

# pUNO1-SpikeV2-dfur

Expression vector encoding the SARS-CoV-2 U.K. variant (B.1.1.7 lineage) Spike (delta furin) gene

Catalog code: p1-spike-v2-df

<https://www.invivogen.com/uk-b117-spike-expression-vectors>

For research use only

Version 21C15-ED

## PRODUCT INFORMATION

### Contents

- 20 µg of lyophilized pUNO1-SpikeV2-dfur (plasmid DNA)
- 2 x 1 ml of **Blasticidin** (10 mg/ml)

### Storage and Stability

- Product is shipped at room temperature.
- Store lyophilized DNA 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

### United Kingdom (U.K.) SARS-CoV-2 variant Spike cassette

• **EF-1α/HTLV hybrid promoter** is a composite promoter comprised of the Elongation Factor-1α (EF-1α) core promoter<sup>1</sup> and the 5' untranslated region of the Human T-Cell Leukemia Virus (HTLV). EF-1α 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-1α 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-SpikeV2-dfur contains the Spike (S) coding sequence from the United Kingdom (U.K.) variant (B.1.1.7 lineage). This variant is characterized by a number of deletions (del) and mutations within the Spike coding sequence (see below)<sup>3</sup>. The furin cleavage site in pUNO1-SpikeV2-dfur has been inactivated (dfur) by the inclusion of two mutations (R683/5A). Furthermore, to improve expression of the S protein in cell lines, the gene is codon-optimized and the last 19 amino acids, which contain an ER-retention motif (KxHxx), have been removed<sup>4,5</sup>.

pUNO1-SpikeV2-dfur includes the following sequence features:

- **S1 domain:** Del-H69-V70, Del-Y144, A570D, D614G, and P681H
- **RBD:** N501Y
- **S1/S2 boundary:** R683A, R685A
- **S2 domain:** T716I, S982A, and D1118H

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 β-Globin pAn** is a strong polyadenylation (pAn) signal placed downstream of *bsr*. The use of β-globin pAn minimizes interference and possible recombination events with the SV40 pAn signal<sup>8</sup>.

### General features of pUNO1-SpikeV2-dfur

- **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.

### Antibody screening by flow cytometry

pUNO1-SpikeV2-dfur has been specifically designed for mammalian cell expression of the SARS-CoV-2 S protein. Notably, due to the inactivated furin cleavage site, when this plasmid is expressed by a host cell (e.g. 293T cells) there is high surface expression of the full-length S protein<sup>4,9</sup>. Ideal for SARS-CoV-2 S-specific antibody screening by flow cytometry (*in-house data*).

## 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 DH5α.

### • 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. Shen, X. *et al.* 2021. SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral Spike vaccines. *Cell Host and Microbe*, doi: 10.1016/j.chom.2021.03.002. 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. Walls, A.C. *et al.* 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.*

