# A one-enzyme RT-qPCR assay for SARS-CoV-2, and procedures for reagent production

## Sanchita Bhadra, Andre C. Maranhao, and Andrew D. Ellington<sup>1</sup>

Center for Systems and Synthetic Biology University of Texas at Austin Austin, TX 78703

<sup>1</sup> Inquiries about manuscript to <u>ellingtonlab@gmail.com</u> <sup>2</sup> Inquiries about vectors to <u>https://reclone.org/</u>

## ABSTRACT

Given the scale of the ongoing COVID-19 pandemic, the need for reliable, scalable testing, and the likelihood of reagent shortages, especially in resource-poor settings, we have developed a RT-qPCR assay that relies on an alternative to conventional viral reverse transcriptases, a thermostable reverse transcriptase / DNA polymerase (RTX)<sup>1</sup>. Here we show that RTX performs comparably to the other assays sanctioned by the CDC and validated in kit format. We demonstrate two modes of RTX use – (i) dye-based RT-qPCR assays that require only RTX polymerase, and (ii) TaqMan RT-qPCR assays that use a combination of RTX and Taq DNA polymerases (as the RTX exonuclease does not degrade a TaqMan probe). We also provide straightforward recipes for the purification of this alternative reagent. We anticipate that in low resource or point-of-need settings researchers could obtain the available constructs from Addgene or our lab and begin to develop their own assays, within whatever regulatory framework exists for them.

We lay out protocols for dye-based and TaqMan probe-based assays, in order to best compare with 'gold standard' reagents. These protocols should form the basis of further modifications that can simplify the assay to the use of overexpressing cells themselves as reagents.

## Developing dye-based and TaqMan probe-based RT-qPCR assays with RTX

### MATERIALS

- 10X RTX buffer: 600 mM Tris-HCl, 250 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM KCl, 20 mM MgSO<sub>4</sub>, pH 8.4 @25°C
- 10X ThermoPol buffer (New England Biolabs): 200 mM Tris-HCl, 100 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 100 mM KCl, 20 mM MgSO<sub>4</sub>, 1% Triton<sup>®</sup> X-100, pH 8.8 @25°C
- Forward PCR primer<sup>a</sup>
- Reverse PCR primer<sup>a</sup>
- TaqMan probe<sup>a</sup>

- 20X EvaGreen solution
- 4 mM deoxyribonucleotides (dNTP)
- 5M Betaine
- 20 U/µL SUPERase.In RNase inhibitor
- 5 U/µL Taq DNA polymerase
- 0.2 mg/ml RTX polymerase<sup>b</sup>
- Nuclease-free water
- qPCR tubes or plates<sup>c</sup>
- Real-time PCR machine<sup>c</sup>
- Cold-block or ice
- Barrier tips

<sup>a</sup>Individual primer and probe stocks or pre-made primer-probe mixes for CDC N1, N2 and N3 assays (**Table 1**) available from Integrated DNA Technologies may be used.

<sup>b</sup>RTX polymerase with proofreading capability (RTX), and RTX polymerase without proofreading capability (RTX Exo-) have been compared.

<sup>c</sup>LightCycler96 real-time PCR machine and 96-well plates with optical plastic film cover designed for use with LightCycler platform were used for these experiments.

Name	Description	Sequence <sup>b</sup>
2019-nCoV_N1-F	2019-nCoV_N1 Forward Primer	5'-GAC CCC AAA ATC AGC GAA AT-3'
2019-nCoV_N1-R	2019-nCoV_N1 Reverse Primer	5'-TCT GGT TAC TGC CAG TTG AAT CTG-3'
2019-nCoV_N1-P	2019-nCoV_N1 Probe	5'-FAM-ACC CCG CAT TAC GTT TGG TGG ACC-BHQ1-3'
2019-nCoV_N2-F	2019-nCoV_N2 Forward Primer	5'-TTA CAA ACA TTG GCC GCA AA-3'
2019-nCoV_N2-R	2019-nCoV_N2 Reverse Primer	5'-GCG CGA CAT TCC GAA GAA-3'
2019-nCoV_N2-P	2019-nCoV_N2 Probe	5'-FAM-ACA ATT TGC CCC CAG CGC TTC AG-BHQ1-3'
2019-nCoV_N3-F	2019-nCoV_N3 Forward Primer	5'-GGG AGC CTT GAA TAC ACC AAA A-3'
2019-nCoV_N3-R	2019-nCoV_N3 Reverse Primer	5'-TGT AGC ACG ATT GCA GCA TTG-3'
2019-nCoV_N3-P	2019-nCoV_N3 Probe	5'-FAM-AYC ACA TTG GCA CCC GCA ATC CTG-BHQ1-3'

Table 1. CDC TaqMan qRT-PCR primers and probes for SARS-CoV-2 <sup>a</sup> (adapted from
https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html

<sup>a</sup>According to CDC, oligonucleotide sequences are subject to future changes as the 2019-Novel Coronavirus evolves. Refer to CDC website for latest updates. <sup>b</sup>EAM: 6-carboxyfluorescein: BHO-1: Black Hole Quencher 1

<sup>b</sup>FAM: 6-carboxyfluorescein; BHQ-1: Black Hole Quencher 1

### RESULTS

Dye-based RT-qPCR reaction set up using only RTX DNA polymerase

EvaGreen dye-containing RT-qPCR reaction set up shown in **Table 2** was used for assembling CDC SARS-CoV-2 N1, N2, and N3 assays. The three assays target three different regions within the nucleocapsid phosphoprotein (N) gene near the 3'-end of the ~30 kb viral genome. In particular, N1 primers (**Table 1**) amplify a 72 nt long region from position 28287 to 28358 of the genome (GenBank Accession No. MN985325.1). The N2 primers produce a 67 bp amplicon from the genomic position 29164 to 29230. Meanwhile, the N3 primers amplify a 72 nt long region from genomic position 28681 to 28752. All reaction mixes were assembled on a cold-block and directly transferred to a real-time PCR machine programmed to cycle through the steps depicted in **Table 3**. Melt curve analysis was included after PCR amplification in order to distinguish target amplicons from spurious products.

	Volume	Final
	(μL)	concentration
RNA sample	5	
Forward PCR primer	1.5	500 nM
Reverse PCR primer	1.5	500 nM
20X EvaGreen	1.25	1X
10X RTX buffer	2.5	1X
4 mM dNTP	2.5	400 µM
20 U/µL SUPERase.In RNase inhibitor	1	0.8 U/μL
RTX or RTX Exo- 0.2mg/ml	0.5	4 ng/μL
5M betaine	2.5	0.5M
Water	6.75	
Total reaction volume	25	

Table 2. EvaGreen RT-qPCR reaction mix using only RTX DNA polymerase

Table 3. RTX polymerase-based EvaGreen RT-qPCR cycling conditions

Step	Temperature Duration		Temperature	Signal
Step	remperature	Bulation	acquisition	
1 <sup>a</sup>	60 °C	30 min	None	
<b>2</b> <sup>a</sup>	95 °C	10 min	None	
	55 cycles of	Steps 3 and	4	
3	95 °C	15 sec	None	
4 <sup>b</sup>	55 °C	30 sec	Single	
Me	elt curve analysi	s in Steps 5,	. 6, and 7	
5	95 °C	10 sec	None	
6	65 °C	60 sec	None	
7	97 °C	1 sec	Continuous <sup>c</sup>	

<sup>a</sup>Assay time may be reduced by shortening reverse transcription step (Step 1) to 15 min and the following denaturation step (Step 2) to 2 min.

