

Graphical Tools for RFLP Measurement Quality Assurance: Single-Locus Charts*

REFERENCE: Duewer DL, Liu H-K, Reeder DJ. Graphical tools for RFLP measurement quality assurance: single-locus charts. *J Forensic Sci* 1999;44(5):969-977.

ABSTRACT: Forensic restriction fragment length polymorphism analyses typically provide two results for each sample, one result for each unique allele, at each genetic locus probed. In collaboration with the member laboratories of the Technical Working Group for DNA Analysis Methods, we have developed graphical techniques that compactly summarize even large numbers of such paired measurements. This paper provides a detailed description of the basic tool, a modified bivariate control chart for data from one sample at one locus. We demonstrate how various modifications and combinations of these "single-locus charts" can be used for within- and among-laboratory quality assurance activities.

KEYWORDS: forensic science, DNA typing, data analysis, graphical analysis, interlaboratory comparison, statistical intervals, restriction fragment length polymorphism, quality assurance

The National Institute of Standards and Technology (NIST) of the U.S. Department of Commerce collaborates with member laboratories of the Technical Working Group for DNA Analysis Methods (TWGDAM) in the study and documentation of restriction fragment length polymorphism (RFLP) measurement technologies (1-7). From 1992 to 1996, more than 40 U.S. and Canadian forensic laboratories provided over 100,000 RFLP data from cell line control and other well characterized materials. These data were collected during routine casework, population and offender database development, quality assurance activities, proficiency demonstrations, and validation studies.

We have developed several graphical analysis techniques for compact, yet quantitative, summary of large numbers of related RFLP data. These graphics are adaptations of the familiar scatterplot, displaying the bivariate structure of typical RFLP data. Two DNA fragments (bands) are expected for every sample at each forensically important polymorphic genetic locus, although single- and multiple-banded patterns do occur. The approximate molecular weight (size) of these bands are measured as $\{x_1, x_2\}$ ordered pairs, where " x_1 " is the base pair (bp) size of the larger ("high") band and " x_2 " the bp size of the smaller ("low") band.

¹ Analytical Chemistry Division, Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

² Statistical Engineering Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

³ Biotechnology Division, Chemical Science and Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

* This work was supported in part by the Federal Bureau of Investigation and by the National Institute of Justice, U.S. Department of Justice.

Received 15 May 1998; and in revised form 27 Oct. 1998; accepted 25 Jan. 1999.

For any DNA sample that has been repeatedly analyzed, whether within-laboratory cell line control or among-laboratory proficiency demonstration unknown, plotting x_1 against x_2 efficiently captures the information present in the data. With appropriate centering, scaling, representation of various tolerance (also called "acceptance") criteria, and data labeling, these single locus $\{x_1, x_2\}$ scatterplots document measurement performance in a compact yet rigorously quantitative manner.

Applied to single-laboratory cell line control data, these plots can be used as bivariate control charts to confirm the acceptability of particular data as well as document measurement performance over time. The Federal Bureau of Investigation's (FBI's) Combined DNA Index System (CODIS) incorporates a version of these charts (8).

Applied to multi-laboratory proficiency demonstration data, these charts unambiguously display the extent of measurement concordance among the participants and permit rapid identification of unacceptably discordant data. The Forensic Identity Proficiency Testing Program (College of American Pathologists, 325 Waukegan Road, Northfield, IL 60093) now includes these charts as part of their test documentation (9). Collaborative Testing Services, Inc. (PO Box 1049, 340 Herndon Parkway, Herndon, VA 20170) has developed similar bivariate scattergrams and intends to soon use them in their DNA Profiling Forensic Testing Program (10,11).

We present here our procedures for preparing what we term single-locus charts (SLC) for RFLP measurements. The essential components of the SLC can be constructed with pen and graph paper; all aspects are easily implemented as spreadsheet macro-programs. We hope that, however they are constructed, interested forensic laboratories will find these graphical techniques useful tools for measurement quality assurance and performance documentation.

