

# Polymorphisms around the *SHOX* gene of the Xp/Yp pseudoautosomal region

Six non-contiguous regions were analysed (5'→3': SEQ-6, -2, -4, -1, -35, & -7). The relevant sequences, and coordinates shown to the left, are taken from Genbank entry U82668. Exons are delineated by "+", Alu repeats are shown in yellow, other dispersed repeats in magenta, inverted repeats in green and tandem repeats in blue. The locations of PCR primers used for amplification of genomic DNA and for resequencing for SNP discovery are indicated. The alternative allele at each SNP is shown above the top strand of the sequence.

---

## *SEQ-6:*

U82668

```
-----> SEQ6A
4261  GGATCCTCAAGCCGTAGACATTTTCACGTACAATCAAAATCAACTGAAAAGGCAGGCGACG
-----+-----+-----+-----+-----+-----+
4321  TTTCTGAGGAACTTTGCTGAAAGAAAAGTCATGAGCCATAACAAACCTGGAACCGGC
-----+-----+-----+-----+-----+-----+
4381  TTCTTAACCCCAACCGCAAACCTCCCAAGCTTCGGCGTGTGTTGAAAAACAGAAAGACAG
-----+-----+-----+-----+-----+-----+
4441  CCTTCTCTAAATAGGATCCACATCTTCATCTTTCTCTCAGGACCCCAAAGACACCTCCTG
-----+-----+-----+-----+-----+-----+
4501  GCTGTTTCTTGCTGGAGAGAAAGCATTTCACGTGCAGTACACCCTGTTTGAAACATCGCT
-----+-----+-----+-----+-----+-----+
          G +6/610
4561  TGGGTATTTGATAACATTGTTTTGTTGTAGTCAGAAACAGCGCATCAGAAAACACAGTTG
-----+-----+-----+-----+-----+-----+
          +6/597      T
4621  TACAAATACTCTTTTAAATGGTACTCGGTGGGGGCTGTGTGTGTGTTTTAAATAACAAA
-----+-----+-----+-----+-----+-----+
```

ACTGCTAAACTCCATGGGAATTACTCAAATGAATATCAGGGAAAGGTTTCCTTCGTTAC  
 4681 -----+-----+-----+-----+-----+-----+-----+

TTAAACAGATGCCTTTGCAACCTACTATTTTATTTTATTATGTCATTTCTATGGGAAGAG  
 4741 -----+-----+-----+-----+-----+-----+-----+

ATTATTTAAAAGAAGAAAAGGGGCCCCAGGTTTTCTCTCCTAAATTTCTAGCCCCAAACC  
 4801 -----+-----+-----+-----+-----+-----+-----+

TTCGTGATATTGAAAAGCAAGTATATCACAAAAGACTGTCTTCTACAAATTATATTTTAA  
 4861 -----+-----+-----+-----+-----+-----+-----+

ATTGCAGTGTGCGTGTGTGTTTTAAATAACAAAAGCTGCTAAACTCCATGGGAATTAGTC  
 4921 -----+-----+-----+-----+-----+-----+-----+

AAAATGAATATCGGGGAAGGGTTTCCTTCGATACTTAAATAGGTACCTTTGCAACCTACT  
 4981 -----+-----+-----+-----+-----+-----+-----+

A +6/1127  
 ATTTTATTTTATTATGTCATTTCTATGGGAGGAGATTATTTAAAAGGAGGAGCTCACGTT  
 5041 -----+-----+-----+-----+-----+-----+-----+

TTTTCGTACGAAATTTCTAGCCCCAACCCCTCCATGATACTGAAAAGCAAGTATATCAGAAA  
 5101 -----+-----+-----+-----+-----+-----+-----+

AGACTGTGTTCTACAAATTATGTTTTTAATTGCACTGTGTGTGTGTGTGTGTGTGTGTGT  
 5161 -----+-----+-----+-----+-----+-----+-----+

GTTTTAAATAACAAAAGCTGCTAAACTCCATGGGAATTACTCAAATGAATATCAGGGAA  
 5221 -----+-----+-----+-----+-----+-----+-----+

C +6/1323  
 AGGTTTCCTTCGTTACTTAAACAGATGCCTTTGCAACCTACTATTTTATTTTATTATGTC  
 5281 -----+-----+-----+-----+-----+-----+-----+

ATTTCTATGGGAGGAGATTATTTAAAAGGAGGAGCCCACTTTCTTGCTTTTAAATTTCCA  
 5341 -----+-----+-----+-----+-----+-----+-----+

GCCCAACCCTTCTGGACATAACTCGTCTTCCTCTGCAGATTGGAACAGATGATTTTGCC  
 5401 -----+-----+-----+-----+-----+-----+-----+

GTCCAAGCATCCACCGGTGCTGTAGGAGAAGACGCACCGTGTAGGTAGGTCAGCTGGCT  
 5461 -----+-----+-----+-----+-----+-----+-----+

CTGAGTTGAAAGTGTACCCAAACCCAGGAGATCCAGGGCAGAAATTGTTTTAGTTCCAC  
 5521 -----+-----+-----+-----+-----+-----+-----+

←----- SEQ6C

5581 AGAGCAGCCACGGGTGCACCGTCTGTGTATCTCTGTCAGTGATGCAATCCAATGCATCTG  
 -----+-----+-----+-----+-----+-----+-----+  
 5641 GCAAAGGGAACGCAGGGGGCTGCGAGGAAGGCCATTCGAGTGGCCTGTAAAAATGCTGA  
 -----+-----+-----+-----+-----+-----+-----+  
 5701 AAGCTCCATACACCTTCCCTGTCTCTGGGATTCTCAGCATCCTGCTCTGCCCTCCAAGTA  
 -----+-----+-----+-----+-----+-----+-----+  
 T +6/1808  
 5761 CAGGTCGCTGACCCACCTCCCACCCCTGCCACCTCCAATCACATGCCTTGCACCCCTTC  
 -----+-----+-----+-----+-----+-----+-----+  
 5821 CTACCATTAGGCCTGCCGCTAAGGACCTGATTAGCTATAAATTTATGGCAATCCGCAACC  
 -----+-----+-----+-----+-----+-----+-----+  
 5881 ACCGGCCTTCCCTCCTGATGTTTTCTGACACATTTTCAGGAGACGCAGAATACAGGCTCG  
 -----+-----+-----+-----+-----+-----+-----+  
 5941 CTGGTCCACCCGAGACCCCTGGCTGCAGAAACAGCCCGGCTCGTGAACACAGTTGCCGTG  
 -----+-----+-----+-----+-----+-----+-----+  
 6001 GTAACATGCCCCCAGCTGTAGCCGGCCTAGGAAAACACTACGTGGACAAGGATGATTTTTTT  
 -----+-----+-----+-----+-----+-----+-----+  
 6061 CACAAGTTTTATTTTTTATTTTACTTGGGACCTTTTGTGGCAGCCGGACGGTGCA  
 -----+-----+-----+-----+-----+-----+-----+  
 6121 GCTTTTTCTCTGCACGGAGTCTGGTTGTGTCTACACACACAGGCAACACGCATAGATGGA  
 -----+-----+-----+-----+-----+-----+-----+  
 6181 CATAAACTATCCCCTCCCAAGACTTGCAAAGGATGCACGTTTTTTCAGGCTGCAGCCAGG  
 -----+-----+-----+-----+-----+-----+-----+  
 <----- SEQ6B

---

## SEQ-2:

**U82668**

-----> SEQ2A  
 13161 AAGGGAGATGCCTCCAAAAGGAGGAAGTACTCTCTTTGACCAACGTACATGATAGCAA  
 -----+-----+-----+-----+-----+-----+-----+  
 13221 CCTGCCAGGAACCAGGAACCAGAGACACGTATCCAACAAAAGGCACCTGCCTCTTCCCC  
 -----+-----+-----+-----+-----+-----+-----+

```

AGCTCCTGCTGGTTTTGTGGGGCAACCCTGGGAGTCCTTGTCTTGCAAATGTGTTGTCTAG
13281 -----+-----+-----+-----+-----+-----+
CTCTCCTCCTGAATGAGGCATCACCTGCATTCTCCACCTGCGTTCTTTGGAAGAGCTTTG
13341 -----+-----+-----+-----+-----+-----+
CAGCCTCTCAGGAGAGCTGCTCACCTGTCCCCCTGCCTGCTCATTGCCCTGTGCAAACC
13401 -----+-----+-----+-----+-----+-----+
AGAGTCCTGAGACACTCCCCAAAGGTTGGTCCCTACCAACCCAGGTTGCCTCTTATTTCT
13461 -----+-----+-----+-----+-----+-----+
TGCAACAAGACAGCAAAGACGCTGCAGAAGGATCGCTTACCTTGTTTGGAAAGAGAGCC
13521 -----+-----+-----+-----+-----+-----+
CACCCTTTCACCTCGCTGGCAAACATTTTCTTGTGTCCTAATGAGAACCCCAAGGAATAC
13581 -----+-----+-----+-----+-----+-----+
-----+-----+-----+-----+-----+-----+
CCC GGGATGTGAGGAGAGGGAGCGTCCGCCCGGCGGCCCGCCCAAGGAGCCTGAGGCTG
13641 -----+-----+-----+-----+-----+-----+
> SEQ2C
TCCTGCACACATCACCACACACTCTCTGCATTGCAGTCCCATGTTTATAGAGAAGATT
13701 -----+-----+-----+-----+-----+-----+
ACAGGTGAGTTTCCAAAATCCCAGCCCATGCTGCCCTACCTGACGATGTTTGATTTCGG
13761 -----+-----+-----+-----+-----+-----+
AATTCCTTTAATTTTCAGCAAATGGATATTGGCTGTGGGTCAGCATAGATTTTTTCCGGTT
13821 -----+-----+-----+-----+-----+-----+
ATTTAAACAACATTTTCAGATGCTTTTCTGGCATCATTAGCTGAATGCTTCTACTGACAG
13881 -----+-----+-----+-----+-----+-----+
CTGGAATCATTCAAACATCTGAGACAAGAGAGTGAACAGGACGTTCACTTGTTACCGTGT
13941 -----+-----+-----+-----+-----+-----+
CTCCTCGCTTCTCAACCACCACCCCTGAAATTGCGAGACTGTCACTCTGGGGAGCCTGTG
14001 -----+-----+-----+-----+-----+-----+
AACGTGGCTGAACGCAAACGGATTTAAAACCTTGACAGGTGTTTCCAGCATGGAAAGGAAC
14061 -----+-----+-----+-----+-----+-----+
AAGGACTATAACGGCGAATAATAATAATAATAATAAAAAGACAAATTAGGGATCAAGATCC
14121 -----+-----+-----+-----+-----+-----+
AGTCTGTTTTTGTCTCATTATAGTCTTTTCTCCAAGAGGTTGTAGGCTGGTACCATTTC
14181 -----+-----+-----+-----+-----+-----+

```

ACTAGGTTAAAAGCCTACGGTGTCAAGAGGTGCTAAGCTTTGCTGCTGTCTCTTGCTTGG  
 14241 -----+-----+-----+-----+-----+-----+-----+  
 CAAGACAAGCTCCGAGTCTTTCTGGAAGTACAAATTGTGTGGGGCATTTCCTTGTCTTCTG  
 14301 -----+-----+-----+-----+-----+-----+-----+  
 GGGTTCCTTGCATGCATAGGGGCTGATCTTTGATATCCACTCTGGACTCCTCTCTGCAAA  
 14361 -----+-----+-----+-----+-----+-----+-----+  
 CCTCTGGACATGCCACAGACAGCAGCTGTAACGTAGTGGGGTCAGCCTGCTTACAACCCA  
 14421 -----+-----+-----+-----+-----+-----+-----+  
 CGTACCAACCCCATGGAGGTGATTGTTCGTATGACTAACAGCAGGCAGTTAAGGAGAAAGT  
 14481 -----+-----+-----+-----+-----+-----+-----+  
 CATGCTTTACAACGAGTCTGTGACTAGAAATCACAGGACCTGATGAATCCCTCAGAAAC  
 14541 -----+-----+-----+-----+-----+-----+-----+  
 CCAAGAGGGATGGACTAACATTAAGGGGGTTGTGGACTGTCCAGCCGTGCAGGGACGGTC  
 14601 -----+-----+-----+-----+-----+-----+-----+  
 TCCAAGAAGATGCTTTAGACACAACGGTGAATTCTGCAGGCGCTAGTTTGCATGCATATT  
 14661 -----+-----+-----+-----+-----+-----+-----+  
 <----- SEQ2B

