The Evolution of Short Tandem Repeat (STR) Multiplex Systems

Ann M. Lins, Cynthia J. Sprecher, Katherine A. Micka, Dawn R. Rabbach, Jennifer A. Taylor, Jeffrey W. Bacher, Laura F. Hennes, Allan Tereba, Trent Gu, Elizabeth A. Amiott, Nadine A. Nassif and James W. Schumm
Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711 USA.

INTRODUCTION

In the past five years, polymorphic short tandem repeat (STR) loci have been developed into one of the most rapid, efficient, and precise methods for human identification (1-6). The amplification of these STR loci has progressed from single locus reactions to megaplex reactions, amplifying as many as sixteen loci in one reaction tube. In parallel, the detection of STR loci has advanced from manual radioactive and silver stain detection to automated detection using fluorescent instrumentation.

We describe the evolution of STR multiplex systems within our laboratory for both manual and automated detection formats. The loci CSF1PO, TPOX, TH01, vWA, D16S539, D7S820, D13S317, D5S818, F13A01, FESFPS, F13B, LPL have been characterized and incorporated into various multiplex amplification sets. Three triplex systems (CTT Multiplex, FFv Multiplex, and SilverSTR™ III Multiplex) have been developed for silver stain detection. Five multiplex systems (CTTv Multiplex, FFFL Multiplex, GammaSTR™ Multiplex, PowerPlex™ 1.1 System, and PowerPlex™ 1.2 System) have been optimized for fluorescent detection. Most recently, we have incorporated additional tetranucleotide loci (FGA, D21S11, D8S1179, D18S51, D3S1358) and new pentanucleotide loci into two megaplex systems, PowerPlex™ 2 and PowerPlex™ 16.

The loci have been well characterized and allelic ladders have been or are being constructed for each locus. Validation studies have been completed for the CTT Multiplex, FFFL Multiplex, and the PowerPlex™ 1.1 System. We have collaborated with other laboratories to determine the genotypes for at least 200 individuals in each of four population groups to calculate characteristic allele frequencies, power of discrimination, power of exclusion, and paternity indices for each locus and multiplex system.

MATERIALS AND METHODS

All materials necessary for amplification except Taq DNA Polymerase were supplied in the appropriate *GenePrint*™ STR Multiplex System (CTT Multiplex, FFv Multiplex, SilverSTR™ III Multiplex, CTTv Multiplex, FFFL Multiplex, GammaSTR™ Multiplex, PowerPlex™ 1.1 System or PowerPlex™ 1.2 System) (Promega, Madison, WI). The amplification and detection protocols are provided in the corresponding *GenePrint*™ STR Systems Technical Manual (Promega, Madison, WI). For the multiplex systems detected by silver stain analysis, the amplified fragments are separated in a 4% denaturing polyacrylamide gel with 7M urea and 0.5X TBE as the running buffer. The gel is temporarily bound to one glass plate to simplify handling during the silver stain procedure of Bassam *et al.* (7). A permanent record may be obtained using Automatic Processor Compatible Film (Promega, Madison, WI). The fragments amplified using the CTTv Multiplex, FFFL Multiplex, GammaSTR™ Multiplex or PowerPlex™ 1.1 System are separated in a 4% denaturing polyacrylamide gel (32cm or 43cm long) containing 7M urea and 0.5X TBE. Following electrophoresis, the gel is scanned using the FMBIO® II Fluorescent Scanner (Hitachi Software Engineering America, Ltd., South San Francisco, CA). Samples amplified using the PowerPlex™ 1.2 System are detected using the ABI PRISM® 377 DNA Sequencer or ABI PRISM® 310 Genetic Analyzer (PE Applied Biosystems, Foster City, CA).

CHARACTERISTICS OF STR LOCI

All twelve STR loci which have been developed (Table 1) into currently available multiplex sets are tetranucleotide repeat loci except for Amelogenin, which is used as a gender identification locus. Each locus is located on a different chromosome except for CSF1PO and D5S818 which are both located on chromosome 5. Each locus has a well defined size range and the common alleles have been identified and sequenced. Few microvariants have been found for this set of STR loci. Two of the common microvariants are the F13A01 allele 3.2 (8) and TH01 allele 9.3 (9). A few additional microvariants, such as TH01 allele 8.3, have been reported but are rare with a frequency of less than 0.001.

The STR multiplexes have been developed to allow shared data across detection formats, including both silver stain and fluorescent detection. Table 2 shows

the loci that are included in the current *GenePrint*™ Multiplex Systems. For silver stain detection, there are three triplexes available (CTT Multiplex, FFv Multiplex, and SilverSTR™ III Multiplex) which include a total of nine loci. Twelve loci are included in three fluorescent multiplexes (CTTv Multiplex, FFFL Multiplex, and GammaSTR™ Multiplex) which utilize one fluorescent dye, fluorescein. The PowerPlex™ 1 System is a two-color multiplex which allows for the co-amplification of eight loci in a single tube. In this system, the loci in the CTTv Multiplex are labeled with carboxy-tetramethylrhodamine (TMR), and the loci contained in the GammaSTR™ Multiplex are labeled with fluorescein.

