

Differential evolution of duplicated medakafish *mitf* genes

Mingyou Li, Feng Zhu, Yunhan Hong*

Department of Biological Sciences, National University of Singapore, Singapore 117543, Singapore;

Supplementary figure legends

Figure S1. Nucleotide and deduced amino acid sequence of medaka *mitf1* and *mitf2*. (A) *mitf1*. (B) *mitf2*. Start codon and stop codon are shown in bold. Broken and solid arrows depict sequences of primers for PCR analysis and cDNA cloning, respectively. Sequences are deposited in GenBank under accession numbers KC249979 (*mitf1*) and KC249980 (*mitf2*).

Figure S2. Mitf protein sequence alignment. Common species names, Mitf proteins or isoforms and percentage sequence similarity values compared to the medaka Mitf1 and Mitf2 are seen. The basic helix-loop-helix leucine zipper sequences are indicated. For accession numbers see Figure 1B.

Figure S3. Phylogenetic tree construction by maximum likelihood. For accession numbers and more details, see the legend to Figure 1B.

Figure S4. Phylogenetic tree construction by minimum evolution. For accession numbers and more details, see the legend to Figure 1B.

A

1 CCCCAGTTTTGTGTTTACTTGTGTTGGGTTTTTTTTACGTTCTGGAAAAGTTGACGCTCGCAAGT**ATG**CAGTCCGAATCCGGAATAGTTCCTG
1 M Q S E S G I V P
100 ACTTTGAGGTCGGCGAGGACCTTCAGGACGAGCCTAAAACGTACTAGAGCTGAAGAGTCAACCTTTGAAAAACAGCCAAACCAACATCTCTCTCCA
10 D F E V G E D L Q D E P K T Y Y E L K S Q P L K N S P N Q Q H I S S
199 AGCCCCACTCGACTCCCGCATGACATCAGCATCTGCTACGCCAGCCTGATGAGGAGCAGCTTCAAGAACAGGAGCGGCTGAGCAGCAGC
43 K P P L S S S A M T S R I L L R Q Q L M R E Q L Q E Q E R R E Q Q
298 GGCAGCAGTCTCTCATTCCACAAACCACGCCACCCCGCCATCAATGTCTAGTGTCCCGACAGTTCGTGCGCCCTCCGACAGGTCCTCCCA
76 R Q Q S S H F P Q T T A T Q T P A I N V S V P T S S S P S A Q V P
397 TGGAGGTGCTCAAGGTTCAAGACTCACTTAGAGAACCAACCAAGTACCACATCCAACGCTCCAGCAGCAACAAGTGAACGATACCTGGGGAAGCTTG
109 M E V L K V Q T H L E N P T K Y H I Q R S Q Q Q Q V R R Y L G K L
496 GCTCCAGGCGTGAAGTTCGCTGCCCCAACCAGTCTCTGACCACGGGGCATGCCCGAGGGCCAGGCAACAGCGCCCCAACAGCCCTATGGCCT
142 G S Q A L S L P C P N Q S S D H G G M P P G P G N S A P N S P M A
595 TACTGACCTCAACTCAACTGCGAGAAAGAGATGGATGATGTCATGACGACATTATTAGTCTGGAGTCCAGTTACAGTGTGATGACATCCTTGGTCTGA
175 L L T L N S N C E K E M D D V I D D I I S L E S S Y S D D I L G L
694 TGGACCCAGGACTTCAGATGCCAAATACGATCCCTGTGCCTGCAAACTCATGGACATGTACGGTAACCCAGGGCCTCCCGCAGCAAGGTCGCCATAA
208 M D P G L Q M P N T I P V P A N L M D M Y G N Q G L P Q Q G L P I
793 GCAACTCCTGCCCTCCAAACCTACCCAACATCAAAAGGGAATACTCAGTGTCAACAATCTCTGCCATCATGCACATGTTGGATAAACTGGATCTTGTG
241 S N S C P P N L P N I K R E Y S V S Q S P A I M H M L D K S G S C
892 GCAAATTTGACACCTATCAAAGACCAGAGGGCTTCCTGTTGAAGTGTAGGTAAGAGCGTGGCAAGGAAAGACAGAAAAAGGACAACCATAATTTGA
274 G K F D T Y Q R P E G F P V E A E V R A L A K E R Q K K D N H N L
991 TTGAAAGAAGCAAGGTTCAACATCAACGATCGGATCAAGAEACTAGGAATCTGATTCAAAATCCAATGACCCAGACATGCGCTGGAACAAAAGGCA
307 I E R R R R F N I N D R I K E L G T L I P K S N D P D M R W N K G
1090 CCATCTGAAGGCATCGGTAGATTACATCCGCAAGCTGCAACGGGAAACAACAGCGACCCAAAGAGCTGGAGAACCAGCAAAAAAATGGAGCAGCCCA
340 T I L K A S V D Y I R K L Q R E G Q Q R A K E L E N R Q K K L E H A
1189 ACCGACCTGATGCTGAGAATACAGGAATGCAAGCAAGGGCTCATGGCCTGCCATCAGCTTCCAGCTGCTGCTGTTGAACCTGGGG
373 N R H L M L R I Q E L E M Q A R A H G L P I T S S A L C S V E L G
1288 TGAGAGCCATCAAGCAGGAGCCCGCTCTGGAGGACTGCCAACAGACATGTATACACTCCACCCCATCACCAACACCACCCAGCCTGCATCCAGAGC
406 V R A I K Q E P A L E D C H Q D M Y T L H P H H Q T L H P A C T P E
1387 AGCCAGGCAGCTGGAGCTCAACGACGGCCACTCAAACCTCCCTGAAGGGCACTACAACGTTTACAGCAAAGCAGGCTCCAACTCAACGACATCCTCA
439 Q P G T L E L N D G H S N F P E G H Y N V H S K A G S K L N D I L
1486 TGAAGACACCTTAGCCCTGTGAGAGGGGGGACCCCTGCTCTCGTCCGCTCTCCAGACGCTCAAAGGACAGCAGCCGTAAGGACAGCGTCAGCA
472 M E D T L A P V R G G D P L L S S V S P D A S K D S S R K S S V S
1585 TGGATGAAAACGAGCAGGGCTGTTAG
505 M D E N E Q G C *

