J Forensic Sci, Jan. 2004, Vol. 49, No. 1
Paper ID JFS2003001_491
Available online at: www.astm.org

TECHNICAL NOTE

Jason R. Gilder,[1] *M.S.; Simon Ford,*[2] *Ph.D.; Travis E. Doom,*[3] *Ph.D.; Michael L. Raymer,*[3] *Ph.D.; and Dan E. Krane,*[4] *Ph.D.*

Systematic Differences in Electropherogram Peak Heights Reported by Different Versions of the GeneScan® Software

ABSTRACT: DNA profiling using STRs on the 310 and 3100 Genetic Analyzers routinely generates electropherograms that are analyzed with the GeneScan® software available from the instrument's manufacturer, Applied Biosystems. Users have been able to choose from three different smoothing options that have been known to result in significant differences in the peak heights that are reported. Improvements in the underlying algorithm of the most recent version of the software also result in significant and somewhat predictable differences in peak height values. Laboratories that have performed validation studies using older versions of GeneScan® should either reanalyze the data generated in those validation studies with the newest version of the software or otherwise take into consideration the systematically higher peak height values obtained as they begin following the recommendation of the manufacturer and use the new algorithm.

KEYWORDS: forensic science, DNA typing, GeneScan®, peak heights, smoothing, validation

Short tandem repeat (STR) sequences have increasingly become the genetic markers of choice for the purposes of human identification by crime laboratories (1,2). STR typing typically involves a polymerase chain reaction (PCR) amplification step followed by size fractionation of the resulting products and fluorescent signal detection and processing. While alternatives are available, separation of alleles from different STR loci is most commonly performed in the United States with Perkin Elmer-Applied Biosystems capillary electrophoresis equipment such as the 310 and 3100 Genetic Analyzers. In such instances, the data are automatically analyzed by Applied Biosystems, Inc.'s (ABI) Prism® GeneScan® and GenoTyper® software. GeneScan® uses internal size standards that are co-injected with each sample to estimate the size of the PCR products and also reports the relative amount of material observed in terms of peak heights and areas (expressed in relative fluorescent units or RFUs). This information is compiled into the tabular data that are then imported into GenoTyper® . GenoTyper® then looks for correspondence in size between the peaks present in samples and in allelic ladders that contain DNA fragments representing the alleles most commonly encountered in most human populations.

This general approach has become popular for forensic analyses because of its numerous advantages over other methods of DNA analysis (2,3) and is exclusively used for the Combined DNA Index System (http://www.fbi.gov/hq/lab/codis/index1.htm) that includes DNA profiles generated from over one million convicted sexual offenders and other felons. Peak height (or, for some laboratories, area) is the primary objective factor used to distinguish meaningful peaks from background noise. Individual testing laboratories conduct validation studies to determine the minimum RFU value that a peak must exceed in order to be reported based on the performance of the test kit and analytical platform in their laboratory environment. Minimum reportable peak height (MRPH) values vary substantially among U.S. laboratories, ranging from approximately 50 to 200 RFU.

The original versions of GeneScan® and GenoTyper® were available only for use with the Apple Macintosh operating system. In June of 2002, Applied Biosystems released a new version of the GeneScan® Analysis Software (v. 3.7.1) that is designed for use on Windows NT platforms (4). Several features of the new Windows NT version of GeneScan® have been enhanced relative to the final supported version of the Macintosh-based software (v. 3.1.2) including: smoothing options, minimum peak half width calculation, increasing the robustness of the size caller, and improving the baselining function to eliminate negative peak area. The Windows NT version also has several new features that allow users to modify parameters such as: polynomial degree, peak window size, slope threshold for peak start/end, and window size (4). All of these improvements imply potential differences in outputs reported between the versions, particularly with regard to the reporting of peak heights.

