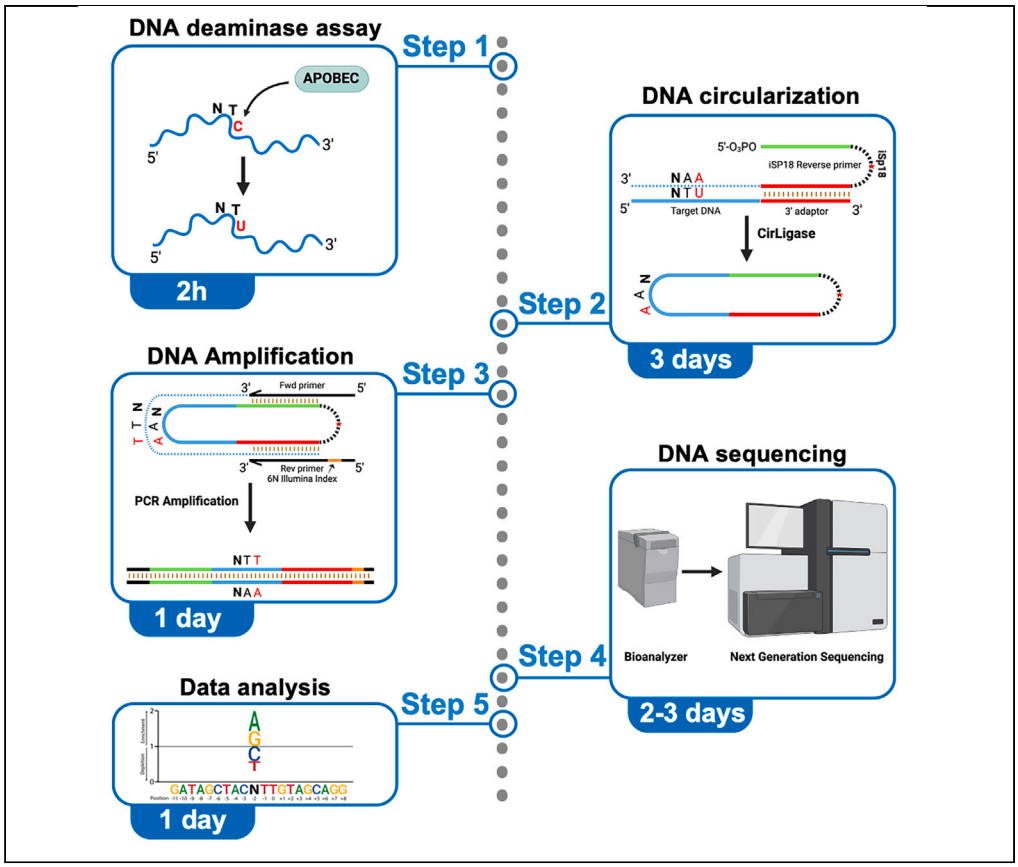
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Oligo-seq protocol for mapping DNA motifs targeted by base editors



Determining which DNA sequences are preferentially targeted by base editors is critical for understanding how APOBECs, AID, and other CRISPR-Cas9 base editors edit DNA in cells or improve their editing efficiency. We have developed Oligo-seq, an *in vitro* sequencing-based method to identify the preferred sequence motifs targeted by these enzymes. This assay monitors DNA deaminase activity on DNA oligonucleotides containing random nucleotides and/or DNA structures and determines by sequencing which sequences are preferentially deaminated.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Pedro Ortega, Ambrocio Sanchez, Marcus Seldin, Rémi Buisson

rbuisson@uci.edu

Highlights

Protocol for identifying DNA motifs targeted by base editors

Preparation of cell lysates containing APOBEC3B and assessment of its activity

Generation of oligonucleotide libraries following APOBEC3B editing

Data analysis and visualization of Oligo-seq results

Ortega et al., STAR Protocols 6, 103758
June 20, 2025
https://doi.org/10.1016/j.xpro.2025.103758

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Oligo-seq protocol for mapping DNA motifs targeted by base editors

Pedro Ortega,[1,2] Ambrocio Sanchez,[1,2] Marcus Seldin,[1,2] and Rémi Buisson[1,2,3,4,5,*]

[1]Department of Biological Chemistry, School of Medicine, University of California, Irvine, Irvine, CA, USA

[2]Chao Family Comprehensive Cancer Center, University of California, Irvine, Irvine, CA, USA

[3]Department of Pharmaceutical Sciences, School of Pharmacy and Pharmaceutical Sciences, University of California, Irvine, Irvine, CA, USA

[4]Technical contact

[5]Lead contact

*Correspondence: rbuisson@uci.edu
https://doi.org/10.1016/j.xpro.2025.103758

SUMMARY

Determining which DNA sequences are preferentially targeted by base editors is critical for understanding how APOBECs, AID, and other CRISPR-Cas9 base editors edit DNA in cells or improve their editing efficiency. We have developed Oligo-seq, an *in vitro* sequencing-based method to identify the preferred sequence motifs targeted by these enzymes. This assay monitors DNA deaminase activity on DNA oligonucleotides containing random nucleotides and/or DNA structures and determines by sequencing which sequences are preferentially deaminated.

For complete details on the use and execution of this protocol, please refer to Sanchez et al.[1]

BEFORE YOU BEGIN

This protocol aims to identify DNA sequence preferences specifically targeted by DNA deaminase enzymes. To illustrate this assay, we focus on the characterization of the DNA sequences recognized by APOBEC3B, a cytosine deaminase enzyme that causes mutations in tumors and viral genomes. However, this protocol can be applied to monitor other types of deaminase enzymes, including members of the APOBEC family, AID, or CRISPR-cas9 base editors.

The antiviral DNA cytosine deaminase APOBEC3B is a major source of mutations in tumor cells by catalyzing the deamination of cytosine to uracil (C to U) in DNA.[2–9] APOBEC3B deaminates cytosines that are preceded by a thymidine (TpC motifs) on single-stranded DNA (ssDNA).[9] The Oligo-seq assay revealed that APOBEC3B deaminase activity is strongly influenced by the specific sequences surrounding the targeted TpC motif.[1,10] We demonstrated that APOBEC3B preferentially targets stem-loop structures with a loop of five nucleotides and 5′-CCG or 5′-CCA sequence preceding the TpC motif.[1] Importantly, results obtained from Oligo-seq allow us to identify mutations induced by APOBEC3B in patient tumors present in similar DNA structures,[1] underscoring the importance of understanding the substrate preferences of DNA deaminase enzymes to explain how mutations form in tumors. This study highlights how results from Oligo-seq can be used to better understand the function of base editors against viral infection, promoting mutations in tumors, and their use in base editor technologies.

The Oligo-seq assay is based on the use of a base editor of interest that induces an alteration in a pool of synthetic ssDNA oligonucleotides containing random nucleotides around the target base. This modified nucleotide is then converted by a DNA polymerase to a different base during the

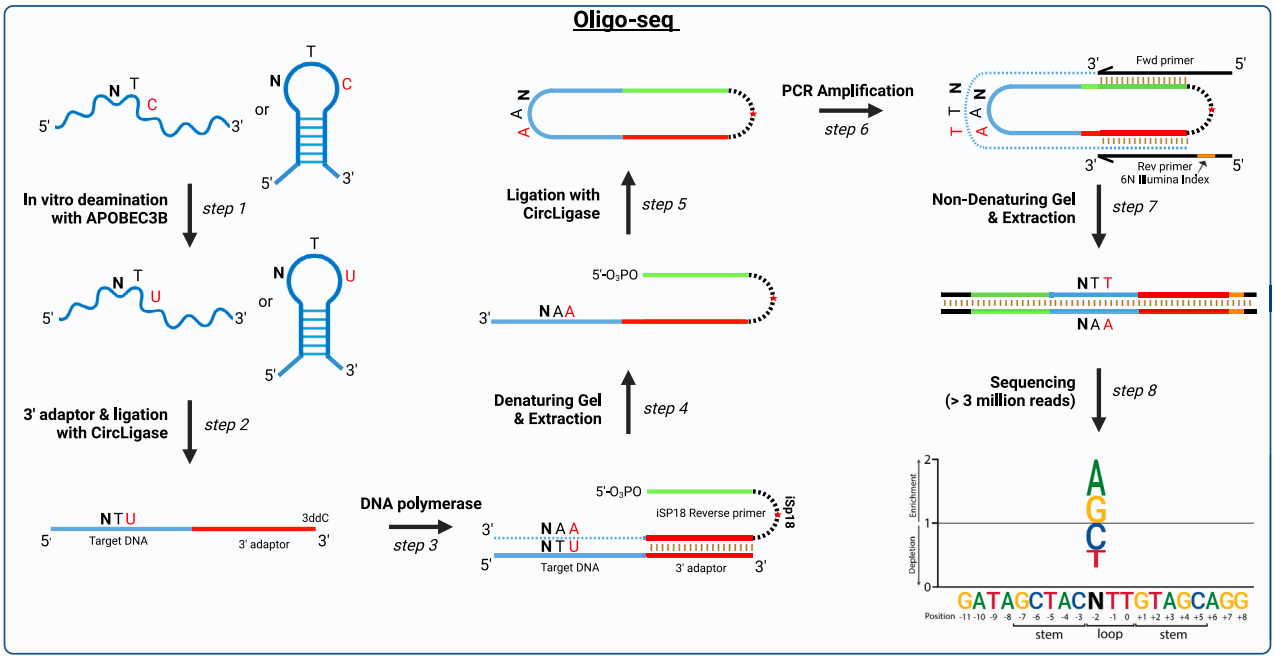


Figure 1. Oligo-seq assay overview

The Oligo-seq assay is based on the use of a base editor of interest that induces an alteration in a pool of synthetic ssDNA oligonucleotides containing random nucleotides around the target base. This modified nucleotide is then converted by a DNA polymerase to a different base during the sequencing library preparation, leading to a traceable footprint that can be quantified after sequencing (Figure 1). In this protocol, we focused on the deamination of the TpC motif by APOBEC3B, which is converted to TpT by the passage of the DNA polymerase. This assay can be used to monitor the DNA deaminase activity of other members of the APOBEC family but can also be adapted to adenine deaminase enzymes such as TadA deaminase and TadA-evolved base editors.

