

Supplemental Table 3: The MAST2 5' transduction group of SVA elements (SVA_F1 subfamily)

A. SVA_F1 elements without additional transductions (group1): 53

Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	Remarks	Source element	Polymorphic state
H15_5	Chr.15:73559835-73562195	aaaatacaaaaa	382	1080	AluY insertion in SINE-R Part		
H17_2	Chr.17:39612544-39614492	aaaaaaaaaaaaaaaa	379	1246	premature polyadenylation at pos.248		fixed
H3_9	Chr.3:11503287-11505162	aaacagaacca	362	904	19bp extra 5' ¹		
H12_9	Chr.12:40906890-40909515	aaaacattaaata	362	1664	8bp extra 5' ¹	H3_9	
H6_F_70	Chr.6:43655601-43657642	aagattcttcactc	362	1077	POLH* 6bp extra 5' ¹	H3_9	fixed
H9_3	Chr.9:38231219-38233623	aaaattacaatgga	361	1450			
H9_17	Chr.9:120336579-120339295	aaagaatcatca	352	1758			
H5_1	Chr.5:75502336-75504673	aaaaattaaaaaaact	335	1398	RAP1B*		polymorphic
H5_17	Chr.5:177877469-177878742	aaagaaccgctgttg	335	339	1 mismatch in TSDs		fixed
H5_F_158	Chr.5:167218110-167220665	gaaaacaacgtcaa	335	1885	NEURESTIN*		fixed
H4_17	Chr.4:158914566-158916811	gaaaagtt	300	1363			
HX_2	Chr.X:71683453-71685840	aaagaaagttttc	295	1491			
H1_F_95	Chr.1:46417889-46420661	gaaagtcagacaatta	295	1885			polymorphic
H8_9	Chr.8:128397069-128399331	no	286	1411			
H19_11	Chr.19:22899898-22903267	gaaattatactgtc	283	2492			
H1_F_281	Chr.1:211430305-211432415	ggaggagcagaga	250 (230)	1255	20nts internal insertion in MAST-2		fixed
H1_7	Chr.1:26827864-26829997	aacatctcccag	234	1645			
H14_2	Chr.14:49610281-49611411	no	180	378			
H7_F_35	Chr.7:24830685-24833110	aaagaatataaaagtc	158	1672			polymorphic
H17_1	Chr.17:55982862-55985041	aaattttt	153	1450			
HX_3	Chr.X:62190984-62193262	no	148	1546			
H6_F_57	Chr.6:34472623-34474910	gaaaaat	148	1558			fixed
H8_16	Chr.8:9551517-9553375	gaaaattctttctt	148	1123	TNKS*		polymorphic
H8_F_89	Chr.8:95671494-95673779	aaataatttaaaaaaaaa	148	1562			fixed
H1_F_160	Chr.1:113152208-113154115	aaagaacttggtttt	128	1126	source element for H6_6		fixed
H22_16	Chr.22:39220810-39223350	aaaagaaagcaatag	125	1753			fixed
H12_1	Chr.12:112772309-112774247	agaaaaggcactgtg	92	1246			fixed
H11_17	Chr.11:45715624-45717399	taaatccagcaaaaaata	89	1081			fixed
H1_8	Chr.1:54738766-54740851	taaaaaatcag	89	1397			
H13_17	Chr.13:73046347-73048394	aacaggatgiga	89	1354			
H3_4	Chr.3:102640328-102642142	aaaaaagaatgagtac	88	1129			fixed
H7_F_22	Chr.7:10277540-10279390	aaagaaaagaggtttaa	88	1137	BC002644*		polymorphic
H8_F_65	Chr.8:65760244-65762269	aacattaggca	88	1355			fixed
H9_4	Chr.9:36604406-36606104	caaaaaataattttt	86	1010	MELK*		fixed
H1_2	Chr.1:24064489-24066855	aaagaccacaatcagt	86	1680	source element for H6_2		
H3_18	Chr.3:57707655-57709588	aaaagcataatcaag	86	1228			
H2_F_129	Chr.2:113867557-113869857	aaaaacagttgctg	82	1615			polymorphic
H15_17	Chr.15:39461355-39463561	aaaaaaaaaaagatcaag	82	1528			
HX_1	Chr.X:114404045-114405955	aaagttaggc	76	1248			fixed
HX_17	Chr.X:24486146-24487906	aaaaaaaaaaaaaaaaag	76	1097			
H2_F_134	Chr.2:127592678-127594386	aaagaaggacttga	76	1031			fixed
H6_F_119	Chr.6:87775471-87777327	aagaagtgtggcgtatag	74	1179			fixed
H3_1	Chr.3:180319931-180322610	agaactggttaatt	68	2015			
H14_1	Chr.14:34577452-34579273	aaagctaccgaatag	67	1140	C14orf24*		polymorphic
H21_18	Chr.21:18012804-18015373	agaaatattccactc	64	1901			
H10_2	Chr.10:76762222-76764027	gaaaaattta	64	1155			
H2_F_65	Chr.2:43435616-43437419	aaaattgtaattaac	64	1150			fixed
H20_1	Chr.20:16155154-16157130	aaaaagaattggagag	63	1304			fixed
H10_22	Chr.10:24984450-24987019	aaaaaatgctcatcatc	49	1928			
H15_1	Chr.15:33735022-33737117	aagtagtgggtgtt	49	1434			fixed
H10_9	Chr.10:94423427-94425776	gagagatggatgtg	47	1688	HHEX*		fixed
H3_6	Chr.3:38097285-38098992	aagaacacccgaacc	35	1073			
H16_17	Chr.16:1588754-1590491	aaatgtacacaggtg	14	1130			

