A Promising New STR Marker System for DNA Fingerprinting

Jian Qing Tang, Mireille Laframboise, Marie-Claude Morse and Serge B. Melançon
Procrea BioSciences Inc. 1100 Beaumont Ave, Suite 305, Mount-Royal, Quebec, Canada H3P 3H5

SUMMARY

Seven newly discovered DNA markers, including 4 simple and 3 complex highly polymorphic STR markers, were combined into a multiple marker system intended for human identification. Protocols were designed to allow for simultaneous multiplex amplification and analysis of predetermined combinations of markers using an automatic DNA analyzer or a standard high resolution PAGE method as an alternative. With either procedure, the markers displayed a high power of discrimination and altogether reached a matching probability level of 1 in $8x10^{10}$, in the upper range of available marker systems. Using the automatic DNA analyzer procedure, results of a routine genetic identity test may be released within 24 hours and reach a typical paternity probability of 99.988%.

INTRODUCTION

Polymorphisms at the DNA level convey unequivocal information about the segregation pattern of parental chromosomes following the mating process and best disclose a person's biological origin and genetic identity. With the advent of the polymerase chain reaction (PCR) technique in the late eighties, a new generation of polymorphic DNA markers, the short tandem repeats (STR), became more easily accessible. First used to create maps of the human genome and to help locate genes, STR markers were rapidly adopted by forensic investigation laboratories because they offered significant advantages in sensitivity and speed of analysis over existing techniques (1). Most STR markers being used today consist of the *simple* one-variable repeat unit type. *Complex* STR markers are made of hypervariable repeat units and carry a much greater informativeness then simple STR. When pooled and amplified together in one PCR reaction, multiple markers of both STR types provide an optimized power of discrimination. An increasing number of such multiplex systems have been designed and are currently being validated as forensic investigation tools and for genetic identity testing as well (2, 3). The likelihood ratio of each individual system is expected to reach levels of $1X10^8 – 1X10^9$. Owing to the limited number of available simple STR markers with high heterozigosity and taking into account the need for internationally standardized databanks, most of these multiplex systems share the same 8 to 12 markers with one or two of them differing between systems.

We perceived the need to develop a distinct, original and highly discriminative STR marker system that could be offered as an extension of or as an alternative to currently available DNA forensic/paternity testing systems. In addition, this new performant DNA profiling system could serve in situations where an independent counter-expertise is needed.

The 3'ends of *Alu* elements are highly polymorphic and their use as a potential source of new DNA markers was considered. The polymorphism in these STR regions consists of mostly trimeric repeats with some tetrameric and occasional pentameric motifs. Using the Alumorphs approach, we have established sequence libraries from a collection of 3'ends of Alu elements. A pool of oligonucleotides with trimeric, tetrameric and pentameric repeats was synthesized and hybridized to the libraries. Positive colonies were selected and sequenced. Primers were designed from the repeat's flanking sequences and used to assess the polymorphic characteristics of the amplified fragments.

In addition to *Sextolets* 100 and 900 that had been previously reported (4), five new *Sextolet* markers, 110, 150, 180, 700 and 800 were identified and further studied. Using PCR amplification of diverse combinations of these seven markers, we designed a DNA profiling system suitable for different levels of laboratory equipment. A conventional radiolabelled multiplex PCR procedure can be used in routine diagnostic laboratories where a DNA sequence analyzer is not available. A fluorescent-dye labelled multiplex PCR procedure was also set up to improve the convenience of manipulation, allow for automation and considerably accelerate single sample output. Data from a limited population study have showed a matching probability in the 8×10^{10} scale and a typical paternity index of over 8000.

ANALYTICAL PROCEDURES

We have used two methods to amplify and resolve the seven *Sextolet* markers. The first method relies on radiolabeled primers in a (hot) multiplex PCR amplification followed by a standard high-resolution polyacrylamide gel electrophoresis (PAGE). The second method

uses fluorescent primers and the alleles are separated and quantified using capillary electrophoresis.