STR MULTIPLEX SYSTEMS FOR SILVER STAIN DETECTION

Three multiplex systems have been developed for manual detection using a silver staining procedure (7) and are shown in Figure 1. The CTT Multiplex which contains the loci CSF1PO, TPOX, and TH01 is shown in panel A. The FFv Multiplex (F13A01, FESFPS, vWA) is shown in panel B. The third multiplex called the SilverSTR™ III Multiplex contains the loci D16S539, D7S820, and D13S317 and is displayed in panel C. All three multiplex systems may be used in conjunction to obtain genotype data for nine loci using a manual detection format.

Population studies have been completed for all three multiplex sets for three population groups (Table 3) (10). The matching probability for each individual multiplex is not very impressive. However, when all three multiplexes are combined for a total of nine loci, the matching probability exceeds 1 in 1 billion. For paternity testing, the typical paternity index is greater than 500 and the power of exclusion is above 99% when all three multiplexes are employed.

In 1994, Promega collaborated with the FBI Academy, Palm Beach Sheriff's Crime Laboratory, The Blood Center of Southeastern Wisconsin, and DNA Berkeley Laboratory to validate the CTT Multiplex. The validation work was published in the July 1996 issue of the Journal of Forensic Sciences (11). The study included experiments to determine sensitivity, evaluate annealing temperatures, analyze mixed samples, and complete additional work recommended by TWGDAM.

STR MULTIPLEX SYSTEMS FOR FLUORESCENT DETECTION

Shortly after silver stain multiplexes were introduced, the next generation of multiplex systems were developed using fluorescent detection. Figure 2 shows three fluorescent multiplex systems, each containing four loci. Panels A, B, and C display the CTTv Multiplex, FFFL Multiplex, and GammaSTR™ Multiplex, respectively. One primer for each locus is labeled with fluorescein. The four loci for each multiplex are simultaneously amplified, loaded on a gel, and subsequently detected with a fluorescent detection device, such as the Hitachi FMBIO® II fluorescent scanner (Hitachi Software Engineering America, Ltd., South San Francisco, CA) or the ABI PRISM® 310 Genetic Analyzer.

Multicolor detection allows for more loci to be analyzed in one reaction. The PowerPlex™ 1 System is a two-color multiplex system which allows for the simultaneous amplification of eight loci and detection in a single gel lane. The PowerPlex™ 1.1 System, designed and customized for the FMBIO® Fluorescent Scanner, is shown in Figure 3. Four loci (D16S539, D7S820, D13S317, D5S818) are labeled with fluorescein (green fragments in panel A) and the other four loci are labeled with TMR (red fragments in panel A). A third dye shown in blue is used for the Fluorescent Ladder (CXR), 60-400 Bases which is an internal lane size standard. The PowerPlex™ 1.2 System contains the same loci as PowerPlex 1.1 plus Amelogenin, but is designed for detection using the ABI PRISM® 377 DNA Sequencer or ABI PRISM® 310 Genetic Analyzer. Figure 4 shows one sample amplified using the PowerPlex™ 1.2 System and detected using the ABI PRISM® 310 Genetic Analyzer. The blue peaks represent the fluorescein-labeled loci, the black peaks are the TMR labeled loci, and the red peaks represent the Fluorescent Ladder (CXR), 60-400 Bases. Allelic ladders have been developed for each locus in the PowerPlex™ 1.2 System (Figure 5). Individual alleles have been balanced within each locus and allelic ladders have been balanced within the same color and between colors. Each component of the allelic ladder has been sequenced to confirm the repeat number and fragment length.

The PowerPlex™ 1.1 System and the FFFL Multiplex were recently validated according to TWGDAM guidelines. Both systems performed well for all the different parameters tested and the results have been submitted for publication (12).

Population studies have been completed for all twelve loci contained in the PowerPlex™ 1 System and FFFL Multiplex for three population groups (10). Table 3 displays the matching probability for the FFFL Multiplex which exceeds 1 in 2600 while the PowerPlex™ 1 System is at least 1 in 114,000,000. When the two multiplexes are combined for a total of

twelve loci, the matching probability exceeds 1 in over 300 billion. For paternity testing, the typical paternity index is greater than 2600 and the power of exclusion is above 99.9% for all three population groups when using both the FFFL Multiplex and the PowerPlex™ 1 System.

NEW STR MULTIPLEX SYSTEMS

Two new multiplex systems are currently in development in our laboratory to allow the amplification of additional loci including the remaining loci selected as the thirteen CODIS core loci (13) and the six loci contained in the Second Generation Multiplex (SGM) (14). In addition, the new multiplex systems will incorporate one or two of the new pentatnucleotide repeat loci. Pentanucleotide loci contain a repeat structure of 5 base pairs in length, have less stutter artifact than tetranucleotide loci, are highly discriminating loci, and have greater spatial separation of individual alleles (15). These advantages make the pentanucleotide loci ideal for inclusion in the new multiplex systems.

