
Supplementary Information

Chimeric RNA design principles for RNA-mediated gene fusion

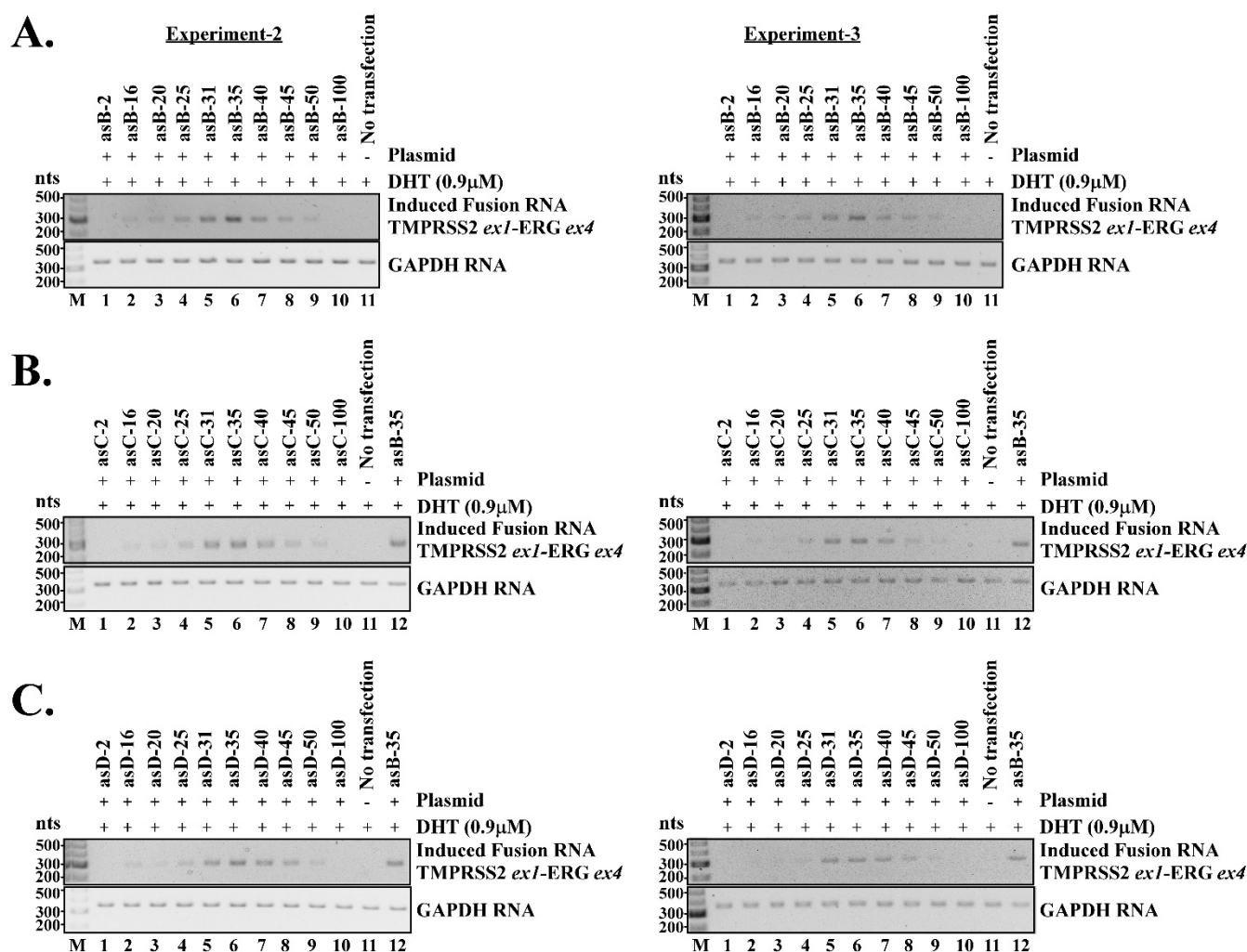
Sachin Kumar Gupta and Laising Yen*

¹ Department of Pathology & Immunology, Baylor College of Medicine, Houston, TX 77030, USA; sachinkumar.gupta@bcm.edu

² Department of Molecular & Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA

³ Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA

* Correspondence: author: yen@bcm.edu; Tel.: +1-713-798-1180



Supplementary Figure S1. The bulge size regulates the efficiency of RNA-mediated gene fusion.