Hot multiplex PCR

Multiplex PCR The reaction was carried out in a 200µl thin wall tube in a Robocycler (Stratagene). The reaction mixture included 20ng of genomic DNA extracted from blood samples according to conventional methods or from dry blood spots (DBS). One unit of *Taq* DNA polymerase was combined with 12µl of PreMix A for the PENTAPlex, the QUATROPlex A or the QUATROPlex B and 4µl of labeled PreMix B for the corresponding multiplex sets. The final volume was adjusted to 20µl. PreMix B was labeled with ^{32}P using a T4 Kinase reaction. 16µl of PreMix B will generate 50-60µl of labeled mix suitable for 15 reactions. The PCR amplification was carried out with 30 cycles at 94°C for 45sec., 52°C for 60sec., and 72°C for 60sec.

Gel electrophoresis To the 20µl of post PCR reaction mixture, 10µl of loading buffer (95%formamide, 0.04% Bromophenol blue, 0.04% Xylene cyanol and 5mM EDTA) was added. The sample was heated at 94°C for 5 min. and immediately cooled to 4-6°C. Five to six µl of the sample was loaded into a sequencing type gel (6% polyacrylamide:bis-acrylamide 19:1, 8M urea and 0.5x TBE). The gel was preheated and run under 1x TBE (90mM Tris-borate pH 8.3, 2mM EDTA) at constant power of 70W for 2-3 hrs. The gel was dried and autoradiographed overnight at room temperature.

Marker sets for hot multiplex PCR Due to the two dimensional illustration capacity of autoradiography, blank spaces should be available between markers to avoid overlapping between allelic fragments. Here, the first dimension is the migration interval while the second dimension is the density of the bands. Three sets of multiplex were designed to amplify all seven markers. The best results were obtained with the following primer sets (table 1)

Table 1. Marker Sets for hot Multiplex PCR

Sextolets Markers	Labeling	PENTAPlex	QUATROPlex A	QUATROPlex B
900	^{32}P	+	+	
800	^{32}P			+
700	^{32}P		+	
180	^{32}P	+		+
150	^{32}P	+		
110	^{32}P	+	+	+
100	^{32}P	+	+	+

Cold multiplex PCR

Multiplex PCR Two µl of DNA sample equivalent to about 4ng DNA or more were added to a PCR reaction tube. To this DNA sample was added 18µl of a premixed M1, M2 or M3 cocktail containing 1 unit *Taq* Polymerase and 0.1unit Pwo Polymerase. The reaction was carried out with 35 cycles at 94°C for 45 sec., 52°C for 60 sec., and 72°C for 60 sec.

Gel electrophoresis. Two µl of each multiplex PCR were mixed with 20.5µl of 100% formamide containing 0.5µl of GS500 (available from PE-ABI, Cat. No. 401733). The sample was mixed and heated to 94°C for 4 min. and then cooled to 4-6°C. Samples were loaded onto an ABI 310 Genetic Analyzer. Electrophoresis was performed with POP6 (PE-ABI) using a 47cm capillary column. The migration module for POP6 is set on filter C and other parameters are as follow:

Table 2. Capillary Electrophoresis Analytical Parameters

Injection Time	5 sec.
Injection Voltage	15.0 Kv.
Collection Time	50 min.
EP Voltage	15.0 Kv.
Heatplate Temperature	50 °C
Syringe Pump Time	120 sec.
Pre-Injection EP	300 sec

Marker sets for cold multiplex PCR Due to the three dimensional illustration capacity of ABI 310, overlapping allelic fragments can be distinguished from each other. Here, the first dimension is the migration interval; the second dimension is the peak height of each allelic fragment and the third dimension is the colour of each specific marker. In principal, many markers can be amplified with one multiplex PCR and analyzed during a single run. Since the *Sextolets* 150 and 700 cannot be included in the same PCR mixture, two separate amplifications are needed to amplify all seven markers (M1 and M2). How-

ever, to amplify all six markers without *Sextolet*150 (M3), a single multiplex amplification is required.