sequencing library preparation, leading to a traceable footprint that can be quantified after sequencing (Figure 1). In this protocol, we focused on the deamination of the TpC motif by APOBEC3B, which is converted to TpT by the passage of the DNA polymerase. This assay can be used to monitor the DNA deaminase activity of other members of the APOBEC family but can also be adapted to adenine deaminase enzymes such as TadA deaminase and TadA-evolved base editors.

Preparation of cell lysate expressing APOBEC3B

Timing: 3–4 days before the experiment

To study endogenous APOBEC3B substrate preference, we selected U2OS cells that express high protein levels of APOBEC3B.[11] Importantly, U2OS cells do not express APOBEC3A, which is also known to target TpC motif and could impact the results obtained from the Oligo-seq assay.[11] In addition, we depleted UNG2 from U2OS cells using siRNA to prevent uracil removal caused by APOBEC3B activity in ssDNA. Thus, UNG2 depletion prevents the formation of abasic sites, which can ultimately lead to DNA breakage during the sequencing library preparation.

1. Plate U2OS cells in a 10 cm culture plate (0.6×10^6 cells) in Dulbecco's modified Eagle's medium supplemented (DMEM) with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin.

   Note: The number of cells must be adjusted to obtain between 30% and 40% of cell confluency the following day.

Note: The number of 10-cm cell plates can be increased depending on the cell lines used and the level of APOBEC3B expression or other enzymes of interest.

2. 16–24 h after plating, transfect U2OS cells with a siRNA against UNG2 (siUNG2) using Lipofectamine RNAiMax following the manufacturer's protocol. Per reaction:
   a. Change the cell media approximately 1 h before transfection with 10 mL of fresh culture media without penicillin/streptomycin.
   b. Mix 1250 μL of Opti-MEM medium with 25 μL of Lipofectamine RNAiMax in a 1.5 mL microcentrifuge tube (Tube 1).
   c. Mix 1250 μL of Opti-MEM medium with 5 μL of 10 μM siUNG2 in a 1.5 mL microcentrifuge tube (Tube 2).
   d. Add the contents of Tube 1 to Tube 2 and mix thoroughly by pipetting or slowly vortexing for 5 s.
   e. Incubate for 20 min at room temperature (20°C–24°C).
   f. Add the transfection complex dropwise to the cells.
   g. Change the cell media with fresh media containing 1% penicillin/streptomycin 8–16 h after transfection.

Note: Other transfection methods can be used but have not been tested.

3. 48 h following transfection, U2OS cells were washed once with 7 mL of PBS 1X (4°C), trypsinized for 2–3 min, and collected in a 15 mL conical tube.

⚠ CRITICAL: All steps must be performed on ice (4°C) to prevent protein degradation.

4. Centrifuge the cells at 300 x g for 5 min at 4°C and remove supernatant.
5. Wash the cells twice by resuspending them in 10 mL of PBS, centrifuge them at 300 x g for 5 min at 4°C, and remove the supernatant.
6. Add 300 μL of Deaminase Lysis Buffer (See materials and equipment setup) to the cells, mix gently, and incubate for 5 min at 4°C.

⚠ CRITICAL: The volume of the lysis buffer should be adjusted based on the size of the cell pellet. In this study, we used five 10-cm U2OS cell culture plates, applying 300 μL of lysis buffer per plate.

7. Sonicate each sample twice at 60% amplitude for 10 s at 4°C.
8. Centrifuge the cell lysate at 20,000 x g for 5 min at 4°C to remove insoluble material.
9. Transfer supernatant to a 1.5 mL microcentrifuge tube and incubate with RNase A (0.2 mg/mL final concentration) for 20 min at 4°C.

⚠ CRITICAL: Incubation of cell extract with RNase A is essential to remove RNAs, which inhibit APOBEC3B deaminase activity.[12]

10. Centrifuge the cell lysate at 20,000 x g for 10 min at 4°C to remove insoluble material.
11. Transfer the supernatant to a new 1.5 mL microcentrifuge tube and quantify protein concentration using the Bradford Assay (Bio-Rad or equivalent) per the manufacturer's instructions.
12. Aliquot the cell lysates in 1.5 mL microcentrifuge tubes (20 μL aliquots), flash freeze, and store at −80°C.

Pause Point: The cell lysate can be stored at −80°C for up to one month without compromising the results.

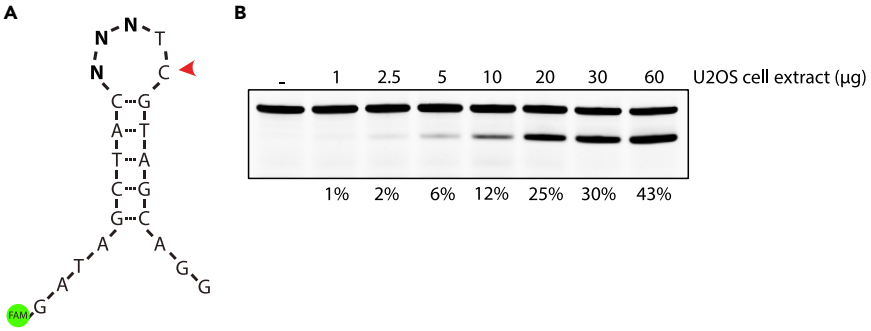


Figure 2. DNA deaminase assay

(A) Schematic of the 5-nt hairpin loop oligonucleotide used to perform the Oligo-seq assay. N refers to any of the four DNA bases.

(B) APOBEC3B deamination activity was performed with U2OS whole cell extract (1–60 µg) on the 5-nt hairpin loop oligonucleotide shown in A.

Alternatives: Not all cell lines express APOBEC3B or express APOBEC3B at levels too low to be detected using the Oligo-seq assay. Therefore, APOBEC3B can be ectopically expressed in cells using a vector expressing APOBEC3B. In addition, some cell lines express other APOBEC3 members, such as APOBEC3A, which can interfere with APOBEC3B activity. In this situation, we recommend purifying APOBEC3B from cell extract by immunoprecipitation as previously described.[1,13,14]

Optimization of the DNA deaminase reaction

Timing: 1 day before the experiment

The objective of this step is to determine the optimal concentration of cell lysate expressing APOBEC3B that achieves 5–10% deamination on the DNA target oligonucleotide. To evaluate the DNA deamination efficiency, we incubate an increasing amount of cell lysate in a reaction buffer containing purified recombinant Uracil DNA Glycosylase (UDG), which removes deaminated cytosines to generate an abasic site. The abasic site is then subjected to cleavage through treatment with NaOH and heat. The resulting oligonucleotide fragments are finally analyzed by gel electrophoresis, and the extent of deamination is determined through the quantification of the amount of cleaved DNA to assess the level of APOBEC3B activity in the cell extract.

13. Incubate the cell lysate expressing APOBEC3B with the DNA oligonucleotide containing a TpC motif and labeled with a 6-FAM fluorophore for detection and quantification (Figure 2A).

Note: The DNA sequence used in this protocol is: 5′-(6-FAM)-GATAGCTACNNN**TC**GTAGC AGG-3′. The underlined nucleotides indicate the sequence forming the stem. The TC motif targeted by APOBEC3B is indicated in bold. N indicates random nucleotides.

a. Add 42 µL of Deaminase Reaction Buffer (see materials and equipment setup) containing 60 pmol of DNA oligonucleotide and 1.5 U of purified recombinant UDG to a 1.5 mL microcentrifuge tube.
b. Add an increased amount of cell lysate (1 µg–60 µg) with a maximum volume of 8 µL per reaction. If less than 8 µL of cell lysate is used in each reaction, complement the reaction up to 8 µL with Deaminase Lysis Buffer (see materials and equipment setup) before adding the cell lysate to the reaction.

Note: Add a control sample that does not contain any cell lysate.

Note: The amount of cell extract should be adjusted based on the APOBEC3B expression levels.

14. Incubate the samples for 1 h at 37°C.

Note: We recommend using a thermocycler to maintain the proper temperature level.