¹ H3_9 contains 19bp untemplated sequence at its 5' end; the 3'-most 8 and 6bps of those are found in H12_9 and H6_F_70, respectively.

* nomenclature according to Bennett et al. (2005)

B. SVA_F1 elements with additional 5' transductions (group 2): 12

Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	5' Transduction	Source locus 5'TD	Source element	Polymorphic state
H19_56	Chr.19:13673729-13676405	aaaagcctcatca	364	1178	385 (MER115/AluJo)	Chr.9:130695439-130695822		
					108 (AluSc)	Chr.9:33445301-33445422	H10_1 ²	
H6_17	Chr.6:134026663-134029051	aagtagggcgaa	364	1305	128 (AluSc)	Chr.9:33445291-33445422	H10_1 ²	
H1_F_67	Chr.1:33185642-33188195	aagaatag	364	1578	27	Chr.17:74054840-74054866		fixed
H9_2	Chr.9:112354395-112357759	aagactqaatgct	334 (313) ³	1585	879 (L2/AluSg/x/AluSx/L2)	Chr.17:7210133-7211008		
H22_17	Chr.22:39256526-39259105	aaaaaattaaacaagaagca	150	1555	274 (AluJo)	Chr.12:48035454-48035724		fixed
H8_2	Chr.8:88807-90852	gaaagtatcagatata	148	1176	107	Chr.5:138470814-138470923		
H12_3	Chr.12:47011734-47014662	aataaacttaattac	148	1495	693 (contains AluJb in antisense)	Chr.1:39056881-39057573		fixed
H12_2	Chr.12:3074262-3075599	aaagattattgattt	147	558	21	Chr.2:201388069-201388089		fixed
H6_2	Chr.6:15296388-15298817	aagagqaaattcaa	91	1535	211 (Mer21A)	Chr.1:24066845-24067055		H1_2
H8_F_116	Chr.8:129906382-129908969	aaaaatagcttatgttca	89	1578	213 (AluY/L1M4)	Chr.19:11583835-11584045		fixed
					99 (AluSp; 3' part)	Chr.21:33014335-33014433		
H1_F_163 (LRIG2*)	Chr.1:113323365-113325836	aaaaacacatttgcg	76	1771	26	Chr.19:20122585-20122610		H19_47 polymorphic
H3_5	Chr.3:17720382-17722747	acaatatcaatgtggt	68	1419	204 (spliced RNA-NPEPPS)	chr17:42,963.766-42,965.384		
					56 (AluY; 3' part)	Chr.12:124056563-124056623		

² could have been derived from either a MAST-2 SVA previously existing on 9p13.3 or from H10_1

