Supplementary Information (SI) for Chemical Science.
This journal is © The Royal Society of Chemistry 2025

Supplementary Information

Ligase-catalyzed transcription and reverse-transcription of XNA-containing nucleic acid polymers using T3 DNA ligase

Natalie Khamissi, Christopher Korfmann, Areeba Chaudhry, and Ryan Hili*

Department of Chemistry, Centre for Research on Biomolecular Interactions, York University
4700 Keele Street, Toronto, ON M3J 1P3, Canada

Table of Contents

Supporting Methods

Unless otherwise noted, water was purified with Milli-Q purification system. DNA oligonucleotides that are not XNA modified were purchased from Integrated DNA Technologies. DNA oligonucleotides with amine modification were synthesized on an ABI 394 DNA/RNA synthesizer. All materials and reagents used for oligonucleotide synthesis were purchased from Glen Research. All oligonucleotides were synthesized and deprotected according to the manufacturer's protocols. Oligonucleotides were purified by reverse-phase high-pressure liquid chromatography (HPLC, Agilent 1260) using a C18 stationary phase (Eclipse-XDB C18, 5 µm, 9.4 x 200 mm) and an acetonitrile/100 mM triethylammonium acetate gradient. Oligonucleotide concentrations were quantitated by UV spectroscopy using a Nanodrop ND2000 spectrophotometer.

SEQUENCES

LOOPER Sequences

Poly_pr1-PEG
/phos/TGCGACGGCAGGCGAATC/iSp18/AACAACAACAACAA

3P-F
/56-FAM/AGGATCCGAGCTCCACGTG

3P5P(AAA)13
GATTCGCCTGCCGTCGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACACGTGGAGCT
CGGATTCT

3P5P(NNN)13
GATTCGCCTGCCGTCGCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACGTG
GAGCTCGGATTCT

Duplex Sequencing

PR5
/phos/GGATCCGAGCTCCACGTG

PR6
/phos/TGCGACGGCAGGCGAATCT

PRIMER B
AATGATACGGCGACCACCGAG

iTruS_i7_D701
CAAGCAGAAGACGGCATACGAGATATTACTCGGTGACTGGAGTTCAG

iTruS_i7_D702
CAAGCAGAAGACGGCATACGAGATTCCGGAGAGTGACTGGAGTTCAG

iTruS_i7_D703
CAAGCAGAAGACGGCATACGAGATCGCTCATTGTGACTGGAGTTCAG

iTruS_i7_D704
CAAGCAGAAGACGGCATACGAGATGAGATTCCGTGACTGGAGTTCAG

iTruS_i7_D705
CAAGCAGAAGACGGCATACGAGATATTCAGAAGTGACTGGAGTTCAG

AdapterA
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

AdapterB
/phos/ACTGNNNNNNNNNNNNNNAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC3'

TP(NNN)13P
/phos/GATTCGCCTGCCGTCGCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNC
ACGTGGAGCTCGGATCCT

Hairpin Sequencing
5HP_39mer_IT1
/phos/TGCGACGGCAGGCGAATCCAATGATTCGCCTGCCGTCGCANNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNCACGTGGAGCTCGGATCCGCGTGCGGACTCCAGCTA

3P_IT1
AGGCACGGCGAGCGATCTGGATCCGAGCTCCACGTG

3P-F_IT1 **used to confirm hairpin product – not for sequencing*
/56-FAM/AGGCACGGCGAGCGATCTGGATCCGAGCTCCACGTG

3HP_39polymerase
/phos/NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCACGTGGAGCTCGGATCC
CAATGGATCCGAGCTCCACGTG

HPprimerA
/phos/TGCGACGGCAGGCGAATCGCGTGCGGACTCCAGCTA

HPprimerArev
AGGCACGGCGAGCGATCTGATTCGCCTGCCGTCGCA

Fl-HPprimerArev **used to confirm hairpin product – not for sequencing*

/56-FAM/AGGCACGGCGAGCGATCTGATTCGCCTGCCGTCGCA

Hairpin Sequencing PCR Primers

IT1_3PRev_Seq

CCACTACGCCTCCGCTTTCCTCTCTATGGGCAGTCGGTGAT TAGCTGGAGTCCGCACGC

IT1_IC_0130

CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCCAAGCTGC GGTGAT AGGCACGGCGAGCGATCT

IT1_IC_0131

CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCTTACACAC GGTGAT AGGCACGGCGAGCGATCT

IT1_IC_0132

CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TTCTCATTGAAC GGTGAT
AGGCACGGCGAGCGATCT

IT1_IC_0133

CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TCGCATCGTTC GGTGAT AGGCACGGCGAGCGATCT

IT1_IC_0134

CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG TAAGCCATTGTC GGTGAT
AGGCACGGCGAGCGATCT

Reverse Transcription – XNA templates

RTLtemp_unmod

TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

F-RTLtemp Green=2'F modified

TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

LNA-RTLtemp Red=LNA modified

TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

allOMe-RTLtemp Blue=2'OMe modified

TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

midOMe-RTLtemp Blue=2'OMe modified

TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

allFANA-RTLtemp Orange=FANA modified
TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

3'FANA-RTLtemp Orange=FANA modified
TCCAAGCGTGCACGAGCAG CAT GGG GTA GAG TGC TCA ATC CGA TAC TAG CTG GCT AAA
AGGCTCGCGACGGCATAG

RTL_3P
CTATGCCGTCGCGAGCCT

RTL_5P_peg
/phos/CTGCTCGTGCACGCTTGGA/iSp18/AACAACAACAACAA

RTL_AmpEZp1
ACACTCTTTCCCTACACGACGCTCTTCCGATCT
GGAGTCTGATCGATCGTCGAACGGTCGCCTTGACGTGGGCTAGAGCG TCCAAGCGTGCACGAGCAG

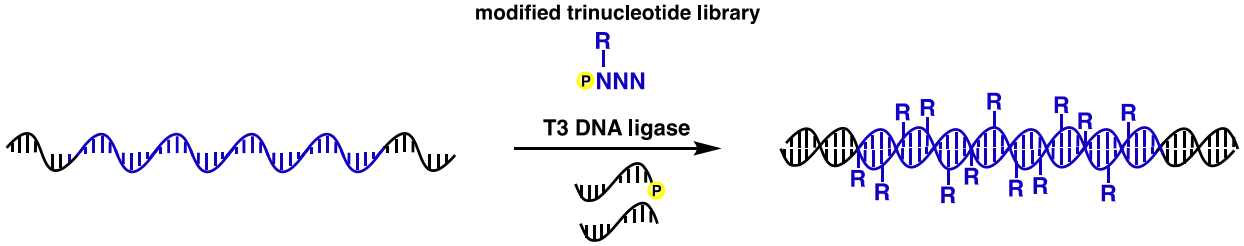
RTL_AmpEZp3
GACTGGAGTTCAGACGTGTGCTCTTCCGATCT GGAGTCTGATCGATCGTCGAACGGTCG
CTATGCCGTCGCGAGCCT