We used three fluorescent dyes to distinguish between all seven individual markers; green (the fluorescent dye is TET), blue (6-FAM) and yellow (HEX), The red colour (TAMRA) available for ABI 310 is used for the reference size standard, GENESCAN 500. Each markers allelic fragment is identified according to its size.

Table 3. Marker Sets for the cold Multiplex PCR

Sextolet Markers	Labeling	M1	M2	M3
900	6-FAM	+		+
800	HEX		+	+
700	TET		+	+
180	TET	+		+
150	6-FAM	+		
110	TET	+		+
100	6-FAM	+		+

RESULTS AND DISCUSSION

In previous studies we were able to demonstrate that multiple STR markers can be amplified together using single primer pairs (4). However, the method usually generated a high background with incomplete resolution of all individual markers The strategy proposed here uses a shared 5' primer in combination with individual 3' primers. Each 3' primer is designed independently so the size of the corresponding marker can be adjusted and the background amplification can be reduced to the minimum. Seven markers were accordingly designed. (table 4).

Table 4. Characteristics of the Sextolet Markers.

Sextolet Markers	H	STR Motif	Chrom. Location	Base pair	Alleles
100	0.765	AAT	5q31.2-33.2	88 to 118	10
110	0.768	AAT	7p12.3-13	127 to 148	8
150	0.662	AAAAC	1q11.2-1q12	155 to 180	6
180	0.729	ACAT	2q14.1-14.3	152 to 196	10
700	0.925	(AAAG)(GAAG)(AAAG)*	5q22.2-23.1	350 to 380	>40
800	0.919	(AAAG)(AG)(AAAG)*	10p11.23-12.2	180 to 328	>42
900	0.967	(AAAG)(AG)(AAAGG)	17q11.2-21.2	233 to 408	> 129

H: Heterozygosity
*: only one allele has been sequenced

As shown above, we determined from the allele segregation data using 8 large CEPH families, the chromosomal location of the seven *Sextolet* markers. Five of them were localized on different chromosomes, with the exception of *Sextolets* 100 and 700, 26 cM apart on chromosome 5.

Sextolets 100 and 900 had been originally described as Q120 and Q900, respectively (4). However, new primer sets and procedure were used to resolve *Sextolet* 100 and *Sextolet* 900 in the present study. A more efficient amplification, a lower background and appropriate fragment sizes were obtained with the current methodology.

Sextolet 110, *Sextolet* 150, *Sextolet* 180, *Sextolet* 700 and *Sextolet* 800 are newly discovered STR markers. They were either cloned from a marker enriched library or from simultaneously amplified multiple markers. After cloning, the markers were entirely sequenced. Simple STR motifs with trimeric, tetra- or pentameric units were found in *Sextolet* 100 and *Sextolet* 110, *Sextolet* 180 and *Sextolet* 150, respectively. Results are analyzed using a standard PAGE migration with a resolution of 3 bp.

Complex STR motifs, consisting of an AG-rich stretch of DNA (about 200 bp) were found in *Sextolet* 700, *Sextolet* 800 and *Sextolet* 900. Their STR were made of hypervariable repeat units with motifs that may vary independently between individuals. The hypervariable repeat units result in one base pair differences between alleles for *Sextolet* 900 and *Sextolet* 700; two base pairs for *Sextolet* 800. Although a high resolution sequencing gel may be able to resolve 1 base pair differences, it is hard to tell the alleles size between lanes or between runs. In this case a bin frequency strategy is used by which a bundle of alleles are grouped and called together. However in order to achieve a precise allelic

frequency calling, an improved separation technique is needed with one base pair resolution in the 200 to 400 bp range and an alleles size reference scale present in each run.

