J Forensic Sci, May 2008, Vol. 53, No. 3
doi: 10.1111/j.1556-4029.2008.00720.x
Available online at: www.blackwell-synergy.com

TECHNICAL NOTE

Magdalena Spolnicka,[1] Ph.D.; Agata Jagiello,[2] M.S.; Ireneusz Soltyszewski,[1] Ph.D.; Renata Jacewicz,[3] Ph.D.; Andrzej Plucienniczak,[2] Ph.D.; and Jaroslaw Berent,[3] Ph.D.

Sequence Analysis of a New Short Tandem Repeat Locus D17S2266E Located in the Human Growth Hormone Gene Cluster HumGH@*

ABSTRACT: D17S2266E is a new, variable genetic marker exhibiting polymorphism of the number of repeats of four- and two-nucleotide motifs. This study, carried out on a group of 250 unrelated persons from various regions of Poland, revealed the presence of 24 different alleles ranging in size from 232 to 290 base pairs. Analysis of the sequenced fragments demonstrated that the alleles consisted of two flanking regions and two variable blocks that were separated by a consensus sequence. There were $(AAAG)_5(AG)_1(AAAG)_{3–4}$ repeats in the first block, and $[(AAAG)_2(AG)_1]_{0–1}[(AAAG)_6(AG)_1]_{0–1}(AAAG)_n$ repeats in the second block. On the basis of the allele frequencies in the population, we were able to do biostatistical calculations, which gave the following results: expected heterozygosity 0.8947 ± 0.0137, power of discrimination 0.9793, polymorphism information content 0.8837, probability of exclusion (PE) 0.7859, PE for motherless cases 0.6473, and an average paternity index of 4.7470. These biostatistical parameters show that the marker D17S2266E can find a wide range of applications in forensic testing.

KEYWORDS: forensic science, DNA typing, DNA polymorphism, short tandem repeat marker, D17S2266E, HumGH@

Microsatellite short tandem repeat (STR) markers have been used for personal identification purposes for over 10 years and fulfill this function exceptionally well. Although many STR markers have been precisely described and are routinely used, many more are being identified (1–4).

We analyzed the sequence structure and population genetics of an STR marker located within the GH@ locus on chromosome 17 (5). Although polymorphism of locus D17S2266E has been known for a few years, proper genetic analysis and forensic application are presented in this report. The locus GH@ contains two growth hormone genes interspersed with three chorionic somatomammotropin genes, all in the same transcriptional orientation. The STR marker is situated between the growth hormone 1 gene (GH1) and the chorionic somatomammotropin hormone-like gene (CSHL1) (6).

Materials and Methods

Sample Preparation

The study material consisted of buccal samples taken from 250 nonrelated persons from different regions of Poland. The isolation of genomic DNA was carried out with the QIAGEN™ (Hilden, Germany) system. Additionally, genomic DNA was investigated from 50 mother–father–child trios (100 meioses) with genetically confirmed parentage, using Amp Fl STR Identifiler PCR amplification kit (Applied Biosystems, Foster City, CA). The quantity of DNA was determined by fluorometric measurement using a Fluroskan Ascent FL 2.5 (Venta, Finland) apparatus.

PCR Amplification

Optimization of the PCR amplification reaction was performed, i.e., primer concentrations, quantity of template DNA, and amplification conditions (temperature of hybridization, number of cycles). The PCR reactions were carried out using the GeneAmp PCR System 9700 (Applied Biosystems), volume 25 μL. Amplification conditions were determined to be as follows: 2.5 μL 10x PCR Buffer (Promega, Madison, WI), 1.5 μL 25 mM $MgCl_2$, 0.2 μL dNTPs 0.25 mM (Promega), forward and reverse primers 1.4 nM, 1 U *Taq* DNA polymerase (Promega), 2.5–5 ng genome DNA. Reaction temperature profile consisted of: initial denaturation 94°C—2 min, the subsequent 30 cycles: 94°C—30 sec, 64°C—30 sec, 72°C—30 sec, final extension 72°C—45 min. Forward primer 5′-GACAGAATGAGACTCCAACTG-3′ (marked with blue ester dye FAM) and reverse primer 5′-CAGAACAGGTAGTGTGGTGT-3′ were used for amplification (1).

