

Av7-1 TTGTGAAGCCAAGAAGGAAACATCAACAAAGAAAACAGTCAAATTTTGATTCAA 54
 Av7-2 ***** 54
 Av7-3 -----
 Av10 -----

Av7-1 GTAAGTTTGATGGATGTCATTTCTTCTTACGATTTTATTTAAAGAATATATACT 108
 Av7-2 *****_* 107
 Av7-3 -----
 Av10 -----

Av7-1 GTATTTATACTCATTATTATATCATGATTCCCATCAAACCGACCCTTTTCCTTT 162
 Av7-2 *****_* 160
 Av7-3 -----
 Av10 -----

Av7-1 CGAAGCAACAGGCGGGTTAGGACTGTACTAAACCCTCCTGAGTTATAATTATAA 216
 Av7-2 ***** 214
 Av7-3 -----
 Av10 -----

Av7-1 AATTAAGAGATTTTATTTCTGTTTGTTTCCTACAATCTAATCAAACCAAACAC 270
 Av7-2 ***** 268
 Av7-3 -----
 Av10 -----

Av7-1 ATCAGGAGTGACCAAGAGCGACTTTTGTTTATCAGTATGTTCTTTTCTTTATTT 324
 Av7-2 *****_* 321
 Av7-3 -----
 Av10 -----

Av7-1 CCACGCTTACCCTCAAGGGCTGGGTATGTCCACAGAACAGAACTTCTCCACTAC 378
 Av7-2 ***** 375
 Av7-3 -----
 Av10 -----

Av7-1 GGGTAGCACTAGCACTTTTTTATCACTAAAGGAGCATTTTCAATGAATAAGCGA 432
 Av7-2 *****_A* 429
 Av7-3 -----
 Av10 -----

Av7-1 AATATCTTTATGATGTCACCATTTTACTGTTGGCCATCTAGAATGGCATCTCTC 486
 Av7-2 ***** 483
 Av7-3 -----
 Av10 -----

Av7-1 TCTGGACGTGTCTTTTGTAAAAAAAATAAAAAATTTAGCCGCATGCGTCTAC 540

Av7-2 ***** 537
Av7-3 -----
Av10 -----

Av7-1 TTTTGTAAATCTGTTAAATACACAAAAAATGGTTCATACATTCTTTTTTAACAG 594
Av7-2 *****_***** 591
Av7-3 -----
Av10 -----

[M] M N R L L V F L M L G A/- A F

Av7-1 AAAGCTACAGTGATGAACCGTCTTTTGGTCTTTCTGATGCTCGGGGCAGCTTTT 648
Av7-2 ***** 645
Av7-3 -----A*****C*****_*****C 42
Av10 -----A*****C*****_*****C 42

Av7 SID (part 1)

M L V V S A N D A Y G

Av7-1 ATGCTTGTCTTTT CAGCCAATGATGCTTATGGCGTAAGATTCTAAATGGGGACA 703
Av7-2 ***** 700
Av7-3 *****G***** 97
Av10 *****G***** 97

Av7-1 TAGAGAGAGTAGAAGAGAGGGAGAGAGAGGGAGTTTGAGAGA----- 744
Av7-2 *****AG*TT***----- 741
Av7-3 *****AGA*****---AATTGAGT 146
Av10 *****AGA*****GAGAAATTGAGT 150

Av7-1 -----GAGAGAGGCAAGTTGTGTACTCACAAATATGTGCACAATT 784
Av7-2 -----GA*****TC*_***** 782
Av7-3 TAGGGAGATTGAGA*****C*_***** 199
Av10 TAGGGAGATTGAGA*****C*_***** 203

Av7-1 GTTCGTCGCTATATCGGATCTTTATCGGATAATTCATGCTGAAATTAATGCCAG 838
Av7-2 *****G*****----- 828
Av7-3 *****G*****GT***--- 250
Av10 *****G*****GT***--- 254

Av7-1 TAATAGATGGATACAAAGCTATCTTGACCAGAAATGTGGAACACAACCTCAACTC 892
Av7-2 ----- 828
Av7-3 -----*C**G**A**GCAAC*CT**T**GATT*T**AA**A**G***CTA*T**-- 296
Av10 -----*C**G**A**GCAAC*CT**T**GATT*T**AA**A**G***CTA*T**-- 300

