



Second generation sequencing of three STRs D3S1358, D12S391 and D21S11 in Danes and a new nomenclature for sequenced STR alleles

Gelardi, Chiara; Rockenbauer, Eszter; Dalsgaard, Sigrun; Børsting, Claus; Morling, Niels

Published in:
Forensic science international. Genetics

DOI:
[10.1016/j.fsigen.2014.04.016](https://doi.org/10.1016/j.fsigen.2014.04.016)

Publication date:
2014

Citation for published version (APA):
Gelardi, C., Rockenbauer, E., Dalsgaard, S., Børsting, C., & Morling, N. (2014). Second generation sequencing of three STRs D3S1358, D12S391 and D21S11 in Danes and a new nomenclature for sequenced STR alleles. *Forensic science international. Genetics*, 12, 38-41. <https://doi.org/10.1016/j.fsigen.2014.04.016>



Second generation sequencing of three STRs D3S1358, D12S391 and D21S11 in Danes and a new nomenclature for sequenced STR alleles



Chiara Gelardi^{a,b,1}, Eszter Rockenbauer^{a,1,*}, Sigrun Dalsgaard^a, Claus Børsting^a, Niels Morling^a

^a Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

^b Faculty of Mathematical, Physical and Natural Sciences, University of Palermo, Palermo, Italy

ARTICLE INFO

Article history:

Received 24 February 2014
Received in revised form 29 April 2014
Accepted 30 April 2014

Keywords:

Second generation sequencing
STR
Nomenclature
Forensic genetics
Frequency database

ABSTRACT

Second generation sequencing (SGS) may revolutionize the field of forensic STR typing. Two of the essential requirements for implementation of an SGS based approach for forensic investigations are (1) establishment of adequate frequency databases and (2) adoption of a new STR nomenclature. We report the STR sequences and allele frequencies of three STR loci: D3S1358, D12S391 and D21S11 in 197 unrelated Danes. We used a new STR nomenclature that depicts the locus name used in forensic genetics, the length of the repeat region divided by the repeat length (typically 4 nucleotides) and detailed sequence information of possible sub-repeats and SNPs within the amplified fragment.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

STR typing and detection of PCR products by capillary electrophoresis (PCR-CE) has been the preferred DNA profiling method in forensic genetics for the past two decades. However, second generation sequencing (SGS) methods provide an interesting alternative to the PCR-CE methods. It was possible to investigate the true variation in STR loci with SGS and identify previously unknown alleles and mutational events in relationship cases [1–12]. It was also possible to identify SNP-STR haplotypes by sequencing both the STR and closely located SNPs [1,2,4,7]. Furthermore, it was possible to identify alleles that were inseparable by PCR-CE and thus, individuals could be correctly typed as heterozygote at a particular locus [1,10,11]. SGS also offers the possibility of combining STR typing with typing of forensically relevant SNPs or indels in the same assay. It may be an advantage in relationship case work to use polymorphic SNPs in order to reduce the risk of mutation events. In crime case work, STRs may be

supplemented with ancestry informative markers (AIMs) and/or markers for physical traits that may provide important information for the police investigation.

Implementation of SGS based assays in forensic genetics requires complete reevaluation of the current STR frequency databases, because more alleles may be identified by SGS. Also, it requires a new nomenclature for sequenced STR alleles, because some alleles with the same sizes have different sequences.

Here, we report an STR sequence database with sequence information and allele frequencies of two compounds (D3S1358 and D12S391) and one complex STR (D21S11) [13–17]. We used a nomenclature for sequenced STR alleles that may be used for future studies and applications of STR profiling.

2. Materials and methods

2.1. Samples and DNA extraction

A total of 197 samples from unrelated individuals in Danish paternity cases were sequenced for three STR loci (D3S1358, D12S391 and D21S11). DNA was extracted from anticoagulated blood, buccal swabs or FTA cards using the BioRobot EZ1 Workstation (Qiagen) following the recommendations of the manufacturer. The work was approved by the Danish ethical committee (H-1-2011-081).

* Corresponding author at: Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Frederik V's Vej 11, DK-2100 Copenhagen, Denmark. Tel.: +45 35 32 61 10; fax: +45 35 32 61 20.