<sup>b</sup>Signal acquisition occurs in the FAM (fluorescein) channel during step 4.

<sup>c</sup>20 measurements/<sup>°</sup>C temperature change



Representative results of SARS-CoV-2 RT-qPCR tests performed using synthetic RNA templates and only RTX or RTX Exo- DNA polymerases are depicted in **Figures 1 and 2**.

**Figure 1.** CDC SARS-CoV-2 N1, N2, and N3 RT-qPCR assays performed using indicated copies of synthetic RNA and RTX DNA polymerase. Panels A, B, and C depict N1, N2, and N3 assays measured in real-time using EvaGreen dye. Amplification curves from reactions containing 100,000 (black traces), 10,000 (blue traces), 1,000 (red traces), and 100 (pink traces) copies of SARS-CoV-2 synthetic N RNA are depicted. Negative control reactions either contained no templates (gray traces) or contained 100,000 copies of synthetic N RNA from MERS-CoV (green traces). Panels D, E, and F depict melting peaks of amplicons determined using the 'Tm calling' analysis in the LightCycler96 software. Average Cq values of all assays are tabulated.



**Figure 2.** CDC SARS-CoV-2 N1, N2, and N3 RT-qPCR assays performed using indicated copies of synthetic RNA and RTX Exo- DNA polymerase. Panels A, B, and C depict N1, N2, and N3 assays measured in real-time using EvaGreen dye. Amplification curves from reactions containing 100,000 (black traces), 10,000 (blue traces), 1,000 (red traces), and 100 (pink traces) copies of SARS-CoV-2 synthetic N RNA are depicted. Negative control reactions either contained no templates (gray traces) or contained 100,000 copies of synthetic N RNA from MERS-CoV (green traces). Panels D, E, and F depict melting peaks of amplicons determined using the 'Tm calling' analysis in the LightCycler96 software. Average Cq values of all assays are tabulated.

These results demonstrate that RTX DNA polymerase alone, with or without proofreading capability, can support dye-based RT-qPCR analyses. In our hands the full-length RTX DNA polymerase was slightly better than the Exo- version, especially in the N3 assay.

## RTX polymerase-based TaqMan qRT-PCR reaction set up

In order to perform TaqMan RT-qPCR we developed a one-pot two enzyme system comprising RTX and Taq DNA polymerases. TaqMan qRT-PCR reaction set up shown in **Table 4** was used for assembling CDC SARS-CoV-2 N1, N2, and N3 assays using either RTX buffer or ThermoPol buffer. Control reactions lacking RTX polymerase were set up using the same recipe except RTX was

replaced with water. All reaction mixes were assembled on a cold-block and directly transferred to a real-time PCR machine programmed to cycle through the steps depicted in **Table 5**.

	Volume	Final
	(μL)	concentration
RNA sample	5	
Forward PCR primer	1.5	402 nM
Reverse PCR primer	1.5	402 nM
TaqMan probe	1.5	102 nM
10X Assay buffer <sup>a</sup>	2.5	1X
4 mM dNTP	2.5	400 µM
20 U/µL SUPERase.In RNase inhibitor	1	0.8 U/μL
5 U/ μL Taq DNA polymerase	0.5	0.1 U/μL
RTX or RTX Exo- 0.2mg/ml or water	0.5	4 ng/μL
5M betaine	2.5	0.5M
Water	6	
Total reaction volume	25	

Table 4. RTX polymerase-based TaqMan qRT-PCR reaction mix

<sup>a</sup> Assay buffer: Duplicate TaqMan RT-qPCR reactions were performed either in 1X ThermoPol buffer or in 1X RTX buffer as indicated.

Step 1 <sup>a</sup>	60 °C	30 min
Step 2 <sup>a</sup>	95 °C	10 min
55 cycles of Steps 3 and 4		
Step 3	95 °C	15 sec
Step 4 <sup>b</sup>	55 °C	30 sec

**Table 5.** RTX polymerase-based TaqMan qRT-PCR cycling conditions

<sup>a</sup>Assay time may be reduced by shortening reverse transcription step (Step 1) to 15 min and the following denaturation step (Step 2) to 2 min.

<sup>b</sup>Signal acquisition occurs in the FAM (fluorescein) channel during step 4.

Representative results of RTX polymerase-based SARS-CoV-2 TaqMan qRT-PCR tests performed using synthetic RNA templates are depicted in **Figures 3 and 4**. Both RTX and RTX Exo- when combined with Taq DNA polymerase were able to support one-pot TaqMan RT-qPCR assays for all three CDC SARS-CoV-2 assays. Under both buffer conditions tested, 1X ThermoPol versus 1X RTX buffer, RTX polymerase yielded more consistent amplification curves for all three CDC assays compared to RTX Exo- polymerase.



**Figure 3.** CDC SARS-CoV-2 N1, N2, and N3 TaqMan RT-qPCR assays performed in 1X ThermoPol buffer using indicated copies of synthetic RNA and RTX or RTX Exo- and Taq DNA polymerases. Panels A, B, and C depict TaqMan assays containing both RTX and Taq DNA polymerases. Panels D, E, and F depict TaqMan assays containing only Taq DNA polymerase. Panels G, H, and I depict TaqMan assays containing RTX Exo- and Taq DNA polymerases. Amplification curves from reactions containing 100,000 (black traces), 10,000 (blue traces), 1,000 (red traces), and 100 (pink traces) copies of SARS-CoV-2 synthetic N RNA are depicted. Negative control reactions either contained no templates (gray traces) or contained 100,000 copies of synthetic N RNA from MERS-CoV (green traces). Cq values of all assays are tabulated.



**Figure 4.** CDC SARS-CoV-2 N1, N2, and N3 TaqMan RT-qPCR assays performed in 1X RTX buffer using indicated copies of synthetic RNA and RTX or RTX Exo- and Taq DNA polymerases. Panels A, B, and C depict TaqMan assays containing RTX and Taq DNA polymerases. Panels D, E, and F depict TaqMan assays containing RTX Exo- and Taq DNA polymerases. Amplification curves from reactions containing 100,000 (black traces), 10,000 (blue traces), 1,000 (red traces), and 100 (pink traces) copies of SARS-CoV-2 synthetic N RNA are depicted. Negative control reactions either contained no templates (gray traces) or contained 100,000 copies of synthetic N RNA from MERS-CoV (green traces). Cq values of all assays are tabulated.

