Rapid Analysis of Genome Editing Efficiency using PCR Amplicons processed by the Exo-CIP™ Rapid PCR Cleanup Kit Followed by Sanger Sequencing

Pei-chung Hsieh, Ph.D. and Eric Cantor, Ph.D., New England Biolabs, Inc.

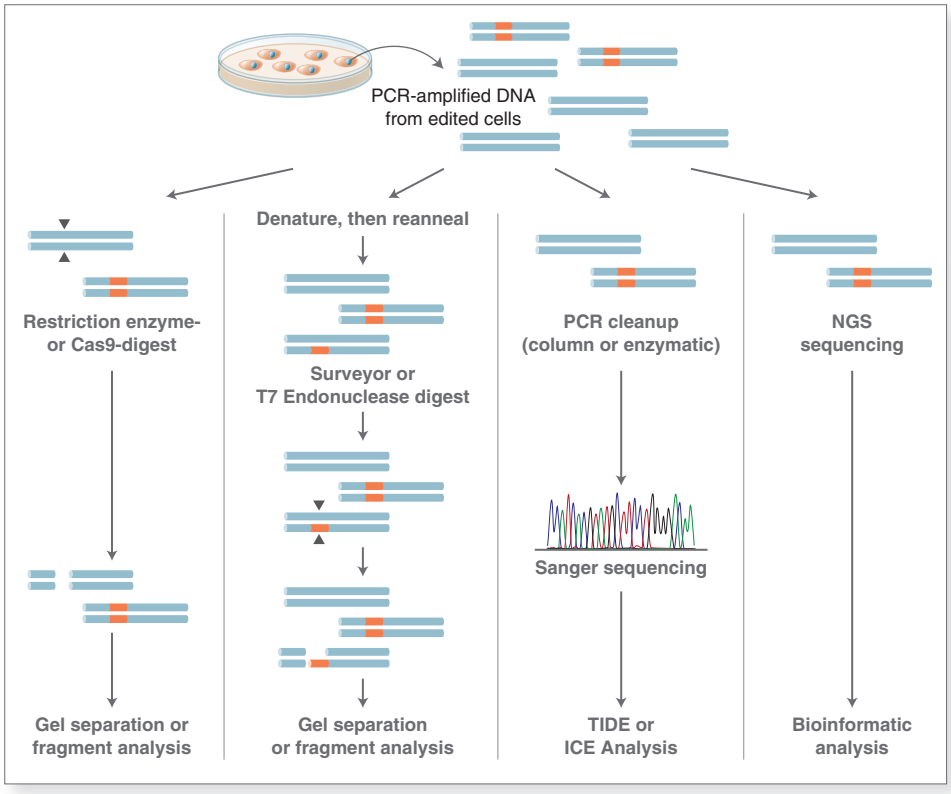
INTRODUCTION

The creation of targeted genome modifications is an exciting area of the life sciences. Originally enabled by zinc-finger nucleases (ZFNs) and transcription-activator-like effector nucleases (TALENs), recent advancements employ a bacterial CRISPR-associated protein-9 nuclease (Cas9) or CRISPR-Cas12 (Cpf1) (1,2). These tools induce sequence-specific double-strand breaks (DSBs) in cellular genomic DNA, which are then repaired by cellular repair pathways. Non-homologous end joining (NHEJ) is one DNA repair mechanism that often results in insertions or deletions (indels) at the break site. The repair of these genomic DNA breaks enables rapid knock-out of gene function. Homology-directed repair (HDR) is another genomic DNA repair pathway that is also stimulated by dsDNA breaks. In contrast to NHEJ, HDR uses a nucleic acid that spans the sequence of the break site as a template to repair the cleaved genomic DNA. The design and introduction of the HDR template enables precise insertion of any sequence at any location (knock-in) with base-pair resolution. These techniques have enabled robust genome editing, capable of initiating precise, targeted changes to the genome of many types of living cells (3).

