

# EXHIBIT 10



# Low Copy Number Typing Still Lacks Robustness and Reliability

## Article

**Bruce Budowle**

Institute of Investigative Genetics, Department of Forensic and Investigative Genetics, University of North Texas Health Science Center, Fort Worth, Texas

2010

*Editor's Note: At the 20th International Symposium on Human Identification, prominent figures in the DNA typing field were asked to respond to a subset of specific questions posed during the low copy number (LCN) session. Bruce Budowle from the University of North Texas Health Science Center shares his response below. Points of view expressed in this article are those of the author and are not necessarily shared by Promega.*

## Introduction

Forensic DNA typing applies a reliable and robust suite of technologies to the analysis of a wide range of biological materials for direct and indirect identity testing. With the PCR and short tandem repeat (STR) loci the methodologies are extremely sensitive and highly discriminating. This technology is robust, yields reliable, reproducible results and is validated and well documented<sup>(1) (2) (3)</sup>. Because STR typing yields reproducible results, the interpretation of the profiles is based on scientifically sound foundations. Due to the success of DNA typing, the envelope of the technology is being pushed to type ever smaller amounts of DNA, even down to the equivalent of DNA contained in a single cell. This technology that attempts to type extremely limited DNA samples is known as low copy number (LCN) typing<sup>(4) (5)</sup>. While it may seem appealing to the forensic scientist to be able to analyze LCN samples, there are a number of robustness and reliability issues that LCN typing has yet to address adequately.

## Definition of LCN Typing

The recommended amount of DNA for STR analysis using manufactured kits typically ranges from 200 picograms (pg) to 2ng<sup>(6) (7) (8)</sup>. Approximately 0.5–1ng is considered the optimum amount for most commercial kits. When the amounts of DNA used for analysis fall below 100–200pg of DNA, stochastic effects are more noticeable and result in increased heterozygote imbalance and allele drop-out. Because of these effects the concept of a stochastic threshold was instituted<sup>(9) (10)</sup>. Typing of DNA samples at or below 100–200pg of DNA has been termed LCN typing<sup>(11) (12) (13)</sup>. As described in the Caddy Report<sup>(14)</sup> LCN typing refers to a particular technique where sensitivity of detection for low-level DNA analyses was increased substantially by increasing the number of cycles during PCR from 28 to 34 cycles<sup>(15)</sup>. However, Budowle *et al.*<sup>(16)</sup> used the same terminology—LCN typing—to describe any method that increases sensitivity of detection (e.g., reduced PCR volume, post-PCR cleanup, increased injection time, use of low-conductivity formamide). While some advocates have sought different names for LCN typing, such as high sensitivity or low template DNA, a different name does

**EXHIBIT 10**

not change the intent for use and performance limitations of the methodology. The intent of all LCN-like assays is to increase the sensitivity of detection of very minute amounts of DNA and use a consensus approach to determine what alleles should be called in a profile. The same issues of robustness and reliability apply to all current low template DNA typing approaches.

*"While it may seem appealing to the forensic scientist to be able to analyze LCN samples, there are a number of robustness and reliability issues that LCN typing has yet to address adequately."*

LCN samples contain so little DNA that they inherently yield nonreproducible results. In the scientific arena reliability has traditionally been based on reproducibility. As such, reliability is a significant challenge to meet for LCN typing. The nonreproducible results are due to stochastic effects during the PCR portion of the analysis. The lack of robustness is manifested as: 1) a greater potential for error (compared with conventional STR typing protocols) due to interpretational challenges caused by allele drop-in, allele drop-out, peak height imbalance, and large stutter peaks; and 2) the probative value of the results may not be estimated reliably. Defining a sample as LCN should be based on at least two parameters: the amount of input DNA and the resultant allele peak heights. For the current forensically validated STR typing systems the amount of input DNA for considering a sample as LCN is less than 100–200pg. Target DNA amount is a good first approximation, but it is a simplified definition of a much more complex process. The other criterion is the peak height of alleles that imply the conditions are such that reliable interpretations may not be possible. However, it is obvious in casework that it cannot be known with LCN typing whether an allele peak is truly from a homozygous individual or one of a pair of alleles from a heterozygous individual (a problem exacerbated with mixtures and allele drop-out). Carragine and Prinz<sup>(17)</sup> reported that "peak heights above 2000RFUs were observed for samples amplified with as little as 6pg of DNA" and "a peak height threshold is not adequate for assignment of a homozygous allele." Thus, allele drop-out is a concern for every LCN sample regardless of peak heights. LCN typing laboratories do not appear to have a threshold for peak height imbalance.