To compare the RTX-based TaqMan qRT-PCR with qRT-PCR performed using commercial master mix, duplicate reactions (**Table 6**) were set up using TaqPath<sup>™</sup> 1-Step RT-qPCR Master Mix, CG (Thermo Fisher Scientific, Waltham, MA, USA). All reactions were assembled on cold block and then incubated at room temperature (25 °C benchtop) for 2 min prior to being loaded on a real-time PCR machine set to cycle through the steps indicated in **Table 7**.

	Volume
	(μL)
RNA sample	5
6.7 μM Forward PCR primer	1.5
6.7 μM Reverse PCR primer	1.5
1.7 μM TaqMan probe	1.5
4X TaqPath <sup>™</sup> master mix	5
Water	5.5
Total reaction volume	20

 Table 6. CDC N1, N2, N3 TaqMan qRT-PCR assay set up using commercial qRT-PCR master mix

**Table 7**. TaqPath<sup>™</sup> TaqMan qRT-PCR cycling conditions

Step 1	50 °C	15 min
Step 2	95 °C	2 min
45 cycles of Steps 3 and 4		and 4
Step 3	95 °C	3 sec
Step 4	55 °C	30 sec

**Figure 5** depicts typical results from CDC N1, N2, and N3 TaqMan qRT-PCR reactions performed using synthetic RNA and TaqPath<sup>™</sup> commercial qRT-PCR mastermix.



**Figure 5.** CDC SARS-CoV-2 N1, N2, and N3 TaqMan qRT-PCR assays performed using indicated copies of synthetic RNA and TaqPath<sup>™</sup> commercial qRT-PCR mastermix. Amplification curves from reactions containing 100,000 (black traces), 10,000 (blue traces), 1,000 (red traces), and 100 (pink traces) copies of SARS-CoV-2 synthetic N RNA are depicted. Negative control reactions either contained no templates (gray traces) or contained 100,000 copies of synthetic N RNA from MERS-CoV (green traces). Cq values of all assays are tabulated.

The results with either a viral RT (in TaqPath<sup>™</sup> commercial mastermix) or RTX are summarized in **Table 8**. As can be seen, RTX-based TaqMan assays are of comparable sensitivity and specificity to the gold standard TaqPath assay. While all three RTX-based TaqMan assays performed in ThermoPol buffer consistently yielded Cq values, these were somewhat delayed compared to Cq values obtained with the commercial TaqPath<sup>™</sup> Master Mix. In contrast, RTX-based TaqMan assays when executed in RTX buffer yielded Cq values that were closer to Cq values obtained with the commercial mastermix.

N1 qRT-PCR	RTX+Taq in ThermoPol	RTX+Taq in RTX buffer	TaqPath™
100,000 nCoV-2019 N RNA	26.19	22.18	22.02
10,000 nCoV-2019 N RNA	29.71	25.54	25.55
1,000 nCoV-2019 N RNA	33.21	28.88	28.79
100 nCoV-2019 N RNA	36.29	32.06	32.21
100,000 MERS N RNA			
No template			
N2 qRT-PCR	RTX+Taq in ThermoPol	RTX+Taq in RTX buffer	TaqPath™
100,000 nCoV-2019 N RNA	28.81	23.77	21.57
10,000 nCoV-2019 N RNA	32.47	27.46	25.16
1,000 nCoV-2019 N RNA	35.78	31.09	28.90
100 nCoV-2019 N RNA	39.53	34.97	32.58
100,000 MERS N RNA			
No template	an an		
N3 qRT-PCR	RTX+Taq in ThermoPol	RTX+Taq in RTX buffer	TaqPath™
100,000 nCoV-2019 N RNA	28.03	23.28	21.20
10,000 nCoV-2019 N RNA	31.53	26.55	24.56
1,000 nCoV-2019 N RNA	35.16	29.77	27.98
100 nCoV-2019 N RNA	38.51	35.44	31.59
100,000 MERS N RNA			
No template			

 Table 8. Comparison of Cq values of RTX-based and TaqPath-based SARS-CoV-2 TaqMan qRT-PCR assays

## Summary of assay performance

Overall, RTX performs well either as a single enzyme for RT-qPCR with intercalating dyes, such as EvaGreen, or as the RT component of a TaqMan based assay. At the lowest RNA concentrations examined (100 copies), the gold standard TaqPath assay general gave a signal at a Cq value of ca. 32. The more robust full-length RTX on its own actually performed slightly better, with C<sub>q</sub> values typically from 29-30 (although the N3 primer set gives higher signal and higher background). When used as a substitute for a viral RT, RTX is slightly less sensitive overall, with Cq values typically from 32-35.

# Purification of RTX

For sites wishing to produce their own RTX, the following protocol can be used.

	• 50 mM Phosphate Buffer, pH 7.5
	• 300 mM NaCl
Decusponsion Duffer	• 20 mM imidazole
Resuspension Buller	• 0.1% Igepal CO-630
	(non-toxic Non-idet P40 equivalent)
	• 5 mM MgSO4
	• 50 mM Phosphate Buffer, pH 7.5
Equilibration Buffer	• 300 mM NaCl
-	• 20 mM imidazole
	30 mL Resuspension Buffer
	+
Lysis Buffer	<ul> <li>1x EDTA-free protease inhibitor tablet</li> </ul>
-	+
	• 30-60 mg HEW Lysozyme
	• 50 mM Phosphate Buffer, pH 7.5
Wash Buffer	• 300 mM NaCl
	• 50 mM imidazole
	• 50 mM Phosphate Buffer, pH 7.5
Elution Buffer	• 300 mM NaCl
	• 250 mM imidazole
	• 40 mM Tris-HCl, pH 7.5
Llonovin Duffer A	<ul> <li>100 mM NaCl</li> </ul>
Heparin Butter A	• 0.1% Igepal CO-630
	(non-toxic Non-idet P40 equivalent)
	• 40 mM Tris-HCl, pH 7.5
Honorin Puffor P	• 2 M NaCl
neparin buller b	• 0.1% Igepal CO-630
	(non-toxic Non-idet P40 equivalent)

Table 7. Buffers for RTX purification

	• 40 mM Tris-HCl, pH 7.5
Ni-NTA Dialvsis Buffer	<ul> <li>10 mM beta-mercapto ethanol (BME)</li> </ul>
(22.1)	or
(2x 2 L)	1 mM DTT
	<ul> <li>0.1% Igepal CO-630</li> </ul>
	(non-toxic Non-idet P40 equivalent)
	• 50 mM Tris-HCl, pH 8.0
	• 50 mM KCl
Heparin Dialysis Buffer (2 L)	• 0.1% Tween-20
Final Dialysis Buffer (1 L)	• 0.1% Igepal CO-630
	(non-toxic Non-idet P40 equivalent)
	• 50% Glycerol
	• 50 mM Tris-HCl, pH 8.0

• 50 mM KCl
• 0.1% Tween-20
• 0.1% Igepal CO-630
(non-toxic Non-idet P40 equivalent)

 Table 8. Buffers for dialysis of RTX

In brief, a T7-based *E. coli* expression strain such as BL21(DE3) was transformed with a RTX expression plasmid that contained an carboxyl terminal six histidine tag for purification. On the next day, an individual transformant colony was picked into desired media with appropriate antibiotics (e.g. ampicillin) and 1-2% glucose. This starter culture was grown overnight in shaking incubator at 37°C.