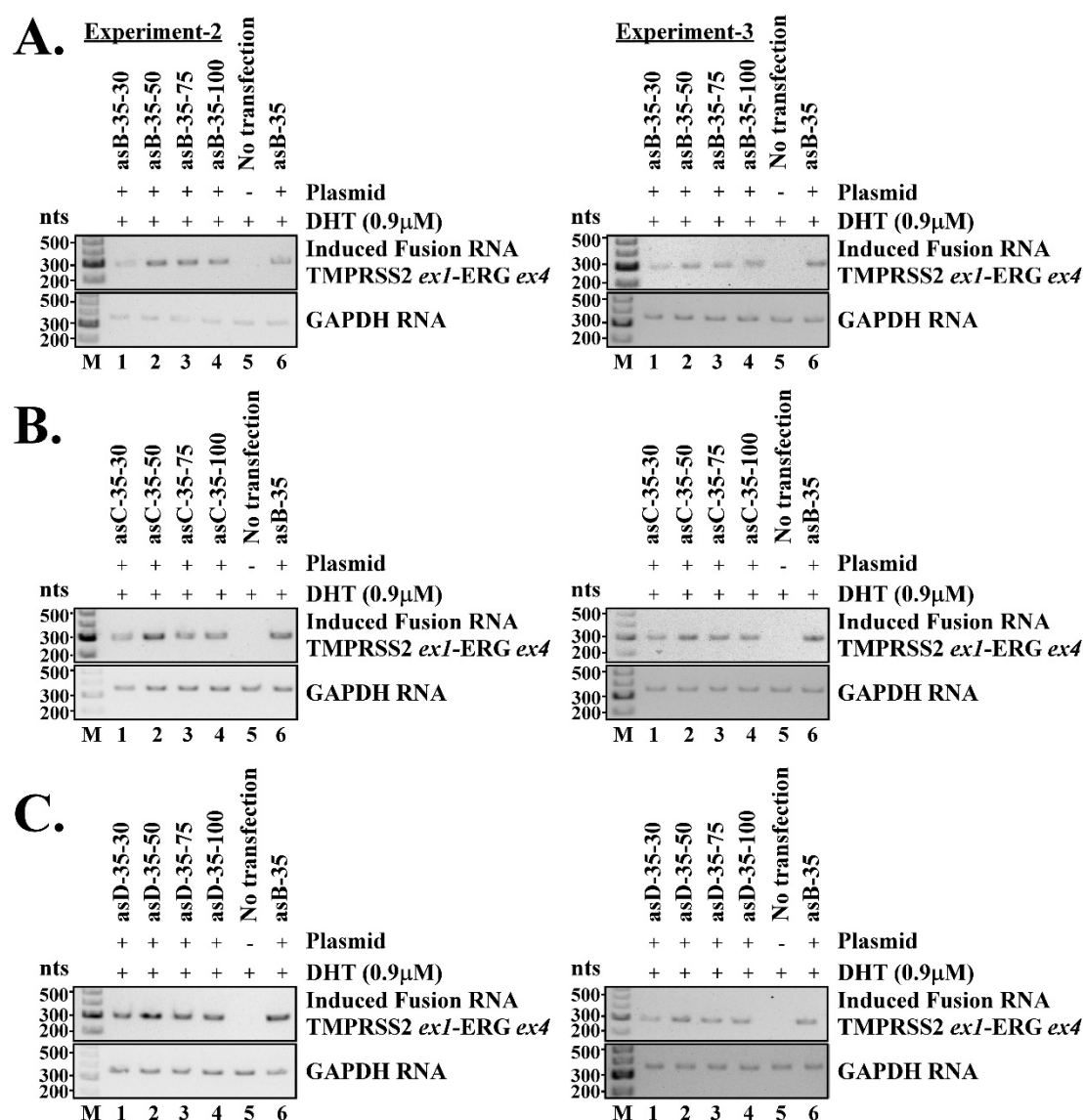
The RT-PCR gels shown in Figure 1 in the main text were obtained from a single experiment. Here we show the RT-PCR gels from additional two independent experiments. **A.** RT-PCT results of induced TMPRSS2-ERG transcripts by chimeric RNAs designed to target location B. **B.** RT-PCT results by chimeric RNAs designed to target location C. **C.** RT-PCT results by chimeric RNAs designed to target location D. GAPDH RNA was used as loading control. No transfection was used as the negative control for RT-PCR reactions. Left panel: Experiment-2, Right panel: Experiment-3.

A.	Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region
asB-2	85	75
asB-16	86	74
asB-20	87	73
asB-25	88	73
asB-31	88	74
asB-35	86	75
asB-40	87	76
asB-45	87	76
asB-50	87	76
asB-100	79	75

B.	Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region
asC-2	81	73
asC-16	83	75
asC-20	84	74
asC-25	84	73
asC-31	83	73
asC-35	83	74
asC-40	84	75
asC-45	82	73
asC-50	80	74
asC-100	78	75

C.	Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region
asD-2	88	72
asD-16	86	72
asD-20	87	74
asD-25	86	74
asD-31	85	73
asD-35	85	73
asD-40	83	72
asD-45	82	73
asD-50	84	72
asD-100	86	75

Supplementary Figure S2. The melting temperature (Tm) of chimeric RNA designed to create different bulge sizes. Tm was calculated using formula: $Tm = 64.9 + 41 * (nG + nC - 16.4) / (nA + nT + nG + nC)$ where “n” stands for total number of particular nucleotide. The Tm calculation can be found at this link: <http://insilico.ehu.es/tm.php?formula=basic>.