---

## SEQ-4:

**U82668**

-----> SEQ4A  
 AAAGTGGGCCAGTCTCAGAGTTCCTTCCTCTTCGCAGCCATCCTAAGCGGTTCTCTCGAG  
 36208 -----+-----+-----+-----+-----+-----+-----+  
 TTTCCCCCAGCAGGTGCGTTCTAGAAAGCGCTGGTCAGGACGCATGGCTGCCCCATAGA  
 36268 -----+-----+-----+-----+-----+-----+-----+  
 TTTCTTGGGGACGGGAGTGAGAGACGCCTGGACGCCCAAGCGCAAGCCTCCAGCCTCCAA  
 36328 -----+-----+-----+-----+-----+-----+-----+  
 GGTCTACAACAGATACAGCAGCTACTGAGTCTCCTCTCTGGTTTCTTTTCCGCCCTGG  
 36388 -----+-----+-----+-----+-----+-----+-----+

GGAGAGAAAAGAGAAGGGGAGAAACGTCACCACCAGCAACTGCTCCCAGCCAGCAAGACG  
 36448 -----+-----+-----+-----+-----+-----+-----+

T +4/375

CACGGCGGGGAGGCCCGGGGCTGGCCGCGTCCTGCAGCTCCTGAGCTGCGTGTCTGAGCTT  
 36508 -----+-----+-----+-----+-----+-----+-----+

TTCCCAGCCCCGCCTGCGTCTCCAGGGGCGAAATAAGAAGCAAATCTTGCTGGGACCGAG  
 36568 -----+-----+-----+-----+-----+-----+-----+

GGGG +4/500L

TTCAGGTCGAGGAACTCCGGGACCCTCTGCAGCCTCAGCCCTGCCGCGCTCCCCGAGCT  
 36628 -----+-----+-----+-----+-----+-----+-----+

TCGGGCCTGTTGCGGGGAGTTGGGGAGGAATTCAGGGAAGGGGGGAACCAGGCCGGGATA  
 36688 -----+-----+-----+-----+-----+-----+-----+

-----> OX-M

AGAGGCTNCCGGCTTGCTGTGGGCTGCATGACGGGGTGCAAAGGCGAGGAGAAGCCCCTG  
 36748 -----+-----+-----+-----+-----+-----+-----+

GGATCCTTGCTTGTGTTGCTGCGGGACCCCCAGCCCCGGCGGGCGGACCGCGGAGAGCGC  
 36808 -----+-----+-----+-----+-----+-----+-----+

TGTCCCGGATCCGGAGATTCGCTTTTCTTTGCGACTGCGGGGCTGGCGCGCGCGCGGA  
 36868 -----+-----+-----+-----+-----+-----+-----+

GGGGTCAGCACCCCTGGGCATTGGGGTCAATGCAACCAGGGCAGGTCCCTGGGGGGCCCCC  
 36928 -----+-----+-----+-----+-----+-----+-----+

† EXON I (untranslated)

AGGAGCCAATAGGGGTTTTTGTAGTCACCCTGCGGCCGCTGCTTTTGCCCGGGTCTGAG  
 36988 -----+-----+-----+-----+-----+-----+-----+

AACAGGGGCTCCCCACACTTTGTTTTTTTTTTTTTGGTTTTGTTTTATTTTCGTTTCCGCGC  
 37048 -----+-----+-----+-----+-----+-----+-----+

GTCTCTTTCTACTGCAAACAGAAATGGGAGGGTGGACAGGCGGGTAGGAGCGGATCAGAC  
 37108 -----+-----+-----+-----+-----+-----+-----+

C +4/1063

GCCCAGGACGCAGCAGCCCCGAGTCCGCACAGGGTTTGCGGGAGGTGGTGACCGCGCTGGG  
 37168 -----+-----+-----+-----+-----+-----+-----+

†

GACGCCAGGACGCGAATGAACCTCCGGGGCGCGCTCGGGGCCTGCGCTCAGAGCTTGGGT  
 37228 -----+-----+-----+-----+-----+-----+-----+

<-----

37288 AGCTCGGTTTTGGAACCTTCCTTTTCTTTTCCAGAAAGCTTTGCCTTCTTTCTTTCTTC  
 -----+-----+-----+-----+-----+-----+-----+  
 ----- SEQ4B

---

## *SEQ-1 (recombination assay interval):*

**U82668**

OX-A ----->

42331 ACCCACCCTCACTCATGTCCCCCACTGCTGTGCCATCTCACACAAGTTCACAGCTCAGC  
 -----+-----+-----+-----+-----+-----+-----+

42391 TGTCACTCCTGGGTCCCCAGGCCCGCCGGGGAGGAAGATGCGCCGTGGGGTTACGGGAGG  
 -----+-----+-----+-----+-----+-----+-----+

42451 AAGGGGACTCCGGGCCTCCTGGTGCCCCACTTTATTTGCAGAAGGTCCTTGGCAGGAACC  
 -----+-----+-----+-----+-----+-----+-----+

Ch3 ----->

42511 GTGACGCGTTTGGTTTCCAGGACTTGGAAAACGAATTTTCAGGTCGCGATGGCGAGCACCG  
 -----+-----+-----+-----+-----+-----+-----+

42571 GCTTCCCCTGAAGCACATTCAATAGCGAGAGGCCGGGAGGGAGCGAGCAGGAGCATCCCAC  
 -----+-----+-----+-----+-----+-----+-----+

S1 ----->

42631 CATGAAAACCAAAAACACAAGTATTTTTTTTACCCCGGTAAATACCCAGACGCCAGGGTG  
 -----+-----+-----+-----+-----+-----+-----+

42691 ACAGCGCGGCGCTAAGGGAGGAGGCCTCGCGCCGGGGTCCGCCGGGATCTGGCGCGGGCG  
 -----+-----+-----+-----+-----+-----+-----+

42751 GAAAGAATATAGATCTTTACGAACCGGATCTCCCGGGGACCTGGGCTTCTTTCTGCGGGC  
 -----+-----+-----+-----+-----+-----+-----+

6I1 (95% homology with 6I2)  
 +701 T

42811 GCTGGAGACCCGGGAGGCGGCCCGGGGATCCTCGGCCTCGCCGCGCGCTCCCAAG  
 -----+-----+-----+-----+-----+-----+-----+

42871 CGCCCGCGTCCCGGTTTGGGGACACCCGGCCCTTCTTCTCACTTTGCGGGATTCTCCAG  
 -----+-----+-----+-----+-----+-----+-----+

<----- S17

42931 CCGCGTTCATCTCACCAACTCTCCATCCAAGGGCGCGCCGCCACCAACTTGGAGCTCAT  
 -----+-----+-----+-----+-----+-----+-----+

+876 A

42991 CTTCTCCCAAGATCGTGCCTCCCCGGGGCGCCGGGTCCCCCCCCCTCGCCATCTCAACCC  
-----+-----+-----+-----+-----+-----+-----+

43051 CGGCGCGACCCGGGCGCTTCTGGAAAGATCCAGGCGCCGGGCTCTGCGCTCCTCCCGGG  
-----+-----+-----+-----+-----+-----+-----+

43111 AGCGAGGGCGGCCGGACGACTGGGACCCTCCTCTCTCCAGCCGTGAACTCCTTGTCTCTC  
-----+-----+-----+-----+-----+-----+-----+

43171 TGTCTCTCTCTGCAGGAAAAGTGGAGTTTGTCTTTTCTCCGGCCACGGAGAGAACCGGG  
-----+-----+-----+-----+-----+-----+-----+

+1096 A

43231 TAACCTGTGTGGGGGGCTCGGGCGCCTGCGCCCCCTCCTGCGCGCGCTCTCCCTTCC  
-----+-----+-----+-----+-----+-----+-----+

43291 AAAAATGGGATCTTTCCCCCTTCGCACCAAGGTGTACGGACGCCAAACAGTGATGAAATG  
-----+-----+-----+-----+-----+-----+-----+

<----- S18

43351 AGAAGAAAGCCAATTGCCGGCCTGGGGGGTGGGGGAGACACAGCGTCTCTGCGTGCGTCC  
-----+-----+-----+-----+-----+-----+-----+

43411 GCCGCGGAGCCCGGAGACCAGTAATTGCACCAGACAGGCAGCGCATGGGGGGCTGGGCGA  
-----+-----+-----+-----+-----+-----+-----+

EXON II †

43471 GGTCGCCGCGTATAAATAGTGAGATTTCCAATGGAAAGGCGTAAATAACAGCGCTGGTGA  
-----+-----+-----+-----+-----+-----+-----+

43531 TCCACCCGCGCGCACGGCCGTCCTCTCCGCGCGGGGAGACGCGCGCATCCACCAGCCCC  
-----+-----+-----+-----+-----+-----+-----+

start of translation

43591 GGCTGCTCGCCAGCCCCGGCCCCAGCCATGGAAGAGCTCACGGCTTTTGTATCCAAGTCT  
-----+-----+-----+-----+-----+-----+-----+

612

43651 TTTGACCAGAAAAGCAAGGACGGTAACGGCGGAGGC GGAGGCGGCGGAGGTAAGAAGGAT  
-----+-----+-----+-----+-----+-----+-----+

43711 TCCATTACGTACCGGAAGTTTTGGAGAGCGGACTGGCGCGCTCCCGGGAGCTGGGGACG  
-----+-----+-----+-----+-----+-----+-----+

<----- S2

43771 TCGGATTCCAGCCTCCAGGACATCACGGAGGGCGGCGGCCACTGCCCCGTGCATTTGTTC  
-----+-----+-----+-----+-----+-----+-----+



AAGGACCACGTAGACAATGACAAGGAGAAACTGAAAGAATTTCGGCACCGCGAGAGTGGCA  
 43831 -----+-----+-----+-----+-----+-----+-----+

†  
 GAAGGTAAGTTCCTTTGCGCGCCGGCTCCAGGGGGCCCTCCTGGGGTTCGGCGCCTCCT  
 43891 -----+-----+-----+-----+-----+-----+-----+

CGCCACGGAGTCGGCCCCGCGCGCCCTCGCTGTGCACATTTGCAGCTCCCGTCTCGCCA  
 43951 -----+-----+-----+-----+-----+-----+-----+

S3 ----->  
 GGGTAAGGCCCGGGCCGTCAGGCTTTGCCTAAGAAAGGAAGGAAGGCAGGAGTGGACCCG  
 44011 -----+-----+-----+-----+-----+-----+-----+

ACCGGAGACGCGGGTGGTGGGTAGCGGGGTGCGGGGGACCCAGGGAGGGTTCGCAGCGGG  
 44071 -----+-----+-----+-----+-----+-----+-----+

GGCCGCGCGCTGGGCACCGACACGGGAAGTCCCGGGCTGGGGTGGATCCGGGTGGCTG  
 44131 -----+-----+-----+-----+-----+-----+-----+

<----- S4

TGCCTGAAGCCGTAGGGCCTGAGATGCTTTTTTCATTTTCTTTTTCTTTTCCTTTT  
 44191 -----+-----+-----+-----+-----+-----+-----+

TTTGTGTTGTTGTTGTTGTTTTTTGTGTTGAGACAGAGTCTCGCTCTGTCCCCAGGCTGGAGTGC  
 44251 -----+-----+-----+-----+-----+-----+-----+

+2190 T  
AGTGGTTCGATCTCGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCC  
 44311 -----+-----+-----+-----+-----+-----+-----+