MATERIALS

- Q5® High-Fidelity 2x Master Mix (NEB #M0492) or One*Taq*® 2X Master Mix (NEB #M0482)
- Exo-CIP™ Rapid PCR Cleanup Kit (NEB #E1050)
- Monarch® PCR & DNA Cleanup Kit (5 µg) (NEB #T1030)
- DNA templates and primers to generate amplicons with insertion/deletion/mismatch variants (Figure 2)

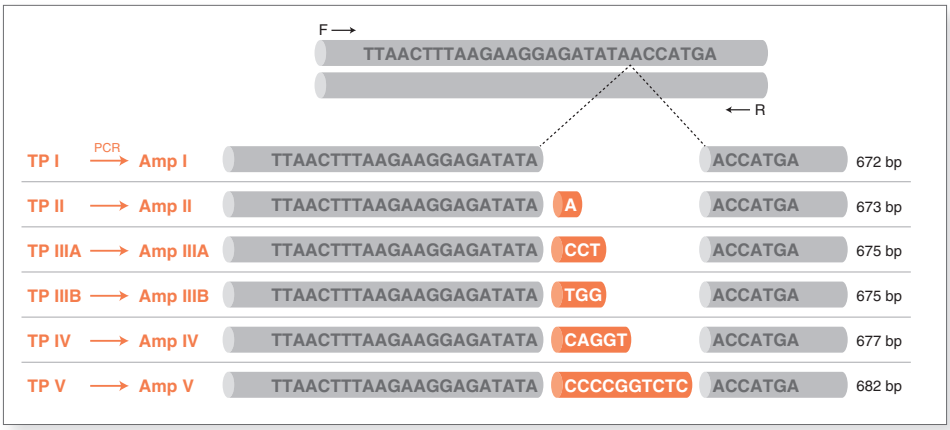
FIGURE 1:
Workflows for the analysis of genome editing outcomes



For CRISPR/Cas workflows, the choice of the nuclease and corresponding gRNAs directly affects the calculation of indel frequency after genome editing takes place. Current methods for evaluating editing efficiency use PCR amplification of the indel regions from pooled gDNA from transfected cells followed by either sequencing-based or mismatch cleavage-based analysis of denatured and reannealed indel DNA (4). To expedite the determination of editing efficiency and to avoid costly NGS sequencing, the 'tracking of indels by decomposition' (TIDE) (5) and 'inference of CRISPR edits' (ICE) methods (6) were developed to deconvolute the heterogenous populations of edits in PCR amplicons using Sanger DNA sequencing data. However, for these methods to be reliable, the PCR product analyzed must be of high quality (e.g., single band, free of primer and dNTPs). Herein, we demonstrate that the quality of amplicons cleaned up by the Exo-CIP Rapid PCR Cleanup Kit method matches that achieved by a traditional spin-column-based kit when used for batch analysis with the ICE software tool, thereby enabling a faster, higher-throughput method for preparing samples for Sanger sequencing post-amplification.

FIGURE 2: **Sequence of templates (TP) used to generate amplicons (Amp) with variants**

Letters in the orange boxes indicate the sequences of variance among templates, which are used for PCR to generate amplicons with expected population mixes.



QUICK HIGH-THROUGHPUT (HT) PCR CLEANUP PROTOCOL

In this study, amplicons were created by PCR using commercially-sourced DNA oligos. In genome editing workflows, these protocols can be applied after genome editing and extraction of genomic DNA.

Enzymatic cleanup with Exo-CIP Rapid PCR Cleanup Kit

1. Amplify target region using flanking primers and Q5 High Fidelity DNA Polymerase.
2. Confirm amplification was successful by running 1/10th of reaction on an agarose gel. Amplicon specific DNA band should be observed in 8–40 range/µl reaction compared to mass of a known gel standard.
3. Add 2 µl of Exo-CIP tube A and 2 µl of Exo-CIP tube B to 10 µl of PCR product. Save remainder of PCR reaction for spin column cleanup.
4. Incubate the mix at 37°C for 4 minutes.
5. Inactivate the reaction at 80°C for 1 minute.
6. Use 3-5 µl of the reaction (containing 40–200 ng) for Sanger sequencing. Design a primer region that is 150-200 bases from the PAM sequence site. Triplicate sequencing reactions for each PCR amplicon are suggested for ICE analysis.
7. Batch analyze each sequencing result using ICE software and average the percentage of each of the 3 variants predicted by ICE to assign a percentage of editing efficiency.

Samples for spin column cleanup

1. Use remaining PCR reaction (from step 3 above) for cleanup with a traditional spin column kit (e.g., Monarch PCR & DNA Cleanup Kit, NEB #T1030).
2. Analyze 2 µl of eluted material on microvolume spectrophotometer to quantitate.
3. Use 40-200 ng DNA for Sanger sequencing. Design a primer region that is 150–200 bases from the PAM sequence site. Triplicate sequencing reactions for each PCR amplicon are suggested for ICE analysis.
4. Batch analyze each sequencing result using ICE software and average the percentage of each of the 3 variants predicted by ICE to assign a percentage of editing efficiency.