15. Add 0.5 μL of NaOH (10 N) to the reaction and incubate the deaminated samples for 40 min at 95°C.

Note: We recommend using a thermocycler with a heated lid set up at 105°C to minimize evaporation.

16. Add 50 μL of Formamide (100%) containing Bromophenol Blue (0.03% m/v) to each sample.
17. Heat the samples at 95°C for 10 min, followed by 5 min incubation on ice.
18. Run the samples for 60 min at 250 V on a TAE-Urea (8 M) 20% denaturing polyacrylamide gel using a gel electrophoresis apparatus (Bio-Rad Protean II xi Cell, or equivalent).

⚠ CRITICAL: It is important to wash the wells extensively before loading the samples to remove urea residue.

Note: We recommend using a gel electrophoresis apparatus that can be connected to a circulating bath to heat the system at 60°C. This improves the migration of the DNA.

Note: We recommend pre-running the gel for 20 minutes before loading the samples.

Note: The percentage of the denaturing polyacrylamide gel needs to be optimized according to the length of the DNA oligonucleotide used.

19. Following migration, wash the gel for 5 min in H_2O.
20. Reveal the gel using a gel imaging system that detects FAM fluorescence (e.g., Bio-Rad Chemidoc MP Imaging System or equivalent) (Figure 2B).

Alternatives: Stain the DNA by incubating the gel in 1X SYBR Gold diluted in 1X TBE on a shaker for 10 minutes. This staining method can be used when the target oligonucleotide is not labeled with a dye such as 6-FAM.

Troubleshooting, problem 1.

21. Quantify the percentage of cleaved DNA products by comparing it to the amount of uncleaved DNA oligonucleotide in the control sample (Figure 2B).
22. Select the amount of cell lysate expressing APOBEC3B that yields between 5 to 10% of cleaved DNA products.

⚠ CRITICAL: It is important to limit the total amount of DNA oligonucleotide deaminated by APOBEC3B to 10%. High levels of deamination will decrease the sensitivity of the assay because poor DNA substrates for APOBEC3B will also be deaminated at high levels.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
CircLigase ssDNA ligase	Avantor	Cat# CL4115K
N,N,N′,N′-Tetramethylethylenediamine (TEMED)	Fisher Scientific	Cat# BP150-100
Bis-Acrylamide 19:1	Fisher Scientific	Cat# BP1406-1
Sodium chloride	Thermo Fisher Scientific	Cat# AM9760G
EDTA (pH 8)	Thermo Fisher Scientific	Cat# AM9260G
Tris-HCl (pH 7.4)	Thermo Fisher Scientific	Cat# AM9855G
2-propanol	Fisher Scientific	Cat# BP26181
SYBR Gold nucleic acid gel stain	Thermo Fisher Scientific	Cat# S11494
dNTP set	Thermo Fisher Scientific	Cat# 10297018
Phusion high-fidelity DNA polymerase	New England Biolabs	Cat# M0530S
GlycoBlue coprecipitant	Thermo Fisher Scientific	Cat# AM9515
TBE (10X), RNase-free	Thermo Fisher Scientific	Cat# AM9863
TAE (50X), RNase-free	Thermo Fisher Scientific	Cat# B49
Formamide	Fisher Scientific	Cat# BP227-500
Bromophenol blue	Fisher Scientific	Cat# B392-5
Novex Hi-Density TBE sample buffer	Thermo Fisher Scientific	Cat# LC6678
Phosphate-buffered saline (PBS) pH 7.4	Thermo Fisher Scientific	Cat# 10010023
DMEM/F12 GlutaMAX-I	Thermo Fisher Scientific	Cat# 10565-018
Fetal bovine serum (FBS)	Thermo Fisher Scientific	Cat# 10437-036
Triton X-100	Thermo Fisher Scientific	Cat# 85111
Magnesium chloride	Thermo Fisher Scientific	Cat# AM9530G
Halt protease and phosphatase inhibitor cocktail (PIC) 100X	Thermo Fisher Scientific	Cat# 78446
PureLink RNase A	Thermo Fisher Scientific	Cat# 12091021
Lipofectamine 2000 transfection reagent	Thermo Fisher Scientific	Cat# 11668019
Lipofectamine RNAiMAX reagent	Thermo Fisher Scientific	Cat# 13778075
Penicillin/streptomycin solution	Corning	Cat# 30-002-CI
Opti-MEM reduced serum medium	Thermo Fisher Scientific	Cat# 3198570
Glycerol	Thermo Fisher Scientific	Cat# G33-1
Uracil-DNA glycosylase	New England Biolabs	Cat# M0280S
Ammonium persulfate (APS)	Sigma-Aldrich	Cat# A682-500
Urea	Thermo Fisher Scientific	Cat# AC424580025
Sodium hydroxide solution (NaOH) 10 N	Fisher Scientific	Cat# SS255-1
Critical commercial assays		
Bradford assay	Bio-Rad	Cat# 5000002
Oligo clean & concentrator kit	Zymo Research	Cat# D4061
DNA clean & concentrator-5	Zymo Research	Cat# 11-302C
Qubit dsDNA high-sensitivity kit	Thermo Fisher Scientific	Cat# Q32851
Agilent high-sensitivity DNA kit	Agilent	Cat# 5067-4626
KAPA library quantification kit	Roche	Cat# 07960140001
Oligonucleotides		
20/100 oligo ladder	IDT	Cat# 51-05-15-02
Ultra low range DNA ladder	Thermo Fisher Scientific	Cat# 10597012
siUNG (s14678)	Thermo Fisher Scientific	Cat# 4427038
Other		
Qubit 4 fluorometer	Thermo Fisher Scientific	Cat# Q33238
PowerPac high-voltage power supply	Bio-Rad	Cat# 1645056
C1000 touch thermal cycler	Bio-Rad	Cat# 1851148
PROTEAN II xi cell for 2-D electrophoresis	Bio-Rad	Cat# 1651815
Agilent Bioanalyzer 2100	Agilent	Cat# G2938A.
Platform rotator	Thermo Fisher Scientific	Cat# 88861045

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Cell lines		
U2OS	Lee Zou laboratory	N/A
Software and algorithms		
Oligo-seq pipeline	This study	https://github.com/ambrociosanchezphd/Oligo-Seq-Pipeline
RStudio (2024.04.1+748)	https://www.r-project.org/	https://www.r-project.org/
R v.4.4.0	https://www.r-project.org/	https://www.r-project.org/

MATERIALS AND EQUIPMENT

Oligonucleotides sequences

Target oligonucleotide: 5′-GATAGCTACNNNTCGTAGCAGG-3′

The underlined nucleotides indicate the sequence forming the stem. The TpC motif targeted by APOBEC3B is indicated in bold. N indicates random nucleotides.

⚠ CRITICAL: The target oligonucleotide must have a GG motif at the 3′ end of the sequence. The GG motif is essential for the ligation using CircLigase (see steps 5 to 22 of the section step-by-step method details).

Note: The target oligonucleotide can be adapted in function of the base editor of interest and the context of DNA folding to be studied. Here, we selected a single-stranded DNA that forms a 5-nucleotide (nt) loop hairpin with a stem of 5 base pairs (bp), which we previously described as one of the preferred substrates for APOBEC3B.[1]

Oligo 3′ adaptor: /5Phos/AGATCGGAAGAGCACACGTCTGAA/3ddC/

Note: 3′ Dideoxycytidine (3ddC), a 3′ chain terminator, is added to the oligonucleotide to prevent 3′ extension by DNA polymerases.

Note: 5′ Phosphorylation (5Phos) is necessary for DNA ligation by CircLigase.

iSP18 reverse primer:

5′-Phos/AGATCGGAAGAGCGTCGTGTAGGGAAAGAG/iSp18/GTGACTGGAGTTCAGACGTGTGCTC-3′

Note: 18-atom hexa-ethyleneglycol spacer (iSp18) is an ethylene glycol spacer that blocks the elongation of DNA polymerases to avoid rolling-circle amplification of circular DNA during the library construction PCR.

P5 Forward primer: 5-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC-3.

P7 Reverse primer:

5′-CAAGCAGAAGACGGCATACGAGATNNNNNNGTGACTGGAGTTCAGACGGTGTGCTCTTCCGATCT-3′

(NNNNNN denotes barcode sequences)	
Index barcode sequence (NNNNNN)	Sequence
Index 1	CGTGAT
Index 2	ACATCG
Index 3	GCCTAA
Index 4	TGGTCA
Index 5	CACTGT
Index 6	ATTGGC
Index 7	GATCTG
Index 8	TCAAGT
Index 9	CTGATC
Index 10	AAGCTA
Index 11	GTAGCC
Index 12	TACAAG

Note: Different indexed reverse primer sequences can be used to barcode different samples during the library construction. This will allow the sequencing of up to 12 different samples at the same time.