On day three, 1 L Superior Broth<sup>M</sup> (Athena Enzyme Systems<sup>M</sup>) or other rich media with appropriate antibiotics in a 4 L Erlenmeyer flask was inoculated with 250-350 µL of the overnight starter culture. Incubate at 37°C, 250 rpm until the culture reaches OD<sub>600</sub> of 0.4-0.7 (0.5-0.6 is ideal). When desired OD<sub>600</sub> is reached, place flask(s) in 4°C cold room for 30-45 min and set shaking incubator to 18°C. Following cold incubation, induce expression of RTX by adding 1 mL of 1 M IPTG to each 1 L of culture. Place flask(s) in the 18°C shaking incubator for 16-18 hours.

Following overnight expression, cells were harvested by centrifuging cultures at 4°C, 5000 xg for 20 min. On ice, resuspend cell pellet in 30 mL cold **Lysis Buffer**. Transfer the resuspended cell pellet to a small 50 mL beaker with clean stir bar and place securely in an ice bath. With moderate stirring, sonicate the sample using 40% amplitude and 1 sec ON / 4 sec OFF for 4 min total sonication time. Centrifuge the resulting lysate at 4°C, 35,000 xg for 30 min. Carefully transfer the supernatant to a clean ultracentrifugation tube, shake in thermomixer at 400 rpm, 85°C for 10 min, and then place on ice for 10 min. Centrifuged the heat-treated lysate at 4°C, 35,000 xg for 30 min. Carefully transfer the supernatant to a clean tube and filter the clarified lysate using a 0.2  $\mu$ m filter.

Perform IMAC column purification in a cold room or allow gravity flow to occur in a refrigerator. Prepare Ni-NTA agarose columns with final column volume (CV) of 1 mL. Apply the clarified lysate from 1 L of expressed culture to a previously equilibrated column and collect the flow-through. Wash column with 20 CV **Equilibration Buffer** and collect the flow-through. Wash column with 5 CV **Wash Buffer** and collect flow-through. Elute RTX from column with 5 mL **Elution Buffer** and transfer to a dialysis cassette with appropriate molecular weight cut-off. Dialyze eluate into 2 L of **Ni-NTA Dialysis Buffer** for 3-4 hours at 4°C. Then dialyze eluate into a second 2 L of **Ni-NTA Dialysis Buffer** overnight at 4°C.

Dialyzed eluate was then passed over an equilibrated 5 mL heparin column (HiTrap<sup>™</sup> Heparin HP) and eluted along a sodium chloride gradient (100 mM to 2 M NaCl). The peak corresponding to RTX or its variants could be expected between 40-60% Buffer B. RTX fractions were collected, pooled, and dialyzed into 2 L **Heparin Dialysis Buffer** for 3-4 hours. Then dialysis cassettes were transferred into 2 L **Final Dialysis Buffer** overnight where it is expected that the protein sample

volume will decrease significantly. Following the final dialysis, protein sample was recovered and protein concentration was determined before storing at -20°C.

A similar protocol should work for Taq and other thermostable polymerases, but may require different dialysis and storage buffers.

## Summary and opportunities for distribution

By obtaining a RTX expression vector and purifying the thermostable reverse transcriptase it should prove possible to carry out RT-qPCR reactions with a sensitivity similar to that already observed for approved kits, with few or no false positive results. The sequence information for plasmids for RTX [pET\_RTX-6xHis] and its exonuclease deficient variant, RTX(exo-) [pET\_RTX(exo -)-6xHis], bearing a C-terminal 6xHis tags can be found in the **Supplemental Information**. These plasmids will be uploaded to Addgene in short order, or can be obtained via <u>https://reclone.org/</u>. The Board of Regents of The University of Texas has licensed IP covering RTX to Promega Corporation.

Acknowledgements: We acknowledge the extraordinary contributions of the Schoggins lab in making synthetic RNA available on short notice for these assays.



Figure S1. Plasmid map of pET\_RTX-6xHis

LOCUS pE	T_RTX-6xHis	8232 bp	DNA	circular UNA 26-MAR-2020	
DEFINITION A new nucleotide sequence entered manually.					
ACCESSION urn.localf-bq4nr1q					
VERSION urn.localf-bq4nr1q					
KEYWORDS .					
SOURCE					
ORGANISM					
FEATURES	Location/Qua	lifiers			
gene	1401000				

# Supporting Information

/created_by="User"				
/label="bla (AmpR)"				
rep_origin 10421555				
/created_by="User"				
/modified_by="User"				
/label="M13 ori"				
rep_origin 16662254				
/created_by="User"				
/modified_by="User"				
/label="pMB1 ori"				
gene complement(26232814)				
/created_by="User"				
/label="rop"				
gene complement(33714453)				
/created_by="User"				
/label="laci"				
promoter 56375654				
/created_by="User"				
/label="T7 Promoter"				
operon 56565680				
/created_by="User"				
/label="lacO"				
misc_feature 57225727				
/created_by="User"				
/modified_by="User"				
/label="Ndel"				
CDS 57258085				
/created_by="User"				
/modified_by="User"				
/label="RTX"				
misc_difference 58365838				
/created_by="User"				
/label="F38L"				
misc_difference 60136015				
/created_by="User"				
/label="R97M"				
misc_difference 60766078				
/created_by="User"				
/modified_by="User"				
/label="K118I"				
misc_difference 61336135				
/created_by="User"				
/label="M137L"				
misc_difference 68656867				

/created by="User" /label="R381H" misc difference 6874..6876 /created by="User" /label="Y384H" misc\_difference 6889..6891 /created by="User" /label="V389I" misc difference 7120..7122 /created by="User" /label="K466R" misc difference 7201..7203 /created by="User" /label="Y493L" misc\_difference 7264..7266 /created by="User" /label="T514I" misc difference 7285..7287 /created by="User" /label="I521L" misc\_difference 7483..7485 /created by="User" /label="F587L" misc difference 7714..7716 /created by="User" /label="E664K" misc difference 7855..7857 /created by="User" /label="G711V" misc difference 7927..7929 /created by="User" /label="N735K" misc difference 8026..8028 /created by="User" /label="W768R" misc feature 8047..8052 /created by="User" /label="Xhol" misc feature 8053..8058 /created by="User" /label="BamHI" 8104..8226 terminator /created by="User" /label="T7 Tphi transcription terminator"

