

# pUNO1-SpikeV5

Expression vector encoding the SARS-CoV-2 Brazilian variant (P.1 lineage) Spike gene

Catalog code: p1-spike-v5

<https://www.invivogen.com/brazil-p1-spike-expression-vectors>

For research use only

Version 21E07-ED

## PRODUCT INFORMATION

### Contents

- 20 µg of lyophilized pUNO1-SpikeV5 (plasmid DNA)
- 2 x 1 ml of **Blasticidin** (10 mg/ml)

### Storage and Stability

- Product is shipped at room temperature.
- Lyophilized DNA should be stored at -20°C.
- Resuspended DNA is stable for 1 year at -20°C.
- Store Blasticidin at 4°C or -20°C. The expiry date is specified on the product label.

### Quality control

- Plasmid construct is confirmed by restriction analysis and full-length open reading frame (ORF) sequencing.
- After purification by ion exchange chromatography, predominant supercoiled conformation is verified by electrophoresis.

## PLASMID FEATURES

### Brazilian Variant SARS-CoV-2 Spike cassette

• **EF-1 $\alpha$ /HTLV hybrid promoter** is a composite promoter comprised of the Elongation Factor-1 $\alpha$  (EF-1 $\alpha$ ) core promoter<sup>1</sup> and the 5' untranslated region of the Human T-Cell Leukemia Virus (HTLV). EF-1 $\alpha$  utilizes a type 2 promoter that encodes a "house-keeping" gene. It is expressed at high levels in all cell cycles and lower levels during the G0 phase. Additionally, since the promoter is not tissue-specific it is highly expressed in all cell types. The R segment and part of the U5 sequence (R-U5') of the HTLV Type 1 Long Terminal Repeat<sup>2</sup> has been coupled to the EF-1 $\alpha$  promoter to enhance stability of DNA and RNA. This modification not only increases steady state transcription, but also significantly increases translation efficiency.

- **Codon-optimized Spike ORF**

pUNO1-SpikeV5 contains the Spike coding sequence from the Brazilian (BRA.) SARS-CoV-2 variant (P.1 lineage). This variant is characterized by a number of mutations within the the Spike coding sequence (see below)<sup>3</sup>. Additionally, to improve expression of the S protein in cell lines, the gene is codon-optimized and the last 19 amino acids, which contain an endoplasmic reticulum (ER)-retention motif (KxHxx), have been removed<sup>4,5</sup>.

pUNO1-SpikeV5 includes the following sequence features:

- **S1 domain:** L18F, T20N, P26S, D138Y, R190S, D614G, H655Y
- **RBD:** K417T, E484K, N501Y
- **S1/S2 boundary:** Functional furin cleavage site
- **S2 domain:** T1072I, V1176F

Spike (S) is a structural glycoprotein expressed on the surface of SARS-CoV-2. It mediates membrane fusion and viral entry into target cells upon binding to the host receptor ACE2 and the proteolytic activity of host proteases such as furin and TMPRSS2<sup>6</sup>.

For more information visit: <https://www.invivogen.com/sars2-spike>

- **SV40 pAn** is the Simian Virus 40 late polyadenylation (pAn) signal and it enables efficient cleavage and polyadenylation reactions resulting in high levels of steady-state mRNA<sup>7</sup>.

### Antibiotic selection cassette

- **hCMV (human cytomegalovirus) enhancer & promoter** drive the expression of the blasticidin resistance gene (*bsr*) in mammalian cells.
- **EM7** is a bacterial promoter that enables the constitutive expression of the blasticidin resistance gene (*bsr*) in *E. coli*.
- ***bsr* (blasticidin resistance gene)** encodes a deaminase from *Bacillus cereus* that confers resistance to the antibiotic blasticidin. The expression of the *bsr* gene is driven by the CMV promoter/enhancer and the bacterial EM7 promoter. Therefore, **Blasticidin** can be used to select stable clones in mammalian cells and *E. coli* transformants.
- **Human  $\beta$ -Globin pAn** is a strong polyadenylation (pAn) signal placed downstream of *bsr*. The use of  $\beta$ -globin pAn minimizes interference and possible recombination events with the SV40 pAn signal<sup>8</sup>.