The new PowerPlex™ 2 System is a two-color multiplex which contains Amelogenin, eight tetranucleotide STR loci (TPOX, TH01, vWA, FGA, D21S11, D8S1179, D18S51, D3S1358), and one new pentanucleotide locus (Penta E). As designed, this multiplex contains the six SGM loci and, when used in combination with PowerPlex™ 1, contains all thirteen CODIS core loci. Three STR loci (TPOX, TH01, vWA) are shared between PowerPlex™ 1 and PowerPlex™ 2 for cross-checking purposes to confirm that the same sample has been amplified. Population studies for the PowerPlex™ 2 loci are in progress and will be published when completed.

The new PowerPlex™ 16 System is a three-color multiplex system which allows the simultaneous amplification of sixteen loci in a single amplification reaction. It contains all the loci in PowerPlex™ 1 and PowerPlex™ 2 plus an additional pentanucleotide locus. The loci FGA, TPOX, D8S1179, vWA, and Amelogenin are labeled with one dye color; the loci Penta E, D18S51, D21S11, TH01, and D3S1358 are labeled in a second dye color; and the loci Penta D, CSF1PO, D16S539, D7S820, D13S317, and D5S818 are labeled in a third dye color. The advantages of the PowerPlex™ 16 system include a one tube amplification, no overlap across loci labeled with the same dye color, contains the six SGM loci, contains the thirteen CODIS core loci, five STR loci have alleles less than 200 bases, and contains two new low stutter pentanucleotide loci. Preliminary population studies estimate the matching probability to exceed 1 in 10^{17} and the Power of Exclusion to exceed 99.999%.

SUMMARY

The amplification of short tandem repeat loci has progressed from single locus reactions to megaplex reactions, amplifying as many as sixteen loci in one reaction tube. STR multiplex systems have been developed for both manual silver stain detection and fluorescent detection formats. The one-color multiplex systems are compatible with both the Hitachi FMBIO and the ABI PRISM instrumentation, while the two-color multiplexes are customized for a respective instrument. The new low-stutter pentanucleotide loci which are highly discriminating and have few microvariants have been incorporated into the PowerPlex™ 2 and PowerPlex™ 16 Systems. The PowerPlex™ 1 and PowerPlex™ 2 Systems may be used in combination to support the CODIS core loci or the SGM loci. Alternatively, the three-color PowerPlex™ 16 System may be used to amplify sixteen loci (including the SGM loci and CODIS core loci) in a single amplification reaction.

REFERENCES

1. Edwards, A. *et al.* (1991) "DNA Typing and Genetic Mapping with Trimeric and Tetrameric Tandem Repeats," *Am. J. Hum. Genet.* **49**, 746.

2. Edwards, A. *et al.* (1992) "Genetic Variation at Five Trimeric and Tetrameric Tandem Repeat Loci in Four Human Population Groups," *Genomics* **12**, 241.

3. Hammond, H. *et al.* (1994) "Evaluation of 13 Short Tandem Repeat Loci for Use in Personal Identification Applications," *Am. J. Hum. Genet.* **55**, 175.

4. Kimpton, C. *et al.* (1994) "Evaluation of an Automated DNA Profiling System Employing Multiplex Amplification of Four Tetrameric STR Loci," *Int. J. Leg. Med.* **106**, 302.

5. Urquhart, A. *et al.* (1995) "Highly Discriminating Heptaplex Short Tandem Repeat PCR System for Forensic Identification," *BioTechniques* **18**, 116.

6. Lins, A.M. *et al.* (1996) "Multiplex Sets for the Amplification of Polymorphic Short Tandem Repeat Loci – Silver Stain and Fluorescent Detection," *BioTechniques* **20**, 882.
7. Bassam, B.J., Caetano-Anolles, G. and Gresshoff, P.M. (1991) "Fast and Sensitive Silver Staining of DNA in Polyacrylamide Gels," *Anal. Biochem.* **196**, 80.

8. Puers, C. *et al.* (1994) "Allelic Ladder Characterization of the Short Tandem Repeat Polymorphism Located in the 5' Flanking Region to the Human Coagulation Factor XIII A Subunit Gene," *Genomics* **23**, 260.

9. Puers, C. *et al.* (1993) "Identification of Repeat Sequence Heterogeneity at the Polymorphic STR Locus HUMTH01 $[AATG]_n$ and Reassignment of Alleles in Population Analy-

sis Using a Locus-Specific Allelic Ladder," *Amer. J. Human Genet.* **53**, 953.

10. Lins, A.M. *et al.* (1998) "Development and Population Study of an Eight-Locus Short Tandem Repeat (STR) Multiplex System," *J. Forensic Sci.* **43**, 1168.

