TECHNICAL NOTE

Henrietta Margolis-Nunno[1] Ph.D.; Lewis Brenner,[2] M.S.; Jennifer Cascardi,[1] B.S.; and Lawrence Kobilinsky,[1] Ph.D.

A New Allele of the Short Tandem Repeat (STR) Locus, CSF1PO

REFERENCE: Margolis-Nunno H, Brenner L, Cascardi J, Kobilinsky L. A new allele of the short tandem repeat (STR) locus, CSF1PO. J Forensic Sci 2001;46(6):1480–1483.

ABSTRACT: CSF1PO is one of the thirteen core loci used for the CODIS database, and alleles reported for this short tandem repeat (STR) locus contain from 6 to 15 repeats of the tetranucleotide AGAT. Screening of DNA from 76 individuals by gel electrophoresis and silver stain detection yielded one sample that contained a rare, off-ladder CSF1PO allele; an allele larger than $CSF1PO_{15}$ was detected in a heterozygote that also contained a $CSF1PO_{10}$ allele. Capillary electrophoresis analysis using GeneScan™ software demonstrated that the variant allele contained four bases more than $CSF1PO_{15}$. Following agarose gel electrophoresis to separate the two alleles of the heterozygote and cycle sequencing using dye terminators, sequence analysis showed that the variant, which was otherwise identical to the CSF1PO GenBank sequence, contained exactly 16 AGAT repeats. These results demonstrate the existence of an additional CSF1PO allele, a previously unreported size variant, $CSF1PO_{16}$.

KEYWORDS: forensic science, DNA typing, short tandem repeats, CSF1PO, variants, alleles

Tandemly repeated DNA sequences are common in the human genome and show considerable variability among individuals. Such polymorphisms have become useful tools for DNA analysis and typing, and are currently employed for human identification, paternity testing, and genetic mapping (1–3).

Short tandem repeats (STRs) contain tandemly repeated DNA sequences ranging in length from about two to seven base pairs (3–5). Of the STR loci used in identity testing, most have tetranucleotide repeats and allele sizes between 100 and 350 base pairs. STR alleles are defined and named by the number of repeated sequences they contain, and as more samples and populations are being examined, the number of reported alleles is increasing. (For recent published and unpublished updates on STR alleles, see Ref 6.)

The locus CSF1PO (7) (proto-oncogene for CSF-1 receptor gene) is included in the thirteen core loci used by the FBI for the Combined DNA Index System (CODIS) database. Reported alleles of this STR locus contain 6 to 15 repeats of the tetranucleotide AGAT. Herein, we describe our finding of a novel CSF1PO allele.

Materials and Methods

DNA was extracted from a 3 mm^2 cutting from each bloodstain following modifications of the Chelex method of Walsh et al. (8) as described in Ref 9. The final supernatant was used directly for PCR.

Initially, bloodstain samples from 76 random individuals were screened for STRs at the DNA Identification Laboratory of the Philadelphia Police Department. Silver stain detection of CSF1PO, TPOX, and THO1 alleles utilized the GenePrint™ STR System (Promega, Madison, WI) and polyacrylimide gel electrophoresis. The manufacturer's Technical Manual (9) was followed and Protocol 2 was used for PCR (96°C for 2 min; 10 cycles of 94°C for 1 min, 64°C for 1 min, and 70°C for 1.5 min; 20 cycles of 90°C for 1 min, 64°C for 1 min, and 70°C for 1.5 min). Amplification product (2.5 μL per sample mixed with an equal volume of STR 2X loading solution) was electrophoresed at 1500 V for 1.5 h in a 19.5 cm by 32 cm, 0.4 mm thick denaturing polyacrylamide gel. Alleles were determined by visual comparison of band positions with those of the allelic ladder run next to each sample lane, after silver staining the gel. All further DNA characterizations were performed by capillary electrophoresis at John Jay College of Criminal Justice using the ABI Prism™ 310 Genetic Analyzer (PE Applied Biosystems, Division of Perkin Elmer, Foster City, CA).

For GeneScan™ analysis, approximately 1.0 to 2.0 ng of DNA was amplified in 50 μL reaction volumes using the AmpFISTR™ Green I PCR Amplification Kit (Perkin Elmer), following the manufacturer's protocol for PCR (10) (95°C for 11 min; 29 cycles of 94°C for 1 min, 59°C for 1 min, and 72°C for 1 min; then 60°C for 30 min). Where possible, DNA was quantitated spectrophotometrically (i.e., 1 OD unit = 50 μg/mL).

