

Additional file 2

(A) Gene and aligned protein sequence of wild-type *xyr1/Xyr1*

```
1 ATGTTGTCCAATCCTCTCCGTCGCTATTCTGCCTACCCCGACATCTCCTCGGCGTCATTTGACCCGAAGTACCATGGCTC
1 M L S N P L R Y S A Y P D I S S A S F D P N Y H G S
81 ACAGTCGCATCTCCACTCGATCAACGTCAACGATTCGGCAACAGCCACCCCTATCCCATGCAGACCTCGCACAGCATG
27 Q S H L H S I N V N T F G N S H P Y P M Q H L A Q H
161 CGGAGCTTTTCAGTTTCACGCATGATAAGGGCCAGTCCGGTCGACGCCAAGCAGCGCCAGGGCTCTCTTATGCTGCCAGG
54 A E L S S S R M I R A S P V Q P K Q R Q G S L I A A R
241 AAGAATTCAACGGGTACTGCTGGGCCATTTCGGCGGAGGATCAGTCGCGCTTGTGACCAAGTGAACACAGCTTCGTACCAA
81 K N S T G T A G P I R R R I S R A C D Q C N Q L R T K
321 GTGCGATGGCTTACACCCATGTGCCATTGTATAG gtatgtoccttttctctacacagtgatgtgctgcgctcaagcacat
107 C D G L H P C A H C I
401 gtactgatcgatcttgttttag AATTCGGCCTTGGATGCGAATATGTCCGAGAGAGAAAGAAGCGTGGCAAAGCTTCGCGC
119 E F G L G C E Y V R E R K K R G K A S R
481 AAGGATATTGTGCCCCAGCAGCCGCGGCGCTGCAAGCACAACACTCCGGCCAGGTCCAGGATGGTCCAGAGGATCAACA
139 K D I A A Q Q A A A A A A Q H S G Q V Q D G P E D Q H
561 TCGCAAATCTTCACGCCAGCAAAGCGAATCTTCGCTGGCAGCGCTGAGCTTGCCAGCCTGCCACAGCCCGCTCATG
165 R K L S R Q Q S E S S R G S A E L A Q P A H D P P H
641 GCCACATTGAGGGCTCTGTACGCTCCTTCAGCGACAATGGCCTTTCCAGCATGTGCCATGGGCGGCATGGATGGCCTG
192 G H I E G S V S S F S D N G L S Q H A A M G G M D G L
721 GAAGATCACCATGGCCACGTCGGAGTTGATCCTGCTGGGCCGAAGCTCAGCTGGAAGCGTCATCAGCAATGGGCCCTGGG
219 E D H H G H V G V D P A L G R T Q L E A S S A M G L G
801 CGCATACGGTGAAGTCCACCCCGGCTATGAGAGCCCGGCGATGAATGGCCATGTGATGGTGGCCCGCTCGTATGGCGCGC
245 A Y G E V H P G Y E S P G M N G H V M V P P S Y G A
881 AGACCACCATGGCCGGGTATTCCGGTATCTCGTATGCTGCGCAAGCCCCGAGTCCGGCTACGTATAGCAGGACGGTAAC
272 Q T T M A G Y S G I S Y A A Q A P S P A T Y S S D G N
961 TTTCGACTCACCGGTACATCCATGATTACCCGCTGGCAATGGGAGCTCGCCCTCATGGGGAGTCTCGTGGCCCTCGCC
299 F R L T G H I H D Y P L A N G S S P S W G V S L A S P
1041 TTCGAACCAATTCCAGCTTCAGCTCTCGCAGCCCATCTTCAAGCAAAGCGATTTGCGATATCCTGTGCTTGAGCCTTG
326 S N Q F Q L Q L S Q P I F K Q S D L R Y P V L E P L
1121 TGCCTCACCTGGGAAACATCCTCCCCGTGTCTTTGGCGTGCATCTGATTGACCTGTACTTCTCCTCGTCTTCATAGCA