## RELATED PRODUCTS

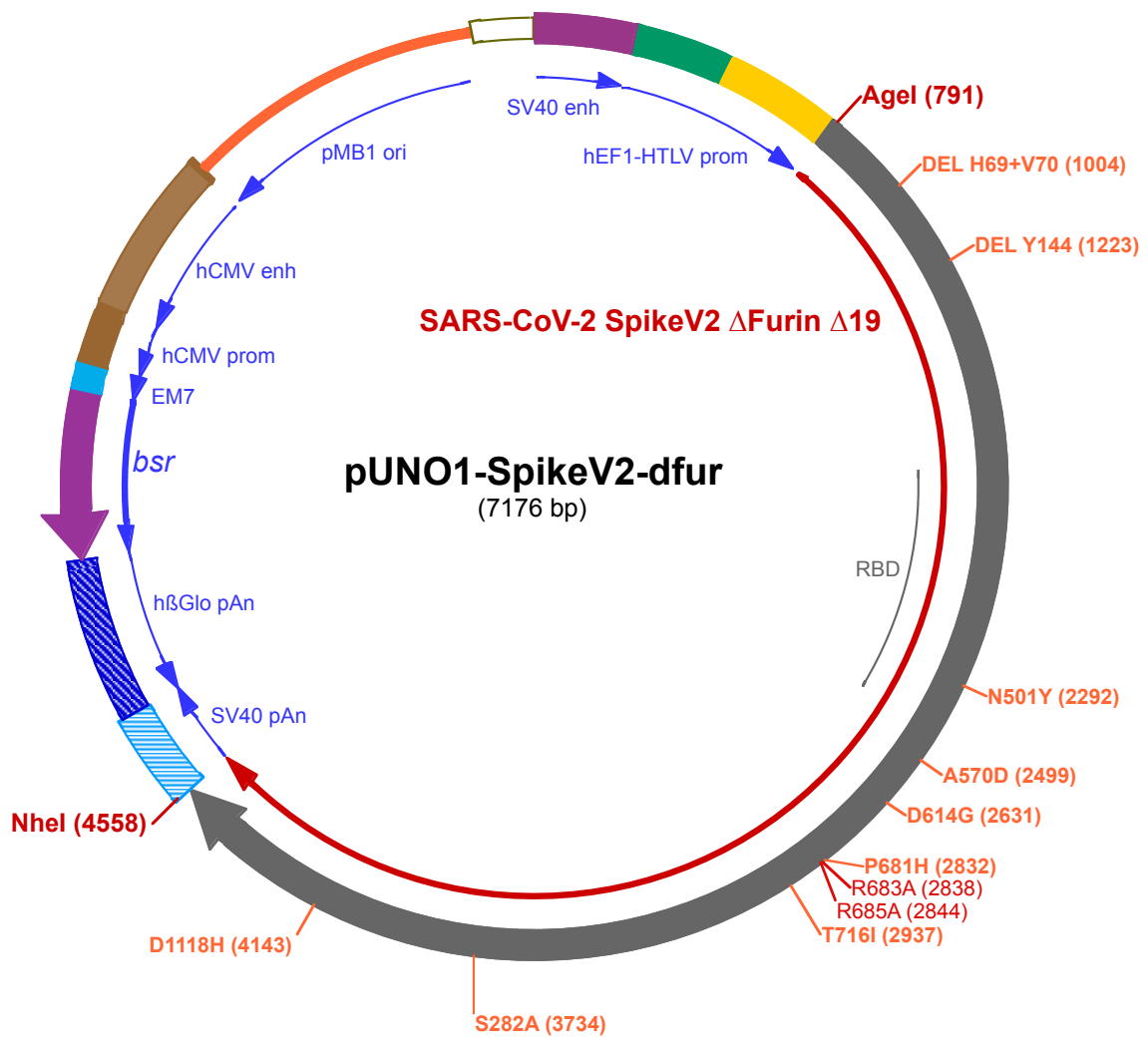
Product	Description	Cat. Code
Blasticidin	Selection antibiotic	ant-bl-1
ChemiComp GT116	Competent <i>E. coli</i>	gt116-11
<b>COVID-19 Product Range</b>		
HEK-Blue™ hACE2 Cells	Cell line	hkb-hace2
A549-hACE2-TMPRSS2 Cells	Cell Line	a549-hace2-tpsa
pUNO1-hACE2	Expression vector	puno1-hace2
pUNO1-hTMPRSS2a	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 GGACCTGCAGGGCCTGAAATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATGTGTGTAGTTAGGGTGTG  
101 GAAAGTCCCAGGCTCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCAGGCTCCCAGCAGGCAG  
201 AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCACTAGTCTCCGGTGCCCGTCACTGGGAGAGCGCACATCGCCACAGTCCCAG  
301 GAAGTTGGGGGGAGGGGTGCGCAATTGAACGGGTGCTTAGAGAAGGTGGCGGGGTAACAGTGGAAAGTGTGCTGTACTGGCTCCGCTTTTTCC  
401 GAGGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGGGTTTGGCCGAGAACAGCTGAAGCTTCGAGGGGCTC  
501 GCATCTCTCTTACGCGCCCGCCCTACCTGAGGCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCTCCCGCTGTGGTGCCTCCTGAACTGC  
601 GTCCGCGCTTAGGTAAGTTTAAAGCTCAGGTCGAGACCGGGCCTTTGTCCGGCGCTCCCTTGAGGCTACCTAGACTCAGCCGGCTCCTCACGCTTTGC