11. Micka, K. *et al.* (1996) "Validation of Multiplex Polymorphic STR Amplification Sets Developed for Personal Identification Applications," *J. Forensic Sci*. **41**, 582.

12. Micka, K. *et al.* (submitted) "TWGDAM Validation of a Nine-Locus and Four-Locus Fluorescent STR Multiplex System."

13. Budowle, B. (1997) "Studies for Selecting Core STR Loci for CODIS", at DNA Forensics: Science, Evidence, and Future Prospects, McLean VA, November 17-18, 1997 (oral presentation).

14. Sparkes, R. *et al* (1996) "The Validation of a 7-locus Multiplex STR Test for use in Forensic Casework," *Int. J. Legal Med.* **109**, 195.

15. Schumm, J.W. *et al* (in press) "Pentanucleotide Repeats: Polymorphic Genetic Markers Displaying Minimal Stutter Artifact," In: *Proceedings from the Second European Symposium on Human Identification 1998*, Promega Corporation.

Table 1. Characteristics of *GenePrint*™ STR Systems

Locus	Chromosome Location	Allele Sizes (bases)	Alleles (frequency >0.001)
Amelogenin	X:Y	212, 218	X, Y
CSF1PO	5q33.3-34	291-327	6-15
D5S818	5q23.3-32	119-151	7-15
D7S820	7q11.21-22	215-247	6-14
D13S317	13q22-q31	165-197	7-15
D16S539	16q24-qter	264-304	5, 8-15
FESFPS	15q25-qter	222-250	7-14
F13A01	6p24.3-25.1	281-331	3.2, 4-16
F13B	1q31-q32.1	169-193	6-12
LPL	8p22	105-133	7-14
TH01	11p15.5	179-203	5-9, 9.3, 10-11
TPOX	2p25.1-pter	224-252	6-13
vWA	12p12-pter	127-167	11, 13-21

Table 2. Shared Loci Among *GenePrint*™ Silver and Fluorescent STR Multiplex Systems

	CSF1PO	TH01	TPOX	vWA	F13A01	FESFPS	F13B	LPL	D5S818	D7S820	D13S317	D16S539
Silver STR Multiplexes												
CTT	■	■	■									
FFv				■	■	■						
SilverSTR III										■	■	■
Fluorescent STR Multiplexes												
CTTv	■	■	■	■								
FFFL					■	■	■	■				
GammaSTR									■	■	■	■
PowerPlex 1	■	■	■	■					■	■	■	■

Table 3. Population Statistics

	Matching Probability (1 in …)			Typical Paternity Indices			Power of Exclusion		
	Caucasian-American	African-American	Hispanic-American	Caucasian-American	African-American	Hispanic-American	Caucasian-American	African-American	Hispanic-American
CTT Multiplex	435	1590	549	6.88	10.17	5.15	0.869	0.906	0.830
FFv Multiplex	927	2828	1343	9.81	16.04	7.76	0.904	0.938	0.881
SilverSTR™ III	2552	1152	2493	7.72	7.56	14.08	0.880	0.877	0.929
CTT, FFv, and SilverSTR™ III (combined)	1.03×10^{9}	5.18×10^{9}	1.84×10^{9}	521	1233	563	0.9985	0.9993	0.9986
FFFL Multiplex	2658	16802	3276	15.28	16.83	8.23	0.941	0.946	0.902
PowerPlex™ 1	1.14×10^{8}	2.74×10^{8}	1.45×10^{8}	260	498	319	0.9969	0.9982	0.9973
FFFL and PowerPlex™ 1 (combined)	3.03×10^{11}	4.61×10^{12}	4.75×10^{11}	3976	8373	2627	0.99981	0.99990	0.99974

