and an optimized 5'phosphate Tag-specific (red) primers, followed by the enrichment of targeted Tag-specific amplicons; 4) Amplicons were sequenced using Ion-Torrent NGS platform; 5) Data analysis using in-house developed plugin.

Figure 2: Off-targets detected by TEG-seq and GUIDE-seq



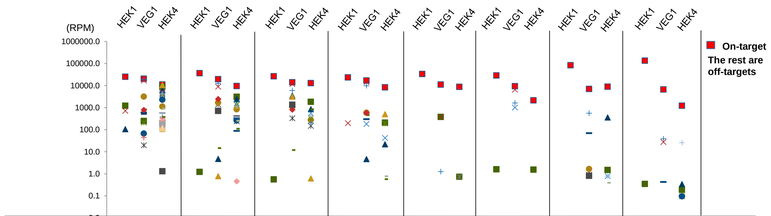
TEG-seq detected more off-targets and more (~10 fold) reads per target than the GUIDE-seq under similar total read output. 252 off-targets were detected by TEG-seq, and 142 off-targets were detected by GUIDE-seq for the gRNA targeting HEK4 gene. The on-targets were indicated by red color.

Figure 3: Detection level of off-target by TEG-seq and Targeted Amplicon-seq



The RPM (Reads Per Million) of off-targets detected by TEG-seq was plotted against the percentage of cleavage measured by Targeted Amplicon-seq. The correlation factor R value is 0.9070. With a single PGM run, the lowest off-target detected by TEG-seq and Targeted Amplicon-seq is 1 RPM and 0.0004% respectively. The on-target activity is indicated by red color.

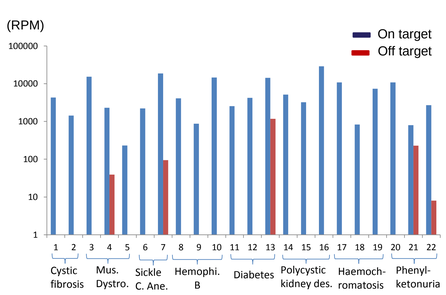
Figure 4: TEG-seq screening for HF Cas9 candidates



		wtCas9	Mutant-1	Mutant-2	Mutant-3	Mutant-4	Mutant-5	Mutant-6	Mutant-7	MIT score
Ratio of on/highest off	HEK1	20.3	29754.4	47157.3	118.1	>33635.4	17989.3	>83615.8	>135890	76
	VEG1	1.1	1.5	1.2	1.2	29.1	1.4	12.7	171.0	55
	HEK4	1.1	3.1	7.1	16.5	11848.2	1385.1	24.6	46.0	24

Genome-wide off-targets were screened for three gRNAs (HEK1, VEG1 and HEK4) co-transfected with different Cas9 mutants. Mutant-4 was identified to contain much less off-targets, but maintained relatively high on-target activity.

Figure 5. Genome-wide off-target detection for a set of high-score gRNAs targeting eight genetic disorder relevant SNPs



To test the gRNAs targeting the biological relevant SNPs that cause genetic disorders, we designed and chose 2-3 gRNAs with high *in silico* score (to minimize off-target) for each SNP and screened off-target using TEG-seq. Of the 22 gRNA, 17 (77%) have no detectable off-target, five (23%) resulted in detectable off-targets by TEG-seq that were subsequently confirmed by Targeted Amplicon-seq.