### General features of pUNO1-SpikeV5

- **pMB1 ori** is a minimal *E. coli* origin of replication.

## APPLICATIONS

### Stable gene expression in mammalian cells.

pUNO1 plasmids are designed for both transient and stable transfection in mammalian cell lines by selection with **Blasticidin**. Furthermore, they facilitate high levels of expression of the gene of interest.

### Cell fusion assays

pUNO1-SpikeV5 has been specifically designed for mammalian cell expression of the SARS-CoV-2 S protein. This plasmid features a functional furin cleavage site, and therefore when expressed by a host cell (**293-hMyD88 cells**) it facilitates cell-cell fusion assays with ACE2-expressing cells (**HEK-Blue™ hACE2 Cells**)<sup>9</sup>.

## METHODS

### Plasmid resuspension

- Quickly spin the tube containing the lyophilized plasmid to pellet the DNA.
- To obtain a plasmid solution at 1 µg/µl, resuspend the DNA in 20 µl of sterile water.
- Store the resuspended plasmid at -20°C.

### Plasmid amplification and cloning

Plasmid amplification and cloning can be performed in *E. coli* **GT116** or other commonly used laboratory *E. coli* strains, such as DH5 $\alpha$ .

### Blasticidin usage

Blasticidin should be used at 25-100 µg/ml in bacteria and 1-30 µg/ml in mammalian cells. Blasticidin is supplied as a 10 mg/ml colorless solution in HEPES buffer.

## TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873

InvivoGen USA (International): +1 (858) 457-5873

InvivoGen Europe: +33 (0) 5-62-71-69-39

InvivoGen Hong Kong: +852 3622-3480

E-mail: [info@invivogen.com](mailto:info@invivogen.com)

## REFERENCES

1. Kim D. *et al.*, 1990. Use of the human elongation factor 1 $\alpha$  promoter as a versatile and efficient expression system. *Gene* 91(2):217-23 2. Takebe Y. *et al.*, 1988. SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1 long terminal repeat. *Mol Cell Biol.* 8(1):466-72. 3. Faria, N.R. *et al.* 2021. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science.* doi:10.1126/science.abh2644. 4. Johnson, M.C. *et al.* 2020. Optimized Pseudotyping Conditions for the SARS-COV-2 Spike Glycoprotein. *J Virol* 94. 5. Ou, X. *et al.* 2020. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat Commun* 11, 1620. 6. Hoffmann M. *et al.*, 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* 181:1-16. 7. Carswell S. & Alwine J., 1989. Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol Cell Biol.* 9(10):4248-58. 8. Yu J. & Russell J., 2001. Structural and functional analysis of an mRNP complex that mediates the high stability of human  $\beta$ -globin mRNA. *Mol Cell Biol.* 21(17):5879-88. 9. Papa, G. *et al.* 2021. Furin cleavage of SARS-CoV-2 Spike promotes but is not essential for infection and cell-cell fusion. *PLoS Pathog* 17, e1009246.