The RT-PCR gels shown in Figure 2 in the main text were obtained from a single experiment. Here we show the RT-PCR gels from additional two independent experiments. **A.** RT-PCT results of induced TMPRSS2-ERG transcripts by chimeric RNAs designed to target location B. **B.** RT-PCT results by chimeric RNAs designed to target location C. **C.** RT-PCT results by chimeric RNAs designed to target location D. GAPDH RNA was used as loading control. No transfection was used as the negative control for RT-PCR reactions. Left panel: Experiment-2, Right panel: Experiment-3.

A.			B.			C.		
	Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region		Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region		Tm (°C) for targeting TMPRSS2 genomic region	Tm (°C) for targeting ERG genomic region
asB-35-30/30	77	65	asC-35-30/30	79	63	asD-35-30/30	75	61
asB-35-50/50	86	69	asC-35-50/50	83	72	asD-35-50/50	84	69
asB-35-75/75	88	75	asC-35-75/75	85	74	asD-35-75/75	88	73
asB-35-100/100	91	77	asC-35-100/100	86	76	asD-35-100/100	91	75

Supplementary Figure S4. The melting temperature (Tm) of chimeric RNA with different target lengths. Tm was calculated using formula: $Tm = 64.9 + 41 * (nG + nC - 16.4) / (nA + nT + nG + nC)$ where “n” stands for total number of particular nucleotide. The Tm calculation can be found at this link: <http://insilico.ehu.es/tm.php?formula=basic>.

Location B:

TMPRSS2

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

■: targeted ERG sequence, ■: bulge, ■: stem

TMPR.SS2

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG

■: targeted ERG sequence, ■: bulge, ■: stem ■

TMPRSS2

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Ccctcccagctcacgaggggtggccccacttctatgcttttccttgcagtttctccatctggaaatgctccccaccattttgattt
 acccagtaacccccattcttcaaggctcaggccagtgacaccagttcctttaatccctccttccttcttctgtcatccagtcgaca
 cacttagcatga

■: targeted ERG sequence, ■: bulge, ■: stem

B-25

TMPRSS2

Tgagattaaagcgagagccagggcgggccgggcccagtaggcgcgagctaagcaggaggcgaggcgaggcggaagggcgaggggagc
 gggagcgccgcctggagcgccgcaggtgagcggcgccgggtaccaggggtcccggctcggggtccgggctggggagggg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Ccctcccagctcacgaggggtggccccacttctatgcttttccttgcagtttctccatctggaaatgctccccaccattttgattt
 acccagtaacccccattcttcaaggctcaggccagtgacaccagttcctttaatccctccttccttcttctgtcatccagtcgaca
 cacttagcatga

■: targeted ERG sequence, ■: bulge, ■: stem

B-31

TMPRSS2

Tgagattaaagcgagagccagggcgggccgggcccagtaggcgcgagctaagcaggaggcgaggcgaggcggaagggcgaggggaggggagc
 gggagcgccgcctggagcgccgcaggtgagcggcgccgggtaccaggggtcccggctcggggtccgggctggggagggg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Ccctcccagctcacgaggggtggccccacttctatgcttttccttgcagtttctccatctggaaatgctccccaccattttgattt
 acccagtaacccccattcttcaaggctcaggccagtgacaccagttcctttaatccctccttccttcttctgtcatccagtcgaca
 cacttagcatga

■: targeted ERG sequence, ■: bulge, ■: stem

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Ccctccccagc t c a c g a g g g t g g c c c c a c t t c t a t g c t t t t c c t t g c a g t t t t c t c a t c t g g a a a t g t c c c c a c c a t t t t g a t t t
a c c c a g t a c a c c c a t t c t t c a a g g c t c a g g c a g t g c a c a c c a g t t c c t t t a a t c c c t c c t t c c t t c t t g t c a t c c a g t c g a c a
c a c t t a g c a t g a

■: targeted ERG sequence, ■: bulge, ■: stem

B-50

TMPRSS2

[illegible]

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Ccctccccagcctcacgaggggtggccccacttctatgcttttcccttgagtttctccatctggaaatgctccccaccattttgattt
accagtacccccattcttcaaggctcaggccagtgcacaccagttctttaatccctccttccttcttctgtcatccagtcgaca
cacttagcatga

■: targeted ERG sequence, ■: bulge, ■: stem

B-100

TMPRSS2

Tgagattaaagcgagagccagggcgggcccgggagtaggcgcgagctaagcaggaggcggaggcggaggcggaggcgcaggggcccggagcgccgccttgagcgcgccaggtgagcggcgccggtaccagggctccgggctcgggggtccgggctgggggaggg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG





Ccctccccagc t c a c g a g g g t g g c c c c a c t t c t a t g c t t t t c c t t g c a g t t t c t c c a t c t g g a a t g c t c c c c a c c a t t t t g a t t t
a c c c a g t a c a c c c a t t c t t c a a g g c t c a g g c a g t g c a c a c c a g t t c c t t a a t c c c t c c t t c c t t c t t g t c a t c c a g t c g a c a
c a c t t a g c a t g a

■: targeted ERG sequence, ■: bulge, ■: stem, ■: 3' UTR

Location C:**C-2**





TMPRSS2

Catcgcgaggttggggcaggtggtcctgcgagtccttagccagttggtggaagagagtagccccggggtccccaagctggctcctag
tccgcctgccctccacggccccgcctgggagcaccgggtgcgccttttctctttggggaggaggaactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG





Cagcactcccagagtcataaaaaatcaagtctcactggcaaccaacttgtagtgcataaaaagtgttaccaggacggggtctacca
gtgctcatctatgtcataggaagcatgagctcagcgtataacattagcaaagggtctgatgacctggctacttagtcccccttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-16





TMPRSS2

Catcgcgaggttggggcaggtggtcctgcgagtccttagccagttggtggaagagagtagccccggggtccccaagctggctccctag
tccgcctgccctccacggccccgcctgggagcaccgggtgcgccttttctctttggggaggaggaactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG





Cagcactcccagagtcataaaaaatcaagtctcactggcaaccaacttgtagtgcataaaaagtgttaccaggacggggtctacca
gtgctcatctatgtcataggaagcatgagctcagcgtataacattagcaaagggtctgatgacctggctacttagtcccccttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-20





TMPRSS2

Catcgcgaggttggggcaggtggtcctgcgagtccttagccagttggtggaagagagtagccccggggtccccaagctggctcctag
tccgcctgccctccacggccccgcctgggagcaccgggtgcgccttttctctttggggaggaggaactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 


ERG





Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttggttagtgcatgaaaagtgttaccaggacggggtctaccca
gtgctcatctatgtcataggaaaagcatgagctcagcgtataacattagcaaagggctctgatgacctggctacttagtcccctttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-25





TMPRSS2

Catcgcgaggttggggcaggtggtcctgcgagtcacctagccagttggtggaagagagatcccccggggtcccaaa
tccgctgcctccacggcccccgctgggagcaccggtgcgccttttctctttggggaggaggactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 


ERG





Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttggttagtgcatgaaaagtgttaccaggacggggtctaccca
gtgctcatctatgtcataggaaaagcatgagctcagcgtataacattagcaaagggctctgatgacctggctacttagtcccctttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-31





TMPRSS2

Catcgcgaggttggggcaggtggtcctgcgagtcacctagccagttggtggaagagagatcccccggggtc
tccgctgcctccacggcccccgctgggagcaccggtgcgccttttctctttggggaggaggactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG





Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttggttagtgcatgaaaagtgttaccaggacggggtctaccca
gtgctcatctatgtcataggaaaagcatgagctcagcgtataacattagcaaagggctctgatgacctggctacttagtcccctttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-35





TMPRSS2

Catcgcgaggttggggcaggtgggtcctgcgagtccttagccagttggtggaagagagtccccggggtccccaaagctggctcctag
tccgcctgccctccacggcccccgctgggagcaccgggtgcgccttttctctttggggaggaggactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG





Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttggttagtgcatagaaaagtgttaccaggacggggtctaccca
gtgctcatctatgtcatagggaaaagcatgagctcagcgataacattagcaaagggtctgatgacctggctacttagtcccccttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-40





TMPRSS2

Catcgcgaggttggggcaggtgggtcctgcgagtccttagccagttggtggaagagagtccccgggggtccccaaagctggctcctag
tccgcctgccctccacggcccccgctgggagcaccgggtgcgccttttctctttggggaggaggactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG





Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttggttagtgcatagaaaagtgttaccaggacggggtctaccca
gtgctcatctatgtcatagggaaaagcatgagctcagcgataaacattagcaaagggtctgatgacctggctacttagtcccccttc
catgcagcaacca

: targeted ERG sequence, : bulge, : stem 

C-45

TMPRSS2

Catcgcgaggttggggcaggtgggtcctgcgagtccttagccagttggtggaagagagtcccccgggggtccccaaagctggctcctag
tccgcctgccctccacggcccccgctgggagcaccgggtgcgccttttctctttggggaggaggactgggagtgctg

: targeted TMPRSS2 sequence, : bulge, : stem 

ERG

Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttgtagtgcataaaaagtgttaccaggacggggtctaccca
 gtgctcatctatgtcatagggaaaagcatgagctcagcgtataacaattagcaaagggctctgatgacctggctacttagtcccccttc
 catgcagcaacca

■: targeted ERG sequence, ■: bulge, ■: stem ■

C-50

TMPRSS2

Catcgcgagggtggggcaggtgggtcctgcgagtccttagccagttggtagagagagtcccccggggtccccaaagctgggtcctag
 tccgctgcctccacggcccccgctgggagcaccgggtgcgccttttctctttggggaggaggaatgggagtgctg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG

Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttgtagtgcataaaaagtgttaccaggacggggtctaccca
 gtgctcatctatgtcatagggaaaagcatgagctcagcgtataaacattagcaaaagggctctgatgacctggctacttagtcccccttc
 catgcagcaacca

■: targeted ERG sequence, ■: bulge, ■: stem ■

C-100

TMPRSS2

Catcgcgagggtggggcaggtgggtcctgcgagtccttagccagttggtagagagagtcccccggggtccccaaagctgggtcctag
 tccgctgcctccacggcccccgctgggagcaccgggtgcgccttttctctttggggaggaggaatgggagtgctg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG

Cagcactcccagagtcataaaaatcaagtctcactggcaaccaacttgtagtgcataaaaagtgttaccaggacggggtctaccca
 gtgctcatctatgtcatagggaaaagcatgagctcagcgtataaacattagcaaaagggctctgatgacctggctacttagtcccccttc
 catgcagcaacca

■: targeted ERG sequence, ■: bulge, ■: stem ■

Location D:**D-2**

TMPRSS2

Gctggggaggggaacctggggcgctgggacccgccgatgccccctgccccgccggaggtgaaagcgggtgtgaggagcgcgcg
gcaggtgagtgcgccccaggggtcgagcgctggggccagccgggcagggcgctcccgggggtgctgggagagtgcctggg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Cccagcactctccgtttaccatccagatcctttattccatccagatccagggtctggacacttatcccttcacatctttcttctgt
ctcttccctctaccagcctctctctctctgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccacc
gtgtgacttttctt

■: targeted ERG sequence, ■: bulge, ■: stem

D-16

TMPRSS2

Gctggggaggggaacctggggcgctgggacccgccgatgccccctgccccgccggaggtgaaagcgggtgtgaggagcgcgcg
gcaggtgagtgcgccccaggggtcgagcgctggggccagccgggcagggcgctcccgggggtgctgggagagtgcctggg

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Cccagcactctccgtttaccatccagatcctttattccatccagatccagggtctggacacttatcccttcacatctttcttctgt
ctcttccctctaccagcctctctctctctgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccacc
gtgtgacttttctt

■: targeted ERG sequence, ■: bulge, ■: stem

D-20

TMPRSS2

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

■: targeted ERG sequence, ■: bulge, ■: stem ■

D-25

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

■: targeted ERG sequence, ■: bulge, ■: stem, ■

D-31

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG

Cccagcactctccggtttaccatccagatcctttattccatcccagatccagggttctggacacttatcccttcattctttcttctgtctcttccctctaccagcctctctctctctgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccaccgtgtgacttttctt

■: targeted ERG sequence, ■: bulge, ■: stem ■

D-35

TPRSS2

Gctggggaggggaacctgggagcctgggacccgccgatgccccctgccccgcccggaggtgaaagcgggtgtgaggagcgcgggcgcggcaggtgagtgcgccccgagggtcgagcgctggggccagccgggcagggcgccctcccggggtgctgggagagtgtctggg

■: targeted TPRS2 sequence, ■: bulge, ■: stem ■

ERG

Cccagcactctccggtttaccatccagatcctttattccatcccagatccagggttctggacacttatcccttcattctttcttctgtctcttccctctaccagcctctctctctctgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccaccgtgtgacttttctt

■: targeted ERG sequence, ■: bulge, ■: stem ■

D-40

TPRSS2

Gctggggaggggaacctgggagcctgggacccgccgatgccccctgccccgcccggaggtgaaagcgggtgtgaggagcgcgggcgcggcaggtgagtgcgccccgagggtcgagcgctggggccagccgggcagggcgccctcccggggtgctgggagagtgtctggg

■: targeted TPRS2 sequence, ■: bulge, ■: stem ■

ERG

Cccagcactctccggtttaccatccagatcctttattccatcccagatccagggtattctggacacttatcccttcattctttcttctgtctcttccctctaccagcctctctctctctgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccaccgtgtgacttttctt

■: targeted ERG sequence, ■: bulge, ■: stem ■

D-45

TMPRSS2

Gctggggaggggaacctggg'gcctggg'acccgccgatgccccctgccccgcgcg'gaggtgaaagcgggtgtgaggagcgcggcgcg'gaggtgagtgcgcccaggg'gtcgagcgcctggggccagccgggcagggcgcctcccggggtgctg'ggagagtgcctggg'

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Cccagcactctcc'gtttaccatccagatcctttattccatcccagatccaggattct'ggacacttatcccttcattctttcttctgt'ctcttccctctaccagcctctctctctctgtctctctctaggtgat'tccatctctcatggcaagagcaggcaccaaaacactccacc'gtgtgacttttctt'