PROTOCOLS

Synthesis of trinucleotide libraries and templates

Trinucleotides for LOOPER were synthesized on a ABI 394 DNA/RNA synthesizer using a DMT-ON protocol on a 1µmol scale (1000 Å CPG column). 2'F-modified phosphoramidites (Glen Research, 10-3430, 10-3420, 10-3415, 10-3400), LNA-modified phosphoramidites (Glen Research, 10-2030, 10-2000, 10-2029, 10-2011), 2'OMe-modified phosphoramidites (Glen Research, 10-3121, 10-3130, 10-3115, 10-3100), FANA-modified phosphoramidites (Glen Research, 10-3800, 10-3820, 10-3830, 10-3815), dA+dC+dG+dT-CE Phosphoramindite (Glen Research 10-1000, 10-1010, 10-1020, 10-1030), Chemical Phosphorylation Reagent II (Glen Research 10-1901) were incorporated as specified by the manufacturer. Following synthesis, the oligonucleotide was cleaved from the resin according to the manufacturer's protocols. The cleaved resin was filtered off, and the oligonucleotide was concentrated under reduced pressure using a speedvac. The residue was then taken up into 100 µL of H_2O and purified using reverse-phase HPLC purification using a [10 % acetonitrile in 0.1 M TEAA, pH 7] to [80 % acetonitrile in 0.1 M TEAA, pH 7] solvent gradient with a column temperature of 45 °C. The purified oligonucleotide was then incubated at room temperature in 1 mL of 40 % aqueous acetic acid for 1 hour to cleave the DMT group, and then frozen and lyophilized. The oligonucleotide was incubated in 500 µl 30 % ammonium hydroxide at room temperature for 15 min to cleave the CPRII linker. Following deprotection, the oligonucleotide was concentrated under reduced pressure using a speedvac. The dried product was dissolved into 100 µL H_2O and subjected to reverse-phase HPLC purification using a [10 % acetonitrile in 0.1M TEAA, pH 7] to [80 % acetonitrile in 0.1M TEAA, pH 7] solvent gradient with a column temperature of 45 °C. The purified oligonucleotide was resuspended in H_2O for subsequent use in LOOPER.

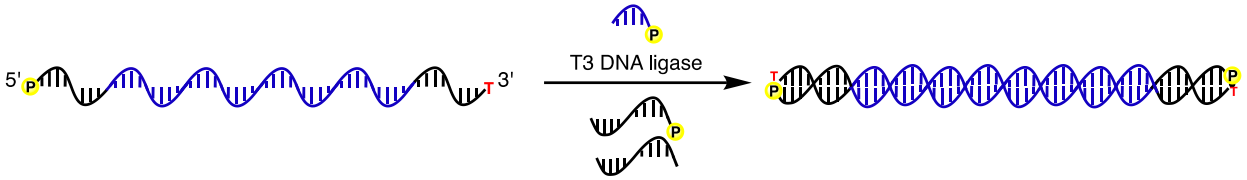
Ligase-catalyzed oligonucleotide polymerizations (LOOPER) and efficiency analysis



The ligation protocol begins with mixing 1 µL of 3P5P(NNN)13 (10 µM), 0.5 µL 3P-F (30 µM), 0.5 µL Polu_pr1-PEG (30 µM), 1.3 µL of the modified trinucleotide library (1 mM), and 5 µL of 2X T3 DNA ligase buffer (NEB M0317L) in a PCR tube. The mixture is then incubated at 95°C 10s, 65°C for 4 min, and cooled using a ramp from 65°C to 4°C at 0.1°C/s. After incubation, 1 µL T3 DNA ligase (3000 U/µL, NEB M0317L) and 0.7 µL of water were added. The reaction was then incubated in a thermal cycler at 4°C for 24 h. The product was then purified using the E.Z.N.A. Omega Biotek® purification kit (Omega Bio- tek, D6492-02) and eluted in 30 µL of water. The samples were denatured at 98°C for 2 min and loaded while still hot onto a 10% denaturing PAGE gel (55 °C, 150V, 40 min). The gels were imaged before and after ethidium bromide staining. The full-length product yield was quantified by gel densitometry against the template band after ethidium bromide staining.

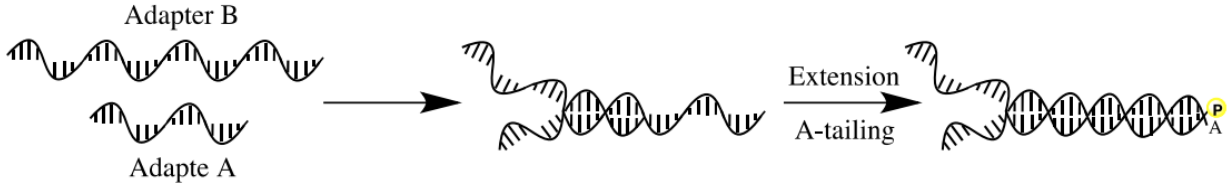
Duplex Sequencing

LOOPER for Duplex Sequencing



The LOOPER protocol begins with mixing 1 µL of TP(NNN)13P (10 µM), 0.5 µL PR5 (30 µM), 0.5 µL PR6 (30 µM), 1.3 µL of the modified trinucleotide library (1 mM), and 5 µL of 2X T3 DNA ligase buffer (NEB M0317L) in a PCR tube. The mixture is then incubated at 95°C 10s, 65°C for 4 min, and cooled using a ramp from 65°C to 4°C at 0.1°C/s. After incubation, 1 µL T3 DNA ligase (3000 U/µL, NEB M0317L) and 0.7 µL of water were added. The reaction was then incubated in a thermal cycler at 4°C for 24 h. The product was then purified using the E.Z.N.A. Omega Biotek® purification kit (Omega Bio- tek, D6492-02) and eluted in 30 µL of water.

Adapter duplex synthesis and A-tailing



In a PCR tube, 15 µL of Adapter A (100 µM in H_2O) and 15 µL of Adapter B (100 µM in H_2O) were mixed. The mixture was incubated for 5 minutes at 95 °C, then cooled to room temperature over 1 hour. In the same PCR tube, 4 µL NEBuffer2 (10X, New England Biolabs, M0212L), 25U Klenow Fragment (3'→5' exo-, New England Biolabs, M0212L), and 1 uL dNTP mix were added. The mixture was incubated at 37 °C for 1 hour and the adapter duplex was purified using the Monarch® PCR and DNA Cleanup Kit (New England Biolabs, T1030S).

In a PCR tube, 30 µL of the purified adapter duplex, 5 µL of NEBuffer2 (10X, New England Biolabs, M0212L), 25U Klenow Fragment (3'→5' exo-, New England Biolabs, M0212L), 5 µL dATP (10 mM, Thermo Scientific), and 5 µL H_2O were added. The mixture was incubated at 37 °C for 1 hour, and the A-tailed adapter duplex was purified using the Monarch® PCR and DNA Cleanup Kit (New England Biolabs, T1030S) and eluted in 30 µL of water.

Adapter ligation

In a PCR tube, 10 pmol of LOOPER product, 200 pmol of the A-tailing adapter duplex, 10 µL NEBNext® Quick Ligation Reaction Buffer, 2.5 µL BSA (2 mg/mL in H_2O), 1000U T4 DNA ligase, and 21 µL of water were mixed and incubated at 16 °C for 16 hours. The ligated products were purified with the E.Z.N.A. Omega Biotek® purification kit (Omega Bio- tek, D6492-02), and native

PAGE (5% native, 150 V, 45 minutes). The band containing the correct product was extracted and incubated overnight at 37 °C in 100 µL 0.3 M NaCl solution. The product was lastly purified using Centri-Sep columns (Princeton Separations, CS-901) and quantified by Qubit.

PCR and sequencing

In a PCR tube, 50 attomoles of the purified ligation product in 20 µL water, 2.5 µL Primer B (10 µM), 2.5 µL of the corresponding iTruS_i7_D7XX primer (10 µM), 25 µL Q5® High-Fidelity Master Mix (2X, New England Biolabs, M0492S) were transferred to a thermocycler where it was incubated for 10s 98 °C, 30s 55 °C, and then 30s 72 °C for the first two cycles, and then 10s 98 °C, 30s 71 °C, and then 72 °C for the rest of the cycles (20-22 cycles). The PCR products were purified with the E.Z.N.A. Omega Biotek® purification kit (Omega Bio- tek, D6492-02), and native PAGE (5% native, 150 V, 40 minutes). The band containing the correct product was extracted and incubated overnight at 37 °C in 100 µL 0.3 M NaCl solution. The products were lastly purified using Centri-Sep columns (Princeton Separations, CS-901) and quantified by Qubit.