#### ORIGIN

1 aactacgtca ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttattttt 61 ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata 121 atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt 181 tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc 241 tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat 301 ccttgagagt tttcgccccg aagaacgttt cccaatgatg agcactttta aagttctgct 361 atgtggcgcg gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca 421 ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 481 catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa 541 cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg 601 ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga 661 cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg 721 cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt 781 tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg 841 agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc 901 ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca 961 gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc 1021 atatatactt tagattgatt taccccggtt gataatcaga aaagccccaa aaacaggaag 1081 attgtataag caaatattta aattgtaaac gttaatattt tgttaaaatt cgcgttaaat 1141 ttttgttaaa tcagctcatt ttttaaccaa taggccgaaa tcggcaaaat cccttataaa 1201 tcaaaagaat agcccgagat agggttgagt gttgttccag tttggaacaa gagtccacta 1261 ttaaagaacg tggactccaa cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca 1321 ctacgtgaac catcacccaa atcaagtttt ttggggtcga ggtgccgtaa agcactaaat 1381 cggaacccta aagggagccc ccgatttaga gcttgacggg gaaagccggc gaacgtggcg 1441 agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc 1501 acgctgcgcg taaccaccac acccgccgcg cttaatgcgc cgctacaggg cgcgtaaaag 1561 gatctaggtg aagatccttt ttgataatct catgaccaaa atcccttaac gtgagttttc 1621 gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt 1681 tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt 1741 gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat 1801 accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 1861 accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa 1921 gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg 1981 ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 2041 atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa aggcggacag 2101 gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 2161 cgcctggtat ctttatagtc ctgttgggtt tcgccacctc tgacttgagc gtcgattttt 2221 gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg 2281 gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc 2341 tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac 2401 cgagcgcagc gagtcagtga gcgaggaagc tatggtgcac tctcagtaca atctgctctg 2461 atgccgcata gttaagccag tatacactcc gctatcgcta cgtgactggg tcatggctgc 2521 gccccgacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc tcccggcatc

2581 cgcttacaga caagctgtga ccgtctccgg gagctgcatg tgtcagaggt tttcaccgtc 2641 atcaccgaaa cgcgcgaggc agctgcggta aagctcatca gcgtggtcgt gcagcgattc 2701 acagatgtct gcctgttcat ccgcgtccag ctcgttgagt ttctccagaa gcgttaatgt 2761 ctggcttctg ataaagcggg ccatgttaag ggcggttttt tcctgtttgg tcactgatgc 2821 ctccgtgtaa gggggatttc tgttcatggg ggtaatgata ccgatgaaac gagagaggat 2881 gctcacgata cgggttactg atgatgaaca tgcccggtta ctggaacgtt gtgagggtaa 2941 acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc aatgccagcc 3001 gaacgccagc aagacgtagc ccagcgcgtc ggccgccatg ccggcgataa tggcctgctt 3061 ctcgccgaaa cgtttggtgg cgggaccagt gacgaaggct tgagcgaggg cgtgcaagat 3121 tccgaatacc gcaagcgaca ggccgatcat cgtcgcgctc cagcgaaagc ggtcctcgcc 3181 gaaaatgacc cagagcgctg ccggcacctg tcctacgagt tgcatgataa agaagacagt 3241 cataagtgcg gcgacgatag tcatgccccg cgcccaccgg aaggagctga ctgggttgaa 3301 ggctctcaag ggcatcggtc gagatcccgg tgcctaatga gtgagctaac ttacattaat 3361 tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 3421 aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgccagggtg gtttttcttt 3481 tcaccagtga gacgggcaac agctgattgc ccttcaccgc ctggccctga gagagttgca 3541 gcaagcggtc cacgctggtt tgccccagca ggcgaaaatc ctgtttgatg gtggttaacg 3601 gcgggatata acatgagctg tcttcggtat cgtcgtatcc cactaccgag atgtccgcac 3661 caacgcgcag cccggactcg gtaatggcgc gcattgcgcc cagcgccatc tgatcgttgg 3721 caaccagcat cgcagtggga acgatgccct cattcagcat ttgcatggtt tgttgaaaac 3781 cggacatggc actccagtcg ccttcccgtt ccgctatcgg ctgaatttga ttgcgagtga 3841 gatatttatg ccagccagcc agacgcggac gcgccgagac agaacttaat gggcccgcta 3901 acagcgcgat ttgctggtga cccaatgcga ccagatgctc cacgcccagt cgcgtaccgt 3961 cttcatggga gaaaataata ctgttgatgg gtgtctggtc agagacatca agaaataacg 4021 ccggaacatt agtgcaggca gcttccacag caatggcatc ctggtcatcc agcggatagt 4081 taatgatcag cccactgacg cgttgcgcga gaagattgtg caccgccgct ttacaggctt 4141 cgacgccgct tcgttctacc atcgacacca ccacgctggc acccagttga tcggcgcgag 4201 atttaatcgc cgcgacaatt tgcgacggcg cgtgcagggc cagactggag gtggcaacgc 4261 caatcagcaa cgactgtttg cccgccagtt gttgtgccac gcggttggga atgtaattca 4321 gctccgccat cgccgcttcc actttttccc gcgttttcgc agaaacgtgg ctggcctggt 4381 tcaccacgcg ggaaacggtc tgataagaga caccggcata ctctgcgaca tcgtataacg 4441 ttactggttt cacattcacc accctgaatt gactctcttc cgggcgctat catgccatac 4501 cgcgaaaggt tttgcgccat tcgatggtgt ccgggatctc gacgctctcc cttatgcgac 4561 tcctgcatta ggaagcagcc cagtagtagg ttgaggccgt tgagcaccgc cgccgcaagg 4621 aatggtgcat gccggcatgc cgccctttcg tcttcaagaa ttaattccca attccccagg 4681 catcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgttttat ctgttgtttg 4741 tcggtgaacg ctctcctgag taggacaaat ccgccgggag cggatttgaa cgttgcgaag 4801 caacggcccg gagggtggcg ggcaggacgc ccgccataaa ctgccaggaa ttaattcccc 4861 aggcatcaaa taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt 4921 ttgtcggtga acgctctcct gagtaggaca aatccgccgg gagcggattt gaacgttgcg 4981 aagcaacggc ccggagggtg gcgggcagga cgcccgccat aaactgccag gaattaattc 5041 cccaggcatc aaataaaacg aaaggctcag tcgaaagact gggcctttcg ttttatctgt 5101 tgtttgtcgg tgaacgctct cctgagtagg acaaatccgc cgggagcgga tttgaacgtt 5161 gcgaagcaac ggcccggagg gtggcgggca ggacgcccgc cataaactgc caggaattaa 5221 ttccccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt tcgttttatc 5281 tgttgtttgt cggtgaacgc tctcctgagt aggacaaatc cgccgggagc ggatttgaac 5341 gttgcgaagc aacggcccgg agggtggcgg gcaggacgcc cgccataaac tgccaggaat 5401 taattcccca ggcatcaaat aaaacgaaag gctcagtcga aagactgggc ctttcgtttt 5461 atctgttgtt tgtcggtgaa cgctctcctg agtaggacaa atccgccggg agcggatttg 5521 aacgttgcga agcaacggcc cggagggtgg cgggcaggac gcccgccata aactgccagg 5581 aattggggat cggaattaat tcccggttta aaccggggat ctcgatcccg cgaaattaat 5641 acgactcact ataggggaat tgtgagcgga taacaattcc cctctagaaa taattttgtt 5701 taactttaag aaggagatat acatatgatt cttgacacgg attacatcac ggaagacggc 5761 aagccggtta tccgtatttt caagaaagaa aacggcgaat tcaagattga atacgatcgg 5821 acatttgaac cgtacctgta cgctctcctc aaggatgata gcgcaatcga agaagtgaaa 5881 aaaatcaccg cagagcggca tggcacagtg gtaacagtta agcgggtcga gaaagtgcag 5941 aagaagttet taggeeggee agtegaagta tggaaattat aetteacaea teeacaggae 6001 gttccggcga tcatggataa gattcgggag catccggcgg taatcgatat ctatgaatac 6061 gatattccgt tcgctattcg ctaccttatt gacaaaggtt tagttccaat ggagggtgat 6121 gaggaactta aactgttagc attcgatatc gaaacacttt atcacgaagg tgaagagttt 6181 gccgaaggtc cgattttaat gatctcgtac gccgatgaag aaggcgcacg cgtaattacg 6241 tggaaaaatg tggacctccc gtacgtagac gtagtgagca ctgagcgcga gatgattaaa 6301 cgtttccttc gggtagtaaa agaaaaagac ccagacgtgc tgattacgta taacggcgac 6361 aactttgatt ttgcctatct caagaagcgt tgcgaaaagt taggcattaa tttcgccctg 6421 ggtcgggacg gttcagagcc gaaaattcag cggatgggcg accgctttgc tgtggaggta 6481 aaaggtcgca tccatttcga tttatatccg gttatccggc gcaccatcaa cttgccgact 6541 tacacacttg aagcagttta cgaagcggtg ttcggccaac caaaagaaaa ggtttatgcc 6601 gaggagatta ccaccgcatg ggaaactggc gaaaacttgg agcgggtggc tcggtattcc 6661 atggaagatg ccaaggtgac ctacgaactg ggcaaagagt ttttaccgat ggaagcacaa 6721 ttaagccgcc ttattggtca gtccctctgg gatgtgtcgc gttcttcaac gggcaattta 6781 gtcgaatggt ttcttcttcg gaaagcgtac gagcgtaacg agcttgctcc aaataagcca 6841 gacgaaaaag aattggctcg gcgccatcag tcacatgagg gcggctacat taaggagcca 6901 gaacggggct tgtgggagaa catcgtctac cttgattttc ggtctcttta tccgtctatt 6961 atcatcacac ataacgtctc gccagatacc ctgaaccgtg aaggctgtaa agaatatgat 7021 gtggcaccac aggtcggcca tcgtttttgt aaagacttcc cgggcttcat tccatctctt 7081 ctgggtgatt tgttagaaga gcgtcaaaag atcaagaaac gtatgaaagc gacaattgac 7141 ccaattgaac gcaaattact tgattaccgt cagcgtgcaa tcaagatcct cgcgaactct 7201 ctgtacggtt attacggcta cgcacgcgcc cggtggtatt gcaaagaatg tgcagaatca 7261 gtcattgctt ggggtcggga gtacctgacc atgacgatta aggaaattga ggagaaatac 7321 ggtttcaagg tcatctatag tgacacggat ggtttctttg caacgattcc aggtgcggac 7381 gcagaaactg taaagaaaaa ggcaatggag ttcttgaagt atattaatgc gaagttgcca 7441 ggcgccctgg aattagagta cgaaggtttt tataagcgtg gcctgttcgt gacaaagaag 7501 aaatacgcgg taattgacga ggaaggcaag atcacaactc gtggcttgga aattgttcgt 7561 cgcgattgga gcgagatcgc aaaggagacc caagctcgtg tgttggaggc cctcctgaag 7621 gatggtgacg tcgaaaaagc cgtacgcatc gttaaggagg ttacagagaa gcttagcaag 7681 tatgaggtcc caccagagaa acttgttatt cataaacaaa tcactcgcga ccttaaagac 7741 tataaggcca ctggtccaca cgtcgccgta gcaaagcggc ttgcggctcg gggcgtcaag 7801 attcggccag gcacggttat tagttacatc gtcctcaaag gctcaggccg gattgttgat