STR Typing

The PCR reaction products were subjected to capillary electrophoresis using the ABI PRISM 3100 genetic analyzer (Applied Biosystems). The products were separated in a 36-cm capillary filled with POP4 gel (Applied Biosystems). The results obtained were analyzed using the programs GeneScan 3.7 NT and Genotyper 3.7 NT with respect to the internal size standard ILS 600 (Promega).

Allele Sequencing

For sequencing, alleles were chosen from homozygous individuals and heterozygous individuals differing by at least 8 bp. The

[1]Department of Biology, Central Forensic Laboratory, General Headquarters of Police, Iwicka 14, 00-735 Warsaw, Poland.
[2]Institute of Biotechnology and Antibiotics, Staroscinska 5, 02-516 Warsaw, Poland.
[3]Department of Forensic Medicine, Medical University of Lodz, Sedziowska 18a, 91-304 Lodz, Poland.

*Poster presentation at the 4th European Academy of Forensic Science Conference, Helsinki, June 13–16, EFAS 2006.

Received 24 Mar. 2007; and in revised form 13 Oct. 2007; accepted 13 Oct. 2007

samples were amplified in a PCR reaction using a specific forward primer 5′-GACAGAATGAGACTCCAACTG-3′ and a reverse primer 5′-CAGAACAGGTAGTGTG GTGT-3′. The PCR reactions were carried out using Pfu DNA Polymerase (EURx, Gdansk, Poland), in accordance with the manufacturer's recommendations. Each DNA sample reacted in two 25-μL Eppendorf tubes. The thermal conditions were as follows: initial denaturation 95°C—2 min, the following 32 cycles: 94°C—30 sec, 64°C—30 sec, 72°C—30 sec, final extension 72°C—10 min. The obtained products were separated in an 8% polyacrylamide gel (bis-acrylamide:acrylamide = 1:59) in TAE buffer for 3–4 h under an electric field of 10 V/cm. The separated DNA fragments were stained in ethidium bromide (2 μg/mL) and portions of the gel containing the DNA fragments were cut out and each was placed in a separate Eppendorf tube. The alleles were eluted from the gel using the method described by Dybczynski and Plucienniczak (7). The resultant DNA was dissolved in 10 μL of H_2O and then used in the sequencing reaction, which was carried out using the reagent system of ABI Prism BigDye Terminator 1.1 (Applied Biosystems), 50–100 ng DNA and 5 pM unlabelled primers. The following thermal conditions were applied in a 25-cycle sequencing reaction: 95°C—20 sec, 52°C—15 sec, 60°C—1 min. The products were separated in a 36-cm capillary filled with POP7 gel in an ABI PRISM 3730 apparatus.

Statistical Analysis

Statistical processing of the results was carried out using the computer programs GDA (8) and DLP (9), the data for which were prepared in the form of a standard NEXUS (10) computer file. With these programs, the observed and expected heterozygosity as well as its error (11), power of discrimination (PD) parentage (12), polymorphism information content (13), probability of exclusion for complete (14) and partial (15) parentage trios, and the minimum, maximum (16), and average/mean paternity index (17) were calculated.

Results

Location on the Chromosome

The STR marker analyzed is situated on chromosome 17 (at band 17q23) in the region of the *Homo sapiens* growth hormone locus GH@ (NCBI accession number NG_001334 base pairs 11,027–11,290), between gene GH1 and gene CSHL1. It should be noted that there are other STR markers located on chromosome 17. Klintschar et al (18) studied STR markers at the locus D17S976 in the region 17p11.2, and Polymeropulos et al. (19,20) studied another marker situated in the *H. sapiens* growth hormone locus GH@ (NCBI accession number NG_001334 base pairs 25,725–25,967) between gene CSHL1 and gene CSH1, at a distance of 14,435 bp from the D17S2266E marker.