E-mail addresses: eszter.rockenbauer@sund.ku.dk, rockenbauer@hotmail.com (E. Rockenbauer).

¹ These authors contributed equally to the work.

2.2. PCR-CE

The samples were typed with the AmpFISTR[®] NGM SELECT[™] PCR Amplification kit (Life Technologies (LT)) according to the guidelines of the manufacturer. Prior to electrophoresis, 1 μ L of the amplified product and 0.5 μ L of GeneScan[™] 600 LIZ[®] Size Standard v2.0 (LT) was added to 9.5 μ L of deionized Hi-Di[™] formamide (LT), denatured at 95 °C for 3 min and placed on an Eppendorf[®] PCR-Cooler (Eppendorf) for 3 min. Samples were injected for 10 s at 3 kV and separated by electrophoresis in POP-4[®] Polymer (LT) using the HIDFragmentAnalysis36_POP4 Module (LT) and a 1500 s run time on the 3130xl Genetic Analyzer (LT).

Collection files were analyzed with GeneMapper[®] ID-X Software 1.1.1 (LT) using the default settings except for the following changes: in the Analysis Method Editor, the peak amplitude threshold was 50 RFUs for all colors, the Peak Height Ratio was 0.5 and the Min/Max Peak Height was 150 RFU for homozygotes and 50 RFU for heterozygotes.

The allele calls determined by the AmpFISTR[®] NGM SELECT[™] PCR Amplification kit (LT) were used for comparison with the allele calls from the sequencing data.

2.3. 454 sequencing

A total of 24 sets of fusion primers with multiplex identifiers (MIDs) designed according to the protocol “454 Sequencing System Guidelines for Amplicon Experimental Design” (Roche Diagnostics, Basel, Switzerland, July 2011) and [18] were used for each locus. The sequences of the D3S1358 fusion primers are presented in Supplementary Table ST1. The sequences of the D21S11 and D12S391 fusion primers were previously described [10,11]

For the PCR amplification, two different approaches were used: the D3S1358 and D21S11 loci were amplified together, while the D12S391 locus was amplified alone. Each reaction contained 1 \times AmpliTaq Gold buffer (LT), 2.5 mM MgCl₂ (LT), 0.5 μ M forward and 0.5 μ M reverse primers (DNA Technology, Denmark), 0.2 mM of each dNTP, 0.2 μ L AmpliTaq Gold[®] (LT), and 2 μ L DNA. The total reaction volume was 25 μ L. The PCR cycling conditions were as follows: 94 °C for 10 min, 30 cycles of 94 °C for 30 s, 63 °C for 30 s, 72 °C for 30 s and a final extension at 72 °C for 10 min.

The PCR products were visualized with LabChip[®] GX with the HT DNA High Sensitivity LabChip[®] Kit (Caliper Life Sciences) as recommended by the manufacturer. Excess primers and nucleotides were removed using the Agencourt AMPure XP PCR Purification system (Beckman Coulter) following the recommendations of the manufacturer. The PCR products were quantified using the Qubit[®] 2.0 Fluorometer with the Qubit[®] dsDNA HS Assay Kit (LT). Sample concentrations were calculated following the guidelines of the manufacturer.

All samples were diluted to a final concentration of 10⁸ molecules/ μ L. The amplified samples were pooled into libraries and sequenced on a 454 GS Junior System (Roche). Prior to sequencing, an emulsion PCR-step was performed according to the manual, emPCR Lib-A (Roche Diagnostics). The 197 samples were analyzed in ten experiments.

2.4. Data sorting

An algorithm described in Fordyce et al. [1] was used to analyze the GS Junior FASTA reads. The reads were sorted by MID-tags, primer sequences and STR-ends. Three output files, one for each STR locus, were created for each sample. Sequences containing erroneous base calls were removed and consensus sequences for each allele were created using the alignment software BioEdit (URL: <http://www.mbio.ncsu.edu/bioedit/bioedit.html>). Allele

calls (compatible with http://www.cstl.nist.gov/strbase/str_fact.htm) were made by comparing the consensus sequences to GenBank reference sequences (rcAC099539 (reverse complement sequence of AC099539), AP000433 and G08921 for D3S1358, D12S391 and D21S11, respectively).

