Int J Leg Med (1994) 106:262–267

International Journal of
Legal Medicine

Locus ACTBP2 (SE33)

Sequencing data reveal considerable polymorphism

A. Möller, B. Brinkmann

Institute of Legal Medicine, Westfälische Wilhelms-Universität, Von-Esmarch-Strasse 86, D-48149 Münster, Germany

Received December 20, 1993

Summary. A total of 50 different ACTBP2 (human beta-actin related pseudogene) alleles and the cell line K562 were sequenced and analysed. Sequence data revealed not only length polymorphism but also a highly variable sequence polymorphism. 2 different sequence structures were characterised. Type I (223–259 bp) contained the known regular 4 bp repeat AAAG. Type II (265–309 bp) revealed a further hexanucleotide unit AAAAAG in addition to the common AAAG which only occurs once in the repeat region. The position of this insertion showed considerable variation. To obtain a regular spaced allelic ladder 20 of the sequenced alleles were selected. Denaturing gels and high resolution/non-denaturing gels were compared and striking differences could be seen between the 2 gel systems. Separation of the alleles on a 6% denaturing gel and analysis on the ABI 373A Sequencer revealed fragment sizes which corresponded to the sequencing data but were in general 6–10 bp longer. In contrast, in non-denaturing gels some alleles showed different electrophoretic mobilities compared to the sequenced allelic ladder which could indicate different fragment length and/or different sequence structure.

Key words: ACTBP2 polymorphism – Sequencing – Sequence variation – Allelic ladder

Zusammenfassung. 50 verschiedene ACTBP2 Allele und die Zell-Linie K562 wurden sequenziert und analysiert. Die Sequenzierergebnisse zeigen nicht nur einen Längen- sondern auch einen hoch variablen Sequenzpolymorphismus. 2 verschiedene Sequenzstrukturen wurden charakterisiert. Typ I (223–259 bp) enthält den regulären 4 bp-Repeat AAAG. Im Typ II (265–309 bp) tritt zusätzlich die Hexanukleotideinheit AAAAAG auf, deren Position innerhalb der Repeatregion stark variabel ist. 20 sequenzierte Allele wurden für die Konstruktion einer allelischen Leiter ausgewählt. Denaturierende und hochauflösende, nicht denaturierende Gele wurden ausgewertet und Unterschiede im Wanderungsverhalten der Allele festgestellt. In denaturierenden Gelen ist die elektrische Mobilität abhängig von den Fragmentlängen der Allele. Dagegen zeigen in nicht denaturierenden Gelen einige Allele unterschiedliches Wanderungsverhalten im Vergleich zur sequenzierten allelischen Leiter, welches durch unterschiedliche Fragmentlängen und/oder unterschiedliche Sequenzstrukturen erklärt werden kann.

Schlüsselwörter: ACTBP2 Polymorphismus – Sequenzierung – Sequenzvariation – Allelische Leiter

Introduction

STR (short tandem repeat) systems are highly polymorphic and sensitive markers for human identification (Edwards et al. 1992; Brinkmann 1992; Kimpton et al. 1992; Polymeropoulos et al. 1992a; Wiegand et al. 1993). ACTBP2 (human beta-actin related pseudogene; Moos and Gallwitz 1983; Polymeropoulos et al. 1992b) is one of the most informative systems with a heterozygosity index of at least 93% (Wiegand et al. 1993). Original data (Polymeropoulos et al. 1992b) suggested that the basic sequence structure was $(AAAG)_{11}$-(AAAAAG)-$(AAAG)_{14}$ and alleles differed by a regular 4 bp repeat. Recent validation studies have shown a more complex sequence structure with additional inter alleles differing by 1–3 bp (Urquhart et al. 1993).

The aim of this study was to determine the variation of sequence structure and the possibilities of interlaboratory comparison.

Materials and methods

DNA was extracted from blood (Brinkmann et al. 1991) and quantified using the slot blot technique and the human specific probe D17Z1 (Gibco BRL).

