Incorporating High Quality Genetic Markers into Forensically Useful Multiplexes

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

INTRODUCTION

In recent years, short tandem repeat (STR) polymorphisms have become the standard genetic markers used for forensic casework and database development throughout the world. This shift to STRs has been made possible by the development of reliable multiplex systems that allow simultaneous amplification and detection of several loci, providing high throughput analysis and large powers of discrimination.

Presented herein is a review of the current *GenePrint*® STR multiplex systems currently in use as well as a description of the ongoing development of new STR multiplexes. Included in developments plans is a system that allows amplification and detection of fifteen STR loci and the gender identification marker Amelogenin giving a power of discrimination of approximately 1 in 10^{17} which far exceeds that of any system available today.

OVERVIEW OF *GENEPRINT*® STR SYSTEMS

The *GenePrint*® STR Systems have undergone a steady evolution over the years, allowing each new system to become more powerful with increased throughput capacity by taking advantage of the latest technologies. The first silver staining monoplex (TH01) in 1993 lead the way for three silver triplexes (CTT, FFv and Silver STR™ III). These were then expanded and converted to fluorescent technology with the introduction of fluorescein-labeled quadriplexes, the CTTv, FFFL and GammaSTR™ Systems. In 1996, by adding a second dye, the loci contained in the CTTv and GammaSTR™ Systems were incorporated into a single eight-locus multiplex, the *GenePrint*® PowerPlex™ 1.1 System, compatible with the Hitachi FMBIO® II Fluorescent Scanner. Following the introduction of the PowerPlex™1.1 System, the multiplex was reformulated to create the PowerPlex™1.2 System for compatibility with ABI PRISM® instrumentation so that optimal results could be achieved on a variety of detection platforms. At this time, a total of nine STR loci are available in three silver triplexes and twelve STR loci can be analyzed in as little as two fluorescent multiplex systems, the FFFL and the PowerPlex™1 Systems.

Currently, there are two new STR multiplex systems being developed in formats compatible with multiple detection instruments. The first of these is the *GenePrint*® PowerPlex™ 2 System (Figure 1), a 10-locus multiplex containing five STR loci labeled with fluorescein (Penta E, D18S51, D21S11, TH01 and D3S1358) and four loci labeled with tetramethyl-rhodamine (TMR), (FGA, TPOX, D8S1179 and vWA), as well as Amelogenin (also labeled with TMR). The second multiplex, the *GenePrint*® PowerPlex™ 16 System (Figure 2) is a complete system that allows for simultaneous amplification with 32 primers and single lane or capillary analysis of fifteen STR loci and Amelogenin. Included in this system are the same dyes and ten loci found in the PowerPlex™ 2 System configuration and six additional loci (Penta D, CSF1PO, D16S539, D7S820, D13S317 and D5S818) labeled with a third fluorescent dye (Figure 3). As with the existing fluorescent multiplexes, each system includes a premixed and balanced allelic ladder and an internal lane standard (ILS) comprised of evenly-spaced fragments labeled with carboxy-X-rhodamine (CXR) that can be included with each sample for accurate sizing of the amplification products.

NEW STR MULTIPLEX SYSTEM DEVELOPMENT

With the widespread increase in the use of STRs for databasing and forensic investigation, the PowerPlex™ 2 and PowerPlex™ 16 Systems are being developed to expand and improve upon existing STR systems in order to meet the growing needs of the DNA Typing community. Several key factors contributing to the design and features of both new systems are described below.

Locus Selection

The standardization of loci for national databases worldwide has a significant influence on locus selection for new multiplexes. In the United States, compliance with the required 13 core loci for the national CODIS (Combined DNA Indexing System) database is essential. In addition, the six STR loci included in the Second Generation Multiplex (SGM) from the FSS are standards for the United Kingdom and four of these have also been selected by the international police organization INTERPOL. Another European organization (ENSFI) is currently evaluating several systems in order to select a

standard set of loci to be used throughout the European community. Consequently, it is important to create systems that can satisfy the U.S. requirements and provide compatibility among databases in many nations throughout the world. Figure 4 illustrates how the 13 core CODIS loci can be analyzed in a single PowerPlex™ 16 System reaction or with the combination of the PowerPlex™ 1 and PowerPlex™ 2 Systems. In addition, all of the SGM and selected INTERPOL loci are contained within both the PowerPlex™ 2 System and the PowerPlex™ 16 System.

