

 gcacccccctccagtcgaggatatgicggtgtgaakaiagttagang tcgaiang $\square$ TARDBP






 TARDBP

## TARDBP-201



 $\square$ TARDBP TARDBP-201


 TARDBP

## TARDBP-201













TARDBP
TARDBP-201




TARDBP
TARDBP-201


















TARDBP-201





TARDBP-201





TARDBP
TARDBP-201
TARDBP-201








## TARDBP

TARDBP-201

TARDBP-201






 TARDBP
TARDBP-201
TARDBP-201





TARDBP-201




TARDBP
TARDBP-201

TARDBP-201




TARDBP $>$
TARDBP-201

TARDBP-201


 TARDBP TARDBP-201

## TARDBP-201





## TARDBP

TARDBP-201




TARDBP-201


 TARDBP TARDBP-201

TARDBP-201





TARDBP-201




TARDBP
TARDBP-201

TARDBP-201




TARDBP-201


 TARDBP TARDBP-201

TARDBP-201






 TARDBP
TARDBP-201









TARDBP
TARDBP-201
TARDBP-201



TARDBP

TARDBP-201




TARDBP-201




TARDBP-201




TARDBP
TARDBP-201

TARDBP-201


 TARDBP-201





 TARDBP TARPBP-201







 TARDBP-201





TARDBP-201





 TARDBP $>$ TARDBP-201

TARDBP-201








TARDBP-201


 TAROBP $>$ TARDBP-201
 m

 TARDBP-201


 TARDBP TARDBP-201

TARDBP-201












TARDBP-201




TARDBP
TARDBP-201

TARDBP-201




TARDBP
TARDBP-201

TARDBP-201





TARDBP-201




TARDBP
TARDBP-201
TARDBP-201




## TARDBP

TARDBP-201






PCR Forward ACTAAAAGCTGTATTGGGGGTT

 TARDBP
TARDBP-201
TARDBP-201







 TARDBP-201

## AGGCAGCACTA ARNA Protospacer





 GGTACTACCGACGGCGGGTCCGTCGTGATGTCTCGTCAACCCCATACTACCCGTACAATCGGTCGGTCGTCTTGGTCAGTCCGGGTAGCCCATTATTG



—nem




TARDBP-201




TARDBP
TARDBP-201












TARDBP
TARDBP-201






















TARBBP
TAROBP-201






 $\square$ TARDBP-201 $>$


 $\square$ TARDBP
TARBBP-201




TARDBP-201




## TARDBP









TARDBP
TARDBP-201




TAROBP

## TARBBP-201






#### Abstract

  




 TARDBP

















 TARDBP








 TARDBP
















 ACCTGACCTACTACACGTATCCTTAATCTTCCTATATTGGAAAAGTAACAAAGACAAAAGGATGAACATTTACCCATAAATACATTAGGATCCTTICG TARDBP




TARDBP










 TARDBP


 TARDBP







TARDBP











## $3^{\prime}$

18,185

| TARDBP |  |  | .. 18,185 | 18,185 bp | $\square$ | $\rightarrow$ | gene |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| /note | $\begin{aligned} & =\begin{array}{l} \text { gene ENSG00000120948 } \\ \text { Protein coding } \end{array} \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-205 |  |  | .. 8146 | 8146 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000473118 |  |  |  |  |  |  |
| TARDBP-236 |  | 81 | .. 11,266 | 11,186 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000639083 |  |  |  |  |  |  |
| TARDBP-237 |  | 82 | .. 11,840 | 11,759 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000639599 <br> Nonsense mediated decay |  |  |  |  |  |  |
| TARDBP-235 |  | 279 | .. 10,905 | 10,627 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000629725 |  |  |  |  |  |  |
| TARDBP-203 |  | 310 | .. 10,515 | 10,206 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000439080 <br> Nonsense mediated decay |  |  |  |  |  |  |
| TARDBP-201 |  | 311 | .. 13,149 | 12,839 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000240185 |  |  |  |  |  |  |
| TARDBP-204 |  | 311 | .. 9833 | 9523 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST00000472476 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-207 |  | 311 | .. 8246 | 7936 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000476201 |  |  |  |  |  |  |
| TARDBP-219 |  | 311 | .. 6110 | 5800 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000613864 |  |  |  |  |  |  |
| TARDBP-221 |  | 319 | .. 18,185 | 17,867 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} & =\text { primary transcript ENST00000614757 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-231 |  | 327 | .. 9967 | 9641 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000621715 |  |  |  |  |  |  |
| TARDBP-206 |  | 1373 | .. 12,676 | 11,304 bp | $\square$ | $\rightarrow$ | prim_transcript |
| / note | $\begin{aligned} = & \text { primary transcript ENST00000473869 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-202 |  | 1373 | .. 11,840 | $10,468 \mathrm{bp}$ | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000315091 |  |  |  |  |  |  |
| TARDBP-222 |  | 1385 | .. 10,998 | 9614 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000616545 |  |  |  |  |  |  |
| TARDBP-232 |  | 1385 | .. 10,998 | 9614 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000621790 |  |  |  |  |  |  |
| TARDBP-222 |  | 1385 | .. 10,984 | 9600 bp | $\square$ | $\rightarrow$ | CDS |