Methods and Materials

Demonstration Data

The within-laboratory K562 cell line control data used here are representative of casework and population studies conducted from 1990 to 1993. The data were provided by 15 different laboratories. All 15 laboratories reported data for loci D1S7 and D2S44; most laboratories also reported values for loci D4S139, D10S28, and/or D17S79.

The among-laboratory data used were collected as part of the FBI supervised post-graduate training ("Postgrad") exercises, during the period 1990 to 1992. These Postgrad data were reported by 19 different non-FBI analysts. The data consist of band sizes at loci D1S7, D2S44, D4S139, and D17S79 for four different

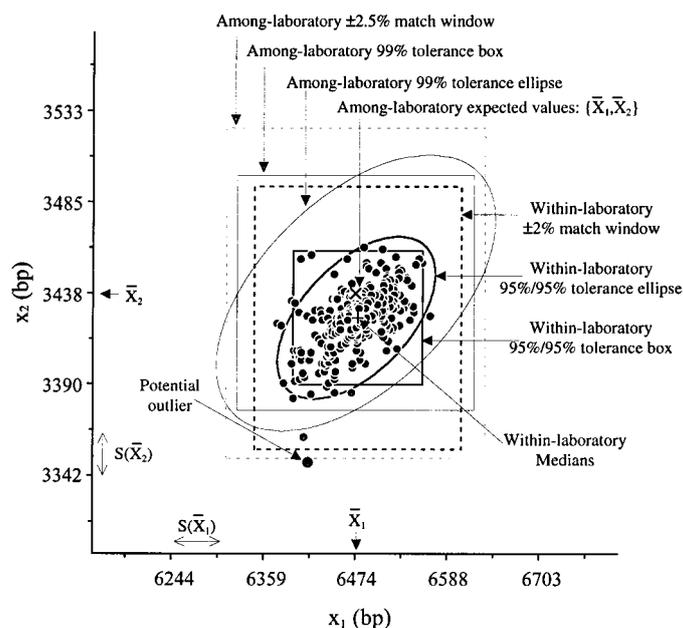


FIG. 1—Single-locus chart (SLC), displaying and labeling all components. These locus D4S139 K562 cell line control data were collected by laboratory A14 from 1990 to 1993.

blood stain sources—DA, JA, KH, and SC as well as the K562 control.

The above data were published as Supplementary Material to Ref 2.

Computation

All graphics were generated using commercial spreadsheet software.

Design and Construction

Construction of a Single-locus Chart (SLC)

Figure 1 displays an SLC for locus D4S139 K562 cell line control data, showing within- and among-laboratory data tolerance criteria and labeling all the graphical elements. (These data were provided by the anonymous laboratory labeled “A14” in Ref 2.) Figure 2 details the major stages of the chart construction: raw scatterplot, scaled and centered scatterplot, among-laboratory tolerance criteria, and within-laboratory tolerance criteria.

Figure 2A is a basic scatterplot, showing just “high” band values (x_1) plotted versus their paired “low” band (x_2) with default settings for the axes. The beginning and end points of the x_1 and x_2 axes are dictated by the minimum and maximum values of the data.