Deaminase lysis buffer			
Reagent	Stock concentration	Final concentration	Amount
HEPES (pH 7.4)	1 M	25 mM	25 μL
Glycerol	100%	10%	100 μL
NaCl	5 M	150 mM	30 μL
$MgCl_2$	1 M	1 mM	1 μL
$ZnCl_2$	0.5 M	1 mM	2 μL
Triton X-100	100%	0.5%	5 μL
PIC	100X	2X	20 μL
Nuclease-free H_2O	N/A	N/A	798 μL
Total	**N/A**	**N/A**	**1 mL**

Note: We recommend making the buffer one day before each use and storing it at 4°C. Protease and Phosphatase Inhibitor Cocktail (PIC) and Triton X-100 should be added just before lysing the cells (see step 6, in the section before you begin).

Deaminase reaction buffer			
Reagent	Stock concentration	Final concentration	Amount
Tris-HCl (pH 7.4)	1 M	50 mM	500 μL
EDTA (pH 8.0)	0.5 M	10 mM	200 μL
NaCl	4 M	30 mM	75 μL
ZnCl2	0.5 M	0.25 mM	5 μL
Nuclease-free H_2O	N/A	N/A	9.3 mL
Total	**N/A**	**N/A**	**10 mL**

Note: Store at 4°C for up to 1 month.

CircLigase reaction mix			
Reagent	Stock concentration	Final concentration	Amount
CircLigase reaction buffer	10X	1X	2 μL
ATP	1 mM	0.05 mM	1 μL
$MnCl_2$	50 mM	2.5 mM	1 μL
CircLigase	100 U/μL	5 U/μL	1 μL
Nuclease-free H_2O	N/A	N/A	5 μL
Total	**N/A**	**N/A**	**10 μL**

Note: To be made just before each use.

DNA gel extraction buffer			
Reagent	Stock concentration	Final concentration	Amount
NaCl	5 M	300 mM	600 μL
Tris-HCl (pH 7.4)	1 M	10 mM	100 μL
EDTA (pH 8.0)	0.5 M	1 mM	20 μL
Nuclease-free H_2O	N/A	N/A	9280 μL
Total	**N/A**	**N/A**	**10 mL**

Note: Store at room temperature (20°C–24°C) for up to 1 month.

2X denaturing loading buffer			
Reagent	Stock concentration	Final concentration	Amount
EDTA (pH 8.0)	0.5 M	10 mM	200 μL
Formamide	100%	98%	9.8 mL
Bromophenol blue	N/A	0.03%	3 mg
Total	**N/A**	**N/A**	**10 mL**

Note: Store at −20°C for up to 1 year.

TAE-Urea denaturing polyacrylamide gel			
Reagent	Stock concentration	15% gel	20% gel
Urea	N/A	24 g	24 g
TAE	50X	1 mL	1 mL
Bis-Acrylamide 19:1	40%	18.75 mL	25 mL
H_2O	N/A	qs to 50 mL	qs to 50 mL
TEMED	100%	50 μL	50 μL
APS	30%	500 μL	500 μL
Total	**N/A**	**50 mL**	**50 mL**

Note: Store at room temperature (20°C–24°C) for up to 1 month.

Note: APS and TEMED should be added just before adding the solution to the gel casting system.

⚠ CRITICAL: Do not store the solution at temperatures below 20°C–24°C, as it would cause urea precipitation.

⚠ CRITICAL: The solution should be filtered to remove any urea precipitation or impurities by using a bottle-top vacuum filter system (0.45 μm).

TBE native polyacrylamide gel		
Reagent	Stock concentration	10% gel
TBE	10X	5 mL
Bis-Acrylamide 19:1	40%	12.5 mL
H_2O	N/A	qs to 50 mL
TEMED	100%	50 μL
APS	30%	500 μL
Total	**N/A**	**50 mL**

Note: APS and TEMED should be added just before adding the solution to the gel casting system.

Note: Store at room temperature (20°C–24°C) for up to 1 month.

STEP-BY-STEP METHOD DETAILS

Note: We selected APOBEC3B deaminase as an example to illustrate the Oligo-seq method. However, this assay can be easily adapted to other APOBEC family members or DNA deaminase enzymes. Moreover, we selected a DNA substrate that forms a hairpin structure that is preferentially targeted by APOBEC3B.[1]

Deaminase reaction

Timing: 2 h

These steps describe how to perform the deaminase reaction on the target oligonucleotide of interest (Figure 1, step 1).

We selected a 22-nt DNA oligonucleotide (target oligonucleotide) containing a TpC motif within a 5-nt loop of stem-loop structure. We have previously shown that APOBEC3B preferentially targets this type of structured DNA.[1]

Note: A change in the length of the oligonucleotide will result in a change in the size of the DNA in the steps described below.

1. Incubate the cell lysate expressing APOBEC3B (see steps 1 to 12 of the section before you begin) with the target oligonucleotide.
   a. Add 42 μL of reaction buffer containing 60 pmol of target oligonucleotide to a 1.5 mL microcentrifuge tube.
   b. Add the appropriate amount of cell lysate to deaminate 5–10% of the target oligonucleotide (determined in steps 13 to 22 before beginning section). Before adding the cell lysate to the reaction, add Deaminase Lysis Buffer (see materials and equipment setup) to complement to a total volume of 8 μL with the cell lysate volume.

⚠ CRITICAL: Purified recombinant UDG should not be used in the reaction.

2. Incubate the samples for 1 h at 37°C.

Note: We recommend using a thermocycler to maintain the proper temperature level.

Pause Point: The samples can be stored at −20°C or −80°C for at least a month without compromising the results.

Purification of the deaminated oligonucleotide

Timing: 10 min

This step describes how to purify the deaminated target oligonucleotide of interest.

3. Purify the DNA oligonucleotide from the reaction using the Oligo Clean & Concentrator kit following the manufacturer's protocol.

Alternatives: Purification methods from other manufacturers can also be used but has not been tested.

4. Elute the target oligonucleotide from the columns with 6 μL of 10 mM Tris-HCl (pH 7.4).

Pause Point: Purified oligonucleotide can be stored long-term at −20°C without compromising the results.

Ligation of an adaptor to the deaminated oligonucleotide

Timing: 2 days

This step describes how to ligate the deaminated target oligonucleotide of interest to an oligonucleotide adaptor (Figure 1, step 2).

5. Add 4 μL of the Oligo 3′ adaptor (50 μM) to the 6 μL of the purified oligonucleotide.
6. Perform the ligation of the adaptor to the target oligonucleotide by adding 10 μL of the CircLigase Reaction Mix (see materials and equipment setup) to the samples (10 μL) in 200 μL microcentrifuge tubes.
7. Mix thoroughly by pipetting, then briefly centrifuge to collect the sample at the bottom of the microcentrifuge tube.
8. Incubate the CircLigase reaction in a thermocycler overnight (16 h) at 60°C (lid at 95°C).

Note: A shorter incubation time with CircLigase will lead to a significant decrease in ligation products.

9. Add 20 μL of 2X Denaturing Loading Buffer (see materials and equipment setup).
10. Denature the samples for 1 min at 90°C and incubate them on ice until they are loaded into the gel.
11. Run the samples for 3 h at 300 V on a TAE-Urea (8 M) 20% denaturing polyacrylamide gel (see materials and equipment setup) using a gel electrophoresis apparatus (Bio-Rad Protean II xi Cell, or equivalent).

⚠ CRITICAL: It is essential to wash the wells extensively before loading the samples to remove urea residue.

⚠ CRITICAL: In addition to loading the samples, it is important to add a DNA ladder for the identification of the correct band sizes in the subsequent steps. Therefore, load 20 μL (100 ng) of 20/100 DNA oligo length standards (0.1 μg/μL) that was prediluted (1:20) in 2X denaturing loading buffer.

Note: We recommend using a gel electrophoresis apparatus that can be connected to a circulating bath to heat the system at 60°C. This improves the migration of the DNA.

Note: Select a comb size that can accommodate a 40 μL sample.

Note: We recommend pre-running the gel for 20 min at 300 V before loading the samples.

Note: The expected ligated product is 47 nt. If the template DNA oligonucleotide used has a different length, the percentage of acrylamide should be adapted.

Alternatives: Other electrophoresis gel system apparatus can be used but have not been tested. We recommend using a large gel apparatus (16 cm × 20 cm) and 1.5 mm spacers

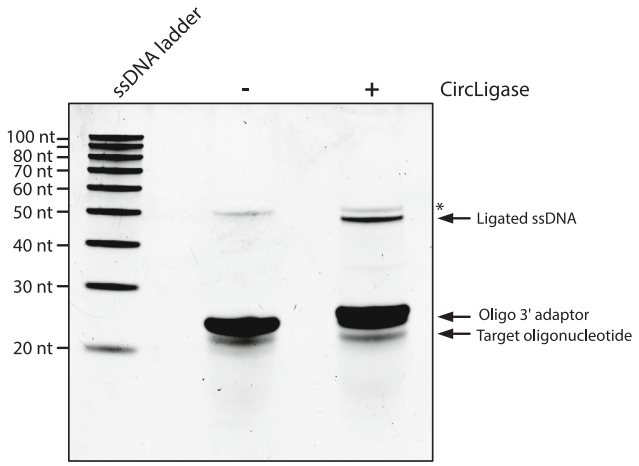


Figure 3. Ligation of an adaptor to the deaminated oligonucleotide

Example of a TAE-Urea 20% denaturing polyacrylamide gel showing the ligation of the Oligo 3′ adaptor to the target oligonucleotide by the CircLigase. The asterisk (*) indicates a nonspecific band. ssDNA ladder: 20/100 oligo ladder.

for optimal separation of the DNA products and to obtain sharper bands, which will facilitate the subsequent band excision process.