| - 6 segments $=$ | 906 bp |
| :--- | :--- |
| /note | $=$ coding sequence ENSP 00000484722 |





TARDBP-202 $1385 \ldots 10,905 \quad 9521 \mathrm{bp} \quad \square \rightarrow \mathrm{CDS}$


- 5 segments $=1245 \mathrm{bp}$

Inote $=$ coding sequence ENSP00000240185

 SKGFGFVRGELGVKVMSRMIDGRWCI GFGNSRGGGAGLGNNQGSNMGGGMNFGAFSINPAMMAAAQAALQSSWGMMGMLASQQNQSGPSGNNQNQGNMQREPNQAFGSGNNSYSGSNSGAAIGWGSASNAGSGSGFNGGFGSSMDSKSSGWGM* 414 amino acids $=44.7 \mathrm{kDa}$

| TARDBP-236 | 1385 .. 10,311 |  | 8927 bp | $\square$ | $\rightarrow$ | cDs |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| - 5 segments $=1245 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | = coding sequence EnSP00000491203 |  |  |  |  |  |
| /translation | MSEYIRVTEDENDEPIEIPSEDDGTVLLST SKGFGFVRFTEYETQVKVMSQRHMIDG GFGNSRGGGAGLGNNQGSNMGGGMNF 414 amino acids $=44.7 \mathrm{kDa}$ | CMRGVRLVEGILHAP RKV FV GRCTEDMTED WGMM GMLASQQNQ | A GWGNLV LREFFSQY GPSGNNQN | $\begin{aligned} & \text { VVNY } \\ & \text { SDVM } \\ & \text { QGNM } \end{aligned}$ | $\begin{aligned} & \text { PK,,DN } \\ & \text { DVFIP } \end{aligned}$ QQREP | KRKMDETDASS PFRAFAFVTFA QQAFGSGNNSY |
| TARDBP-231 |  | 1385 .. 9967 | 8583 bp | $\square$ | $\rightarrow$ | cos |
| - 4 segments $=730 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | = coding sequence ENSP00000480690 |  |  |  |  |  |
| /translation | MSEYIRVTEDENDEPIEIPSEDDGTVLLS SKGFGFVRFTEYETQVKVMSQRHMIDGR 243 amino acids $=26.7 \mathrm{kDa}$ | CMRGVRLVEGILHAP LIIKGISVHISNAEPK | A GWGNLV NSNRQLER | VVNY GRF | $\mathrm{PK}_{t,} \mathrm{Dr}$ GNP G | KRKMDETDASSA FGNQGGFGNS |
| TARDBP-207 |  | 1385 .. 8246 | 6862 bp | $\square$ | $\rightarrow$ |  |
| - 5 segments $=827 \mathrm{bp}$ |  |  |  |  |  |  |
| /note <br> /translation | = coding sequence ENSP00000466842 |  |  |  |  |  |
|  | MSEYIRVTEDENDEPIEIPSEDDGTVLLS TSDLIVLGLP WKTTEQ DLKEYFSTFGEV 275 amino acids $=31.4 \mathrm{kDa}$ | CMRGVRLVEGILHAP RFTEYETQVKVMSQR | A GWGNLV <br> MIDGRWCD | CKLP | $\begin{aligned} & \mathrm{PK},, \mathrm{GI} \\ & \mathrm{NSK}, \mathrm{Q} \end{aligned}$ | TWWLTSVIPAT SQDEPLRSRKV |
| TARDBP-205 |  | 1385 .. 8146 | 6762 bp | $\square$ | $\rightarrow$ | cDs |
| - 4 segments $=604 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | = coding sequence ENSP00000465240 |  |  |  |  |  |
| /translation | = MSEYIRVTEDENDEPIEIPSEDDGTVLLS <br> SKGFGFVRFTEYETQVKVMSQRHMIDGR | CMRGVRLVEGILHAP RKVFV GRCTED | AGWGNLV | vvnr | PK,,D | KRKMDETDASSA |
| TARDBP-233 |  | $4516 . .10,944$ | 6429 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000622057 |  |  |  |  |  |
| TARDBP-233 |  | $4516 . .10,840$ | 6325 bp | $\square$ | $\rightarrow$ | cDs |
| - 5 segments $=602 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | $=$ coding sequence ENSP00000481206 |  |  |  |  |  |
| /translation | $\begin{aligned} = & \text { WMRQMLHQQ*K*KEQSRKHPI**CWV } \\ & \text { LPLLHLQMIR,,LRSLFVERT*SLKESAFIY } \end{aligned}$ | CR,,SRKILRLVIQRGL NMKPSF | LFV LRNMK |  | Chs | **mddgutanf |