Av7-1 ATGAATAATGTTACGATTCTCTTGTTAAAATTCTTTTTTCTGAGCAAACCTTTTT 946
Av7-2 ----- 828
Av7-3 ----- 296
Av10 ----- 300

Av7-1	CATG TTCAGTTGGTTAGTGTGCGTCCTTCGTAGATGAAGGCCCTCGGTTCTGACT	1000
Av7-2	-----	828
Av7-3	-----	296
Av10	-----	300
Av7-1	CCCAGGATCCACAAATCTATCTGCTTTGACTTCTCTTTTCCGTGTAGCTACAG	1054
Av7-2	-----	828
Av7-3	-----A*GT*G****C**GCGTGTC*A*T**CC*T*T**C*TAC***A	339
Av10	-----A*GT*G****C**GCGTGTC*T*T**CC*T*T**C*TAC***A	343
Av7-1	CTTAAATAACTTTGTAAGGAGCACTGATGGTTTGAGGGGGTCAGTAAAGTGCTC	1108
Av7-2	-----	828
Av7-3	G***C*ATT***C*G-----*CC***	361
Av10	G***C*ATT***C*G-----*CC***	365
Av7-1	ACCTTCGATTGC-TTAAACTATTTGGTAATCAACCTCTCCTTTACCTATGCCTT	1161
Av7-2	-----	828
Av7-3	G*****CATG***C*****-----	383
Av10	G*****CATG***C*****-----	387
Av7-1	TTTTTTCAGACGAGCCGGCATT TAAAGACCTCAAACGAGGTATCGTATAAAGTG	1215
Av7-2	-----	828
Av7-3	-----***TCT*GGA*G**AGGGG*GCCCC**A	411
Av10	-----***TCT*GGA*G**AGGGG*GCCCC**A	415
Av7-1	TTTAAATTATCGAATTTTAGTGAAACTGACAATGTGAAAATGATATATTCAA	1269
Av7-2	-----	828
Av7-3	C*A*C*A*G*-----**CA*****ATCC*G*C**C*	442
Av10	C*A*C*A*G*-----**CA*****ATCC*G*C**C*	446
Av7-1	GAACAAAT-CCATTAATATAATGTCATTCAGGGGAAGAAAACCTGAAGGA-ATTA	1321
Av7-2	-----**G**G*****G*T-----**ATG**C*****C*T*A*T*A*G	867
Av7-3	A**TT**G**G*****G*T-----**ATG**C*****C*T*A*T*A*G	486
Av10	A**TT**G**G*****G*T-----**ATG**C*****C*T*A*T*A*G	490
Av7-1	TTAAACTCCTACCTTTAATGTTTTCTTTCTGCCATTTTCATCATTGTTTTTTTT	1375
Av7-2	***CGA*G**CTTG****AA**C*T**T-CTG*GC*-----AGC*****	910
Av7-3	***CGA*TG*CTTG****AA**C*T**T*CTG*GC*-----AGC*****	530
Av10	***CGA*TG*CTTG****AA**C*T**T*CTG*GC*-----AGC*****	534
Av7-1	TTATG TTCAGTTGGTTAGTATGCGGCCCT-GTAGATGAAGGCCCTCGGTTCTGAC	1428
Av7-2	*C*****G*****T*C*****	964
Av7-3	GC*****G*****T*C*A*****A*****	584
Av10	GC*****G*****T*C*A*****A*****	588
Av7-1	TCCC---AGGATGCACAAATCTGCTTTGACTTCTCTTTGTC-----	1466

Av7-2 ****--****C****G*****CGTGCTAGCTACAG 1015
 Av7-3 ****TGC***T*T****G*****CGTGCTAGCTACAG 638
 Av10 ****TGC***T*T****G*****CGTGCTAGCTACAG 642

Av7-1 -TTAAATACCTTTGTAAAGAGCACTGATGGTTTGAGGGGGTCAGTAAAGTGCT 1519
 Av7-2 C*****A***G-*****A*****C*T***** 1068
 Av7-3 C*****A*****G-***** 691
 Av10 C*****A*****G-***** 695

Av7-1 CACCT-----CGATTGCTTTTAAACTATTTGGTAATCAACCTACTCCTTTAC 1566
 Av7-2 *****TAAAATT*****_*****_***** 1119
 Av7-3 *****TAAAATT*****_*****_***** 742
 Av10 *****TAAAATT*****_*****_***** 746