Paired-end Illumina® sequencing was performed on an Illumina® MiSeq system using the kit v2 with 300 cycles (150bp PE sequencing) at the TCAG Seuquencing Facility at the Hospital for Sick Children.

Hairpin Sequencing

Formation of the hairpin product

In a PCR tube, 1 µL 5HP_39mer_IT1 (10 µM), 0.5 µL nuclease-free water, 0.5 µL 3P_IT1 (30 µM), 1.3 µL trinucleotide (1 mM), and 5 µL StickTogether DNA Ligase Buffer (2X, NEB M0317L) were added. The mixture was then incubated at 98 °C for 10 s, 65 °C for 4 min, then a ramp from 65 °C to 4 °C at 0.1 °C/s. Lastly, 0.7 µL nuclease-free water and 1 µL T3 DNA ligase (3000 U/µL, NEB M0317L) were added to the mixture. The PCR tube was then incubated at 4 °C for 24 hours. The resulting mixture was then purified with Omega E.Z.N.A.® Cycle Pure Kit (Omega Bio- tek, D6492-02) and eluted in 30 µL nuclease-free water. A 15% denaturing PAGE gel (55 °C, 150 V, 80 min, denatured at 98 °C for 2 min and loaded hot) was then run to visualize and extract the product. The product was then purified using a Centri-Sep column (Princeton Separations, CS-901) and then quantified by Qubit and diluted to 10 pM before PCR.

Hairpin Sequencing – polymerase control

On ice and in a PCR tube, 1 µL 3HP_39polymerase (10 µM), 24 uL nuclease-free water, and 25 µL Q5 High-Fidelity Master Mix (2X, NEB M0492S) were added and transferred to a preheated thermocycler at 98 °C. The mixture was incubated at 98 °C for 10s, 71 °C for 30s, then 71°C for 5 min. The product was then purified with Omega E.Z.N.A.® Cycle Pure Kit (Omega Bio-tek, D6492-02) and eluted in 30 µL nuclease-free water. The sample was then dried and reconstituted in 11.5 µL of nuclease-free water.

In a PCR tube, 2 µL HPprimerA (100 uM in nuclease-free water), 2 µL HPprimerArev (100 µM in nuclease-free water). The primers were denatured at 95 °C for 5 min, then annealed at room temperature for 1 h. Then, 11.5 µL of polymerized sample, 2 µL T4 DNA ligase buffer (10X, NEB

B0202S), and 2.5 µL T4 DNA ligase (400U/µL, NEB M0202S) were added to the primer mixture. The mixture is then incubated at 16 °C for 16h, then 65 °C for 10 min for enzyme deactivation). Lastly the product was purified with Omega E.Z.N.A.® Cycle Pure Kit (Omega Bio-tek, D6492-02) and eluted in 30 µL nuclease-free water.

Hairpin Sequencing – PCR and Ion Torrent sequencing

In a PCR tube and on ice, 10 µL corresponding hairpin product (10 pM), 2.5 µL IT1_3PRev_Seq (10 µM), 2.5 µL corresponding IT1_IC primer (10 µM), 10 µL nuclease-free water, and 25 µL Q5 High-Fidelity Master Mix (2X, NEB M0492S) or NEBNext® Q5U MasterMix (2X, NEB M0597S) (2'F modified products required Q5U polymerase) were added and transferred to a preheated thermocycler at 98 °C. The mixture was then incubated at 98 °C for 2 min, then remain incubating for PCR [98 °C for 10s, 55 °C for 30s, 72 °C for 30s] for the first 2 cycles, then [98 °C for 10s, 70 °C for 30s, 72 °C for 30s] for the rest of the cycles. The product was then purified with Omega E.Z.N.A.® Cycle Pure Kit (Omega Bio-tek, D6492-02) and eluted in 30 µL nuclease-free water before gel extraction (15% native PAGE, 150 V, 100 min). All purified samples were subjected to Ion Torrent sequencing using the 530 chip and kit.

Hairpin Sequencing – script development

The script pipeline was developed to determine the total number of occurrences of each trinucleotide codon found across all template reads. Its reverse complement, being the trinucleotide anticodon, was also determined for each codon occurrence. If any anticodons contained an error, that would be +1 occurrence of trinucleotide anticodon error. The total number of errors across all trinucleotide anticodons is divided by the total occurrence of all codons and multiplied by 100% to achieve the %error. The fidelity is calculated by 100% - %error. The total formula is shown below.

$$\%fidelity = 100\% - \left(\frac{total\ number\ of\ anticodon\ error\ occurrences}{total\ number\ of\ codon\ occurences} \times 100\%\right)$$

Calculating the mononucleotide fidelities were determined using the same method, except we analysed the total number of mononucleotide error and dividing that by three times the total number of codon occurrences. The formula is shown below.

$$\%fidelity = 100\% - \left(\frac{total\ number\ of\ mononucleotide\ error\ occurrences}{(total\ number\ of\ codon\ occurences \times 3)} \times 100\%\right)$$

The analysis pipeline is as follows: sequences are read from a FASTQ file. Reads with lengths outside the specified tolerance (read length ±(N-1), where N is the modified oligonucleotide length) are rejected. For example, if the specified line length is 190 bases and N=3, reads <188 bases and >192 bases are rejected. This is to ensure that the reads represent the accurate full-length product that may or may not have indels. Reads are further filtered to ensure the primer sequences are present verbatim.

For each accepted read, the LOOPER coding region and its corresponding template are extracted and global alignment is performed. Alignment is done using the 'PairwiseAligner' object in Biopython's 'Align' module ('mode="global"', 'scoring="blastn"'). The read is rejected if 'PairwiseAligner' is unable to align the two sequences. Insertions are removed from the LOOPER sequence.

Each sequence is then split into Nmer codons. Comparison is performed between each codon pair. Codon pairs are skipped if the template contains an indel. Unique global LOOPER codon and unique global template codon instance counts are tracked across all accepted and aligned reads. For each template codon, the Nmer error count (i.e. in a codon pair, the template codon does not equal the LOOPER codon) is tracked, as well as the mononucleotide error count (i.e. in a codon pair, the positional mismatch for each base between template codon and LOOPER codon). From these metrics, the LOOPER codon bias is calculated for each codon by dividing the global LOOPER codon instance count by the respective template codon instance count. Template codon fidelity is calculated by dividing the template codon error count by the global template codon instance count and subtracting the result from 1.

Reverse Transcription LOOPER, PCR, and sequencing

The ligation protocol begins with mixing 1 µL of the modified DNA template (10 µM), 0.5 uL RTL_3P, 0.5 uL RTL_5P_peg, 1.3 µL of the unmodified trinucleotide library (1 mM), and 5 µL 2X T3 DNA ligase buffer (New England Biolabs, M0317L) in a PCR tube. The mixture is then incubated at 95°C 10s, 65 °C for 4 min, and cooled using a ramp from 65°C to 4°C at 0.1°C/s. After incubation, 0.7 µL nuclease-free water and 1 µL T3 DNA ligase (NewEngland Biolabs, M0317L) were added. The reaction was then incubated at 4°C for 24 h. The product was then purified using the E.Z.N.A. Omega Biotek® purification kit. The samples were denatured at 98°C for 2 min and loaded while still hot onto a 10% denaturing PAGE gel. The gels were imaged before and after ethidium bromide staining. The full-length product yield was gel extracted (10% denaturing PAGE, 55 °C, 150 V, 40 min).