| TARDBP-223 |  | $4522 . .11,247$ | 6726 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000617172 |  |  |  |  |  |
| TARDBP-223 |  | $4522 . .10,905$ | 6384 bp | $\square$ | $\rightarrow$ | cDs |
| - 5 segments $=638 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | = coding sequence ENSP00000479219 |  |  |  |  |  |
| /translation | $\begin{aligned} = & \text { RQMLHQQ*K*KEQSRKHPI**CWVSHG } \\ & \text { LHLQMIR,,LRSLFVERT*SLKESAFIYPM } \end{aligned}$ | ,,SRKILRLVIQRGLALF ,FISFQMFMEEALH*K | LRNMKHK <br> C |  |  | MDDGVTANFLIL |
| TARDBP-225 |  | 4535 .. 6843 | 2309 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} & =\begin{array}{l} \text { primary transcript ENST00000618606 } \\ \text { Retained intron } \end{array} \end{aligned}$ |  |  |  |  |  |
| TARDBP-215 |  | 4623 . 11,850 | 7228 bp |  | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000611963 <br> Nonsense mediated decay |  |  |  |  |  |
| TARDBP-226 |  | $6437 . .11,841$ | 5405 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000619555 <br> Nonsense mediated decay |  |  |  |  |  |
| TARDBP-229 |  | 6437 .. 11,700 | 5264 bp |  | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST000000620632 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |
| TARDBP-212 |  | 6510 .. 11,700 | 5191 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000610369 |  |  |  |  |  |
| TARDBP-212 |  | 6510 .. 10,984 | 4475 bp |  | $\rightarrow$ |  |
| - 4 segments $=384 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | $=$ coding sequence ENSP00000482559 |  |  |  |  |  |
| /translation | $\begin{aligned} = & \text { CKLPNSK,,QSQDEPLRSRKVFVGRCTED } \\ & 127 \text { amino acids }=14.5 \mathrm{kDa} \end{aligned}$ | PKPFRAFAFVTFADD | ,,IAQSLCG |  |  | ISNAEPKHNSN |
| TARDBP-220 |  | 8090 .. 11,840 | 3751 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST00000614494 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |
| TARDBP-214 |  | 8163 .. 18,087 | 9925 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST000000611136 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |
| TARDBP-208 |  | $8244 . .12,864$ | 4621 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST00000477447 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |
| TARDBP-217 |  | $9801 . .18,183$ | 8383 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000612542 |  |  |  |  |  |
| TARDBP-217 |  | $9801 . .12,843$ | 3043 bp | - | $\rightarrow$ | cDs |
| - 2 segments $=184 \mathrm{bp}$ |  |  |  |  |  |  |
| /note | $=$ coding sequence ENSP00000478249 |  |  |  |  |  |
| /translation | $=R G L D H * R N Q R S Y I Q C R T * A Q * Q * T V R K K$ <br> 61 codons (7 internal stop codons) | NT*NGTRGPPKTKL |  |  |  |  |
| TARDBP-228 | $=$ primary transcript ENST00000620505 <br> Nonsense mediated decay | 9965 .. 12,878 | 2914 bp | - | $\rightarrow$ | prim_transcript |
| /note |  |  |  |  |  |  |
| Donor Template SNV -> REVProtospacer Sequence |  | 10,028 .. 10,127 | 100 bp |  | $\square$ | misc_feature |
|  |  | 10,046 .. 10,065 | 20 bp | $\square$ | - | misc_feature |