B

1 TTTAGCCTTCTCTGAGGTTTGAGAAGCTTGCTCTTAACCTTTGAAACCCATTCAAAGGGATTACATGAACAAAACACATCCGATATTAGCAGTTTCA
100 **ATG**TTGGAGATGTTGGAATACAGCAATTACCAGGTACAATTAATTTAGAAAAATGAAAACAAGCACCACCTTCCAGCAGGCTCATAGACAGCAAGTGAGG
1 M L E M L E Y S N Y Q V Q S N L E N E N K H H F Q Q A H R Q Q V R
199 CAGTATCTGTCAACTACTTTGGTAGTAAAGCTGGGGCCAAATGTGCCATCCAGCCTTCCAGAGCATAGCTTACCACCAGGATCCGAGGGCAGCATGCC
34 Q Y L S T T L V G K A G G Q C A I Q P S E H S L P P G S E G S M P
298 AACAGCCCAATGGCTTTGCTCACCCCTCAGTTCTAATTTGTGAGAAAGAGATGAATGATGTCATTGATGATATATTAGTCTGGAGTCAAGTTCACATGAC
67 N S P M A L L T L S S N C E K E M N D V I D D I I S L E S S Y N D
397 GACGTTCTTGGACTAATGGAATCAGGACTTCAAATCAACAACAGCCTCCAGTGTCCAGTAGTTTAATAGATGTGTACAACAATCAAGGACTAACGCTG
100 D V L G L M E S G L Q I N N Q L P V S S S L I D V Y N N Q G L T L
496 TCCAACCTGCCTACCAGCCTGCTCTGTCATCTGGCATTAAAGAGAATTAACAGCTCCTGGCATGAAGCAAGTGTGGAACAAGCCTGATCTCTGTGGC
133 S N L P T S S S C A S G I K R E L T A P G M K Q V L N K P E S C G
595 CAGTATGAAAGCTATCAAAGGCCAGAGGAGTTTCCAGTAGAGACTGAGGTTAGAGCTCTTGCCAAAGAGAGACAAAAGAAAGACAACCAACTTGATT
166 Q Y E S Y Q R P E E F P V E T E V R A L A K E R Q K K D N H N L I
694 **GAA**CGAAGGCGAAGGTTCAACATTAATGACCCGATTAAGGAACCTTAGGAACCTTAATACCAAAATCCAATGATCCAGATATGCGCTGGAACAAGGGCAGC
199 E R R R R F N I N D R I K E L G T L I P K S N D P D M R W N K G T
793 ATTCTCAAAGCTTCAGTAGACTATATCAGAAAACCTGCAACAGGAGCAACAGAGAATAAGGAGCTGGAGTGCAGGCATAAGAAGCTGGAGCATGCAAAAC
232 I L K A S V D Y I R K L Q Q E Q Q R T K E L E C R H K K L E H A N
892 AGTCATCTTATGCTTCAATACAGGAACCTGAGATCCAGGCACGGGCTCATGGTTTACAGTGGCAGCCGCCCATCTATCTGCACATCAGAGCTTTTG
265 S H L M L R I Q E L E I Q A R A H G L T V A A A P S I C T S E E L L
991 GCACGAGCCATTAACAGGAGCCTATCTTGGCGAATGCCACAGAGATATACCAGCACTGCTCGGCTCCTGACATGTCTCTCAACAATAGGAC
298 A R A I K Q E P I L G E N C E P P E I Y Q H C S A P D M S P S T T M D
1090 TTAACAATGGCATCATTACTTTGACACCATTCTGCGAGAACCTGGGGATTCTAGCTCTTATGGAATCTCCAGAATCTGTAATAATGAAGGAGATGGTA
331 L N N G I I T F D T I P A E P G D S S S Y G I S R I C K M K E M V
1189 AGGGACAAAAGTTTGGGCCACTTTCCACCAAGTAAACCTTTGTTGCTTCAATACCACAGATGTTTCAACAACAGCAGTAACCACTGTAGCTCCAAT
364 R D K S F G P L S P S K P L L S S I P P D V S N N S S N H C S S N
1288 GCAAGCATGGAGGAAAAAGAGAAATGGCTGTTAAACCCCTGCTGGGAGCAACTACACCTG
397 A S M E E K E N G C *