7861 cgcgcgattc catttgatga atttgatccg acgaagcata aatatgatgc ggaatattac 7921 attgaaaaac aggttctgcc ggcggtggag cgcatcttac gtgcgttcgg ctatcgcaag 7981 gaggatttgc ggtaccagaa aactcgtcaa gtcggtttga gtgcccgtct gaagccgaaa 8041 ggtaccctcg agggatccag ccatcaccat caccatcact gataagcggc cgcattgatc 8101 cggctgctaa caaagcccga aaggaagctg agttggctgc tgccaccgct gagcaataac 8161 tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg aaaggaggaa 8221 ctatatccgg at





Figure S2. Plasmid map pET\_RTX(exo-)-6xHis

LOCUS pET\_RTX(exo-)-6x 8232 bp DNA circular UNA 26-MAR-2020 DEFINITION A new nucleotide sequence entered manually.

ACCESSION	urn.local28-bq4yik3				
VERSION urn.local28-bq4yik3					
KEYWORDS .					
SOURCE					
ORGANISM	Л				
•					
FEATURES	Location/Qualifiers				
gene	1401000				
	/created_by="User"				
	/label="bla (AmpR)"				
rep_origin 10421555					
	/created_by="User"				
	/modified_by="User"				
	/label="M13 ori"				
rep_orig	in 16662254				
	/created_by="User"				
	/modified_by="User"				
	/label="pMB1 ori"				
gene	complement(26232814)				
	/created_by="User"				
	/label="rop"				
gene	complement(33714453)				
	/created_by="User"				
	/label="lacl"				
promote	er 56375654				
	/created_by="User"				
	/label="T7 Promoter"				
operon	56565680				
	/created_by="User"				
	/label="lacO"				
RBS	57085714				
	/created_by="User"				
	/label="RBS"				
misc feature 57225727					
	/created_by="User"				
	/modified_by="User"				
	/label="Ndel"				
CDS	57258085				
	/created_by="User"				
	/modified_by="User"				
	/label="RTX"				
, misc difference 58365838					
—	/created by="User"				
	/label="F38L"				

misc difference 6013..6015 /created by="User" /label="R97M" misc difference 6076..6078 /created by="User" /modified\_by="User" /label="K118I" misc difference 6133..6135 /created by="User" /label="M137L" misc difference 6352..6354 /created\_by="User" /label="N210D (exo- mutation)" misc difference 6865..6867 /created\_by="User" /label="R381H" misc difference 6874..6876 /created by="User" /label="Y384H" misc difference 6889..6891 /created by="User" /label="V389I" misc difference 7120..7122 /created by="User" /label="K466R" misc difference 7201..7203 /created by="User" /label="Y493L" misc\_difference 7264..7266 /created by="User" /label="T514I" misc\_difference 7285..7287 /created by="User" /label="I521L" misc difference 7483..7485 /created by="User" /label="F587L" misc\_difference 7714..7716 /created by="User" /label="E664K" misc\_difference 7855..7857 /created by="User" /label="G711V" misc difference 7927..7929

