Int J Legal Med (1997) 110 : 69–72

ORIGINAL ARTICLE

B. Rolf · M. Schürenkamp · A. Junge · B. Brinkmann

Sequence polymorphism at the tetranucleotide repeat of the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus

Received: 19 November 1996 / Received in revised form: 23 January 1997

Abstract The tetranucleotide repeat polymorphism in the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus has become a widely used short tandem repeat (STR) system in paternity testing and human identification. The discrimination index of this locus has been reported to be as high as 99.65%. To study the overall variability of this locus, 222 alleles were sequenced in our laboratory. Here, we report the sequences of the 102 different alleles observed. In addition to the length polymorphism, up to ten different sequence variants for single fragment lengths have been detected.

Key words Short tandem repeat · ACTBP2 · Polymorphism · Sequence variation

Introduction

Since the first description in 1991 by Warne et al., the tetranucleotide polymorphism in the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus (Moss and Gallwitz 1983) has become a widely used short tandem repeat (STR) system in paternity testing and human identification. Due to the high degree of polymorphism, the discrimination index of this locus has been reported to be as high as 99.65% (Möller et al. 1995). Consequently, for many laboratories ACTBP2 is the system of first choice when large numbers of samples have to be compared or individualised. In addition to the length polymorphism based on the number of repeats, sequence variants have also been observed (Urquhart et al. 1993). In this study, we investigated the sequence of 222 alleles to determine the overall variability at this locus.

B. Rolf · M. Schürenkamp · B. Brinkmann (✉)
Institute of Legal Medicine, von-Esmarch-Strasse 86,
D-48149 Münster, Germany
FAX: +49 (251) 8355 158

A. Junge
Institute of Legal Medicine, Stiftsplatz 12,
D-53111 Bonn, Germany

Materials and methods

Genomic DNA of Caucasians from North-West Germany was extracted from blood according to Brinkmann et al. (1991) and quantified using the slot-blot technique and the probe D17Z1 (Waye et al. 1989). PCR amplification was performed as described by Wiegand et al. (1993). Alleles were separated on native gels and isolated from the gels as described elsewhere (Möller and Brinkmann 1994). After reamplification under the same conditions, sequencing was carried out on a ABI 373 sequencer using the Taq Dye-Deoxy-Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, Calif.). Every sequence was determined twice in 5′-3′ direction. For additional confirmation, original DNA was then reamplified under the same conditions except that a 5′Cy5-labelled upstream primer was used. Separation of these samples was carried out in a 6% acrylamide gel in TBE-buffer containing 7 M urea on a Pharmacia ALF-system according to the manufacturer's instructions. Using a Cy5-labelled allelic ladder the fragment length could be determined and compared to the calculated allele size.

Results and discussion

A total of 222 alleles have been sequenced in our laboratory. The sequences of the 102 different alleles observed are summarised in Table 1. Most of the fragments were chosen arbitrarily. As we expected to find more sequence variation in the range above allele no. 19, most of the fragments studied were alleles > 19 and selected because they did not appear in the ladder. Consequently, we did not calculate allele frequencies for the variants found.

The highly variable tandem repeat region of this locus consists of 7–22 units of the repetitive unit AAAG up to a fragment length of 273 bp (Table 1, lines 1–17). The longer fragments (number of repeat units 19–32) contain an additional hexamer unit, in most cases AAAAAG, which might have been the result of a insertion of an AA. The position of this hexamer is variable. Among the 12 alleles with a length of 303 bp, 8 variants with 28 AAAG and 1 AAAAAG unit ranging between $(AAAG)_8AAA-AAG(AAAG)_{20}$ and $(AAAG)_{16}AAAAAG(AAAG)_{12}$ were observed (Table 1, lines 71–82). In addition to these variants, three alleles were observed with deletions or transitions in the 5′ and 3′ region of this central region. Two alleles of