³ internal deletion in MAST-2; actual length in parentheses

C. SVA_F1 elements with additional 3' transductions (group 3): 4

Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	3' Transduction	Source locus 3'TD	Source element	Polymorphic state
H9_1	Chr.9:134009072-134013871	aaataaaaatta	148	3693	289	Chr.15:71059852-71060140		
H20_17	Chr.20:32279590-32282896	gaaagtctta	92	2184	289	Chr.15:71059852-71060140		H9_1
					17	Chr.9:134009067-134009139		
HX_4	Chr.X:100032701-100034660	aaatgcatagaagg	147	1077	128	Chr.4:184345663-184345790		fixed
H3_8	Chr.3:67576386-67579502	gaaattagatctc	132	1659	128	Chr.4:184345663-184345790		HX_4 fixed
					571	Chr.X:100032143-100032747		

D. SVA_F1 elements with additional 5' and 3' transductions (group 4)

H1_F160 → H6_6 transduction group: 1 element

Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	5' Transduction	Source locus 5'TD	Source element
H6_6	Chr.6:112935868-112937012	aaaagaagtctt	128 (58) ⁴	135	54	Chr.1:113154100-113154153	H1_F160
					3' Transduction	Source locus 3'TD	Source element
					241	Chr.1:113151982-113152308	H1_F160

⁴ range in MAST-2 is 333-390; internal deletion of 391-460 (base count from NM_015112) and of the SVA Alu-like sequence

H10_1 transduction group: 14 (13) elements

Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	5' Transduction	Source locus 5'TD	Polymorphic state
H10_1	Chr.10:101587188-101590760	aagaatggactt	364	2086	519 (185-320 AluSc-14)	Chr.9:33444904-33445422	fixed
Locus name	Position	TSD	MAST-2 (nts)	VNTR (nts)	5' Transduction	3' Source locus 3'TD	Source element
H14_17	Chr.14:104647703-104650640	aataatattgata	364	1346	195 (181-AluSc-14)	Chr.9:33445228-33445422	H10_1
H3_17	Chr.3:48226881-48230038	agaaagtcttaggcac	364	1643	400 (19-299 AluSp-82)	Chr.10:101586803-101587235	H10_1
H19_108	Chr.19:35080630-35084954	aaaaataaaaatta	364	2791	167 (153-AluSc-14)	Chr.9:33445256-33445422	H10_1
H19_39	Chr.19:20376996-20379861	gaaactgaatatic	362	1483	400 (17-301 AluSp-82)	Chr.10:101586803-101587235	H10_1
H3_7	Chr.3:107311542-107314321	aacaaactactg	148	1610	148 (134-AluSc-14)	Chr.9:33445289-33445422	H10_1
H17_17	Chr.17:20827790-20830555	aagaatggctattg	148	1534	397 (17-300 AluSp-80)	Chr.10:101586803-101587235	H10_1
H19_47	Chr.19:20120146-20122599	aaattatgggagtc	76	1368	401 (17-302 AluSp-82)	Chr.10:101586803-101587235	H10_1
H10_7	Chr.10:94124326-94127675	no	starts VNTR	2098	479 (17-302 AluSp-160)	Chr.10:101586725-101587235	H10_1
H15_2	Chr.15:88197744-88200214	aaagaatcaatga	starts VNTR	1510	405 (17-302 AluSp-86)	Chr.10:101586803-101587235	H10_1
H4_2	Chr.4:182240060-182241942	aaaaataaatatttttc	starts VNTR	874	398 (17-299 AluSp-82)	Chr.10:101586803-101587235	H10_1
H15_3	Chr.15:69848304-69849771	aaaaaatcttgctgg	starts VNTR	495	402 (17-303 AluSp-82)	Chr.10:101586803-101587235	H10_1
H6_5	Chr.6:39038990-39040220	gaaaaaggactcccagtg	starts VNTR	241	396 (17-299 AluSp-80)	Chr.10:101586803-101587235	H10_1
H4_22	Chr.4:186615370-186615847	taaaggggaaaagtg	3'TD only		402 (17-303 AluSp-82)	Chr.10:101586803-101587235	H10_1
					402 (18-305 AluSp-79)	Chr.10:101586803-101587235	

Highlighted SVAs were tested for their polymorphic state by PCR;