

STR PRIMER CONCORDANCE DATA - VALIDATION STUDIES

Bruce Budowle, Senior Scientist,

FBI Laboratory, FBI, 935 Pennsylvania Ave., NW, Washington, D.C. 20535

The 13 short tandem repeat (STR) loci CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11 have been selected as the core loci required for CODIS (1). One of the benefits of establishing standard loci for typing known reference database samples and for casework analyses is that high quality, commercially available kits have been developed that contain the requisite reagents to amplify by PCR the core CODIS loci (Table 1). Only data generated from CODIS approved kits will be accepted into NDIS. The multiplex kits include: PowerPlex® 1.1 System, PowerPlex® 1.2 System and PowerPlex® 2.1 System (Promega Corp., Madison, WI), AmpF/STR™ Profiler Plus™ and AmpF/STR™ Cofiler™ DNA typing systems (Applied Biosystems, Foster City, CA). To be registered as an NDIS accepted kit (or if a currently accepted kit is modified), data shall be presented to NDIS, that demonstrate that the new kit generates NDIS compatible results (or that the modification is justified). Until recently, use of at least two kits was necessary so that all 13 STR loci could be typed. Commercial multiplex kits are being developed that enable amplification of all 13 CODIS core loci in a single reaction. These kits (i.e., PowerPlex® 16 kit (Promega Corporation) and Identifiler™ (Applied Biosystems)) currently are under validation for use in typing samples from convicted felons for entry into CODIS and for casework analysis.

Table 1. NDIS accepted STR kits.

Manufacturer	Kit Name (Catalog number)
Promega Corp.	PowerPlex® 1.1(DC6091/6090)
Promega Corp.	PowerPlex® 1.2 (DC 6101/6100)
Promega Corp.	PowerPlex® 2.1 (DC 6471/6470)
Applied Biosystems	AmpF/STR™ Profiler Plus™ (PIN 4303326)
Applied Biosystems	AmpF/STR™ Cofiler™ (PIN 4305246)
Applied Biosystems	AmpF/STR™ Profiler Plus and AmpF/STR™ Cofiler™ (PIN 4305979)
Promega Corp.*	Monoplex D5S818 (DC6161)

Promega Corp.*	Monoplex D7S820 (DC6141)
Promega Corp.*	Monoplex D13S317 (DC6151)
Promega Corp.*	Monoplex D16S539 (DC6131)
Promega Corp.*	Monoplex TH01 (DC5081)
Promega Corp.*	Monoplex TPOX (DC5111)
Promega Corp.*	Monoplex CSF1PO (DC5091)
Promega Corp.*	Monoplex vWA (DC5141)

* Monoplexes are all fluorescently-labeled and the chemistry is the same as in the multiplex kits

Because there are different manufacturers producing kits, the sequence of primers and the flanking region where the primers reside differ per locus per manufacturer (and in some cases between different versions of kits from the same manufacturer). Typically, the intent is to design primers such that they will anneal to STR target site flanking regions that are generally invariant in the population. Regardless of design, when a sufficiently large number of individuals are typed, a variant may well be observed in the “invariant” flanking region of the target sequence where a primer resides. If a mismatch in complementary binding is proximal to the 3' end of the primer, the mismatch may be such that primer extension does not occur during PCR. When no product is synthesized, allele dropout results. Profile comparisons within and between laboratories that use the same primer sets usually would be unaffected by potential allele dropout. [Although in some rare casework scenarios where inhibitors or contaminants co-purify with the DNA, it may be possible to observe an allele dropout in an evidence sample and not in a reference sample from the same source using the same primer sets (and vice versa)].

