

gtagaaatctcacttattcagtaagaact

gagtttcatagaaaATGATTGTCATGGTTTTGTTTAAATATCTTCGTTTTTCGTC AATTTCT

1 M I V M V L F N I F V F V N F signal sequence

CATATTCGTTGGCTAATTTAAATGATTCCTAGCATTATATCTCAAGTTGTCTCTAATA

16 S Y S L A N L N D S S I L Y S Q V V S N

AGTGTGCATCTAATAATGGTATATCTGTGAATGTTTCGAAACAAGGCGGAGAAATTAAGg

36 K C A S N N G I F C E C S K Q G G E I K

ttctacattcaaaaagataaaaaccaattcaataaaatcttgatgcagATTTGGAATTGC

56 I W N C

A in u787 (W to *)

ATCCGTC CAGGAGTTGTCAACGTGGTCAAAATGGTCGAAATGTAGAGAAGGTATCCGA

60 I R P G G W S T W S K W S K C R E G I R TSP1 #1

AAAAGAAGACGAACGTGTAATAATCCTTTACCAATTTGGTACCCTTTGTCTGGCCAGAAA

80 K R R R T C N N P L P I G T T C S G Q K

GTTGAGAAGCAATCTGTGCAATTTCTTCAAATGTACCGGAGTATCTTTTGGATCCTGG

100 V E K Q S C A I S S N V P E Y L F G S W

ACATCCTGGAATCCATGGTCTCGATGCGATTGTGATCGTAGCTGAGGATACGgtaggtt

120 T S W N P W S R C D C D R S L R I R TSP1 #2

caacaattgggaaatgtacagtaaccaagtttcagAACTCGGCACTGCAAAGGTAATT

138 T R H C K G N

A in lq37 (C to Y)

CCTGCGAGGGATGTGACAAGGATTACGAAGATTGCCGCCAGATGAGTGTCCAATTAGTA

145 S C E G C D K D Y E D C R P D E C P I S

AAAAATGGTCCGAATGGACAGATTGGGTGAATTATGgtatgggtctcattacgctgaaat

165 K K W S E W T D W V N Y

atatcaacaatctgtaatgatttaatttcagGAATTGAGCAAGTACGGTTTTTCAGCCTG

177 G I E Q V R F S A W

A in lq78 (G to E)

GTGTTTCGTCATCAAACGTGGCGAATACTGAAGTGGGAATACGGAAGGAGACTCAAGACTC

187 C S S S N V A N T E V G I R K E T Q D S

AAATGAAGCATGgtattaatcttagttaaaaaatctcgaagaaccataactgaaagactaa

207 M K H

atttgaaagtcttcaagttacaatgtccactttactgtctaaatctttgtgtagttaaca

a in lq84 (3' splice site)

acatgataaaaaatataatgagaataatctcagCCAAGTGGTCTGAATGGCATATGCATC

210 A N W S E W H M H

CAGGAGTTGCATACCGTTATCGTCTTCTTCAACTCTTCTATTTCCATCGAACACCATC

219 P G V A Y R Y R L L H N S S I S I E H H

TCTTATCAAGATTCACGTCATCATGTCTTCCCTTGCATTTTGCAATCCAATATTTGTT

239 L L S R F T S S C L P L H F A I P I F C Transmembrane

TTTGATTCTCACGGGTTTCTTCTCAAATATTTATTTATTGTTGTGAATCGTTTTTA

259 F C I L T G F L L Q N I I Y C V V N R F

AAAGGAGATTCATAAGATTGAATTATTCGTATGATTCAAATCCACGTGACTATCCTTCTC

279 K R R F I R L N Y S Y D S N P R D Y P S

ATTTGATTCGTTCTCCGGTTCCTCCGAAAGATGAAAGTTTTTGGTGAatcttcggatgctt

299 H L I R S P G S P K D E S F W *

tcgagaacagtctctgtctgcccattctcagccacgataataaaaagttatcattgatc

Figure S2 The *mig-21* coding region. The DNA sequence of the *mig-21* coding region as determined from cDNAs and transcript sequencing runs on Wormbase, and from our RNA seq analysis. The predicted 5' and 3' untranslated regions and introns are lowercase, and the open reading frame is upper case. A putative poly-A site in the 3' UTR is underlined. Nucleotide lesions associated with *mig-21* alleles are shown in red above the DNA sequence. The putative polypeptide encoded by *mig-21* is shown below the DNA sequence. A potential signal sequence, two potential thrombospondin type I domains, and a potential transmembrane domain and highlighted in grey.