## RELATED PRODUCTS

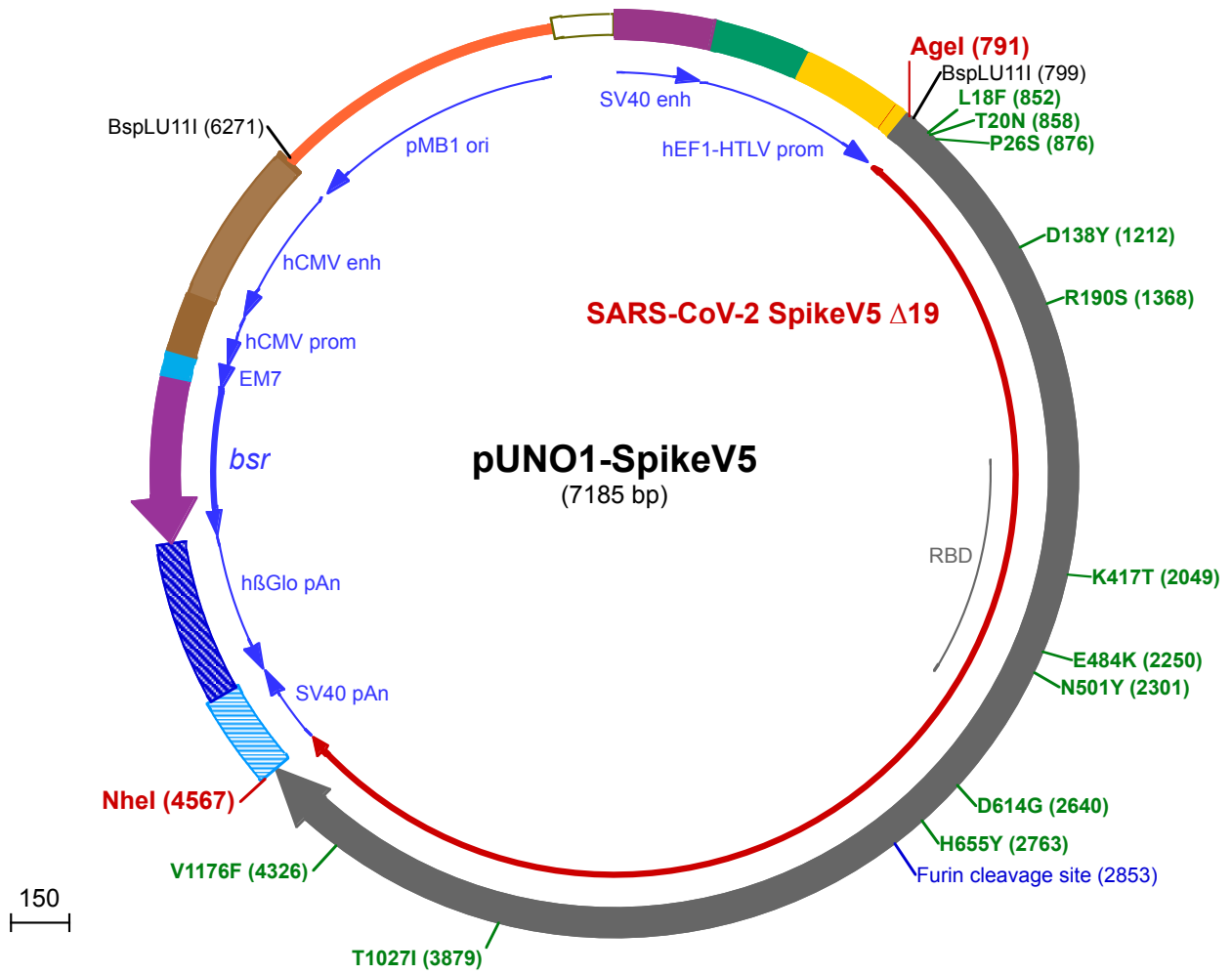
Product	Description	Cat. Code
Blasticidin	Selection antibiotic	ant-bl-1
ChemiComp GT116	Competent <i>E. coli</i>	gt116-11
<b>COVID-19 Product Range</b>		
HEK-Blue™ hACE2 Cells	Cell line	hkb-hace2
293-hMyD88 Cells	Cell line	293-hmyd
A549-hACE2-TMPRSS2 Cells	Cell Line	a549-hace2-tpsa
pUNO1-hACE2	Expression vector	puno1-hace2
pUNO1-hTMPRSS2a	Expression vector	puno1-htp2a
Anti-CoV2RBD-c1-hlgG1	Recombinant Antibody	cov2rbdc1-mab1

For a complete list of InvivoGen's COVID-19 related products visit:  
<https://www.invivogen.com/covid-19>

### TECHNICAL SUPPORT

InvivoGen USA (Toll-Free): 888-457-5873  
InvivoGen USA (International): +1 (858) 457-5873  
InvivoGen Europe: +33 (0) 5-62-71-69-39  
InvivoGen Hong Kong: +852 3622-3480  
E-mail: [info@invivogen.com](mailto:info@invivogen.com)