One major difference between the Macintosh versions of GeneScan® and the Windows NT version of GeneScan® is that in the Macintosh version smoothing is applied prior to exporting

[1] Forensic Bioinformatics, Inc., 2850 Presidential Drive, Suite 150, Fairborn, OH.
[2] Lexigen Science and Law Consultants, Inc., 2261 Market Street, No. 302, San Francisco, CA.
[3] Department of Computer Science and Engineering, Wright State University, Dayton, OH.
[4] Department of Biological Sciences, Wright State University, Dayton, OH.
Received 1 Jan. 2003; and in revised form 10 Aug. 2003 and 16 Sept. 2003; accepted 19 Sept. 2003; published 23 Dec. 2003.

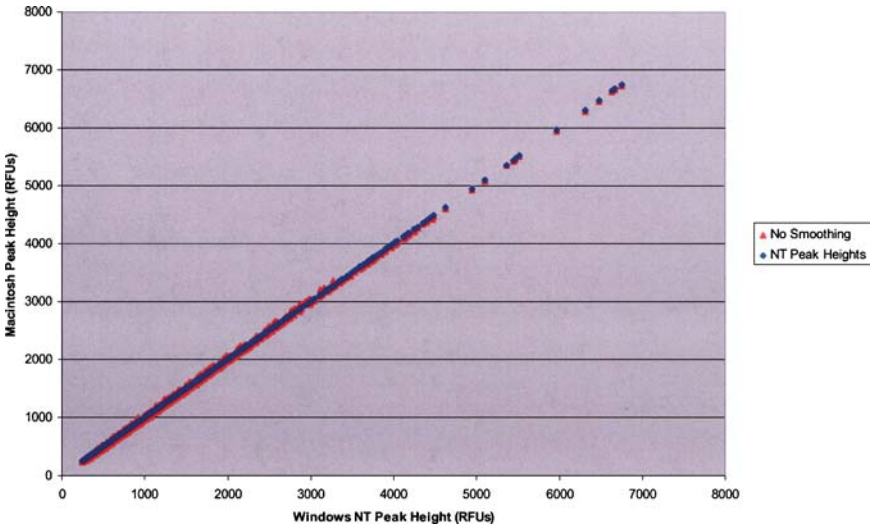

FIG. 1—*Macintosh GeneScan Analysis v. 3.1.2 versus Windows NT GeneScan Analysis v. 3.7.1.*

to tabular data, whereas in the NT version smoothing is applied after exporting to tabular data. Therefore, choice of smoothing type directly affects all subsequent stages of analysis in the Macintosh version of the software; however, this is not the case in the Windows NT version (4). GeneScan's® manufacturer is recommending that DNA testing laboratories update to the NT version of the software. As GeneScan® is among the most widely used software packages relied upon by DNA testing laboratories, it is important that forensic scientists are aware of the significance of the systematic differences between these software versions so that they may be taken into appropriate consideration when upgrading.

Methods

A total of 2,979 peaks associated with 41 ABI 310 allelic ladders run during routine casework by the Tucson Police Department Crime Laboratory (26 ladders) and Federal Bureau of Investigation DNA Analysis Unit (15 ladders) are analyzed in this study. Allelic ladder peaks contain all the commonly observed peaks for each locus, and thus peaks spanning the entire size range typically considered in genotyping experiments are considered. All allelic ladders are analyzed once with the Windows NT version of GeneScan® (v. 3.7.1) and three times using the most recent Macintosh version of GeneScan® (v. 3.1.2)—once with each of the three smoothing options available in the Macintosh version (none, light, and heavy). In order to minimize the impact of stochastic effects and other noise in the data, only peaks with heights reported as being greater than 200 RFUs in all four runs of the software are considered. The analysis parameters on the Windows NT analysis use the default settings for minimum peak half width, polynomial degree, peak window size, and baseline window size detailed in the ABI user bulletin (4). Microsoft Excel is used for standard statistical analyses and plotting.