In a PCR tube and on ice, 10 µL corresponding reverse transcription product (10 pM), 2.5 µL RTL_AmpEZp1 (10 µM), 2.5 µL RTL_AmpEZp3 (10 µM), 10 µL nuclease-free water, and 25 µL Q5 High-Fidelity Master Mix (2X, NEB M0492S) were added. The mixture was then incubated for PCR [98 °C for 10s, 55 °C for 30s, 72 °C for 30s] for the first 2 cycles, then [98 °C for 10s, 70 °C for 30s, 72 °C for 30s] for the rest of the cycles. The product was then purified with Omega E.Z.N.A.® Cycle Pure Kit (Omega Bio-tek, D6492-02) and eluted in 30 µL nuclease-free water. The sample was then concentrated to 25 µL and submitted for paired-end Illumina sequencing.

Supporting Data

FIGURES

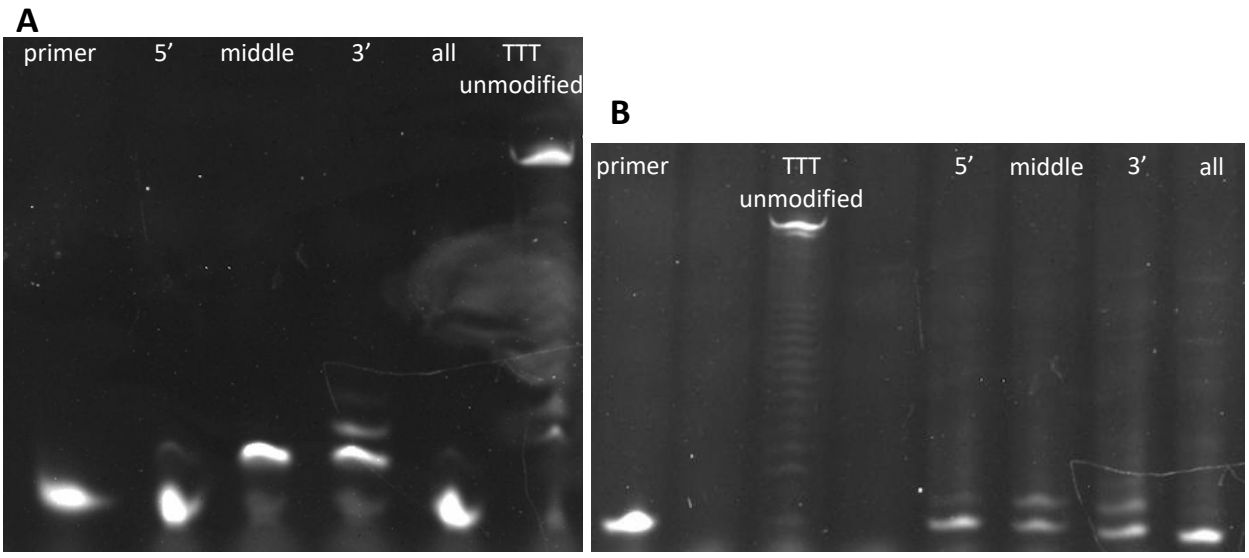


Figure S1. LOOPER yields involving FANA modifications. A) LOOPER products involving TTT trinucleotides with FANA modifications. B) LOOPER products involving NNN trinucleotide libraries with FANA modifications.

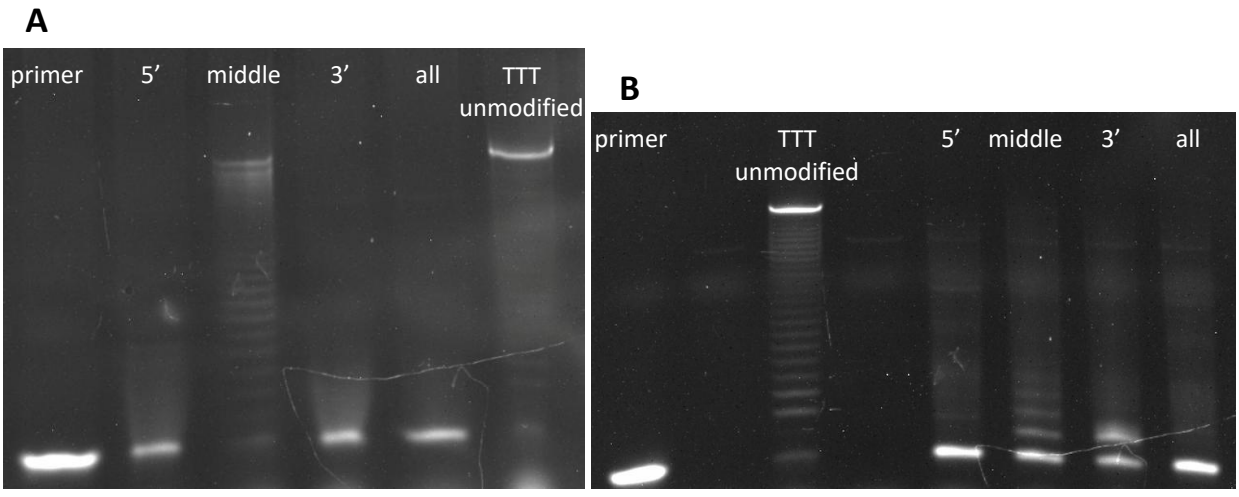


Figure S2. LOOPER yields involving 2’OMe modifications. **A**) LOOPER products involving TTT trinucleotides with 2’OMe modifications. **B**) LOOPER products involving NNN trinucleotide libraries with 2’OMe modifications.

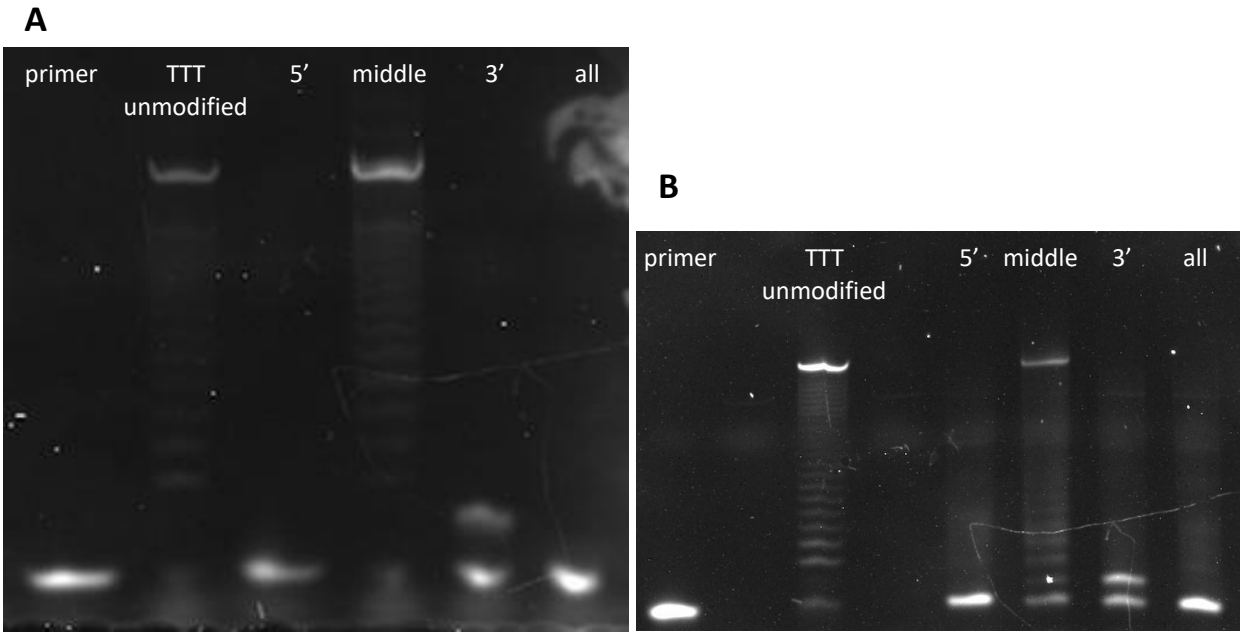


Figure S3. LOOPER yields involving LNA modifications. **A**) LOOPER products involving TTT trinucleotides with LNA modifications. **B**) LOOPER products involving NNN trinucleotide libraries with LNA modifications.