1 GGACCTGCAGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTAGTTAGGGTGTG  
 101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGAAAGTCCCAGGCTCCCAGCAGGCAG  
 201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCACTAGTGTCTCGGTGCCCGTCACTGGGCGAGCGCACATCGCCACAGTCCC  
 301 GAAGTTGGGGGGAGGGTTCGCAATTGAACGGGTGCCTAGAGAAGGTGGCGGGGTAACAGTGGAAAGTGTGCTGTACTGGCTCCGCCCTTTTCCC  
 401 GAGGGTGGGGGAGAACCGTATATAAGTCAGTAGTGCCTGAAACGTTCTTTTTCGAACGGGTTTGGCCGAGAACAGCTGAAGCTTCGAGGGGCTC  
 501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCGGTTGAGTCGCGTTCTGCGCCTCCCGCTGTGGTGCCTCCTGAACTGC  
 601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCCTCACGCTTGC

Agel (791)

701 CTGACCTGCTTGTCAACTCTACGTCTTTGTTTCTGTTTCTGTTCTGCGCAGTTACAGATCCAAGCTGTGACCGGCGCTACTCTGAGATCACCGGTCAA

T20N (858)

L18F (852)

P26S (876)

801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAGTCCCAATGCGTAAATTTACACCAACCGAACTCAACTCCCATCCGCATATACAAATTCCTTACC  
 1▶ M F V F L V L L P L V S S Q C V N F T N R T Q L P S A Y T N S F T  
 901 AGAGGAGTGTACTATCCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTCTTTCTGCCTTTCTTTTCTAACGTTACATGGTTTCATG  
 34▶ R G V Y Y P D K V F R S S V L H S T Q D L F L P F F S N V T W F H  
 101 CAATCCATGTGTCTGGGACAAACGGCACCAACGCTTCGACAACCCTGTATTGCCATTCAATGATGGGTGTACTTTGCCTCCACAGAGAAATCCAACAT  
 67▶ A I H V S G T N G T K R F D N P V L P F N D G V Y F A S T E K S N I  
 1101 CATTGAGGATGGATTTTCGGGACTACTCTGGACTCAAAGACACAGAGCCTGCTGATCGTTAAACAACGCCACAAACGTTGTCATCAAAGTGTGCGAATTC  
 100▶ I R G W I F G T T L D S K T Q S L L I V N N A T N V V I K V C E F

D138Y (1212)

1201 CAGTTTTGCAATTAATCCCTTCTGGGAGTGTACTATCACAAGAATAACAAGTCTGGATGGAGAGCGAATTTCCGGTCTACAGCAGCGCAAACAACCTGCA  
 134▶ Q F C N Y P F L G V Y Y H K N N K S W M E S E F R V Y S S A N N C

R190S (1368)

1301 CCTTCGAGTACGTGAGTCAACCTTTCTGATGGACCTGGAAGGAAACAGGAAACTTCAAGAACCTGAGTGTGTTTGAAGAACATCGACGGCTA  
 167▶ T F E Y V S Q P F L M D L E G K Q G N F K N L S E F V F K N I D G Y  
 1401 TTTAAGATCTATAGTAAGCATACGCCATCAACCTGGTAAGGGATCTTCCCAGGGCTTTTCCAGCCCTGGAACTTTGGTTGACTTGCCTATTGGTATC  
 200▶ F K I Y S K H T P I N L V R D L P Q G F S A L E P L V D L P I G I  
 1501 AATATCACCAGATTTACAGACCTTCTGGCATTGCAICGGTCTTACTTACTCCAGGTATTCTCTCCGGTGGACTGCCGGCGCCGCTGCCTACTATG  
 234▶ N I T R F Q T L L A L H R S Y L T P G D S S S G W T A G A A A Y Y  
 1601 TCGGCTATCTGCAACCAAGAAGTCTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTGATTGTGCCCTGGACCCTGTCTGAGACTAA  
 267▶ V G Y L Q P R T F L L K Y N E N G T I T D A V D C A L D P L S E T K  
 1701 ATGCACCCTCAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGTCCAACCCACGAAAGCATTGTGCGGTTCCCAAATATACC  
 300▶ C T L K S F T V E K G I Y Q T S N F R V Q P T E S I V R F P N I T