Figure S1. Nucleotide and deduced amino acid sequence of medaka *mitf1* and *mitf2*.

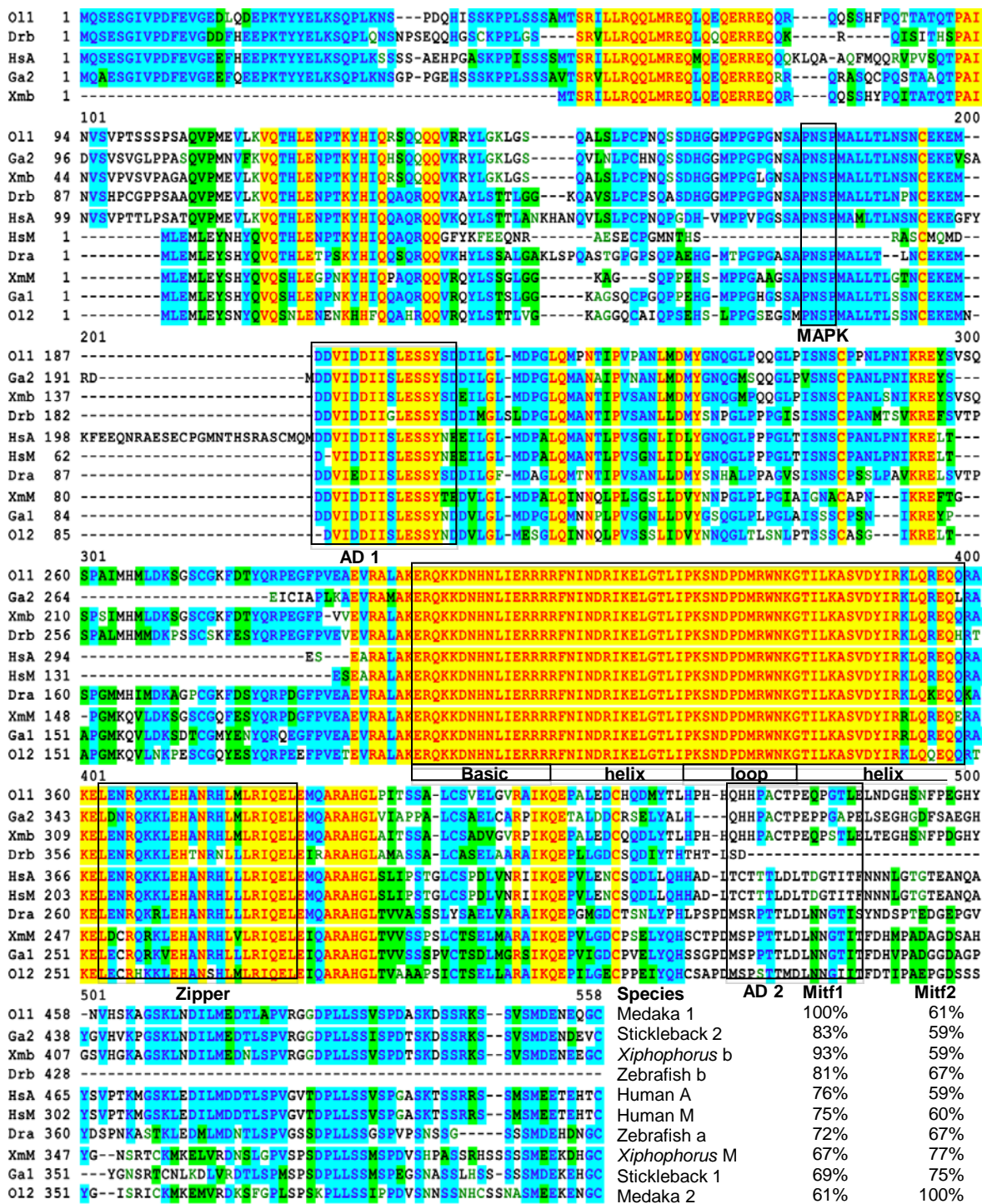


Figure S2. Mitf protein sequence alignment.