Amplification products were electrophoresed on the ABI Prism™ 310 Genetic Analyzer following manufacturer's recommendation (10). Briefly, 1 μL of PCR product, or AmpFISTR™ Green I allelic ladder, was combined with 18 μL of deionized formamide and 1 μL ROX-500 size standard. Following denaturation (2 min at 95°C and immediate chilling on ice) samples were loaded and run using Module GS STR POP4 A (10). Results were analyzed using GeneScan™ Analysis Software, Version 3.1 using de-

[1] Science Department, John Jay College of Criminal Justice, City University of New York, New York, N.Y.

[2] DNA Identification Laboratory, Philadelphia Police Department, Philadelphia, PA.

Received 7 July 2000; and in revised form 15 Feb. 2001; accepted 16 Feb. 2001.

fault parameters with the exception of the following: range 2950 to 10 000, minimum peak half width − 3 pts., and no split peak correction.

Prior to cycle sequencing, PCR was carried out as described for the AmpFISTR™ Green I Kit above (10). The same reagents and quantities were utilized except for unlabeled, custom synthesized forward and reverse primers (final concentration of 1 pmol/μL each, in 11 μL). CSF1PO primers (Custom Oligo Synthesis Service, PE Applied Biosystems) were as follows: Forward: AACCTGAGTCTGCCAAGGACTAGC; Reverse: TTCCACACACCA CTGGCCATCTTC (GenBank accession #X14720). Amplification parameters were as described above.

The PCR product was purified using the QIAquick™ PCR Purification Kit (Qiagen, Valencia, CA) following the manufacturer's instructions. For the separation of the two CSF1PO alleles, the samples were subjected to agarose gel electrophoresis using a 3% agarose gel (Agarose for the Separation of GeneAmp PCR Products; Perkin Elmer, #N930-2774) in Tris-borate EDTA buffer (TBE, 89 mM Tris, 89 mM borate, 2 mM EDTA pH 8.0) containing ethidium bromide (final concentration 0.5 μg/mL in gel and buffer). PCR products (20 μL), containing about 0.1 to 2.0 μg of DNA, were combined with 3 μL gel loading buffer per lane and run on a mini gel (5.5 by 9 cm, about 3 mm thick) at 200 volts for 1.25 h at room temperature.

Bands comprising the 10 repeat and 16 repeat were cut out of the gel and purified using the QIAquick™ Gel Extraction Kit (Qiagen) following the manufacturer's instructions. Purified band DNA was then either used directly for cycle sequencing, or subjected to another round of amplification and purification as described above.

Cycle sequencing of DNA used PE Applied Biosystem's BigDye Terminator Cycle Sequencing Ready Reaction Kit following the manufacturer's directions (11). Briefly, 20 μL reaction volumes containing 8 μL Reaction Mix, 3.2 pmol of CSF1PO forward or reverse primer, and about 30 to 90 ng of purified DNA were amplified (25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min; then 4°C) using rapid thermal ramp (1°C/s) for temperature changes. Extension products were purified using the DyeEx™ Spin Kit (Qiagen) following manufacturer's directions.

Samples were electrophoresed on the ABI Prism™ 310 Genetic Analyzer following rapid sequencing instructions (11). Briefly, purified samples were dried (SpeedVac; medium, 20 min), resuspended in 18 μL Template Suppression Reagent, denatured, loaded, and run using Module Seq POP6 (1 mL) Rapid E. Results were analyzed using ABI Prism™ DNA Sequencing Software, Version 2.1.1. Sequence Navigator Software, Version 1.01, was used to compare and align sequences and to obtain reverse complements.

Results and Discussion

DNA samples were initially screened for alleles of the STR loci CSF1PO, TPOX, and THO1 by a silver stain detection method. As determined by a visual comparison of sample band positions relative to those of known allelic ladders (not shown), the allele bands for all samples matched those of the standard alleles for TPOX and THO1, while for CSF1PO, bands for all samples except one matched the standards. The CSF1PO variant was found in a heterozygote, which, in addition to a $CSF1PO_{10}$ allele, appeared to contain an allele that was larger than the 15 repeat which is the largest in the CSF1PO ladder.