The ABI 310 Genetic Analyzer uses a capillary column to analyze DNA profiles. Two types of polymers are proposed. The POP6 (Performance Optimized Polymer 6) for DNA sequencing applications and POP4 for fragment analysis. We observed poor resolution of *Sextolet* 900, *Sextolet* 800 and *Sextolet* 700 using POP4. Alleles differing by one base pair came out as a single wider peak instead of two separate peaks in the high molecular weight range. A new module was designed using POP6 (see below). Using this new module, one base pair resolutions at 400 bp could be achieved, allowing for complete analysis of *Sextolet* 900, *Sextolet* 800 and *Sextolet* 700. We have been able to perform over a thousand runs using this module, without any significant loss in resolution and quality of the detection signal.

Multiplex PCR amplification procedures were designed to accelerate the analysis procedures of the *Sextolet* markers. Diverse combinations of the seven markers were tested in order to reach optimal amplification conditions. With the exception of *Sextolet* 150 which can not be included in the multiplex PCR with *Sextolet* 700, all other associations were operational. Therefore, to amplify all seven markers, two PCR amplifications are needed. A single multiplex PCR is required however for the six markers without *Sextolet* 150.

Radio-labelled multiplex PCR products were resolved using a high resolution PAGE sequencing gel. Results were analyzed after autoradiography of the dry gel (Fig 1). Each marker's position on the gel was set according to the alleles sizes observed following a population. Allelic fragment sizes and predetermined positions allow for an easy identification as the markers were designed to cluster in definite areas of the gel. In very rare instances, allelic fragments of adjacent markers may overlap. In these cases, a simplex PCR amplification and electrophoresis are then needed to resolve the individual markers.

Fluorescent multiplex PCR profiles were analyzed using an ABI 310 Genetic Analyzer. Examples of typical *Sextolet* profiles are shown in figures 2 and 3. The ABI 310 has the advantage over gel electrophoresis apparatus of being highly automated. Factors that may adversely affect gel electrophoresis performance such as temperature, gel filling, sample loading and power level are computer-controlled under the ABI 310 operating conditions. Results of the STR profiles are scored using the computer program GeneScan 2.1(PE-ABI). With the use of different fluorescent dyes, the different markers were easy to locate and identify. A reference size scale was included in each migration. The size of each allelic fragment was obtained. Both size and colour data were combined to help identify each marker's allele.

Table 5. Multiplex *Sextolet* Marker Systems

Multiplex PCR	Markers	Labelling	PM	Typical PI
PENTAPlex	100, 110, 150, 180, 900	^{32}P	$3.8x10^6$	230
QUATROPlex A	100, 110, 700, 900	^{32}P	$5.4x10^6$	448
QUATROPlex B	100, 110, 180, 800	^{32}P	$4.2x10^4$	45
M1	100, 110, 150, 180, 900	Fluorescent	$3.8x10^6$	230
M2	700, 800	Fluorescent	$3.0x10^3$	29
M1+M3	100, 110, 150, 180, 700, 800, 900	Fluorescent	$8.03x10^{10}$	8500
M3	100, 110, 180, 700, 800, 900	Fluorescent	$1.9x10^9$	4297

REFERENCES

1. Committee on DNA Forensic Science: An Update: in The Evaluation of Forensic DNA Evidence, 1996 National Academy press, Washington D.C. Chapter 4: Population Genetics pp89-123.

2. Schumm J.W., Lins A.M., Micka K.A. , Sprecher C.J., Rabbach D., Bacher J. Automated fluorescent detection of STR multiplexes-development of the GenePrint™ PowerPlex™ and FFFL multiplexes for forensic and paternity applications. In : Proceedings from the Seventh International Symposium on Human Identification; 1996; Promega Corporation, pp70-88.

3. Walsh P.S., Lazaruk K., Wallin J., Robertson J. AmpFISTR™: Complete systems for amplification and typing of STR loci. In: Proceedings from the Seventh International Symposium on Human Identification; 1996; Promega Corporation, pp68.

4. Tang J.Q., Korab-Laskowska M., Jarnik M. Cardinal G., Vanasse M., Melançon S.B., Labuda, D. Alu-PCR combined with non-Alu primers reveals multiple polymorphic loci; 1995; *Mammalian Genome* **6**, 345-349.

