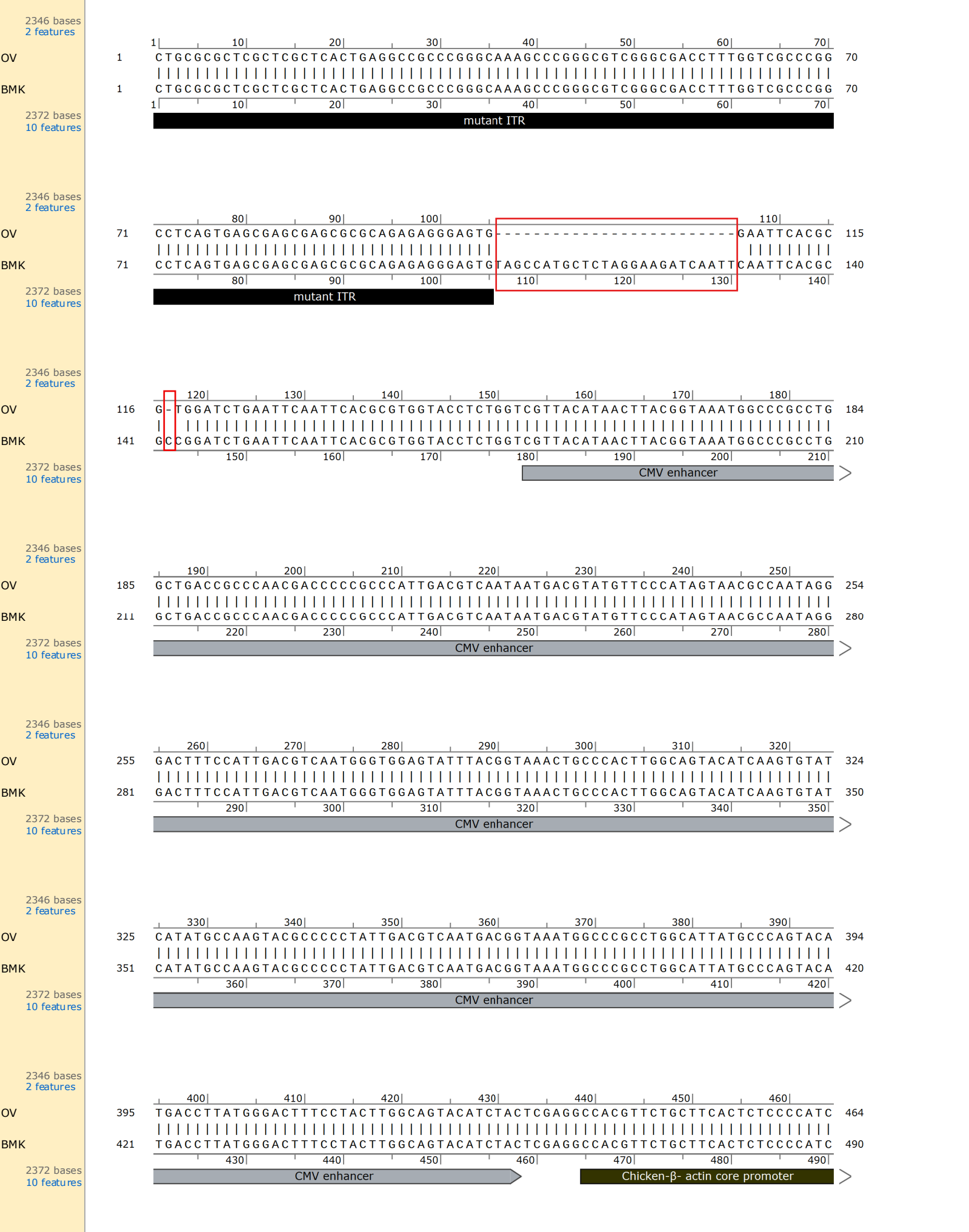


Appendix Table of Contents

1. Appendix Fig S1
2. Appendix Fig S2



2372 bases
10 features

2346 bases
2 features

2346 bases
2 features

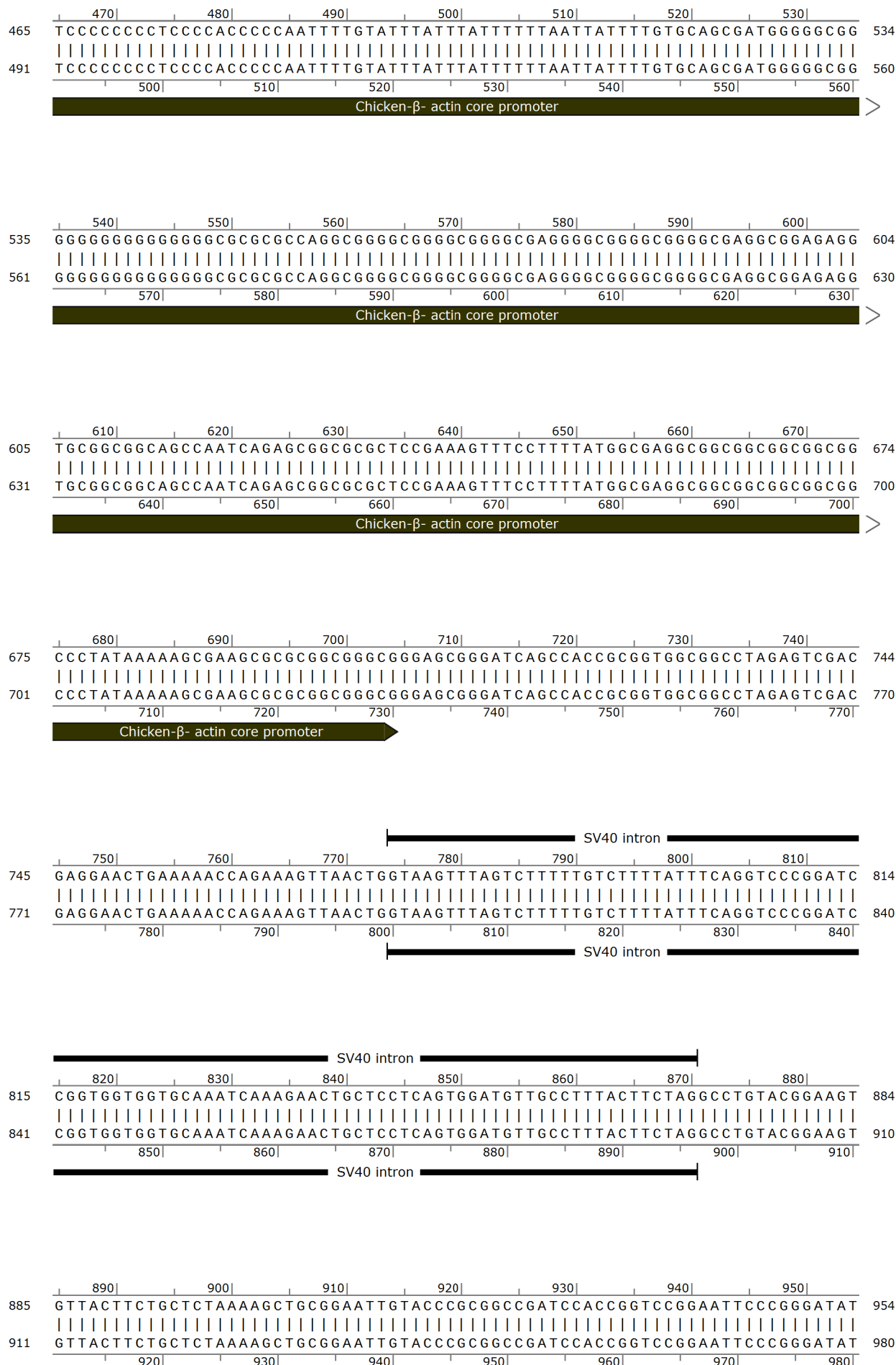
2346 bases
2 features

2346 bases
2 features

2346 bases
2 features

2346 bases
2 features

2372 bases
10 features



2346 bases
2 features

OV

BMK

2372 bases
10 features

960|970|980|990|1000|1010|1020|

955CGTTCGACCCACGCGTCCGGGGCCCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGGCGGCA1024