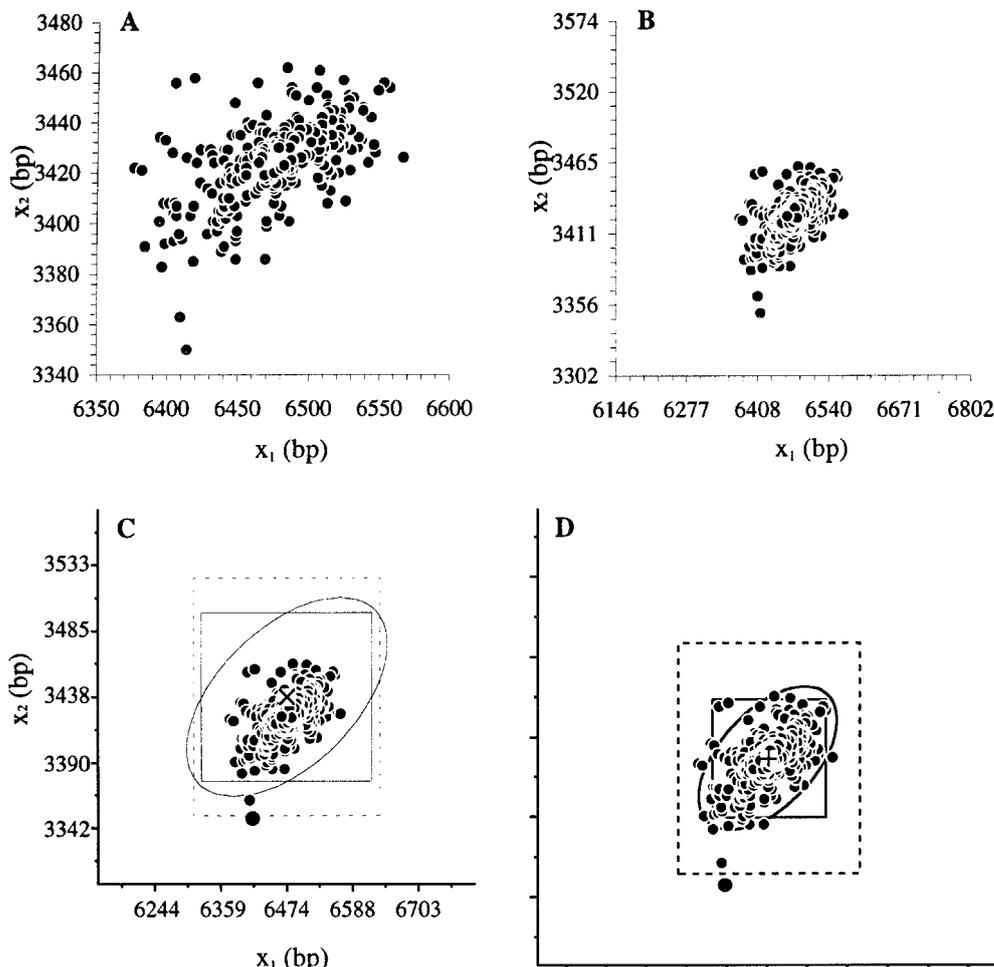


FIG. 2—Construction and graphical elements of the SLC, using the data displayed in Fig. 1. Figure 2A is the default scattergram; both axes have been standardized in Fig. 2B; the within-laboratory tolerance criteria are displayed in Fig. 2C with standard deviation-based axis tick-spacing; and the among-laboratory tolerance criteria are displayed in Fig. 2D with axis tick-labels removed.

While filling the graphical space, this default range-scale is entirely dependent upon the particular data displayed.

Figure 2B displays data with the x_1 and x_2 axes set to span a fixed number of expected among-laboratory standard deviations (S) about their expected sizes (\bar{X}_1 and \bar{X}_2). We use the NIST-certified values for K562 cell line control bands at locus D4S139 as the \bar{X}_1 and \bar{X}_2 for these data: 6474 bp and 3438, respectively (1). We expect the among-TWGDAM laboratories measurement uncertainty for any RFLP band of size \bar{X} to be:

$$S(\bar{X}) = 7.5 \left(1 + \frac{\bar{X}}{19,500} \right)^{7.1} \quad (1)$$

where \bar{X} ranges from about 1000 bp to 22,000 bp (4). The expected values of $S(\bar{X}_1)$ and $S(\bar{X}_2)$ for K562 D4S139 are 57 bp and 24 bp, respectively.

For normally distributed data, the number of S about the expected value required to display all valid data with a two-tailed probability of $(1 - \alpha/2)$, $Z_{1-\alpha/2}$, can be calculated from (or looked up in tables of) the standardized cumulative normal distribution. We chose $\alpha = 1.0 \times 10^{-6}$ so that all data having a one-in-a-million possibility of being within measurement uncertainty of the \bar{X}_1 and \bar{X}_2 values are included in the scatterplot. The $Z_{1-\alpha/2}$ corresponding to this 0.999999 inclusion probability is approximately 5.74. Thus, the beginning and endpoints of the x_1 and x_2 axes are:

$$\begin{aligned} \bar{X}_1 \pm 5.74 \times S(\bar{X}_1) \\ \bar{X}_2 \pm 5.74 \times S(\bar{X}_2) \end{aligned} \quad (2)$$

For K562 D4S139, the x_1 endpoints are 6144 bp and 6804 bp, and the x_2 endpoints are 3302 bp and 3574 bp.