Av7 SID (part 2) → ● Av7 S2D (part 1)

(D) E P A F K D L N Q

Av7-1 CTATGCCTTTTTTTTTC-AGACGAGCCGGCATTTAAAGACCTCAACCAAGGTAT 1619
 Av7-2 *****C*C***** 1173
 Av7-3 *****CA--***** 794
 Av10 *****_***** 798

Av7-1 CGTATAAAGTGTTTAAG-TTATCGAATTTTAGTGGAACCTGACAATGCGAAAAT 1672
 Av7-2 *****_*****T*T*****A 1226
 Av7-3 *****C*****G** 848
 Av10 ***** 852

Av7-1 AATATATTC-----TTGCAATAAAGTCACTTTGGTGTTTCCAAAGAACAATCC 1721
 Av7-2 *****ATGTC***** 1280
 Av7-3 *****_***** 897
 Av10 *****_***** 901

Av7-1 ATTAATATGATGTCATTCAGGAGAAGAAAAGTGAAGGAATTATTAACCTCCTG- 1774
 Av7-2 *****A*****A***G***AG 1334
 Av7-3 *****_ 950
 Av10 *****_ 954

Av7-1 -CTTTAATGTTTTCTTCTGCCATTTTCATCATTTGTT-GTTTTTTTATACTT 1826
 Av7-2 CT*****T*****T*****TA***** 1388
 Av7-3 -*****_***** 1002
 Av10 -*****C*****_***** 1006

Av7 S2D (part 2) →

(G)D E S L G K

Av7-1 GCATTTTAGGTGACGAAAGTTTGGGAAAAAGTAAGTATTCATGCAAAGCATGCC 1880
 Av7-2 *****T 1442
 Av7-3 ***** 1056
 Av10 ***** 1060

Av7-1 ATTCGTTGCCTTAATTAATATTTTTCAATATTTT-AGTGAGAA-GGTAAATGTG 1932
 Av7-2 *****T*****A***** 1496
 Av7-3 *****_*****_*****_***** 1107
 Av10 *****_*****_***** 1112

Av7-1 ATGTAATCGAAGGACTCAAGGACTTTGATGGATGTGACACGTCAGTTTGAGAAG 1986
 Av7-2 ***** 1550
 Av7-3 ***** 1161
 Av10 ***** 1166

Av7-1 TGTGAGAAGAGTAGCCGCTTGTGTGTTAATGCCACGGCGTGCTTGAAAACGGG 2040
 Av7-2 *****G*****G*****C*****G***** 1604
 Av7-3 *****G*****G***** 1215
 Av10 *****G***** 1220

Av7-1 GGGGAGATAGGGGATGTAGCCAAGTACCGGGCGTCAGTGAAGGGCATCGGATT 2094
 Av7-2 **_***** 1656
 Av7-3 **_***** 1267
 Av10 **_***** 1272