By comparing typing results on samples that have been analyzed using the different primer sets from the different manufacturers, the degree of allele dropout can be assessed for the respective kits from each manufacturer. More importantly, the results of such a comparison can be used as an additional mechanism to validate the typing reliability of the primer sets. In fact, a number of primer concordance validation studies have been carried out. Budowle, et al. (2) typed several hundred individuals for the loci D5S818, D7S820, D13S317, CSF1PO, TPOX, and TH01 using different manufacturer’s kits. They compared typing results for the loci CSF1PO, TPOX, and TH01 for African American, Caucasian and Hispanic samples using the COfiler kit and the GenePrint CTTV kit (the same primers are used in the PowerPlex® 1.1 kit). All typing results were concordant, except for one sample in Caucasians at the CSF1PO locus. The PowerPlex® 16 primers yielded 11,15 as the type; but with the COfiler primers, the type appeared as a 15,15. The TPOX and TH01 types were the same with either primer set for this sample, and the types were 8,8 and 6,9.3, respectively. The presence of the CSF1PO null allele was confirmed by comparison of the CSF1PO 15 peak height and peak area with those of the TPOX and TH01 alleles generated

using the COfiler kit. For good quality samples, heterozygous allele peaks at one locus generally exhibit peak heights and peak areas similar to heterozygous allele peaks at other loci, labeled with the same fluorophore, in a profile. Homozygous peak heights and peak areas generally are approximately twice those of heterozygous allele peaks. The CSF1PO 15 peak height (211 rfu) and peak area (1884 rfu) were similar to the TH01 6 peak height (250 rfu) and peak area (1576 rfu), and TH01 9.3 peak height (258 rfu) and peak area (1694 rfu), but half that of the TPOX 8 peak height (505 rfu) and peak area (3662 rfu), respectively. Thus, allele dropout of the CSF1PO 11 allele due to a primer mismatch is the most plausible explanation.

The primers for the D5S818, D7S820, and D13S317 loci are included in both the PowerPlex® 1.1 and Profiler Plus kits. In the study by Budowle, *et al.* (2), there were no typing discrepancies for any samples at the D5S818 and D7S820 loci. However, at the D13S317 locus, nine samples (in Bahamians, Trinidadians, and Jamaicans) were typed as heterozygotes using the Profiler Plus kit but appeared as homozygotes (i.e., single peaks at the locus) using the PowerPlex® 16 kit. The allele drop out was due to a 4-base deletion in the region where one of the D13S317 primers resides (personal communication, Promega Corporation). The D13S317 primers have been modified in subsequent generation PowerPlex® kits so that amplification of the variant alleles can be accomplished.

Kline, *et al.* (3) reported an observed allele drop-out at the vWA locus by typing 600 samples using both the PowerPlex® 1.1 kit and AmpF/STR Blue™ kit (Applied Biosystems, Foster City, CA). Only one sample was typed differently when using these kits. Thus, the presence of null alleles at the vWA locus also is a rare occurrence (4).

Budowle and Sprecher (5) typed over 500 population database samples comprising African Americans, Bahamians, and Southwestern Hispanics using both the PowerPlex® 16 and the Profiler Plus™ /COfiler™ kits. There was only one sample in African Americans in which a typing difference was observed. An FGA heterozygote profile was observed using the PowerPlex® 16 primers, and a single allele FGA profile was observed using Profiler Plus™ primers. The presence of the FGA null allele with Profiler Plus™ primers was confirmed by comparison of the peak height and peak area of the FGA single peak with those of the vWA and D3S1358 alleles generated using the Profiler Plus™ kit. The FGA single peak's height (762 rfu) and area (8424 rfu) were similar to the heterozygous vWA peak heights (762 and 829 rfu) and peak area (8739 and 9336 rfu), and approximately half that of the homozygous D3S1358 peak height (1860 rfu) and peak area (18232 rfu) for the sample.

