

PAX6 coding region with three-frame translation

```

    A E Q S Q R S E S A R W C L C Q R A A T A G L H P A E D C R
    C * R T V T A E * I S S V V S L S T G G H C R T P P G R R L *
1 M Q N S H S G V N Q L G G V F V N G R P L P D S T R Q K I V 30
1 ATGCAGAACAGTCACAGCGGAGTGAATCAGCTCGGTGGTGTCTTTGTCAACGGGCGGCCACTGCCGGACTCCACCCGGCAGAAGATTGTA 90
    *5 * * * * * * * *
    A S S Q R G P A V R H F P N S A G V Q R M C E * N S G Q V L
    S * L T A G P G R A T F P E F C R C P T D V * V K F W A G I
31 E L A H S G A R P C D I S R I L Q V S N G C V S K I L G R Y 60
91 GAGCTAGCTCACAGCGGGGCCCGCCGTGCGACATTTCCCGAATTCTGCAGGTGTCCAACGGATGTGTGAGTAAAATTCTGGGCAGGTAT 180
    * * * * * 6 * * * *
    R D W L H Q T Q G N R W * * T E S S D S R S C K Q N S P V *
    T R L A P S D P G Q S V V V N R E * R L Q K L * A K * P S I
61 Y E T G S I R P R A I G G S K P R V A T P E V V S K I A Q Y 90
181 TACGAGACTGGCTCCATCAGACCCAGGGCAATCGGTGGTAGTAAACCGAGAGTAGCGACTCCAGAAGTTGTAAGCAAATAGCCAGTAT 270
    * * * * * * * * * *
    A G V P V H L C L G N P R Q I T V R G G L Y Q R * H T K R V
    S G S A R P S L L G K S E T D Y C P R G S V P T I T Y Q A C
91 K R E C P S I F A W E I R D R L L S E G V C T N D N I P S V 120
271 AAGCGGGAGTGCCCGTCCATCTTTGCTTGGGAAATCCGAGACAGATTACTGTCCGAGGGGGTCTGTACCAACGATAACATAACCAAGCGTG 360
    * * * * * * * * * 7 *
    I N K Q S S S Q P G * R K A T D G R R R H V * * T K D V E R
    H Q * T E F F A T W L A K S N R W A Q T A C M I N * G C * T
121 S S I N R V L R N L A S E K Q Q M G A D G M Y D K L R M L N 150
361 TCATCAATAAACAGAGTTCTTCGCAACCTGGCTAGCGAAAAGCAACAGATGGGCGCAGACGGCATGTATGATAAACTAAGGATGTTGAAC 450
    * * * * * * * * *

```

A D R K L G H P P W L V S G D F G A R A T Y A R W L P A T G
G R P E A G A P A L V G I R G L R C Q G N L R K M A A S N R
151 **G Q T G S W G T R P G W Y P G T S V P G Q P T Q D G C Q Q Q** 180
451 GGCAGACCGGAAGCTGGGGCACCCGCCCTGGTTGGTATCCGGGGACTTCGGTGCCAGGGCAACCTACGCAAG**ATGGCTGCCAGCAACAG** 540
* * * * * 8 * *

R R G R E Y Q L H Q F Q R R R F R * G S N A T S A E A E A A
K E G E R I P T P S V P T E K I Q M R L K C D F S * S G S C
181 **E G G G E N T N S I S S N G E D S D E A Q M R L Q L K R K L** 210
541 **GAAGGAGGGGGAGAGAATACCAACTCCATCAGTTCCAACGGAGAAGATTTCAGATGAGGCTCAAATGCGACTTCAGCTGAAGCGGAAGCTG** 630
* * * * * * * * *

K K * N I L Y P R A N * G P G E R V * E N P L S R C V C P R
K E I E H P L P K S K L R P W R K S L R E P I I Q M C L P E
211 **Q R N R T S F T Q E Q I E A L E K E F E R T H Y P D V F A R** 240
631 **CAAAGAAATAGAACATCCTTTACCCAAGAGCAAATTGAGGCCCTGGAGAAAG**AGTTTGAGAGAACCCATTATCCAGATGTGTTTGCCCGA 720
* * * * * 9 * * *