TCAGCCTCCCCAGTAGCTGGGATTACAGGCATGCACCACCACGCCTGGCTAATTTTTGTG  
 44371 -----+-----+-----+-----+-----+-----+-----+

CTTTTAGTAAAGACGGGGATTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCA  
 44431 -----+-----+-----+-----+-----+-----+-----+

+2378 C  
GGTGATCCACCCGCTCGGCCTCCCAAAGTGCTGGGATGACAGGCGTGAGGCACCGCGCC  
 44491 -----+-----+-----+-----+-----+-----+-----+

CGGCCTGGGTCCTGACGGCTTAGGATGTGTGTTTCTGTCTCTGCCTGTCTGCCTTGTATT  
 44551 -----+-----+-----+-----+-----+-----+-----+

<----- S19

TACGGTCACCCAGACGCACAGAGGAGCCGTCTCCACGCGCCTTCCCAGCGCTCAGCGCCT  
 44611 -----+-----+-----+-----+-----+-----+-----+

44671 GCCGGGCCCCCGGAGATCACGGGAAGACTCGAGGCTGCGTGGTAGGAGACGGGAAGGCC  
 -----+-----+-----+-----+-----+-----+  
 44731 CGGGTCAGCTCGGTTCTGTTTCCTTTAAGGAACCCCTTCATTATTATTTTCATTGTTTTCT  
 -----+-----+-----+-----+-----+-----+  
 44791 TTGAACGTCGAGGCTTGATCTTGGCGAAAGCTGTTGGGTCCATAAAAAACCACTCCCGTGA  
 -----+-----+-----+-----+-----+-----+  
 OX-C ----->  
 <----- OX-CR  
 44851 GCGGAGGTGGCCGGGATCTGGATGGGGCGCGAGGGGCCCGGGGAAGCTGGCGGCTTCGC  
 -----+-----+-----+-----+-----+-----+  
 44911 GGGCGCGTCCTAAGTCAAGGTTGTCAGAGCGCAGCCGGTTGTGCGCGGCCCGGGGAGCT  
 -----+-----+-----+-----+-----+-----+  
 +2784 A  
 <----- OX-B  
 44971 CCCCTCTGGCCCTTCCTCCTGAGACCTCAGTGGTGGGTGCGTCCCGTGGTGGAAATCGGGG  
 -----+-----+-----+-----+-----+-----+  
 45031 AGTAAGAGGCTCAGAGAGAGGGGCTGGCCCCGGGGATCTCTGTGCACACACGACAAGTGG  
 -----+-----+-----+-----+-----+-----+  
 +2900 A  
 45091 GCGGCATACATCTTAAGAATAAAATGGGCTGGCTGTGTGCGGGGCACAGCTGGAGACGGCT  
 -----+-----+-----+-----+-----+-----+  
 45151 ATGGACGCCTGTTATGTTTTTCATTACAAAGACGCAGAGAATCTAGCCTCGGCTTTTGCTG  
 -----+-----+-----+-----+-----+-----+  
 45211 ATTCGCAGAGTTGAGGTGCGAGGGTGAATGCCCAAAGGTAATTCTTCTAAGACTCTGG  
 -----+-----+-----+-----+-----+-----+  
 45271 GGCTACCTGCTCTCCGGGGCCCTGCATTTGGGGTGTGGAGTGGCCCCGGGAAATAGCCCT  
 -----+-----+-----+-----+-----+-----+  
 45331 TGTATTCGTAGGAGGCACCAGGCAGCTTCCCAAGGCCCTGACTTTGTGCAAGCAGAAAGC  
 -----+-----+-----+-----+-----+-----+  
 45391 TGTGGCTACGGTTTACAAAGCAGTCCCCGGTTTCTGACCGTCTAAGAGGCAGGAGCCCAG  
 -----+-----+-----+-----+-----+-----+  
 45451 CCTGCCTTTGACAGTGAGAGGAGTTCCCTCCCTACACACTGCTGCGGGCACCCGGCACTGT  
 -----+-----+-----+-----+-----+-----+  
 45511 AATTCATACACAGAGAGTTGGCCTTCCTGGACGCAAGGCTGGGAGCCGCTTGAGGGCCTG  
 -----+-----+-----+-----+-----+-----+

CGTGTAATTTAAGAGGGTTCGCAGCGCCCGGCGCCGCTTCTGTGGGGTTGCTTTTTTGGT  
 45571 -----+-----+-----+-----+-----+-----+-----+

TGTCTTCGCAGACACCGTTTTGCTCCTCTGAACTCTCTCTTCTCCCCCTGGCCGTGGAC  
 45631 -----+-----+-----+-----+-----+-----+-----+

+3571 C  
 CCGGGAGAGCAAAGTGTCTCCAGACCTTTTGAAAGTGAGAGGAAAATAAAGACCAGGCC  
 45691 -----+-----+-----+-----+-----+-----+-----+

AAAGACCCAGGGCCACAGGAGAGGAGACAGAGAGTCCCCGTTACATTTTCCCCTTGGCTG  
 45751 -----+-----+-----+-----+-----+-----+-----+

<----- S5

GGTGCAGAAAGACCCCCGGGCCAGGACTGCCACCCAGGCTACTATTTATTCATCAGATCC  
 45811 -----+-----+-----+-----+-----+-----+-----+

AAGTTAAATCGAGGTTGGAGGGCAGGGGAGAGTCTGAGGTTACCGTGGAAGCCTGGAGTT  
 45871 -----+-----+-----+-----+-----+-----+-----+

+3837 T  
 TTTGGGAACAGCGTGTCCCCGCCGAGCCTGGGAGCCCGTGGGTTCTGCAAAGCCTGC  
 45931 -----+-----+-----+-----+-----+-----+-----+

TGTGTTGAGGACTTTGAAGACCAGTTTGTGAGTTGGGCTCAATTCTGGGGTTCAGACTTA  
 45991 -----+-----+-----+-----+-----+-----+-----+

+3955 A  
S6 -----  
 GAGAAATGAAGGAGGGAGAGCTGGGGTCGTCTCCAGGAAACGATTCACTTGGGGGGAAGG  
 46051 -----+-----+-----+-----+-----+-----+-----+

+3983 G      A    +3990  
----->  
 AATGGAGTGTCTTGCAGGCACATGTCTGTTAGGAGGTGAAACAGAATGTGAAATCCACG  
 46111 -----+-----+-----+-----+-----+-----+-----+

TTGGAGTAAGCGTCCAGCGCTGAATGTAGCTCGGGGTGGGGTGGGAGGGCCCTGGTGTGG  
 46171 -----+-----+-----+-----+-----+-----+-----+

<----- S7

ATCGTGGAAGGAAGAAAGACAGAACAGGGTGCTAGTATTTACCCCGTTCCCTGTAGACAC  
 46231 -----+-----+-----+-----+-----+-----+-----+

CCTGGATTTGTCAGCTTTGCAAGCTTCTTGGTTGCAGCGGCCTTGCCCTGTGCCCTTTGA  
 46291 -----+-----+-----+-----+-----+-----+-----+

GACTGTTTCCAGACTAAACTTCCAAATGTCAGCCCCTTACCCTTGACAGCAAGGGACATC  
 46351 -----+-----+-----+-----+-----+-----+-----+

TCATTAGGGCATCGCGTGCTTCTCATCTGTGCTCAGCAGGCCCGAGATAGGAACAGAGGG  
 46411 -----+-----+-----+-----+-----+-----+-----+

GCGTTGGAGATGCCACTTCCACCAGCCCTGGGTTGAAGGGGAGCGAGGGAGACACCTTTT  
 46471 -----+-----+-----+-----+-----+-----+-----+

ACTTAAACCCCTGAGCTTGGTCAGAGAGGCTGAATGTCTAAAATGAGGAAGAAAAGTTT  
 46531 -----+-----+-----+-----+-----+-----+-----+

TTCACCTGGAAACGCTTGAGGGCTGAGTCTTCTGCCCTTCTGACTCCCCAGCAAATACA  
 46591 -----+-----+-----+-----+-----+-----+-----+

GACAGGTCACCAACTACTGGAGATGAGAAAGTGCCATTTTTGGCACACTCTGGTGGGGTA  
 46651 -----+-----+-----+-----+-----+-----+-----+

GGTGCCCCACCGCGTGTGAAAAAGTGGGAAGGAGAGATTTCTGCGCACGCGGTTTCAGCCC  
 46711 -----+-----+-----+-----+-----+-----+-----+

CCAGGCGCGGTGGCGCATTTCAGGTA CT CAGACGCGGTTCTGCTGTTCTGCTGAGAAACAG  
 46771 -----+-----+-----+-----+-----+-----+-----+

S20 ----->  
 GCTTCGGGTAGGGGCTCCTAGCTCCGCCAGATCGCGGAGGGACCCCCAGCCCTCCTGCGC  
 46831 -----+-----+-----+-----+-----+-----+-----+

TGCAGCGGTGGGGATAGCGTCTCTCCGTAGGCCTAGAATCTGCAACCCGCCCCGGGTCTT  
 46891 -----+-----+-----+-----+-----+-----+-----+

CCCCCTGTCCTTCCCGGGCGTCCCGCCGGGGATCCCACAGTTGGCAGCTCTTCTCAAAT  
 46951 -----+-----+-----+-----+-----+-----+-----+

TCTTCCCTTAAAAATAGGATTTGACACCCCACTCTCCTTAAAAAAAAAAAAATAAGAAAA  
 47011 -----+-----+-----+-----+-----+-----+-----+

<----- S8  
OX-E -----  
 AAAGGTTAGGTTATGTCAACAGAGGTGAAGTGGATAATTGAGGAAACGATTCTGAGATGA  
 47071 -----+-----+-----+-----+-----+-----+-----+

----->  
 GGCCAAGAAAACAACGCTCGTGCAAAGCCCAGGTTTTTGGGAAAGCAGCGAGTATCCTCC  
 47131 -----+-----+-----+-----+-----+-----+-----+

+5053 C  
 TCGGCTTTTGC**G**TTATGGACCCACGCAGTTTTTTCGTCAAAGCGCATTGGTTTTTCGAGG  
 47191 -----+-----+-----+-----+-----+-----+-----+

GCCCCCTTTCACCGCGGGATGCACGAAGGGGTTTCGCCACGTTGCGCAAAACCTCCCCGG  
 47251 -----+-----+-----+-----+-----+-----+-----+

† EXON III OX-DR -----

CCTCAGCCCTGTGCCCTCCGCTCCCCACGCAGGGATTTATGAATGCAAAGAGAAGCGCGA  
 47311 -----+-----+-----+-----+-----+-----+-----+

<-----

----->

GGACGTGAAGTCGGAGGACGAGGACGGGCAGACCAAGCTGAAA**CAGAGGGCGCAGCCGCAC**  
 47371 -----+-----+-----+-----+-----+-----+-----+

----- OX-D

homeobox

CAACTTCACGCTGGAGCAGCTGAACGAGCTCGAGCGACTCTTCGACGAGACCCATTACCC  
 47431 -----+-----+-----+-----+-----+-----+-----+

CGACGCCTTCATGCGCGAGGAGCTCAGCCAGCGCCTGGGGCTCTCCGAGGGCGCGGTGCA  
 47491 -----+-----+-----+-----+-----+-----+-----+

†  
GGTAGGAACCCGGGGGCGGGGGCGGGGGGCCCGGAGCCATCGCCTGGTCTCCGGGAGCGC  
 47551 -----+-----+-----+-----+-----+-----+-----+

ACAGCACGCGTACAGCCACCTGCGCCCGGGCCGCCCGCTCCCCTTCCCGGAGCGCGGGG  
 47611 -----+-----+-----+-----+-----+-----+-----+

AGGTTGGGTGAGGGACGGGCTGGGGTTCCTGGACTTTTGGAGACGCCTGAGGCCTGTAGG  
 47671 -----+-----+-----+-----+-----+-----+-----+

7I1 (78% homology to 7I2)