The PCR protocol and electrophoresis for allele isolation was carried out according to Wiegand et al. (1993).

Isolation of DNA fragments from heterozygous individuals after silver staining: Gel pieces containing DNA bands were cut from the gel and transferred to microfuge tubes. Elution of DNA was carried out using the "crush and soak" method (Maniatis et al.

Correspondence to: B. Brinkmann

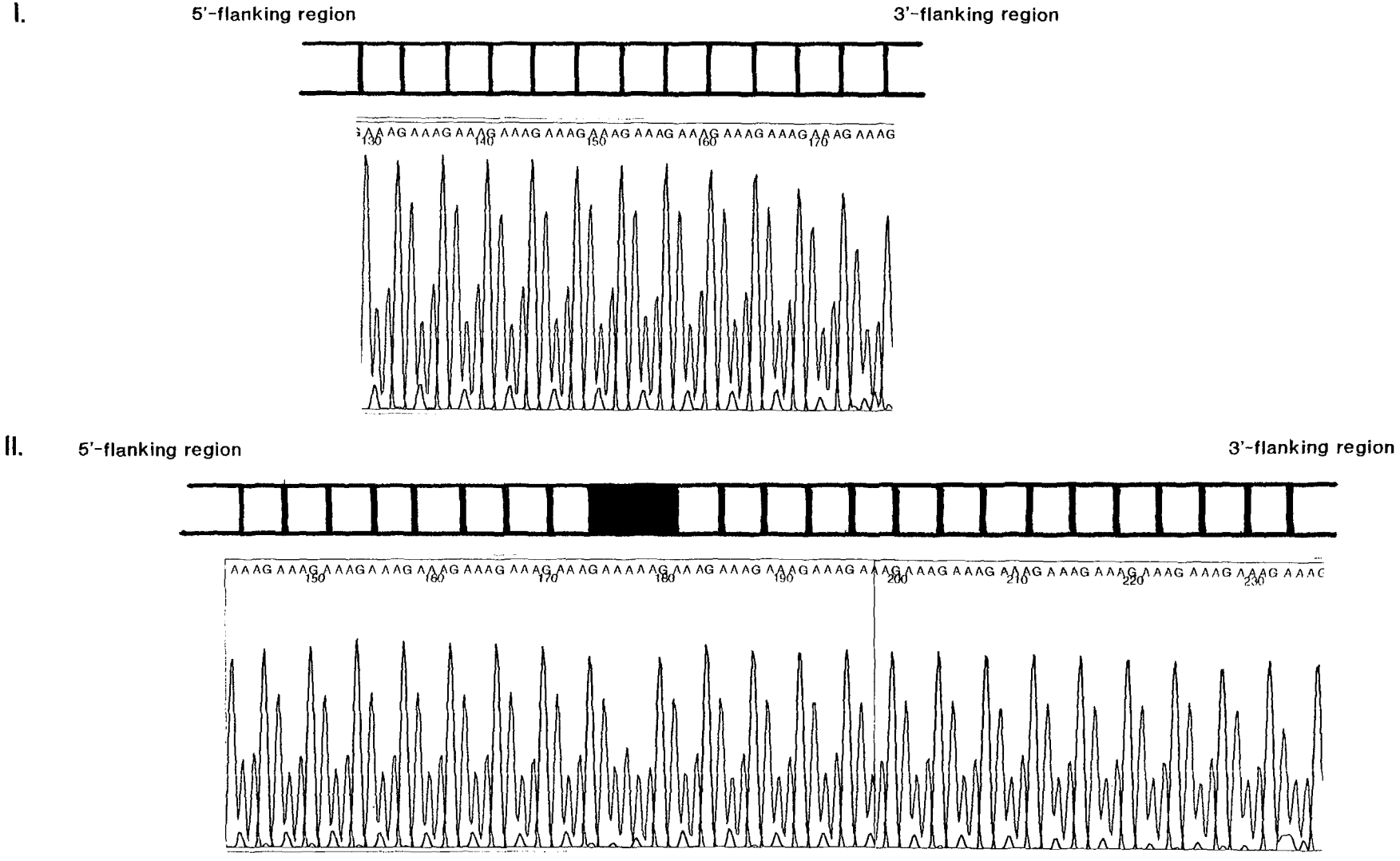

Fig. 1. Electropherograms of the 2 different sequence structures. Samples were separated on a 6% denaturing PAA gel and analysed automatically on the ABI 373A Sequencer. High peaks represent the base G, the lower peaks represent the base A. I: repeat region of type I with the regular AAAG repeat (□). II: repeat region of type II with the hexanucleotide unit AAAAAG (■)