Table1: Example of TEG-seq report

			Sequence	PAM	(RNP control)		BC2 (Tag control)		BC3 (RNP+Tag)		
Target#	Chromsome_Position	Mismatch	GGGATTAGGCCAGAGAAGAG	GGG	Reads	RPM	Reads	RPM	Reads	RPM	TEG-Seq analysis
1	chr3:121618077	2	A...C..............	GGG	117	129	424	12	9,155,315	124,456	Potential Off-Target
2	chr7:105639123	6	TCCTA....T.........	AGG	132	146	281	8	6,768,854	92,010	Potential Off-Target
3	chr15:84172742	0	....	GGG	66	73	250	7	4,700,045	63,889	On-target
4	chr8:25205391	4	A...CA...T.........	AGG	55	61	232	7	4,688,236	63,724	Potential Off-Target
5	chr8:94912312	4	A..G.....T......G...	GGG	40	44	175	5	3,847,558	52,312	Potential Off-Target
6	chr16:94913434	2	.....G...T.........	TGG	58	64	170	5	3,531,743	48,016	Potential Off-Target
7	chr1:91482606	4	.TAGG..............	AGG	10	11	108	3	2,455,243	33,392	Potential Off-Target
8	chr7:102499203	4	..A..C...T.........A	TGG	16	18	93	3	1,729,138	23,518	Potential Off-Target
9	chr5:131458805	4	..TT.C.....G......	GGG	23	25	82	3	1,724,879	23,452	Potential Off-Target
10	chr4:125629412	5	AA.GGC.............	GGG	12	13	81	3	1,553,805	21,141	Potential Off-Target
11	chr4:119187375	5	CATG.........g.....	AGG	5	6	67	2	1,330,050	18,094	Potential Off-Target
12	chr12:116892219	2	A........A.........	AAG	21	23	35	1	1,035,280	14,063	Potential Off-Target
13	chr15:65307666	4	AA.C....T..........	TGG	7	8	55	2	945,067	12,848	Potential Off-Target
14	chr9:103848093	5	CA.G.C.......T.....	CGG	0	0	41	1	845,076	11,475	Potential Off-Target
15	chr19:40995245	4	A...CC...........G.	AGG	12	13	33	1	646,580	8,788	Potential Off-Target
16	chr13:71715265	3	AT..........A.....	TGG	3	3	31	1	615,346	8,369	Potential Off-Target
17	chr9:119929517	6	CCTG.C.....G......	AGG	5	6	29	1	500,693	6,803	Potential Off-Target
18	chr14:27301076	3	..C..C...G.........	GGG	0	0	34	1	491,162	6,686	Potential Off-Target
19	chr14:25669238	4	CA..C....T.........	AGG	0	0	24	1	490,135	6,669	Potential Off-Target
20	chr6:98727613	3	..C..C.............A	AAG	13	14	24	1	442,451	6,008	Potential Off-Target