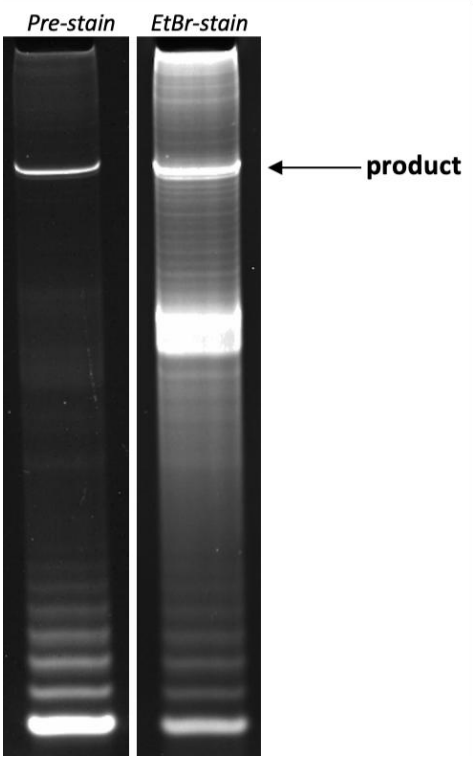


Figure S4. The hairpin product after LOOPER before and after staining. The product band is fluorescently labelled.

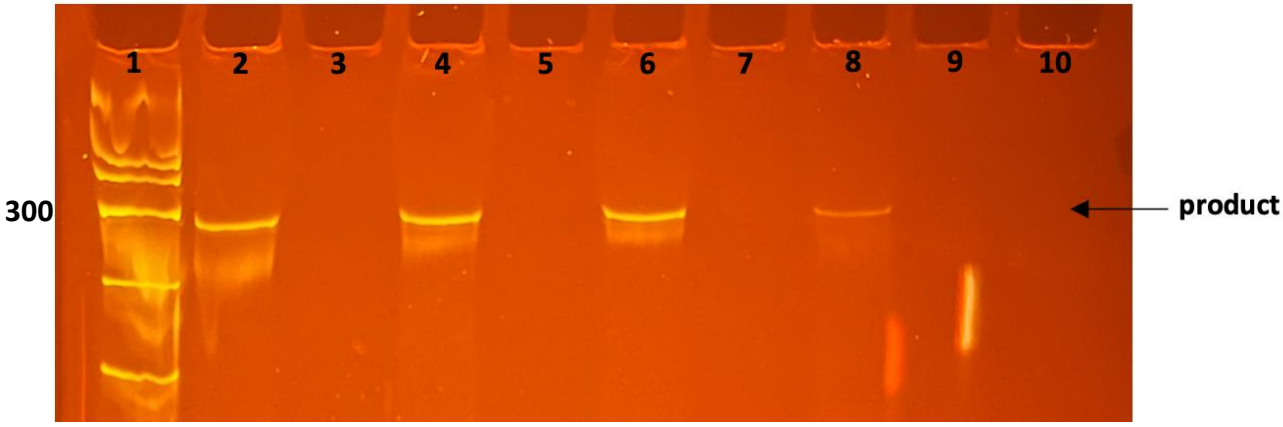


Figure S5. PCR amplified products of the LOOPER hairpins. Lane 1: dsDNA ladder (ThermoFisher, SM1193), Lane 2: NNN unmodified LOOPER – IC_0130, Lane 4: 5'F-NNN LOOPER – IC_0131, Lane 6: midF-NNN LOOPER – IC_0132, Lane 8: 3'F-NNN LOOPER – IC_0133, Lane 10: allF-NNN LOOPER – IC_1034

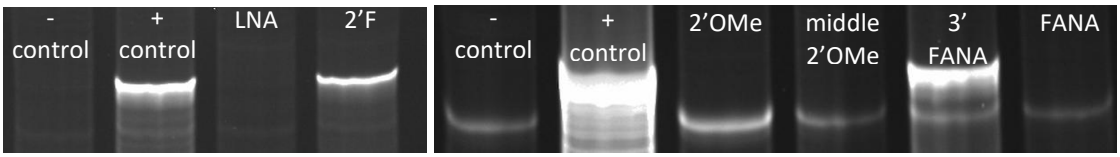


Figure S6. PCR amplification of the XNA-modified templates that were subsequently involved in reverse-transcription LOOPER

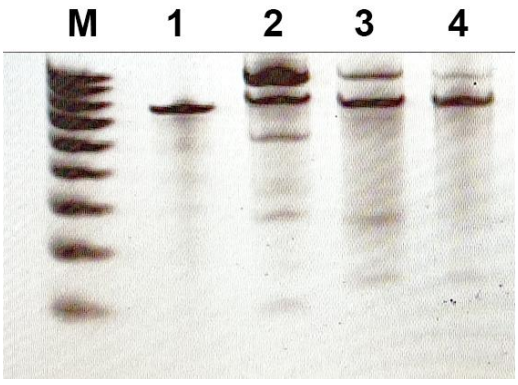


Figure S7. Comparison of T3 DNA ligase vs T4 DNA during LOOPER with trinucleotides modified with XNA at the middle position. M: dsDNA ladder (ThermoFisher, SM1193), Lane 1: template, Lane 2: control with T3 DNA ligase and unmodified NNN, Lane 3: T3 DNA ligase with NNN modified with XNA at the middle position, Lane 4: T3 DNA ligase with NNN modified with XNA at the middle position.

TABLES

Table S1. Duplex sequencing matched pairs and fidelity of each codon within each read from both sequencing runs attempted.

	Round 1			Round 2		
	Matched Reads	**Codon Occurrence**	**Fidelity**	**Matched Reads**	**Codon Occurrence**	**Fidelity**
R8-L1	15	195	97.9%	4	52	92.3%
5'F-NNN	11	143	94.4%	5	65	98.5%
middle F-NNN	40	520	78.1%	25	325	79.7%
3'F-NNN	1	13	92.3%	3	39	97.4%
all F-NNN	1	13	92.3%	1	13	100%

Table S2. Hairpin Sequencing output data for the unmodified control

Observed template	Template occurences	1mer errors	3mer errors	3mer bias	Expected LOOPER
TTT	925	67	64	0.947027	AAA
GTT	1063	178	169	0.907808	AAC
CTT	617	72	63	0.951378	AAG
ATT	788	122	112	0.878173	AAT
TGT	1312	162	156	0.90625	ACA
GGT	919	250	238	0.817193	ACC
CGT	720	119	111	0.938889	ACG
AGT	732	128	121	0.912568	ACT
TCT	763	92	88	0.954128	AGA
GCT	559	84	79	1.121646	AGC
CCT	366	49	44	1.071038	AGG
ACT	465	76	70	0.984946	AGT
TAT	600	53	44	0.98	ATA
GAT	591	79	73	0.939086	ATC
CAT	522	63	63	0.944444	ATG
AAT	465	47	44	0.956989	ATT
TTG	931	74	68	0.941998	CAA
GTG	819	60	56	1.059829	CAC
CTG	529	30	24	1.100189	CAG
ATG	626	100	95	0.875399	CAT
TGG	782	86	76	0.952685	CCA
GGG	490	149	145	0.777551	CCC
CGG	414	47	42	1.016908	CCG
AGG	439	53	46	0.984055	CCT
TCG	537	97	88	0.888268	CGA
GCG	346	68	64	0.956647	CGC
CCG	206	34	29	1.208738	CGG
ACG	311	61	55	0.980707	CGT
TAG	280	22	20	1.196429	CTA
GAG	439	33	29	1.116173	CTC
CAG	374	37	37	1.005348	CTG
AAG	339	39	36	0.967552	CTT
TTC	638	67	61	0.935737	GAA
GTC	521	87	82	0.992322	GAC
CTC	339	51	50	1.0059	GAG