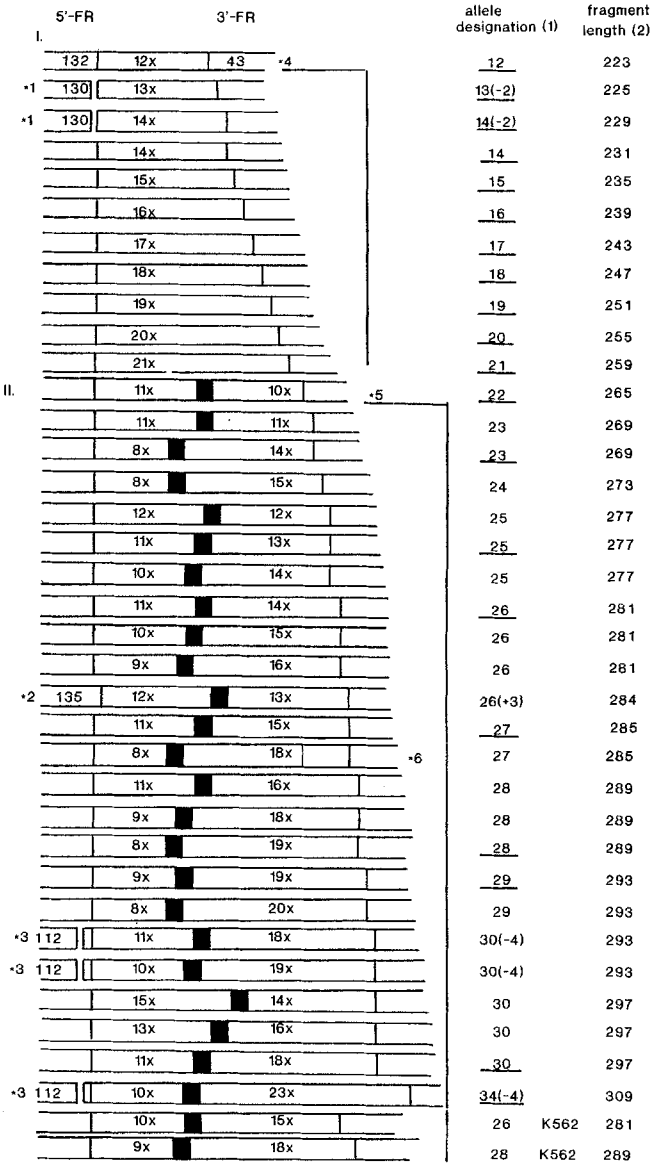

Fig. 2. Sequence structure and fragment lengths of different ACTBP2 alleles and the cell line K562. Underlined alleles are used to construct the allelic ladder. (1) numbers reflect the total of repeats including the unit AAAAAG; deletions (–) and insertions (+) in brackets, together with the number of bp. (2) fragment lengths in base pairs (bp), as determined by Taq-Cycle-Sequencing, *1 AG deletion at position 131–132, *2 AAG insertion at position 93–95, *3 AAAG deletion at position 113–116, *4 GAAAG at position 1–5, *5 GAAGG at position 1–5, *6 transversion from A to T in repeat number 22, FR flanking region

1989) but modified using Centricon 100 (Amicon, Beverly, USA) purification instead of ethanol precipitation.