352 L P H L G N I L P V S L A C D L I D L Y F S S S S A
1201 CAGATGCACCAATGTCCCATACGTTCTGGGCTTCTGCTTCCGGAAGCGCTCCTTCTTGACCCCAAGACCCCAAG
379 Q M H P M S P Y V L G F V F R K R S F L H P T N P R R
1281 GTGCCAGCCCGCGCTGCTTGCAGCATGCTGTGGGTGGCGGCACAGACTAGCGAAGCGTCTTCTTGACGAGCCTGCCGT
405 C Q P A L L A S M L W V A A Q T S E A S F L T S L P
1361 CGGCGAGGAGCAAGTCTGCCAGAAGCTGCTGAGCTGACCGTTGGGCTTCTTCAGCCCTGACACACCGGCAAC
432 S A R S K V C Q K L L E L T V G L L Q P L I H T G T N
1441 AGCCCGTCTCCAAGACTAGCCCGTGTGCTGCTGCTGCGCTGGGAGTCTTGGGGTGGCCATGCCGGGCTCGCTGAA
469 S P S P K T S P V V G A A A L G V L G V A M P G S L N
1521 CATGATTCACTGGCCGCGAAACGGGTGCTTTTGGGGCCATAGGGAGCCTTGACGACGTACATCACTATGTGCACTCG
485 M D E T G A G E T G A F G A I G S L D D V I T Y V H L
1601 CCACGGTCTGCTCGGCCAGCAGTACAAGGGCGCCAGCCTGCGGTGGTGGGGTGGCGCATGGTCTCTCGCCAGAGAGCTC
512 A T V V S A S E Y K G A S L R W W G A A W S L A R E L
1681 AAGCTTGGCCGTGAGCTGCCGCTGGCAATCCACCTGCCAACCAGGAGGACGGCGAGGGCCTTAGCGAAGACGTGGATGA
539 K L G R E L P P G N P P A N Q E D G E G L S E D V D E
1761 GCACGACTTGAACAGAAACAACACTCGCTTCTGTGACGGAAGAGGAGCGGAAGAGCGACGGCGAGCATGGTGGCTCGTTT
565 H D L N R N N T R F V T E E E R E E R R R A W W L V
1841 ACATCGTCGACAGGCACCTGGCGCTCTGCTACACCCGCCCTTGTCTTCTTGACAGCGAGTGCAGCAGCTTGTACCAC
592 Y I V D R H L A L C Y N R P L F L L D S E C S L K R
1921 CCGATGGACGACATCAAGTGGCAGGCAGGCAATTTTCGAGCCACGATGCAGGGAAGTCCAGCATCAACATCGATAGCTC
619 P M D D I K W Q A G K F R S H D A G N S S I N I D S S
2001 CATGACGGACGAGTTTGGCGATAGTCCCCGGGCGGCTCGCGGCGCACACTACGAGTGCCGCGGTCTGATGCTTTTGGCT
645 M T D E F G D S P R A G G A H Y E C R G R S I F G
2081 ACTTCTGTCTTGTATGACAATCCTGGGCGAGATTGTGATGTCCACCATGCTAAAAGCCACCCCGGTTCCGGCTTGG
672 Y F L S L M T I L G E I V D V H H A K S H P R F G V G
2161 TTCCGCTCCGCGCGGATTGGGACGAGCAGGTGTGTAATCACCCGACACCTGGACATGTATGAGGAGAGCCTCAAGAG
699 F R S A R D W D E Q V A E I T R H L D M Y E E S L K R
2241 GTTCGTGGCCAAGCATCTGCCATTGTCTCAAAGGACAAGGAGCAGCATGAGATGCACGACAGTGGAGCGGTAACAGACA
725 F V A K H L P L S S K D K E Q H E M H D S G A V T D
2321 TGCAATCTCCACTCTCGGTGCGGACCAACGCTCCAGCCGATGACGGAGAGCGAGATCCAGGCCAGCATCTGTGTGGCT