RESULTS

Genome editing creates both deletions and insertions at various frequencies. Editing efficiency can be scored with a variety of methods, and the use of sequencing-based workflows is popular. These analyses typically require the user to amplify the targeted region with flanking primers, clean up the reaction, and submit it for Sanger sequencing before applying the analytical software packages described earlier (TIDE and ICE).

The results below examine the use of both spin-column or enzymatically cleaned up PCR products for sequencing and the resulting quality of the DNA in support of the ICE predictions for editing efficiency.

Abundance of edits predicted by ICE correlates well with expected variant distribution (either insertion or deletion) using samples sequenced from both column cleanup and Exo-CIP enzymatic cleanup.

PCR amplicons (p1–p11) were amplified by Q5 High-Fidelity 2X Master Mix using defined mixed template ratios (Figure 3) to generate mixed populations of insertion or deletion variants (determined by the identity of the template sequence). Each PCR product was cleaned up by either a Monarch spin column process (MC) or enzymatic cleanup with the Exo-CIP reagent (EC). As shown in Figure 3 (using TP II as the control sequence for analysis of the insertion efficiency of Amp IV), the edits for sequenced samples prepared from both methods, shown in light and dark orange circles, show strong agreement. In addition, these samples (based on EC prepared samples) also show consistency in the correlation (dark orange circle R^2=0.99) between the predicted percentages (determined by ICE) and expected percentages based on PCR template ratios (ranging from 5–95%). Conversely, when TP IV is used as the control sequence, the deletion population from Amp II (prepared from MC and EC) not only shows strong agreement but also shows a strong correlation (dark blue squares, R^2=0.99) between the ICE predicted percentage and expected percentage. We observe a 3–10% predicted variance depending on the insertion-based or deletion-based analysis, which we attribute to variance in how the software processes the signal. We observed similar results when we generated PCR amplicons using One*Taq* 2X Master Mix (data not shown).

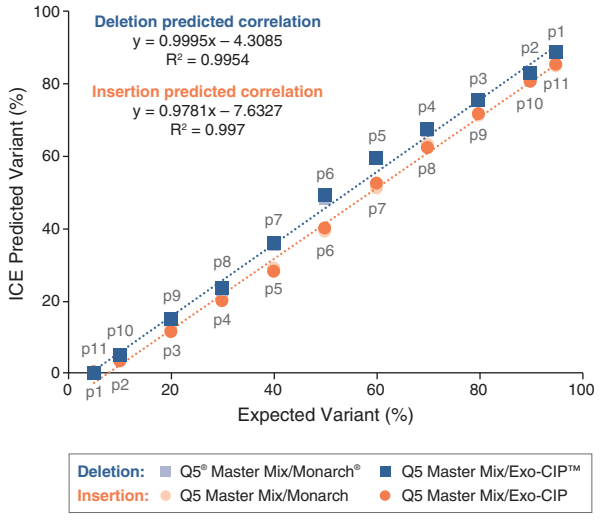
FIGURE 3: **ICE variant (InDel) predictions correlate well with expected variant distributions**

A. Tables depicts the composition of each pool (p1–p11) of amplification products which are treated as the expected insertion/deletion variant population.
B. Samples of ICE-predicted variants from column cleanup (light blue square for deletion variants and light circle for insertion variants) and enzymatic Exo-CIP cleanup (dark blue square for deletion variants and dark circle for insertion variants) are plotted against the expected variants. Both insertion-based and deletion-based analyses yield good correlation between expected and predicted variants population.

A

Defined Ratio of Mixed TP for PCR		Pool of Amp. Products	Insertion Percentage (Expected)	Deletion Percentage (Expected)
TP II	TP IV			
95	5	p1	5	95
90	10	p2	10	90
80	20	p3	20	80
70	30	p4	30	70
60	40	p5	40	60
50	50	p6	50	50
40	60	p7	60	40
30	70	p8	70	30
20	80	p9	80	20
10	90	p10	90	10
5	95	p11	95	5

B



Mixed edit populations containing both insertions and deletions are analyzed accurately using ICE with little difference observed from samples sequenced using either column or enzymatic cleanup of the amplicon.