ATGGGTTCAATTGCGTTTTGTTTTTCACCAACAGCAAACAAATATATATACATATATATTAT  
 47731 -----+-----+-----+-----+-----+-----+-----+

+5683 C

ACAAATAACAAATAAATATATATGTTATACAGATGGGTATAT**TGTATATATTATAGATAT**  
 47791 -----+-----+-----+-----+-----+-----+-----+

7I2                      OX-R ----->

TTGTTTCGTCCTTGGTGCAAAGACACCCGGTGAACCCATATATTGGCTCCTGACTGCCTTC  
 47851 -----+-----+-----+-----+-----+-----+-----+

GGTCCCCTGGGATTGGTTATAGGGCAACACATGCAAACAAAACCTTCCCTGGATTATA  
 47911 -----+-----+-----+-----+-----+-----+-----+

<----- OX-V

CTTAGGAGACGAAGCTACAGATGCGTTTTGATCCAGAGTGTTTTACAAGATTTTTTCATTTA  
 47971 -----+-----+-----+-----+-----+-----+-----+  
 <----- S9  
 OX-S ----->  
 AAAAAAATGTGTCTTTTGGCCCCTGATTCCCCTCCGTCTTCCCGTGTGGCTGCATTGAA  
 48031 -----+-----+-----+-----+-----+-----+-----+  
 AAGGTTTCCTTAGGATGAAAGGAGAGGGGTGTCCTCTGTCCCTAGGTGGAGAGAAACAGG  
 48091 -----+-----+-----+-----+-----+-----+-----+  
 GTCTTCTCTTTCCCTCCGTTTTTTTACCTACCGTTTCTATCTCCCTCCTCCCCTCTCCAGC  
 48151 -----+-----+-----+-----+-----+-----+-----+  
 <----- OX-T  
 +6085 T  
 CCTGTCTCTGCTACAAACCACCCCTCCTCCCTCCGGCTGTGGGGAGCGCAGGAGCACG  
 48211 -----+-----+-----+-----+-----+-----+-----+  
 +6139 A  
 TTGGGCATCTGGATGAGCGGAGACTATTAGCGGGGCACGGGGGCTCCCCGAGGAGCGCGC  
 48271 -----+-----+-----+-----+-----+-----+-----+  
 +6195 T                  +6211 T                  +6236 G  
 GAATTCACGCTGCCCATGAGACCAGGCACCGGGGGCGGAGGGGCCTTGGGTGTCGCA  
 48331 -----+-----+-----+-----+-----+-----+-----+  
 s23 <-----S-----  
 +6260 A  
 s22 ----->  
 GAGGGACGGGCGGGCAGAGCCTTCCCTCCGCATTCTAAACATTCACTTAAAGGTATGAGTT  
 48391 -----+-----+-----+-----+-----+-----+-----+  
 -----  
 TATTTAGGGGTGCTGCTGGGAGAGCCTCCAAATGGCTTCTTCCAGCCCCTGCCTGACAG  
 48451 -----+-----+-----+-----+-----+-----+-----+  
 S10 ----->  
 TTCAGCTCCCCTGGAAGGTCAACTCCTCTAGTCCTTTCTCCTGGTTCTGGGCAGGACAGA  
 48511 -----+-----+-----+-----+-----+-----+-----+  
 <-----  
 AGTGGGGGGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCGGTCAGGATCCCCGGACCCTG  
 48571 -----+-----+-----+-----+-----+-----+-----+  
 --- S11

+6485 C  
 48631 GGGAAACCCGTCAAAAATAAATGAAATTAAGATTGCCGACCAGAGAGAGAACCGTGACAAA  
 -----+-----+-----+-----+-----+-----+-----+  
 <----- OX-P

48691 GCAAACGGCGTTCAAAGCAAAGAGACGAACTGAAAGCCCCTTCCCCTAGGACTGGTTATG  
 -----+-----+-----+-----+-----+-----+-----+  
 <----- OX-J

G +6630  
 48751 AGGTCAACACATTCAAACACAGCTTGCTCTGGATTTTGCTGAGCAGAGGAAGATACAGAT  
 -----+-----+-----+-----+-----+-----+-----+  
 48811 GCATTTGATCCAAAGTGTGTTACATCTTTCATTATATGTGTGTCTATATATATAAACATA  
 -----+-----+-----+-----+-----+-----+-----+  
 tandem repeat/8I1  
 48871 TATAAATATATAAACATACATAAATGTATGTAAATATATATAATCTATATACATATATAA  
 -----+-----+-----+-----+-----+-----+-----+  
 8I1 (69% homology to 8I2)  
 48931 ATATATAAACACATATATAAATATATAAATCTATAAACATATATAAATATATAAACATAAAT  
 -----+-----+-----+-----+-----+-----+-----+  
 48991 ATATAAACATATATAAATATATAAATATATTAACATATATAAAATATGTATAAATATATAT  
 -----+-----+-----+-----+-----+-----+-----+  
 tandem repeat/8I1  
 49051 AAACATATAAACATATATAAATATATAAACATATAAATATATAAACATATATAAATATAT  
 -----+-----+-----+-----+-----+-----+-----+  
 8I2  
 49111 ACAAACATATTGTATATATATAAATATATATAAAAACATATATATACATATAAAAATATA  
 -----+-----+-----+-----+-----+-----+-----+  
 49171 TATAAACATATATACATATAAAGAAATATATATAAACATATATACATATAAATATACATA  
 -----+-----+-----+-----+-----+-----+-----+  
 49231 TATAAACATATATATACATAAATATATATAAACATATATACATATAAAAATATATATAT  
 -----+-----+-----+-----+-----+-----+-----+  
 49291 ATTAACATATATATACATATAAAAATATATATATTAACATATATATACATATAAAAATAT  
 -----+-----+-----+-----+-----+-----+-----+  
 49351 ATATATATTTTGGCCCCTGATTCCCTTCGGTTCCTGTGGGATGGGTGATTGAGTCAACA  
 -----+-----+-----+-----+-----+-----+-----+  
 49411 CATTCAAACACAACCTTTCCATCGATGTTGCTTAGGAGATGAGGATACAGATGCGTTTGA  
 -----+-----+-----+-----+-----+-----+-----+

OX-G ----->  
TGGAGAGGGTTTTACAAGCTCTTTCATTTAAATATATATATATATATATATATATTTTT  
49471 -----+-----+-----+-----+-----+-----+-----+  
<----- S12

OX-K ----->  
GGCTCCTGATTCTCTTCCGTCTTCCCATGTGGCTGCATTTTAAAAGGCTTCCCTAAGATC  
49531 -----+-----+-----+-----+-----+-----+-----+  
G +7445  
GTTACGATTAAATCAACCCTCCCCAGGCATCTTTACCGAGGGCTGTGGTCCCCAAAGCGA  
49591 -----+-----+-----+-----+-----+-----+-----+  
TACAGCCCAGGAGGGAGAGAGGCTTTGGTGACTTGGAGGAAGGACTGTGTCCCTCCTTAG  
49651 -----+-----+-----+-----+-----+-----+-----+  
GGCGTCTGTGGCCTCAGTGAGGGAAGGAAGCTGCATCAGACAGGGGTTTCTCGCTGTCC  
49711 -----+-----+-----+-----+-----+-----+-----+  
ACCCCTCTGGCAGAAGATGGATTGGGCTGCCCCGTATAAAATTAATGAAAAGATTAAAGTT  
49771 -----+-----+-----+-----+-----+-----+-----+  
TCGCTAAAGGGGACATCGAGTTTATGTGTCATCTCCTGGTGTCTGTGTGCCTGGGATCTG  
49831 -----+-----+-----+-----+-----+-----+-----+  
CAATATATCCCAGCCCTTGATGTACTGTTTCTATAAAAAATAAATTACTTGTAATTTAATT  
49891 -----+-----+-----+-----+-----+-----+-----+  
CCACACTATTTCTTTCCGTAGTCTATTACCGACGAGAGCACGTTAGTTCAGCTGCGGAAA  
49951 -----+-----+-----+-----+-----+-----+-----+  
ATTGGTTGTGGGGTGTGTGCGGACCCCGAGAACGCCCTAAAATAAAGACAAATCGGGGAC  
50011 -----+-----+-----+-----+-----+-----+-----+  
AAGCTGGGGGTTATCGATTGCAGGGGTCGCATGAAAATTTAACGACGGTAAATAATAATA  
50071 -----+-----+-----+-----+-----+-----+-----+  
C +7982  
AAACAACATGGGAATGCAATAAAAGACATAATTCTCCATCGCCGCGGGGGAAAGGAT  
50131 -----+-----+-----+-----+-----+-----+-----+  
CCTATAGTAAAGGCGAGTGCCTTTGAGGGGTCATAAAAATCAATTAGTTCCAACACCCA  
50191 -----+-----+-----+-----+-----+-----+-----+  
CGTCCCGCGTTGAGGGGACGGGGACGAGCAGGGACAGAAAAAGAAACCATATTTGAATCC  
50251 -----+-----+-----+-----+-----+-----+-----+



S13 ----->

50311 CATCTCTCTGTGAATTCTTGGGTACATGCGTCTCAGTACAGCCCGTCCCGTGCTGTGAC  
-----+-----+-----+-----+-----+-----+

50371 CGGATAGAGTTTCAATTTACTGTGGAAATTTGCTGTAAATAAATTGAGCATCCGATAGAA  
-----+-----+-----+-----+-----+-----+-----<

50431 GCTGTTGCTGATTAACCTTTTATTTTTAGCGTGGCCCTGCAAAGTCGTATCACCCAGCTG  
-----+-----+-----+-----+-----+-----+-----

----- S14

50491 TCAGGCTTCTAATCGAAAGTTATGAGACCACGGTGAGGGGCAGGCGGTAATTTAATTACA  
-----+-----+-----+-----+-----+-----+-----

50551 ACAAATATCTTTGGGTTTATGGCGCAGAGCTAAATTAATGTCATTATTCACTGTCTGTA  
-----+-----+-----+-----+-----+-----+-----

50611 ATGGAAATCAAAGGAAATCGCATTACGGCATTGAAAGAAAGCGGGGAGTGCTCTTTT  
-----+-----+-----+-----+-----+-----+-----

50671 AATGAAGAAATAACTGTCTTAAGCAGTGTACACACTTCACTTACCATATTCGGGCCTAA  
-----+-----+-----+-----+-----+-----+-----

50731 TTGGAATGGATCGTGAATCACTCCAAGACTGATTTATTAGCGCTTCACGCAGCGGCTAAT  
-----+-----+-----+-----+-----+-----+-----

50791 TCATCACTTGTATTCTTCATCATTTTTTTTTTCTCTCGCCGTGTTGAAGGGAGAGTGA  
-----+-----+-----+-----+-----+-----+-----

50851 ATGAGGCTTCCACGTTTCAGGAGGATTTCTTTTTTGAAAAATGCCCTTCCAGAGGCTT  
-----+-----+-----+-----+-----+-----+-----

50911 TTGGGTGGCTGGCTTGCTTTCTGGGCCCTGGAGGAGACAGGCGGAGAGTCCAGGTGGGCA  
-----+-----+-----+-----+-----+-----+-----

50971 TGGAGAGGCACAGTGGCAGGTCACCTGGATGGTCAGTGGAGGTGGAGGTCTGAAGGCCCC  
-----+-----+-----+-----+-----+-----+-----

51031 AGCTTTGAAATTATTGGTGAATTTGATGTCAGCACCAGGCAGGGGCCTTTTTGGCGGG  
-----+-----+-----+-----+-----+-----+-----

+8965 C

51091 GGTGTGAGGGAGGATGACTTTGCTGGGAAACAGGATCAGGTTCTCCAGGCGCACTGCAGC  
-----+-----+-----+-----+-----+-----+-----

51151 CCGGTAGGACCCACTTTGAAATGAAAAGCCAGTTCCGAAAGCTGGGCTGGAAGCTTCCG  
-----+-----+-----+-----+-----+-----+-----

51211 TGTGGGTTCAAGAGCAAGTTCACGTTGCGCTGTGTAGACTCCTGGCTGCTCCCAAATC  
 -----+-----+-----+-----+-----+-----+-----+

S15 -----> +9177 T  
 51271 TGAGGGTTTTCTGAGGTTCCCTTCATAGGGGCACCGGCCCTGGGCCATGCACAGTGCCTA  
 -----+-----+-----+-----+-----+-----+-----+