Taq-Cycle-Sequencing: Eluted DNA fragments were reamplified with 1–5 µl Centricon 100 eluate as described above using a 100 µl volume instead of 25 µl. PCR products were desalted and concentrated using Centricon 100 (Amicon, Beverly, USA). Sequencing reactions were carried out using the Taq Dye-Deoxy-Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA), one third of the centricon eluate (100–400 ng template DNA) and 0.5 µM sequencing primer (ACTBP2-primer 1: 5´AAT CTG GGC GAC AAG AGT GA 3´).

Table 1. Allele composition of the ACTBP2-cocktail. Alleles are characterised by their repeat number (including the AAAAAG unit) and their fragment sizes as measured by the ABI Sequencer 373A in bp in combination with the structure as shown in Fig. 2.

Allele designation	Fragment length (bp) (Genescan)
12	230
13(–2)	232
14(–2)	236
14	238
15	242
16	246
17	250
18	254
19	258
20	262
21	266
22	271
23	275
25	284
26	288
27	293
28	297
29	302
30	306
34(–4)	318
K562	
26	289
28	297

Amplification conditions: 96°C – 15 s, 60°C – 15 s, 60°C – 4 min: 25 cycles using a Perkin Elmer 9600

DNA was purified by phenol/chloroform extraction and ethanol/NaAc (pH 5.8) precipitation. The DNA pellet was dried and 4 µl loading buffer was added (formamide/50 mM EDTA (5:1)). Heat denaturation was performed at 90°C for 2 min.

Sequencing gel and analysis of sequence data: The samples were separated in a 6% PAA gel with a separation distance of 24 cm. The gel solution consisted of 8.3 M urea, 6% acrylamide/bis-solution and 1 × tris/borate buffer. The gel solution was filtered and degased before the addition of 180 µl 10% ammonium peroxodisulfate solution and 24 µl TEMED (N´N´N-´N´-Tetramethylethylendiamine). Electrophoresis was carried out on the ABI 373A Sequencer at 2500 V, 40 mA and 30 W for 10–12 hours depending on the length of the fragments. Sequence data was analysed automatically on the ABI 373A Sequencer using the Data collection and SeqEd software (Applied Biosystems, Foster City, CA).

Construction of the ACTBP2 ladder: The allelic ladder was obtained after reamplification and combination of 20 sequenced alleles. Separation of the alleles was carried out as described by Wiegand et al. (1993) as well as on a 6% denaturing gel (8.3 M urea, 6% acrylamide/bis-solution and 1 × tris/borate buffer, separation distance 12 cm, 800 V, 45 mA, 30 W) using the fluorescence detection system (Applied Biosystems, Foster City, CA). The ACTBP2 primer 1 (5´ AAT CTG GGC GAC AAG AGT GA 3´) was labelled with FAM (5-carboxy-fluorescein). The fragment sizes were automatically analysed using the Genescan software 672 and the internal standard Genescan 2500 labelled with ROX (6-carboxyrhodamin X) (Applied Biosystems, Foster City, CA).