■: targeted ERG sequence, ■: bulge, ■: stem

D-50

TMPRSS2

Gctggggaggggaacctggg'gcctggg'acccgccgatgccccctgccccgcgcg'cccggaggtgaaagcgggtgtgaggagcgcggcgcg'gaggtgagtgcgccc'cgagggtcgagcgcctggggccagccgggcagggcgcctcccggggtgctg'ggagagtgcctggg'

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem

ERG

Cccagcactctcc'gtttaccatccagatcctttattccatcccagatccaggattctggaca'cttatcccttcattctttcttctgt'ctcttccctctaccagcctctctctctctgtctctctctaggtgattccat'ctctcatggcaagagcaggcaccaaaacactccacc'gtgtgacttttctt'

■: targeted ERG sequence, ■: bulge, ■: stem

D-100

TMPRSS2

Gctggggaggggaacctggg'gcctggg'acccgccgatgccccctgccccgcgcg'ccggaggtgaaagcgggtgtgaggagcgcggcgcg'gaggtgagtgcgcccaggggtcgagcgcctggggccagccgggcagggcgcctcccggggtgctg'ggagagtgcctggg'

■: targeted TMPRSS2 sequence, ■: bulge, ■: stem ■

ERG

Cccagcactctccgtttaccatccagatcctttattccatcccagatccagggattctggacacttatcccttcacatctttcttctgt
ctcttccctctaccagcctctctctctctctgtgtctctctctaggtgattccatctctcatggcaagagcaggcaccaaaacactccacc
gtgtgactttttctt

■: targeted ERG sequence, ■: bulge, ■: stem ■

Supplementary file S2: Chimeric RNA sequences

asB-2

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTAAATCAAAATGGTGGGAGCATGTCCAGATGGAGAAA
CTGCAAGGAAAAGCATAGAAGTGGGGCCACCCCTCGTGGACCCCGAGCCGGGACCCCTGGTACCGGCGCCGCTCACCTGCCGCGCTCC
AG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-16

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGGGGTGTACTGGGTAAATCAAAATGGTGGGGAGCATT
TCCAGATGGAGAACTGCAAGGAAAAGCATAGAAGTGGACCCCTGGTACCGGCGCCGCTCACCTGCCGCGCTCCAGGCGGCGCTCCCC
GC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-20

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAGAATGGGGTGTACTGGGTAAATCAAAATGGTGGGG
AGCATTTCAGATGGAGAACTGCAAGGAAAAGCATAGACTGGTACCGGCGCCGCTCACCTGCCGCGCTCCAGGCGGCGCTCC
CCGCCCCCT

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-25

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCTTGAAGAATGGGGTGTACTGGGTAAATCAAAATGGTG
GGGAGCATTTCAGATGGAGAACTGCAAGGAAAAGCACCGGCGCCGCTCACCTGCCGCGCTCCAGGCGGCGCTCCCCGCCCTCGC
CC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-31

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCTGAGCCTTGAAGAATGGGGTGTACTGGGTAAATCA
AAATGGTGGGGAGCATTTCAGATGGAGAACTGCAAGGCGCGCTCACCTGCCGCGCTCCAGGCGGCGCTCCCCGCCCTCGC
CCTCCGCC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-35

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGGCCTGAGCCTTGAAGAATGGGGTGTACTGGGTAA
ATCAAAATGGTGGGGAGCATTTCAGATGGAGAACTGCCTCAGCTGCGCGCTCCAGCGCGGCTCGCGCGCTCGCGCTG
CGCCTCCG

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-40

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGCACTGGCCTGAGCCTTGAAGAATGGGGTGTACTGGG
TAAATCAAAATGGTGGGGAGCATTTCAGATGGAGAACTGCCGCGCTCCAGGCGGCGCTCCCCGCCCTCGCCCTCCGCCTCCGCCT
CC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-45

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGGTGTGCACTGGCCTGAGCCTTGAAGAATGGGGTGT
CTGGGTAAATCAAAATGGTGGGGAGCATTTCAGATGGCGCTCCAGGCGGCGCTCCCCGCCCTCGCCCTCCGCCTCCGCCTCCGCC
TC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-50

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGGAAC TGGTGTGCACTGGCCTGAGCCTTGAAGAATGGG
GTGTACTGGGTAAATCAAAATGGTGGGGAGCATTTCCTCAGGCGGCGCTCCCCGCCCTCGCCCTCCGCCTCCGCCTCCGCCTCCTG
CT

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTCATGCTAAGTGTGTCGACTGGATGACAGAAGAAGGAA
GGAGGGATTAAAGGAACTGGTGTGCACTGGCCTGAGCCTTAGCTCGCGCCTACTCGGCCCGGCCCGCCCTGGCTCTCGCTTTAATCT
CA

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asC-2

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagACTGGGTAGACCCCGTCCTGGTAACACTTCTCATGCAC
TAACAAGTTGGTTGCCAGTGAGACTTGATTCTTATGACTCCTCCCCAAAGAGAAAAGGCGCACCCGGTGCTCCCAGGCGGGGGCCGTG
GA