Results and Discussion

Historically, the Macintosh-based versions of the GeneScan® software used to analyze the raw electronic data generated for most forensic STR-DNA profiling have given analysts the ability to choose between one of three smoothing options for the presentation of electropherogram peaks: none, light, and heavy (4). The "none" option performs a minimal (but measurable) amount of smoothing. As the amount of smoothing is increased, rough edges are smoothed out, making the peaks appear more uniform. As the amount of smoothing is increased, reported peak heights become increasingly lower. In contrast, the NT-based version of GeneScan® captures peak height information from electropherograms before any smoothing is performed.

Figures 1, 2, and 3 present the peak height values observed for the 2,979 comparable peaks in this study for the NT-based version and the Macintosh-based version under each of the three smoothing options. Table 1 presents the maximum, minimum, and average of these peak height values.

Differences between the peak heights reported by GeneScan® v. 3.7.1 and v. 3.1.2 are least when the "no smoothing" option was chosen for the Macintosh-based analysis. Because peak heights are captured only after a small amount of smoothing has occurred even when the "no smoothing" option is chosen on the Macintosh-based versions of the software, values for the same peaks are systematically less (−2.12%, on average) than those reported by the NT-based version (Table 2). More than 95% of the Macintosh values fall within two standard deviations (+2.56% and −6.8%) of the mean Windows NT values. Similar differences are observed throughout the range of peak height sizes (i.e., for the 1,490 smallest and for the 1,490 largest peak heights) (see Fig. 1). Some percentage peak height differences are more extreme, Macintosh-based values ranging from extremes of +9.8% to −12.1% of the NT-based values are observed from some individual peaks. Peak areas reported by both versions of the software are generally equivalent regardless of smoothing option chosen (see Table 2). As increased smoothing lowers reported peak heights, it also increases peak width proportionally. Thus, the reported peak areas show no significant systematic difference.

The most extreme differences in peak heights under the "no smoothing" option are generally associated with low peak heights (<500 RFU). Small changes in absolute RFU level between peak heights reported by the two versions have the potential to be observed as an extreme relative (percent) difference for small peaks. Most of the observed extreme percent differences appear to be the result of small changes in absolute RFU calls due to changes in the calculated sample baseline. Improvements in the baselining algorithm used in the NT-based version of GeneScan® are believed to be responsible for the differences in baseline calls. Differences in baseline calls between the two versions of the software are generally

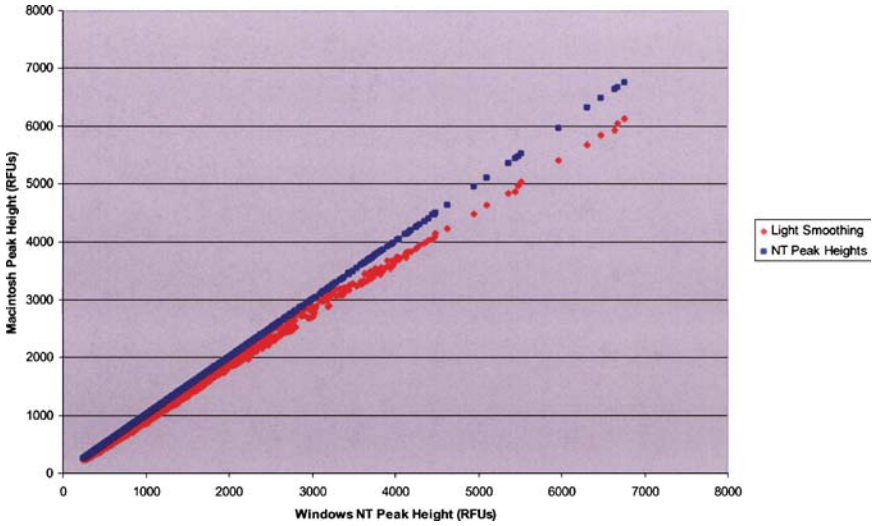

FIG. 2—*Macintosh GeneScan Analysis v. 3.1.2 versus Windows NT GeneScan Analysis v. 3.7.1.*