Improvements to System Design

Improvements to overall system configuration have been incorporated into the new multiplexes to help prevent allele overlap, improve allelic ladders and generate small amplification products for compatibility with degraded DNA samples. Throughout the process of generating thousands of DNA profiles with all the STR systems that currently exist, several new alleles have been discovered beyond those known at the time of development. As a result, these newly identified alleles fall above or below the range of their corresponding allelic ladders and overlap into adjacent loci. Collaborations with organizations such as the FSS, the FBI and the Bode Technology Group have allowed information on newly identified alleles to be incorporated into the new systems' design to help prevent allele overlap and develop more complete allelic ladders. Also, with knowledge of the full range of alleles for each locus, primers can be selected to produce allele fragments that fall within a specified size range. When working with forensic samples it is particularly desirable to have smaller amplification products (<400 bases) due to the possibility of degraded DNA samples. The PowerPlex™ 16 System contains twelve CODIS loci in which all of the known alleles migrate below 372 bases as well as eight CODIS loci and five SGM loci with amplification products smaller than 265 bases. The FGA locus is not included in these smaller size categories only because of a few larger alleles that are rare (less than 2% of those observed).

New Polymorphic Loci for Minimal Stutter

Another factor influencing system development is the occurrence of stutter bands one repeat length smaller than the major allele. The tetranucleotide repeat STR loci used today typically exhibit an average stutter artifact in the range of 4-6%. It is well known that the appearance of these stutter bands in mixed forensic samples can complicate interpretation. Several new STR loci containing pentanucleotide repeat units have been identified which are highly polymorphic and display minimal stutter artifact, typically in the range of ≤1-2%. Inclusion of these new loci in the multiplex systems adds the benefits of increased discrimination power and more straightforward mixture interpretation due to less stutter artifact. In addition, the five-base repeat gives greater separation between alleles. Therefore, primers can be designed to produce larger fragments that can be separated in gels or capillaries. A more detailed description of pentanucleotide repeat loci was presented at the *Ninth International Symposium on Human Identification*. (1)

CONCLUSION

It cannot be overstated that that nature of forensic analysis and databasing requires universal systems that are accurate and easy to use, efficient and powerful. This can be achieved by generating systems with high quality standard components in the form of evenly spaced size markers and premixed and balanced allelic ladders. In addition, primers should be designed to prevent allele overlap, produce small amplification products and to reach optimal performance with multiple detection platforms. The introduction of the *GenePrint*® PowerPlex™ 16 System will provide a system with extraordinary discrimination power that maintains the high standards of previous *GenePrint*® Systems and increases efficiency by essentially doubling current throughput capabilities.

ACKNOWLEDGEMENTS

We would like to extend our thanks to Rebecca Sparkes of the Forensic Science Service for her assistance in identifying alleles for the systems currently under development. In addition, special thanks to Robert A. Bever, Kevin McElfresh and the Bode Technology Group (Sterling, VA) and Steven D. Creacy (Creacy Enterprises, Inc., Greensboro, N.C.) for their assistance in the generation of population databases.

REFERENCES

1. Bacher, Jeffrey W., *et al*. (In Press). Pentanucleotide Repeats: Highly Polymorphic Genetic Markers Displaying Minimal Stutter Artifact. In: *Proceedings from the Ninth International Symposium on Human Identification.* 1998 Oct.7-10; Orlando (FL); Promega Corporation.