981CGTTCGACCCACGCGTCCGGGGCCCCACGCTGCGCACCCGCGGGTTTGCTATGGCGATGAGCAGCGGGCGGCA1050

990|1000|1010|1020|1030|1040|1050|

MetAlaMetSerSerGlyGly

SMN cDNA

2346 bases
2 features

OV

BMK

2372 bases
10 features

1030|1040|1050|1060|1070|1080|1090|

1025GTGGTGGCGGCGTCCCGGAGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTC1094

1051GTGGTGGCGGCGTCCCGGAGCAGGAGGATTCCGTGCTGTTCCGGCGCGGCACAGGCCAGAGCGATGATTC1120

1060|1070|1080|1090|1100|1110|1120|

SerGlyGlyGlyValProGluGlnAspSerValLeuPheArgArgGlyThrGlyGlnSerAspAspSer

SMN cDNA

2346 bases
2 features

OV

BMK

2372 bases
10 features

1100|1110|1120|1130|1140|1150|1160|

1095TGACATTTGGGATGATACAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTA1164

1121TGACATTTGGGATGATACAGCACTGATAAAAGCATATGATAAAGCTGTGGCTTCATTTAAGCATGCTCTA1190

1130|1140|1150|1160|1170|1180|1190|

AspIleTrpAspAspThrAlaLeuIleLysAlaTyrAspLysAlaValAlaSerPheLysHisAlaLeu

SMN cDNA

2346 bases
2 features

OV

BMK

2372 bases
10 features

1170|1180|1190|1200|1210|1220|1230|

1165AAGAATGGTGACATTTGTGAACTTCGGGTAAACCAAAAACACACCTAAAAGAAAACCTGCTAAGAAGA1234

1191AAGAATGGTGACATTTGTGAACTTCGGGTAAACCAAAAACACACCTAAAAGAAAACCTGCTAAGAAGA1260

1200|1210|1220|1230|1240|1250|1260|

LysAsnGlyAspIleCysGluThrSerGlyLysProLysThrThrProLysArgLysProAlaLysLys

SMN cDNA

2346 bases
2 features

OV

BMK

2372 bases
10 features

1240|1250|1260|1270|1280|1290|1300|

1235ATAAAAGCCAAAAGAAGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCAT1304

1261ATAAAAGCCAAAAGAAGAATACTGCAGCTTCCTTACAACAGTGGAAAGTTGGGGACAAATGTTCTGCCAT1330

1270|1280|1290|1300|1310|1320|1330|

AsnLysSerGlnLysLysAsnThrAlaAlaSerLeuGlnGlnTrpLysValGlyAspLysCysSerAlaIle

SMN cDNA

2346 bases
2 features

OV

BMK

2372 bases
10 features

1310|1320|1330|1340|1350|1360|1370|

1305TTGGTCAGAAGACGGTTGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTT1374