ATC	467	61	59	0.950749	GAT
TGC	611	45	39	1.155483	GCA
GGC	383	46	42	1.592689	GCC
CGC	278	26	25	1.294964	GCG
AGC	307	15	13	1.429967	GCT
TCC	389	53	51	1	GGA
GCC	230	29	28	1.465217	GGC
CCC	227	29	26	1.202643	GGG
ACC	204	40	35	1.161765	GGT
TAC	460	40	39	0.986957	GTA
GAC	305	24	22	1.206557	GTC
CAC	243	7	7	1.572016	GTG
AAC	295	22	18	1.071186	GTT
TTA	597	46	42	0.988275	TAA
GTA	527	93	85	0.979127	TAC
CTA	371	48	42	0.967655	TAG
ATA	356	63	63	0.898876	TAT
TGA	646	64	59	0.955108	TCA
GGA	433	78	72	0.942263	TCC
CGA	278	34	30	1.039568	TCG
AGA	397	35	33	1.010076	TCT
TCA	484	63	62	0.952479	TGA
GCA	341	47	43	1.184751	TGC
CCA	232	35	31	1.133621	TGG
ACA	344	53	52	0.976744	TGT
TAA	410	40	37	1.012195	TTA
GAA	328	41	38	0.984756	TTC
CAA	263	28	25	1.060837	TTG
AAA	222	34	30	0.941441	TTT

Table S3. Hairpin Sequencing output data for 5'-F NNN library

Observed template	Template occurences	1mer errors	3mer errors	3mer bias	Expected LOOPER
TTT	1731	249	233	0.883882	AAA
GTT	1963	429	404	0.817117	AAC
CTT	1251	165	153	0.947242	AAG
ATT	1437	208	200	0.90675	AAT
TGT	2530	372	364	0.877075	ACA
GGT	1695	393	373	0.812389	ACC
CGT	1334	154	143	0.988756	ACG
AGT	1269	164	156	0.919622	ACT
TCT	1630	242	234	0.953374	AGA
GCT	1180	242	233	1.105932	AGC
CCT	858	106	105	1.165501	AGG
ACT	949	167	156	1.139094	AGT
TAT	1062	146	136	0.925612	ATA
GAT	969	121	113	1.048504	ATC
CAT	874	107	106	1.033181	ATG
AAT	942	864	830	0.141189	ATT
TTG	1860	134	130	0.947849	CAA
GTG	1661	174	163	0.951836	CAC
CTG	1069	82	77	1.086997	CAG
ATG	1211	173	168	0.885219	CAT
TGG	1885	216	204	0.922016	CCA
GGG	868	189	186	0.860599	CCC
CGG	922	76	69	1.093275	CCG
AGG	917	102	93	0.955289	CCT
TCG	1142	191	182	0.904553	CGA
GCG	773	110	108	1.028461	CGC
CCG	507	45	43	1.64497	CGG
ACG	709	207	203	0.834979	CGT
TAG	708	49	47	1.035311	CTA
GAG	732	34	32	1.061475	CTC
CAG	747	51	45	1.042838	CTG
AAG	628	47	45	0.977707	CTT
TTC	1387	147	141	0.981255	GAA
GTC	1151	186	178	0.961772	GAC
CTC	712	93	91	1.16573	GAG

ATC	881	163	157	0.905789	GAT
TGC	1264	137	127	1.110759	GCA
GGC	737	103	100	1.313433	GCC
CGC	556	32	31	1.370504	GCG
AGC	600	76	72	1.093333	GCT
TCC	939	173	166	1.096912	GGA
GCC	618	77	74	1.605178	GGC
CCC	423	33	31	1.815603	GGG
ACC	552	156	148	1.03442	GGT
TAC	903	76	75	1.024363	GTA
GAC	590	42	38	1.169492	GTC
CAC	480	27	26	1.420833	GTG
AAC	469	50	46	1.857143	GTT
TTA	1142	126	118	0.950963	TAA
GTA	1046	239	230	0.844168	TAC
CTA	670	70	69	1.010448	TAG
ATA	725	108	105	0.910345	TAT
TGA	1346	147	136	0.946508	TCA
GGA	857	72	67	0.97783	TCC
CGA	552	43	40	1.11413	TCG
AGA	665	47	40	1.007519	TCT
TCA	957	135	129	0.984326	TGA
GCA	589	72	68	1.307301	TGC
CCA	463	41	37	1.336933	TGG
ACA	580	63	61	1.086207	TGT
TAA	677	61	60	0.986706	TTA
GAA	542	55	52	0.983395	TTC
CAA	399	34	31	1.050125	TTG
AAA	363	25	22	1.008264	TTT

Table S4. Hairpin Sequencing output data for mid-F NNN library

Observed template	Template occurences	1mer errors	3mer errors	3mer bias	Expected LOOPER
TTT	26504	24468	23402	0.124321	AAA
GTT	28544	8073	7549	0.791935	AAC
CTT	18708	3611	3465	1.106372	AAG
ATT	16672	15842	14994	0.105086	AAT
TGT	34250	10050	9775	0.748847	ACA
GGT	30758	10717	10495	0.701931	ACC
CGT	23724	5887	5794	0.863261	ACG
AGT	18418	4873	4707	0.786622	ACT
TCT	20784	5049	4864	1.200491	AGA
GCT	20391	6619	6417	0.921975	AGC
CCT	13285	2833	2772	1.117125	AGG
ACT	13210	3515	3290	1.258668	AGT
TAT	13144	8359	7902	0.407638	ATA
GAT	17208	4462	4235	0.798291	ATC
CAT	14219	2193	2115	1.047683	ATG
AAT	9104	6147	5738	0.3817	ATT
TTG	28856	2600	2452	0.974459	CAA
GTG	28695	3895	3664	0.903224	CAC
CTG	18182	1151	1078	1.116269	CAG
ATG	19065	3288	3139	0.878259	CAT
TGG	26776	3553	3418	0.906894	CCA
GGG	17684	2922	2879	0.92513	CCC
CGG	17651	1606	1525	1.070534	CCG
AGG	16514	1859	1744	0.934601	CCT
TCG	20414	2360	2279	0.978299	CGA
GCG	16871	2667	2581	1.028155	CGC
CCG	11094	895	856	1.334325	CGG
ACG	11991	2191	2137	0.973063	CGT
TAG	9003	1054	977	0.9679	CTA
GAG	13748	1050	937	0.983343	CTC
CAG	12852	411	359	1.079365	CTG
AAG	9516	1179	1098	0.916982	CTT
TTC	20232	2058	1952	1.754893	GAA
GTC	18772	3174	3044	1.163115	GAC
CTC	12183	972	941	1.343347	GAG