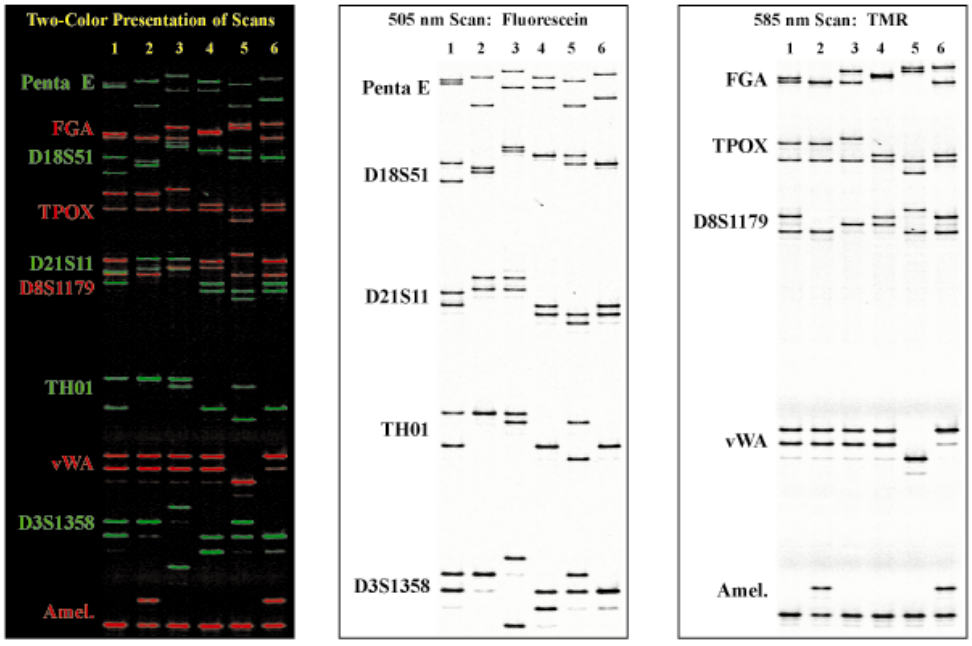

Figure 1. The *GenePrint*® PowerPlex™ 2.1 System Protoype detected using the Hitachi FMBIO® II Fluorescent Scanner. Six human DNA samples derived from blood (lanes 1-6) were amplified using a prototype version of the PowerPlex™ 2.1 System. The image on the left simultaneously displays two scans of the same gel at 505nm (displayed in green) and 585nm (displayed in red) and the two images to the right display each scan individually in black and white. The 505nm scan represents the loci labeled with fluorescein (D3S1358, TH01, D21S11, D18S51 and Penta E) and the 585nm scan represents TMR-labeled loci (Amelogenin, vWA, D8S1179, TPOX and FGA).

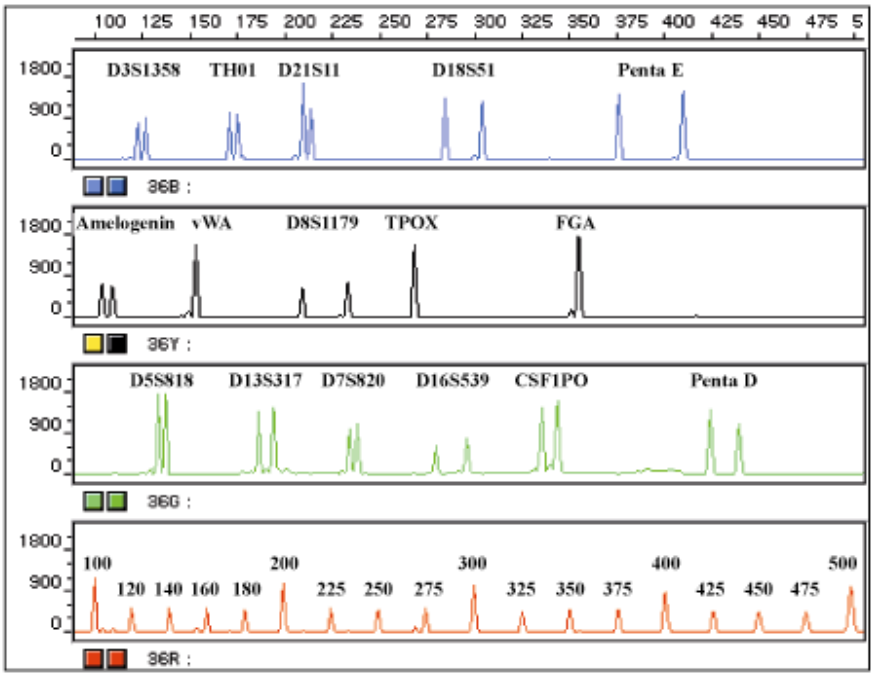

Figure 2. The *GenePrint*® PowerPlex™ 16.2 System Protoype Illustrated with Internal Lane Standard 500. The electropherogram of a DNA Sample co-amplified at sixteen loci is shown. The amplified products of the fluorescein-labeled loci, D3S1358, TH01, D21S11, D18S51 and Penta E are shown as blue peaks, the TMR-labeled loci, Amelogenin, vWA, D8S1179, TPOX and FGA are shown as black peaks and the products of loci labeled with a third dye currently under development, D5S818, D13S317, D7S820, D16S539, CSF1PO and Penta D are shown as green peaks. Fragments of the Internal Lane Standard 500 (ILS 500) are illustrated in red. All materials were separated using the ABI PRISM® 310 Genetic Analyzer.