Table 1 Sequence of the polymorphic region of the ACTBP2 alleles

No.	5′-flank. region				Central region							3′-flank. region				No. found	Allel name	Length bp
	AAAG	AG	AAAG	AG	AAAG	AAAAAG	AG	AGAAAG	AAAG	AAAAAG	AAAG	G	AAGG	AAAG/ANAG	AG			
1	2	1	3	1	7	0	0	0	0	0	0	0	0	3	1	1	7	212
2	2	1	3	1	12	0	0	0	0	0	0	1	0	3	1	1	12	233
3	2	1	3	0	13	0	0	0	0	0	0	1	0	3	1	1	12.2	235
4	2	1	3	0	14	0	0	0	0	0	0	1	0	3	1	1	13	239
5	2	1	3	1	14	0	0	0	0	0	0	1	0	3	1	1	14	241
6	2	1	3	1	15	0	0	0	0	0	0	1	0	3	1	1	15	245
7	2	1	3	1	16	0	0	0	0	0	0	1	0	3	1	1	16	249
8	2	1	3	1	17	0	0	0	0	0	0	1	0	3	1	1	17	253
9	2	1	3	1	18	0	0	0	0	0	0	1	0	3	1	1	18	257
10	2	1	3	1	19	0	0	0	0	0	0	1	0	3	1	1	19	261
11	2	1	3	1	10	1	0	0	8	0	0	1	1	2	1	1	19.2	263
12	2	1	3	1	20	0	0	0	0	0	0	1	0	3	1	7	20	265
13	2	1	3	1	11	1	0	0	8	0	0	1	1	2	1	2	20.2	267
14	2	1	3	1	21	0	0	0	0	0	0	1	0	3	1	9	21	269
15	2	1	3	1	9	1	0	0	11	0	0	1	1	2	1	1		
16	2	1	3	1	11	1	0	0	9	0	0	1	1	2	1	2	21.2	271
17	2	1	3	1	22	0	0	0	0	0	0	1	0	3	1	4	22	273
18	2	1	3	1	7	1	0	0	14	0	0	1	1	2	1	1		
19	2	1	3	1	8	0	5	0	12	0	0	1	1	2	1	1		
20	2	1	3	1	9	1	0	0	12	0	0	1	1	2	1	1		
21	2	1	3	1	10	1	0	0	11	0	0	1	1	2	1	3		
22	2	1	3	1	11	1	0	0	10	0	0	1	1	2	1	1		
23	2	1	3	1	12	1	0	0	9	0	0	1	1	2	1	1	22.2	275
24	2	1	3	1	7	1	0	0	15	0	0	1	1	2	1	2		
25	2	1	3	1	8	1	0	0	14	0	0	1	1	2	1	1		
26	2	1	3	1	8	1	0	0	14	0	0	1	1	2	1	2		
27	2	1	3	1	9	1	0	0	13	0	0	1	1	2	1	1		
28	2	1	3	1	10	0	3	0	12	0	0	1	1	2	1	1		
29	2	1	3	1	10	1	0	0	12	0	0	1	1	2	1	1		
30	2	1	3	1	11	1	0	0	11	0	0	1	1	2	1	5		
31	2	1	3	1	12	1	0	0	10	0	0	1	1	2	1	1	23.2	279
32	2	1	3	1	5	1	0	0	18	0	0	1	1	2	1	1		
33	2	1	3	1	7	1	0	0	16	0	0	1	1	2	1	2		
34	2	1	3	1	8	1	0	0	15	0	0	1	1	2	1	1		
35	2	1	3	1	10	1	0	0	13	0	0	1	1	2	1	2		
36	2	1	3	1	11	1	0	0	12	0	0	1	1	2	1	1		
37	2	1	3	1	12	1	0	0	11	0	0	1	1	2	1	1	24.2	283
38	2	1	3	1	9	1	0	0	15	0	0	1	1	2	1	10		
39	2	1	3	1	10	0	0	1	14	0	0	1	1	2	1	1		
40	2	1	3	1	10	1	0	0	14	0	0	1	1	2	1	2		
41	2	1	3	1	11	1	0	0	13	0	0	1	1	2	1	1		
42	2	1	3	1	12	1	0	0	12	0	0	1	1	2	1	1		
43	2	1	3	1	14	1	0	0	10	0	0	1	1	2	1	1	25.2	287
44	2	1	3	1	8	1	0	0	17	0	0	1	1	2	1	2		
45	2	1	3	1	9	1	0	0	16	0	0	1	1	2	1	4		
46	2	1	3	1	10	1	0	0	15	0	0	1	1	2	1	1		
47	2	1	3	1	11	0	0	1	14	0	0	1	1	2	1	2		
48	2	1	3	1	11	1	0	0	14	0	0	1	1	2	1	4		
49	2	1	3	1	14	1	0	0	11	0	0	1	1	2	1	1	26.2	291
50	2	1	3	1	8	0	0	1	18	0	0	1	1	2	1	1		
51	2	1	3	1	8	1	0	0	18	0	0	1	1	2	1	3		
52	2	1	3	1	9	0	0	1	17	0	0	1	3	0	1	1		