PCR targets (p21–p28) were amplified by Q5 High-Fidelity 2X Master Mix (Q5MM) and/or One*Taq* 2X Master Mix (OTMM) using defined combinations of templates TPII , TP IIIA, and TP IV (ratios shown in Figure 4) to create heterogeneous populations of indels. Again, each PCR product was cleaned up by either a Monarch spin column (MC) or enzymatic treatment with the Exo-CIP reagent (EC). As shown in Figure 4, using TP IIIA set as the control sequence, the indel efficiency predictions of samples prepared from both cleanup methods (MC = light and EC = dark) are in good agreement regardless of polymerase used (Q5MM = orange circles; OTMM = blue squares). Additionally, these samples are well correlated (R^2=0.99 and R^2=0.99) with respect to the predicted and expected percentages of indel frequencies, ranging from 10–95%.

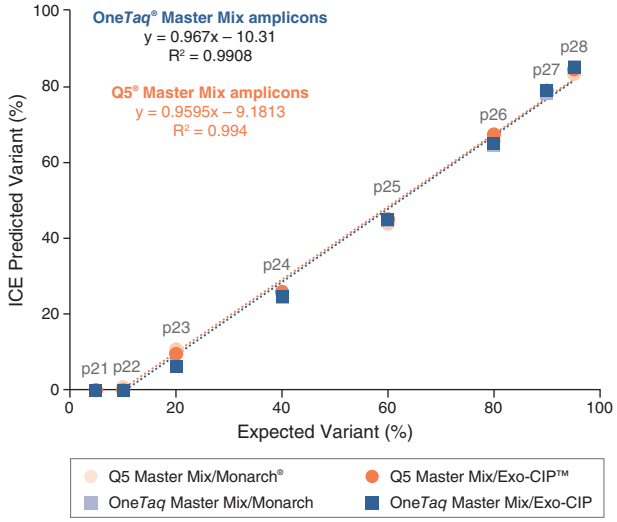
FIGURE 4: **ICE Indel variant predictions correlate well with expected InDel variant starting ratios**

A. Table depicts the composition of each pool (p21-p28) of amplification products which are treated as the expected variant population.
B. Samples of ICE-predicted variants from column cleanup (light blue square from One*Taq* DNA Polymerase Master Mix amplified variants and light circle from Q5 High Fidelity DNA Polymerase Master Mix amplified variants) and enzymatic Exo-CIP cleanup (dark blue square from One*Taq* DNA Polymerase Master Mix amplified variants and dark circle Q5 High Fidelity DNA Polymerase Master Mix amplified variants) are plotted against the expected variants. Analyses from both sample preparation yield good correlation between expected and predicted variants population.

A

Defined Ratio of Mixed TP for PCR			Pool of Amp. Products	Insertion/Deletion Percentage (Expected)
TP II	TP IIIA	TP IV		
2.5	95	2.5	p21	5
5	90	5	p22	10
10	80	10	p23	20
20	60	20	p24	30
30	40	30	p25	40
40	20	40	p26	50
45	10	45	p27	60
47.5	5	47.5	p28	70

B



Underestimation of ICE prediction increases for low abundance variants but is independent of methods of purification.

A known limitation of the ICE workflow can be seen when analyzing populations containing edits in lower abundance. To examine the influence of the purification method on the quality of the predictions, PCR amplicons were amplified by Q5 High-Fidelity 2x Master Mix and/or OneTaq 2X Master Mix using defined combinations of templates TP I, II, IIIA, IIIB, IV and V (ratios shown in Figure 5) to create PCR Amp pools (p31–38) with deletions or insertions as low as 5% of the total. These combinations rely on the use of insertion variants (p31–38) with Amp I as the control sequence or mismatch variants (p39–41) with TP IIIA as the control sequence. Each PCR product was cleaned up by either a Monarch spin column (MC) or enzymatically with the Exo-CIP enzyme cleanup reagent (EC). Additionally, cleanup was also performed with a widely-used spin column from another vendor to ensure our results aligned with generally accepted performance expectations. As shown in Figure 5, ICE predictions show that samples p31, 33, 34 and 35 contain several low-abundance variants as designed, resulting in more underestimation of predicted percentage as compared to p32, 36, 37 and 38, which have indels designed to be present at higher levels. We attribute this discrepancy to low peak heights in the electropherograms for each low-abundance variant that confounds calculations by the ICE software tool. Nevertheless, all three cleanup methods produce high-quality DNA samples for Sanger sequencing with similar performance (less than 5% difference) from running the ICE software. This concordance demonstrates that the higher-throughput enzymatic cleanup method (EC) is robust and matches the performance seen using the traditional spin column method.