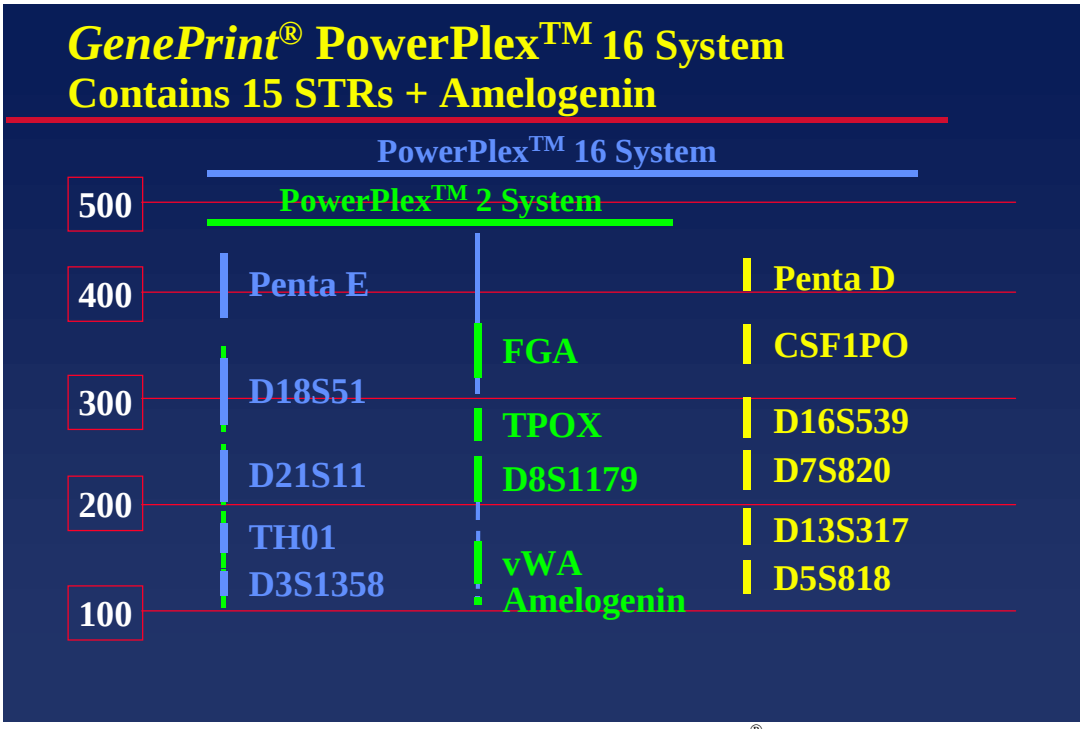

Figure 3. System configuration for the development of the *GenePrint*® PowerPlex™ 2 System and the *GenePrint*® PowerPlex™ 16 System. Illustrates the locus and dye configuration of both systems and the approximate size range for each locus. Narrow lines extending above and/or below each locus represent rare alleles.

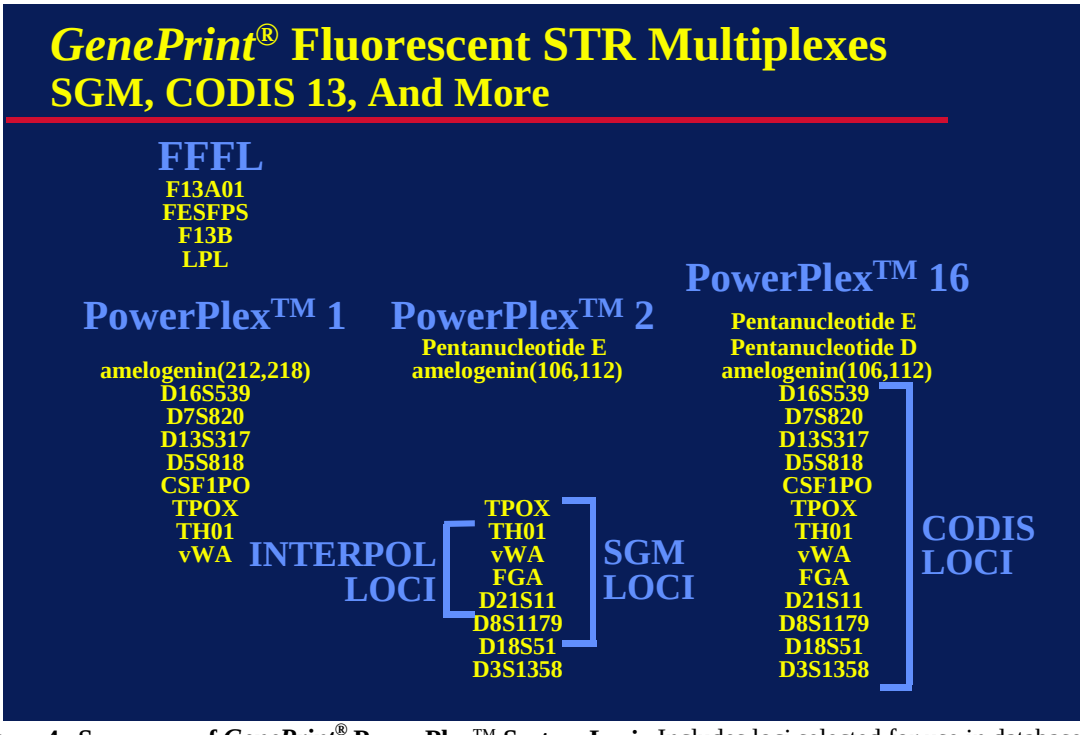

Figure 4. Summary of *GenePrint*® PowerPlex™ System Loci. Includes loci selected for use in database systems for the United States (CODIS) the United Kingdom (SGM multiplex) and Europe (INTERPOL).