These observations were confirmed by capillary electrophoresis using the ABI Prism™ 310 Genetic Analyzer. Figure 1 illustrates

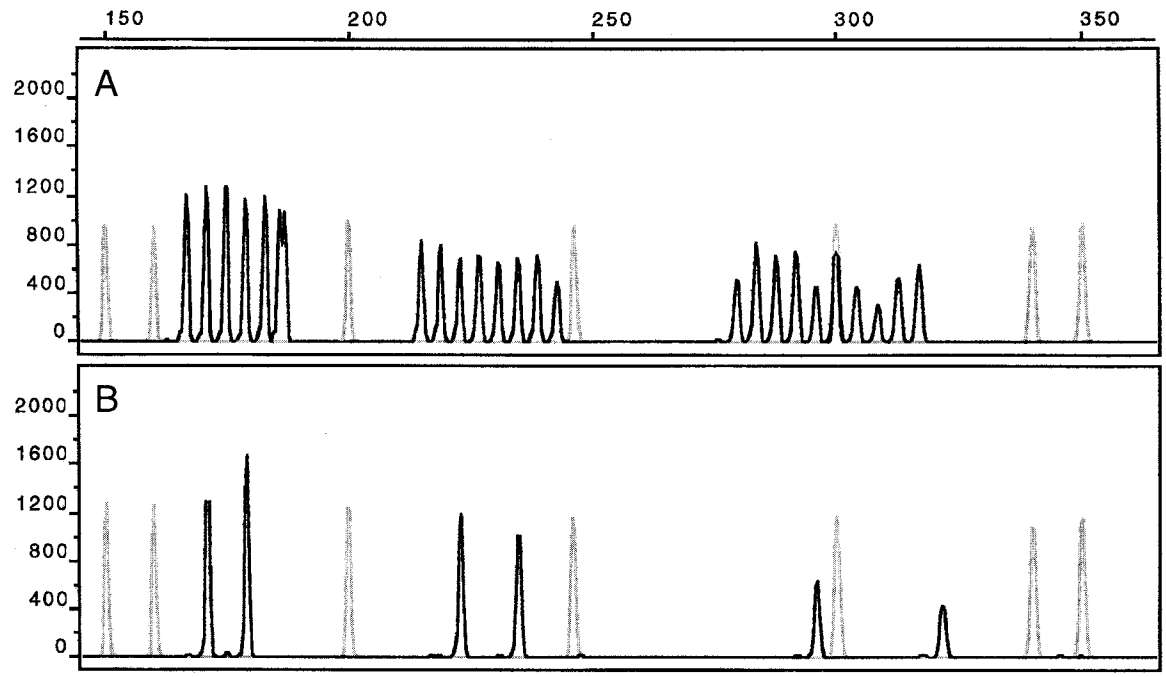

FIG. 1—*GeneScan™ electropherograms of THO1, TPOX, and CSF1PO alleles. In both panels, peaks produced by ROX-500 size standards are shown in gray at 150, 160, 200, 250, 300, 340, and 350 base pairs.*

Panel A: Peaks produced by allelic ladder for AmpFISTR™ Green I alleles; (shown in black) from left to right: THO1 (alleles 5, 6, 7, 8, 9, 9.3, and 10), TPOX (6, 7, 8, 9, 10, 11, 12, and 13) and CSF1PO (6, 7, 8, 9, 10, 11, 12, 13, 14, and 15). (Position 300 contains peaks for both the size standard and the $CSF1PO_{11}$ allele.)

Panel B: Peaks for THO1, TPOX, and CSF1PO alleles (shown in black) obtained using DNA from the individual containing the $CSF1PO_{16}$ *allele, prepared for analysis as described in Materials and Methods.*