Currently, additional concordance studies are being compiled on samples typed using both the PowerPlex® 16 kit and the Profiler Plus™ /COfiler™ kits (manuscript in preparation). Preliminary concordance data are available on approximately 1300 samples (African Americans N=277, Caucasians N=481, Hispanics N=234, Native Americans N=313). There were 6 examples of allele dropout observed (Table 2). Thus, in more than 1800 samples typed (by combining these results with those of Budowle and Sprecher (5)) using both manufacturer's kits, there were only 7 examples of allele drop-out, and these were observed in only 4 of the 13 core

STR loci. The loci are CSF1PO, D16S539, FGA, and vWA, and there were no more than 2 observations in any of the 4 loci.

Table 2. Total number of discrepant types due to allele drop-out observed in approximately 1300 individuals typed with both the PowerPlex® 16 kit and the Profiler Plus™/COfiler kits™^a

Population	Locus	PowerPlex® 16	Profiler Plus/ COfiler
Caucasian	D16S539	13,13	12,13
Caucasian	vWA	17,18	18,18
Caucasian	vWA	15,16	16,16
Native American	CSF1PO	12,14	12,12
Native American	CSF1PO	13,14	13,13
Native American	D16S539	11,11	11,12

a) Data are part of a consortium study. Contributors to the data in the Table are: Barbara Llewellyn, Sindy Schueler, Tim Kupferschmid, Jennifer Hamill, Charles Barna, Susanne Brenneke, Chris Tomsey, Chris Palaski, Stacey Anderson, Jill Cramer, Gretchen DeGroot, Dan Hasse, and Dirk Janssen.

In conclusion, the current primer sets from the commercial manufacturers Applied Biosystems and Promega Corporation do not produce significant levels of allele dropout in the populations studied. Very few typing discrepancies have been observed. Typing samples with different primer sets and comparing the results also contribute to the data that support the validity of STR typing. The knowledge of the primer sequences was not needed to perform the validation study. The kit(s) described in this concordance can be used for STR typing and data can be used in CODIS (as of the date of this manuscript, the Powerplex® 16 kit is approved only for typing offender samples). Reliable results can be obtained, as long as proper protocols are used and appropriate validation studies are carried out to define the limitations of the analytical systems (see FBI Quality Assurance Standards (6)). If new primer sets are developed for any of the 13 core STR loci, concordance studies should be performed as part of the validation process.

One cannot assure that allele dropout will never occur; however, the phenomenon appears to be relatively rare for the current primer sets/STR systems and for the major population groups residing in the United States. Since allele dropout occurs at a very low frequency, common sense and the current searching algorithms in CODIS can accommodate the matching requirements when an allele dropout may be considered. CODIS matches will still occur with the low-level allele dropout that has been observed. When using different kits/primer sets, a mismatch at only one locus characteristic of allele dropout should still be strong evidence for an investigative lead.

This is publication number 01-03 of the Laboratory Division of the Federal Bureau of Investigation. Names of commercial manufacturers are provided for identification only, and inclusion does not imply endorsement by the Federal Bureau of Investigation.

REFERENCES

- 1) Budowle B, Moretti TR, Niezgoda SJ, Brown, BL. CODIS and PCR-based short tandem repeat loci: Law enforcement tools. In: Second European Symposium on Human Identification 1998, Promega Corporation, Madison, Wisconsin pp 73-88, 1998.
- 2) Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forens Sci 1999;44:1277-86.
- 3) Kline MC, Jenkins B, Rodgers S. Non-amplification of a vWA allele. J Forens Sci 1998;43:250.
- 4) Walsh S. Commentary on Kline, M.C., Jenkins, B., Rodgers, S. Non-amplification of a vWA allele. J Forens Sci 1998;43:1103-4.
- 5) Budowle B, Sprecher C. Concordance study on population database samples using the PowerPlex™ 16 Kit and AmpF/STR® Profiler Plus™ Kit and AmpF/STR® COfiler™ Kit. J Forens Sci (in press).
- 6) Quality Assurance Standards for Forensic DNA Testing Laboratories. Forensic Science Communications July 200; 2(3). Available at: <http://www.fbi.gov/programs/lab/fsc>.