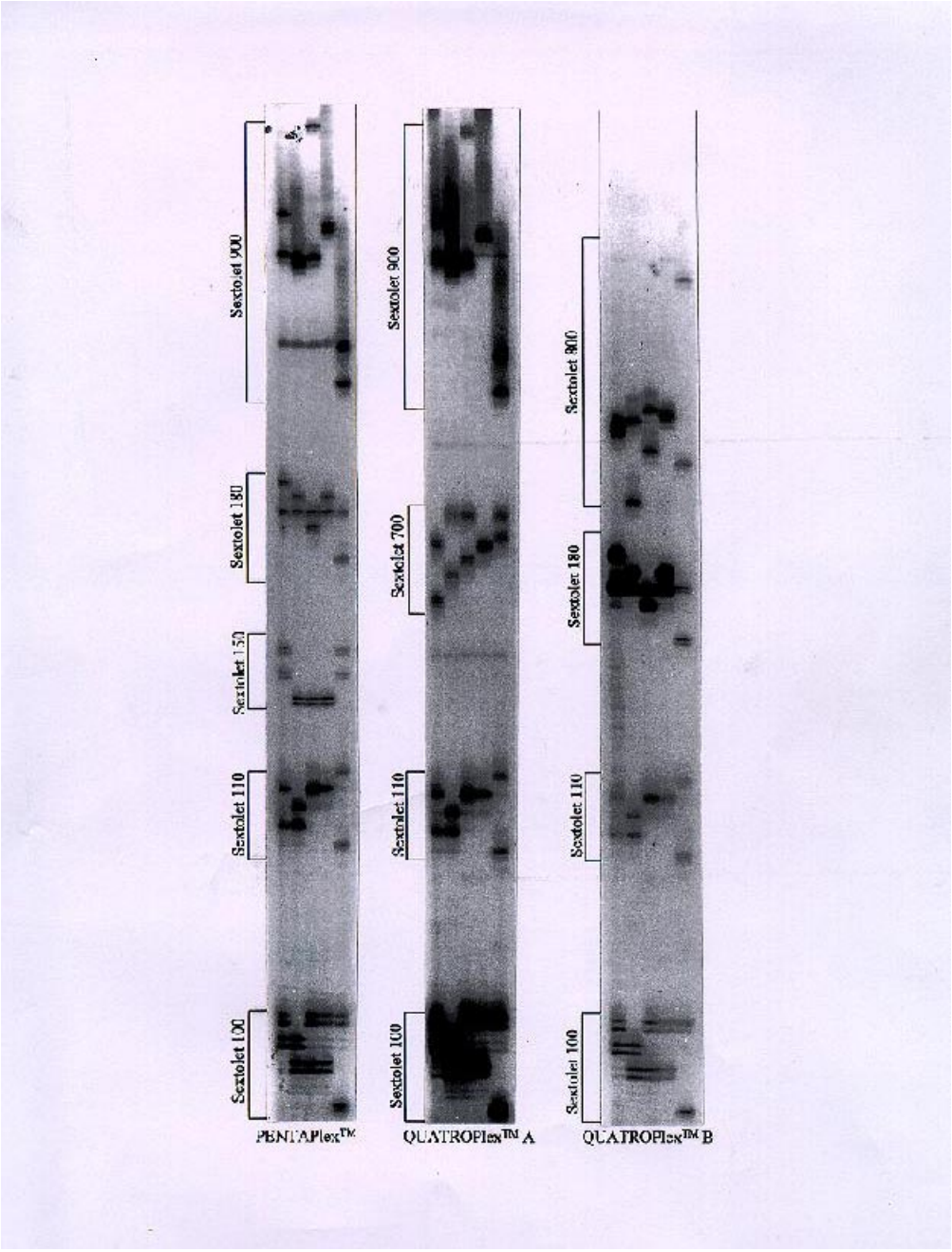

Figure 1. Multiplex PCR for PENATPlex, QUATROPlex A and QUATROPlex B