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asC-16

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGACATAGATGAGCACTGGGTAGACCCCGTCCTGGT
AAACTCTTCATGCACTAACAAGTTGGTTGCCAGTGAGAGAAAAGGCGCACCCGGTGCTCCCAGGCGGGGGCCGTGGAGGGCAG
GCGGACTA

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI


asC-20

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCCTATGACATAGATGAGCACTGGGTAGACCCCGTCC
TGGTAACACTCTTCATGCACTAACAAGTTGGTTGCCAGTAGGCGGAGGGGTCTCCCAGGCGGGGGCCGTGGAGGGCAGGCGG
ACTAGGAG

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asC-25

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCTCTTCCTATGACATAGATGAGCACTGGGTAGACCCCG
TCCTGGTAACACTTCTCATGCACTAACAAGTTGGTTGCACCGGTGCTCCCAGGCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCA
GC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-31

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCTCATGCTCTTCCTATGACATAGATGAGCACTGGGT
AGACCCCGTCCTGGTAACACTCTTCATGCACTAACAAGTTGCTCCCAAGCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCCA
GCTTTGGG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-35

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGAGCTCATGCTCTTCCTATGACATAGATGAGCACT
GGGTAGACCCCGTCCTGGTAACACTCTTCATGCACTAACCCAGGCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCAGCTT
TGGCGAGG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-40

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTACGCTGAGCTCATGCTCTTCCTATGACATAGATGA
GCACTGGGTAGACCCCGTCCTGGTAACACTCTTCATGCACCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCAGCTTTGGGG
ACCCCGGG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-45

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGTTATACGCTGAGCTCATGCTCTTCCTATGACATA
GATGAGCACTGGGTAGACCCCGTCCTGGTAACACTCTTCGGCCGTGGAGGGCAGGCGGACTAGGAGCCAGCTTTGGGGACCCG
GGGGGACT

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-50

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGCTAATGTTATACGCTGAGCTCATGCTGTTCCCTATGAC
 ATAGATGAGCACTGGGTAGACCCCGTCTGGTAACACTTGGAGGGCAGGCGGACTAGGAGCCAGCTTTGGGGACCCCGGGGACTCTC
 TT

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asC-100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGGTTGCTGCATGGAAAGGGGACTAAGTAGCCAGGTCA
 TCAGAGCCCTTTGCTAATGTTATACGCTGAGCTCATGTTCACCAACTGGCTAGGGACTCGCAGGACCACCTGCCCAACCTCGCGA
 TG

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asD-2

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGAGACAGAAGAAAGATGAAGGGATAAGTGTCCAGAA
 TCCCTGGATCTGGGATGGAATAAAGGATCTGGATGGTAAGCACCGCGGAGGCGCCCTGCCCGGCTGCGCCAGCGCTCGACC
 CTCGGGCG

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asD-16

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGCTGGTAGAGGGAAGAGACAGAAGAAAGATGAAGGGAT
 AAGTGTCCAGAATCCCTGGATCTGGGATGGAATAAAGGCCCTGCCCGGCTGGCCCCAGCGCTCGACCCCTCGGGCGCACTCACCTGCC
 GC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asD-20

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAGAGGCTGGTAGAGGGAAGAGACAGAAGAAAGATGA
AGGGATAAGTGTCCAGAATCCCTGGATCTGGGATGGAATTGCCCCGCTGGCCCCAGCGCTCGACCCCTCGGGCGCACTCACCTG
CCGGGGGG

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asD-25

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGAGAGAGAGGCTGGTAGAGGGAAGAGACAGAAGAAA
GATGAAGGGATAAGTGTCCAGAATCCCTGGATCTGGGATGGCTGGCCCCAGCGCTCGACCCCTCGGGCGCACTCACCTGCCGCG
CCGCGCTG

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asD-31

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGACAGAGAGAGAGAGGCTGGTAGAGGGAAGAGACAG
AAGAAAGATGAAGGGATAAGTGTCCAGAATCCCTGGATCCCCAGCGCTCGACCCCTCGGGCGCACTCACCTGCCGCGCCGCGCG
TCCTCACA

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asD-35

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGAGAGACAGAGAGAGAGAGGCTGGTAGAGGGAAGAGAC
AGAAGAAAGATGAAGGGATAAGTGTCCAGAATCCCTGAGCGCTCGACCCCTCGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACC
CG

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI


asD-40

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCTAGAGAGAGACAGAGAGAGAGAGGCTGGTAGAGGGAA
GAGACAGAAGAAAGATGAAGGGATAAGTGTCCAGAATTGACCCCTCGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACCCGCTT
TC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-45

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagATCACCTAGAGAGAGACAGAGAGAGAGAGGGCTGGTAGA
GGGAAGAGACAGAAGAAAGATGAAGGGATAAGTGTCCCTCGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACCCGCTTTCACC
TC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-50