Agel (791)

701 CTGACCCTGCTTGTCAACTCTACGCTTTTGTCTTTTCTGTTCTGCGCGTACAGATCCAAGCTGTGACCGGCGCTACCTGAGATCACCGGTCAA  
801 CATGTTTGTGTTCTTGGTGTGCTTCCACTGGTCAAGTCCCAATGCGTAAATCTCACCACCCGAACTCAACTCCCACCCGATATACAAATTCCTTACC  
1 M F V F L V L L P L V S S Q C V N L T T R T Q L P P A Y T N S F T  
901 AGAGGAGTGTACTATCTGACAAAGTGTTCGGTCAAGTGTCTCCACTCTACTCAGGACCTTTTCTGCCTTTCTTTTCTAACGTTACATGGTTTCATG  
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  
DEL H69+V70 (1004)  
1001 CAATCTCTGGGACAAACGGCACAAACGCTTCGACAACCTGTATTGCCATCAATGATGGGGTGTACTTTGCTCCACAGAGAAATCCAACATCATTGC  
67 A I 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 I R  
1101 AGGATGGATTTTCCGGACTACTCTGGACTCAAAGACACAGAGCCTGTGATCGTTAAACAACGCCACAAACGTTGTCATCAAAGTGTGCGAATTCCAGTTT  
100 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 Q F  
DEL Y144 (1223)  
1201 TGCAATGATCCCTTCTGGGAGTGTATCACAGAATAACAAGTCTGGATGGAGAGCGAATTTGGGTCTACAGCAGCGCAAACAACCTGCACCTTCGAGT  
134 C N D P F L G V Y H K N N K S W M E S E F R V Y S S A N N C T F E  
1301 ACGTGAAGTCAACCTTTCTGATGGACCTGGAAGGGAAACAGGGAAACTTCAAGAACCTGAGAGAGTTTGTCTTTAAGAACATCGACGGCTATTTTAAGAT  
167 Y V S Q P F L M D L E G K Q G N F K N L R E F V F K N I D G Y F K I  
1401 CTATAGTAAGCATAACGCTATCAACCTGGTAAGGGATCTTCCCAGGGCTTTTCCAGCCTGGAACCTTTGGTTGACTTGCCTATTGGTATCAATATCACC  
200 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 N I T  
1501 AGATTTGAGACCTTCTGGCATTGCATCGGTCTTATCTTACTCCAGTGATTCTCCTCCGGTGGACTGCCGGCGCCGCTGCCTACTATGTCGGGTATC  
234 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 V G Y  
1601 TGCAACCAAGAAGCTTCTGCTCAAGTACAACGAAACGGCACTATTACGGATGCTGTTGATTGTGCCCTGGACCCTGTCTGAGACTAAATGCACCT  
267 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 C T L  
1701 CAAGAGCTTTACCGTTGAGAAGGGGATTTACCAAACAGTAATTTCCGGGTCCAACCCACCGAAAGCATTGTGCGGTTCCCAAATATCACCAATCTGTGT  
300 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 N L C

1801 CCCTTTGGCGAAGTGTCAATGCTACAAGGTTTGTCTTGTGTACGCATGGAATAGGAAACGCATCTCCAATTGTGTGCGTATTACTCCGTGCTGTACA  
334 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 V L Y  
1901 ATTCCGCTCTTTTCTCAACCTTCAAGTGTATGGCGTTTACCTACCAAACCTTAACGACCTGTGCTTCACTAATGTGTATGCCGACTTTTTGTGATACG  
367 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 V I R  
2001 AGCGATGAAGTGAACAGATTGCACCAGGCGAGACCGGCAAATTCGGACTACAACCTACAAGCTTCCAGATGACTTTACCGGATGTGTTATTGCATGG  
400 G D E V R Q I A P G Q T G K I A D Y N Y K L P D D F T G C V I A W  
2101 AACTCAAACAATCTGGATTCCAAGTGGGTGGCAACTATAACTACCTGTATAGACTGTTCCAGAAATCCAACCTGAAACATTTCGAGCGAGATATAAGCA  
434 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 D I S