ATC	12855	2102	1992	1.42746	GAT
TGC	21191	2048	1968	1.434288	GCA
GGC	15742	1342	1269	1.785097	GCC
CGC	12298	388	362	1.760774	GCG
AGC	10858	748	694	1.518143	GCT
TCC	13614	1222	1179	1.466579	GGA
GCC	11761	1115	1057	1.885724	GGC
CCC	7614	201	199	1.988442	GGG
ACC	8512	1297	1247	1.382989	GGT
TAC	12839	1385	1318	1.327829	GTA
GAC	11222	791	707	1.414543	GTC
CAC	8921	279	251	1.379442	GTG
AAC	7248	676	621	1.501932	GTT
TTA	16016	15337	14575	0.09528	TAA
GTA	16625	5204	4847	0.783459	TAC
CTA	11337	1834	1769	1.277939	TAG
ATA	10019	9915	9365	0.075057	TAT
TGA	20131	4450	4315	0.833838	TCA
GGA	17036	4881	4793	0.767434	TCC
CGA	11473	1663	1616	1.025713	TCG
AGA	11267	2546	2502	0.825686	TCT
TCA	15046	2320	2263	1.268643	TGA
GCA	12455	3119	3024	1.005379	TGC
CCA	9049	1049	1003	1.307437	TGG
ACA	9206	1925	1871	1.275038	TGT
TAA	8819	2718	2564	0.741354	TTA
GAA	10359	2589	2466	0.818515	TTC
CAA	8615	731	686	1.085781	TTG
AAA	5176	1270	1201	0.827859	TTT

Table S5. Hairpin Sequencing output data for 3'-F NNN library

Observed template	Template occurences	1mer errors	3mer errors	3mer bias	Expected LOOPER
TTT	5215	750	696	0.8907	AAA
GTT	5823	1094	1019	0.851623	AAC
CTT	3799	509	467	0.947355	AAG
ATT	3681	528	488	0.888346	AAT
TGT	9679	2130	2063	0.803905	ACA
GGT	6855	1894	1835	0.754778	ACC
CGT	5307	754	729	1.028453	ACG
AGT	4161	509	491	0.9118	ACT
TCT	5275	770	734	0.995071	AGA
GCT	4049	958	940	0.996542	AGC
CCT	2754	399	385	1.102033	AGG
ACT	2902	437	408	0.954859	AGT
TAT	4633	4437	4259	0.087632	ATA
GAT	3838	619	567	0.895518	ATC
CAT	3250	315	292	1.600923	ATG
AAT	2347	274	250	0.927567	ATT
TTG	5568	536	497	0.925108	CAA
GTG	5036	538	500	0.936259	CAC
CTG	3302	232	211	1.13507	CAG
ATG	3137	593	563	0.842843	CAT
TGG	5966	813	767	0.893899	CCA
GGG	3032	537	514	0.91095	CCC
CGG	3289	255	233	1.134387	CCG
AGG	2950	429	393	0.898644	CCT
TCG	3672	426	402	0.962963	CGA
GCG	2586	264	248	1.104022	CGC
CCG	1671	80	73	1.43447	CGG
ACG	1947	400	389	0.894196	CGT
TAG	2081	267	244	0.983662	CTA
GAG	2460	193	177	0.993902	CTC
CAG	2523	138	124	1.086405	CTG
AAG	1820	229	214	0.911538	CTT
TTC	3878	516	476	0.946622	GAA
GTC	3421	581	560	0.978077	GAC
CTC	2069	241	234	1.147414	GAG

ATC	2186	371	353	0.910796	GAT
TGC	4529	495	452	1.273129	GCA
GGC	2585	246	214	1.728046	GCC
CGC	2072	75	69	1.626448	GCG
AGC	1881	174	159	1.222222	GCT
TCC	2620	296	278	1.264504	GGA
GCC	1861	134	118	1.895755	GGC
CCC	1222	39	38	1.900164	GGG
ACC	1285	225	213	1.189105	GGT
TAC	3111	343	330	1.469945	GTA
GAC	2100	202	180	1.167619	GTC
CAC	1840	69	62	1.434239	GTG
AAC	1503	162	155	1.02994	GTT
TTA	3180	455	417	0.892453	TAA
GTA	2776	610	566	0.842579	TAC
CTA	2120	302	277	0.981604	TAG
ATA	1713	385	356	0.813193	TAT
TGA	4806	731	693	0.882855	TCA
GGA	3246	508	484	0.895564	TCC
CGA	1988	104	93	1.226358	TCG
AGA	2219	223	209	0.936458	TCT
TCA	3328	367	351	1.009916	TGA
GCA	2184	274	252	1.139652	TGC
CCA	1614	95	90	1.346344	TGG
ACA	1837	257	248	0.961894	TGT
TAA	2092	1888	1859	0.123805	TTA
GAA	1934	228	211	0.975181	TTC
CAA	1667	96	88	1.754649	TTG
AAA	946	101	100	0.959831	TTT

Table S6. Hairpin Sequencing output data for the Q5 DNA polymerase control

Expected LOOPER	Observed template	Template occurrences	3mer errors
AAA	TTT	86	1
AAC	GTT	64	1
AAG	CTT	96	6
AAT	ATT	155	7
ACA	TGT	73	0
ACC	GGT	17	2
ACG	CGT	46	7
ACT	AGT	84	2
AGA	TCT	107	5
AGC	GCT	31	0
AGG	CCT	76	12
AGT	ACT	114	3
ATA	TAT	118	7
ATC	GAT	62	2
ATG	CAT	138	14
ATT	AAT	187	5
CAA	TTG	64	3
CAC	GTG	20	2
CAG	CTG	61	8
CAT	ATG	71	3
CCA	TGG	29	3
CCC	GGG	7	1
CCG	CGG	25	2
CCT	AGG	25	2
CGA	TCG	41	13
CGC	GCG	24	4
CGG	CCG	28	6
CGT	ACG	47	2
CTA	TAG	56	10
CTC	GAG	27	1
CTG	CAG	62	6
CTT	AAG	59	2
GAA	TTC	115	18
GAC	GTC	46	7
GAG	CTC	91	22
GAT	ATC	130	22

GCA	TGC	71	2
GCC	GGC	9	1
GCG	CGC	36	12
GCT	AGC	42	2
GGA	TCC	100	13
GGC	GCC	35	9
GGG	CCC	50	11
GGT	ACC	115	17
GTA	TAC	130	1
GTC	GAC	51	2
GTG	CAC	116	16
GTT	AAC	160	6
TAA	TTA	124	7
TAC	GTA	63	1
TAG	CTA	122	7
TAT	ATA	146	6
TCA	TGA	76	2
TCC	GGA	25	1
TCG	CGA	52	3
TCT	AGA	71	2
TGA	TCA	149	22
TGC	GCA	36	4
TGG	CCA	102	29
TGT	ACA	170	1
TTA	TAA	169	8
TTC	GAA	79	4
TTG	CAA	147	18
TTT	AAA	189	5

UNPROCESSED GEL IMAGES

Figure 3 (Lanes 1-6)
Lane 1: 20/100 ssDNA ladder (IDT), Lane 2: fluorescent primer, Lane 3: N39 template, Lane 4: NNN control (previously validated in lab), Lane 5: N39 template, Lane 6: 5'modified 2'-F NNN, Lane 7: middle-modified 2'-F NNN, Lane 8: 3'modified 2'F NNN, Lane 9: all-modified 2'-F NNN, Lane 10: middle-modified LNA NNN.