Table 1 (continued)

No.	5′-flank. region				Central region							3′-flank. region				No. found	Allel name	Length bp
	AAAG	AG	AAAG	AG	AAAG	AAAAAG	AG	AGAAAG	AAAG	AAAAAG	AAAG	G	AAGG	AAAG/ANAG	AG			
53	2	1	3	1	10	1	0	0	16	0	0	1	1	2	1	3		
54	2	1	3	1	11	1	0	0	15	0	0	1	1	2	1	5		
55	2	1	3	1	12	1	0	0	14	0	0	1	1	2	1	6		
56	2	1	2	1	12	1	0	0	15	0	0	1	1	2	1	1		
57	2	1	3	1	13	0	0	1	13	0	0	1	3	0	1	1		
58	2	1	3	1	13	1	0	0	13	0	0	1	1	2	1	5		
59	2	1	3	1	15	1	0	0	11	0	0	1	1	2	1	1	27.2	295
60	2	1	3	1	8	1	0	0	19	0	0	1	1	2	1	3		
61	2	1	3	1	9	0	0	1	18	0	0	1	1	2	1	1		
62	2	1	3	1	9	0	9	0	15	0	0	1	1	2	1	1		
63	2	1	3	1	9	1	0	0	18	0	0	1	1	2	1	2		
64	2	1	3	1	10	1	0	0	17	0	0	1	1	2	1	1		
65	2	1	3	1	11	1	0	0	16	0	0	1	1	2	1	8		
66	2	1	3	1	12	1	0	0	15	0	0	1	1	2	1	3		
67	2	1	3	1	13	1	0	0	14	0	0	1	1	2	1	5		
68	2	1	3	1	14	1	0	0	13	0	0	1	1	2	1	3		
69	2	1	3	1	14	1	0	0	13	0	0	1	3	0	1	1		
70	2	1	3	1	16	1	0	0	11	0	0	1	1	2	1	1	28.2	299
71	2	1	3	1	8	1	0	0	20	0	0	1	1	2	1	2		
72	2	1	3	1	9	0	0	1	19	0	0	1	1	2	1	2		
73	2	1	3	1	9	1	0	0	19	0	0	1	1	2	1	1		
74	1	1	3	1	10	1	0	0	19	0	0	1	1	2	1	1		
75	2	1	3	1	11	0	5	0	16	0	0	1	1	2	1	1		
76	1	1	3	1	11	1	0	0	18	0	0	1	1	2	1	1		
77	2	1	3	1	11	1	0	0	17	0	0	1	1	2	1	3		
78	2	1	3	1	12	1	0	0	16	0	0	1	1	2	1	6		
79	2	1	3	1	13	0	0	1	15	0	0	1	3	0	1	1		
80	2	1	3	1	13	1	0	0	15	0	0	1	1	2	1	10		
81	2	1	3	1	14	1	0	0	14	0	0	1	1	2	1	5		
82	2	1	3	1	16	1	0	0	12	0	0	1	1	2	1	1	29.2	303
83	2	1	3	1	11	1	0	0	18	0	0	1	1	2	1	1		
84	2	1	3	1	12	1	0	0	17	0	0	1	1	2	1	4		
85	1	1	3	1	12	1	0	0	18	0	0	1	1	2	1	1		
86	2	1	3	1	13	1	0	0	16	0	0	1	1	2	1	4		
87	2	1	3	1	14	1	0	0	15	0	0	1	1	2	1	3		
88	2	1	3	1	15	1	0	0	14	0	0	1	1	2	1	1	30.2	307
89	1	1	3	1	9	1	0	0	22	0	0	1	1	2	1	1		
90	1	1	3	1	10	1	0	0	21	0	0	1	1	2	1	1		
91	2	1	3	1	12	1	0	0	18	0	0	1	1	2	1	2		
92	2	1	3	1	13	1	0	0	17	0	0	1	1	2	1	2		
93	2	1	3	1	14	1	0	0	16	0	0	1	1	2	1	3	32	311
94	2	1	3	1	13	1	0	0	18	0	0	1	1	2	1	1		
95	2	1	3	1	14	1	0	0	17	0	0	1	1	2	1	2	32.2	315
96	2	1	2	1	10	1	0	0	12	1	9	1	1	2	1	2		
97	2	1	3	1	10	1	0	0	11	1	9	1	1	2	1	1	33	317
98	1	1	3	1	10	1	0	0	23	0	0	1	1	2	1	2	33.2	319
99	2	1	3	1	9	1	0	0	13	1	9	1	1	2	1	1	34	321
100	1	1	3	1	13	1	0	0	22	0	0	1	1	2	1	1	35.2	327
101	2	1	3	1	10	1	0	0	14	1	9	1	1	2	1	1	36	329
102	2	1	3	1	9	1	0	0	16	1	9	1	1	2	1	1	37	333