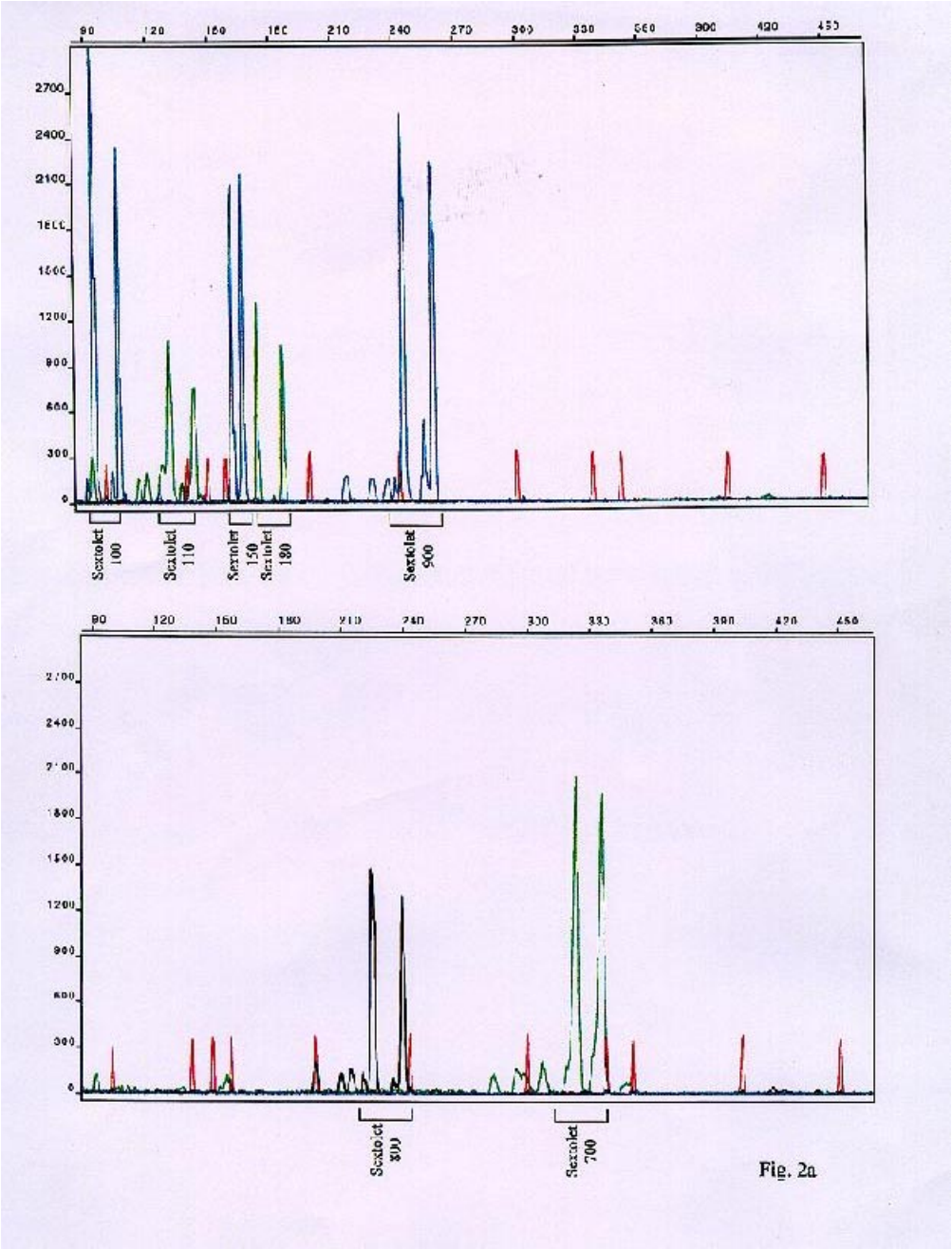

Figure 2a. Multiplex PCR fluorescent detection of Sextolet markers. 2a, upper: M1; lower: M2.

