

Supplementary Figures for

Complex Alternative Splicing of the Smarca2 Gene Suggests the Importance of Smarca2-B Variants (Min Yang, et al.)

ACCCCTGAACCAGATGATTGCTGCCGGGAGGAAGAATTGATCTTTTATGCATGGACATGGACCGGGAGGA
TGCCCAGAACCGAAGCGAAACCCGCTTGATGGAGGAAGATGAGCTGCCCTCCTGGATTATCAAGGATGACGCCGAAG
TGGAAAGGCTCACCTGTGAAGAAGAGGAGGAGAAGATATTGGGAGGGCTCGCCAGGCCGGATGTGGACTACAGT
GATGCCCTCACCGAGAACGATGGCTCAGGCCATCGAACGGCAATTGAAAGAAATGAAAGAGGAGGTACGGCTTAA
GAAGAGAAAAAGACGAAGAAATGTGGATAAAGACCCGTGAAGGAAGATGTGGAAAAGCGAACGAAAGAGGAGGCC
CTCCGGCTGAGAAGTGTCAACCAAATCCCCAAAACTAACGAAGCAGATGAACGCCATCATTGATACTGTGATAAACTAC
AAAGACAGGTAACTGGAGAAGGTGCCCTGTAATTCTCAGTTGAAATAGAAGGAAACAGTTCAGGGCGACAGCTCAG
TGAAGTCTCATTAGTTACCTTCCAGGAAAGACTTACCAAGAATACTATGAATTAAATTAGGAAGCCAGTGGATTCAAA
AGATAAAAGGAGCGAATCCGTAATCATAAGTAT

Fig. S1. Sequence of the region between exon 26 and exon 31 of the mSmarca2-a1 cDNA. Exon 26 and exon 31 are incomplete. Every other exon is presented in grey color to show the boundary of the neighbor exons. The underlined part is the 54-bp exon 29 that is skipped in the mSmarca2-a2 (NM_011416.2) in the NCBI database.

mSmarca2-b3	CCCCACTAAGTGTATTCCGATAGCCGGCTGCTGATAGTGGTAAGGCCATCGAAGACG
mSmarca2-b4	CCCCACTAAGTGTATTCCGATAGCCGGCTGCTGATAGTGGTAAGGCCATCGAAGACG
mSmarca2-b3	GCAATTGGAAGAAATGGAAGAGGGAGGTACGGCTTAAGAAGAGAAAAAGACGAAGAAATG
mSmarca2-b4	GCAATTGGAAGAAATGGAAGAGGGAGGTACGGCTTAAGAAGAGAAAAAGACGAAGAAATG
mSmarca2-b3	TGGATAAAAGACCCCGTGAAGGAAGATGTGGAAAAGCGAAGAAAAGAAGAGGCCGCCCTC
mSmarca2-b4	TGGATAAAAGACCCCGTGAAGGAAGATGTGGAAAAGCGAAGAAAAGAAGAGGCCGCCCTC
mSmarca2-b3	CGGCTGAGAAGTTGTCACCAATCCCCAAAACTAACGAAGCAGATGAACGCCATCATTC
mSmarca2-b4	CGGCTGAGAAGTTGTCACCAATCCCCAAAACTAACGAAGCAGATGAACGCCATCATTC
mSmarca2-b3	ATACTGTGATAAACTACAAAGgtgtaacgtggagaaggtgccctgttaattctcagttgga
mSmarca2-b4	ATACTGTGATAAACTACAAAG-----
mSmarca2-b3	aatagaaggaaacagACAGTTAGGGCGACAGCTCAGTGAAGTCTTCATTCAAGTTACCTT
mSmarca2-b4	-----ACAGTTAGGGCGACAGCTCAGTGAAGTCTTCATTCAAGTTACCTT
mSmarca2-b3	CCAGGAAAGACTTACCAGAATACTATGAATTAATTAGGAAGCCAGTGGATTCAAAAAGA
mSmarca2-b4	CCAGGAAAGACTTACCAGAATACTATGAATTAATTAGGAAGCCAGTGGATTCAAAAAGA
mSmarca2-b3	TAAAGGAGCGAATCCGTAATCATAAGTA
mSmarca2-b4	TAAAGGAGCGAATCCGTAATCATAAGTATCGGAGC

mSmarca2-b10:

Fig. S2: Part of the sequences of the three novel mSmarca2-b variants. **Top panel:** Sequences of the region between exon 2 and exon 7 of the mSmarca2-b3 and mSmarca2-b4. The exon 2 and exon 7 is incomplete. The exon 2 is presented in grey color to show the boundary between the 96-bp exon 2 and the 218-bp exon 4. Different from the mSmarca2-b2 documented in the NCBI database (NM_026003.2), both mSmarca2-b3 and mSmarca2-b4 lack the 35-bp exon 3 and the mSmarca2-b3 retains the 54-bp exon 5. **Bottom panel:** part of the mSmarca2-b10 sequence. The sequences underlined at the 5'- and 3'-ends are forward primer mSmarca2F'43 and reverse primer mSmarca2R4524, respectively. The lowercase letters are the 408 nucleotides of intron 1 retained by exon 1, within which the underlined atg is the start codon that adds two amino acids to the proteins starting from the ATG1 at exon 2.