/created\_by="User" /label="N735K" misc\_difference 8026..8028 /created\_by="User" /label="W768R" misc\_feature 8047..8052 /created\_by="User" /label="Xhol" misc\_feature 8053..8058 /created\_by="User" /label="BamHI" terminator 8104..8226 /created\_by="User" /label="T7 Tphi transcription terminator"

### ORIGIN

1 aactacgtca ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttattttt 61 ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata 121 atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt 181 tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc 241 tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat 301 ccttgagagt tttcgccccg aagaacgttt cccaatgatg agcactttta aagttctgct 361 atgtggcgcg gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca 421 ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 481 catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa 541 cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg 601 ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga 661 cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg 721 cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt 781 tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg 841 agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc 901 ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca 961 gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc 1021 atatatactt tagattgatt taccccggtt gataatcaga aaagccccaa aaacaggaag 1081 attgtataag caaatattta aattgtaaac gttaatattt tgttaaaatt cgcgttaaat 1141 ttttgttaaa tcagctcatt ttttaaccaa taggccgaaa tcggcaaaat cccttataaa 1201 tcaaaagaat agcccgagat agggttgagt gttgttccag tttggaacaa gagtccacta 1261 ttaaagaacg tggactccaa cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca 1321 ctacgtgaac catcacccaa atcaagtttt ttggggtcga ggtgccgtaa agcactaaat 1381 cggaacccta aagggagccc ccgatttaga gcttgacggg gaaagccggc gaacgtggcg 1441 agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc 1501 acgctgcgcg taaccaccac acccgccgcg cttaatgcgc cgctacaggg cgcgtaaaag 1561 gatctaggtg aagatcettt ttgataatet catgaccaaa atecettaac gtgagtttte 1621 gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt 1681 tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt

1741 gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat 1801 accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc 1861 accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa 1921 gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg 1981 ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag 2041 atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa aggcggacag 2101 gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa 2161 cgcctggtat ctttatagtc ctgttgggtt tcgccacctc tgacttgagc gtcgattttt 2221 gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg 2281 gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc 2341 tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac 2401 cgagcgcagc gagtcagtga gcgaggaagc tatggtgcac tctcagtaca atctgctctg 2461 atgccgcata gttaagccag tatacactcc gctatcgcta cgtgactggg tcatggctgc 2521 gccccgacac ccgccaacac ccgctgacgc gccctgacgg gcttgtctgc tcccggcatc 2581 cgcttacaga caagctgtga ccgtctccgg gagctgcatg tgtcagaggt tttcaccgtc 2641 atcaccgaaa cgcgcgaggc agctgcggta aagctcatca gcgtggtcgt gcagcgattc 2701 acagatgtct gcctgttcat ccgcgtccag ctcgttgagt ttctccagaa gcgttaatgt 2761 ctggcttctg ataaagcggg ccatgttaag ggcggttttt tcctgtttgg tcactgatgc 2821 ctccgtgtaa gggggatttc tgttcatggg ggtaatgata ccgatgaaac gagagaggat 2881 gctcacgata cgggttactg atgatgaaca tgcccggtta ctggaacgtt gtgagggtaa 2941 acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc aatgccagcc 3001 gaacgccagc aagacgtagc ccagcgcgtc ggccgccatg ccggcgataa tggcctgctt 3061 ctcgccgaaa cgtttggtgg cgggaccagt gacgaaggct tgagcgaggg cgtgcaagat 3121 tccgaatacc gcaagcgaca ggccgatcat cgtcgcgctc cagcgaaagc ggtcctcgcc 3181 gaaaatgacc cagagcgctg ccggcacctg tcctacgagt tgcatgataa agaagacagt 3241 cataagtgcg gcgacgatag tcatgccccg cgcccaccgg aaggagctga ctgggttgaa 3301 ggctctcaag ggcatcggtc gagatcccgg tgcctaatga gtgagctaac ttacattaat 3361 tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 3421 aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgccagggtg gtttttcttt 3481 tcaccagtga gacgggcaac agctgattgc ccttcaccgc ctggccctga gagagttgca 3541 gcaagcggtc cacgctggtt tgccccagca ggcgaaaatc ctgtttgatg gtggttaacg 3601 gcgggatata acatgagctg tcttcggtat cgtcgtatcc cactaccgag atgtccgcac 3661 caacgcgcag cccggactcg gtaatggcgc gcattgcgcc cagcgccatc tgatcgttgg 3721 caaccagcat cgcagtggga acgatgccct cattcagcat ttgcatggtt tgttgaaaac 3781 cggacatggc actccagtcg ccttcccgtt ccgctatcgg ctgaatttga ttgcgagtga 3841 gatatttatg ccagccagcc agacgcggac gcgccgagac agaacttaat gggcccgcta 3901 acagcgcgat ttgctggtga cccaatgcga ccagatgctc cacgcccagt cgcgtaccgt 3961 cttcatggga gaaaataata ctgttgatgg gtgtctggtc agagacatca agaaataacg 4021 ccggaacatt agtgcaggca gcttccacag caatggcatc ctggtcatcc agcggatagt 4081 taatgatcag cccactgacg cgttgcgcga gaagattgtg caccgccgct ttacaggctt 4141 cgacgccgct tcgttctacc atcgacacca ccacgctggc acccagttga tcggcgcgag 4201 atttaatcgc cgcgacaatt tgcgacggcg cgtgcagggc cagactggag gtggcaacgc 4261 caatcagcaa cgactgtttg cccgccagtt gttgtgccac gcggttggga atgtaattca 4321 gctccgccat cgccgcttcc actttttccc gcgttttcgc agaaacgtgg ctggcctggt