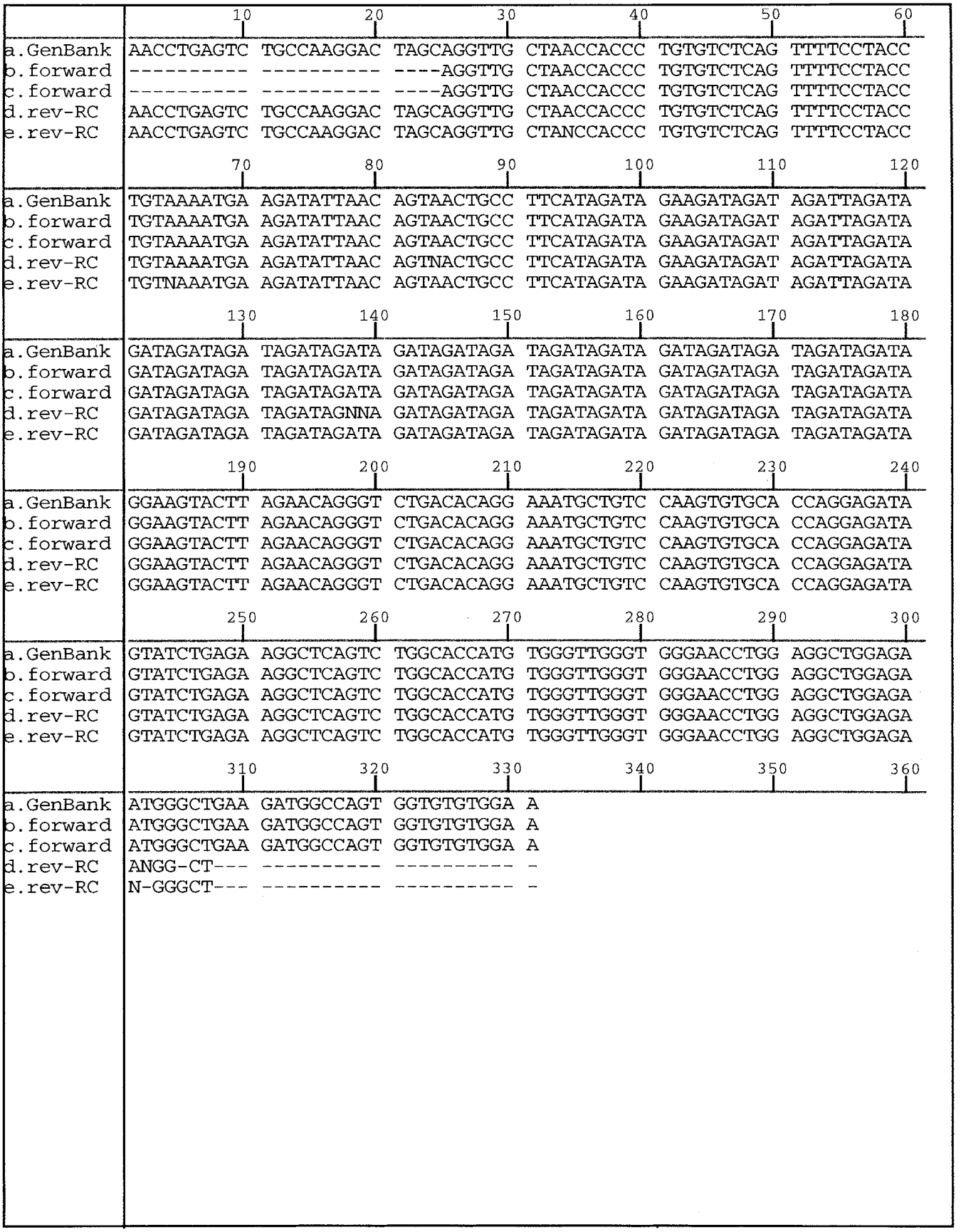

FIG. 2—*Sequence layout for CSF1PO$_{16}$. Sequences were compared and aligned (and, where indicated, their reverse complements obtained) using Sequence Navigator Software, Version 1.01. Row 2a is the Genbank sequence for CSF1PO$_{16}$ as described in Results and Discussion. Sequences illustrated in Rows 2b to 2e were obtained by analysis of the CSF1PO$_{16}$ allele contained in the sample under study, prepared using the BigDye Terminator Cycle Sequencing Ready Reaction Kit. The sequence running from base number 116 through 179 reflects the 16 core repeats.*

