

S20 File. The gene model for *RcWRKY54*. 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 tctacgctttctgttcttcttgcaactgatcttggatccattagctttttcagctctt
61 ttcattctttataactcaacattcttctcttttctcttttaagcccaactttcatgtttgtt
121 gcttttaattcggttatcagagctttatcttagccttctttttcacttctcttctcccac
181 atagcaatagctctgtttaatatcttctatccaaagcctgggtacaagccacagctgtag
1
1 M E K
241 ctgtttcgaagtatatcagtagtttgaacttgtccttggctgcatttttctaaATGGAAA
4 H T L I G E L T Q G K E L T E Q L K H H
301 AGCATACTTCATCGGTGAACTAACACAAGGGAAGGAGCTCACGGAACAGCTCAAACACC
24 L N P A S S L Q T R E F L V G K I I S S
361 ATCTTAACCCTGCATCGTCACTTCAAACCTCGTGAATTCCTTGTGGAAAGATAATTTCTT
44 Y E K A L S L L N C G A S F D V H H N P
421 CTTATGAGAAAAGCACTTTCTTTGCTGAATTGTGGTGCTAGTTTTGATGTTACCATAACC
64 N I C I S E S M H P I H K R S P R S E I
481 CCAATATTTGCATATCAGAATCGATGCACCCAATTCATAAGCGTAGCCCCGAGAAGTGAAA
84 P D E D C K D Q A Q K N V F K K R
541 TCCCTGATGAAGACTGCAAAGACCAAGCTCAGAAAAATGTATTCAAGAAAAGgtaaatta
601 attagtttgtaatttgagtcattgagctaagctatattttcagttttctttatctgatt
101
101 K T Q Q R C T
661 tcgcggttgatattcatcgtatattgtttctttgcagAAAGACACAGCAAAGATGTACA
108 E Q V K V C S G T G L D G P L G D G Y S
721 GAACAAGTGAAGGTGTGTTTCAGGAACAGGACTAGATGGTCCGCTTGGTGATGGTTATAGC
128 W R K Y G Q K D I L G A K F P R
781 TGGAGGAAATATGGTCAAAAAGATATTCTTGGAGCTAAATTTCCAAGgtgtgtgagatac
841 tctgttcatctttatcttcacattttaatttcctaagaaaaagaatagaaaaagtacag
901 gtggacttgaattatatgcagcagctttgattaaatggtagtgtgtagatgcagcttgg
961 aacatgaaccagtaataggagcatatgtcactgatattaataaatataatctattggatg
1021 tacgttatgaaaggtaaaaagtaattataggtaaagtttttaagataaaaacacacgtc
1081 tggtagttttttattaacttacttaacagtaccaaagtctgtatgtatgttctctaagaa
1141 gctggaaggattggcttatttaagtaacttccgtttgcataattattttcttcttgaga
144
144 G Y Y
1201 aaaagaagtaaagaaaggattgctcctaacttggagagttatgcaattacagGGGATATTA
147 R C T H R H S Q G C L A I K Q V Q R S D
1261 TAGATGCACTACCGTCACTCGCAAGGTTGTTTGGCGATAAAGCAAGTACAAAGATCAGA
167 E N P S I F E V T Y R R K H T C V Q A S
1321 CGAGAACCCTCAATTTTTGAAGTGACCTATCGACGAAAGCATACTGCGTCCAAGCTTC
187 H S A M S N D N P K Q N Q G H P Q L Q Q
1381 GCACTCAGCAATGTCTAATGACAACCCTAAACAGAACCAAGGACATCCTCAGCTACAACA
207 E E K P I P T T M E E F L N S E G L I V

```

1441 AGAAGAGAAGCCAATACCAACAACAATGGAGGAGTTCCTTAACTCGGAAGGACTTATAGT
227 K T E N L D T G E E I F P S F S F L D A
1501 TAAAACTGAGAATTTGGATACTGGGGAGGAAATCTTCCATCCTTTTCCTTCTGGATGC
247 L I D K H N E E N S F V G S L S P T F I
1561 ATTGATCGACAAACATAATGAAGAAAACAGTTTCGTGGGCAGCCTTTCTCCGACATTTAT
267 S P A T S E S N Y F S T S S C Q F D T L
1621 ATCTCCAGCGACTTCCGAGTCCAATAATTTCTCGACGTCGTCTTGCCAGTTTGACACCCT
287 G I G Y N V Q T P E S E L T E I G S V P
1681 AGGAATAGGCTACAATGTGCAGACTCCAGAGTCGGAACACGGAGATAGGCTCGGTTCC
307 T S V T N S P I G D F D F S N M D L E T
1741 GACTTCGGTCACTAATTCTCCAATAGGGGATTTTGACTTCTCTAACATGGATCTCGAAAC
327 S F Q F G N L E L E F L A *
1801 CAGTTTCCAATTTGGCAACTTGGAGTTGGAATTCCTTGCTTAGgagtgttagccacatac
1861 tgcattaaacgtatcaatataccttagaaaggggatagagtagaaatccagctga
1921 ttaggtagctgaatcttgaaataatgtgtagcctgtagtccatctctctgtttcttt
1981 cttttctctcactcaacttattggtatttcaagatgctgcagtgttgaattattg
2041 tttcacctgtttcttaatcttgaatgtattcataatctcatgcgattcattctc