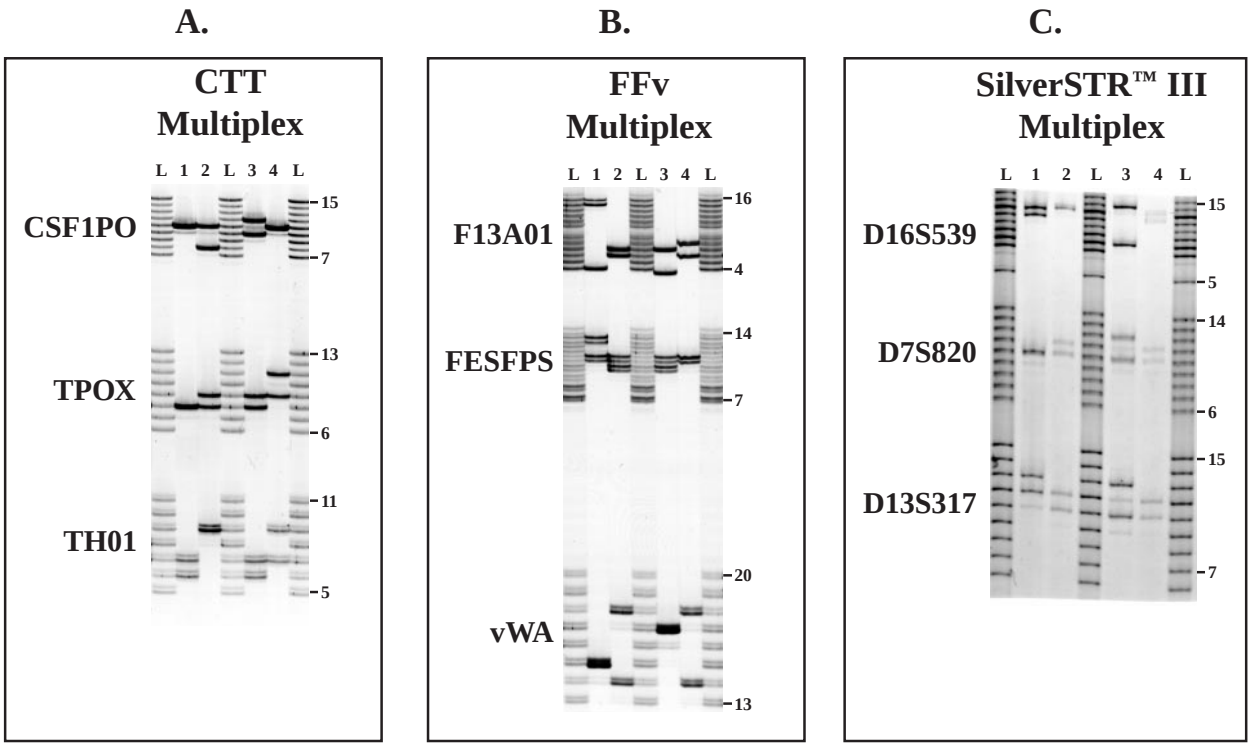

Figure 1. *GenePrint*™ STR Multiplexes for Silver Stain Detection. The CTT Multiplex, FFv Multiplex, and SilverSTR™ III Multiplex are displayed in panels A, B and C, respectively. Lanes 1-4 contain individually amplified DNA samples and lanes labeled (L) contain a mixture of allelic ladders for the respective loci in each multiplex system. The amplified fragments and allelic ladders were separated in a 4% denaturing polyacrylamide gel and detected by silver stain analysis.