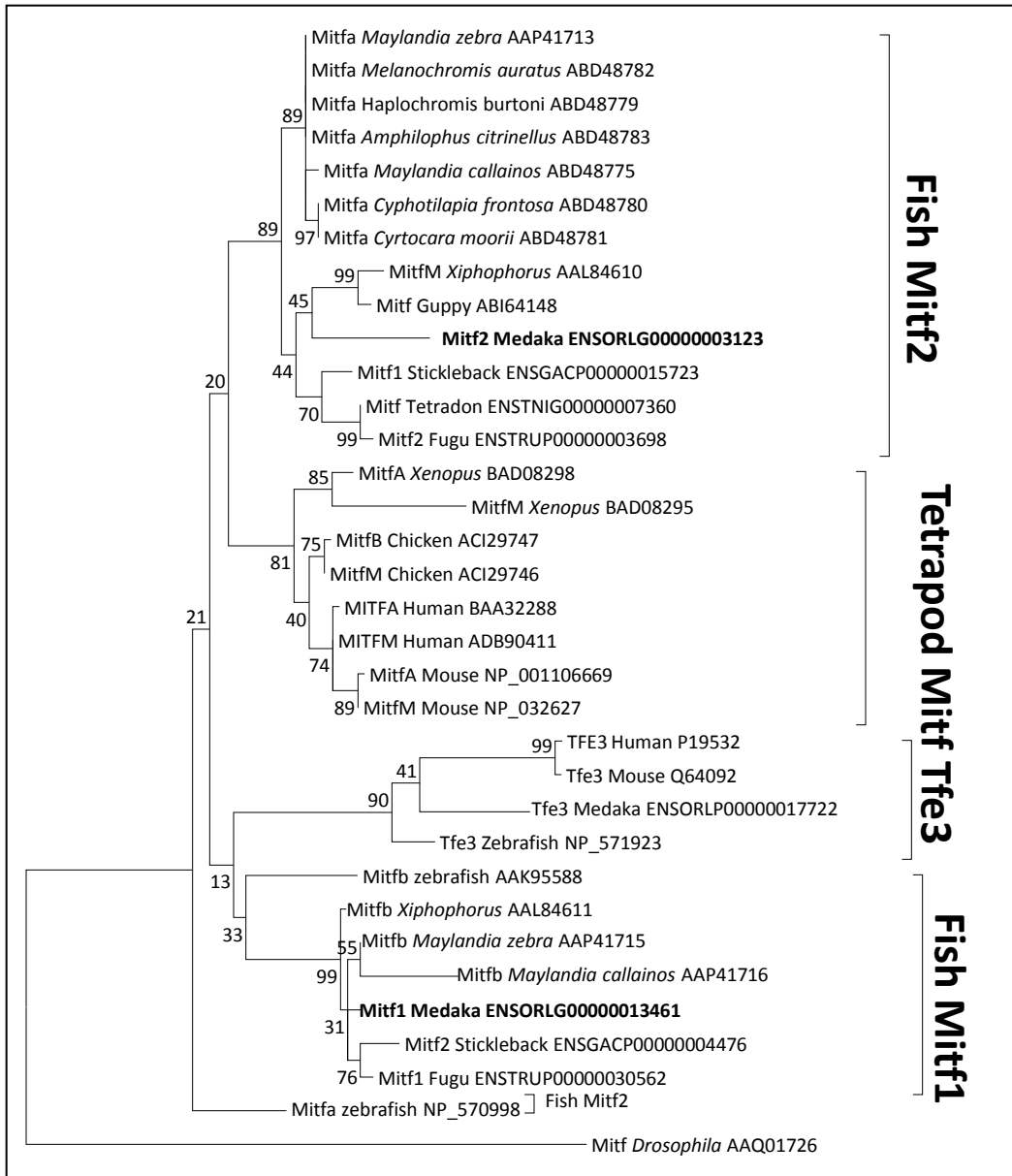


Figure S3. Phylogenetic tree construction by maximum likelihood.