5' RACE-a -----AGAGGCAGTTGAGCATGGGAGCCGTCTGTATGTAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 5' RACE-b ATTAATCTAGAGGCAGTTGAGCATGGGAGCCGTCTGTATGTAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b1 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b2 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b3 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b4 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b5 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b6 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b7 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b8 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG
 hSmarca2-b9 -----CTGTATGTTGAATTAGGGCTCGCACTCTTGCACAAACAGTCACCAGTCGGAAACTG

 5' RACE-a GGGGTTTGCTCTGTGATTATTTCATTATTGTCG-----CTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 5' RACE-b GGGGTTTGCTCTGTGATTATTTCATTATTGTCGTTGAAAAGCTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b1 GGGgttgctctgtgatttattcattattgtgctggtaaaagctgatgaagagactagcagctcgctgtttgtggctttaatt
 hSmarca2-b2 GGGGTTGCTCTGTGATTATTTCATTATTGTCGTTGAAAAGCTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b3 GGGGTTGCTCTGTGATTATTTCATTATTGTCGTTGAAAAGCTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b4 GGG-----CTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b5 GGGGTTGCTCTGTGATTATTTCATTATTGTCG-----CTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b6 GGGGTTGCTCTGTGATTATTTCATTATTGTCG-----CTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b7 GGGGTTGCTCTGTGATTATTTCATTATTGTCGTTGAAAAGCTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b8 GGGGTTGCTCTGTGATTATTTCATTATTGTCGTTGAAAAGCTGATGAAGAGACTAGCAGCTCGCTGTTGGCTGTTAATT
 hSmarca2-b9 GGG-----

 5' RACE-a TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAGTGCCTGGCTCTGGTTAAAGCAAGCGTTGCAGGCCAT
 5' RACE-b TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b1 tatccccactaaactgtgatttctgatagccggcctgctgatagtggtaagtgcgtctggtaaaagcaagcggttcagGCCAT
 hSmarca2-b2 TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b3 -----GCCAT
 hSmarca2-b4 TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b5 TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b6 TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b7 TATCCCCACTAACTGTGATTCTGATAGCCGGCCTGCTGATAGTGGTAAG-----GCCAT
 hSmarca2-b8 -----GCCAT
 hSmarca2-b9 -----GCCAT

 5' RACE-a CGAACCGGCAATTGGAGGA
 5' RACE-b CGAACCGGCAATTGGAGGA
 hSmarca2-b1 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b2 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b3 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b4 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b5 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b6 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b7 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b8 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA
 hSmarca2-b9 CGAACCGGCAATTGGAGGAATGGAAGAGGAAGTACGGCTTAAGAACGGAAAAAGACGAAGAAATGGATAAAGATCCTGAAAGA

 hSmarca2-b1 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b2 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b3 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b4 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b5 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b6 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b7 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b8 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC
 hSmarca2-b9 AGATGTGAAAAGCTAAGAACGAGAGGCGCCCTCCCGCTGAGAAACTGTACCAAATCCCCCAAAGTCACAAAGCAGATGAACGC

 hSmarca2-b1 TATCATCGATACTGTGATAAAACTACAAGATAGgtgtacgtggagaagggtcccagtaatttcagttggaaatagaaggaaacagTTC
 hSmarca2-b2 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC
 hSmarca2-b3 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC
 hSmarca2-b4 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC
 hSmarca2-b5 TATCATCGATACTGTGATAAAACTACAAGATAGGTGTAACGTGGAGAAGGTGCCAGTAATTCTCAGTTGGAAATAGAAGGAAACAGTTC
 hSmarca2-b6 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC
 hSmarca2-b7 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC
 hSmarca2-b8 TATCATCGATACTGTGATAAAACTACAAGATAGGTGTAACGTGGAGAAGGTGCCAGTAATTCTCAGTTGGAAATAGAAGGAAACAGTTC
 hSmarca2-b9 TATCATCGATACTGTGATAAAACTACAAGATAG-----TTC

 hSmarca2-b1 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b2 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b3 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b4 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b5 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b6 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b7 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b8 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b9 AGGGCGACAGCTCAGTGAAGTCTTCAATTCAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC

 hSmarca2-b1 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b2 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b3 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b4 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b5 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b6 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b7 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b8 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC
 hSmarca2-b9 CAAAAAAATAAGGAAGGATTCTGTAATCATAAGTACCGGAGCCTAGGGACCTGGAGAAGGATGTCATGCTTCTGTCACAAACGCTC

Fig. S3: Sequences of the nine hSmarca2-b variants and two of the original 5'RACE sequences. These sequences are aligned to project out the differences among the variants. Every other exon is presented in grey color to show the boundary of the neighbor exons. The lowercase letters in the hSmarca2-b1 indicate the parts that may be involved in alternative splicing.