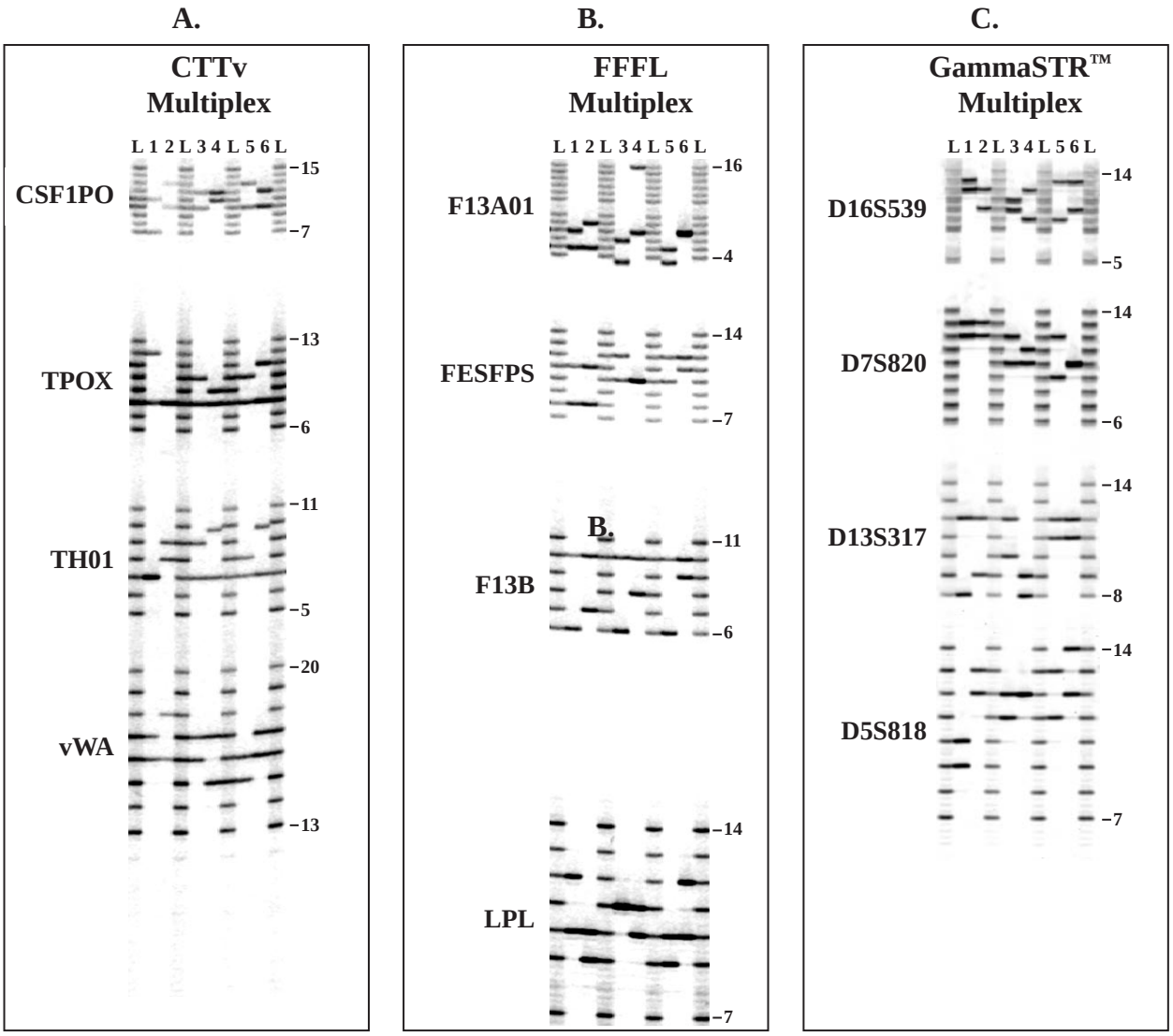

Figure 2. *GenePrint*™ STR Multiplexes for Fluorescent Detection. The CTTv Multiplex, FFFL Multiplex, and GammaSTR™ Multiplex are displayed in panels A, B and C, respectively. Lanes 1-6 contain individually amplified DNA samples and lanes labeled (L) contain a mixture of allelic ladders for the respective loci in each multiplex system. The amplified fragments and allelic ladders were separated in a 4% denaturing polyacrylamide gel and detected using the Hitachi FMBIO® II Fluorescent Scanner.