Budowle *et al.*<sup>(18)</sup> recognize the value of peak heights and advocate defining LCN typing based on the values where peak height imbalance becomes exaggerated, and these values are relative to specific assays, kits and methodologies. The value will change with technology and genetic markers typed. Heterozygote peak height imbalance is a better criterion for defining the conditions where stochastic effects occur because it is technique-oriented (defined by in-house validation studies). Not all DNA typing technologies are the same, and thus one should not construe the reliability and robustness of one methodology to be equal to that of another approach nor that the same technique can be applied to all amounts of DNA. Due to the lack of reproducibility inherent in all LCN methods, the interpretation and statistical weight issues are inherent in the reliability of the methods. Therefore, interpretation guidelines or rules should be based on a validation process that is sufficiently comprehensive to apply to the conditions under which LCN typing is performed.

### General Acceptance

In legal proceedings under the adversarial system, admissibility of scientific evidence can be challenged. General acceptance is the primary criterion under the Frye Standard in the United States. If it is deemed by the court that the methodology is generally accepted, then the jury is permitted to hear the evidence. It should be noted that this is a legal concept and not necessarily a scientific concept. Budowle *et al.*<sup>(19)</sup> pointed out that legal acceptance of a scientific method does not make it valid and reliable, and conversely, lack of acceptance in the legal arena does not necessarily make a method unreliable. Leventhal<sup>(20)</sup> described the typical general acceptance criteria under the Frye standard and argued that LCN typing should be considered admissible under

this legal standard. However, he omitted one criterion that was essential to scientists who supported other types of DNA evidence challenged under the Frye standard over the past two decades. General acceptance was supported by validity and reliability testing of the DNA methodology. This concept of reliability is an important distinction, is necessary for scientific general acceptance, should not be ignored by scientists, and has been emphasized by the recent National Academy of Sciences report on the forensic sciences<sup>(21)</sup>. Scientists' opinions should be scientifically based and not just equate meeting a legal threshold of general acceptance as a justification for scientific reliability. If scientists do not advocate reliability, then the gold standard status of forensic DNA typing will likely decay.

A flaw with general acceptance as a metric on scientific validity is that it only requires stating that other labs use LCN typing—not how they use it. Yet, stating solely that others use it is pervasive among the LCN typing laboratories. It is analogous to saying "they have houses in New Zealand, they have houses in the United Kingdom—therefore the house I built in the United States is sound." Soundness is based on how it was built, whether it was built to code, what materials were used, etc. The community does not know how LCN typing is being used by the LCN laboratories, as they proclaim their protocols are proprietary<sup>(22) (23) (24)</sup>. This is an odd position because in numerous cases the LCN practitioners claim that the methods are the same as standard STR typing (except that the sensitivity is increased). Validation studies have been published, and the protocols must be in concert with these studies; otherwise they are not validated. In addition, the interpretation and statistical weight applied should be described openly, especially when the life and liberty of individuals are at stake. However, there are different practices in the interpretation and statistical weight applied by the LCN laboratories, and some of the practices advocated by some LCN scientists would be soundly rejected by other LCN scientists. For example, the Office of the Chief Medical Examiner in New York City (OCME) advocates the use of the probability of exclusion for assessing the weight of evidence and has applied this approach for the past four years<sup>(25) (26) (27)</sup>, while Balding and Buckleton<sup>(28)</sup> reject this approach as inappropriate and untenable for addressing allele drop-out.