two electropherograms showing the peaks produced by allelic ladders for THO1, TPOX, and CSF1PO (left to right respectively, Fig. 1*a*) and those produced by the sample which contains the CSF1PO variant (Fig. 1*b*). The sample is heterozygous for all three STRs and contains peaks that match the 6 and the 8 repeat alleles for THO1, and the 8 and 11 repeat alleles for TPOX. For CSF1PO, one of the peaks matches that of the 10 repeat allele, while the other peak is larger than all alleles present in the ladder. The 15 repeat allele on the allelic ladder had a size of 317.37 ± 0.11 ($n = 5$) bases, while that of the larger allele of the heterozygote was 321.59 ± 0.12 ($n = 8$) bases, indicating an increase in 4 bases or one tetranucleotide repeat. The size of the 10 repeat allele was similar in both the ladder and the heterozygote, i.e., 295.85 ± 0.08 ($n = 5$) and 295.89 ± 0.05 ($n = 8$) bases, respectively.

Because the DNA sample containing the $CSF1PO_{16}$ allele was that of a heterozygote, conclusive sequencing results required the prior separation of the 10 and 16 repeat alleles by agarose gel electrophoresis (see Materials and Methods). When the 16 repeat purified band was used directly in cycle sequencing, no single run of a sample gave definitive results; however, pooling results from multiple runs of the same sample indicated that the sequence contained 16 AGAT repeats and that the portion of the gene that was sequenced was otherwise equivalent to the GenBank sequence for CSF1PO.

The short reads and ambiguity encountered when purified bands were used directly were probably due to low DNA concentrations. By amplifying the purified band DNA prior to cycle sequencing to increase DNA quantity, we were able to obtain unambiguous, long sequences. Results are exactly as one would have predicted from the GenBank sequence for CSF1PO given in Fig. 2, Row a.

Figure 2 shows a layout of sequences for the 16 repeat allele obtained using the forward primer (Rows b and c), and also shows reverse complements of sequences obtained using the reverse primer (Rows d and e). Taken together, these results show that the sequence of the 16 AGAT repeat is exactly as one would have predicted from the GenBank CSF1PO sequence (Row a).

Occasionally, analysis of specimens for various STR loci reveals the presence of an off-ladder allele, that is, an allele banding between alleles of the allelic ladder or an allele banding higher or lower than the upper or lower bands of the ladder (12). Where a novel allele is detected, it is necessary to demonstrate that it is authentic rather than an amplification artifact. Such alleles are, by definition, relatively rare and therefore could have great significance in human identification. Verification of the authenticity of a novel allele is best demonstrated by sequencing forward and backward. Our results demonstrate the existence of an additional CSF1PO allele, a previously unreported size variant, $CSF1PO_{16}$.

Acknowledgments

The authors would like to thank the U.S. Department of Education, MSEIP program for providing support for this work through Grant P120A80031. We would also like to acknowledge the Graduate Research and Training Initiative Program of the City University of New York for providing equipment support for this project.

References

1. Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Gen 1989;44:388–96.
2. Beckman JS, Weber JL. Survey of human and rat microsatellites. Genomics 1992;12:627–31.
3. Edwards A, Civitello A, Hammond HA, Caskey T. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 1991;49:746–56.
4. Edwards A, Hammond HA, Jin L, Caskey T, Chakraborty R. Genetic variation at five trimeric and tetrameric repeat loci in four human population groups. Genomics 1992;12:241–53.
5. Murray JC, Buetow KH, Weber JL, Ludwignsen S, Scherpbier-Heddema T, Manion F, et al. A comprehensive human linkage map with centimorgan density. Science 1994;265:2049–54.
6. National Institute of Standards and Testing. STR Website: www.cstl.nist.gov/div831/strbase.
7. Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R. Evaluation of thirteen STR loci for use in personal identification applications. Am J Hum Genet 1994;55:175–89.
8. Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 1991;10:506.
9. Promega GenePrint™ STR Systems: Technical Manual, TMD004, 1996.
10. AmpFlSTR™ Green I PCR Amplification Kit: User's Manual, PE Applied Biosystems, PN 402944 Rev. A., 1997.
11. ABI Prism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit: Protocol, PE Applied Biosystems, 1998.
12. 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. [published erratum appears in J Forensic Sci 1999; 44(3)] J Forensic Sci 1999;44(1):87–94.

Additional information and reprint requests:
Lawrence Kobilinsky, Ph.D.
Associate Provost
John Jay College of Criminal Justice
The City University of New York
899 Tenth Avenue
New York, NY 10019