| Featur |  | Locati |  | Size | $\square$ | \# | Type variation |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SNV |  | 10,057 | .. 10,057 | 1 bp | $\square$ | $\mapsto$ |  |
| /note | $\begin{aligned} = & R E V \\ S N V & =A \end{aligned}$ |  |  |  |  |  |  |
| PAM |  | 10,066 | .. 10,068 | 3 bp | $\square$ | $\mapsto$ | misc_feature |
| TARDBP-230 |  | 10,205 | .. 11,930 | 1726 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENSTO0000621573 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-213 |  | 10,475 | .. 13,149 | 2675 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000611008 |  |  |  |  |  |  |
| TARDBP-218 |  | 10,702 | .. 12,860 | 2159 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000613177 |  |  |  |  |  |  |
| TARDBP-234 |  | 10,855 | .. 18,154 | 7300 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} & =\text { primary transcript ENST00000622108 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-210 |  | 10,855 | .. 13,396 | 2542 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $\begin{aligned} = & \text { primary transcript ENST00000496840 } \\ & \text { Nonsense mediated decay } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-224 |  | 10,980 | .. 11,809 | 830 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000617757 |  |  |  |  |  |  |
| TARDBP-209 |  | 11,586 | .. 13,149 | 1564 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000480464 |  |  |  |  |  |  |
| TARDBP-227 |  | 12,450 | .. 18,182 | 5733 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000620028 |  |  |  |  |  |  |
| MASP2 |  | 14,180 | .. 34,896 | 20,717 bp | $\square$ | $\leftarrow$ | gene |
| /note | $\begin{aligned} & =\text { gene ENSG00000009724 } \\ & \text { Protein coding } \end{aligned}$ |  |  |  |  |  |  |
| MASP2-201 |  | 14,180 | .. 34,896 | 20,717 bp | $\square$ | * | prim_transcript |
| /note | $\begin{aligned} &= \text { primary transcript ENST00000400897 } \\ & \text { Protein coding } \end{aligned}$ |  |  |  |  |  |  |
| TARDBP-211 |  | 17,316 | .. 18,185 | 870 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000607145 |  |  |  |  |  |  |
| TARDBP-216 |  | 17,415 | .. 18,185 | 771 bp | $\square$ | $\rightarrow$ | prim_transcript |
| /note | $=$ primary transcript ENST00000612387 |  |  |  |  |  |  |

[^0]Binding Sit
Donor Template SNV-> REV $\quad 100-\mathrm{mer}$ ○ 10,028 .. 10,127 - $82^{\circ} \mathrm{C}$ Apr 22,2022
$=$ TGGTTATTACCCGATGGGCCTGACTGGTTCTGCTGGCTGGCTAACATGCCCATCATACCCCAACTGCTCTGTAGTGCTGCCTGGGCGGCAGCCATCATGG 58\% GC / 30,751.9 Da
gRNA Protospace
20-mer $\quad 10,046$.. $10,065 \rightarrow 54^{\circ} \mathrm{C}$ Apr 22, 2022
/sequence $\quad=$ AGGCAGCACTAAAGAGCAGT $50 \%$ GC / 6184.1 Da

| PCR Reverse |  |
| :--- | :--- |
|  |  |
| /sequence | $=$GCTGCACCAGAATTAGAGCC  <br>  $55 \% \mathrm{GC} / 6111.0 \mathrm{Da}$ | $55 \%$ GC / 6111.0 Da

Sanger Sequencing Primer $\quad 20-\mathrm{mer} \quad 10,192 \ldots 10,211 \sim 59^{\circ} \mathrm{C}$ Apr 22,2022
sequence $=$ GCTGCACCAGAATTAGAGCC $55 \%$ GC / 6111.0 Da

| 20-mer | 10,192 .. 10,211 | - | $59^{\circ} \mathrm{C}$ | Apr 22, 2022 |
| :---: | :---: | :---: | :---: | :---: |


[^0]:    Length

