

Human alpha satellite DNA – consensus sequence and conserved regions

B.Vissel and K.H.Choo

Murdoch Institute, Royal Children's Hospital, Flemington Road, Parkville 3052, Australia
Submitted June 29, 1987

Since the previous human alpha satellite consensus sequence (1) was derived by direct sequencing of genomic DNA containing the EcoRI-dimer structure, the other major alpha satellite species were naturally excluded. We report a modified human alpha satellite DNA consensus sequence (Table 1) derived from published data for 145 individually cloned and sequenced, complete alphoid monomers: 4 monomers from chromosomes 13 and 21 (2), 2 from acrocentric chromosomes (3), 3 from chromosome 22 (4), 5 from the Y chromosome (5), 16 from chromosome 17 (6), 12 from the X chromosome (7), 32 from chromosome 7 (8) and 71 from the 340 bp band in EcoRI-digested genomic DNA (8). Since these monomeric units were derived from a range of chromosomal sources, the new consensus sequence (Table 1) should be an improved representative of the human alpha satellite DNA (Table 2). Direct comparison of individual sequence also allows visualisation of nucleotides or regions of nucleotides which are well conserved (Table 1). We propose that the present consensus sequence should serve as a base for future updating as more sequence data becomes available.

Table 1 : Derivation of new consensus sequence

	10										20										30									
I	A	A	T	T	C	T	C	A	G	T	A	A	C	T	T	C	C	T	T	G	T	G	T	T	G	T	G	T	G	T
II	52	141	135	139	140	131	131	143	134	42	137	133	92	136	98	134	21	134	135	138	139	141	42	142	140	124	106	135	134	36
III	93	4	10	6	5	14	14	2	11	103	8	12	53	9	47	11	124	11	10	7	6	4	103	3	5	21	39	10	11	109
IV	76C	*	*	*	*	*	*	*	*	101A	*	*	39A	*	*	119T	*	*	*	*	*	102A	*	*	*	*	*	*	*	104C
V	(C)	A	T	T	C	T	C	A	G	(A)	A	A	C	T	T	C	(T)	T	T	G	T	G	(A)	T	G	T	G	T	G	(C)
	40										50										60									
I	A	T	T	C	A	A	C	T	C	A	C	A	G	A	G	T	T	G	A	A	C	G	A	T	C	C	T	T	T	A
II	86	142	144	136	140	130	128	144	142	132	137	133	125	143	144	136	116	87	138	142	131	36	12	119	73	139	132	133	141	46
III	59	3	1	9	5	15	17	1	3	13	8	12	20	2	1	9	29	58	7	3	14	109	133	26	72	6	13	12	4	99
IV	47C	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	51T	*	*	*	73C	127T	63T	*	*	*	*	*	85T	
V	A	T	T	C	A	A	C	T	C	A	C	A	G	A	G	T	T	G	A	A	C	(C)	T	T	C	C	T	T	T	T
	70										80										90									
I	C	A	C	A	G	A	G	C	A	G	A	C	T	T	G	A	A	A	C	A	C	T	C	T	T	T	T	T	G	T
II	71	137	46	126	142	134	135	142	139	141	38	12	123	66	142	141	142	143	133	133	132	133	139	136	90	142	134	140	139	111
III	74	8	99	19	3	11	10	3	6	4	107	133	22	79	3	4	3	2	12	12	13	12	6	9	55	3	11	5	6	34
IV	71G	*	84T	*	*	*	*	*	*	*	91T	130T	59G	*	*	*	*	*	*	*	*	*	*	50G	*	*	*	*	*	*
V	C	A	(T)	A	G	A	G	C	A	G	(T)	(T)	T	T	G	A	A	C	A	C	T	C	T	T	T	T	T	G	T	
	100										110										120									
I	G	G	A	A	T	T	T	G	C	A	A	G	T	G	G	A	G	A	T	T	T	C	A	G	C	C	G	C	T	T
II	49	90	137	91	141	50	141	136	137	143	141	135	135	139	132	138	45	141	135	118	123	43	26	15	83	120	71	113	119	135
III	96	55	9	54	4	95	4	9	8	2	4	10	10	6	13	7	100	4	10	27	22	102	119	130	62	25	74	32	26	10
IV	91A	50A	*	49G	*	91C	*	*	*	*	*	*	*	*	*	*	90T	*	*	*	72G	98G	129A	55G	45T	*	*	*	*	
V	(A)	G	A	A	T	(C)	T	G	C	A	A	G	T	G	G	A	(T)	A	T	T	T	(G)	(G)	(A)	C	C	G	C	T	T