Figure 2C displays the data with standardized axes and two among-laboratory tolerance criteria. The tick-marks of the x_1 and x_2 axes are in S units about the \bar{X}_1 and \bar{X}_2 values. The values of \bar{X}_1 and \bar{X}_2 are provided by the central tick-label along both axes. The difference between neighboring tick-labels along each axis is twice the S for the given band. The symbol “x” marks the plot’s center, $\{\bar{X}_1, \bar{X}_2\}$.

The light dotted-line box represents a $\pm 2.5\%$ univariate match window applied independently to each band, enclosing all $\{x_1, x_2\}$ data that are simultaneously within the intervals:

$$\begin{aligned} 0.975 \times \bar{X}_1 \leq x_1 \leq 1.025 \times \bar{X}_1 \\ 0.975 \times \bar{X}_2 \leq x_2 \leq 1.025 \times \bar{X}_2 \end{aligned} \quad (3)$$

For K562 D4S139, the $x_1 \pm 2.5\%$ window extends from 6312 bp to 6636 bp and the $x_2 \pm 2.5\%$ window extends from 3352 bp to 3524 bp. The choice of $\pm 2.5\%$ of expected band size as match window reflects common forensic practice (12). Since the expected among-laboratory measurement S is not simply proportional to band size, there is no simple probabilistic interpretation of any such percent-of-size criterion.

In contrast, the light solid-line box represents a tolerance bound designed to enclose 99% of all data pairs that are valid members of their separate univariate normal distributions:

$$\begin{aligned} \bar{X}_1 - [Z_{0.99} \times S(\bar{X}_1)] \leq x_1 \leq \bar{X}_1 + [Z_{0.99} \times S(\bar{X}_1)] \\ \bar{X}_2 - [Z_{0.99} \times S(\bar{X}_2)] \leq x_2 \leq \bar{X}_2 + [Z_{0.99} \times S(\bar{X}_2)] \end{aligned} \quad (4)$$

where (two-tailed) $Z_{0.99}$ is again calculated from (or looked up in tables of) the standardized cumulative normal distribution. The value for $Z_{0.99}$ is about 2.58; for K562 D4S139, the x_1 99% tolerance interval extends from 6326 bp to 6622 bp and the x_2 99% tolerance interval extends from 3377 bp to 3499 bp.

The light solid-line ellipse represents a bivariate tolerance bound designed to enclose 99% of all data pairs that are valid members of their joint bivariate normal distribution. For each $\{x_1, x_2\}$, the standardized bivariate distance, K_{bi} , from $\{\bar{X}_1, \bar{X}_2\}$ is:

$$K_{bi} = \frac{\left(\frac{x_1 - \bar{X}_1}{S(\bar{X}_1)} \right)^2 + \left(\frac{x_2 - \bar{X}_2}{S(\bar{X}_2)} \right)^2 - 2 \times R(\bar{X}_1, \bar{X}_2) \times \left(\frac{x_1 - \bar{X}_1}{S(\bar{X}_1)} \right) \left(\frac{x_2 - \bar{X}_2}{S(\bar{X}_2)} \right)}{(1 - R(\bar{X}_1, \bar{X}_2)^2)} \quad (5)$$

where $R(\bar{X}_1, \bar{X}_2)$ is the expected correlation between measurements. Since x_1 and x_2 are replicate size estimates for the same two DNA fragments, any observed non-zero correlation between x_1 and x_2 cannot indicate covariation in fragment size but rather covariation in the measurement of the size of fragments located in the same lane of a given gel. While there are several aspects of the analytical system that influence this measurement covariation, the distance between the two bands in the agarose gel is the dominant factor. Our current estimate of the expected within-laboratory measurement correlation is (6):

$$R(\bar{X}_1, \bar{X}_2) = 0.72 - \left[0.65 \times \log_{10} \left(\frac{\bar{X}_1}{\bar{X}_2} \right) \right] \quad (6)$$

For K562 D4S139 data, the value of $R(\bar{X}_1, \bar{X}_2)$ is 0.53.