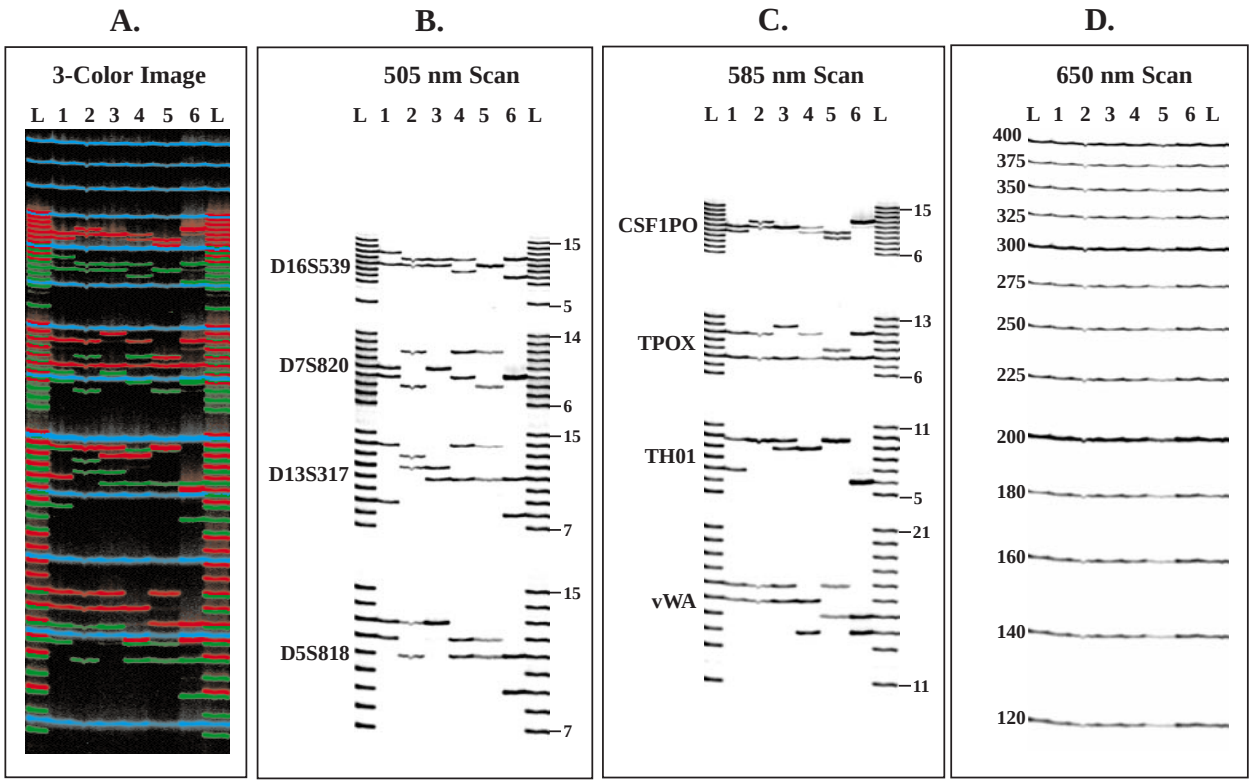

Figure 3. The *GenePrint*™ PowerPlex™ 1.1 System. Six genomic DNA samples (lanes 1-6) were amplified using the PowerPlex™ 1.1 System, separated in a 4% denaturing polyacrylamide gel, and detected using the Hitachi FMBIO® II Fluorescent Scanner. Panel A displays the three-color image. The color image can be separated in the individual scans or dye colors (Panels B, C, D). Panel B (505nm scan) represents the fluorescein-labeled loci (D16S539, D7S820, D13S317, D5S818) which are also displayed as green in panel A. Panel C (585nm scan) represents the TMR-labeled loci (CSF1PO, TPOX, TH01, vWA) which are also displayed as red in panel A. Panel D (650 nm scan) represents the Internal Lane Standard (CXR), 60-400 Bases which is also displayed as blue in panel A. Lanes labeled (L) contain a mixture of the allelic ladders for the respective loci in each dye color.

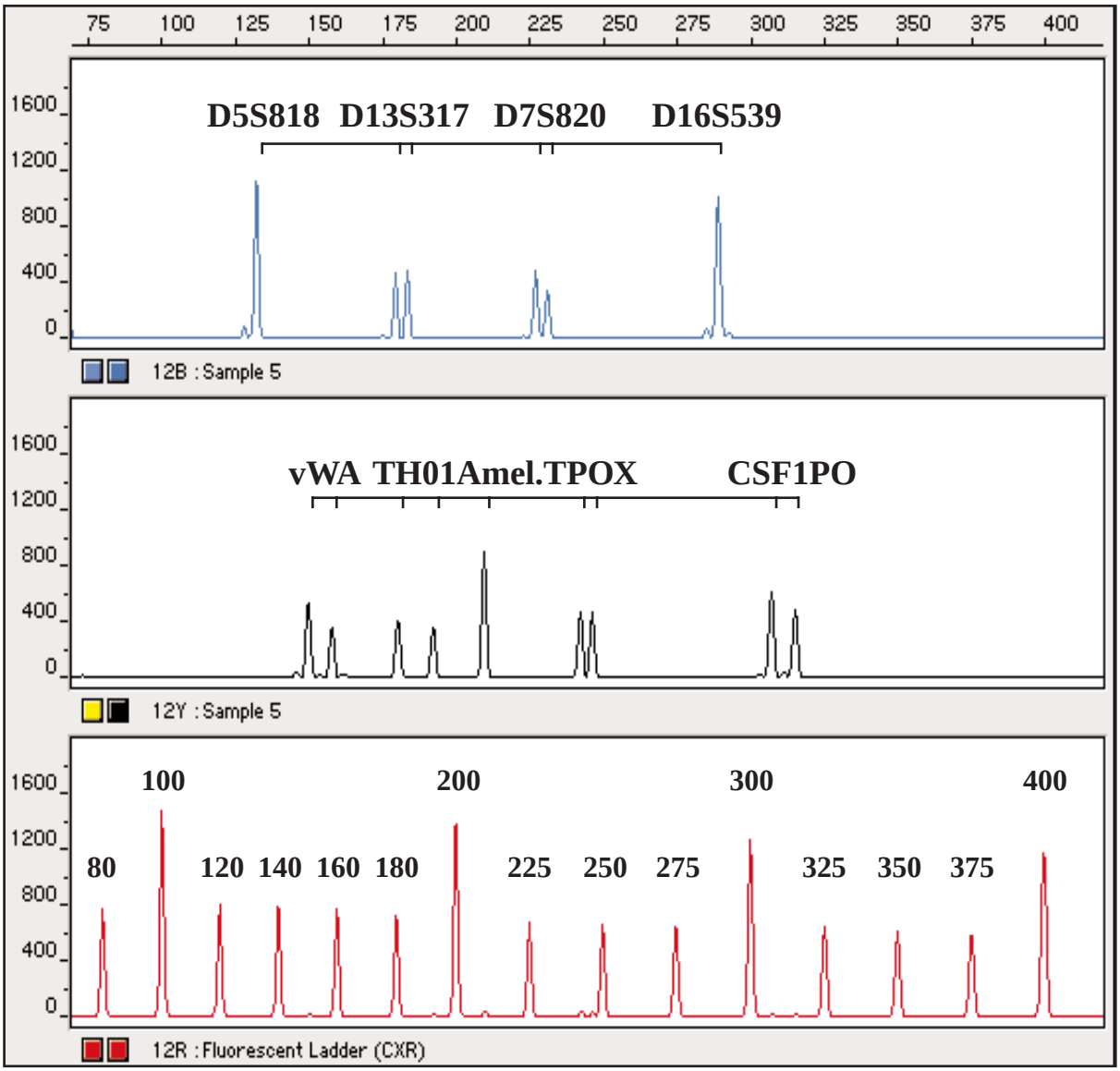

Figure 4. The *GenePrint*™ PowerPlex™ 1.2 System. A single genomic DNA sample was amplified using the PowerPlex™ 1.2 System and detected using the ABI PRISM® 310 Genetic Analyzer. The top panel displays the fluorescein-labeled loci (D5S818, D13S317, D7S820, D16S539), the middle panel displays the TMR-labeled loci (vWA, TH01, Amelogenin, TPOX, CSF1PO), and the bottom panel displays the Fluorescent Ladder (CXR), 60-400 Bases, which is used as an internal lane standard.