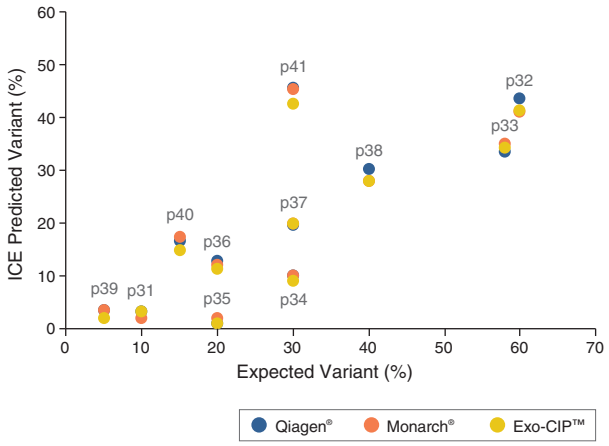
FIGURE 5: **ICE predicted variants percentage outcomes correlate well with InDel samples from Q column cleanup, M column cleanup and Exo-CIP cleanup**

A. Tables depicts the composition of each pool (p31–p41) of amplification products which are treated as the expected variant population.
B. Samples of ICE-predicted variants from column cleanup (blue circles by Qiagen DNA cleanup column and orange circles by Monarch DNA cleanup column) and enzymatic Exo-CIP cleanup (yellow circles) are plotted against the expected variants. Analyses from all sample preparation yield good correlation between expected and predicted variants population.

A

Defined Ratio of Mixed TP for PCR						Pool of Amp. Products	Insertion Percentage of Amp II, IIIA, IV, V	Mismatch Percentage of Amp IIIB
TP I	TP II	TP IIIA	TP IV	TP V	TP IIIB			
90	10					p31	10	
40	30	20				p32	60	
40	20	20	10	5		p33	58	
70	10	10	10			p34	30	
80	5	5	5	5		p35	20	
80	20					p36	20	
70	30					p37	30	
60	40					p38	40	
		95			5	p39		5
		85			15	p40		15
		70			30	p41		30

B



SUMMARY

While spin columns are a great option for preparing reproducibly clean and concentrated samples, their throughput is limited by the use of multiple steps for processing and reliance on microcentrifuges and/or vacuum manifolds. In an effort to provide our customers with solutions for higher throughput sample prep upstream of Sanger sequencing, NEB has developed the Exo-CIP Rapid PCR Cleanup kit. This kit uses thermolabile versions of two enzymes (Exo I nuclease and Calf Intestinal Phosphatase) to destroy ssDNA primers and dephosphorylate dNTP's in advance of the Sanger sequencing workflow in only 5 minutes. This approach provides a high-quality template for sequencing requiring fewer manipulations by the user, thereby increasing throughput for a larger number of samples. Enzymatic cleanup is a convenient way of preparing a PCR product for downstream applications or analysis. It combines the advantages of minimal hands-on time with virtually no sample loss, and enables high-quality sequencing results ready for analysis by the ICE software for rapid determination of genome editing efficiency. As such, NEB recommends the use of Exo-CIP PCR Cleanup Kit for high throughput applications requiring Sanger sequencing.

References

1. Jinek, M. et al. (2012) *Science*, 337, 816–21.
2. Fonfara, I. et al. (2016) *Nature*, 532, 517–21.
3. Nambiar, TS. (2019) *Nature Comm*, 10, 3395.
4. Germini, D. et.al. (2018) *Trends in Biotechnology*, 36, 147–159.
5. Brinkman, E.K. et al. (2014) *Nucleic Acids Research*, 42, e168.
6. Hsiau, T. et al. (2019) *BioRxiv*: 251082

New England Biolabs, Inc., 240 County Road, Ipswich, MA 01938-2723 Telephone: (978) 927-5054 Toll Free: (USA Orders) 1-800-632-5227 (USA Tech) 1-800-632-7799 Fax: (978) 921-1350 e-mail: info@neb.com