A more troubling problem is that many of the recommendations for LCN typing in the scientific literature are not those that are practiced by the LCN laboratories. As examples:

- Caragine *et al.*<sup>(29)</sup> reported based on their validation studies that "plus 4 stutter was rare but was observed particularly with the 100pg samples." Since most LCN sample replicates are typically 33pg or less, the OCME data would support that a peak in the plus 4 stutter position is more likely a real allele and not stutter. However, in casework there are a number of examples of plus 4 stutter position peaks being filtered out. By not including these "alleles" the consequences are that the statistical weight of the evidence (as practiced by the OCME) will be overstated.
- The often-cited paper by Gill *et al.*<sup>(30)</sup> described an approach for combining the probabilities of allele drop-out, allele drop-in and stutter into the weight of LCN evidence. A decade later these approaches have not been implemented by the LCN laboratories.
- Gill and Buckleton<sup>(31) (32)</sup> strongly criticize the use of thresholds for identifying regions where allele drop-out may occur in part because there is no absolute cut-off and its use may unfairly bias against a defendant. Yet in their more recent "Review of DNA Reporting Practices by Victoria Police Forensic Services Division" (April 2010) they (and one other reviewer) to the contrary recommended the use of a 250RFU stochastic threshold (Appendix 1, No. 5)<sup>(33)</sup>. These conflicting positions between their publications and recommendations for practice are difficult to reconcile.

As can be seen from just the above few examples, one may have difficulties in supporting general acceptance and peer review based on the scientific literature. What is purported in the scientific literature does not appear to be what is done in practice. Access to protocols, examples of casework and evaluations of validation studies are needed. Butler and Hill, National Institute of Standards and Technology<sup>(34) (35)</sup>, attempted to justify the validity of LCN typing by counting the number of publication citations on LCN typing. At face value the number of citations

might seem to support the validity of LCN typing, but it is far too superficial an assessment. The Gill *et al.*<sup>(36)</sup> paper, for example, discussed above, constituted almost one-third of the citations<sup>(37)</sup>, and yet the statistical approaches they recommended have not been used even a decade later. There is a significant and problematic discordance between practice, publication recommendations and reliance, which calls into question the reliability of the interpretation of LCN analysis results. It may seem appealing to consider some abstract proposal for interpretation and statistical weight, but when it comes to implementation, there are practical constraints that have yet to be disclosed or considered.

Leventhal<sup>(38)</sup> attempted to justify that a qualitative statement of, for example, "cannot exclude" (and as practiced by the OCME) is acceptable because other forensic disciplines carry out similar practices. The OCME provide the interpretation that the suspect "cannot be excluded" as a contributor of the evidence when a few of his/her alleles are not observed in the evidence; there is no statistical analysis accompanying the interpretation. The other forensic disciplines are under significant scrutiny, and part of the criticism is that they do not quantify the result of a comparison<sup>(39)</sup>. If the criticism is warranted (and to some degree that may be true), then there is no justification for LCN typing interpretation to follow similar practices. In other words, Leventhal's and the OCME's position suggests that whether it is done correctly or not is moot; it only matters that others do it. Leventhal's suggestion that other comparative disciplines do not quantify evidentiary results is not necessarily correct. Most pattern-comparison disciplines attempt to convey the weight of the evidence, although qualitatively. Fingerprints, toolmarks and handwriting explain to some degree (varying from individualization to degrees of likelihood) the strength of the evidence. In contrast, the practice of presenting evidence solely as "cannot exclude" provides no guidance even qualitatively of the significance of the results. This is a very different practice than most pattern-comparison disciplines.

The most serious issue of providing no statistical weight with an interpretation of "cannot exclude" is that there is inherent bias built into this practice<sup>(40)</sup>. There are at least two hypotheses to consider: 1) the missing alleles are due to allele drop-out, and the suspect cannot be excluded; and 2) the alleles are not in the mixture sample, and the suspect cannot be a contributor of the evidence. The "cannot exclude" approach basically supports the first hypothesis and discards discordant data. Such bias should not be advocated by scientists.