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagATGGAATCACCTAGAGAGAGACAGAGAGAGAGAGGGCTG
GTAGAGGGAAGAGACAGAAGAAAGATGAAGGGATAAGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACCCGCTTTCACCTCCGG
GC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAAGAAAAGTCACACGGTGGAGTGTGTTGGTGCCTGCTC
TTGCCATGAGAGATGGAATCACCTAGAGAGAGACAGAGCGGGGCAGGGGGCATCGGCGGGTCCCAGGCGCCAGGTTCCTCCCTCCCCA
GC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-35-100/100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAGGGATTAAAGGAACTGGTGTGCACTGGCCTGAGCCTT
GAAGAATGGGGTGTACTGGGTAAATCAAAATGGTGGGGAGCATTTCCAGATGGAGAAAAGTGCCTCACCTGCCGCGCTCCAGGCGGCG
CTCCCCGCCCCCTCGCCCTCCGCCTCCGCCTCCGCCTCCTGCTTAGCTCGCGCCTACTCGGCCCCGGCCCCGCCCCCTGG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asB-35-75/75

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGGCCTGAGCCTTGAAGAATGGGGTGTACTGGGTAAAT
 CAAAATGGTGGGGAGCATTTCAGATGGAGAACTGCCTCACCTGCCGCGCTCCAGGCGGCGCTCCCCGCCCCCTCGCCCTCCGCCTC
 CGCCTCCGCCTCCTGCTTAGCTCGC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-35-50/50

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGTACTGGGTAAATCAAAAATGGTGGGGAGCATTTCCTCA
 GATGGAGAAACTGCCTCACCTGCCGCGCTCCAGGCGGCGCTCCCCGCCCCCTCGCCCTCCGCCTC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asB-35-30/30

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGTGGGGAGCATTTCAGATGGAGAACTGCCTCACCTG
 CCGCGCTCCAGGCGGCGCTCCC

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asC-35-100/100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAGAGCCCTTTGCTAATGTTATACGCTGAGCTCATGCTC
 TTCCTATGACATAGATGAGCACTGGGTAGACCCCGTCCTGGTAACACTCTTCATGCACTAACCCCAGGCGGGGGCCGTGGAGGGCAG
 GCGGACTAGGAGCCAGCTTTGGGGACCCCGGGGACTCTCTTCCACCAACTGGCTAGGGACTCGCAGGACCACCT

: TMPRSS2, : ERG, +1: Transcription start, ctgcag: PstI

asC-35-75/75

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagTGAGCTCATGCTCTTCCTATGACATAGATGAGCACTGG
 GTAGACCCCGTCCTGGTAACACTCTTCATGCACTAACCCCAGGCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCAGCTTTGGGGGA
 CCCCAGGGGACTCTCTTCCACCAAC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-35-50/50

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagAGATGAGCACTGGGTAGACCCCGTCCTGGTAACACTCT
TCATGCACTAACCCCAGGCGGGGGCCGTGGAGGGCAGGCGGACTAGGAGCCAGCTTTGGGGA

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asC-35-30/30

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagCCGTCCTGGTAACACTCTTCATGCACTAACCCCAGGCG
GGGGCCGTGGAGGGCAGGCGGA

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-35-100/100

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGCCATGAGAGATGGAATCACCTAGAGAGAGACAGAGAG
AGAGAGGCTGGTAGAGGGAAGAGACAGAAGAAAGATGAAGGGATAAGTGTCCAGAATCCCTGAGCGCTCGACCCTCGGGCGCACTCA
CCTGCCGCGCCGCGCTCCTCACACCCGCTTTCACCTCCGGGCGGGCAGGGGGCATCGGCGGGTCCCAGGCGCCC

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-35-75/75

+1gtgctcgcttcggcagcacatataactaacattggaacgatcctgcagGAGAGACAGAGAGAGAGAGGCTGGTAGAGGGAAGAGAC
AGAAGAAAGATGAAGGGATAAGTGTCCAGAATCCCTGAGCGCTCGACCCTCGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACC
CGCTTTCACCTCCGGGCGGGGAGG

 : *TMPRSS2*,  : *ERG*, +1: Transcription start, ctgcag: PstI

asD-35-50/50

+1gtgctcgcttcgggcagcacatataactaacattggaacgatcctgcagAGAGGGAAGAGACAGAAGAAAGATGAAGGGATAAGTGT
CCAGAATCCCTGAGCGCTCGACCCTCGGGCGCACTCACCTGCCGCGCCGCGCTCCTCACACC

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI

asD-35-30/30

+1gtgctcgcttcgggcagcacatataactaacattggaacgatcctgcagAGATGAAGGGATAAGTGTCCAGAATCCCTGAGCGCTCG
ACCCTCGGGCGCACTCACCTGC

: *TMPRSS2*, : *ERG*, +1: Transcription start, ctgcag: PstI