752 M Q S P L S V R T N A S S R M T E S E I Q A S I V A
2401 TACAGCACCCATGTGATGCATGTCTCCACATCCTCCTTGGGATAAGTGGGATCCCATCAACCTTCTAGACGACGACGA
779 Y S T H V M H V L H I L L A D K W D P I N L L D D D D
2481 CTTGTGGATCTCGTCGGAAGGATTTCGTGACGGCAGCAGCCAGCGGTATCGGCTGCCGAAGCTATTAGCCAGATTCTCG
805 L W I S S E G F V T A C T S H A V S A A E A I S Q I L
2561 AGTTTGACCTGGCCTGGAGTTTATGCCATTCTTCTACGGCGTCTATCTCTGACAGGGTCTCTCTCTCTCTGCTCATC
832 E F D P G L E F M P F F Y G V Y L L Q G S F L L L L I
2641 GCCGACAAGCTGCAGGCCGAAGCGTCTCAAGCGTCAAGGCTTGCGAGACCATTGTTAGGGCACACGAAGCTTGCCT
859 A D K L Q A E A S P S V I K A C E T I V R A H E A C V
2721 TGTGACGCTGAGCACAGAGTATCAG gtaagccctatcagttcaaacgtctatcttctgtgtgaatcaaagactgacttgg
885 V T L S T E Y Q
2801 catcag CGCAACTTTAGCAAGGTTATGCGAAGCGCGCTGGCTCTGATTCCGGGCGGTGTGCCGGAAGATTAGCTGAGCA
894 R N F S K V M R S A L A L I R G R V P E D L A E Q
2881 GCAGCAGCGACGACGCGAGCTTCTTGCACTATACCGATGGACTGGTAACGGAACCGGTCTGGCCCTCTAA
918 Q Q R R R E L L A L Y R W T G N G T G L A L -
```

Additional file 2 (continued)**(B) Gene and aligned protein sequence of mutated *xyr1/Xyr1***

```
1 ATGTTGTCCAATCCTCTCCGTCGCTATTCTGCCTACCCCGACATCTCCTCGGCGTCATTTGACCCGAACCTACCATGGCTC
1 M L S N P L R R Y S A Y P D I S S A S F D P N Y H G S
81 ACAGTCGCATCTCCACTCGATCAACGTCAACACATTCGGCAACAGCCACCCCTATCCCATGCAGCACCTCGCACAGCATG
27 Q S H L H S I N V N T F G N S H P Y P M Q H L A Q H
161 CGGAGCTTTTCGAGTTCACGCATGATAAGGGCCAGTCCGGTGCAGCCAAAGCAGCGCCAGGGCTCTCTTATTGCTGCCAGG
54 A E L S S S R M I R A S P V Q P K Q R Q G S L I A A R
241 AAGAATTCAACGGGTACTGCTGGGCCCATTCGGCGGAGGATCAGTCGCGCTTGTGACCAGTGCAACCAGCTTCGTACCAA
81 K N S T G T A G P I R R R I S R A C D Q C N Q L R T K
321 GTGCGATGGCTTACCCCATGTGCCATTGTATAG gtatgtcccttttctctacacagtgatgtgcgtcaagcacat
107 C D G L H P C A H C I
401 gtactgatcgatcttgttttag AATTCCGGCCTTGGATGCGAATATGTCCGAGAGAGAAAGAAGCGTGGCAAAGCTTCGCGC
119 E F G L G C E Y V R E R K K R G K A S R
481 AAGGATATTGCTGCCCAGCAAGCCGCGGCGGCTGCAGCACAACACTCCGGCCAGGTCCAGGATGGTCCAGAGGATCAACA
139 K D I A A Q Q A A A A A A Q H S G Q V Q D G P E D Q H