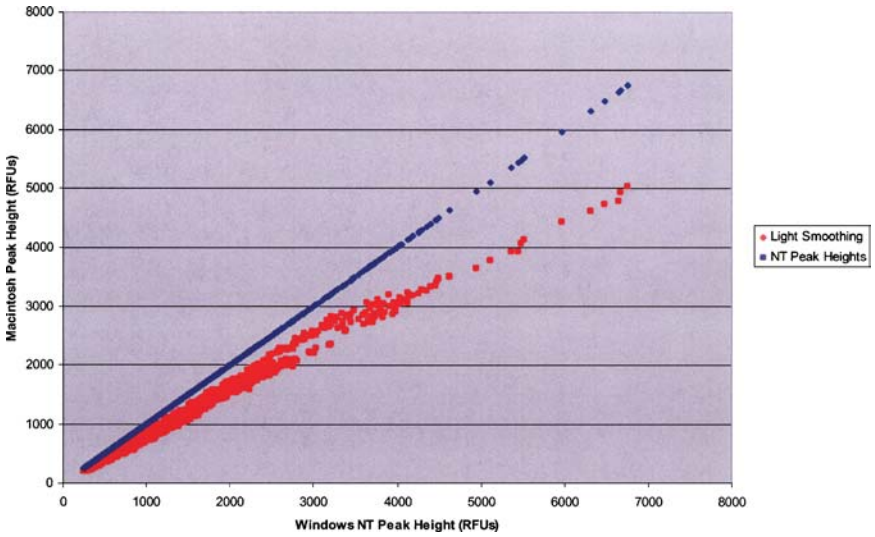

FIG. 3—*Macintosh GeneScan Analysis v. 3.1.2 versus Windows NT GeneScan Analysis v. 3.7.1.*

TABLE 1—*Peak height and peak area values from directly comparable allelic ladder peaks included in this study.**

Software/ Smoothing	Peak Heights				Peak Areas			
	Ave.	Min.	Max.	Std. Dev.	Ave.	Min.	Max.	Std. Dev.
Windows NT	1051	250	6753	785	7610	1839	41206	5629
Macintosh/none	1036	242	6734	786	7256	1674	41863	5695
Macintosh/light	964	231	6123	730	7302	1656	41825	5762
Macintosh/heavy	820	200	5030	622	7253	1591	41745	5765

* Values for peaks are shown in RFUs. $N = 2{,}979$.

TABLE 2—*Differences between peak heights and areas reported by GeneScan® v. 3.1.2 with each of the three smoothing options relative to those reported for the same 2,979 peaks by GeneScan® v. 3.7.1.**

Smoothing	Peak Heights				Peak Areas			
	Ave.	Std. Dev.	Slope	r^2	Ave.	Std. Dev.	Slope	r^2
None	−2.12	2.34	1.001	0.9997	−6.63	7.08	1.008	0.9962
Light	−8.97	3.01	0.929	0.9990	−6.23	6.93	1.020	0.9963
Heavy	−22.73	4.48	0.788	0.9955	−7.17	7.30	1.020	0.9962

* Average percent differences are shown in terms of RFUs; the slope and correlation coefficients (r^2) are for the best-fit linear regression.

small (on the order of tens of RFUs) and usually only affect small regions (three to five peaks) of the samples in which this behavior is noted.

Most crime laboratories have used the "light smoothing" option of the Macintosh-based version of the software in their validation studies and in their routine casework. The peak heights reported by the Macintosh-based version with the light smoothing option differ by an average of −8.97% with those reported by the NT-based software for the same peaks (Table 2). Again, differences in the baseline calls produce a small number of extreme differences (Macintosh-based values ranging from extremes of +5.94 to −9.4% of the NT-based values for some individual peaks), but approximately 94% of the Macintosh values are within two standard deviations (−2.97 and

TABLE 3—*Statistical tests of differences observed between the Windows NT version of GeneScan® v. 3.7.1 and each of the three smoothing options used with the Macintosh version of GeneScan® v. 3.2.1.**

Statistical Test	No Smoothing	Light Smoothing	Heavy Smoothing
F-Test	0.937	7.90×10^{-5}	7.02×10^{-37}
2-Tailed Paired *t*-Test	0	0	0

* All values shown are for comparisons to the Windows NT values.