Nomenclature

We submitted the studied marker to the GDB human genome database and it was given number D17S22 66E (GDB:11524910

TABLE 1—*Size, variable regions, nomenclature, and the number of sequenced alleles of D17S2266E.*

Allele size (bp)	Repeat regions	Allele designation	No. of sequenced alleles	GenBank accession number
290	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1(AAAG)_6(AG)_1(AAAG)_{18}$▲	36.2	2	EF010888
286	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1(AAAG)_6(AG)_1(AAAG)_{17}$▲	35.2	2	EF010887
282	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1(AAAG)_6(AG)_1(AAAG)_{16}$▲	34.2	2	EF010889
278	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1(AAAG)_6(AG)_1(AAAG)_{15}$▲	33.2	1	EF010890
276	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{21}$▲	33	1	EF010891
272	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{20}$▲	32	2	EF010892
268	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{19}$▲	31	3	EF010893
264	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{18}$▲	30	6	EF010894
262	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1(AAAG)_6(AG)_1(AAAG)_{11}$▲	29.2	1	EF010895
260	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{17}$▲	29	6	EF010896
260	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{18}$▲	29	1	EF010897
256	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{17}$▲	28	5	EF010898
256	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{16}$▲	28	2	EF010899
254	●$(AAAG)_5(AG)_1(AAAG)_3$■—$(AAAG)_{19}$▲	27.2	2	EF010900
252	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{16}$▲	27	2	EF010901
252	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{15}$▲	27	2	EF010902
250	●$(AAAG)_5(AG)_1(AAAG)_3$■—$(AAAG)_{18}$▲	26.2	3	EF010903
248	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{15}$▲	26	1	EF010904
248	●$(AAAG)_5(AG)_1(AAAG)_4$■$(AAAG)_2(AG)_1$–$(AAAG)_{14}$▲	26	3	EF010905
246	●$(AAAG)_5(AG)_1(AAAG)_4$■—$(AAAG)_{16}$▲	25.2	1	EF010906
244	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{14}$▲	25	2	EF010907
242	●$(AAAG)_5(AG)_1(AAAG)_3$■—$(AAAG)_{16}$▲	24.2	3	EF010908
240	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{13}$▲	24	1	EF010909
236	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{12}$▲	23	1	EF010910
234	●$(AAAG)_5(AG)_1(AAAG)_3$■—$(AAAG)_{14}$▲	22.2	2	EF010911
232	●$(AAAG)_5(AG)_1(AAAG)_3$■$(AAAG)_2(AG)_1$–$(AAAG)_{11}$▲	22	1	EF010912

● Flanking sequence 5′ (63 bp): forward primer sequence underlined
5′GACAGAATGAGACTCCAACTGAAAAAAAAGAAAGAAAGAGAGAGAGAGAGACAAAGAGGAAGG 3′
3′CTGTCTTACTCTGTGGTTGACTTTTTTTTCTTTCTTTCTCTCTCTCTCTCTGTTTCTCCTTCC 5′
■ Consensus sequence (34 bp) located between variable regions:
5′AAAAGGAAAGGAGGAAGGAAAGAAAAAGAAAAAG 3′
3′TTTTCCTTTCCTCCTTCCTTTCTTTTTCTTTTTC 5′
▲ Flanking sequence 3′ (47 bp): reverse primer sequence underlined
5′AAACTAAAATAACTAAATAACTGAGTAACACCACACTACCTGTTCTG 3′
3′TTTGATTTTATTGATTTATTGACTCATTGTGGTGTGATGGACAAGAC 5′

TABLE 2—*Distribution of D17S2266E allele frequencies in the Polish population.*

Allele	Size (bp)	No. in studied population	Frequency
36.2	290	2	0.0040
35.2	286	5	0.0100
34.2	282	2	0.0040
34	280	2	0.0040
33.2	278	5	0.0100
33	276	1	0.0020
32	272	9	0.0180
31	268	44	0.0880
30	264	64	0.1280
29.2	262	1	0.0020
29	260	91	0.1820
28.2	258	4	0.0080
28	256	76	0.1520
27.2	254	13	0.0260
27	252	64	0.1280
26.2	250	14	0.0280
26	248	34	0.0680
25.2	246	14	0.0280
25	244	21	0.0420
24.2	242	8	0.0160
24	240	12	0.0240
23	236	4	0.0080
22.2	234	4	0.0080
22	232	6	0.0120
Total		500	1.0000

and GDB:11524917). For the marker studied, we propose the name D17S2266E, in accordance with the recommendations of the Scientific Working Group on DNA Analysis Methods (21) and ISFH, currently ISFG (22,23).

For the alleles of the STR system, we applied the nomenclature recommended by the ISFG on the basis of the number of repeats (22,23). Sequences of 26 different D17S2266E alleles were submitted to the GenBank (Table 1).