N501Y (2292)

2201 CAGAAATCTACCAGGCTGGAAGTACGCCCTGCAACGGCGTGGAAAGGTTCAACTGCTACTTCCATTGCGAGATTACGGATTCCAGCCTACAATAGGGGT  
467 T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T Y G V  
2301 GGGTTACCAACCTATCGTGTGCTAGTCTGAGTCTTCTGAGTCTTCCATGCCAGCCAGCTGTGGCCCCAAGAAAAGCACCAATCTGGTGAAGAAC  
500 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 V K N

A570D (2499)

2401 AAATGCGTGAACCTTAACTTTAACGGACTCACAGGAACCGGCTATTGACGGAGAGTAACAAGAAGTCTCCTGCCATCCAGCAGTTCCGGTCGCGATATTG  
534 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 R D I  
2501 ATGACACTACCGACTGTCCGAGATCCCAGACATTGGAGATTCTTGATATCACACCTGTAGTTTCCGGCGAGTGAGCGTGATTACGCCCGGAACCAA  
567 D 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 G T N  
D614G (2631)  
2601 TACCAGCAATCAGGTTCCGCTCTGTATCAGGCGTGAATTGACCGAGGTACCTGTCCGATCCAGCTGACCAACTTACCCACATGGCGAGTATAT  
600 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 R V Y