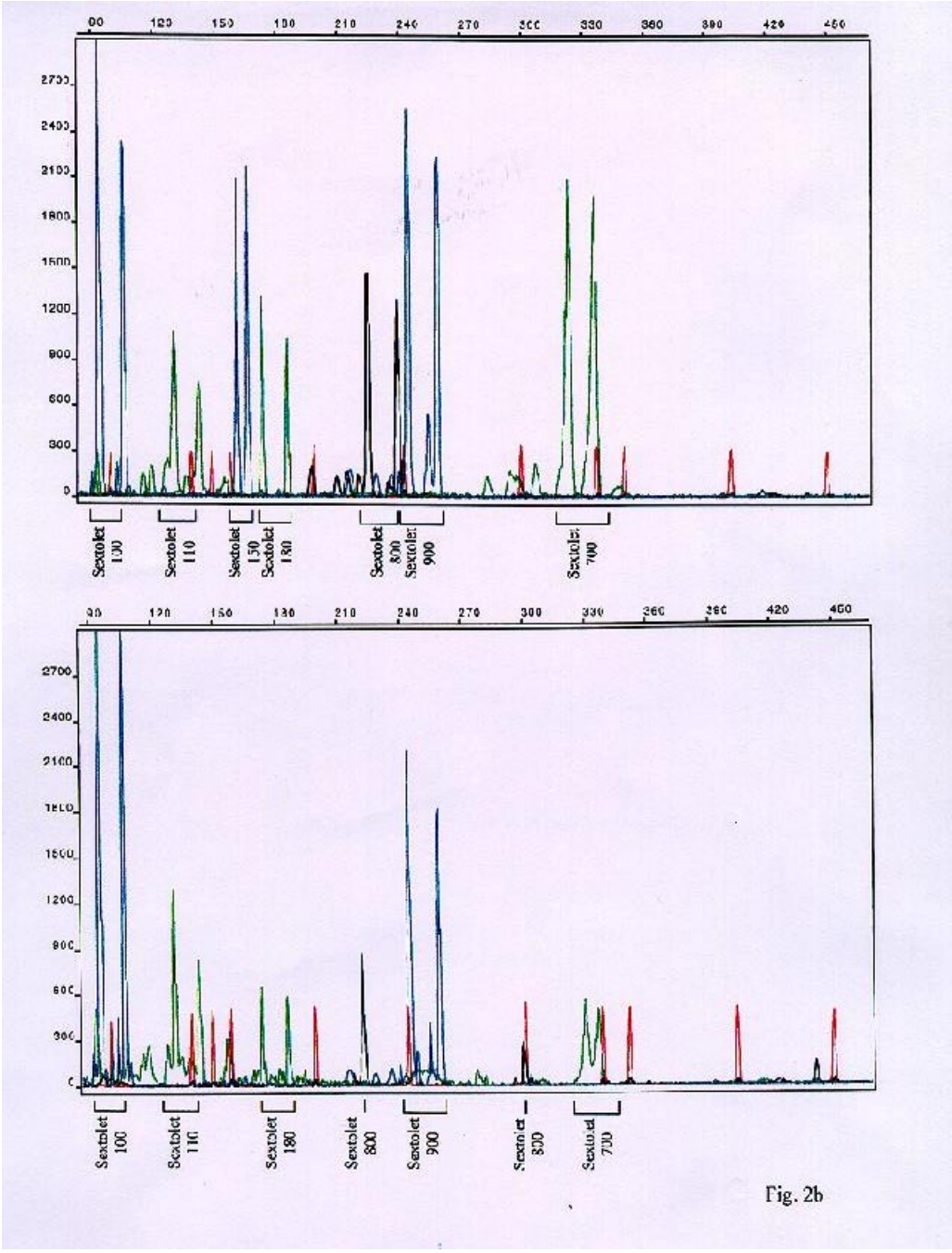

Figure 2b. Multiplex PCR fluorescent detection of Sextolet markers. 2b, upper: M1+M2; lower: M3.