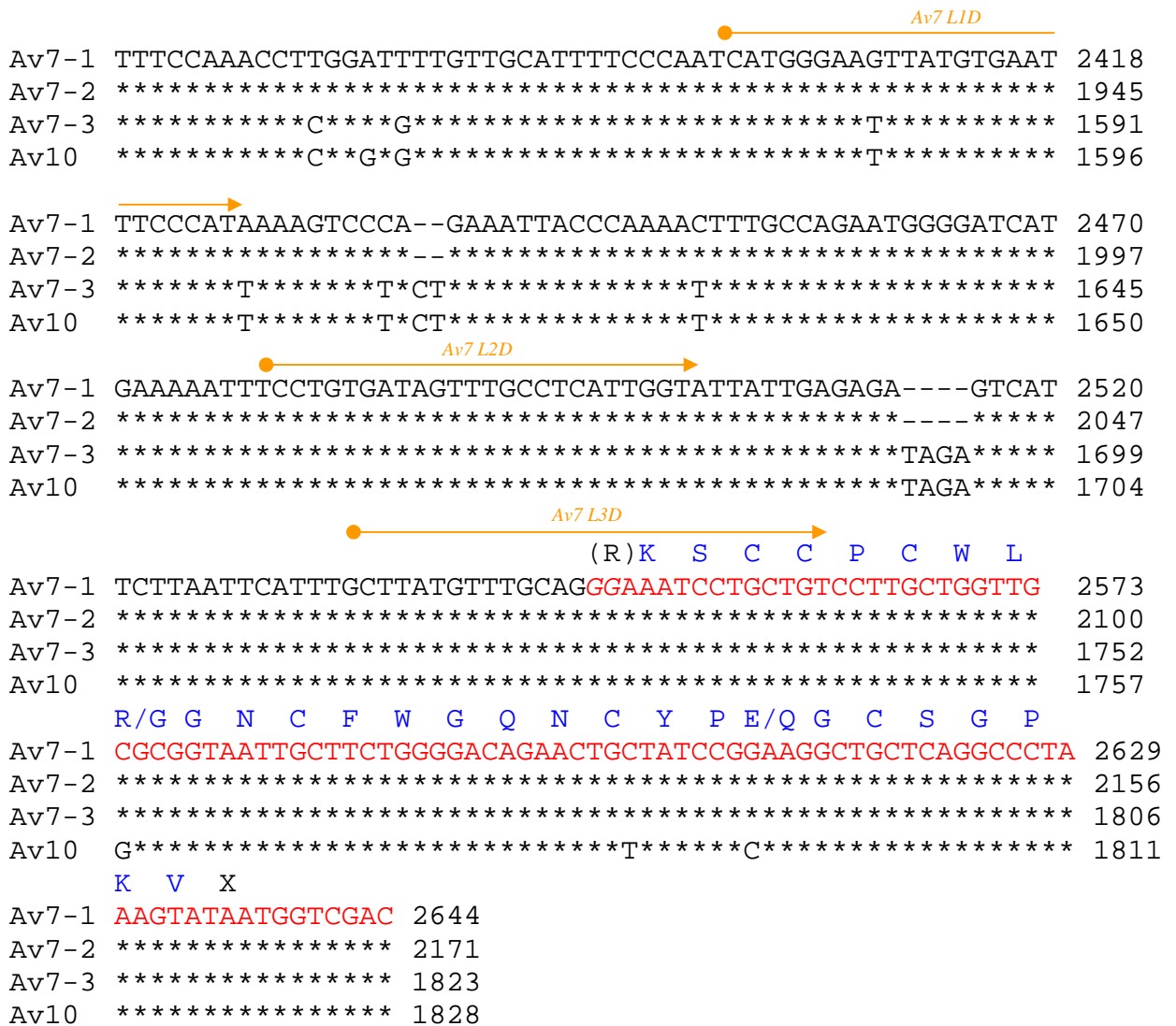
Av7-1 AGTCGTAAGATTAGCTACAGTTTTTTTCATGCGCAGGTTGCATTTTTTAAAAGCTG 2148
 Av7-2 *****_***** *TA* 1675
 Av7-3 *****GC*****T** 1321
 Av10 *****GC*****T** 1326

Av7-1 ACGTACTTCAGCAACCATTACACACTGTGGACCAGGCACTGCTCCATCTTCAA 2202
 Av7-2 *****G** 1729
 Av7-3 *****^{Av7-2D}***** 1375
 Av10 ***** 1380

Av7-1 GCATTTGGGGAAGTGTATGGATCACTGATGCTAATTGTTTATTTCAGTAACTTG 2256
 Av7-2 *****C*****G*****A***** 1783
 Av7-3 ***** 1429
 Av10 ***** 1434

Av7-1 CCAGATTTGCATACACAGGAAACGTTTTCAATTTCCCTAAATGGGAACCTTGGGA 2310
 Av7-2 *****C***** 1837
 Av7-3 ***** 1483
 Av10 ***** 1488

Av7-1 AATTATTGGGAAATCACAAACCAAATTTGGGAAAATCTTTTTGAAATGAAATTT 2364
 Av7-2 *****T***** 1891
 Av7-3 ***** 1540
 Av10 ***** 1545



Supplementary Fig. 2 Nucleotide alignment of Av7 and Av10 genomic sequences. Identical residues are indicated by asterisks and gaps by hyphens. Exons are in red and introns are in black. Amino acid translation appears above the nucleotide sequence. Amino acids of the mature toxin are in blue. Parenthesized amino acids are encoded by nucleotides from two exons. The stop codon is denoted in the translation as "X". The location of the ATG start codon varies between the sequences and is indicated in bold and square parentheses of the first methionine which corresponds only to Av7-3 and Av10. Primers used in genome walking for differentiating between processed pseudogenes and ancestral Av7 genes are marked by magenta and orange arrows, respectively.