2.5. Statistical calculations

Mean match probability (MP), mean paternity exclusion power (A) and typical PI for trios and duos were calculated in DNAVIEW 28.103 based on PCR-CE or SGS results. The mean power of discrimination (PD) was calculated as 1-MP.

3. Results

A total of 197 unrelated Danes were sequenced for the STR loci D3S1358, D12S391 and D21S11 to establish a sequence database for forensic investigations in Denmark. In average, each sample was represented with 631 sequence reads, and of these 33.8%, 33.8% and 32.4% were sorted to D3S1358, D12S391 and D21S11, respectively. Sequence reads containing the sequence of at least one primer site and both STR ends were selected and analyzed as previously described [1]. Length artifacts, including stutters, constituted 10.6%, 17.5% and 9.6% of the analyzed reads of D3S1358, D12S391 and D21S11, respectively. Substitutions, deletions and insertions were observed on average in 17.6%, 24.2% and 21.3% of the analyzed reads for D3S1358, D12S391 and D21S11, respectively. The PCR and sequence errors were randomly scattered across the sequence reads.

Sequences of the identified alleles were grouped for each locus and aligned to the reference sequences (GenBank rcAC099539, AP000433 and G08921, respectively). Allele calls were compared to those determined with the AmpFISTR[®] NGM SELECT[™] PCR Amplification kit (LT). Complete concordance was observed between the PCR-CE and the SGS results (data not shown). As expected from earlier observations [1,10,11], the length of the allele influenced the number of observed reads. The shortest of the two alleles within a system had approximately 1.2 times more reads than the longest of the two alleles (data not shown). Allele annotations and allele frequencies of the 197 individuals are listed in Supplementary Tables ST2–ST4. Of the three STR loci investigated, D3S1358 showed the lowest degree of variation. In D3S1358, 17 alleles with frequencies between 0.0025 and 0.2538 were found (Supplementary Table ST2). Three of these are reported here for the first time (Table 1). No variation was observed in the first sub-repeat. The second sub-repeat contained 1–4 repeats, while the third one contained 8–15 repeats. D12S391 was the most polymorphic locus of the three. A total of 53 different alleles were found, 18 of these are reported here for the first time (Table 1). The allele frequencies of D12S391 were between 0.0025 and 0.1421 (Supplementary Table ST3). The frequent T/C SNP, rs77312049, in the last sub-repeat was observed in the sequenced samples with a minor allele frequency of 0.3046. Three of the 18 new alleles had a rare but known A/G SNP (rs150199795) positioned in the first repeat of the first sub-repeat section, where the AGAT repeat was changed to AGGT. Another new allele, annotated “D12S391[21]AGAT[13]GGAC[1]AGAC[7]”, had a previously unknown variant in the first repeat of the second sub-repeat section that changed the first adenosine in the AGAC repeat unit into a guanine.

Finally, in D21S11, 29 alleles were found and the allele frequencies were between 0.0025 and 0.1777 (Supplementary Table ST4). Nine of the 29 alleles are reported here for the first time (Table 1). Variations were observed in three variable sub-repeat sections with 4–8, 5–7 and 8–15.2 repeats, respectively. As expected, no variation was seen in the intervening sequence. No

Table 1
Novel D3S1358, D12S391 and D21S11 alleles.

CE allele name	SGS allele name	Count	Frequency
14	D3S1358[14]TCTA[1] TCTG[3]TCTA[10]	1	0.0025
14	D3S1358[14]TCTA[1] TCTG[1]TCTA[12]	1	0.0025
20	D3S1358[20]TCTA[1] TCTG[4]TCTA[15]	1	0.0025
18	D12S391[18]AGAT[11] AGAC[7]	3	0.0076
18	D12S391[18]AGAT[12] AGAC[5]AGAT[1]	4	0.0102
19	D12S391[19]AGAT[11] AGAC[7]AGAT[1]	1	0.0025
19	D12S391[19]AGAT[11] AGAC[8]	6	0.0152
20	D12S391[20]AGGT[1] AGAT[10]AGAC[9]	2	0.0051
21	D12S391[21]AGAT[13] GGAC[1]AGAC[7]	1	0.0025
21	D12S391[21]AGGT[1] AGAT[11]AGAC[9]	1	0.0025
22	D12S391[22]AGAT[13] AGAC[8]AGAT[1]	7	0.0178
22	D12S391[22]AGAT[14] AGAC[7]AGAT[1]	1	0.0025
22	D12S391[22]AGAT [14]AGAC[8]	4	0.0102
23	D12S391[23]AGAT[12] AGAC[10]AGAT[1]	3	0.0076
23	D12S391[23]AGAT [12]AGAC[11]	3	0.0076
23	D12S391[23]AGAT[15] AGAC[7]AGAT[1]	1	0.0025
23	D12S391[23]AGAT[15] AGAC[8]	1	0.0025
24	D12S391[24]AGAT[13] AGAC[10]AGAT[1]	1	0.0025
24	D12S391[24]AGGT[1] AGAT[14]AGAC[9]	1	0.0025
25	D12S391[25]AGAT[14] AGAC[11]	1	0.0025
25	D12S391[25]AGAT[15] AGAC[9]AGAT[1]	1	0.0025
28	D21S11[28]TCTA[4] TCTG[5]itvsTCTA[11]	1	0.0025
28	D21S11[28]TCTA[6] TCTG[5]itvsTCTA[9]	2	0.0051
29	D21S11[29]TCTA[4] TCTG[7]itvsTCTA[10]	1	0.0025
29	D21S11[29]TCTA[5] TCTG[6]itvsTCTA[10]	1	0.0025
30	D21S11[30]TCTA[7] TCTG[5]itvsTCTA[10]	2	0.0051
31	D21S11[31]TCTA[4]T CTG[6]itvsTCTA[13]	5	0.0127
32	D21S11[32]TCTA[8] TCTG[5]itvsTCTA[11]	3	0.0076
32.2	D21S11[32.2] TCTA[6]TCTG[6]itvsTCTA[11] TA[1]TCTA[1]	1	0.0025
33	D21S11[33]TCTA[6]TCTG[5]i tvsTCTA[14]	1	0.0025

itvs: TCTA[3]TA[1]TCTA[3]TCA[1]TCTA[2]TCCA[1]TA[1].

SNP was found within the sequences of the D21S11 and D3S1358 STR loci.

Only one pair of individuals had the same three STR types after SGS, while three pairs of individuals had the same haplotype when typed with PCR-CE. Individuals seemingly homozygous with PCR-CE for an STR locus were found to be heterozygous in 23%, 33% and 41% of the cases in D21S11, D12S391 and D3S1358, respectively.

Table 2 shows the forensic statistics based on the PCR-CE and SGS data. The mean match probability was 15 times lower and the

typical PIs for trios and for duos were 7 and 5 times higher, respectively, for the SGS data compared to the PCR-CE data.

4. Discussion

The use of SGS instead of the traditional PCR-CE fragment analysis for D3S1358, D12S391 and D21S11 resulted in the identification of 2.75 times more alleles, higher typical paternity indices and improved power of discrimination. The additional information that is obtained with SGS compared to PCR-CE is only one of the advantages of SGS. Combining genotyping of STRs and SNPs in the same assay will have a large impact on forensic genetics and it will give researchers and case officers important new information to work with.

STR loci, such as D3S1358, D12S391 and D21S11, have several alleles with the same lengths but with different sequence compositions. The variations may be variations in the number of sub-repeats or variations in the nucleotide sequence in form of SNPs or indels in non-repetitive areas. When establishing a SGS-based frequency database, all of these variations must be taken into account and the alleles must be given a unique name. In earlier SGS-studies [1,2,5,7,8] alleles of the same lengths but with different sequences were named e.g. 20a, 20b, 20c etc. or 20, 20' 20'' etc., for alleles with 20 repeats. Obviously, this was not a transparent method for naming alleles and not a perpetual solution. Thus, guidelines from e.g. ISFG or another standardization body are needed to create a new standard for naming sequenced STR alleles.

In this work, the nomenclature was based on the recommendations by the human genome variation society (<http://www.hgvs.org/>) and the currently used nomenclature on STRbase (<http://www.cstl.nist.gov/div831/strbase/>). We used a nomenclature with four elements; (1) the locus name used in forensic genetics, e.g. D12S391, (2) the length of the repeat region divided with the length of the repeat unit, (3) the sequence(s) of the sub-repeat(s) followed by the number of sub-repeats, e.g. AGAT[11]A-GAC[7] and finally (4) variations in the flanking regions preferably identified by an rs number and the base call, e.g. rs6736691[G] [7]. If an rs number is not available, the chromosome position(s) of the variant(s) must be listed. Known SNPs with rs numbers in the STR region should not be mentioned but must be part of element no. 3. The first and second elements of the name will be identical to the nomenclature used for most alleles determined by PCR-CE. However, there will be examples of inconsistencies if there is an indel in the flanking region between the STR and PCR primer binding site. One such example was previously reported [1], where a D21S11[28] allele was typed as a D21S11[27.1] allele by PCR-CE because of a 3 bp deletion downstream of the STR region. The third and fourth elements of the name include the sequence information and may be used to deduce the exact sequence of the allele if standard reference sequence is defined and used. With this nomenclature, some of the allele names will be very long (see Table 1 and Supplementary Tables ST2–4), but that is unavoidable if the additional sequence information provided by SGS is to be used. However, about half of the core STRs used in forensic genetics are simple STRs containing one kind of repeat unit. The allele names for these will be relatively short, e.g. CSF1PO[11]AGAT[11].

The provider of the platform used in this work (Roche) has recently announced that production of the 454 sequencers will be terminated in 2015. Thus, the 454 GS junior platform is not likely to be the future forensic genetic platform. However, two other major providers of SGS equipment, Illumina® and Life Technologies™, are currently working on developing kits for forensic genetics. Software for SGS data analysis will be a vital point for the introduction of SGS into forensic genetics, because the amount of sequence data is so large that the user will be unable to analyze

Table 2

Forensic statistics based on CE and SGS data from 197 unrelated danes.

	CE				SGS			
	D3S1358	D12S391	D21391	Cumulative mean	D3S1358	D12S391	D21391	Cumulative mean
MP	0.0855	0.0211	0.0389	0.0001	0.0372	0.0077	0.0164	0.000005
PD	0.9145	0.9789	0.9611	0.9999	0.9628	0.9923	0.9836	0.999995
A (%)	55.9	78.2	69.9	97.1	70.2	86.6	80.5	99.2
PI	2.62	5.49	4.12	59.2	4.38	12.6	7.53	415.0
mPI	1.86	3.31	2.63	16.1	2.76	6.9	4.32	82.4

MP: mean match probability, PD: mean power of discrimination (1-MP), A: mean paternity exclusion power, PI: typical paternity index (trios), mPI: typical paternity index (duos).

each profile in detail and will have to rely on the software solutions. Several bioinformatics tools are already under development [1,6,9,12]. However, a close collaboration between companies and the forensic community will be essential for the development of better and fully automated SGS software(s).

Acknowledgements

We thank Lena Poulsen and Carmen Thomas Mas for help with the statistical calculations. We thank Nadia Jochumsen and Anja Jørgensen for technical assistance. The work of Chiara Gelardi was partly funded by the Borsa di Perfezionamento all'Estero scholarship granted by MIUR (Ministero dell'Istruzione, Università e Ricerca) and the University of Palermo.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.fsigen.2014.04.016](https://doi.org/10.1016/j.fsigen.2014.04.016).

References

- [1] S.L. Fordyce, M.C. Avila-Arcos, E. Rockenbauer, C. Borsting, R. Frank-Hansen, F.T. Petersen, E. Willerslev, A.J. Hansen, N. Morling, M.T. Gilbert, High-throughput sequencing of core STR loci for forensic genetic investigations using the Roche Genome Sequencer FLX platform, *Biotechniques* 51 (2011) 127–133.
- [2] E. Rockenbauer, M.B. Holgersson, S.L. Fordyce, M.C.Á. Arcos, C. Børsting, A.J. Hansen, R. Frank-Hansen, E. Willerslev, M.T.P. Gilbert, N. Morling, Sequences of microvariant/"off-ladder" STR alleles, *Forensic Sci. Int.: Genet. Suppl. Ser. 3* (2011) e204–e205.
- [3] M. Scheible, O. Loreille, R. Just, J. Irwin, Short tandem repeat sequencing on the 454 platform, *Forensic Sci. Int.: Genet. Suppl. Ser. 3* (2011) e357–e358.
- [4] C. Van Neste, F. Van Nieuwerburgh, D. Van Hoofstat, D. Deforce, Forensic STR analysis using massive parallel sequencing, *Forensic Sci. Int. Genet.* 6 (2012) 810–818.
- [5] D.M. Bornman, M.E. Hester, J.M. Schuetter, M.D. Kasoji, A. Minard-Smith, C.A. Barden, S.C. Nelson, G.D. Godbold, C.H. Baker, B. Yang, J.E. Walther, I.E. Tornes, P.S. Yan, B. Rodriguez, R. Bundschuh, M.L. Dickens, B.A. Young, S.A. Faith, Short-read, high-throughput sequencing technology for STR genotyping, *Biotechniques* 0 (2012) 1–6.
- [6] M. Gymrek, D. Golan, S. Rosset, Y. Erlich, lobSTR: a short tandem repeat profiler for personal genomes, *Genome Res.* 22 (2012) 1154–1162.
- [7] S. Dalsgaard, E. Rockenbauer, C. Gelardi, C. Børsting, S.L. Fordyce, N. Morling, Characterization of mutations and sequence variations in complex STR loci by second generation sequencing, *Forensic Sci. Int.: Genet. Suppl. Ser. 4* (2013) e218–e219.
- [8] E. Rockenbauer, S. Hansen, M.R. Andersen, E.F. Heerup, C. Børsting, S.L. Fordyce, N. Morling, Characterization of sequence variations in the D21S11 locus in Danes, Somalis and Greenlanders by second generation sequencing, *Forensic Sci. Int.: Genet. Suppl. Ser. 4* (2013) e302–e303.
- [9] D.H. Warshauer, D. Lin, K. Hari, R. Jain, C. Davis, B. LaRue, J.L. King, B. Budowle, STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data, *Forensic Sci. Int. Genet.* 7 (2013) 409–417.
- [10] E. Rockenbauer, S. Hansen, M. Mikkelsen, C. Borsting, N. Morling, Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing, *Forensic Sci. Int. Genet.* 8 (2014) 68–72.
- [11] S. Dalsgaard, E. Rockenbauer, A. Buchard, H.S. Mogensen, R. Frank-Hansen, C. Borsting, N. Morling, Non-uniform phenotyping of D12S391 resolved by second generation sequencing, *Forensic Sci. Int. Genet.* 8 (2014) 195–199.
- [12] C. Van Neste, M. Vandewoestyne, W. Van Criekinge, D. Deforce, F. Van Nieuwerburgh, My-Forensic-Loci-queries (MyFLq) framework for analysis of forensic STR data generated by massive parallel sequencing, *Forensic Sci. Int. Genet.* 9 (2014) 1–8.
- [13] H. Li, L. Schmidt, M.-H. Wei, T. Hustad, M.I. Lerman, B. Zbar, K. Tory, Three tetranucleotide polymorphisms for loci: D3S1352, D3S1358, D3S1359, *Hum. Mol. Genet.* 2 (1993) 1327.
- [14] M.V. Lareu, M.d.C. Pestoni, F. Barros, A. Salas, A. Carracedo, Sequence variation og hypervariable short tandem repeat at the D12S391 locus, *Gene* 182 (1996) 151–153.
- [15] V. Sharma, M. Litt, Tetranucleotide repeat polymorphism at the D21S11 locus, *Hum. Mol. Genet.* 1 (1992) 67.
- [16] A. Urquhart, C.P. Kimpton, T.J. Downes, P. Gill, Variation in short tandem repeat sequences – a survey of twelve microsatellite loci for use as forensic identification markers, *Int. J. Leg. Med.* 107 (1994) 13–20.
- [17] P. Gill, C.P. Kimpton, A. Urquhart, N. Oldroyd, E.S. Millican, S.K. Watson, T.J. Downes, Automated short tandem repeat (STR) analysis in forensic casework – a strategy for the future, *Electrophoresis* 16 (1995) 1543–1552.
- [18] M.C. Kline, C.R. Hill, A.E. Decker, J.M. Butler, STR sequence analysis for characterizing normal, variant, and null alleles, *Forensic Sci. Int. Genet.* 5 (2011) 329–332.