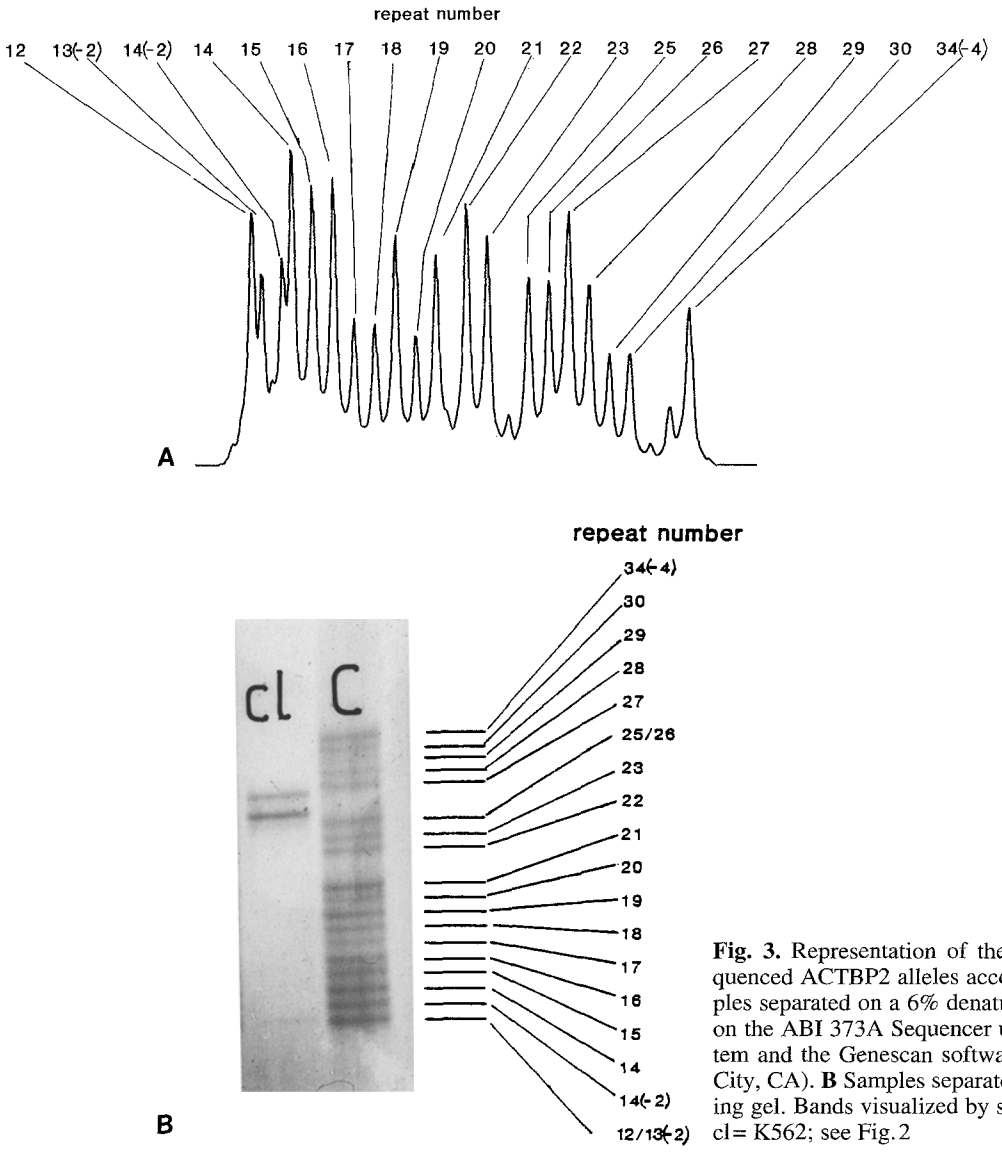

Fig. 3. Representation of the allelic ladder composed of 20 sequenced ACTBP2 alleles according to Fig. 2 and Table 1. **A** Samples separated on a 6% denaturing gel and automatically analysed on the ABI 373A Sequencer using the fluorescence detection system and the Genescan software 672 (Applied Biosystems, Foster City, CA). **B** Samples separated on a high resolution, non-denaturing gel. Bands visualized by silver staining (Budowle et al. 1991). cl = K562; see Fig. 2

Results

Sequence structure and variation of ACTBP2 alleles

A total of 50 different ACTBP2 fragments and the cell line K562 were sequenced and analysed. The fragment lengths ranged from 223 bp to 309 bp. Sequence data of K562 revealed 281 bp and 289 bp alleles (Fig. 2). Two different types of structure variation can be described:

I. The lower allele range up to 259 bp with 11 different alleles which showed a common repeat structure (Fig. 1.I, Fig. 2) however 2 of these alleles had a 2 bp deletion at position 131–132 in the 5′-flanking region (No.13(-2) and No.14(-2), Fig. 2).