51331 AGGGTGGCTGTGGGCCGAGGGACCCAGCACGTGTTTTGCCACAACAGCCGGAGTGAAGT  
 -----+-----+-----+-----+-----+-----+-----+

51391 GTTCACTCACCGCCTTGGCGGAGGACGCCTGTTCTCTGGACGAATCATTCTCTTGGGTG  
 -----+-----+-----+-----+-----+-----+-----+

51451 GTGACTGCCTTGTGGGTCAAGGTGCAGGTTTTCTGCCACAGAAAACCTGTTAGGAGGAAT  
 -----+-----+-----+-----+-----+-----+-----+

51511 TAAGCGACTAAGACTGTCAGGGAGGTGGTGGGGGAGAGGAGGGGGTGGTGTCCAGAT  
 -----+-----+-----+-----+-----+-----+-----+

51571 TACCAGGCATAGGCTAAACTGCCTGCACTCTCCAGCTGGTCTGTCTGTGGAGGAGGGGAT  
 -----+-----+-----+-----+-----+-----+-----+

51631 TGTCAATACTGGGAGAGCAGAGGAGGCTCGTAGGAGGTGAGAGGGGGTGAATTTGCATG  
 -----+-----+-----+-----+-----+-----+-----+

+9595 T  
 51691 CAAATCTTACATGAGGCCTGTGTGAATTTCTCCAGCCTCCTGAGGGTCCCCTGCCTAT  
 -----+-----+-----+-----+-----+-----+-----+

51751 TGCACTCAACTTCTTGATAGTTTACCCCAAGACTCAGAAGTCCTTAGAGGGGCAGAATGC  
 -----+-----+-----+-----+-----+-----+-----+

s21 <-----  
 +9697 T  
 51811 CCCACCACAAAGCCTGCTATCCTTGGGCGTCTCAAGACCCTTGGTTCATGAATGGGACC  
 -----+-----+-----+-----+-----+-----+-----+

51871 CTTTCATGTATGGGACCCTTGGTAATATGAATGGGACGCCTTCAGCTCCCCAGGGCTTC  
 -----+-----+-----+-----+-----+-----+-----+

CGAGGAGGCCGAGAAGGGCAAAGACA  
 51931 -----+-----+-----+-----+-----+-----+-----+

51991 CTGGGCTTGGTGTGTCAAGAGCTAGATTGGAGAAGGGGCTGGATTTGAACTCTTTAGCC  
 -----+-----+-----+-----+-----+-----+-----+

52051 ATCAGCTCACCTCTCCGTTTGTGGCTAAAGTCTGAAGGTGGAAACTTCGGTTCTCCTAC  
 -----+-----+-----+-----+-----+-----+-----+

52111 AGGGTCTACAGGAGTTGGGGGGCGGGGCGCCACACAGAACGCTGGAAAGTTCGACAGTC  
 -----+-----+-----+-----+-----+-----+-----+  
 52171 CACTTCCACTGGCTCGGAACTCACTTTTTTCACCTTAAGTTCATCAGCGGTAACGCATAGG  
 -----+-----+-----+-----+-----+-----+-----+  
 52231 TCTCACTTAGGCAGGGCACGGATGATTTAACAATTTCTACTTCTAGGTCAGGTGCGGTGG  
 -----+-----+-----+-----+-----+-----+-----+  
 --- OX-H

---

## SEQ-35:

### U82668

-----> SEQ3A  
 61282 GCGTGAACATTCAGCATGGAAACAACATACGTCTCTCCACAGGAGGTGAGAAATTGAATT  
 -----+-----+-----+-----+-----+-----+-----+  
 61342 TATGGGGTGGGTGTACGCTGGCGATTCTTGGTGCTTTTTGCTCAAAACAAGTTCTTTTG  
 -----+-----+-----+-----+-----+-----+-----+  
 61402 AAAGTCACGTTCCCTGCTTTCCCTGTGGCTTCCCGGTGAGCTCGCTCGCAGAGCAAGGAAT  
 -----+-----+-----+-----+-----+-----+-----+  
 A +3/341  
 61462 ACCACCCAGAGAGCAACGTGGGCTGTGTTCCGTGTAACGCCGTTGCAGAGAGAGGATTT  
 -----+-----+-----+-----+-----+-----+-----+  
 61522 GGTGTGTGAGATCCGTACCAGCTCCAGCACACTGATAGGAACACGTTGCTGGCCGAACTG  
 -----+-----+-----+-----+-----+-----+-----+  
 61582 AACGATGCTGGGTTGGGTCCTGATTGATACGTATTTTTCTTCCCTCCTCTCCCCAAAACCTT  
 -----+-----+-----+-----+-----+-----+-----+  
 61642 GGCCAAATAGTCCGTGGAGGGTTGTCAGTCGCCGAGTTGAGCAAAAAACACTTCTTCCT  
 -----+-----+-----+-----+-----+-----+-----+  
 61702 TTGAGTGGCTGTTCTGGTGAAATCTGTTTCTGACATATCCACTTTTCTCTCTTTTCTC  
 -----+-----+-----+-----+-----+-----+-----+  
 +3/663 C  
 61762 TCTCTGACTGCGAAGCACCCACAGGGAGAAGGAATTGGATGTATCGGATGTTGGTATT  
 -----+-----+-----+-----+-----+-----+-----+  
 61822 AGATTTTCTTTCTCCGTTTCGAGTCTCTGACTGGTGCATACTTTGCAAAGGTGTGTTCCCTG  
 -----+-----+-----+-----+-----+-----+-----+

```

GCAATTGCCAAGAGTTAGAAAAATGCACCTTCTCTGGTGGCCGTTGGGGTGTGTTTTCAC
61882 -----+-----+-----+-----+-----+-----+
AGGCAGTGGTGACAGGGCCCCTTGGCTGTGGCTGTCTTCTCCAGCGCCGTGGATAAAGAG
61942 -----+-----+-----+-----+-----+-----+
      T  +3/849
ACGGGACAGATTCTGTGCCTCTGTACGATTTAGAGCGTAACTGACCGGTCCAACACCCG
62002 -----+-----+-----+-----+-----+-----+
TTTTTCCACTTACAAAGCTGGTGGTGCACGGGCTTGGTGTCTCCCGTACGGGAAGGAGG
62062 -----+-----+-----+-----+-----+-----+
CCTTTGGGCCGCTCCAAAGACGCCCTGTCGTAGGAATGGCCTCTCCATCCCGCCAAAGTC
62122 -----+-----+-----+-----+-----+-----+
      A  +3/1035
CAGCCAGGCCCCCGAAATGGTCCCATTTCCCTTGAAGCCTGAGTTTCTGTTCTGGTCTTG
62182 -----+-----+-----+-----+-----+-----+
CTGCTGTCCTTGGCCACGTCAGCACGTGGGAGCATCTGTGGATAACCGCAGAGTCTGGGGA
62242 -----+-----+-----+-----+-----+-----+
      ----->  SEQ35B
CAGCTGGGCGTTTAAACCGAAATGAAGCCGAGACGGGTTTTCAGGTTTTGGTGCCAAGCTCT
62302 -----+-----+-----+-----+-----+-----+
      +3/1239      G      A  +3/1248
GGTCAGGATGAAAGGGAAATACCAGAGTCCTCTGTCTCTCGCCTCTGGGTTTCATGCTGAC
62362 -----+-----+-----+-----+-----+-----+
CTTTCTAACATTTGTTTTCCCCTAAGAACAAGCAGAAGCCTCCAGCTCCCTTTAGCTCCA
62422 -----+-----+-----+-----+-----+-----+
CAGTTTTCCCGGGACATAGCGAGGATGGCACACGGCAGCCACTCCCACGACACACATTT
62482 -----+-----+-----+-----+-----+-----+
      <-----  SEQ3B
CGGAGGCACTTTGCTGGAAGCCGCTTGTCTCCTCCAGCTTTGGGAGGTCTGGGGAGGAGA
62542 -----+-----+-----+-----+-----+-----+
      !! sequence termination !!----
GAGGCTTTCGGTGGACACGTTTGACATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAACTGG
62602 -----+-----+-----+-----+-----+-----+
      ----->  SEQ35B1
TGCCTAATTTATTAAGAGAATTAGCTTAGCGAGTATATGCTGATATTCTTCGACACACG
62662 -----+-----+-----+-----+-----+-----+

```

```

TGGGTAAGTTGATGCCATTTATAAATGTTTTATTGAAATTTGATATTTAATGAGAAGCCG
62722 -----+-----+-----+-----+-----+-----+
GTTAAGGAATGTAGACAATATCCCGTTTCAAAGCTATGAAATGTGCTATTTATTGAAAGG
62782 -----+-----+-----+-----+-----+-----+
GGATGTGGCTTCACGAGTTCAGCCCATTGTACGTGCAGGTCCCGTGGGAAGGAGGCAAAA
62842 -----+-----+-----+-----+-----+-----+
GCCCCGCTTCTTACTTTGTGATGTATGTGCATTTGTTATTTATTTTTTTTTTTCCTTGGTC
62902 -----+-----+-----+-----+-----+-----+
-----> SEQ35B2
GGACGTTCAATAATATGTACTATTTTAATTATGTTCGAGTGTAATTTGACATCGCGTTGC
62962 -----+-----+-----+-----+-----+-----+
ATTTATTTTTATATTTCTGAAAACGTTTGCTTTTTCTTTTTCCCTCCCCCATTGACGACA
63022 -----+-----+-----+-----+-----+-----+
TAGCGGCCCCCGCGTCCGGGTTACAAATACATCTACAGATATTTTCAGGGATTGCTTCAG
63082 -----+-----+-----+-----+-----+-----+
ATGAAAACAAATCACACACCGTTTCCCAAACCAACAGTCTTCACATTTCTATCCCTCTGT
63142 -----+-----+-----+-----+-----+-----+
TATTGTCGGCAGGCGGTGAGGGGTAGAAAAAAACAAACAAACAAACAGAAAAAAAAAAC
63202 -----+-----+-----+-----+-----+-----+
-----> SEQ35B3
AAAAAAACCACCCTGAGTTTCTCTGGTGACGCCCTCATTCTCCTAACGTTCAATAATCT
63262 -----+-----+-----+-----+-----+-----+
CAATGTTGAGTTGCAGCAACAGACTGTATTTTTGTGACGCCCCGTAGTATGAATGTACAT
63322 -----+-----+-----+-----+-----+-----+
CTTGTAAAACTGAGATATAAATAAACTTATAAATATTTGTATTCAAGTGTTAAAAAAAA
63382 -----+-----+-----+-----+-----+-----+
AAATTTCTCAACCTCTCCCCTGAGGACAGGCTTATTGGAAAAAAAAAAAAAAAAAAAA
63442 -----+-----+-----+-----+-----+-----+
TCCTGAGTCGGCCGTGGCTGAACACAGAGTGTGTTCTGCTCCGTGCATTTCCAGGGTGG
63502 -----+-----+-----+-----+-----+-----+

```

G +35/2287

63562 GTACCCAGTGTGGCCCCCAGCCTTAGATCGGGAGGTACCATTGACTTTTGCTTGTATCC  
-----+-----+-----+-----+-----+-----+

63622 CATCCCCTTCCTTTACTGAAACCTACCTCCCCGCTTCTCAGCCAACGTCCCCCAGAAGG  
-----+-----+-----+-----+-----+-----+  
<----- SEQ35A2

63682 TGGCAAAAAAACAGAGGAAAAAGCCCTGATTTGAATCAAGTCAGAGCTGCTAATTCTCC  
-----+-----+-----+-----+-----+-----+

63742 ACTTCTTTAATTAATTAATTTATTTTTTTTTTTGAGACTGAGTCTCGCTCTGTGCCCCA  
-----+-----+-----+-----+-----+-----+

63802 GGCCGGAGGAGTGCAGGGGCGCGATCTCGGCTCACCGCGACCTCCGCCTCCCGGGTTCAA  
-----+-----+-----+-----+-----+-----+

A +35/2637

63862 GCGACTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATGACAGTCACTGCACCACCGCGC  
-----+-----+-----+-----+-----+-----+

63922 CCGGCTCATTTTTTGTATTTTTAGTAGCAATGGGGTTTTCACCGTGTGGTCAGGCTGGTCT  
-----+-----+-----+-----+-----+-----+

63982 CGAACTCCTGACCTCGTGATCCACCCGCTCTGGGCCCGCCGGTGATGTGTGTGCTTTT  
-----+-----+-----+-----+-----+-----+

64042 AACTTTTTATTTTGTTCAGTTTTTCGACAGTGGCACGGATTTTCAGCACGGTCTTGCAAG  
-----+-----+-----+-----+-----+-----+

64102 GATGATTGAGTCATTTTTGAGACAAAAATATAATAATAATAATGGAAAAAGAAATCGA  
-----+-----+-----+-----+-----+-----+  
<----- SEQ35A1

!! sequence termination !!