1331TTGGTCAGAAGACGGTTGCATTTACCCAGCTACCATTGCTTCAATTGATTTTAAGAGAGAAACCTGTGTT1400

1340|1350|1360|1370|1380|1390|1400|

TrpSerGluAspGlyCysIleTyrProAlaThrIleAlaSerIleAspPheLysArgGluThrCysVal

SMN cDNA

Printed from SnapGene®: Mar 2, 2023 9:09 PM

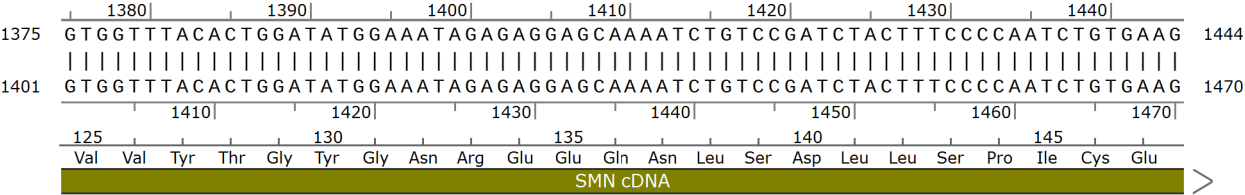
Page 3

2346 bases
2 features

OV

BMK

2372 bases
10 features

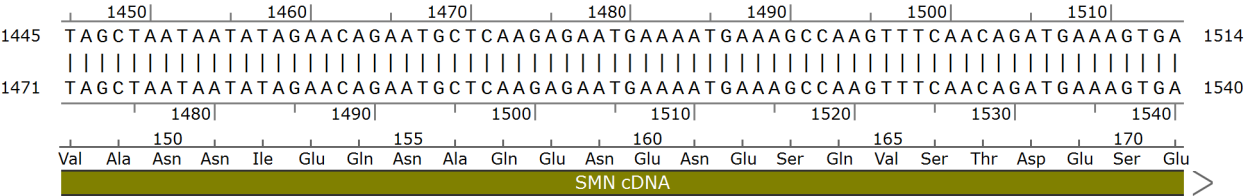


2346 bases
2 features

OV

BMK

2372 bases
10 features

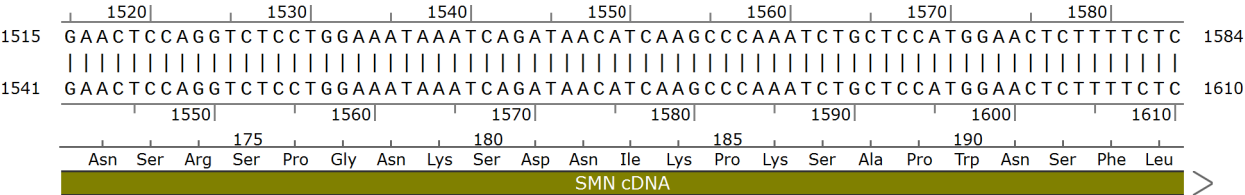


2346 bases
2 features

OV

BMK

2372 bases
10 features

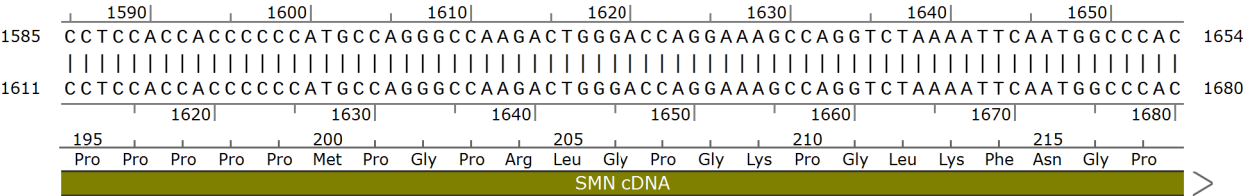


2346 bases
2 features

OV

BMK

2372 bases
10 features

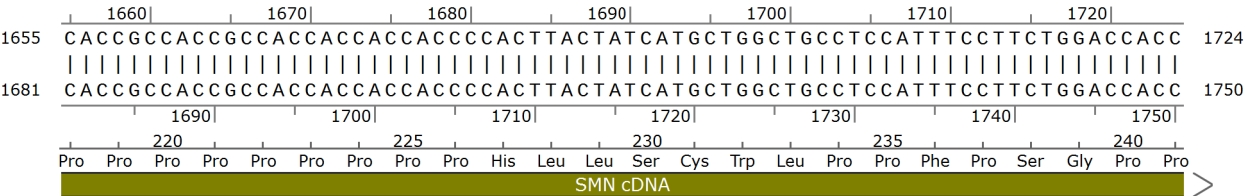


2346 bases
2 features