Table2: Example of targeted amplicon-seq validation report

Target#	Chromsome_Position	Mismatch	GGGATTAGGCCAGAGAAGAG	Off-Target Sequence	GGG	Reads	RPM	% Indel via AmpliSeq	Total Reads (Ampli-seq)	Tag Ins	Large Del (>3bp)	Analysis
1	chr3:121618077	2	A...C..............	AGGACTAGGCCAGAGAAGAGGGG	GGG	9,155,315	124,456	88.39	862,674	y	y	Confirmed
2	chr7:105639123	6	TCCTA....T.........	TCCTATAGGTCAGAGAAGAGAGG	AGG	6,768,854	92,010	44.44	480,457	y	y	Confirmed
3	chr15:84172742	0	....	GGGATTAGGCCAGAGAAGAGGGG	GGG	4,700,045	63,889	85.93	334,609	y	y	Confirmed
4	chr8:25205391	4	A...CA...T.........	AGGACAAGGTCAGAGAAGAGAGG	AGG	4,688,236	63,724	56.65	223,876	y	y	Confirmed
5	chr8:94912312	4	A..G....T.....G....	AGGGTTAGGTCAGAGAGGAGGGG	GGG	3,847,558	52,312	15.33	117,708	y	y	Confirmed
6	chr16:94913434	2	.....G...T.........	GGGATTGGGTCAGAGAAGAGTGG	TGG	3,531,743	48,016	79.23	61,891	y	y	Confirmed
7	chr1:91482606	4	.TAGG..............	GTAGGTAGGCCAGAGAAGAGAGG	AGG	2,455,243	33,392	79.44	25,671	y	y	Confirmed
8	chr7:102499203	4	..A..C...T.....A	GGAATCAGGTCAGAGAAGAATGG	TGG	1,729,138	23,518	12.21	107,468	y	y	Confirmed
9	chr5:131458805	4	..TT.C.....G......	GGTTTCAGGCCAGGGAAGAGGGG	GGG	1,724,879	23,452	9.44	578,262	y	y	Confirmed
10	chr4:125629412	5	AA.GGC.............	AAGGGCAGGCCAGAGAAGAGGGG	GGG	1,553,805	21,141	12.35	94,313	y	y	Confirmed
11	chr4:119187375	5	CATG.........g.....	CATGTTAGGCCAGAGGAAGAGAGG	AGG	1,330,050	18,094	5.1	427,377	y	y	Confirmed
12	chr12:116892219	2	A........A.........	AGGATTAGGACAGAGAAGAAGAAG	AAG	1,035,280	14,063	27.15	16,849	y	y	Confirmed
13	chr15:65307666	4	AA.C....T..........	AAGCTTAGGTCAGAGAAGAGTGG	TGG	945,067	12,848	32.32	388,818	y	y	Confirmed
14	chr9:103848093	5	CA.G.C.......T.....	CAGGTCAGGCCAGATAAGAGCGG	CGG	845,076	11,475	1.8	894,602	y	y	Confirmed
15	chr19:40995245	4	A...CC...........G.	ATGATTAGGCCAGAAAAGAGTGG	TGG	615,346	8,369	62.05	6,312	y	y	Confirmed
16	chr13:71715265	3	AT..........A.....	CCTGTCAGGCCAGGGAAGAGAGG	AGG	500,693	6,803	2.59	197,811	y	y	Confirmed
17	chr9:119929517	6	CCTG.C.....G......	GGCATCAGGGCAGAGAAGAGGGG	GGG	491,162	6,686	3.41	68,462	y	y	Confirmed
18	chr14:27301076	3	..C..C...G.........	CAGACTAGGCAGAGAAGAGAGG	AGG	490,135	6,669	3.85	651,565	y	y	Confirmed
19	chr14:25669238	4	CA..C....T.........	GGCATCAGGCCAGAGAAGAAAAG	AAG	442,451	6,008	4.39	266,926	n	y	Confirmed
20	chr6:98727613	3	..C..C.............A									

CONCLUSION

1. Based on GUIDE-seq (1), we developed a cost-effective, high-throughput workflow of genome-wide CRIPSR off-target detection system (2).
2. TEG-seq improved specificity and sensitivity by ~10 fold compared to GUIDE-seq.
3. The sensitivity level reached 0.1 to 10 Reads Per Million (RPM) depending on the NGS platform used, which was equivalent to the low-frequency off-target at the level of 0.0002% to 1% measured by Targeted Amplicon-seq.
4. A panel of Cas9 mutants were evaluated using TEG-seq, in which a few candidates could be launched as high-fidelity version for potential application of gene and cell therapy.
5. The off-target result from a panel of gRNAs targeting genetic disorder relevant SNPs indicates it is likely to have at least one gRNA without any detectable off-target if 2-3 high score gRNAs can be designed through *in silico designing* tool.
6. TEG-seq was successfully applied to evaluate CRISPR off-target profiling for therapeutic targets in animal model and T-cell collaborated with our pharmaceutical customers (3).

REFERENCES

1. Tsai Q, *et al*. (2015) GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnology* 33:187-197
2. Tang P-Z, et al. (2018) TEG-seq: an ion torrent-adapted NGS workflow for in cellulo mapping of CRISPR specificity. BioTechniques. 65:259-267
3. Anderson K, et al. (2018) CRISPR off-target analysis in genetically engineered rats and mice. *Nat Methods*, 15:512-514

ACKNOWLEDGMENTS

Many thanks to CSD R&D Bioinformatics group (Yutao Fu and others) for developing analysis tools, John Bishop and Steven Roman for discussion of sample preparation and providing the reagents