64162 CTTTTAAAAATGACAAAATTTTTTTTTTTTTTTTTTGCATAGATTTTCTCTCTTTATGTA  
-----+-----+-----+-----+-----+-----+

64222 AAGGAAAGTTCATGATTGGATTTGGCCGGCCTGACTGCTTCCCGGCTGTGATAAAAAACA  
-----+-----+-----+-----+-----+-----+

64282 CATGTGAGCTGGGAGGGAAGTGGGGGAGGGACACAGCTGCCACACAGGGTTCCACCCG  
-----+-----+-----+-----+-----+-----+

64342 GGTACAGGGTGGGCAGTGCTGGGGGAGCTTCTCTGTGGGGGGCTCAGAGCCTGAGGAC  
-----+-----+-----+-----+-----+-----+

64402 AAGTGAGCCTCTCCGACACCTCCCCAGTTGCCTGGAGTCTAAACCGTCCGTTGTCTGTAC  
-----+-----+-----+-----+-----+-----+

64462 CGTCCGTTCTTCCTGCTGACTCCTGGTAGTTCCTGAAAGCTTCTCTTGGCCAGAGAAGGG  
 -----+-----+-----+-----+-----+-----+

SEQ35C ----->

64522 GTTTCAGAGGCCGTGTGTCCAGGCCATTCTGCAAAGTGCAACTTGACCGTTCCTTTTCCTT  
 -----+-----+-----+-----+-----+-----+

64582 TTCTGGCCTGCGTGGTCTGAAGCTCAGAGCCCTCTCTTCACCCAGCCTGTGTGTGTCTTG  
 -----+-----+-----+-----+-----+-----+

C +35/3390

64642 CCGGACAGAAGAAAAATGGTGCTTTTTTGC GTGTTAGCAGAGGTGCTTTTCATGGCTGACC  
 -----+-----+-----+-----+-----+-----+

64702 TCAACGCGTCCATCTCCAGCCTTGACCAAGCTGTTTTTTAGGGGCAAACGCAGGCAAGTT  
 -----+-----+-----+-----+-----+-----+

64762 CTGAATGCACACAGTTATTTTCATGGTTAAACTATTCAGCTTTGGCCGGGCGCAGTGTGGC  
 -----+-----+-----+-----+-----+-----+

<----- SEQ35A

64822 TCTCACGCCTGTCATCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGGTCAG  
 -----+-----+-----+-----+-----+-----+

64882 GAGTTCGAGACCAGCCTGGCCAACACGGTGAAACTCTATCTCTACTAAAAATACAAAAAT  
 -----+-----+-----+-----+-----+-----+

64942 TAGCCGGGCGTGGTGGTGTGTATCTGTAATCCAGCTACTCAGGAAGCTGAGGCAGGAGA  
 -----+-----+-----+-----+-----+-----+

65002 ATCGCTTGGACCCAGGAGGCGGAGGTTGCACTGAGCCGAGATCGCGCCATTGCACTCCAG  
 -----+-----+-----+-----+-----+-----+

65062 CCTGGGCGACAGAGCCAGACGCTGTCTCAAAAAAATGAATAATAAAATAAAATAACAGGA  
 -----+-----+-----+-----+-----+-----+

65122 ACTAAATAAAATAAAACGTTTCAGCTTTGTTCTGCAAATCCACTCCTATTGTTTTACGTGG  
 -----+-----+-----+-----+-----+-----+

65182 TTTGAGAGACTCTGTCCCTTAGAAATAGATGTTTGTGCAATTGTAATGAATCTGTTTC  
 -----+-----+-----+-----+-----+-----+

65242 AAAAATGAACAGAATATTCAAATGGTTTGAGAGATCTTTTCCTTAGAAATAGCTTGTGG  
 -----+-----+-----+-----+-----+-----+

65302 CCAATCACAAGAATGTTTTTCAAAAAATGAATGGAATCTTCCTGGATATCGCTTCCAGAT  
 -----+-----+-----+-----+-----+-----+

65362 CTTCA**TTTTTTTT**GCATAGTTCAACCTGAAAAGTAAGTGTCTCAGCCCTGAATTTCTTTC  
 -----+-----+-----+-----+-----+-----+-----+

65422 TGATTTTCCATGGGTTGTCTTGCAGACTTCTCTGGACTTGACCACATTT**AAAAAAAAAA**  
 -----+-----+-----+-----+-----+-----+-----+

65482 AAATTAACTTTTTTCACACGGACACGGTTTTCAATAGGAATGAGATCTTTGAGTTTTTATGT  
 -----+-----+-----+-----+-----+-----+-----+

65542 AACAGATTCTTACCATCAGTTCTCAGATTCCCAAATTACACACAAAAAGCCACGGACTTC  
 -----+-----+-----+-----+-----+-----+-----+

65602 GCCTCCTGCTAACATGTCCTTCTGTTTTCTGAGGCTTCTGTTGGTGTTAGACTTTCATGTT  
 -----+-----+-----+-----+-----+-----+-----+

65662 TAATAGCAGACAATGTAGGGATTTAAAGAAAAATGCAGAGAAAGCAAAAAACACTGACCAA  
 -----+-----+-----+-----+-----+-----+-----+

65722 ACACACGGAGATAAGCTTTCTAAAGCCTTTGTTCTTGGAGTTGTCGTT**AAAAAAAAAA**G  
 -----+-----+-----+-----+-----+-----+-----+

<----- SEQ35D2

-----> SEQC1

65782 TTGTTTTAACTTTGCAAGCATGCCTATATTGAACTCATAAGCAAGAGACCAAGAAAAA  
 -----+-----+-----+-----+-----+-----+-----+

65842 TAGTGTCGGTCGTCTACTCTACACGTTTTCCCAAACAGACGTATTTAATTTCTTTTGT  
 -----+-----+-----+-----+-----+-----+-----+

A +35/4673

65902 TTGAACTCACAGATGCTGAGAGTTAAAAGTTAAATTTTTGTCATGAACAA**TAGTGGCCAA**  
 -----+-----+-----+-----+-----+-----+-----+

65962 **AACCACAGTTACTTTTGCACT**TATAGCATAATAAGAAAAATACAG**GGCTGGGCTCGGTGGCT**  
 -----+-----+-----+-----+-----+-----+-----+

66022 **CACACCTGTAATCAAAGCACTTTTGGAGGCGAAACAGCCAGATCCCTTGAGCCCAGGAGA**  
 -----+-----+-----+-----+-----+-----+-----+

66082 **TTGAGACCAGCCTGGGCAACATAGCGAGACCCT**CATCTCTACAAAAAAGGTTTGT**ACAT**  
 -----+-----+-----+-----+-----+-----+-----+

66142 **ATGTAACAAACCTGCACATTGTGCACATGTACCCTAAAACCTTAAAGTATAATAA**TAAAAA  
 -----+-----+-----+-----+-----+-----+-----+

66202 AATT**AAAAAAAAA**TTACCAATCAACTGCCTGCTGGTGCCTTCAAGAGACTCACCTAACA  
 -----+-----+-----+-----+-----+-----+-----+



66262 CATAAGGACTTGCATAAACTTATAAAACAATTCAATGGAAGAATCCTTGAAAGTATTCTG  
 -----+-----+-----+-----+-----+-----+  
 66322 AGAAGACAGTATAATAAACTGATTTCTAAAAGGCTATAAAAAATTGAATAAATCATTGT  
 -----+-----+-----+-----+-----+-----+  
 66382 TGGGCATCCTGTGCTGAAATATAATGCAGCCAATAAAAAATTACAAAATGAATAAACATTT  
 -----+-----+-----+-----+-----+-----+  
 66442 TATAACAATAAAAAAAAGTCAAATAATTAGGCAGGCATGGTGGTGCTCTCCTACGGTTGA  
 -----+-----+-----+-----+-----+-----+  
 66502 AGCTATTCAGCAGGCAAGAGGATACTTTGTTTTTGTTTTTTAATTTTTTTTGAGACAGAG  
 -----+-----+-----+-----+-----+-----+  
 66562 TCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGTGATCTCAGCTCACTGTAATTTCTGC  
 -----+-----+-----+-----+-----+-----+  
 66622 CTCCCGGTTCAAGCGATTTTCTGCCCCAGCCTCCCGAGTAGCTGGGATTACAGGTGCC  
 -----+-----+-----+-----+-----+-----+  
 66682 CGCCACCACACCTGGCTAATTTCTTTTGTATTTTGTAGTAGAGACGAGGTTTCCCATGTT  
 -----+-----+-----+-----+-----+-----+  
 66742 GGCCAGGCTGGTTTTGAGCTCCCGACCTCGGGTGATCCACCCGCCTCAGCCTCCCAAAGT  
 -----+-----+-----+-----+-----+-----+  
 66802 GCTGGGATGACAGGCGTGAGCCACCGCGCCTGGCCCAGGAGGATTATTTGATCCCAAGAG  
 -----+-----+-----+-----+-----+-----+  
 66862 GTGGAGGCTGCAGGAAGCCATGATTGCACCACTGCACTCCAGCCTGGCTGACAGAGTGAG  
 -----+-----+-----+-----+-----+-----+  
 66922 ACCACATCTCTAAATAAATGAATAAATACAGGCAGAACTTTTTTTGTTTTGTTTTGATG  
 -----+-----+-----+-----+-----+-----+  
 <----- SEQ35D1  
 -----> SEQ35C2  
 66982 GAGTCTTGCTCTGTCAAGGAGGAGTGCAGTGGTGCCATCTCAGCTCACTGCAACCTC  
 -----+-----+-----+-----+-----+-----+  
 67042 CACCTCCTGGGTTCAAGCAATCCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGT  
 -----+-----+-----+-----+-----+-----+  
 67102 GCCCGCCACCACGCCCGGCTAATTTTTTGTATGTTTGTAGTAGAGACGGGATTTACCGTGT  
 -----+-----+-----+-----+-----+-----+

TAGCCAGGATGGTCTTGATCTCTTGACTTTGTGATCTGCCTGCCTCAGCCTCCCAAAGTG  
 67162 -----+-----+-----+-----+-----+-----+

CTGGGATTACAGGCATGAGCCAGGAGTTCAAGACCAGCCTCAGCAACAAAGTGAGACCT  
 67222 -----+-----+-----+-----+-----+-----+

TTTCTCTCCAAAAAATCAAAAATTTAGCCAGCTGTGGTGGCTCCTGCCCGTGATCCCAGT  
 67282 -----+-----+-----+-----+-----+-----+

ACTGTGGGAGGCTGAGGCAGAATTGCTTGAGCCAGGAGTTCGAGACCAACCTCAGCAAA  
 67342 -----+-----+-----+-----+-----+-----+

AAGGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTATATATATATATATA  
 67402 -----+-----+-----+-----+-----+-----+