12. Stain the DNA by incubating the gel in 1X SYBR Gold diluted in 1X TAE on a shaker for 10 min.
13. Visualize the gel with a gel imaging system with SYBR detection capacity (e.g., Bio-Rad Chemidoc MP Imaging System or equivalent) (Figure 3).

Alternatives: Other electrophoresis gel visualization system apparatus can be used but have not been tested.

Troubleshooting, problem 2.

Troubleshooting, problem 5.

14. Excise the ligated ssDNA (47-nt) band corresponding to the size of the target oligonucleotide (22 nt) ligated to the Oligo 3′ adaptor (25-nt), and collect it in a 1.5 mL microcentrifuge tube.

Note: The size of the ligated ssDNA will vary in function of the target oligonucleotide length.

15. Add 400 μL of DNA Gel Extraction Buffer (see materials and equipment setup) and flash-freeze the samples for 30 min using dry ice.
16. Extract the ligated ssDNA by allowing the samples to thaw overnight at room temperature (20°C–24°C) while gently mixing on a platform rotator.

Alternatives: Column-based DNA gel extraction methods could be used to extract DNA from the get, but have not been tested in this protocol.

17. Centrifuge the samples at 500 x g for 30 s and transfer the supernatants (400 μL) to new 1.5 mL microcentrifuge tubes.
18. **Optional:** Add 1.5 μL of GlycoBlue coprecipitant (15 mg/mL). GlycoBlue coprecipitant is used only to visualize the DNA pellet during DNA precipitation, decreasing the risk of losing it. The GlycoBlue step can be omitted without compromising the quality of the results.

19. Precipitate the extracted ligated ssDNA by adding 500 μL of isopropanol (100%), mix thoroughly, and incubate the samples for 2 h at −20°C.

Note: DNA precipitation can also be performed overnight at either −20 or −80°C.

20. Pellet the ligated ssDNA by centrifugation at 20,000 x g for 30 min at 4°C.
21. Carefully discard the supernatant and air-dry the pellet for 10 min.

⚠ CRITICAL: To prevent potential inhibition of subsequent steps, ensure that there are no residual traces of isopropanol in the samples.

22. Resuspend the pellet in 5 μL of nuclease-free H_2O.

Troubleshooting, problem 3.

Pause Point: Samples can be stored at −20°C or −80°C for a month without compromising the results.

Reverse strand extension

Timing: 1 day

This step describes the reverse extension of the ligated ssDNA oligonucleotide with a primer sequence that will be used to generate a circular template for PCR amplification (Figure 3, step 3–4).

23. Add 20 μL of Extension Reaction Mix (see materials and equipment setup) to the 5 μL of ligated ssDNA obtained in step 22.

Extension reaction mix	
Reagent	Amount
15 μM iSP18 reverse primer	2 μL
10 mM dNTP mix	1 μL
5X Phusion HF buffer	5 μL
Phusion polymerase	0.5 μL
Nuclease-free H_2O	11.5 μL
Total	**20 μL**

Note: The iSP18 reverse primer contains a Spacer 18 that is used to incorporate a long spacer arm into an oligonucleotide. Spacer 18 is a hexa-ethyleneglycol spacer of 18 atoms long (12 carbons + 6 oxygens) and is used to form DNA hairpin looped structures and bulged duplex structures.[15]

Note: The 5X Phusion HF buffer is provided with the Phusion High-Fidelity DNA Polymerase.

24. Perform the primer extension reaction as described in the table below in a thermocycler with the lid temperature set up at 105°C.

Note: When working with different lengths of oligonucleotide templates, we recommend assessing primer extension temperature and time to maximize template extension.

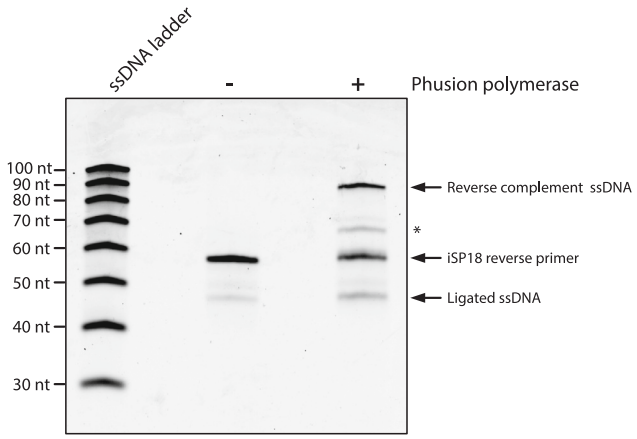


Figure 4. Extension of the ligated ssDNA

Example of a TAE-Urea 15% denaturing polyacrylamide gel showing the extension of the ligated ssDNA by the Phusion polymerase using the iSP18 reverse primer. The asterisk (*) indicates an incomplete extension product. ssDNA ladder: 20/100 oligo ladder.

Extension reaction conditions		
Steps	Temperature	Time
Preheat	98°C	30 s
Denaturation	98°C	5 min
Annealing	55°C	5 min
Extension	72°C	20 min
Hold	4°C	forever

25. Add 25 µL of 2X Denaturing Loading Buffer (see materials and equipment setup) to each sample.
26. Heat the samples for 2 min at 95°C and incubate on ice until loading in the gel.
27. Run the samples for 3 h at 300 V on a TAE-Urea (8 M) 15% denaturing polyacrylamide gel as described in step 11.
28. Stain the gel by incubating it in 1X SYBR Gold diluted in 1X TAE on a shaker for 10 min.
29. Visualize the gel with a gel imaging system with SYBR detection capacity (e.g., Bio-Rad Chem-idoc MP Imaging System or equivalent) (Figure 4).

Alternatives: Other electrophoresis gel visualization system apparatus can be used but have not been tested.

Troubleshooting, problem 2.

Troubleshooting, problem 5.

30. Excise the reverse-complemented ssDNA band (86-nt) corresponding to the size of the ligated ssDNA (47-nt) plus the iSP18 reverse primer (55-nt minus the 16-nt overlapping sequence with the ligated ssDNA)

Note: The size of the reverse-complemented ssDNA will vary in function of the target oligo-nucleotide length.

31. Extract and precipitate the DNA from the polyacrylamide gel slices as described in steps 15 to 21.
32. Resuspend the DNA pellets in 15 µL of 10 mM Tris-HCl (pH 7.4).

Troubleshooting, problem 3.

Pause Point: Samples can be stored at −20°C or −80°C for a month without compromising the results.

Single-stranded DNA circularization

Timing: 16 h

This step describes the circularization of the single-stranded DNA to form the DNA template required to amplify the Illumina-compatible library (Figure 1, step 5).

33. To circularize the DNA oligonucleotide, add 5 μL of the CircLigase Reaction Mix (see materials and equipment setup) to the reverse-complemented DNA samples (15 μL) obtained in step 32.
34. Mix thoroughly by pipetting, then briefly centrifuge to collect the sample at the bottom of the microcentrifuge tube.
35. Incubate the CircLigase reaction in a thermocycler overnight (16 h) at 60°C (lid at 95°C).

Note: A shorter incubation time with CircLigase leads to a significant decrease in ligation products.

Pause Point: Samples can be stored at −20°C or −80°C for a month without compromising the results

Library construction PCR

Timing: 1 day

This step describes how to amplify the target DNA from the circularized ssDNA to generate double-stranded DNA containing flanking Illumina-compatible sequences (Figure 1, step 6).

36. Perform a PCR by adding 45 μL of the PCR reaction mix to 5 μL of circularized ssDNA samples, obtained in step 35, in 200 μL microcentrifuge tubes.

Note: Select a different index reverse library primer sequence for each sample (see materials and equipment setup).

Alternatives: Other index sequences can be selected to adapt to different sequencer systems but have not been tested.

PCR reaction master mix

Reagent	Amount
10 mM dNTP mix	1 μL
10 μM Forward library primer	2 μL
10 μM Reverse library primer	2 μL
5x Phusion HF Buffer	10 μL
Phusion polymerase	0.5 μL
Nuclease-free H_2O	29.5 μL
Total	**45.0 μL**

Note: The 5X Phusion HF buffer is provided with the Phusion High-Fidelity DNA Polymerase.

37. Mix thoroughly by pipetting, then briefly centrifuge to collect the sample at the bottom of the microcentrifuge tube.

38. Put the samples in a thermocycler and run as indicated:

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Preheat	98°C	30 s	1
Denaturation	98°C	15 s	
Annealing	55°C	5 s	8-15 cycles
Extension	65°C	10 s	
Hold	4°C	forever	

Note: The expected PCR product size is 145 bp, corresponding to the size of the reverse complement ssDNA (86-nt) plus the Reverse library primer (65 nt minus 34-nt overlap) and the Forward library primer (49 nt minus 21-nt overlap). The size of the dsDNA will vary in function of the target oligonucleotide length.