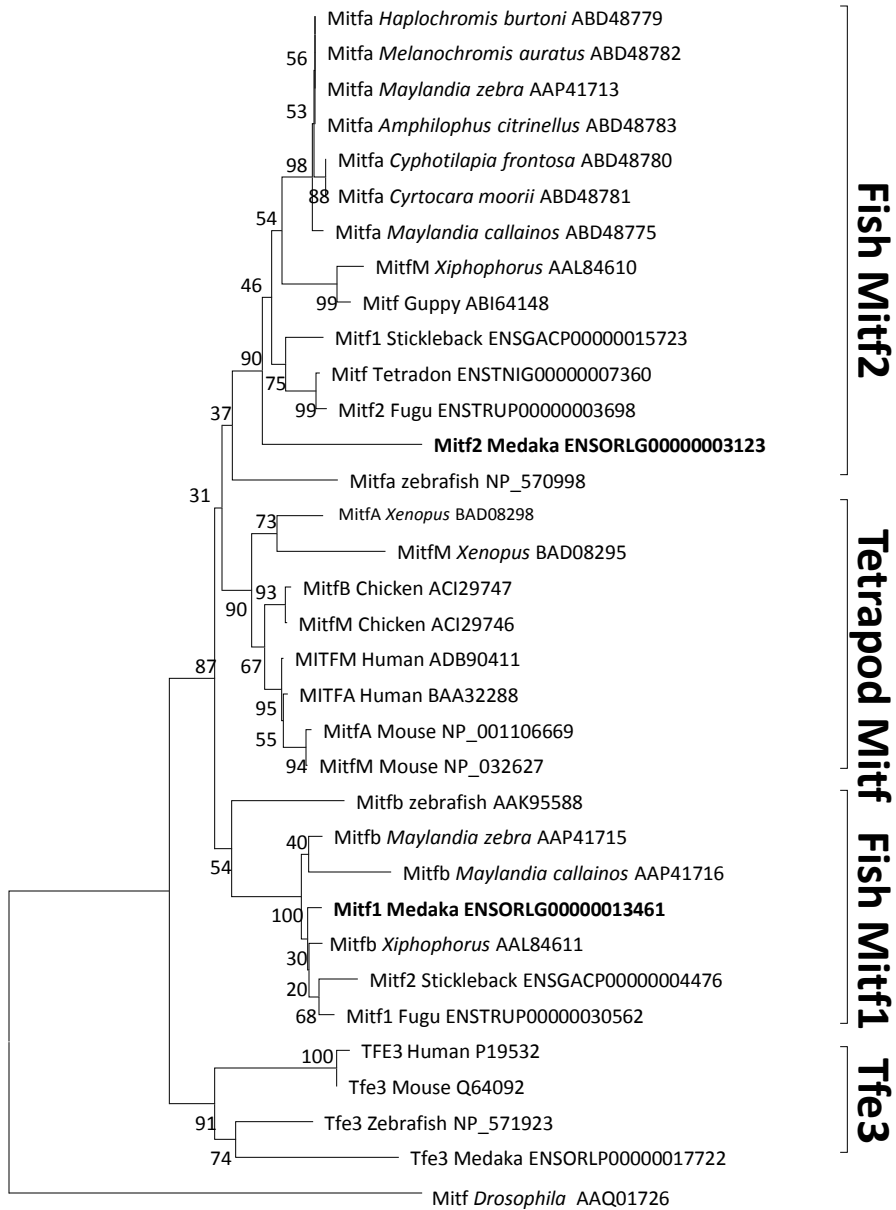


Figure S4. Phylogenetic tree construction by molecular evolution.