2701 TCCACCGCTCCAACGCTTTTCAGACACGTGCTGGATGTCTGATCGGTGCAGAACACGTTAATAATAGCTACGAGTGTGATATCCCCATCGGTGCTGGAA  
634▶ S T G S N V F Q T R A G C L I G A E H V N N S Y E C D I P I G A G  
R683A (2838)  
P681H (2832) R685A (2844)  
2801 TATGCGCCTTTATCAAACCTCAAACCACTCTCACAGGCGGGCAGCTAGTGTAGCATCCCAAAGTATCATTGCCTACACAATGAGCCTCGGTGCTGAGAA  
667▶ I C A S Y Q T Q T N S H R A A A S V A S Q S I I A Y T M S L G A E N  
T716I (2937)  
2901 TTCTGTCGCCTACAGCAACAACCTCATTGCTATCCCTATTAACTTCAACAATCAGTGTGACAACTGAAATTCTGCCGTATCTATGACCAAAAACAGCGTT  
700▶ S V A Y S N N S I A I P I N F T I S V T T E I L P V S M T K T S V  
3001 GACTGCACCATGTACATCTGTGGCGATTCTACCGAATGTAGCAATCTCCTCGCAATACGGATCATTCTGCACTCAGCTGAATCGTCCCTCACAGGTA  
734▶ 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 L T G  
3101 TTGAGTTGAGCAGGACAAGAATACGCAGGAAGTGTGGCCAGGTGAAGCAAATCTACAAAACCTCCACCCATAAAAGACTTTGGCGGATTCAATTTCTC  
767▶ 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 N F S  
3201 ACAGATCTGCCGATCCCTCAAACCTCCAAGCGTAGCTTTATCGAGGATCTGCTCTTCAACAAGGTAACCTCGCAGATGCCGTTTTCATCAAGCAG  
800▶ 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 I K Q  
3301 TATGGCGATTGTCTGGGAGACATCGCCGCTCGGACCTGATCTGTGCACAGAAGTTCAATGGACTGACCGTGTCTCCCTTGTGACCGACGAGATGA  
834▶ 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 L L T D E M  
3401 TAGCCCAATACATAGCGCCCTGTGCGCCGACCTCACTTCTGGTGGACATTCGGAGATTCCGGAGTGGCGTGCCTTCAGATCCCTTTTGTATGTCAGATGGC  
867▶ 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 Q M A  
3501 CTACCGCTTTAACGGCATCGGTGTGACACAAAACGTTCTGTATGAAAACAGAACTCATCGCAACAGTCAACAGTGTCTATCGGTAAGATACAGGAT  
900▶ 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 I Q D  
3601 AGCCTGTATCCACTGCCAGCGCATTGGGAAAGTTGACAGGATGTAGTGAACCAAGATTGCCAGGCACTTAACACCCTGGTGAACAGCTCTCTTCAAATT  
934▶ S L S S T A S A L G K L Q D V V N Q N A Q A L N T L V K Q L S S N  
S282A (3734)  
3701 TTGGTGCCATTTCTAGCGTGTGAATGACATACTGGACCGTTGGACAAGGTGGAGGCTGAAGTGCAGATTGATAGGCTGATAAAGTGGCGCCTTCAGTC  
967▶ F G A I S S V L N D I L A R L D K V E A E V Q I D R L I T G R L Q S  
3801 TCTTCAGACCTATGTGACCCAGCAGCTCATCCGCGCTGCTGAAATTCGCGCATCCGCTAACCTGGCAGCAACCAAATGTCCGAGTGTGTGCTGGGTCAG  
1000▶ L Q T Y V T Q Q L I R A A E I R A S A N L A A T K M S E C V L G Q  
3901 TCTAAGAGAGTGGACTTTTGGCGGAAGGGTATCACCTGATGTCTTTTCTCAGTCTGCACCCATGGTGTGGTCTTTCTGCACGTGACTTATGTCACG  
1034▶ 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 Y V P  
4001 CTCAGGAAAAGAACTTCACTACAGCCCCAGCCATCTGCCACGATGGGAAAGCCCACTTTCCAGGGAAGGCGTATTCTGTCCAATGGTACTCATTGGTT  
1067▶ 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 H W F  
D1118H (4143)  
4101 CGTCACTCAGAGAAATTTCTACGAGCCCAGATTATAACCACTCACAAATACATTTGTATCCGCAATTGTGATGTGGTTATCGGGATTGTGAATAACT  
1100▶ V T Q R N F Y E P Q I I T T H N T F V S G N C D V V I G I V N N T  
4201 GTTTACGATCCTTTCAGCCAGAGCTGGACTCCTTCAAGGAGGAGCTTGACAAATATTTAAGAATCACACATCACCTGACGTGACCTCGGAGATATTT  
1134▶ 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 G D I  
4301 CAGGAATCAATGCTTCCGTGGTCAATATTGAGAAGGAGATAGACAGCTGAATGAGGTTGCCAAGAACCCTCAACGAGTCTCTGATCGATCTGCAGGAGTT  
1167▶ S G I N A S V V N I Q K E I D R L N E V A K N L N E S L I D L Q E L  
4401 GGGCAAGTACGAACAGTATATCAAATGGCCATGGTACATTTGGCTTGGGTTTATTGCTGGGCTGATAGCTATCGTCATGGTACAATTATGTTGTGTTGC  
1200▶ 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 L C C  
NheI (4558)  
4501 ATGACATCCTGTGTAGTTGTCTGAAGGGCTGCTGCTCATCGGCAGCTGTTGCTAAAGCTAGCTGGCCAGACATGATAAGATAACATTGATGAGTTTGGAA  
1234▶ M T S C C S C L K G C C S C G S C C •  
4601 CAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAAACAAGTTA  
4701 ACAACAACAATTGCATTCATTTTATGTTTCAGGTTTCAGGGGAGGTGTGGGAGTTTTTTAAAGCAAGTAAACCTCTACAAATGGGTATGGAATTTCTA  
4801 AAATACAGCATAGCAAACTTTAACCTCAAATCAAGCCTCTACTTGAATCCTTTTCTGAGGGATGAATAAGGCATAGGCATCAGGGGCTGTTGCCAATG  
4901 TGCATTAGCTGTTTGCAGCCTCACCTTCTTTTCATGGAGTTTAAAGATATAGTGTATTTTCCCAAGGTTTGAAGTACTGCTCTTCAATTTCTTTATGTTTTAAAT  
5001 GCACTGACCTCCACATTCCCTTTTTAGTAAAATATTCAGAAATAATTTAAATACATCATTGCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGA  
5101 TGCTCAAGGCCCTTCATAATATCCCCAGTTTAGTAGTTGGACTTAGGGAACAAGGAACCTTTAATAGAAATTTGGACAGCAAGAAAGCGAGCTTCTAGC  
5201 TTTAGTTCCTGGTGTACTTGGGGGATGAGTTCCTCAATGGTGGTTTTGACCAGCTTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCATAGTC  
141▶ • N R T Y K L P I L E E I T T K V L K G N M E I L V F C D P A Y D  
5301 AGAGATGAGCTCTGTCACATGCCACAGGGGCTGACCACCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGCTGACAGCCACAATGGTGTCAAAG  
108▶ 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 T D F  
5401 TCCTTCTGCCCCGTTGCTCACAGCAGACCCAATGGCAATGGCTTCAAGCAGCAGTACCCCTGCAATGTAGGCTCAATGTGACAGCAGAGATGATCT  
74▶ 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 I I E  
5501 CCCCAGTCTTGGTCTGATGGCCGCCCGACATGGTGTCTTGTCTCATAGAGCATGGTGTCTTCTCAGTGGCGACCTCCACCAGCTCCAGATCCTG  
41▶ 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 L D Q  
5601 CTGAGAGATGTTGAAGTCTTCATGATGGCCCTCTATAGTGAAGTCTATTATACTATGCCGATATACTATGCCGATGATTAATTGTCAAACAGCGTGG  
8▶ Q S I N F T K M  
5701 ATGGCGTCTCCAGCTTATCTGACGGTCACTAAACGAGCTCTGCTTATATAGACCTCCACCGTACACGCCTACCGCCATTTGCGTCAATGGGGCGGAG  
5801 TTGTTACGACATTTTGGAAAGTCCCGTTGATTTACTAGTCAAACAAACTCCCATGACGTCATGGGGTGGAGACTTGGAAATCCCCGTGAGTCAAAC