⚠ CRITICAL: The number of PCR cycles needs to be optimized according to the DNA amount obtained in the gel after the PCR amplification (Figure 5). We recommend testing 8, 10, 12, and 15 cycles and selecting the number of cycles that lead to a clear band amplification at 145 bp on the gel while also minimizing non-specific amplification.

39. Purify the DNA oligonucleotide from the PCR reaction using the DNA Clean & Concentrator-5 following the manufacturer's protocol.

Alternatives: DNA purification using kits and protocols from other manufacturers can also be used but has not been tested.

40. Elute the samples in 6 μL of 10 mM Tris-HCl (pH 7.4) in 1.5 mL microcentrifuge tubes.
41. Add 14 μL of 5X Novex Hi-Density TBE Sample Buffer to each sample.

Note: Novex Hi-Density TBE Sample Buffer contains Ficoll, which yields sharper, straighter bands than conventional density agents, and dyes Bromophenol Blue and Xylene Cyanol to track the migration. Other DNA loading buffers could be used but have not been tested.

42. Run a TBE-10% native polyacrylamide gel (see materials and equipment setup) for 2.5 h at 180 V using a gel electrophoresis apparatus (Bio-Rad Protean II xi Cell, or equivalent).

⚠ CRITICAL: In addition to loading the samples, it is important to add a double-stranded DNA (dsDNA) ladder for the identification of the correct band sizes in the subsequent steps. Therefore, load 20 μL (0.5 μg) of Ultra Low Range DNA Ladder (0.5 μg/μL) that was prediluted (1:20) in 5X Novex Hi-Density TBE Sample Buffer.

Note: We recommend pre-running the gel for 20 min at 180 V before loading the samples.

Alternatives: Other electrophoresis gel system apparatus can be used but have not been tested. We recommend the use of a large gel apparatus (16 cm × 20 cm) and 1.5 mm spacers for optimal separation of the DNA products and to obtain sharper bands, which will facilitate the subsequent band excision process.

43. Stain the gel with 1X SYBR Gold in 1X TBE running buffer on a shaker for 10 min.
44. Visualize the gel with a gel imaging system with SYBR detection capacity (e.g., Bio-Rad Chemidoc MP Imaging System or equivalent) (Figure 5).

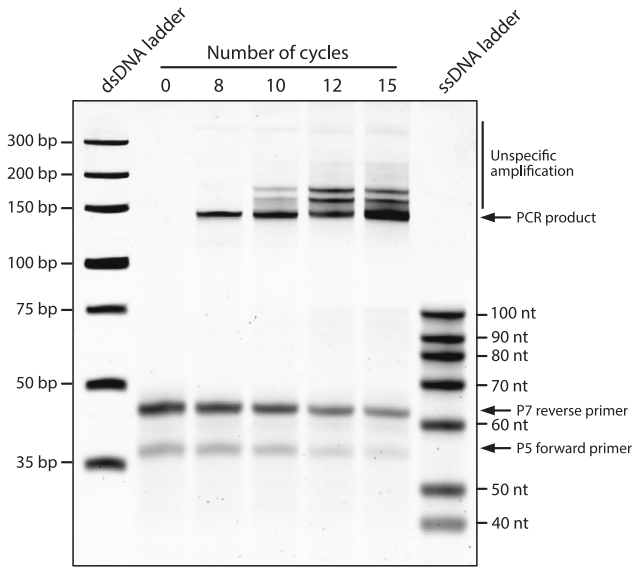


Figure 5. PCR amplification of the circularized ssDNA

Example of a TBE 10% native polyacrylamide gel showing the amplification by PCR of the circularized ssDNA obtained in step 35 by using P5 forward and P7 reverse primers. The migration of P5 forward (49 nt) and P7 reverse primers (66 nt) are migrating with a slight difference from their predicted migration sizes, likely due to the formation of secondary structures. ssDNA ladder: 20/100 oligo ladder. dsDNA ladder: Ultra Low Range DNA Ladder.

Alternatives: Other electrophoresis gel visualization system apparatus can be used but have not been tested.

Troubleshooting, problem 2.

Troubleshooting, problem 5.

45. Excise the PCR product band (145 bp) corresponding to the size of the amplified Illumina-compatible library dsDNA sample.
46. Extract and precipitate the PCR product from the polyacrylamide gel slices as described in steps 15 to 21.
47. Resuspend DNA pellets in 5 μL of 10 mM Tris-HCl (pH 7.4).

Troubleshooting, problem 3.

Pause Point: Samples can be stored at −20°C or −80°C for a month without compromising the results.

48. Quantify the concentration of the PCR products using the Qubit dsDNA high-sensitivity assay kit according to the manufacturer's instructions.

Notes: A concentration between 0.1-0.5 ng/μL of DNA is expected.

Alternatives: Other DNA quantification instruments can also be used but have not been tested. We recommend using a high-sensitivity apparatus system to determine the DNA concentration for samples containing low levels of DNA (below 10 ng/μL).

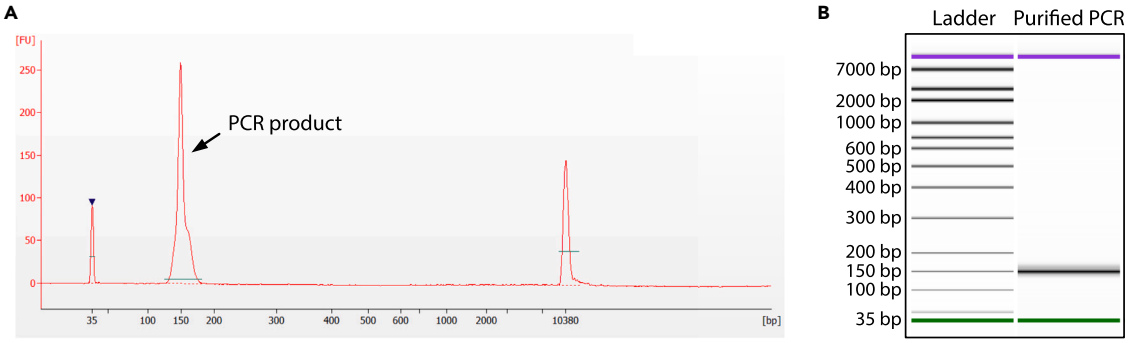


Figure 6. Length of the PCR product monitored by bioanalyzer

(A) Example of an electropherogram generated by Bioanalyzer Agilent 2100 of the purified PCR product.

(B) Electrophoresis representation of the Agilent Bioanalyzer 2100 gel-electropherograms of PCR product.

49. Analyze the DNA length of the PCR products using a Bioanalyzer system (Agilent 2100 Bioanalyzer) with an Agilent High Sensitivity DNA kit for the separation, sizing (35–7000-bp), and quantification of samples with a low concentration of double-stranded DNA according to manufacturer's instructions (Figure 6).

Note: After analyzing the length distribution of the PCR products, the profile should generate a major single peak at the expected length of 145-bp. The peaks at 35 and 10380-bp are reference size markers. The presence of other peaks may indicate the formation of adapter dimers after the PCR and may impair the quality of the sequencing.

Notes: Other apparatuses for analyzing DNA fragments' size and quality can be used but have not been tested.

Troubleshooting, problem 4.

50. Quantify the DNA concentration by qPCR-based quantification using a KAPA Library Quantification Kit according to the manufacturer's instructions.

Note: qPCR-based quantification assays from other companies could be used but have not been tested.

51. Sequence the samples on a high-throughput sequencer. These experiments used a NovaSeq 6000 platform using PE100 cycle chemistry.

Note: We recommend sequencing the sample to obtain 5 to 10 million reads.

Note: Other sequencing platforms could be used but have not been tested.

Data analysis and data visualization

Timing: 1–2 h

The following steps describe the process of analyzing Oligo-seq sequencing data. These steps require the use of a Unix terminal command line system and specific R scripts that were used for this assay are provided below. Additionally, to illustrate the analysis of the sequencing data, we have provided the sequencing file "Oligoseq_read1.fq.gz," a raw fastq file obtained after sequencing the oligonucleotide target used in this protocol. R packages used were installed as

current versions from CRAN using *install.packages()* as of August 2024. The sequencing file and the R scripts can be downloaded at:

https://github.com/ambrociosanchezphd/Oligo-Seq-Pipeline.

52. Download the fastq file(s) obtained from the sequencer.

    Note: If the sample is paired-end, then select the fastq file corresponding to read 1 for the following steps.

    Note: Unzip the fastq file if necessary.

    Note: To illustrate these sections, we will use the following fastq file: "Oligoseq_read1.fq." This file was obtained from an Illumina sequencing run as a fq.gz file

53. Extract the raw sequences from the fastq file using the operating system's terminal command prompt by running the following code:

```
> awk ' NR % 4 == 2 ' Oligoseq_read1.fq > Oligoseq_sequences.fq
```

    Note: This command extracts the sequence reads from the fastq file and creates a new file called "Oligoseq_sequences.fq" output into the directory folder. High-throughput sequencing files are formatted such that starting from the 2nd line, every 4th line of the file corresponds to a full sequence read.