II. The intermediate and upper allele range between 265 and 309 bp which contained the hexanucleotide AAAAAG in addition to the regular 4 bp repeat. This irregular unit occurs only once in each fragment but at varying positions (Fig. 1.II, 2). This type showed 24 different sequence structures which are associated with 12 different fragment lengths. Variations were also observed in the 5′-flanking region of 4 alleles with AAAG deletions at the same site (position 113–116) and an AAG insertion at position 93–95 in the fourth (Fig. 2). Allele No. 27 (8-1-18; designation is according the sequence structure of $(AAAG)_8$-$(AAAAAG)_1$-$(AAAG)_{18}$) showed a transversion from A to T in repeat number 22, leading to TAAG (Fig. 2). Furthermore all alleles belonging to the type II-class showed a transition from A to G at position 4 in the 3′-flanking region (Fig. 2).

Construction of the allelic ladder

The designation of the alleles was according to Fig. 2 and Table 1. A total of 20 sequenced alleles were selected which span the whole allelic range (Table 1).

Separation of the alleles on a 6% denaturing gel and analysis on the ABI 373A Sequencer generated a regularly spaced allelic ladder (Fig. 3A) of fragment lengths corresponding to those obtained by sequencing, but were in general measured to be 6–10 bp longer.

The same alleles were separated on a horizontal high resolution electrophoresis system using non-denaturing

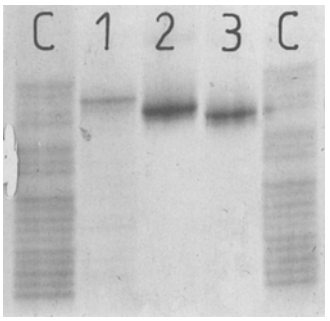

Fig. 4. Electrophoretic mobilities in a non-denaturing gel of 3 fragments with the same lengths (293 bp) but with different structures. C= allelic cocktail of 20 sequenced alleles according to Fig. 3B; 1= allele No. 29 (9-1-19), Fig. 2; 2= allele No. 30(-4) (11-1-18), Fig. 2; 3= allele No. 30(-4) (10-1-19), Fig. 2

gels (Fig. 3B). Despite the different fragment lengths only 18 instead of the expected 20 alleles were visible. The alleles No. 12/No. 13(-2) and No. 25/No. 26 showed the same electrophoretic mobility. Especially in the upper range, alleles differing by 4 bp revealed the same electrophoretic mobility. In addition alleles with the same fragment sizes but different sequence structure can show different electrophoretic mobilities (Fig. 4).

Discussion

Sequence data of 50 different alleles revealed not only length polymorphism but also a highly variable sequence polymorphism leading to a higher possibility of human identification. Length and sequence polymorphism have also been found within the human D11S554 locus (Adams et al. 1993) which contains repeat units similar to ACTBP2 in length and sequence. These findings indicate that the extent of microheterogeneity is correlated with the base composition in the repeat region. In contrast other STR systems such as HumTH01 (repeat: AATG) show a more regular structure (Puers et al. 1993) or little microheterogeneity such as HumVWA (repeat: TCTG, TCTA).

In addition to the published data (Polymeropolous et al. 1992b) 2 different types of sequence structures were found at the locus ACTBP2 depending on the allele sizes (Fig. 2). The tetranucleotide repeats were more characteristic of short alleles (223–259 bp) whereas a more complex structure was found in the upper allelic range. Repeat differences could have evolved from the regular AAAG repeat by insertions and/or deletions. So that this polymorphism could be informative for human evolution biology. Once the hexamer had been acquired it remained a stable component in all alleles above No. 22 (11-1-10) which would indicate that insertion of repeats was the more regular occurrence. Even deletions are sometimes conserved in subsequent alleles such as 30(-4) (11-1-18), 30(-4) (10-1-19) and 34(-4) (10-1-23). If the gap between these observations can be filled in later on such branching effects could become interesting in interracial comparisons.