1801 AATCTGTGTCCCTTTGGCGAAGTGTCAATGCTACAAGGTTTGTCTTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTCGCTGATTACTCCG  
 334▶ N L C P F G E V F N A T R F A S V Y A W N R K R I S N C V A D Y S

1901 TGCTGTACAATCCGCTCTTTCTCAACCTTCAAGTGTATGGCGTTTACCTACCAAACCTAACGACCTGTGCTTACTAATGTGTATGCCGACTCTTT  
 367▶ V L Y N S A S F S T F K C Y G V S P T K L N D L C F T N V Y A D S F

K417T (2049)

2001 TGTGATACGAGGCGATGAAGTGAAGCAGATTGCACCAGGGCAGACCGGCACAATTGCCGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGTGTT  
 400▶ V I R G D E V R Q I A P G Q T G T I A D Y N Y K L P D D F T G C V

2101 ATTCATGGAACCTCAACAATCTGGATTCCAAGGTGGTGGCAACTATAACTACCTGTATAGACTGTTCCAGGAAATCCAACCTGAAACCATTCCGAGCGAG  
 434▶ I A W N S N N L D S K V G G N Y N Y L Y R L F R K S N L K P F E R

E484K (2250)

2201 ATATAAGCACAGAAATCTACCAGGCTGGAAGTACGCCCTGCAACGGCGTGAAGGGTTCAACTGCTACTTCCCATTGCAGAGTTACGGATTCCAGCCTAC  
 467▶ D I S T E I Y Q A G S T P C N G V K G F N C Y F P L Q S Y G F Q P T

N501Y (2301)

2301 ATACGGGGTGGTTACCAACCTATCGTGTGCTAGTCTGAGTTTTGAGTCTCCATGCCAGCCACAGTCTGTGGCCCAAGAAAAGCACCAATCTG  
 500▶ Y G V G Y Q P Y R V V V L S F E L L H A P A T V C G P K K S T N L

2401 GTGAAGAACAAATGCGTGAACCTTAACTTTAACGGACTCACAGGAACCGGCGTATTGACGGAGAGTAACAAGAAGTTCCTGCCATTCCAGCAGTTCGGTC  
 534▶ V K N K C V N F N F N G L T G T G V L T E S N K K F L P F Q Q F G

2501 GCGATATTGCCGACACTACCGACTGTCGAGATCCCAGACATTGGAGATTCTTGATATCACACCTGTAGTTTTCCGGCGGAGTGAGCGTGATTACGCC  
 567▶ R D I A D T T D A V R D P Q T L E I L D I T P C S F G G V S V I T P

D614G (2640)

2601 CGGAACCAATACCAGCAATCAGGTTGCCGCTGTATCAGGGCGTGAATTGCACCGAGGTACCTGTCGCCATCCACGCTGACCAACTTACCCACATGG  
600▶ G T N T S N Q V A V L Y Q G V N C T E V P V A I H A D Q L T P T W

H655Y (2763)

2701 CGAGTATATTCCACCGCTCCAACGCTTTTCAGACAGCTGCTGGATGTCTGATCGGTGCAGAAACGTTAATAATAGCTACGAGTGTGATATCCCCATCG  
634▶ R V Y S T G S N V F Q T R A G C L I G A E Y V N N S Y E C D I P I

Furin cleavage site (2853)

2801 GTGCTGGAATATGCGCCTCTTATCAAACCTCAAACCAACTCTCCTAGCGGGCAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCTCGG  
667▶ G A G I C A S Y Q T Q T N S P R R A R S V A S Q S I I A Y T M S L G

2901 TGCTGAGAATTCTGTCGCTACAGCAACAACCTCATTGCTATCCCTACTAACTTCAACAATCAGTGTGACAACGAAATTCGCCGTATCTATGACCAA  
700▶ A E N S V A Y S N N S I A I P T N F T I S V T T E I L P V S M T K

3001 ACAAGCGTTGACTGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCCTGCAATACGGATCATTCTGCACTCAGCTGAATCGTGCC  
734▶ T S V D C T M Y I C G D S T E C S N L L L Q Y G S F C T Q L N R A