−14.97%) of the mean NT-based values. Similar differences were observed throughout the range of peak height sizes (see Fig. 2).

Differences between the peak heights reported by GeneScan® v. 3.7.1 and v. 3.1.2 are greatest with the "heavy smoothing." The heavily smoothed Macintosh-based peak heights are invariably less (−22.7%, on average) than those reported by the NT-based version (see Table 2). Again, differences in baseline calls produce a small number of extreme differences (Macintosh values ranging from extremes of −6.2 to −35.4% relative to the Windows NT values for some individual peaks), but more than 95% of the Macintosh values were within two standard deviations (−13.8 and −31.7%) of the mean Windows NT values. Similar differences are observed throughout the range of peak height sizes (see Fig. 3).

The specific tests used to provide statistical confidence to our observations of varying means and deviations are summarized in Table 3. *F*-tests performed on the experimental data indicate that the measured variances between the NT-based analysis and the Macintosh-based analysis performed under light/heavy smoothing are statistically significant. Two-tailed paired *t*-tests performed on the sample data indicate that the number of data points included in this study are sufficient to give strong confidence; the results are statistically unlikely to be the results of an unusual population of peaks in our experimental database of peak heights.

Use of the Macintosh-based version of GeneScan® with the "light smoothing" option currently used by many DNA testing laboratories produces results that are significantly dissimilar from the NT-based version's analysis. *F*-tests support the observation that the default Windows NT analysis shares a similar variance with the Macintosh "no smoothing" option and produces significantly dissimilar variances when compared to other smoothing options. In all three comparisons, the means differ significantly, producing two-tailed paired *t*-tests that indicate an exceedingly low probability of the means of these peak height data being equal under any of the smoothing options (see Table 3).

Given these systematic differences, forensic laboratories may wish to reassess their MRPH when changing to the new supported NT-based version of GeneScan, particularly if they do not currently use the "no smoothing" option. Consider, for example, a laboratory that adopted a 150 RFU cut-off using light smoothing for casework analyzed using the Macintosh version of GeneScan®. A peak with a height of 140 RFU that might have been considered questionable or unreliable when analyzed with the older versions of GeneScan® is likely to exceed the MRPH when analyzed with the Windows NT version even though light smoothing was used in both analyses. Failure to take into consideration the differences inherent to the different versions of the analysis software might reduce the laboratory's accuracy in distinguishing true signals from noise.

Conclusions

It is good practice for labs using the ABI 310 and 3100 machines to follow the manufacturer's recommendations and upgrade to the most current version of the software available. The height of electropherogram peaks reported by GeneScan® v. 3.7.1 are consistently greater than those reported by previous versions of this software. While the differences are greatest when users choose the "heavy smoothing" option, the "light smoothing" option that has been used by many crime laboratories in their validation studies and casework also results in differences that are statistically significant. The trends for the differences observed with all levels of smoothing are fairly consistent and might allow a direct conversion to be made between peak height values obtained with the Macintosh-based and the enhanced Windows NT-based versions of the software if individual laboratories determine that 95% confidence intervals associated with converted values are acceptable. Alternatively, laboratories that follow the software provider's recommendations and upgrade to GeneScan® v. 3.7 (or later) could reanalyze the electropherograms associated with their validation studies, particularly those associated with establishing minimum peak height thresholds and those relying upon peak height comparisons within or between samples.

References

1. Edwards A, Hammond HA, Lin J, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 1992;12:241–53. [PubMed]
2. Frégeau CJ, Fourney RM. DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification. BioTechniques 1993;15:100–19. [PubMed]
3. Kimpton CP, Gill P, Walton A, Millican ES, Adams M. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods and Applications 1993;3:13–22. [PubMed]
4. User Bulletin. ABI Prism® GeneScan® Analysis Software for the Windows NT Operating System, June, 2002.