OV

BMK

2372 bases
10 features

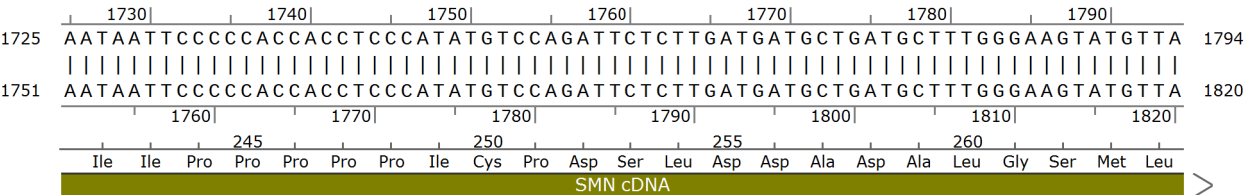


2346 bases
2 features

OV

BMK

2372 bases
10 features





2346 bases
2 features

OV

BMK

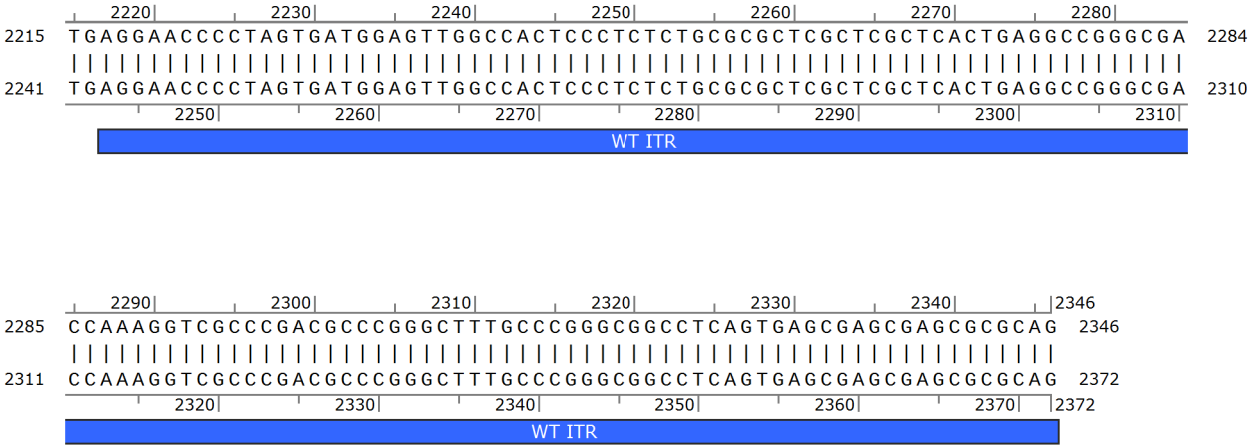
2372 bases
10 features

2346 bases
2 features

OV

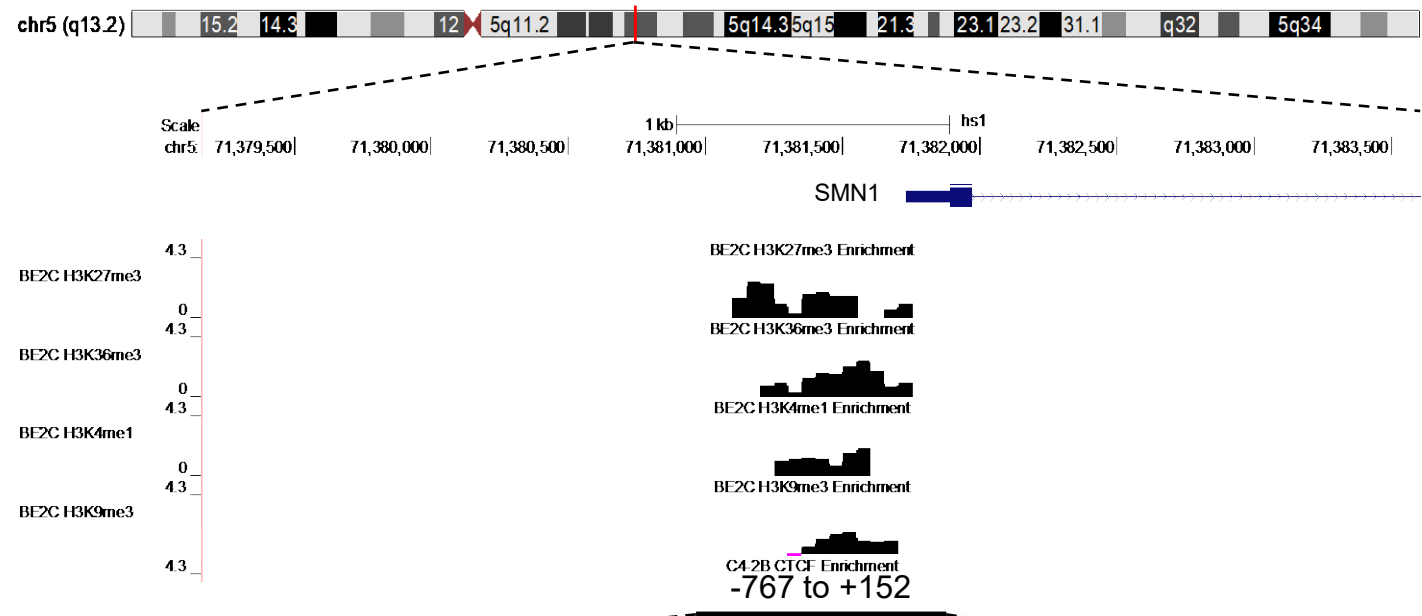
BMK

2372 bases
10 features

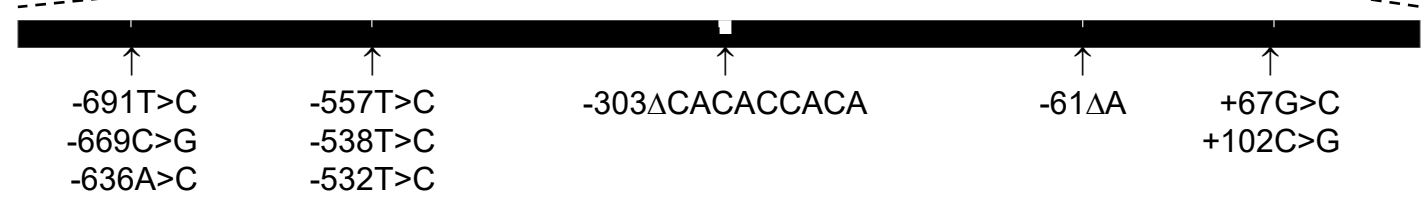


Appendix Figure S1. Vector genome alignment between the benchmark vector and Zolgensma®. Unaligned sequences are indicated by the red box.

A



B



Appendix Figure S2. The SMN promoter (-767 to +152) encompasses enrichment of all epigenetic marks within 2 kb of the transcriptional start site. (A) UCSC genome browser view of epigenetic marks (H3K27me3, H3K36me3, H3K4me1, and H3K9me3) as reported by T2T Encode ENCODE Reanalyses (PMID: 35357915). The four tracks shown for BE2C (neuroblast), are from the most relevant cell line in the T2T CHM13v2.0/hs1 build. **(B)** The SMN promoter, as described in Echaniz-Laguna *et al.*, and used in our 2nd-generation vector design, has ten differences from the T2T build. Each position reported here is in reference to the TSS (+1) reported in Echaniz-Laguna *et al.* For example, “-691T>C” means that the base at position -691 was reported as a thymine (T) in Echaniz-Laguna *et al.* and is annotated as a cytosine (C) in the T2T build; at position -303, there is a 9-nt deletion of CACACCACA