561 TCGCAAACCTCTCAGCCAGCAAAGCGAATCTTCGCGTGGCAGCGCTGAGCTTGGCCAGCCTGCCACGACCCCGCTCATG
165 R Q L S R Q L S E S Q P A L G S A E L A Q P P A H D P P H
641 GCCACATTGAGGGCTCTGTCTCAGCTCCTTCAGCGACAATGGCCTTTCCAGCATGCTGCCATGGGCGGCATGGATGGCCTG
192 G H I E G S V S S F S D N G L S Q H A A M G G M D G L
721 GAAGATCACCATGGCCAGCTCGGAGTTGATCCTGCCCTGGGCGCAACTCAGCTGGAAGCGTCATCAGCAATGGGCGCTGGG
219 E D H H G H V G V L G F V F R K R S F L H P T N P R L
801 CGCATAACGGTGAAGTCCACCCCGGCTATGAGAGCCCCGGCATGAATGGCCATGTGATGGTCCCCCGTCTGATGGCGCGC
245 A Y G E V H P G Y E S P G M N G H V M V P P S Y G A
881 AGACCACCATGGCCGGGTATTCCGGTATCTCGTATGCTGCGCAAGCCCCGAGTCCGGCTACGTATAGCAGCGACGGTAAC
272 Q T T M A G Y S G I S Y A A Q A P S P A T Y S D G N
961 TTTTCGACTCACCGGTACATCCATGATTACCCGCTGGCAATGGGAGCTCGCCCTCATGGGGAGTCTCGCTGGCCTCGCC
299 F R L T G H I H D Y P L A N G S S P S W G V S L A S P
1041 TTGAACACAGTTCCAGCTTCTCAGCTCTCGCAGCCCATCTTCAAGCAAAGCGATTGCGATATCCTGTGCTTGAGCCTCTGC
326 S N Q F Q L Q L S Q P I F K Q S D L R Y P V L E P L
1121 TGCCTCACCTGGGAAACATCCTCCCCGTGTCTTTGGCGTGCATCTGATTGACCTGTACTTCTCCTCGTCTTCATCAGCA
352 L P H L G N I L P V S L A C D L I D L Y F S S S S S A
1201 CAGATGCACCAATGTCCCCATACGTTCTGGGCTTCGTCTTCCGGAAGCGCTCCTTCTTGACCCCCACGAACCCACGAAG
379 Q M H P M S P Y V L G F V F R K R S F L H P T N P R R
1281 GTGCCAGCCCCGCGTCTTGGCAGCATGCTGTGGGTGGCGGCACAGACTAGCGAAGCGTCCTTCTTGACGAGCCTGCCGT
405 C Q P A L L A S M L W V A A Q T S E A S F L T S L P
1361 CGGCGAGGAGCAAGGTCTGCCAGAAGCTGCTCGAGTGAACCGTGGGCTTCTTCCAGCCCCGATCCACACCGGCACCAAC
432 S A R S K V C Q K L L E L T V G L L Q P L I H T N P R L
1441 AGCCCCGTCTCCCAAGACTAGCCCCGTCTCGTGGTGTCTGCTGCCCTGGGAGTTCTTGGGGTGGCCATGCCGGGCTCGTGAA
469 S P S P K T S P V V G A A A L G V L G V A M P G S L N
1521 CATGGATTCACTGGCCGCGGAAACGGGTGCTTTTGGGGCCATAGGGAGCCTTGACGACGTATCACCTATGTGCACCTCG
485 M D S L A G E T G A F A I G S L D D V I T N P R R
1601 CCACGGTCTGCTCGGCCAGCGAGTACAAGGGCGCCAGCCTGCGGTGGTGGGGTGGCGCATGGTCTCTCGCCAGAGAGCTC
512 A T V V S A S E Y K G A S L R W W G A A W S L A R E L
1681 AAGCTTGGCCGTGAGCTGCCGCTGGCAATCCACCTGCCAACCAGGAGGACGGCGAGGGCCTTAGCGAAGACGTGGATGA