T +35/6219  
 TATATATATGAGTTTCAAAAATTGCTGGGTGACCAGCTCATCTACTGGTTTTCCCTTGG  
 67462 -----+-----+-----+-----+-----+-----+

GAAAGTGAAATTGTCATGTATTGAAGATTTCCAAGGAAGTTGTATTGAATGAGAAACAAA  
 67522 -----+-----+-----+-----+-----+-----+

-----> SEQ35E  
 CTCAATCTGTTCTGTTTTAAAGAGCTGCAGTGCCTTTGCTGTGTTCCCATAAAACTGCA  
 67582 -----+-----+-----+-----+-----+-----+

CTTCCAAAAGACACGCTGAGAAAGGAGACCAGGATTTGTAATTCAGAAATTGGAAAGCAA  
 67642 -----+-----+-----+-----+-----+-----+

<-----  
C +35/6445  
 GTTAGGCTGGACGTGGTAGCTCATGCTTGTGTAATCTCAGCACTCTGGGAGGCTGAGGC  
 67702 -----+-----+-----+-----+-----+-----+

----- SEQ35D  
 AGGAGGATCACTTGAGCCAGGAGTTCAAGACCAGCCGTGCCACATGGTGAAACCCTGT  
 67762 -----+-----+-----+-----+-----+-----+

CTCTCCAAAAAATAAAACATTTAGCCAGATGTGGTGAICTCATGCCTGTAATCCCGGTATT  
 67822 -----+-----+-----+-----+-----+-----+

CTGGGAGGCTGAGGCAGAGTTGCTTGAGCCAGGAGTTCAAGACCAGCCTCGGCAACAAA  
 67882 -----+-----+-----+-----+-----+-----+

GTGAGACCCTGTCTCTCCAAAAAATAAAACATTTAGCCAGCTGTGGTGAICTCATGCCTGT  
 67942 -----+-----+-----+-----+-----+-----+

AATCTCAGTACTCTGGGAGGCGGCAGTGAATGGCTTGAGCCAGGAGTTCGAGACCAAC  
 68002 -----+-----+-----+-----+-----+-----+

```

CTCAGCAACAAAGTGAGATCTTGTCTCCAAAAAATCAAAAATTTAGCCAGCTGTGCTG
68062 -----+-----+-----+-----+-----+-----+
GCTCATGCCTGTAATCCCGGTACTCTGGGAGGCTGAGGCAGAATCGTTTGAGCCCAGGAG
68122 -----+-----+-----+-----+-----+-----+
TTCGAGACCAACCTCAGCAACAAAGTGAGATCTTGTCTCCAAAAAATCAAAAATTTA
68182 -----+-----+-----+-----+-----+-----+
      +35/6978      A      SEQ35E1 ----->
GCCAGCTGTGCTGGCTGGTGCCTGTAATCCCGGTACTCTGGGAGGCTGAGGCGGAATTGC
68242 -----+-----+-----+-----+-----+-----+
TTGAGCCCAGGAGTTCAAGACCAGCCTCAGCAACAAAGTGAGATCTTGTCTCCAAAAA
68302 -----+-----+-----+-----+-----+-----+
ATAAAACATTTAGTCAGCTGTGGTGGCTCAAGCCTGTGATCCCAGCATTTTGGGAGGCCG
68362 -----+-----+-----+-----+-----+-----+
AGGCGGGCGGATCACGAGGTCATGAGATCGAGACCATCCTGGCTAACACGGTGAAACCC
68422 -----+-----+-----+-----+-----+-----+
GTCTCTACTAAAAATACAAAGAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGTAGTCCCA
68482 -----+-----+-----+-----+-----+-----+
GCTACTCAGGAGGCTGAGGCAGGAGAATGCCGTGAGCCTGGGAGGCGGACCATGCAGTGA
68542 -----+-----+-----+-----+-----+-----+
GTCAAGATCGCGCCACTGCCCTCCAGCCTGGGCCACAGAGCAAGACTCCGTCTCAAAAAA
68602 -----+-----+-----+-----+-----+-----+
AAAAAAAAAAAACTGCTGCCAACCTGTGTTTGCACCACTGCCCTCCAGCCTGGGCAAC
68662 -----+-----+-----+-----+-----+-----+
AGAGCAAGACTCCGTCTCAAAAAAAAAAAAAATGCTGCCCAAGCTGTGTTTGCACCACTGC
68722 -----+-----+-----+-----+-----+-----+
CCTCCAGCCTGGGCAACAGAGCAAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAATGCTGCC
68782 -----+-----+-----+-----+-----+-----+
CAAGCTGTGTTTGCACCACTGCCCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTCAAA
68842 -----+-----+-----+-----+-----+-----+
AAAAAAAAAAAAATGCTGCCCAAGCTGTGTTTGCACCACTGCCCTCCGGCCTGGGCAACAGA
68902 -----+-----+-----+-----+-----+-----+
GCAAGACTCCGTCTCAAAAAAAAAAAAAATGCTGCCCAAGCTGTGTTTGCACCACTGC
68962 -----+-----+-----+-----+-----+-----+

```

CCTCCAGCCTGGGCAACAAAGCAAGCCTCAGCTTTCTGCCATCTCCACAACCAAGAAAGC  
 69022 -----+-----+-----+-----+-----+-----+

AATTCACACAGAAATCAGTGCATCGTGCAGTGACCTCTTCAGAAAACCAATGAGTTTTCC  
 69082 -----+-----+-----+-----+-----+-----+

ACCTGAGGAAGTGTCTGAGCCCCATTGAGAAAAACACATCCCTGTAAGTGCAGGGCAG  
 69142 -----+-----+-----+-----+-----+-----+

ATTTACTCACTGTATGCCTGTTTTAAATAAAGCTTCCAGCCTCTGCATGGGGTCTGTCTGG  
 69202 -----+-----+-----+-----+-----+-----+

SEQ5A -----  
 AAGCTCCTGTATCTGTCCCACATTCTTGAATCACAATGCACCCTTGGGAGGAAGATATG  
 69262 -----+-----+-----+-----+-----+-----+

->  
 TATTTAAAGGGAGTGGATGTTATGGTGAGAAAATGCTGCCCATCCTTCTAGAAGACAAAA  
 69322 -----+-----+-----+-----+-----+-----+

GCCACACAAAATACATCACAAGAACCAGTTTTTTTTTCAGAGAAGAACCTGCACAAAGAACC  
 69382 -----+-----+-----+-----+-----+-----+

<----- SEQ35F  
 TGCTCCCCCACACCCCCACACACAGGTGAATTAACAGGATGTATGTTTTATCATAAAAG  
 69442 -----+-----+-----+-----+-----+-----+

A +5/507  
 CACAGGTTTGTTCCTATGCACTCTCTGAGGATTTGGCCATATGCAAAGATGTACAAAAA  
 69502 -----+-----+-----+-----+-----+-----+

CCTTCTTTTCCCCAGGGAACCGTAACCCGTCTGAAAAGATGCCCTTCTCAGAAGCGAGT  
 69562 -----+-----+-----+-----+-----+-----+

TGAACGATTGTTGGAAAAGATAAAATACGACGTGCACACACACAGTAGAGAAATGTCACC  
 69622 -----+-----+-----+-----+-----+-----+

CATGCAAATTATGTGTTTGAATGGAACACATTGAGGAAGCTAAATGGGGTATGACCACAC  
 69682 -----+-----+-----+-----+-----+-----+

T +5/772  
 ATTTGGGTTGATTTATTTGACGAGTGAAGGGGCAGATGGAAATGAATACTGCTGTTTTTC  
 69742 -----+-----+-----+-----+-----+-----+

CTTTGAAGGCCATATATGGGAATACCAAGAGGATTACTTTGGAAGTTTAGCTTCTCCAG  
 69802 -----+-----+-----+-----+-----+-----+

GTGGTCTCTCTCTCTCTCTCTTTTTTTGAGACAGAGTCTCACTCTGTACCCAGGCTGCA  
 69862 -----+-----+-----+-----+-----+-----+-----+

A +5/934

GTGCAATGGCGTGCCTCTCGGCTCACTGCAACCTCAGCCTCCCAGGTACAAGCGATTCTCC  
 69922 -----+-----+-----+-----+-----+-----+-----+

TGCCTCAGCCTCCCGAGTAGCTGGGATCACAGGTGTGCACCACCACGCCTGGCTAATGTT  
 69982 -----+-----+-----+-----+-----+-----+-----+

TGTATTTTTCAGTAGAGATGAGGTTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC  
 70042 -----+-----+-----+-----+-----+-----+-----+

CTCAGGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATGACAGACATGAGCTAGCA  
 70102 -----+-----+-----+-----+-----+-----+-----+

SEQ35G ----->  
 CGCCCGGCCCCAGGTGGTCTTTTTAGCGGGTATTAAAGCAGCTTTCTCTCTGAGCCTTAA  
 70162 -----+-----+-----+-----+-----+-----+-----+

ACCATGAAGATAGACAGACTCAGTGTATGGGTTTTAGAGTTGTAATTTTATAAAAAATAAG  
 70222 -----+-----+-----+-----+-----+-----+-----+

AAAAAGTCGACCTATCATTGATGGTTAGT  
 70282 -----+-----+-----+-----+-----+-----+-----+

<----- SEQ5B

ATTTTTTGTAGCAGTTGCATGCAATATTAGGATAAGGCATGTTCTCAAAAAGAACTCTTT  
 70311 -----+-----+-----+-----+-----+-----+-----+

TTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTACCCAGGCTGGAGTGCAGTGGCACGA  
 70371 -----+-----+-----+-----+-----+-----+-----+

TCTCCGCTCACTGCAAGCTCCTCTTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCC  
 70431 -----+-----+-----+-----+-----+-----+-----+

CAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCCGGCTAATTTTTTGTATTTTTAGTA  
 70491 -----+-----+-----+-----+-----+-----+-----+

GAGACGGGGTTTCACCATGTTAGCCAGGAAGGTCTCGATCTCCTGACCTCATGATCCGTC  
 70551 -----+-----+-----+-----+-----+-----+-----+

CGCCTCAGCCTCCCAAAGTGCTGGGACTACAGGCGTGAGCCACTGCACTTGGCCTTTTTTT  
 70611 -----+-----+-----+-----+-----+-----+-----+

-----> SEQ35G1  
 † EXON VIb (SHOXb product)

TTTTTTTAGATGGAGTTTTGCTCTTGTGCGCCAGGCTGGAGTATAATGGCATGATCTCGA  
 70671 -----+-----+-----+-----+-----+-----+-----+

CTCACTGCAACCTCCGCCTCCCGAGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAG  
 70731 -----+-----+-----+-----+-----+-----+

CTGGGATTACAGGTGCCACCACCATGTCAAGATAATGTTTGTATTTTCAGTAGAGATGG  
 70791 -----+-----+-----+-----+-----+-----+

GGTGGTACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCGCCT  
 70851 -----+-----+-----+-----+-----+-----+

TAGCCTCCCAAAGTGCTGGGATGACAGGCGTGAGCCCCTGCGCCCGGCCTTTGTAACCTT  
 70911 -----+-----+-----+-----+-----+-----+

ATTTTAAATTTTTTTTTTTTTTTAAGAAAGACAGAGTCTTGCTCTGTCAACCAGGCTGGA  
 70971 -----+-----+-----+-----+-----+-----+

GCACACTGGTGCATCATAGCTCACTGCAGCCTCAAACCTCCTGGGCTCAAGCAATCCTCC  
 71031 -----+-----+-----+-----+-----+-----+

CACCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACCCACCACCACCCAGCTAATTTT  
 71091 -----+-----+-----+-----+-----+-----+

TTTGATTTTTACTAGAGACGGGATCTTGCTTTGCTGCTGAGGCTGGTCTTGAGCTCCTGA  
 71151 -----+-----+-----+-----+-----+-----+

GCTCCAAAGATCCTCTCACCTCCACCTCCCAAAGTGTTAGAATTACAAGCATGAACCACT  
 71211 -----+-----+-----+-----+-----+-----+

†  
 GCCCGTGGTCTCCAAAAACTGTTACGTGGAGGAATGTTCTAGCTTCCTGTTCTCGTCTT  
 71271 -----+-----+-----+-----+-----+-----+

+35/375 G ---  
 TTCTTTGTTAATTGTACAGTTT**GAGGGTGTGTGTGCGTGTGCGCACGTGTGTGTG**CAG  
 71331 -----+-----+-----+-----+-----+-----+