As reported previously, DNA fragments with AT-rich sequences show anomalous electrophoretic mobilities (Mertz and Berg 1974; Marini et al. 1982; Hagerman 1985) which could be caused by secondary structure and/or DNA-interaction with the gel matrix (Sullivan et al. 1992; Kimpton et al. 1993).

Non-denaturing and denaturing gels systems were tested and it was demonstrated that electrophoretic mobilities are influenced by the gel system applied. In denaturing gels the migration of the alleles correlated with the fragment sizes determined by sequencing but were in general 6–10 bp longer. Fragment sizes were measured using PstI-digested lambda DNA as internal standard and did not represent the absolute length. This phenomenon could be explained by structural differences between human DNA and lambda DNA which can influence their electrophoretic mobilities (Grossman 1989). Difficulties can occur using non-denaturing gels because some fragment lengths determined by side-to-side comparison do not correspond to the appropriate fragment in the sequenced allelic ladder.

As a consequence, further application of the STR system ACTBP2 in forensic work needs the standardization of the experimental conditions e.g. gel systems, gel composition and allelic ladders.

With the proposed allelic ladder of sequenced alleles and a uniform electrophoretic system it would be possible to overcome these problems and to standardize an important feature for exact DNA typing in forensic work.

References

Adams M, Urquhart A, Kimpton C, Gill P (1993) The human D11S554 locus: four distinct families of repeat pattern alleles at one locus. Hum Mol Genet 2:1373–1376

Brinkmann B (1992) The use of STRs in stain analysis. In: Proceedings from the Third International Symposium on Human Identification. Promega Corporation, Madison, USA, pp 357–373

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

Budowle B, Chakraborty R, Giusti AM, Eisenberg AJ, Allen RC (1991) Analysis of the variable number of tandem repeat locus D1S80 by the polymerase chain reaction followed by high resolution polyacrylamide gel electrophoresis. Am J Hum Genet 48:137–144

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253

Grossman LI (1989) Gel electrophoresis of DNA. In: Proceedings of the International Symposium on the Forensic Aspects of DNA Analysis. FBI Academy, Quantico (Virginia), pp 37–46

Hagerman PJ (1985) Sequence dependence of the Curvature of DNA: A Test of the Phasing Hypothesis. Biochem 24:7033–7037

Kimpton CP, Walton A, Gill P (1992) A further tertanucleotide repeat polymorphism in the vWF gene. Hum Mol Genet 1:28

Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Applic 3: 13–22

Maniatis T, Fritsch EF, Sambrook, J (1989) "Molecular cloning: A Laboratory Manual", Cold Spring Harbour Laboratory, Cold Spring Harbor, New York

Marini JC, Levence SD, Crothers DM, Englund PT (1982) Bent helical structure in kinetoplast DNA. Proc Natl Acad Sci USA 79:7664–7668

Mertz JE, Berg P (1974) Viable Deletion Mutants of Simian Virus 40: Selective isolation by means of a Restriction Endonuclease

from hemophilus parainfluenzae. Proc Natl Acad Sci USA 71:4879–4883

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

Polymeropoulos MH, Xiao H, Merril CR (1992a) Tetranucleotide repeat polymorphism at the human myelin basic protein (MBP). Hum Mol Genet 1:658

Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1992b) Tetranucleotide repeat polymorphism at the human beta-actin related pseudogene H-beta-Ac-psi-2 (ACTBP2). Nucleic Acids Res 20:1432

Puers C, Hammond HA, Jin L, Caskey T, Schumm JW (1993) Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HumTH01(AATG) and reassignment of alleles in population analysis by using a locus-specific allelic ladder. Am J Hum Genet 53:953–958

Sullivan KM, Pope P, Gill P, Robertson JM (1992) Automated DNA profiling by fluorescent labeling of PCR products. PCR Methods Applic 2:34–40

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

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