5901 GCTATCCACGCCATTGATGTACTGCCAAAACCGCATCATCATGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCATAAGGT  
6001 CATGTACTGGGCATAATGCCAGGCGGGCATTACCGTCATTGACGTCAATAGGGGGCGTACTTGGCATATGATACACTTGATGTACTGCCAAGTGGGCA  
6101 GTTTACCGTAAATACTCCACCCATTGACGTCAATGGAAAGTCCCTATTGGCGTACTATGGGAACATACGTCATTATTGACGTCAATGGGCGGGGTCGT  
6201 TGGGCGGTCAGCCAGGCGGGCATTACCGTAAGTTATGTAACGCCTGCAGGTTAAITTAAGAACATGTGAGCAAAGGCCAGCAAAGGCCAGGAACCGT  
6301 AAAAAGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGAC  
6401 TATAAAGATACCAGGCGTTTTCCCTGGAAGCTCCCTCGTGCCTCTCTGTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTCTCCCTTCGGG  
6501 AAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCGTTAGCCC  
6601 GACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG  
6701 CGAGGTATGTAGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT  
6801 TACCTTCGAAAAAGAGTTGGTAGCTTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAA  
6901 AAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGAACGAAACTCACGTTAAGGGATTTTGGTCATGGCTAGTTAATTAA  
7001 CATTTAAATCAGCGGCCAATAAAATATCTTTATTTTATTACATCTGTGTGTTGGTTTTTTGTGTAATCGTAACTAACATACGCTCTCCATCAAAC  
7101 AAAACGAAACAAAACAACTAGCAAATAGGCTGTCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTATCGAA