J Forensic Sci, Nov. 2004, Vol. 49, No. 6
Paper ID JFS2003145
Available online at: www.astm.org

Tim M. Clayton,[1] Ph.D.; Jennifer L. Guest,[2] B.Sc.; Andy J. Urquhart,[2] Ph.D.; and Peter D. Gill,[2] Ph.D.

A Genetic Basis for Anomalous Band Patterns Encountered During DNA STR Profiling

ABSTRACT: Since 1995 the Forensic Science Service (FSS) has carried out DNA profiling of reference samples for the UK National DNA Database (1) and in forensic casework using two multiplex STR profiling systems (2,3). During this period, profiles with anomalous banding patterns, although comparatively rare, have been encountered regularly. The FSS has collected instances of triallelic patterns and aberrant diallelic patterns. A systematic examination of these patterns has provided insight into their underlying genetic cause. The triallelic patterns could be classified into two types based on the relative intensities of their component alleles. In the Type 1 pattern the alleles were of uneven intensity, whereas in the Type 2 pattern, all three alleles were of even intensity. Evidence is presented that the more frequent Type 1 pattern is the result of somatic mutation at a heterozygous locus, and the Type 2 pattern is the result of a localized chromosomal rearrangement at a heterozygous locus. Directly from the Type 1 pattern, it was possible to deduce the size difference between the progenitor and mutated allele. All mutational changes were found to be multiples of four nucleotides, suggesting the loss or addition of one or more tetrameric repeat units. Aberrant diallelic patterns were identified by analysts due to an unexpectedly large difference in intensity between alleles at a heterozygous locus. While some of these diallelic patterns are likely caused by the same genetic phenomena described above occurring at a homozygous locus, others are demonstrated to be caused by a mutation in the primer binding sequence, leading to a reduction in amplification efficiency of one allele. It is concluded that based on a visual inspection of a profile, it is possible to infer a likely genetic basis directly from the triallelic pattern. By contrast, the aberrant diallelic patterns can be due to any one of a number of possible genetic effects.

KEYWORDS: forensic science, STR loci, triallelic pattern, somatic mutation, mosaicism, primer binding site mutation, chromosomal rearrangement, aneuploidy

Short Tandem Repeat (STR) loci are widely used as linkage markers in forensic casework and for familial analysis. Most commonly, in forensic and familial applications, autosomal STR loci are amplified using Polymerase Chain Reaction (PCR), and the fluorescent products are separated according to size using gel or capillary electrophoresis. Typically, a heterozygous locus produces similar intensity signals from both alleles although there is always a certain amount of natural variation (4). A perturbation in the balanced signal can be the result of one or more of a number of factors associated with PCR or due to degradation of the DNA template (5). However, if PCR conditions are optimal and the template is not degraded, then the extent of variation is generally small (4), and a balanced diallelic signal can be obtained. Anomalous multibanded signals at a locus can be due to a variety of artifacts associated with amplification and detection (6,7). These effects are diagnosed easily or can be eliminated simply by performing a re-analysis of the sample concerned. Contamination with an exogenous source of DNA also can result in the observation of an additional band or bands at a locus, although this is usually evident at several loci within a multiplex and is therefore relatively straightforward to recognize (8). The presence of non-specific artifact bands has been reported and is apparently the result of random priming during PCR (4). Generally, these artifacts are recognizable because either the fluorescent dye label is the wrong color for the locus concerned or the questioned band is not in an allelic position with respect to an allelic ladder. Moreover, they tend to exhibit atypical peak morphology and to be relatively small (4). Since the advent of "hot start" techniques, these types of artifacts have been practically eliminated.

While carrying out analysis of reference samples for addition to the UK National DNA Database (1) and during routine forensic casework, instances were uncovered where the profile of an individual displayed an anomalous tri- or diallelic banding pattern. Analysts had initially identified an anomalous banding pattern because it could not be reconciled with any of the above explanations. The bands observed were always in an allelic position with respect to an allelic ladder. Individually, they were restricted to a single locus within a profile, but collectively they were distributed across all but one of the loci examined in the multiplex systems employed. Aberrant, diallelic patterns were identified by analysts for one of two reasons: there was an unexpectedly large ($>60\%$) difference in the peak area of two alleles at an apparently heterozygous locus, or a minor band (in the -4 position) in an apparently homozygous locus was unexpectedly large ($>15\%$) (4). The anomalous tri- and diallelic profiles were gathered and studied further, as it was considered that they had an underlying genetic cause. Where additional samples in a case were available, these were also analyzed. The focus of the study was to establish the genetic basis of the observed patterns.

Materials and Methods

The six of the ten STR loci used in this study are contained in a multiplex kit (FSS Second Generation Mulitplex – SGM) developed by the FSS in 1995 (2). The FSS SGM multiplex was used from 1995 until the beginning of 1999 for DNA profiling for the UK National DNA database (1). It contains the vWFA31/A, THO1,

[1] The Forensic Science Service, Paternity Analysis Unit, Sandbeck Way, Wetherby, West Yorkshire, LS22 7DN, UK.
[2] The Forensic Science Service, Research and Development Division, 2690 Trident Court, Birmingham Business Park, Solihull, West Midlands, B37 7YN, UK.

Received 29 Apr. 2003; and in revised form 27 Jan. 2004 and 2 June 2004; accepted 5 June 2004; published 5 Oct. 2004.

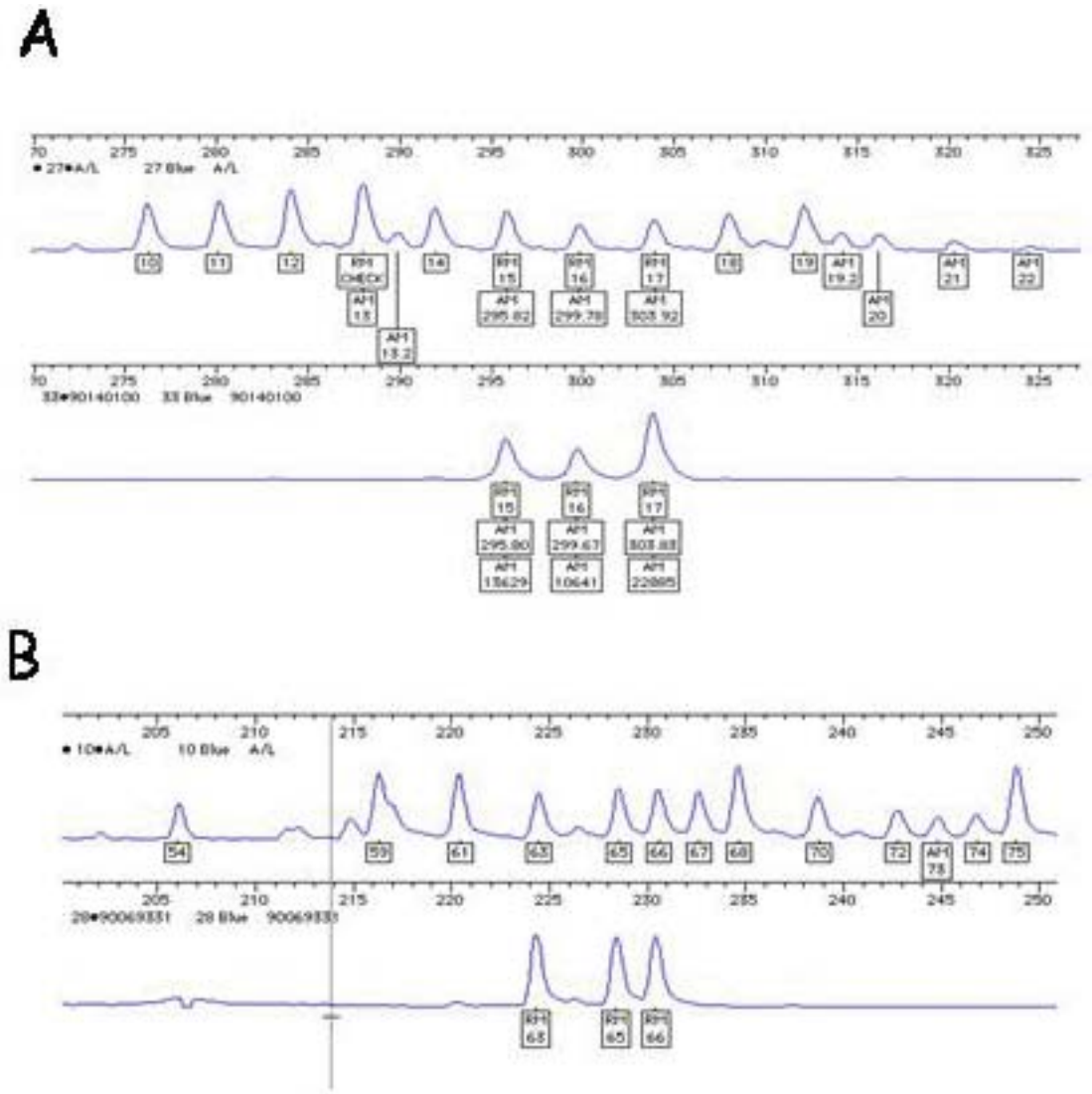

FIG. 1—*Upper panel (A): D18S51 locus in the FSS SGM multiplex. Lane 33 shows a sample displaying a Type 1 triallelic pattern, and lane 27 contains the corresponding allelic ladder with allelic designations in the boxes beneath. Peak area information is provided in the lower set of boxes under lane 33. Note that the peak area of the 15 and 16 allele (13629 and 10641) approximately sum to that of the 17 allele (22885). Lower panel (B): D21S11 locus in the FSS SGM multiplex. Lane 28 shows a sample displaying a Type 2 triallelic pattern, and lane 10 contains the corresponding allelic ladder with allelic designations in the boxes beneath. Note the even signal intensity of the alleles.*

D8S1179, FGA, D21S11, and D18S51 loci together with a sex test based on the XY homologous gene Amelogenin. The ten locus AMP*Fl*STR® SGM Plus™ was developed and introduced in 1999 to replace FSS SGM (3). The kit is manufactured commercially by PE Biosystems (Warrington, England) and incorporates the six FSS SGM loci and Amelogenin together with four additional loci (D2S1338, D3S1358, D16S539, and D19S433). With the exception of the D2S1338 and D19S433 loci, these loci are also contained with the commercially available Powerplex™ 16 kit (Promega, Madison, WI, USA). The Promega kit utilizes different primer sets to those employed in both FSS SGM and AMP*Fl*STR® SGM Plus™.

Extraction and Quantification of DNA

DNA was extracted from blood, cheek (buccal) scrapes, and hair using the Chelex extraction method of Walsh et al. (9). Quantification of the DNA was performed in a dot blot assay using a primate specific probe (D17Z1) according to Walsh et al. (10).

Amplification

DNA amplification was carried out on Perkin Elmer 9600 thermalcyclers using 1 ng of DNA template according to the conditions specified in (2,3) and, in the case of Powerplex™ 16 (Promega), according to the conditions specified by the manufacturer.

Electrophoresis and Detection of Amplified Fragments

Electrophoretic separation of flourescently labeled amplified fragments was achieved using 6% polyarlamide denaturing gels on the PE Biosystems 373 or 377 platforms as described (2,3). Sizing was achieved using GS350/500 Tamra as the internal size standard (PE Biosystems) and GeneScan Analysis v2.1.1 software. Alleles were designated using Genotyper v2.1 software in conjunction with allelic ladders. Nomenclature of alleles follows the recommendations of the International Society for Forensic Genetics (11) with the exception of D21S11, which, historically, used a different form of notation in the FSS SGM kit (12).

Results

Collection of Data

Three hundred and forty two triallelic patterns were collected and studied. These were mainly (97%) profiles produced with the FSS SGM kit. This was due, in part, to the fact that historically more samples were processed with the FSS SGM kit and, in part, to the cessation of an internal policy for systematically collecting anomalous profiles after inception of the AMP*Fl*STR® SGM Plus™ kit into forensic casework. Four hundred instances of aberrant diallelic profiles also were uncovered during the same period.

Investigation of Triallelic Patterns

Analysis of the collection of triallelic patterns revealed two different types that could be differentiated according to the relative intensity of the component alleles. The Type 1 pattern was the more frequent (274/342) and consisted of three alleles of unequal signal intensity. A representative example is shown in the upper panel of Fig. 1. These patterns possessed a second distinctive characteristic in that the peak areas of the two smaller (minor) alleles when summed together were approximately equivalent in peak area to the larger (major) alleles. The 68 Type 2 patterns consisted of three alleles of even signal intensity. A representative example is shown in the lower panel of Fig. 1. The data are summarized in Table 1, broken down by locus and type of pattern. From these data it can be seen that there are clear differences in the distribution of both types of triallellic patterns dependent on which locus is considered. Note that the D8S1179 locus accounts for less than 7% of the total of the Type 1 pattern but about 35% of the Type 2 pattern. This contrasts with the D18S51, where that locus accounts for 33% of the total of the Type 1 pattern but about 12% of the Type 2 pattern.

Are Type 1 Patterns the Result of Somatic Mutation?

It was suspected that the Type 1 patterns were the result of somatic mutation at a heterozygous locus as illustrated in the schematic diagram in Fig. 2. If this hypothesis was correct, then the relative copy number of each of the two minor alleles should be variable from case to case depending upon the particular stage of tissue development at which the mutational event occurred. The data showed that there was indeed variation in the relative ratios of the peak area values between the minor alleles in different individuals. In some instances the two minor alleles were of equal intensity (10:5:5), in others one of the minor alleles was of marginally greater intensity than the other (10:6:4), and in extreme cases one of the minor alleles

TABLE 1—*Data by locus and pattern type.*

Locus	Type 1	Type 2
vWA	42	8
THO1	2	1
D8	19	24
D21	53	13
D18	91	8
FGA	60	12
D2*	3	1
D16*	3	0
D19*	0	0
D3*	1	1
	274	68

* These loci account for only 3% of the total observations. Refer to text for explanation of data.

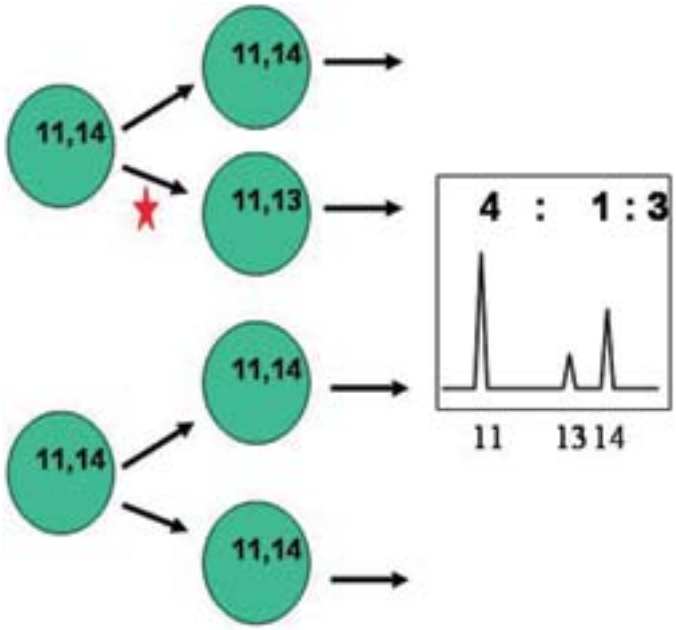

FIG. 2—*Schematic explanation of somatic mutation. This example begins with two cells undergoing mitotic division. One of the two cells undergoes a mutation (denoted by the star), in which the progenitor 14 allele loses one tetrameric repeat unit to create a mutant 13 allele. Assuming the relative proportions of the four daughter cells are maintained throughout the subsequent divisions, then after genotyping, the relative intensities of the wild type 11 and 14 alleles to the mutant 13 allele are indicated schematically on the right.*

was only just detectable (10:9:1). Furthermore, in two cases where hair and blood samples were available for analysis, the Type 1 pattern was tissue dependent as the pattern was confined only to the hair sample. We also discovered that the root sheath cells of individual anagen head hairs could exhibit this effect, while other hairs plucked from the scalp at the same time were unaffected (data not shown). This probably can be attributed to the fact that head hairs develop clonally from a relatively small number of progenitor cells in the follicle bud (13). In two other cases, where both buccal and blood samples were available for analysis, both types of samples displayed the same triallelic pattern, thereby showing that more than one cell lineage can be affected in a single individual. Finally, we encountered a case in which a man and eight members of his immediate family had been tested. A blood sample from the man exhibited a Type 1 pattern at the D18S51 locus (see the upper panel of Fig. 3). This also may have been the situation in his sperm cells because it was observed that, while the triallelic pattern itself was not a directly heritable trait, all three of his D18S51 alleles were apparently transmitted to offspring.

Given the Type 1 patterns were attributable to somatic mutation, the magnitude of each mutational change could be inferred directly from the pattern of alleles. This inference relied on an extrapolation from cases where data were available from either an unaffected cell lineage or from both biological parents. In these situations, it was possible to establish the identity of the two 'wild type' alleles. Consequently, the identity of the mutant allele could be deduced. From such cases it was observed that, as predicted from the model in Fig. 2, the mutated allele was always one of the two minor alleles and never the major allele. Therefore, it was concluded that the size difference between the two minor alleles would reflect the magnitude of the mutational change. Analysis of the data for all the Type 1 patterns revealed that all changes were multiples of four nucleotides. By analogy with germline mutations, this suggests that

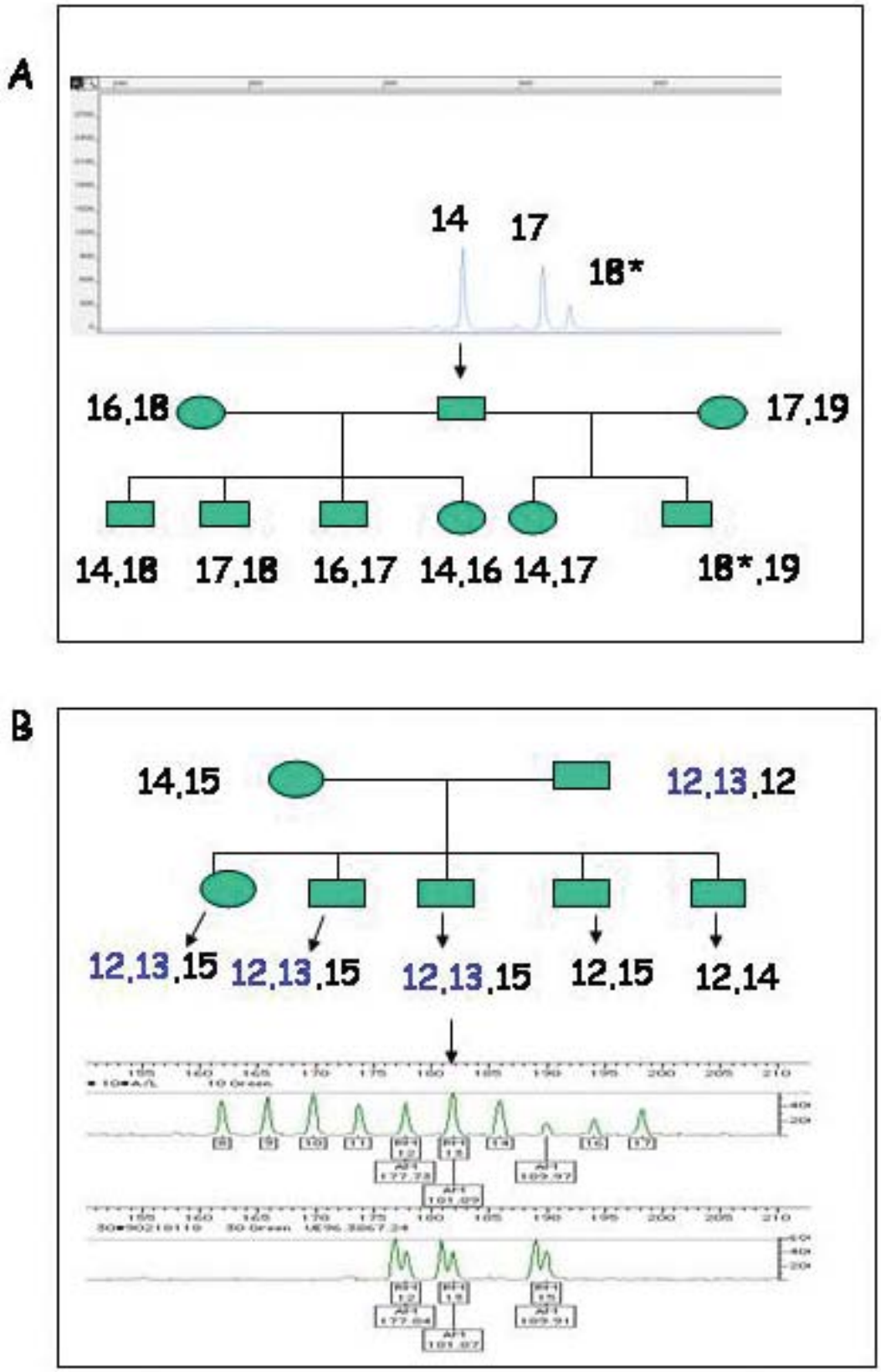

FIG. 3—*Upper panel (A) – A Type 1 triallelic pattern at the D18S51 locus in the FSS SGM kit. The putative mutant allele is denoted by*. The male concerned had six children by two different mothers shown diagrammatically as a family tree. Samples were available from each family member and their genotypes at the D18S51 locus are indicated. Note that three different D18S51 alleles (14, 17, and 18*) are represented amongst the man's six offspring. Lower panel (B) – A Type 2 triallelic pattern at the D8S1179 locus in the FSS SGM kit. The family concerned comprised a mother, father, and five offspring shown diagrammatically as a family tree. From this it can be seen that three of five offspring from the family inherit a triallelic pattern. Beneath the familial tree, lane 30 shows the triallelic pattern obtained from DNA from a buccal scrape from one of five offspring. Lane 10 is an allelic ladder from the same gel run. The 'M' shape of the alleles in lane 30 is due to pronounced 'N' ('A') banding at the D8S1179 locus in the FSS SGM kit.*

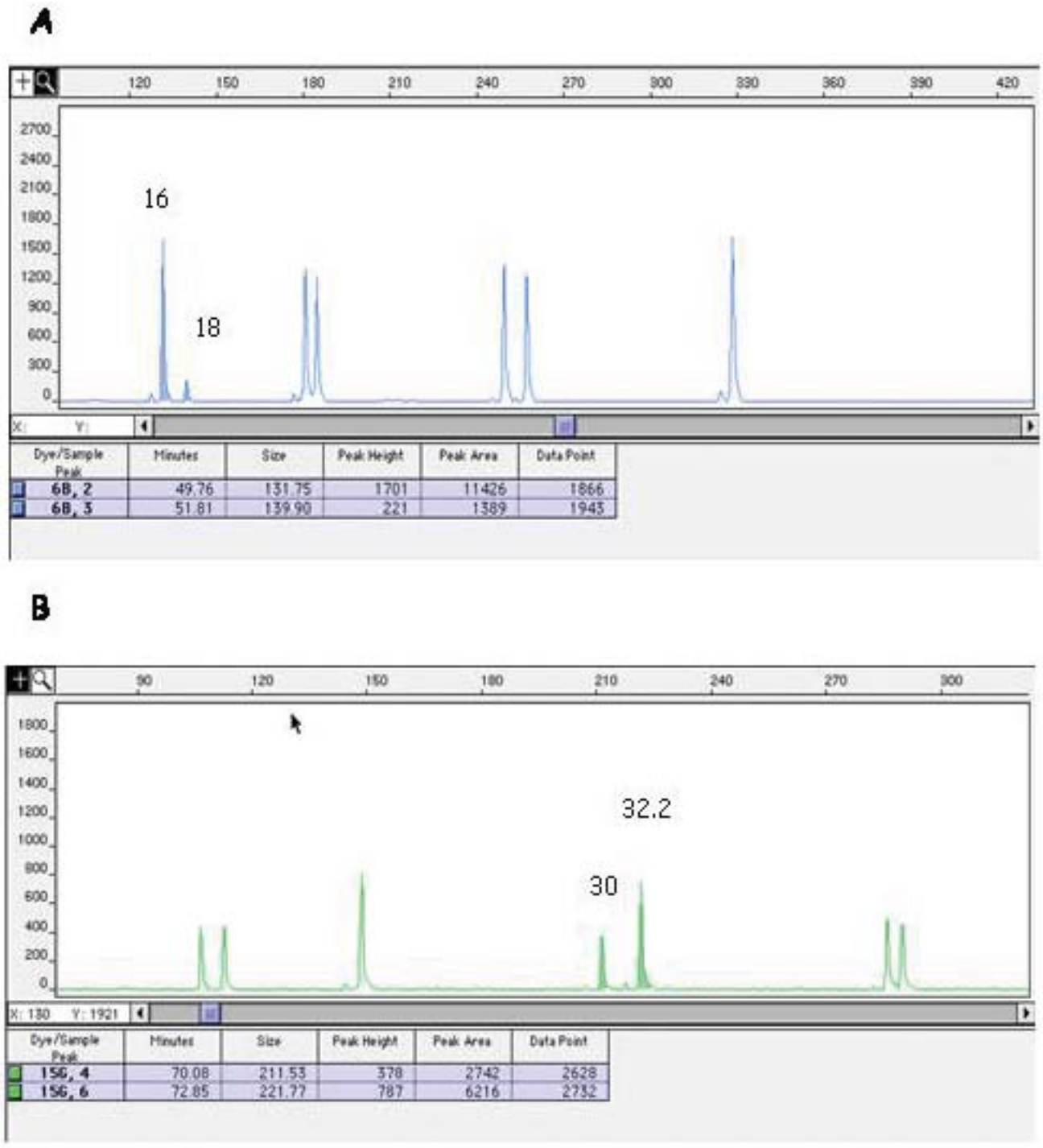

FIG. 4—*Upper panel (A) – D3S1358 locus in the AMPFlSTR® SGM Plus™ kit showing an aberrant diallelic pattern. Peak area information is supplied in column 5 of the table beneath. Note that the 18 allele is about 12% of the 16 allele. Lower panel (B) – D21S11 locus in the AMPFlSTR® SGM Plus™ kit showing an aberrant diallelic pattern. Peak area information is supplied in column 5 of the table beneath. Note that the 30 allele is about 44% of the 32.2 allele.*

each mutational event could be attributed to the loss or gain of one or more of the tetrameric repeat units (14). The majority (89.8%) of the mutational events appeared to involve the loss or gain of a single repeat unit.

Are Type 2 Patterns the Result of Chromosomal Rearrangement?

It was suspected that the Type 2 patterns might reflect the presence of three copies of the STR allele in an individual cell and could be due, for instance, to some type of chromosomal rearrangement at a heterozygous locus. One explanation for this effect could be autosomal aneuploidy (e.g., because of trisomy or a gross chromosomal translocation) involving that chromosome (15). Alternatively, a more restricted localized event could cause a duplication of a smaller region of the chromosome containing the STR locus concerned. However, in order to generate a triallelic pattern, this duplicative event also would have required some kind of slippage or recombination event to generate a distinct third allele. Without this

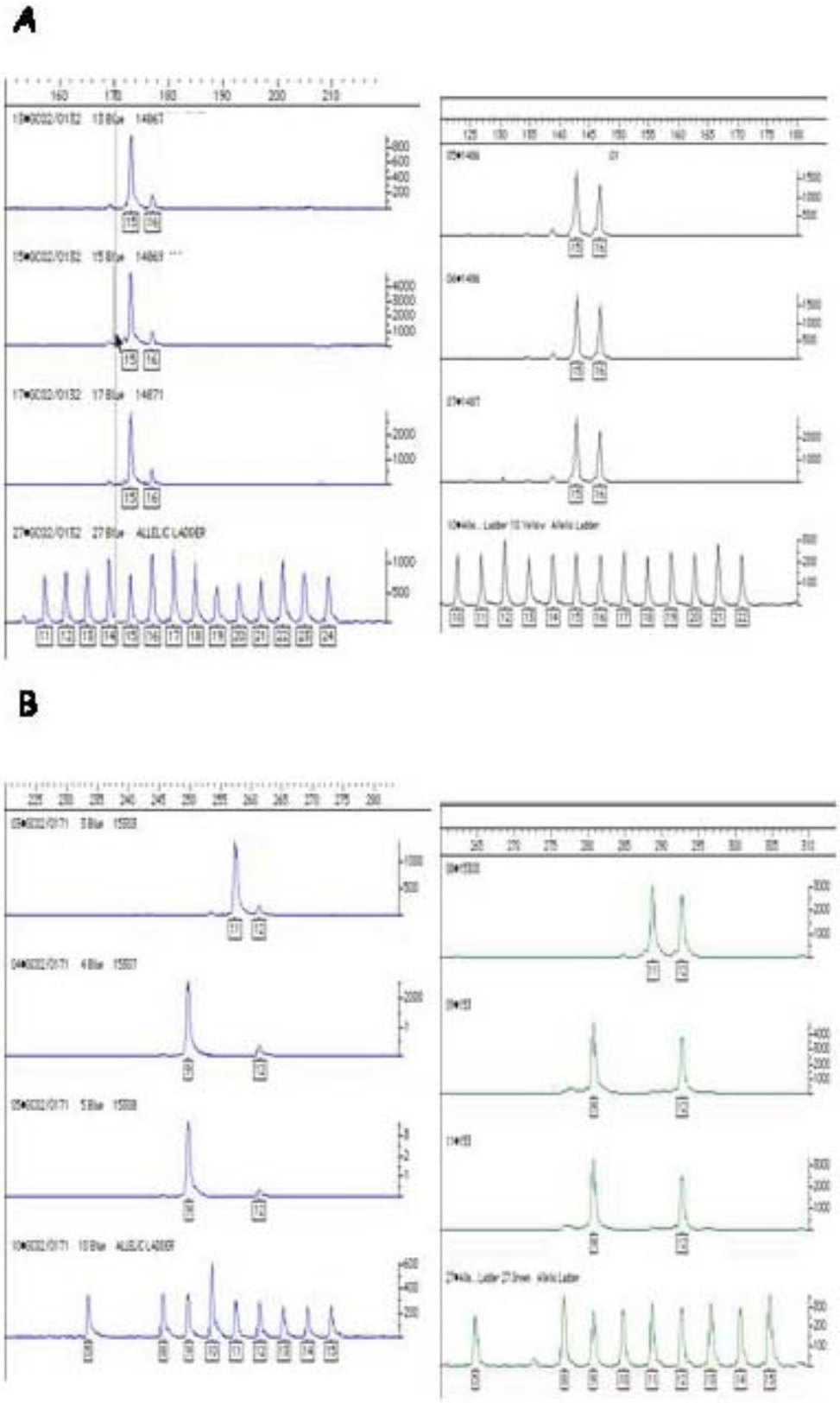

FIG. 5—*Upper panel (A) – vWFA31/A locus in the AMPFlSTR® SGM Plus™ kit. The left box shows the results obtained from the buccal scrapes of three siblings (Lanes 13, 15, 17) and the corresponding allelic ladder (Lane 27). Note the small peak area of the 16 allele in each sibling compared to its partner allele. The right box shows the results obtained from the same DNA after amplification with the Powerplex™ 16 kit (Lanes 5, 6, 7) and the corresponding allelic ladder (Lane 10). Note the restoration of the balanced signal after amplification with Powerplex™ 16 kit. Lower panel (B) D16S539 locus in the AMPFlSTR® SGM Plus™ kit. The left box shows the results obtained from the buccal scrapes of three siblings (Lanes 3, 4, 5) and the corresponding allelic ladder (Lane 10). Note the small peak area of the 12 allele in each sibling compared to its partner allele. The right box shows the results obtained from the same DNA after amplification with the Powerplex™ 16 kit (Lanes 8, 9, 11) and the corresponding allelic ladder (Lane 27). Note the restoration of the balanced signal after amplification with Powerplex™ 16 kit.*

second event, the locus may present as an imbalanced and therefore aberrant, diallelic pattern (15).

Autosomal trisomy is caused by non-disjunction of chromosomes at meiosis. With the exception of chromosome 21, most are lethal to the developing zygote (16). Moreover, in general, most gross chromosomal rearrangements lead to severe developmental defects and medical complications (16). As the individuals tested were apparently healthy adults involved in criminal offenses, and since most events observed did not involve chromosome 21, gross chromosomal rearrangement would not seem to be a likely explanation for the Type 2 patterns. A better proposition would therefore be that these patterns are due to a more localized duplication event. If this was the explanation, the two of the three alleles should be tightly linked and therefore, if an individual's germ cells were affected, the two alleles should be transmitted together to offspring. Anecdotal support for this proposition was obtained from a case in which five members of the same family were found to have inherited the pattern (see the lower panel of Fig. 3). Interestingly, although the biological father did not exhibit a triallelic pattern, he did have an aberrant diallelic pattern in that the signal intensity of his 12 allele was twice that of the 13 allele.

Investigation of Diallelic Patterns

Approximately 400 aberrant diallelic patterns were available for study. By analogy with the explanation for Type 1 triallelic patterns, the diallelic patterns could be the result of somatic mutation at a homozygous locus. Again, in support of this, it was observed that the minor alleles exhibited variation between individuals such that the peak area differences ranged from a maximum 50% (10:5) through to those just detectable at about 10% of the intensity of the larger allele (10:1). See for example the samples shown in Fig. 4. It is possible that some of the 10:5 diallelic patterns encountered may be the result of chromosomal rearrangement rather than somatic mutation (such as the one seen in the lower panel of Fig. 4). Three cases were encountered in which the D21S11 locus showed a 10:5 diallelic pattern, and these could have been due to aneuploidy (15). It is also possible that some of the diallelic patterns could be due to mutations in the primer binding sequences, leading to a reduction in amplification efficiency of one of the alleles. To demonstrate this latter possibility, two cases were reanalyzed using the Promega Powerplex™ 16 multiplex kit (Madison, WI) that contained alternative primer sets for the two loci of interest (vWFA31/A and D16S539). These particular cases were selected because three siblings from the families concerned exhibited the same aberrant diallelic pattern, suggesting that the trait was heritable and not, therefore, attributable to somatic mutation. In both instances, the imbalanced signal could be restored to a balanced heterozygote signal, thus indicating the effect was entirely due to primer binding (see Fig. 5). It is accepted that the number of cases tested in this way represents only a subset of all the diallelic patterns collected, and that they were targeted specifically because of the familial information. Nevertheless, the experiment does show that a proportion of the anomalous, diallelic patterns are due to primer binding site mutations.