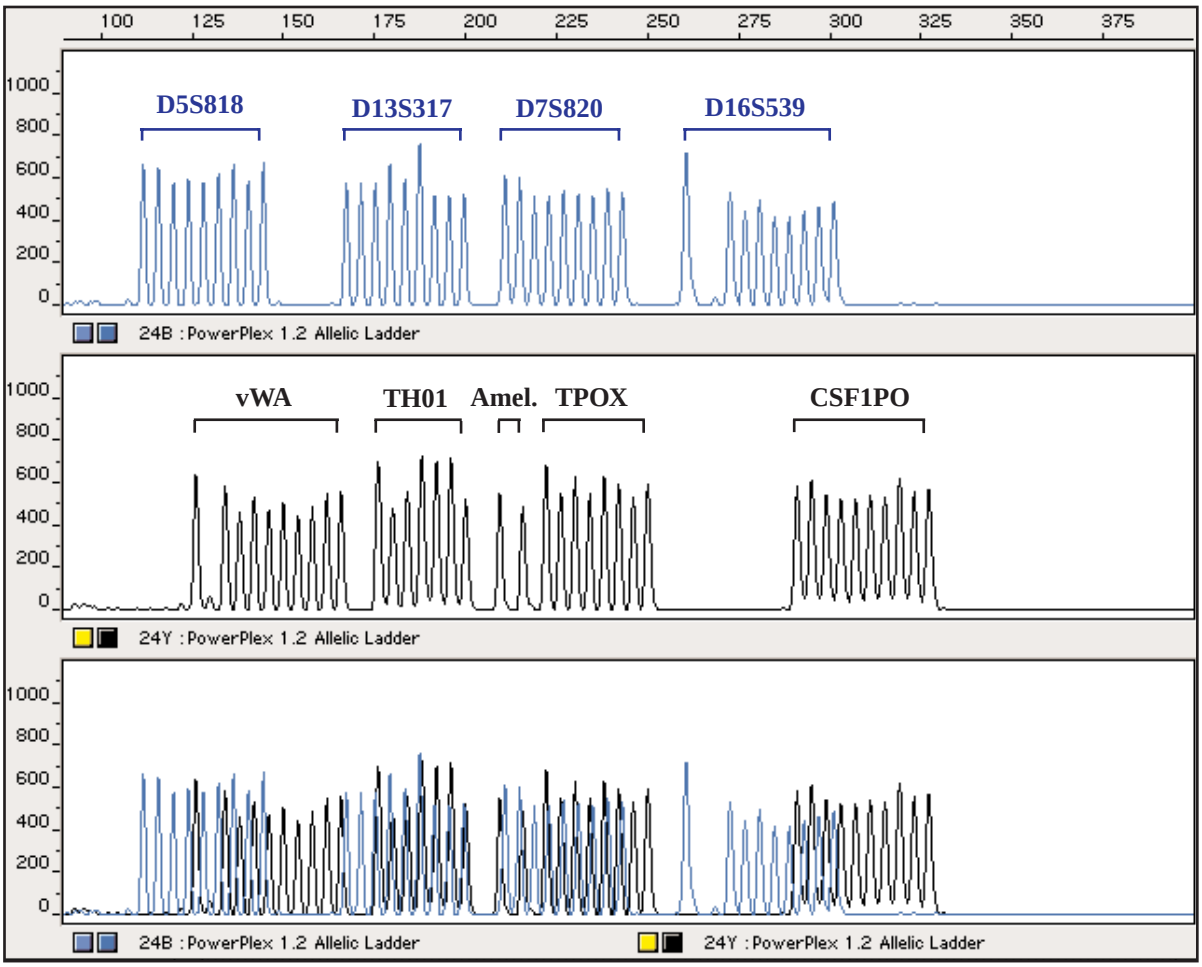

Figure 5. The *GenePrint*™ PowerPlex™ 1.2 Allelic Ladder Mix detected using the ABI PRISM® 310 Genetic Analyzer. The allelic ladders for the fluorescein-labeled loci D5S818, D13S317, D7S820, D16S539 are displayed in the top panel and the allelic ladders for the TMR-labeled loci vWA, TH01, Amelogenin, TPOX, CSF1PO are displayed in the middle panel. The composite mix showing the balance between allelic ladders and dye colors is shown in the bottom panel.