3101 TCACAGGTATTGAGTGTGAGCAGGACAAGAATACGCAGGAAGTGTTCGCCAGGTGAAGCAAACTACAAAACCTCCACCCATAAAAGACTTTGGCGGATT  
767▶ L T G I A V E Q D K N T Q E V F A Q V K Q I Y K T P P I K D F G G F

3201 CAATTTCTCAGATCTGCGGATCCCTCAAACCTCAAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCTCGCAGATGCCGGTTTC  
800▶ N F S Q I L P D P S K P S K R S F I E D L L F N K V T L A D A G F

3301 ATCAAGCAGTATGGCATTGTCTGGGAGACATCGCCGCTCGGGACCTGATCTGTGCACAGAAGTTCAATGGACTGACCTGCTCCCTCCCTGCTGACCG  
834▶ I K Q Y G D C L G D I A A R D L I C A Q K F N G L T V L P P L L T

3401 ACGAGATGATAGCCAACTACTAGCGCCTGCTGGCCGGCACCATCACTTCTGGGTGGACATTGCGAGCTGGCGCTGCCCTCAGATTCTTTTGTAT  
867▶ D E M I A Q Y T S A L L A G T I T S G W T F G A G A A L Q I P F A M

3501 GCAGATGGCCTACCGTTAACGGCCTCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATCGCAACAGTTCAACAGTGTATCGGTAAG  
900▶ Q M A Y R F N G I G V T Q N V L Y E N Q K L I A N Q F N S A I G K

3601 ATACAGGATGCTGTACTCCACTGACAGCGCATTGGGAAAGTGCAGGATGTAGTGAACAGAATGCCAGGCATTAACACCCTGGTGAACAGCTCT  
934▶ I Q D A G S T S T A S A L L G K L Q D V V N Q N A Q A L N T L V K Q L

3701 CTTCAAATTTTGGTGCCATTTCTAGCGTGTGAATGACATACTGAGCCGTTGGACAAGGTGGAGGCTGAAGTGCAGATTGATAGGCTGATAACTGGGCG  
967▶ S S N F G A I S S V L N D I L S R L D K V E A E V Q I D R L I T G R

T10271 (3879)

3801 CCTTCAGTCTCTCAGACCTATGTGACCCAGCAGCTCATCCGCGCTGCTGAAATTCGCGCATCCGCTAACCTGGCAGCAATGAAAATGTCCGAGTGTGTG  
1000▶ L Q S L Q T Y V T Q Q L I R A A E I R A S A N L A A I K M S E C V

3901 CTGGGTGAGTCAAGAGAGTGGACTTTTGGCGGAAGGGGTATCACCTGATGCTTTTCTCAGTCTGCACCCATGGTGTGGTCTTTTGCACGTGACTT  
1034▶ L G Q S K R V D F C G K G Y H L M S F P Q S A P H G V V F L H V T

4001 ATGTCACAGCTCAGGAAAAGAACTTCACTACAGCCAGCCATCTGCCACGATGGGAAAGCCACTTTCCAGGGAAGGCGTATTGCTGTCCAATGGTAC  
1067▶ Y V P A Q E K N F T T A P A I C H D G K A H F P R E G V F V S N G T

4101 TCATTGGTTCGCTCACTCAGAGAAATTTCTACGAGCCAGATTATAACCACTGACAATACATTTGTATCCGCAATTGTGATGTGGTTATCGGGATTGTG  
1100▶ H W F V T Q R N F Y E P Q I I T T D N T F V S G N C D V V I G I V

4201 AATAACTGTTTACGATCCTTTGACGCGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAATATTTAAGAATCACACATCACCTGACGTGCACCTCG  
1134▶ N N T V Y D P L Q P E L D S F K E E L D K Y F K N H T S P D V D L

V1176F (4326)

4301 GAGATATTTAGGAATCAATGCTTCCCTTGTCAATATTGAGAAGGAGATAGACAGGCTGAATGAGGTTGCCAAGAACCTCAACGAGTCTCTGATGATCT  
1167▶ G D I S G I N A S F V N I Q K E I D R L N E V A K N L N E S L I D L