Sequence Structure and Variation of D17S2266E Alleles

The 58 alleles of the locus D17S2266E from 250 unrelated persons were sequenced, and the results are presented in Table 1. The analysis of the sequenced fragments showed that the alleles are composed of two flanking regions (including the primer sites) and two variable blocks, which were separated by a consensus sequence (see Table 1). In the first block there were $(AAAG)_5(AG)_1(AAAG)_{3-4}$ repeats and in the second block $[(AAAG)_2(AG)_1]_{0-1}[(AAAG)_6(AG)_1]_{0-1}$ $(AAAG)_n$ repeats. The size of the alleles ranged from 232 to 290 bp. Alleles characterized by the same length may have different nucleotide sequences (22,23), and in our study, we observed this in the case of alleles 26–29 (Table 1). It is possible that sequence polymorphism may also appear among the other alleles.

Population Genetics

The analysis of 250 DNA samples from males and females in the Polish population revealed the presence of 24 different alleles ranging from 232 to 290 bp in size. The frequency distribution of the alleles was in agreement with that determined on the basis of the Hardy–Weinberg law (Fisher exact test: $p = 0.4678$). The analysis of allele segregation in 50 mother–father–child trios, i.e., in 100 meioses, showed no deviation from Mendel's first law with respect to the length of the fragments studied. The size of the locus D17S2266E alleles and the frequencies observed in the studied group are presented in Table 2. The biostatistical parameters determining the degree of polymorphism of this locus are given in Table 3.

TABLE 3—*Biostatistical parameters determining the degree of D17S2266E locus polymorphism.*

D17S2266E	
No. of persons studied	250
No. of alleles found	24
Heterozygosity—observed	0.8960
Heterozygosity—expected	0.8947
Error in expected heterozygosity	0.0137
Power of discrimination	0.9793
Polymorphism information content	0.8837
Power of exclusion for complete trios	0.7859
Power of exclusion for pairs without mother	0.6473
Minimum paternity index	1.4970
Typical paternity index	4.7470
Maximum paternity index	500

Discussion

It is worth emphasizing that only a few STR markers, such as ACTBP2, possess a number of alleles greater than that of the locus D17S2266E (24,25).

From among the 24 defined length variants at the locus D17S2266E, the most frequent in the studied population were the four alleles 27–30, and their frequency of occurrence was over 12%. By contrast, the least frequent alleles were 29.2 and 33, which appeared only once in the studied group. Only rarely are STR markers encountered that show a polymorphism similar to that of the marker we studied, with a PD value of 0.979, such as ACTBP2, Penta E, and D2S1338, for which the PD values for the Polish population are 0.991, 0.979, and 0.972, respectively (26–28). The favorable parameters of usefulness in forensic medical examinations (Table 3) indicate that D17S2266E could be a valuable forensic marker.

Acknowledgments

This work was supported by the State Committee for Scientific Research Grant No. 2 PO5 A 031 29.

References

1. Reichenpfader B, Leinzinger EP, Klintschar M. D1S1171: a new highly variable short tandem repeat polymorphism. Int J Legal Med 2002;116:195–7.
2. Wiegad P, Klintschar M. Population genetic data, comparison of the repeat structure and mutation events of two short STRs. Int J Legal Med 2002;116:258–61.
3. Henke L, Fimmers R, Reinhold J, Dulmer M, Cleef S, Arnold J, et al. Sequence analysis and population data on the 'new' short tandem repeat locus D5S2360. Forensic Sci Int 2001;116:55–8.
4. Barral S, Lareu MV, Salas A, Carracedo A. Sequence variation of two hypervariable short tandem repeats at the D22S683 and D6S477 loci. Int J Legal Med 2000;113:146–9.
5. Plucienniczak A, Plucienniczk G, Mikiewicz D, Jagiello A, Dabrowska H, Wojtuszek E, et al. (inventors). Institute of Biotechnology and Antibiotics, Central Forensic Laboratory-General Headquarters of Police (assignee). Oligonucleotide starters, method of analysing a genetic material and set of means therefore. Polish Patent PL187335, Jan. 31, 2000.
6. Chen EY, Liao YC, Smith DH, Barrera-Saldana HA, Gelinas RE, Seeburg PH. The human growth hormone locus: nucleotide sequence, biology, and evolution. Genomics 1989;4:479–97.
7. Dybczynski I, Plucienniczak A. A protocol for DNA fragment extraction from polyacrylamide gels. Biotechniques 1998;6:924–6.
8. Lewis PO, Zaykin D. Computer program: Genetic data analysis (GDA) version 1.0d16c, 2001. http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.
9. Berent J, Szram S. DLP—a computer program for calculation of discrete locus parameters. The Forensic Scientist, Ottawa, Canada: Shunderson Communications Inc., 2003;1:46–7.