    Note: Depending on the number of reads, the resulting fastq file may take up a few minutes to be generated.

54. Download the R scripts "Oligoseq_trim.R", "Oligoseq_analysis.R", "Oligoseq_logoplot.R", and "Oligoseq_riverplot.R".
55. Open the "Oligoseq_trim.R" script in R and set the working directory to the folder containing the "Oligoseq_sequences.fq" file.
56. Run the "Oligoseq_trim.R" script, and a new fastq file called "Trimmed_Sequences.fq" will be output into the directory folder.

    Note: This file contains all of the sequence reads trimmed by removing the adaptor sequences to select only the sequences corresponding to the target oligonucleotide that will be used in the following analyses.

57. To generate a sequence enrichment/depletion data used for the bar graph (Figures 7A and 7B).
    a. Open the "Oligoseq_analysis.R" script in R and set your working directory to the folder containing the "Trimmed_Sequences.fq" file produced in "Oligoseq_trim.R".
    b. Run "Oligoseq_analysis.R," and a new Excel file called "Oligoseq Enrichment Analysis.xlsx" will be output into the directory folder.

    Note: This file contains 3 separate tabs. "Enrichment-Depletion" displays the enrichment and depletion values for each possible sequence combination in the target oligonucleotide. "All Sequences" contains the occurrence and frequency values of the possible sequence combinations in sequences that have TC and TT motifs. "Deaminated Sequences" contains the occurrence and frequency values of the possible sequence combinations in sequences that have TT motifs only.

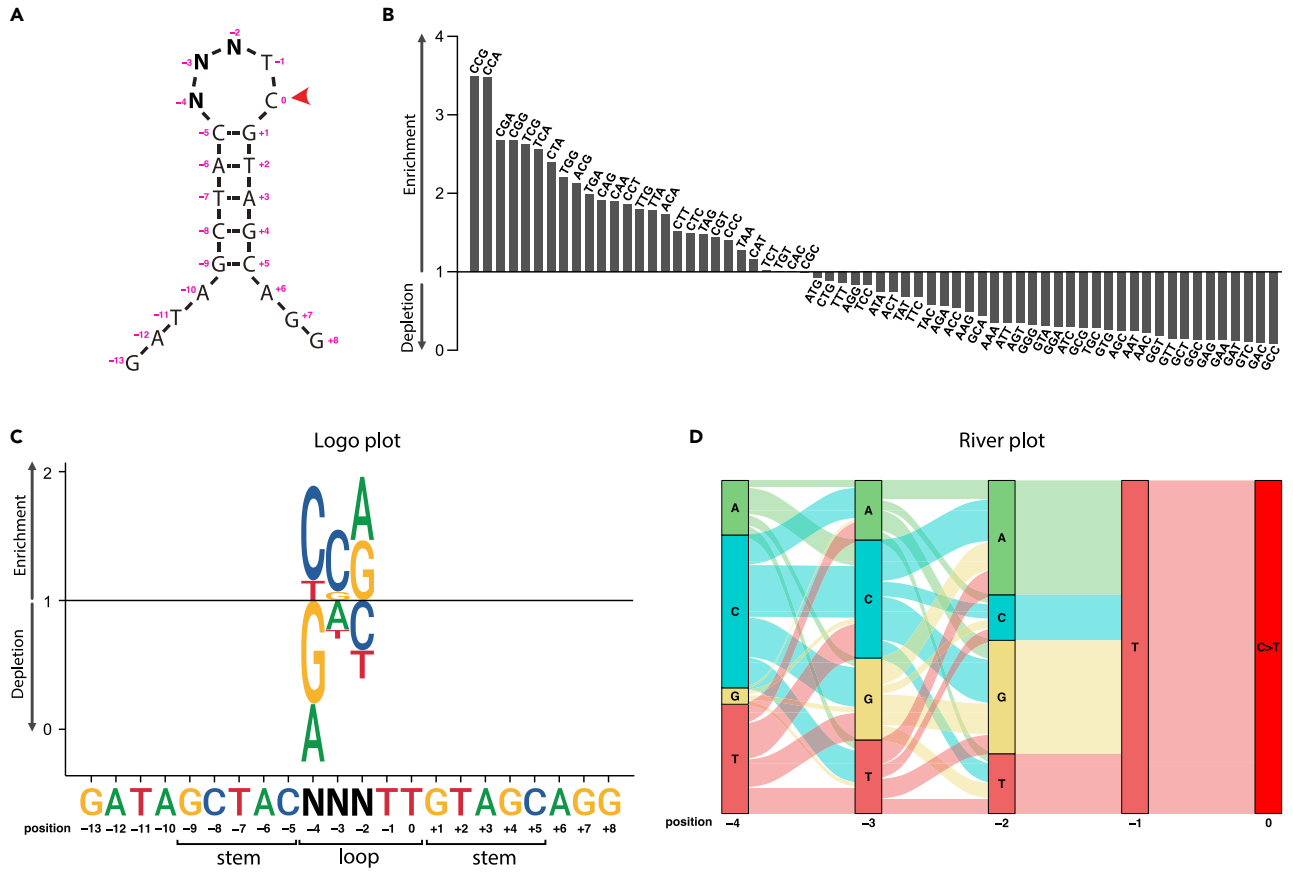


Figure 7. Examples of results obtained after analysis of the Oligo-seq data

(A) Schematic of the 5-nt hairpin loop used to perform the Oligo-seq assay. N refers to any of the four DNA bases. The position of each base relative to the deaminated cytosine is indicated in pink.

(B) Bar graph showing the fold enrichment and depletion for each of the 64 trinucleotide sequences possibly present in the 5-nt hairpin loop shown in **A** after deamination by APOBEC3B.

(C) Example of a sequence logo plot generated in step 58 showing the fold enrichment and depletion for each of the four DNA bases at the −2, −3, and −4 positions after deamination of the cytosine by APOBEC3B.

(D) Example of a river plot generated in step 59 depicting the relative frequency of each nucleotide at the indicated position of the 5-nt hairpin loop shown in **A** and its association with the nucleotides located before or after.

c. The values obtained from this analysis can be visualized by creating a bar plot.

58. To generate a sequence logo plot (Figure 7C),
    a. Open the "Oligoseq_logoplot.R" script in R and set your working directory to the folder containing the "Trimmed_Sequences.fq" file produced in "Oligoseq_trim.R".
    b. Run "Oligoseq_logoplot.R" and a new Excel file called "Oligoseq Logo Plot.xlsx" and a pdf file named "Oligoseq Logo Plot.pdf," which contains a publication-ready logo plot will be output into the directory folder.

Note: The "Oligoseq Logo Plot.xlsx" file contains the values used to generate the logo plot graph.

59. To generate a river plot (Figure 7D),
    a. Open the "Oligoseq_riverplot.R" script in R and set your working directory to the folder containing the "Trimmed_Sequences.fq" file produced in "Oligoseq_trim.R".

b. Run "Oligoseq_riverplot.R" and a pdf file called "Oligoseq River Plot.pdf," which contains a publication-ready river plot will be output into the directory folder.

EXPECTED OUTCOMES

Successful Oligo-seq results should provide between 5 to 10% of deaminated sequences after library sequencing (sequence containing a TT motif instead of a TC motif). After filtering and trimming, we expect that around 60% of the reads obtained after sequencing match the template DNA oligonucleotide sequence (including the sequence with the TC motif and TT motif) after filtering and trimming. Moreover, we expect to detect all the sequence combinations possible within the random sequence from the template oligonucleotides. However, certain sequences should be clearly enriched after APOBEC3B-induced deamination, which is reflected in the target oligonucleotides with the TC motif converted to a TT motif.

QUANTIFICATION AND STATISTICAL ANALYSIS

Each experiment should be repeated three times to ensure the reproducibility of the enrichment for specific sequences deaminated by APOBEC3B for each oligonucleotide template used. The enrichment amplitude levels will vary depending on the percentage of total oligonucleotide template deaminated by APOBEC3B (see before you begin). Therefore, we recommend using the same cell extract when comparing different types of oligonucleotide templates to avoid differences in enrichment amplitude levels caused by differences in APOBEC expression levels in cell extracts.

LIMITATIONS

Although Oligo-seq was developed to study APOBEC deamination activity on DNA substrates, this method can be adapted to study the activity of other DNA modifier enzymes. However, the limitation is that the modified bases by these enzymes need to result in a different base pair after DNA amplification by the DNA polymerase during the library generation. Indeed, a modification of the sequence is essential for the identification and quantification of the preferred sequences targeted by these enzymes. Moreover, the template oligonucleotides used to study APOBEC substrate preference must be sufficiently long to ensure adequate separation of the different DNA products during gel electrophoresis in the library preparation process.

TROUBLESHOOTING

Problem 1 (before you begin: step 20)

Poor DNA band resolution after DNA migration in TEA-Urea denaturing gel electrophoresis.