4401 GCAGGAGTTGGCAAGTACGAACAGTATATCAAATGGCCATGTTACATTTGGCTTGGGTTTCATTGCTGGGCTGATAGCTATCGTCATTGGTGAACATTATG  
1200▶ Q E L G K Y E Q Y I K W P W Y I W L G F I A G L I A I V M V T I M

NheI (4567)

4501 TTGTGTTGCATGACATCTGCTGTAGTTGTCTGAAGGGCTGCTGCTCATCGCGACGCTGTTGCTAAAGCTAGCTGGCCAGACATGATAAGATACATTGAT  
1234▶ L C C M T S C C S C L K G C C S C G S C C •

4601 GAGTTTGGACAAACCACAACCTAGAATGACAGTAAAAAATGCTTTATTGTTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA

4701 AACAAAGTTAACAAACAACAAATTGCATTATTTATGTTTCAGGTTCAAGGGGAGGTGTGGGAGGTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAT

4801 GGAATTTAAATAACAGCATAGCAAACTTTAACCTCCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTG

4901 TTGCCAATGTGCATTAGCTGTTTGCAGCCTCACCTTCTTTTCATGGAGTTAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTCTTCATTCTTTAT

5001 GTTTTAAATGCACTGACCTCCACATCCCTTTTATGATAAATATTAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCA

5101 GAATCCAGATGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACCTTTAATAGAAATTGGACAGCAAGAAAGCGA

5201 GCTTCTAGCTTTAGTTCTGCTGACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCATTCTCAATGAGCACAAAGCAGTCAGG

5301 AGCATAGTCAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCCTGACAGCCACAATG  
111▶ A Y D S I L E R C M G C P S V V R I S R D V E D S Y P H R V A V I

5401 GTGTCAAAGTCTTTGCGCGTGTCTCACAGCAGACCAATGGCAATGGCTTCAAGCAGACAGTACCCTGCCAATGTAGGCCTCAATGTGGACAGCAG  
77▶ T D F D K Q G N S V A S G I A I A E A C V T V R G I Y A E I H V A S

5501 AGATGATCTCCCACTTGGTCTGATGGCCGCCGACATGGTGTGTTGTCCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTC  
44▶ I I E G T K T R I A A G V H H K N D E Y L M T I K E T A V E V L E

5601 CAGATCCTGCTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGTGATTATACTATGCGCATATACTATGCGCATGATTAATTGTCAA  
11▶ L D Q Q S I N F T K M

5701 ACAGCGTGGATGGCGTCTCAGCTTATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAAT

5801 GGGGCGGAGTTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAAACAACTCCCATTGACGTCAATGGGGTGGAGACTTGAAATCCCCGTG  
5901 AGTCAAACCGCTATCCACGCCATTGATGTAAGTCCAAAACCGCATCATGTTAATAGCGATGACTAATACGTAGATGTAAGTCCAAAGTAGGAAAGTC  
6001 CCATAAGGTCATGTAAGTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTAAGTCC  
6101 AAGTGGGCAGTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTTACTATGGAAACATACGTCAATATTGACGTCAATGGGC  
6201 GGGGGTCGTTGGGCGGTGAGCCAGGCGGGCCATTTACCGTAAAGTTATGTAACGCCTGCAGGTTAATTAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC  
6301 AGGAACCGTAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC  
6401 CGACAGGACTATAAAGATACCAGGCGTTTTCCCTGGAAGTCCCTCGTGGCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCTTTCT  
6501 CCTTCGGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGACGAACCCCC  
6601 GTTCAGCCCGACCCTGCGCCTTATCCGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGA  
6701 TTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCT  
6801 GAAGCCAGTTACCTTCGAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTCGAAAGCAGCAGATTACG  
6901 CGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGTCTGACGCTCAGTGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGGCTA  
7001 GTTAATTAACATTTAAATCAGCGGCCGCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGTAACTAACATACGCTCTC  
7101 CATCAAAACAAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCGTCAAGTGCAGGTGCCAGAACATTTCTCTATCGAA