#### Advice

Forensic DNA typing has been labeled the gold standard of forensic science<sup>(41) (42)</sup>. The methodology has been demonstrated to be robust, reproducible and reliable. In contrast, LCN typing has not been well developed and applied appropriately. Moreover, the validation studies do not comport with protocols, assumptions for calculating the weight of the evidence are in question, and the scientific literature recommendations are not necessarily in concert with practices. It would be a shame to abandon the standards in place for forensic DNA typing just to push the envelope with LCN typing. Assisting in solving crime with DNA typing is our desire and our responsibility. However, we should pursue forensic analyses by employing robust and reliable technologies so that we can have the greatest confidence in the reliability of our results. Substantially more work is needed before the conditions are known under which LCN typing should be used for reliable identification purposes.

#### References

1. Collins, P.J. *et al.* (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: The AmpF/STR® Identifier® PCR Amplification Kit. *J. Forensic Sci.* **49**, 1265–77.
2. Cotton, E.A. *et al.* (2000) Validation of the AMPF/STR® SGM Plus™ system for use in forensic casework. *Forensic Sci. Int.* **112**, 151–61.

3. Krenke, B.E. *et al.* (2002) Validation of a 16-locus fluorescent multiplex system. *J. Forensic Sci.* **47**, 773–85.
4. Findlay, I. *et al.* (1997) DNA fingerprinting from single cells. *Nature* **389**, 555–6.
5. Gill, P. *et al.* (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* **112**, 17–40.
6. Collins, P.J. *et al.* (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: The AmpFISTR® Identifier® PCR Amplification Kit. *J. Forensic Sci.* **49**, 1265–77.
7. Cotton, E.A. *et al.* (2000) Validation of the AMPFISTR® SGM Plus™ system for use in forensic casework. *Forensic Sci. Int.* **112**, 151–61.
8. Krenke, B.E. *et al.* (2002) Validation of a 16-locus fluorescent multiplex system. *J. Forensic Sci.* **47**, 773–85.
9. Moretti, T.R. *et al.* (2001) Validation of short tandem repeats (STRs) for forensic usage: Performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J. Forensic Sci.* **46**, 647–60.
10. Moretti, T.R. *et al.* (2001) Validation of STR typing by capillary electrophoresis. *J. Forensic Sci.* **46**, 661–76.
11. Gill, P. *et al.* (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* **112**, 17–40.
12. Budowle, B. *et al.* (2001) Low copy number — consideration and caution. *Proceedings From the Twelfth International Symposium on Human Identification 2001*
13. Caddy, B., Taylor, G.R. and Linacre, A. (2008) A review of the science of low template DNA analysis.
14. Caddy, B., Taylor, G.R. and Linacre, A. (2008) A review of the science of low template DNA analysis.
15. Gill, P. *et al.* (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* **112**, 17–40.
16. Budowle, B. *et al.* (2001) Low copy number — consideration and caution. *Proceedings From the Twelfth International Symposium on Human Identification 2001*
17. Caragine, T. and Prinz, M. (2011) Comment on the paper: "Low copy number typing has yet to achieve 'general acceptance'". *Forensic Sci. Int. Genet.* **5**, 3–4.
18. Budowle, B., Eisenberg, A.J. and van Daal, A. (2009) Validity of low copy number typing and applications to forensic science. *Croat. Med. J.* **50**, 207–17.
19. Budowle, B., Eisenberg, A. and van Daal, A. (2010) Concerns about low copy number typing. *Forensic Sci. Int. Genet.* (epub ahead of print).
20. Leventhal, B. (2009) Presentation, LCN Panel. The Twentieth International Symposium on Human Identification, Las Vegas, Nevada.
21. (2009) Strengthening Forensic Science in the United States: A Path Forward, National Academy of Sciences National Academies Press, 1–254.