Potential solution

We recommend casting the gel the same day and at least 1 h before using the gel to allow proper polymerization of the acrylamide. Moreover, we recommend filtering the Urea polyacrylamide solution before using it to remove a small precipitation aggregate of urea in the solution. It is critical to wash the wells of the gel just before loading the sample to remove residual urea that diffuses rapidly in the wells. Finally, we recommend using a large gel casting device such as Protean II xi Cell apparatus (Bio-Rad) or equivalent that can fit large glass plates (16 cm × 20 cm or 20 cm × 20 cm) and with a spacer between the glass plate of 1–1.5 mm. This improves the migration quality compared to small gel devices and thinner gels.

Problem 2 (step-by-step method details: steps 13, 29, and 44)

Presence of undesired DNA size products or smearing after gel migration.

Potential solution

Oligonucleotides used in the reaction may contain byproducts from the synthesis process. We recommend ordering HPLC or PAGE-purified oligonucleotides. In addition, the ligations step

(incubation time and concentration of CircLigase) should be optimized to improve ligation efficiencies and limit unwanted ligation between oligonucleotides.

Problem 3 (step-by-step method details: steps 22, 32, and 47)

Following precipitation, the DNA pellets may be difficult to resuspend.

Potential solution

We recommend increasing the air-dry time to thoroughly remove all isopropanol residues. Moreover, after adding the buffer to resuspend the pellets, raise the temperature to 37°C–45°C with gentle agitation until the DNA pellets are fully dissolved.

Problem 4 (step-by-step method details: step 49)

The bioanalyzer quality control profile shows a peak of DNA concentration of an undesired base pair product.

Potential solution

Adapter dimers formed during the PCR in step 36 may result in the presence of other peaks in the bioanalyzer profile. We recommend decreasing the number of cycles of the PCR to prevent adapter dimers formation and increasing the migration time of the gel to better separate the DNA fragment after the PCR. Alternatively, the purified DNA could be separated again on a polyacrylamide gel, followed by another DNA extraction.

Problem 5 (step-by-step method details: steps 13, 29, and 44)

Low-level DNA or the absence of bands at the appropriate sizes after linker ligation (steps 5 to 13, Figure 3), reverse strand extension (steps 23 to 29, Figure 4), or amplification of the circularized ssDNA (steps 36 to 44, Figure 5).

Potential solution

We recommend verifying the concentration of the DNA oligonucleotide synthesis by companies using Qubit fluorometric quantification or equivalent before stating the different ligation steps. Moreover, we recommend optimizing ligation time and temperature for each ligation step. Specific oligonucleotide sequences might decrease the efficiency of ligation. Finally, we recommend verifying that the template oligonucleotide does not form secondary structures that could block the 5′-, 3′—end of the oligonucleotide or base pair with other oligonucleotides used in the protocol.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rémi Buisson (rbuisson@uci.edu).

Technical contact

For technical questions regarding the execution of this protocol, please contact Rémi Buisson (rbuisson@uci.edu), who will provide assistance.

Materials availability

This study did not generate new unique reagents.

Data and code availability

Sequencing data generated from Oligo-seq experiments in this study are available at the NCBI's SRA under BioProject ID: PRJNA1010353.

ACKNOWLEDGMENTS

We thank Melanie Oakes for technical assistance. Salary support for P.O. was provided by a California Institute for Regenerative Medicine (CIRM) stem cell biology training grant (TG2-01152) and an EMBO Postdoctoral fellowship (ALTF 213-2023). A.S. was supported by the National Institutes of Health Research Supplement to Promote Diversity

in Health-Related Research (R37-CA252081-S2). R.B. was supported by the National Institutes of Health (R37-CA252081, R21-AI185033, and R21-ES036190) and a Research Scholar Grant (RSG-24-1249960-01-DMC) from the American Cancer Society. M.S. was supported by the National Institutes of Health (DP1 DK130640). This work was also made possible, in part, through access to the UCI Genomics Research and Technology (GRT) Hub, which is supported by NIH grants to the Chao Family Comprehensive Cancer Center (P30CA-062203) and to the GRT Hub for instrumentation (1S10OD010794-01 and 1S10OD021718-01). The cartoons in the graphical abstract and Figure 1 were generated with BioRender.com.

AUTHOR CONTRIBUTIONS

P.O. developed the Oligo-seq method. A.S., with the help of P.O., performed the Oligo-seq experiments and the bioinformatic analysis of the sequencing data. M.S. oversaw the bioinformatic analysis. R.B. oversaw the project. P.O., A.S., and R.B. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Sanchez, A., Ortega, P., Sakhtemani, R., Manjunath, L., Oh, S., Bournique, E., Becker, A., Kim, K., Durfee, C., Temiz, N.A., et al. (2024). Mesoscale DNA features impact APOBEC3A and APOBEC3B deaminase activity and shape tumor mutational landscapes. Nat. Commun. *15*, 2370. https://doi.org/10.1038/S41467-024-45909-5.

2. Mertz, T.M., Collins, C.D., Dennis, M., Coxon, M., and Roberts, S.A. (2022). APOBEC-Induced Mutagenesis in Cancer. Annu. Rev. Genet. *56*, 229–252. https://doi.org/10.1146/ANNUREV-GENET-072920-035840.

3. Burns, M.B., Temiz, N.A., and Harris, R.S. (2013). Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat. Genet. *45*, 977–983. https://doi.org/10.1038/ng.2701.

4. Durfee, C., Temiz, N.A., Levin-Klein, R., Argyris, P.P., Alsøe, L., Carracedo, S., Alonso de la Vega, A., Proehl, J., Holzhauer, A.M., Seeman, Z.J., et al. (2023). Human APOBEC3B promotes tumor development in vivo including signature mutations and metastases. Cell Rep. Med. *4*, 101211. https://doi.org/10.1016/J.XCRM.2023.101211.

5. Hoopes, J.I., Cortez, L.M., Mertz, T.M., Malc, E.P., Mieczkowski, P.A., and Roberts, S.A. (2016). APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. Cell Rep. *14*, 1273–1282. https://doi.org/10.1016/j.celrep.2016.01.021.

6. Caswell, D.R., Gui, P., Mayekar, M.K., Law, E.K., Pich, O., Bailey, C., Boumelha, J., Kerr, D.L., Blakely, C.M., Manabe, T., et al. (2024). The role of APOBEC3B in lung tumor evolution and targeted cancer therapy resistance. Nat. Genet. *56*, 60–73. https://doi.org/10.1038/S41588-023-01592-8.

7. Taylor, B.J., Nik-Zainal, S., Wu, Y.L., Stebbings, L.A., Raine, K., Campbell, P.J., Rada, C., Stratton, M.R., and Neuberger, M.S. (2013). DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. Elife *2*, e00534. https://doi.org/10.7554/eLife.00534.

8. Nik-Zainal, S., Wedge, D.C., Alexandrov, L.B., Petljak, M., Butler, A.P., Bolli, N., Davies, H.R., Knappskog, S., Martin, S., Papaemmanuil, E., et al. (2014). Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat. Genet. *46*, 487–491. https://doi.org/10.1038/ng.2955.

9. Swanton, C., McGranahan, N., Starrett, G.J., and Harris, R.S. (2015). APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. Cancer Discov. *5*, 704–712. https://doi.org/10.1158/2159-8290.CD-15-0344.

10. Butt, Y., Sakhtemani, R., Mohamad-Ramshan, R., Lawrence, M.S., and Bhagwat, A.S. (2024). Distinguishing preferences of human APOBEC3A and APOBEC3B for cytosines in hairpin loops, and reflection of these preferences in APOBEC-signature cancer genome mutations. Nat. Commun. *15*, 2369. https://doi.org/10.1038/S41467-024-46231-W.

11. Jalili, P., Bowen, D., Langenbucher, A., Park, S., Aguirre, K., Corcoran, R.B., Fleischman, A.G., Lawrence, M.S., Zou, L., and Buisson, R. (2020). Quantification of ongoing APOBEC3A activity in tumor cells by monitoring RNA editing at hotspots. Nat. Commun. *11*, 2971. https://doi.org/10.1038/s41467-020-16802-8.

12. Cortez, L.M., Brown, A.L., Dennis, M.A., Collins, C.D., Brown, A.J., Mitchell, D., Mertz, T.M., and Roberts, S.A. (2019). APOBEC3A is a prominent cytidine deaminase in breast cancer. PLoS Genet. *15*, e1008545. https://doi.org/10.1371/journal.pgen.1008545.

13. Buisson, R., Langenbucher, A., Bowen, D., Kwan, E.E., Benes, C.H., Zou, L., and Lawrence, M.S. (2019). Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. Science (1979) *364*, eaaw2872. https://doi.org/10.1126/science.aaw2872.

14. Sanchez, A., and Buisson, R. (2024). An in vitro cytidine deaminase assay to monitor APOBEC activity on DNA. Methods Enzymol. https://doi.org/10.1016/BS.MIE.2024.11.037.

15. Durand, M., Chevrie, K., Chassignol, M., Thuong, N.T., and Maurizot, J.C. (1990). Circular dichroism studies of an oligodeoxyribonucleotide containing a hairpin loop made of a hexaethylene glycol chain: conformation and stability. Nucleic Acids Res. *18*, 6353–6359. https://doi.org/10.1093/NAR/18.21.6353.