K T S S Q N R S T * S K N T G M V F * S K G Q M E K R R K T
K D * Q P K * I Y L K Q E Y R Y G F L I E G P N G E E K K N
241 **E R L A A K I D L P E A R I Q V W F S N R R A K W R R E E K** 270
720 GAAAGACTAGCAGCCAAAATAGATCTACCTGAAGCAAGAATACAG**GTATGGTTTTCTAATCGAAGGGCCAAATGGAGAAGAGAAGAAAA** 810
* * * * * 10 * * *

E E S E K T G Q Q H T * S Y S Y Q Q * F Q H Q C L P T N S T
* G I R E D R P A T H L V I F L S A V V S A P V S T N Q F H
271 **L R N Q R R Q A S N T P S H I P I S S S F S T S V Y Q P I P** 300
811 **CTGAGGAATCAGAGAAGACAGGCCAGCAACACACCTAGTCATATTCCTATCAGCAGTAGTTTCAGCACCAGTGTCTACCAACCAATTCCA** 900
* * * * * * * * *

T H H T G F L L H I W L H V G P N R H S P H K H L Q R S A A
 N P P H R F P P S H L A P C W A E Q T Q P S Q T P T A L C R
301 Q P T T P V S S F T S G S M L G R T D T A L T N T Y S A L P 330
 901 **CAACCCACCACACCGG**TTTCCTCCTTCACATCTGGCTCCATGTTGGGCCGAACAGACACAGCCCTCACAAACACCTACAGCGCTCTGCCG 990
 * 11 * * * * * * * *

Y A Q L H H G K * P A Y A T P S P Q P D L L I L L H A A H Q
 L C P A S P W Q I T C L C N P Q S P A R P P H T P A C C P P
331 P M P S F T M A N N L P M Q P P V P S Q T S S Y S C M L P T 360
 991 CCTATGCCCAGCTTACCATGGCAAATAACCTGCCTATGCAA**CCCCCAGTCCCCAGCCAGACCTCTCATACTCCTGCATGCTGCCACC** 1080
 * * * * 12 * * * * *

P F G E W A E L * Y L H P P T Y A D T H E Q S A N G H L G H
 A L R * M G G V M I P T P P H I C R H T * T V S Q W A P R A
361 S P S V N G R S Y D T Y T P P H M Q T H M N S Q P M G T S G 390
 1081 **AGCCCTTCGGTGAATGGCGGAGTTATGATACCTACACCCCCCACATATGCAGACACACATGAACAGTCAGCCAATGGGCACCTCGGGC** 1170
 * * * * * * * *

H F N R T H F P W C V S S S S S R K * T * Y V S I L A K I
 P L Q Q D S F P L V C Q F Q F K F P E V N L I C L N T G Q D
391 T T S T G L I S P G V S V P V Q V P G S E P D M S Q Y W P R 420
 1171 **ACCACTTCAACAG**GACTCATTTCCCCTGGTGTGTCAGTTCCAGTTCAAGTTCCCAGTGAACCTGATATGTCTCAATACTGGCCAAGA 1260
 * 13 * * * * * * * *

T V K K K K K K R K G N I V L I Q S
 Y S K K K K K K K E R K Y C V N S V S
421 L Q * K K K K K K G K E I L C * F S Q
 1261 TTACAG**TAA** AAAAAAAAAAAAAAAAAAAAGGAAAGGAAATATTGTGTTAATTCACTCAGT
 | | |
 1269 *1 *10 *20 (etc)

This figure shows the cDNA sequence of the coding region of the most common human PAX6 transcript (which lacks the alternatively spliced exon 5a).

cDNA sequence

Alternate exons are unshaded and shaded in yellow.

Exon numbers are shown underneath the first base of each exon.

Every tenth base has an asterisk underneath.

Amino acid sequence

The amino acids of the paired domain are shown in red'

The amino acids of the homeodomain are shown in blue.

Every tenth amino acid is underlined.