this length have transitions in the central region resulting in AGAAAG instead of the AAAAAG or in a $(AG)_5$. As variations outside the central region are less frequent (only 10 of the 222 allele sequenced exhibited a deletion), a nomenclature has been established, in which only the number of repeats in the highly variable region are considered. This nomenclature, based on the ISFH recommendations (1994), was proposed by a satellite group of the GEDNAP (German DNA Profiling Group, H.R. Schneider, personal communication) and results in the allele designation 7–22 for the alleles with the simple repeat region. Fragments showing the dinucleotide insertion, e.g. an incomplete repeat motif (Table 1) are named 19.2 to 33.2. Interestingly, some of the longer fragments (over 317 bp) contain an additional hexanucleotide unit/dinucleotide insertion, resulting in a $(AAAG)_iAAAAAG(AAAG)_jAAAAAG(AAAG)_k$ element (Table 1, 96–102). These fragments with the second dinucleotide insertion are again named regularly 33–37, as two incomplete repeats can be considered as one complete repeat.

The enormous sequence variation at this locus and the known problems with sequence variants under native electrophoretic conditions have led to the detection of a variety of alleles migrating between defined ladder alleles in samples typed on native gels as described by Möller et al. (1995). To achieve a higher reproducibility, we presently separate this locus under denaturing conditions only. All sequenced alleles were retyped under denaturing conditions on a Pharmacia ALF system and compared with a Cy5-labelled allelic ladder. In all cases the fragment length determined differed by less than 0.5 bp from the calculated allele size. Thus the determination of the fragment lengths under denaturing conditions is the only possibility to achieve inter- and intralaboratory reproducibility regarding allele types and frequencies. However, the nomenclature based on the repeat number should still be used and not be replaced by the fragment lengths. Although this nomenclature is somewhat arbitrary considering the complex structure of this system, it is independent of the primer sequence and relates to alleles that are defined by a variable number of tandem repeats as recommended (ISFH recommendations, 1994). A nomenclature based on fragment length only would have to be changed with every new primer applied, e.g. if a new multiplex optimisation would require a different set of primers.

References

Brinkmann B, Rand S, Wiegand P (1991) Population and familiy data of RFLP's using selected single- and multi-locus systems. Int J Legal Med 104:81–86

DNA recommendations - 1994 report concerning further recommendations of the DNA Commision of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Int J Legal Med 107:159–160

Möller A, Brinkmann B (1994) Locus ACTBP2 (SE33): sequencing data reveal considerable polymorphism. Int J Legal Med 106:262–267

Möller A, Schürenkamp M, Brinkmann B (1995) Evaluation of an ACTBP2 ladder composed of 26 sequenced alleles. Int J Legal Med 108:75–78

Moss M, Gallwitz D (1983) Structure of two human β-actin-related processed genes one of which is located next to a simple repetitive sequence. EMBO J 2:757–761

Urquhart A, Kimpton CP, Gill P (1993) Sequence variability of the tetranucleotide repeat of the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus. Hum Genet 92:637–638

Warne D, Watkins C, Bodfish P, Nyberg K, Spurr NK (1991) Tetranucleotide repeat polymorphism at the human beta-actin related pseudogene 2 (ACTBP2) detected using the polymerase chain reaction. Nucleic Acids Res 19: 6980

Waye JS, Presley LA, Budowle B, Shutler GS, Fourney RM (1989) A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. Biotechniques 7:852–855

Wiegand P, Budowle B, Rand S, Brinkmann B (1993) Forensic validation of the STR systems SE33 and TC11. Int J Legal Med 105:315–320