-----> SEQ35G2  
 TCTCCTGATTTTATGTTAATTGTTATTACCACCACCTCCATCTCTCATTCTCTTCTTA  
 71391 -----+-----+-----+-----+-----+-----+

CCCTCACTGTGTAAGATACATGTTGTTTTTAAATTTTATGATTTTATATTTATTTATTT  
 71451 -----+-----+-----+-----+-----+-----+

<----- SEQ35H1  
 T G +35/520 CA, TG, CG, TA  
 GTATTT**CTGAGAC**AGAGTCTCACTCTGTTGCCAGGCTAGTGGCATGATCTCAGCTCACA  
 71511 -----+-----+-----+-----+-----+-----+

71571 GCAACCTTTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGAGA  
 -----+-----+-----+-----+-----+-----+-----+

71631 TTACAGGCACACACCACCACACCCGGCTAGTTTTGTTTTGAGACGGAGTCTCGCTCTGTT  
 -----+-----+-----+-----+-----+-----+-----+

71691 GCAGGCTGCAGTGCAGTGGCGTGATCCTGGCTCACTGCAACCTCTGCCTCCTGGATTCAA  
 -----+-----+-----+-----+-----+-----+-----+

71751 GCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCCACCGCCACACCT  
 -----+-----+-----+-----+-----+-----+-----+

T +35/801

71811 GGCTAATTTTTTATTGGTAGTAGAGACGGGGTTTCTCCATGTTGACCAGACTGGTCTTGA  
 -----+-----+-----+-----+-----+-----+-----+

71871 ACTCCCAACCTCGGGTGATCCACCCACCTGGGCCTCCCAAAGTGCTGGGATGACAGGCGA  
 -----+-----+-----+-----+-----+-----+-----+

A +35/939

71931 GGGCCACCGCGTCCAGCCTTCTTCTTCTTCTTCTTTTTTTTTTTTTTTAAGATGGAGTTTC  
 -----+-----+-----+-----+-----+-----+-----+

71991 ACTCTGTTGCCAGGCTGGAGTGCAGTGGTGCAATCTCGGCTCCCTGCAACCTCCACCTC  
 -----+-----+-----+-----+-----+-----+-----+

72051 CCAGGTTCAAGAAATTCTTTTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGCCCGC  
 -----+-----+-----+-----+-----+-----+-----+

A +35/151

72111 CACCACCCACCTAATGTTTGTATTTTTTTGGTAGAGACGGGGCTTACCACATTGGCC  
 -----+-----+-----+-----+-----+-----+-----+

72171 AGGCTGGTCTTGAACCTGACTTCAGATGATCCTCCTGCCTCAGCCTCCCAGAGTGTG  
 -----+-----+-----+-----+-----+-----+-----+

72231 GGATTACAGGCGTGAGCCACGGTGCCCGGCCAGACGTCATGTCTTAGGAAATCAGAAAGT  
 -----+-----+-----+-----+-----+-----+-----+

A +35/329

72291 GGGTAGTTTCCGCACTCTGAGGAGAAAAAGAGACGTCCGGCGAAGAGAAAGGAGAGTGAA  
 -----+-----+-----+-----+-----+-----+-----+

72351 AGGATGTCTCCTCTTGTCTGTAGCCTGTTCTCAATCGTGAGTGAGCCAATTGCCAGAAAC  
 -----+-----+-----+-----+-----+-----+-----+

72411 TGAGGGTGCTTCATTTGGCCAGGCAAGCTTCTCAACAGAATGTCTAAGTACTTGTTAATG  
 -----+-----+-----+-----+-----+-----+-----+

72471 CTGAGAAGCTCTCCAAGCTACTGCACTCCAGCCTGGGTGACAGAGCACGACCTTGTCTGA  
 -----+-----+-----+-----+-----+-----+  
 72531 AAACAAATTAATTAATCAATTAATTAATATAATGAAATCATACTGAACTCAGGAGACCATT  
 -----+-----+-----+-----+-----+-----+  
 72591 GGGGTGGGCAGGGCTGGGGTTGGAAAGGAACATAAAATATGGTGCAATGGACTTTGCTCC  
 -----+-----+-----+-----+-----+-----+  
 72651 AGTCTCCCTCCCATCTCTTCTCGCCAAGAGTCTCTGGAGGGAGCATGGGGAAGATGCTT  
 -----+-----+-----+-----+-----+-----+  
 <----- SEQ35H

**SEQ-7:**

**U82668**

77041 GAGCCCATGCTCACAGCCTGTCTTAAGGGAGACCTTCTCCAGATCCTGGGTGGTGCCGTT  
 -----+-----+-----+-----+-----+-----+  
 -> SEQ7A  
 77101 TCATGGACAACAAATTCATTCTCAAAAAGCGTAGCCCCAGTGGAAGACCTTGGGTCTTTC  
 -----+-----+-----+-----+-----+-----+  
 77161 CACTCGGAGTGTGTCAGTTCATCACCATAACCGTTTGTGAGGTGTGTTCAAAACATCTA  
 -----+-----+-----+-----+-----+-----+  
 77221 TAGCCAAATTCACAGGAGGAATGGGATGAGGACTCTAAGCTTCTGGACTTGATGTAGAA  
 -----+-----+-----+-----+-----+-----+  
 77281 ACCGGGTCTTTGCAGATGTCTTTAAGGGAAGGATCTGAAATGAGGTCATCCTGGATTAGG  
 -----+-----+-----+-----+-----+-----+  
 77341 GTGGGTCTTAAATCCAATAACAGGTGTCTTGTAAAGGGACAGAAGAGGACCATAAATCTA  
 -----+-----+-----+-----+-----+-----+  
 77401 ATGACTACGGTCTCCTAAGAGATGGAAGAGGAGACATAGACACAGGCACAGACACAGAG  
 -----+-----+-----+-----+-----+-----+  
 77461 GAGAAGGCCACGTGGAGACGGAGGCAGCGACTGGAGTGATCCGGCCACAAGCCCAGGGAC  
 -----+-----+-----+-----+-----+-----+  
 77521 GCCTGGAGCCCCCAGGAGCTGGGAGAGGCAGGAAGGATCCTCCCCTAGAGCCTCCAGAGG  
 -----+-----+-----+-----+-----+-----+



77581 GAACTTCTTATGGTTGAAGTGGATTGAACTATGGTACCCCCAAAAATATGTTTCATATTC  
 -----+-----+-----+-----+-----+-----+-----+

77641 TAACCCAAGAACCTGTGAAAGGAGGTAGAGACTGGAGTGATGGGGCCACAAGCCCAGGGA  
 -----+-----+-----+-----+-----+-----+-----+

77701 TACCTGGAGCACCCAGAAGCTGGGAGAGGCAGGAAGGATCCTCCCCTAGAGCCTCCAGAA  
 -----+-----+-----+-----+-----+-----+-----+

-----> SEQ7C  
 77761 GGAACGGGATACAACCTGAAGTGAAGTGAACCGTGGCCCCCTCAAAGACCTGTCCATGTA  
 -----+-----+-----+-----+-----+-----+-----+

77821 TTGATCATCTGGAATCTGTGAATGGGACATTATTTGGAAATACGGTCTCTGCAGATGTGA  
 -----+-----+-----+-----+-----+-----+-----+

77881 TTAAGTGGGAAGATCTGGAGATGAGATCATCCTAGATTAGGGTGGACCATAAATCTAATGA  
 -----+-----+-----+-----+-----+-----+-----+

77941 CAGGTGTCCTTCTAAGAGACAGAAGAGGAGACACAGACACAGAGGAGGAGGCCACGTGGA  
 -----+-----+-----+-----+-----+-----+-----+

78001 GACAGAGGCAGAGACTGGAGTGATGCAGCCACAAGCCCAGGGATGCC TGGAGCCCCCAGG  
 -----+-----+-----+-----+-----+-----+-----+

78061 AGCTGGGAGAGGCAGGAAGGATCCTCCCTTAGAGTCTCCAGAGGGAGCACAGCCCTGAGA  
 -----+-----+-----+-----+-----+-----+-----+

78121 CTCCTTGACCTCAGACTCCTGGTCTCCAGGACTGGGAGAGGCAGGAAGGACCCTCCCTTA  
 -----+-----+-----+-----+-----+-----+-----+

78181 GAGCCTCTGGAGGGAGCATGGCCCTGAGACACGTTGACCTCAGACTCCTGGTCTCTATGG  
 -----+-----+-----+-----+-----+-----+-----+

78241 CTGGGAGAGGCAGGAAGTCTCCTCCTCAAGAGCCTCTGGAGGGAGCATGGCCCTGAGACT  
 -----+-----+-----+-----+-----+-----+-----+  
 GACCCTCTCCGTCCTCAGAGGAGGAGTTCTCGGAGACCTCCCTCGTACCGGGACTCTGA

78301 CCTTGACCTCAGACTCCTGGTCTCCAGGCCTGGGAGAGGCAGGAAGGACCCTCCCTTAGA  
 -----+-----+-----+-----+-----+-----+-----+

78361 GCCTCTGGAGGGAACATGGTCCTGAGACTCCTTGACCTCAGACTCCTGGTCTCCAGGACT  
 -----+-----+-----+-----+-----+-----+-----+

T +7/1400  
 78421 TGACGGGGATGAATTCCTGCGGCATGCATCGTCATTCTGGTCTTTGTTACGGCTGACT  
 -----+-----+-----+-----+-----+-----+-----+

78481 CAGGAATCTAAGAAACTGACCTGCGGGGTTTTGCCCATCAGGATGGTGGCTTCTCCGGCT  
-----+-----+-----+-----+-----+-----+-----+

78541 GTGGATTTGTTGCTTTTCCATCTTCTCTGCGGCTGACACTCTTACGGGGAGTAATGGTGC  
-----+-----+-----+-----+-----+-----+-----+

T +7/1593

78601 AGAGAATTTTCGGTTGCCTCCCAGAGCTGCAGCCGGAAGACGCAAGGACCAACGAAAACAT  
-----+-----+-----+-----+-----+-----+-----+

78661 TTGCATCGTGGCTCTCTGTGAGGAAGAGTCCTGCTGTGAAAGCATTTTTATTCCGAAAGA  
-----+-----+-----+-----+-----+-----+-----+

78721 GGATCATCTTTGGGAGCAGAACCATGGAAAAGTCTCTTTTCGGCAGCCCACGGAATGATGG  
-----+-----+-----+-----+-----+-----+-----+

<----- SEQ7D

78781 GCCGGCCACATTTTGGATACATCTCATGCTTCTTACTGTTTTCAAATGGTTTTGGCTTGA  
-----+-----+-----+-----+-----+-----+-----+

78841 ATGGAATCCAAGCCATTTTTTTTTTCTCTTTCCCAAGTACTTTTTGTTCTTGAGTCT  
-----+-----+-----+-----+-----+-----+-----+

78901 TGCAATCCTATGAATCCAAGGGAAGCTGAATGTCTATCTTTTTGATACGTGGACACTGT  
-----+-----+-----+-----+-----+-----+-----+

78961 CATCTGCAAACAAAACCGTAGACGTCTTTGGGGGGTTCGTGAACTTGGCAGTGAAGTTC  
-----+-----+-----+-----+-----+-----+-----+

A +7/1985

79021 GAAGGTCTCCAGTGCTTCTCACATCCGAAGTCAGAGCCCACATTTCTCCGTCAGTTTGTA  
-----+-----+-----+-----+-----+-----+-----+

79081 ACCCTGGCAAAGAAGGTCTTCGGCTTGCTTGTTTTCTGAAAAACAGGAAGTTTTATTT  
-----+-----+-----+-----+-----+-----+-----+

79141 TGCCAAGAATGAATGGAATGGGAAGCCAAAGCGTTTGATATGCACAAGGGAAAATGCAGG  
-----+-----+-----+-----+-----+-----+-----+

<----- SEQ7B

-----  
-----