10. Maddison DR, Swofford DL, Maddison WP. NEXUS: an extensive file format for systematic information. Syst Biol 1997;46:590–621.
11. Nei M, Roychoudhury AK. Sampling variants of heterozygosity and genetic distance. Genetics 1974;76:379–90.
12. National Research Council Report II. The evaluation of forensic DNA evidence. Washington, DC: National Academy Press, 1996;96–7.
13. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 1980;32:314–31.
14. Weir BS. Genetic data analysis II. Sunderland, MA: Sinauer Associates Inc. Publishers, 1996;209–11.
15. Fung WK, Chung YK, Wong DM. Power of exclusion revisited: probability of excluding relatives of the true father from paternity. Int J Legal Med 2002;116:64–7.
16. Berent JA, Miścicka-Śliwka D, Czarny J. Średnie wartości szansy ojcostwa—obliczenia dla populacji polskiej. Arch Med Sąd i Kryminol 1999;49:11–15.
17. Brenner C, Morris JW. Paternity index calculations in single locus hypervariable DNA probes: validation and other studies. Proceedings of the International Symposium on Human Identification, Apr. 12–14, 1989, Scottsdale, AZ. Madison, WI: Promega Corporation, 1990;21–53.
18. Klintschar M, Glock B, Dauber EM, Mayr WR. Genetic variation and sequence studies of highly variable short tandem repeat at the D17S976 locus. Int J Legal Med 1998;112:50–4.
19. Polymeropulos MH, Rath DS, Xiao H, Merril CR. A simple sequence repeat polymorphism at the human growth hormone locus. Nucleic Acids Res 1991;19:689.
20. Polymeropulos MH, Merril CR. Three highly informative microsatellite repeat polymorphic DNA markers. US Patent No. 5,468,610, 1995.
21. Scientific Working Group on DNA Analysis Methods (SWGDAM). Revised validation guidelines. Forensic Sci Commun 2004;6(3): 1–6. http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm
22. Bar W, Brinkmann B, Budowle B, Carracedo A, Gill P, Lincoln P, et al. DNA recommendations: further report of the DNA Commission of ISFH regarding the use of short tandem repeat systems. Int J Legal Med 1997;110:175–6.
23. Bar W, Brinkmann B, Lincoln P, Mayer WR, Rossi U. DNA recommendations—report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. Int J Legal Med 1994;107:159–60.
24. Moller A, Brinkmann B. Locus ACTBP2 (SE33): sequencing data reveal considerable polymorphism. Int J Legal Med 1994;106:262–7.
25. Moller A, Meyer E, Brinkmann B. Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. Int J Legal Med 1994;106:319–23.
26. Miścicka-Śliwka D, Czarny J, Grzybowski T, Woźniak M. Population genetics of STRs vWA, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820, D16S539, LPL, F13B, FESFPS, F13A01 and ACTBP2 in the Pomerania-Kujawy region of Poland. Forensic Sci Int 2001;119:119–22.
27. Sŏltyszewski I, Spólnicka M, Kartasińska E, Konarzewska M, Pepiński W, Janica J. Genetic variation of STR loci D2S1358, TH01, D21S11, D18S51, Penta E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta D, vWA, D8S1179, TPOX and FGA by *GenePrint* PowerPlex 16 in a Polish population. Forensic Sci Int 2006;3:241–3.
28. Czarny J. Population genetics of the STRs D3S1358, FGA, D2S1338, D8S1179, D21S11, D18S51 and D19S433 in the Pomerania-Kujawy region of Poland. Forensic Sci Int 2002;125:90–2.

Additional information and reprint requests:
Jaroslaw Berent, Ph.D.
Department of Forensic Medicine
Medical University of Lodz
Sedziowska 18a
91-304 Lodz
Poland
E-mail: j.berent@eranet.pl