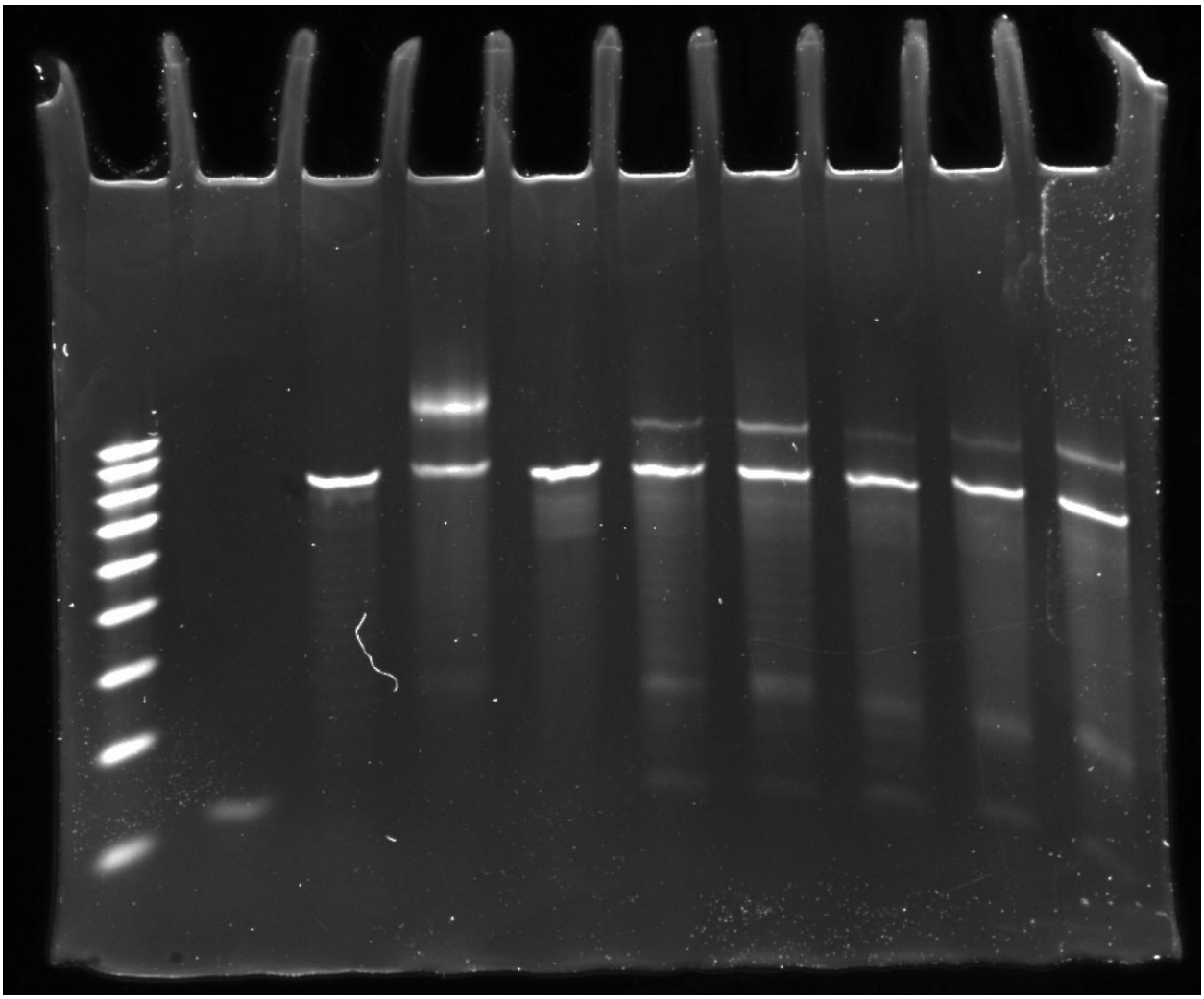


Figure 6a (Lanes 1-5)
Lane 2: 20/100 ssDNA ladder (IDT), Lane 3: no load, Lane 4: unmodified control RT LOOPER, Lane 5: all-modified 2'-OMe RT LOOPER, Lane 6: middle-modified 2'-OMe RT LOOPER, Lane 7: 3'-modified FANA RT LOOPER, Lane 8: all-modified FANA RT LOOPER, Lane 9: all-modified LNA RT LOOPER.

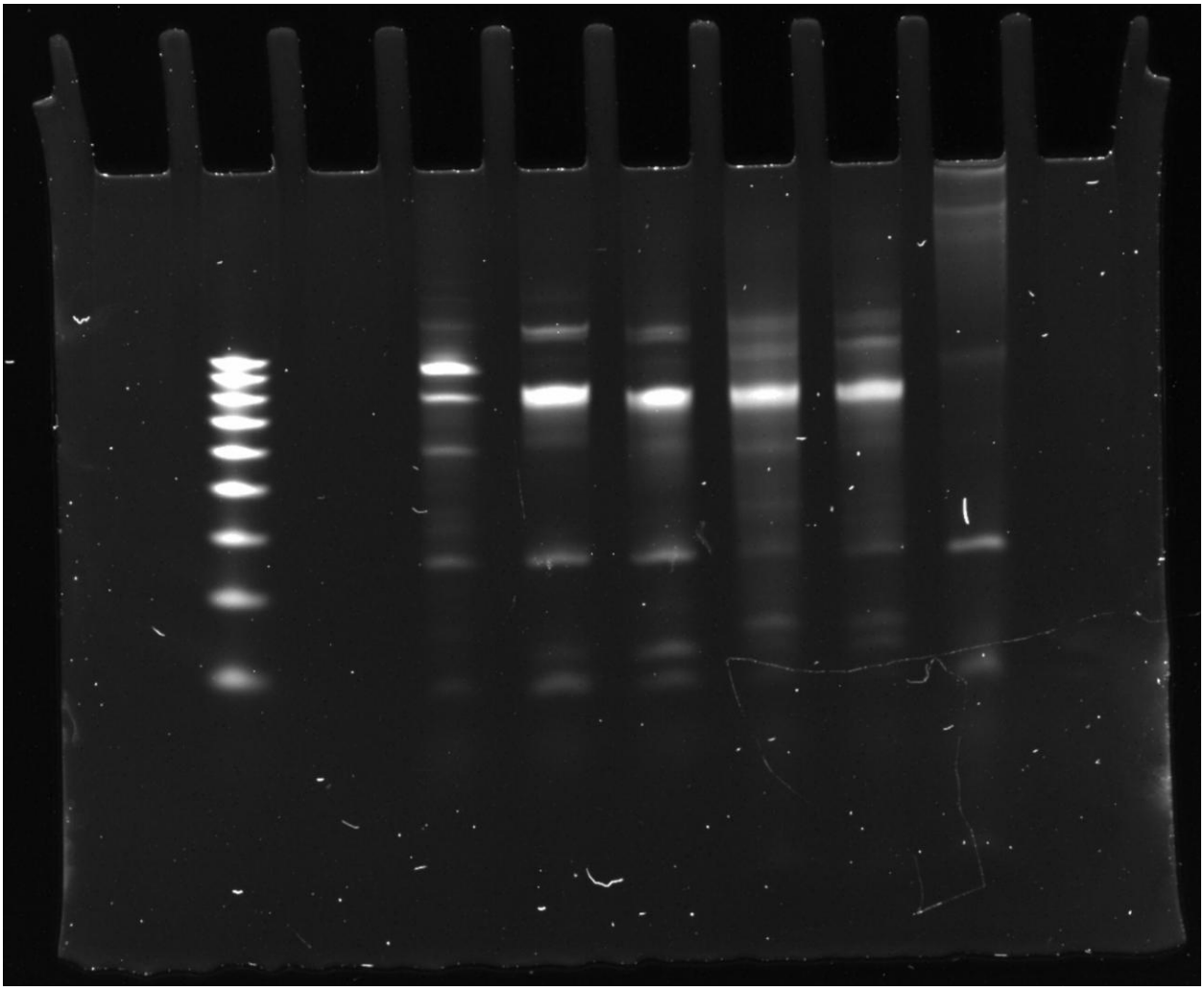


Figure 6a (Lanes 6-7)
Lane1: 20/100 ssDNA ladder (IDT), Lane 2: 3' primer, Lane 3: mid-LNA template, Lane 4: mid-LNA RT product, Lane 5: 2'-F template, Lane 6: 2'-F RT product.