

S16 File. The gene model for *RcWRKY40*. The coding region is marked with uppercase letters, above which is its deduced amino acids. The transcribed untranslated regions, including 5' UTR, intron and 3' UTR sequences, are marked with lowercase letters. The start and stop codons are blacked and the misannotated sequence is boxed.

```

1 cttttactacttttcttttattttatctttttatttttaagtatgatgatttgtttatgt
61 tgttttcgagttaaagagagcctttgttttcttttgggtcattttagcaaaatccaagaa
121 agagagacagctctagtcttctctttcacaccatggtatagctcgcataatctcactgat
181 caacttagcaaggtcttttttttattttttcgtttgggtcattgtttgtctgttctttttc
241 tgtttttctattgttttagccaagcttagttctatcctattataagttattagtgacgag
301 aaaaggctaaattttttgggattattattttttatggatagtttttatttcagtatgg
361 tggttttaactctcactctgcttggagaagaactttttgggtgatgcaaaaattttgag
421 atttgagcaatagttatttgggttttgcatttgggaaatgggttttgtgtgttttggtt
481 ttgtagtggtagtttgaactctcacttgcactgagagtttaagggttttaaaactga
1
M E E V E E A N K A A
541 ggttttggtttcttgagttttgggtgattATGGAGGAGGTTGAGGAGGCAAATAAGGCAGC
12 V E S C R R V I A L L C Q P R D Q V Q A
601 TGTTGAGAGCTGCCGTAGAGTTATTGCTCTCCTATGTCAACCTAGAGATCAAGTTCAGGC
32 R N L V T E T G E T V S K F K R V V S L
661 TAGAACTTGGTGACGGAAACTGGAGAACTGTGTCTAAGTTTAAAAGAGTGGTGTCTCT
52 L S N S L G H G K V R K L K K F K S S L
721 GTTGAGTAATAGTTTAGGTCATGGAAAAGTGAGAAAGTTGAAGAAGTTTAAATCATCTTT
72 P Q N I F L D S P N C T T I L A P K P L
781 GCCTCAAACATCTTTCTAGATAGTCTAATTGCACAACAATTTGGCACCGAAACCCCT
92 Q M V P A S F L E T P I S E M D S K S K
841 CCAAATGGTCTCTGCTAGTTTCTCGAGACCCCTATTCTGAAATGGATTCCAATCTAA
112 S S M Q P T R K I F L E N S V V E M S S
901 ATCTAGTATGCAACCTACCCGAAAATTTTCTTAAAATTCAGTTGTTGAAATGAGCTC
132 N V R P S F Q I S Q I K P T Q Q Y Q F L
961 AAATGTCAGGCCTTCTTTCCAGATTTCTCAAATAAAACCAACACAACAATACCAGTTTCT
152 Q Q Q Q Q I Q R M Q F Q Q Q Q Q M K Y Q
1021 TCAACAACAACAGCAAATACAAAGGATGCAGTTTCAGCAGCAGCAACAATGAAATATCA
172 A D M M Y S R S N S G I N L K F D G S T
1081 AGCTGATATGATGTATTCTAGGAGTAACAGTGGGATTAATCTTAAATTCGATGGGTCTAC
192 C T P A M S T T R S F I S S L S M D G T
1141 CTGCACACCAGCCATGTCAACCACAAGATCATTATATCATCTTTAAGCATGGATGGCAC
212 V T N F D R D S F H L I G V P Q S S D Q
1201 CGTGACTAACTTTGATAGAGATTCATCCATCTGATTGGTGTACCTCAGTCATCTGATCA
232 I S Q Q T R R R C S V R G E D G S V K C
1261 GATCTCGCAGCAGACAAGACGAAGGTGTTCTGTTAGAGGGGAAGATGGCAGTGTTAAATG
252 A S S G K C H C S K R R
1321 CGCCAGCTCCGGTAAATGTCATTGTTCAAAGAGAAAGgtcaaaactcaaatctctgttagtt
1381 tctaattctggttcttttcttttaaaattataattctgatccaaaatacttttgttttag

```

264 K L R V K R S I K V P A I S N K V A D I
1441 GAAACTGAGAGTGAAGAGATCCATAAAGGTGCCTGCTATTAGTAATAAGGTGGCAGATAT
284 P P D E Y S W R K Y G Q K P I K G S P H
1501 TCCTCCTGATGAGTATTCATGGAGAAAATATGGGCAAAAGCCGATTAAGGGCTCTCCACA
304 P R
1561 TCCTAGgtatattatctaattcaaccttttattataagttttaagatacatataacctgtg
1621 atttatattttgtcatgttctaagattgttcctgtgcatgcataaacagtaactctacctt
1681 aagtagttggtaattgcttaatacaaatgaatgacgcatgttaagagatatttgga
1741 acaagaaccttagctgtctttactttctgcaactcttaattggtaggtaactccgtttg
306 G Y Y K C S S L R
1801 tctgacacttttgaaactccacttttactgcagGGGATACTACAAATGTAGCAGCCTTAG
315 G C P A R K H V E R C L E D P S M L I V
1861 AGGTTGCCAGCAAGGAAGCATGTCGAGAGATGCCTGGAAGACCCTTCAATGCTAATTGT
335 T Y E G E H N H S R L L S T Q S A H T *
1921 CACTTATGAAGGTGAGCATAATCATTCCAGATTACTCTCGACACAATCTGCCACATA**TA**
1981 **A**aaactttgcagatgtcttcagtcctctctccaagaattcaggattccaaaaatgtttcg
2041 aatccaaccggaatcgctggcgcttatttgtgaattactactgtatataatattgttg
2101 ttcgaaccagagactagcaatggtggagcagattttatgattagatgacagggggttagt
2161 ccggtttgagctgcttttattatacagaatagggattaaaaattgtaaacttatgcttcc
2221 ctatataatgaagttgtgtttctctgcaaacccatggatttctaatacaaaagcaacg
2281 cctgttaggctaggttttttgttctaatttaataatattatgcttctataggcagcatcagcc
2341 attgtaggtgcaatat