22. Buckleton, J. and Gill, P. (2010) Further comment on Budowle et al.: Low copy number has yet to achieve "general acceptance"; FSI Genetics supplementary series 2 (2009) 551-552. *Forensic Sci. Int. Genet.* (epub ahead of print).
23. Budowle, B., Eisenberg, A. and van Daal, A. (2009) Low copy number typing has yet to achieve "general acceptance". *Forensic Sci. Int. Genet. Suppl. Series 2*, 551-2.
24. Budowle, B., Chakraborty, R. and van Daal, A. (2010) Authors' response. *J. Forensic Sci.* **55**, 269-72.
25. Budowle, B., Eisenberg, A. and van Daal, A. (2010) Concerns about low copy number typing. *Forensic Sci. Int. Genet.* (epub ahead of print).
26. Budowle, B., Eisenberg, A. and van Daal, A. (2009) Low copy number typing has yet to achieve "general acceptance". *Forensic Sci. Int. Genet. Suppl. Series 2*, 551-2.
27. Caragine, T. et al. (2009) Validation of testing and interpretation protocols for low template DNA samples using AmpF/STR® Identifiler®. *Croat. Med. J.* **50**, 250-67.
28. Balding, D. and Buckleton, J. (2009) Interpreting low template DNA profiles. *Forensic Sci. Int.* **4**, 1-10.
29. Caragine, T. et al. (2009) Validation of testing and interpretation protocols for low template DNA samples using AmpF/STR® Identifiler®. *Croat. Med. J.* **50**, 250-67.
30. Gill, P. et al. (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* **112**, 17-40.
31. Buckleton, J. and Gill, P. (2010) Further comment on Budowle et al.: Low copy number has yet to achieve "general acceptance"; FSI Genetics supplementary series 2 (2009) 551-552. *Forensic Sci. Int. Genet.* (epub ahead of print).
32. Gill, P. and Buckleton, J. (2009) Commentary on: Budowle B, Onorato AJ, Callaghan TF, Della Manna A, Gross AM, Guerrieri RA, Luttmann JC, McClure DL. Mixture interpretation: Defining the relevant features for guidelines for the assessment of mixed DNA profiles in forensic casework. *J. Forensic Sci.* **54**, 810-21.
33. Fraser, J., Buckleton, J. and Gill, P. (2010) Review of DNA reporting practices by Victoria Police Forensic Services Division.
34. Butler, J. (2009) Presentation, LCN panel. The Twentieth International Symposium on Human Identification, Las Vegas, Nevada.
35. Butler, J. and Hill, C.R. (2010) Scientific issues with analysis of low amounts of DNA. *Profiles in DNA* **13** (1).
36. Gill, P. et al. (2000) An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* **112**, 17-40.
37. Butler, J. and Hill, C.R. (2010) Scientific issues with analysis of low amounts of DNA. *Profiles in DNA* **13** (1).
38. Leventhal, B. (2010) Low copy number analysis from a legal perspective. *Profiles in DNA* **13**(1).
39. (2009) Strengthening Forensic Science in the United States: A Path Forward, National Academy of Sciences National Academies Press, 1-254.
40. Budowle, B., Eisenberg, A. and van Daal, A. (2010) Concerns about low copy number typing. *Forensic Sci. Int. Genet.* (epub ahead of print).

41. (2009) Strengthening Forensic Science in the United States: A Path Forward, National Academy of Sciences National Academies Press, 1–254.
42. Saks, M.J. and Koehler, J.J. (2005) The coming paradigm shift in forensic identification science. *Science* **309**, 892–5.

#### HOW TO CITE THIS ARTICLE

Scientific Style and Format, 7th edition, 2006

Budowle, B. Low Copy Number Typing Still Lacks Robustness and Reliability. [Internet] 2010. [cited: year, month, date]. Available from: <http://www.promega.com/resources/articles/profiles-in-dna/low-copy-number-typing-still-lacks-robustness-and-reliability/>

American Medical Association, Manual of Style, 10th edition, 2007

Budowle, B. Low Copy Number Typing Still Lacks Robustness and Reliability. Promega Corporation Web site. <http://www.promega.com/resources/articles/profiles-in-dna/low-copy-number-typing-still-lacks-robustness-and-reliability/> Updated 2010. Accessed Month Day, Year.

Contribution of an article to *Profiles in DNA* does not constitute an endorsement of Promega products.