539 K L G S R E L P P G N P P A N Q E D G E G L S E D L K R
1761 GCACGACTTGAACAGAAACAACACTCGCTTCTGTCAGGAAAGAGGAGCGCGAAGAGCGACGCGAGCATGGTGGCTCGTTT
565 H D L N R N N T R F V T E E E R E E R R R A W W L V
1841 ACATCGTCGACAGGCACCTGGCGCTCTGCTACAACCGCCCTTGTCTTCTTCTGGACAGCGAGTGCAGCGACTTGTACCAC
592 Y I V D R H L C Y N P P L F L L D S E C S L K R
1921 CCGATGGACGACATCAAGTGGCAGGCGAGGCAAAATTCGAGCCACGATGCAGGGAACCTCCAGCATCAACATCGATAGCTC
619 P M D D I K W Q A G K F R S H D A G N S S I N I D S S
2001 CATGACGGACGAGTTTGGCGATAGTCCCCGGGCGGCTCGCGGCGCACACTACGAGTGCCGCGGTCTGTAGCATTTTGGCT
645 M T F E G D S P R A R G A H Y E C R F G
2081 ACTTCTTGCTCTGTATGACATCCTGGGCGAGATTGTGATGTCCACCATGCTAAAAGCCACCCCGGTTTCGGCGTTGGA
672 Y F L S L M T I L G E I V D V H H A K S H P R F G V G
2161 TTCCGCTCCGCGCGGGATTGGGACGAGCAGGTTGTGAAATCACCCGACACCTGGACATGTATGAGGAGAGCCTCAAGAG
699 F R S A R D W D E Q V A E I T R H L D M Y E S L K R
2241 GTTCGTGGCCCAAGCATCTGCCATTGTCTCAAAGGACAAGGAGCAGCATGAGATGCACGACAGTGGAGCGGTAACAGACA
725 F V A K H L P L S S K D K E Q H E M H D S G A V T D
2321 TGCAATCTCCACTCTCGGTGCGGACCAACGCGTCCAGCCGATGACGGAGAGCGAGATCCAGGCCAGCATCGTGGTGGCT
752 M Q S P L S V R T N A S S R M T E S E I Q A I V A
2401 TACAGACCCATGTGATGCATGTCTCCACATCCTCCTTCGCGATAAGTGGGATCCCATCAACCTTCTAGACGACGACA
779 Y S T H V M H V L H I L L A D K W D P I N L L D D D D
2481 CTTGTGGATCTCGTCGGAAGGATTCTGTACGGCGACGAGCCACGCGGTATCGGCTGTCGAAGCTATTAGCCAGATTCTCG
805 L W I S S E G F V C T A T S H A V S A V E A I S Q I L
2561 AGTTTGACCCTGGCCTGGAGTTTATGCTTCTTACGGGTCTATCTCCTGCAGGGTTCTTCTCCTCTCTGCTGCTATC
832 E F D P G L E F M P F F Y G V Y L L Q G S F L L L L I
2641 GCCGACAAGCTGCAGGCCGAAGCGTCTCAAGCGTCATCAAGGCTTGCGAGACCATTGTTAGGGCAGACGAAGCTTGCGT
859 A D K L Q A E A S P S V I K A C E T I V R A H E A C V
2721 TGTGACGCTGAGCACAGAGTATCAG gtaagccctatcagttcaaacgtctatcttgcgtgaatcaagactgacttgg
885 V T L S T E Y Q
2801 catcag CGCAACTTTAGCAAGGTTATGCGAAGCGCGCTGGCTCTGATTCCGGGGCCGTGTGCCGGAAGATTAGCTGAGCA
894 R N F S K V M R S A L A L I R G R V P E D L A E Q
2881 GCAGCAGCAGCAGCGAGCTTCTTGCACATACCGATGGACTGGTAACGGAACCGGTCTGGCCCTCTAA
918 Q Q R R R E L L A L Y R W T G N G T G L A L -
```