Discussion and Conclusions

Of the six loci for which most data were available, it appears that there is considerable variation between loci in the propensity for each type of rearrangement (see Table 1). One locus (THO1) appears to be particularly stable in any type of rearrangement considered. Crouse et al. (17) noted a similar effect when genotyping 10 000 individuals at three STR loci (CFS1PO, TH01 and TPOX). Of the 19 (apparently Type 2) triallelic patterns they encountered, 18 were at the TPOX locus, and none occurred at the TH01 locus. It should be noted however, that if the Type 2 pattern is a heritable trait, then since in most cases it was not possible to check whether the affected individuals were related, it could have influenced the distribution. From the evidence presented above, we conclude that the Type 1 triallelic patterns are the result of mutation during mitosis leading to mosaicism. Similar conclusions have been reached independently by Rolf et al. (18), who encountered two paternity cases involving triallelic patterns. Published data for germline mutation (10) suggest that additions or loss of a single repeat unit is the most prevalent type of meiotic event. The data for mitotic events presented here confirm this propensity toward single-step, rather than multiple-step, repeat unit changes. With regard to the Type 2 patterns, Zamir et al. (19) have recently reported on a case involving a triallelic pattern (apparently Type 2) at the D16S539 locus. They demonstrate a localized effect in that other markers on chromosome 16 do not appear to show any unusual effects. It is concluded therefore that the most likely genetic cause of these patterns is a localized duplication event in the region of the chromosome in which the STR locus is found. The data also demonstrate that aberrant diallelic patterns could have at least three potential causes and that visually the diallelic pattern itself offers no information as to its cause. Thus we conclude that it is not possible to infer an underlying genetic basis from the aberrant diallelic pattern alone.

Acknowledgments

We gratefully acknowledge the technical assistance from DNA analysts, past and present, plus many individuals within the FSS Research and Development Department.

References

1. Werrett DJ. The national DNA database. Forensic Sci Int 1997;88:33–42.
2. Sparkes R, Kimpton CP, Gilbard S, Carne P, Andersen J, Oldroyd N, et al. The validation of a 7-locus multiplex STR test for use in forensic casework. Int J Leg Med 1996;109:195–204.
3. Cotton EA, Allsop RF, Guest JL, Frazier RRE, Kuomi P, Callow IP, et al. Validation of the AMPF/STR® SGM Plus™ system for use in forensic casework. Forensic Sci Int 2000;112:151–61. [PubMed]
4. Gill PD, Sparkes RL, Kimpton CP. Development of guidelines to designate alleles using an STR multiplex system. Forensic Sci Int 1997;89:185–97. [PubMed]
5. Whitaker JP, Clayton TM, Urquhart AJ, Millican ES, Downes TJ, Gill PD, et al. STR typing of bodies from a mass disaster: High success rate and characteristic amplification patterns in highly degraded samples. BioTechniques 1995;18(4):670–7. [PubMed]
6. Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterisation of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acid Res 1996;24:2807–12. [PubMed]
7. Clark JM. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acid Res 1998;16:9677–86.
8. Clayton TM, Whitaker JP, Sparkes R, Gill PD. Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int 1998;91:55–70. [PubMed]
9. Walsh PS, Metzgar DA, Higuchi R. Chelex 100 as a medium for the simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 1991;1:91–8.
10. Walsh PS, Valaro J, Reynolds R. A rapid chemiluminescent method for the quantitation of human DNA. Nucleic Acid Res 1992;20;5061–4. [PubMed]
11. DNA recommendations: 1994 Report concerning further recommendations of the DNA Commission of the International Society of Forensic Haemogenetics regarding PCR based polymorphisms in STR systems. Int J Leg Med 1994;107:159–60.

12. Urquhart AJ, Kimpton CP, Downes TJ, Gill PD. Variation in STR sequences – a survey of twelve microsatellite loci for use as forensic identification markers. Int J Leg Med 1994;107:13–20.
13. Linch CA, Whiting DA, Holland MM. Human hair histogenesis for the [PubMed] mitochondrial DNA forensic scientist. J Forensic Sci 2001;46(4):844–53.
14. Brinkmann B, Klintshar M, Neuhuber F, Huhne J, Burkhard R. Mutation rate in human microsatellites: influence of the structure and length of the [PubMed] tandem repeat. Am J Hum Genet 1998;62:1408–15.
15. Pertl B, Weitgasser U, Kopp S, Kroisel PM, Sherlock J, Adinolfi M. Rapid detection of trisomies 21 and 18 by quantitative fluorescent multiplex [PubMed] PCR. Hum Genet 1996;98:55–9.
16. MW Strickberger. Genetics. New York: Macmillan, 1968.
17. Crouse CA, Rogers S, Amiott E, Gibson S, Masibay A. Analysis and Interpretation of Short Tandem Repeat Microvariants and three-banded allele patterns using multiple allele detection systems. J Forensic Sci [PubMed] 1999;44:87–94.
18. Rolf B, Wiegand P, Brinkmann B. Somatic mutations at STR loci – a reason for three-allele pattern and mosaicism. Forensic Sci Int 2002; 126:200–2. [PubMed]
19. Zamir A, Shpitzen M, Oz C, Motro U, Meiner V, Gafny R. Presentation of a three-banded allele pattern – analysis and interpretation. J Forensic Sci. 2002;47:824–6. [PubMed]

Additional information and reprint requests:
Tim Clayton, Ph.D.
Senior Forensic Scientist
Forensic Science Service
Sandbeck Way
Wetherby
West Yorkshire, LS22 7DN
UK
E-mail: tim.clayton@fss.pnn.police.uk