4381 tcaccacgcg ggaaacggtc tgataagaga caccggcata ctctgcgaca tcgtataacg 4441 ttactggttt cacattcacc accctgaatt gactctcttc cgggcgctat catgccatac 4501 cgcgaaaggt tttgcgccat tcgatggtgt ccgggatctc gacgctctcc cttatgcgac 4561 tcctgcatta ggaagcagcc cagtagtagg ttgaggccgt tgagcaccgc cgccgcaagg 4621 aatggtgcat gccggcatgc cgccctttcg tcttcaagaa ttaattccca attccccagg 4681 catcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgttttat ctgttgtttg 4741 tcggtgaacg ctctcctgag taggacaaat ccgccgggag cggatttgaa cgttgcgaag 4801 caacggcccg gagggtggcg ggcaggacgc ccgccataaa ctgccaggaa ttaattcccc 4861 aggcatcaaa taaaacgaaa ggctcagtcg aaagactggg cctttcgttt tatctgttgt 4921 ttgtcggtga acgctctcct gagtaggaca aatccgccgg gagcggattt gaacgttgcg 4981 aagcaacggc ccggagggtg gcgggcagga cgcccgccat aaactgccag gaattaattc 5041 cccaggcatc aaataaaacg aaaggctcag tcgaaagact gggcctttcg ttttatctgt 5101 tgtttgtcgg tgaacgctct cctgagtagg acaaatccgc cgggagcgga tttgaacgtt 5161 gcgaagcaac ggcccggagg gtggcgggca ggacgcccgc cataaactgc caggaattaa 5221 ttccccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt tcgttttatc 5281 tgttgtttgt cggtgaacgc tctcctgagt aggacaaatc cgccgggagc ggatttgaac 5341 gttgcgaagc aacggcccgg agggtggcgg gcaggacgcc cgccataaac tgccaggaat 5401 taattcccca ggcatcaaat aaaacgaaag gctcagtcga aagactgggc ctttcgtttt 5461 atctgttgtt tgtcggtgaa cgctctcctg agtaggacaa atccgccggg agcggatttg 5521 aacgttgcga agcaacggcc cggagggtgg cgggcaggac gcccgccata aactgccagg 5581 aattggggat cggaattaat tcccggttta aaccggggat ctcgatcccg cgaaattaat 5641 acgactcact ataggggaat tgtgagcgga taacaattcc cctctagaaa taattttgtt 5701 taactttaag aaggagatat acatatgatt cttgacacgg attacatcac ggaagacggc 5761 aagccggtta tccgtatttt caagaaagaa aacggcgaat tcaagattga atacgatcgg 5821 acatttgaac cgtacctgta cgctctcctc aaggatgata gcgcaatcga agaagtgaaa 5881 aaaatcaccg cagagcggca tggcacagtg gtaacagtta agcgggtcga gaaagtgcag 5941 aagaagttct taggccggcc agtcgaagta tggaaattat acttcacaca tccacaggac 6001 gttccggcga tcatggataa gattcgggag catccggcgg taatcgatat ctatgaatac 6061 gatattccgt tcgctattcg ctaccttatt gacaaaggtt tagttccaat ggagggtgat 6121 gaggaactta aactgttagc attcgatatc gaaacacttt atcacgaagg tgaagagttt 6181 gccgaaggtc cgattttaat gatctcgtac gccgatgaag aaggcgcacg cgtaattacg 6241 tggaaaaatg tggacctccc gtacgtagac gtagtgagca ctgagcgcga gatgattaaa 6301 cgtttccttc gggtagtaaa agaaaaagac ccagacgtgc tgattacgta tgacggcgac 6361 aactttgatt ttgcctatct caagaagcgt tgcgaaaagt taggcattaa tttcgccctg 6421 ggtcgggacg gttcagagcc gaaaattcag cggatgggcg accgctttgc tgtggaggta 6481 aaaggtcgca tccatttcga tttatatccg gttatccggc gcaccatcaa cttgccgact 6541 tacacacttg aagcagttta cgaagcggtg ttcggccaac caaaagaaaa ggtttatgcc 6601 gaggagatta ccaccgcatg ggaaactggc gaaaacttgg agcgggtggc tcggtattcc 6661 atggaagatg ccaaggtgac ctacgaactg ggcaaagagt ttttaccgat ggaagcacaa 6721 ttaagccgcc ttattggtca gtccctctgg gatgtgtcgc gttcttcaac gggcaattta 6781 gtcgaatggt ttcttcttcg gaaagcgtac gagcgtaacg agcttgctcc aaataagcca 6841 gacgaaaaag aattggctcg gcgccatcag tcacatgagg gcggctacat taaggagcca 6901 gaacggggct tgtgggagaa catcgtctac cttgattttc ggtctcttta tccgtctatt 6961 atcatcacac ataacgtctc gccagatacc ctgaaccgtg aaggctgtaa agaatatgat

7021 gtggcaccac aggtcggcca tcgtttttgt aaagacttcc cgggcttcat tccatctctt 7081 ctgggtgatt tgttagaaga gcgtcaaaag atcaagaaac gtatgaaagc gacaattgac 7141 ccaattgaac gcaaattact tgattaccgt cagcgtgcaa tcaagatcct cgcgaactct 7201 ctgtacggtt attacggcta cgcacgcgcc cggtggtatt gcaaagaatg tgcagaatca 7261 gtcattgctt ggggtcggga gtacctgacc atgacgatta aggaaattga ggagaaatac 7321 ggtttcaagg tcatctatag tgacacggat ggtttctttg caacgattcc aggtgcggac 7381 gcagaaactg taaagaaaaa ggcaatggag ttcttgaagt atattaatgc gaagttgcca 7441 ggcgccctgg aattagagta cgaaggtttt tataagcgtg gcctgttcgt gacaaagaag 7501 aaatacgcgg taattgacga ggaaggcaag atcacaactc gtggcttgga aattgttcgt 7561 cgcgattgga gcgagatcgc aaaggagacc caagctcgtg tgttggaggc cctcctgaag 7621 gatggtgacg tcgaaaaagc cgtacgcatc gttaaggagg ttacagagaa gcttagcaag 7681 tatgaggtcc caccagagaa acttgttatt cataaacaaa tcactcgcga ccttaaagac 7741 tataaggcca ctggtccaca cgtcgccgta gcaaagcggc ttgcggctcg gggcgtcaag 7801 attcggccag gcacggttat tagttacatc gtcctcaaag gctcaggccg gattgttgat 7861 cgcgcgattc catttgatga atttgatccg acgaagcata aatatgatgc ggaatattac 7921 attgaaaaac aggttctgcc ggcggtggag cgcatcttac gtgcgttcgg ctatcgcaag 7981 gaggatttgc ggtaccagaa aactcgtcaa gtcggtttga gtgcccgtct gaagccgaaa 8041 ggtaccctcg agggatccag ccatcaccat caccatcact gataagcggc cgcattgatc 8101 cggctgctaa caaagcccga aaggaagctg agttggctgc tgccaccgct gagcaataac 8161 tagcataacc ccttggggcc tctaaacggg tcttgagggg ttttttgctg aaaggaggaa 8221 ctatatccgg at //

# Table S2. GenBank format of pET\_RTX(exo-)-6xHis

## REFERENCES

1 Ellefson, J. W. *et al.* Synthetic evolutionary origin of a proofreading reverse transcriptase. *Science* **352**, 1590-1593, doi:10.1126/science.aaf5409 (2016).