Given data pairs from a bivariate normal distribution having the expected values for location, dispersion, and correlation, 99% of the data pairs will have:

$$K_{bi} \leq K_{bi_{99}} \quad (7)$$

where the critical constant $K_{bi_{99}}$ is calculated as (or looked up in tables of) the inverse one-tailed probability of the χ^2 distribution with two degrees of freedom (13). The value for $K_{bi_{99}}$ is approximately 9.21.

The 99% bivariate tolerance ellipse is defined by solving Eq 5 for a set of $\{x_1, x_2\}$ values such that $K_{bi} = K_{bi_{99}}$:

$$\begin{aligned} x_2 &= \bar{X}_2 + [S(\bar{X}_2) \times \omega] \\ x_1 &= \bar{X}_1 \\ &+ \left[S(\bar{X}_1) \times \left(\omega \times R(\bar{X}_1, \bar{X}_2) \pm \sqrt{(1 - R(\bar{X}_1, \bar{X}_2)^2)(K_{bi_{99}} - \omega^2)} \right) \right] \end{aligned} \quad (8)$$

where ω is a construction variable of domain $-\sqrt{K_{bi_{99}}} \leq \omega \leq +\sqrt{K_{bi_{99}}}$.

Figure 2D displays the data and two within-laboratory tolerance criteria, using the same standardized axes but with the axis and tick labels suppressed. The symbol “+” marks the data’s center, $\{\bar{X}_1, \bar{X}_2\}$. For these data, the value of \bar{X}_1 is 6477 bp and that of \bar{X}_2 is 3425 bp.

The dark dotted-line box represents at $\pm 2.0\%$ univariate match window applied independently to each band, enclosing all $\{x_1, x_2\}$ data that are simultaneously within the intervals:

$$\begin{aligned} 0.98 \times \bar{X}_1 \leq x_1 \leq 1.02 \times \bar{X}_1 \\ 0.98 \times \bar{X}_2 \leq x_2 \leq 1.02 \times \bar{X}_2 \end{aligned} \quad (9)$$

For these data, the x_1 interval extends from 6347 bp to 6607 bp and the x_2 interval extends from 3357 bp to 3494 bp. The choice of $\pm 2.0\%$ of expected band size as a within-laboratory match window is in accord with recent experience-based recommendations (12). Again, this 2% match window has no simple probabilistic interpretation.

The dark solid-line box represents a tolerance bound designed to enclose 95% of all $\{x_1, x_2\}$ that are valid members of their separate univariate normal distributions with a confidence of 95%:

$$\begin{aligned} \bar{X}_1 - [K_{uni95,95}^n \times S_1] &\leq x_1 \leq \bar{X}_1 + [K_{uni95,95}^n \times S_1] \\ \bar{X}_2 - [K_{uni95,95}^n \times S_2] &\leq x_2 \leq \bar{X}_2 + [K_{uni95,95}^n \times S_2] \end{aligned} \quad (10)$$

where S_1 and S_2 are the observed S for the respective bands and $K_{uni95,95}^n$ is the two-tailed univariate critical tolerance factor for 95% coverage of n data pairs with 95% confidence. For these data, S_1 and S_2 are 38 bp and 17 bp, respectively, and n is 264.

While the integral defining $K_{uni95,95}^n$ is of closed form and can be solved numerically, tables of tolerance factors are available and are much more tractable (14,15). The asymptotic limit of $K_{uni95,95}^n$ at very large n is just (two-tailed) Z_{95} . Using this limit and tabulated values, we have established an empirical estimate of $K_{uni95,95}^n$ for $n \geq 5$:

$$K_{uni95,95}^n \cong 1.96 + \frac{2.77}{n^{0.534} - 1.47} \quad (11)$$

The characteristic expanded uncertainty (95% confidence) of this estimation is ± 0.01 . The value for $K_{uni95,95}^{264}$ is about 2.11.

The dark solid line ellipse is a bivariate tolerance bound designed to enclose 95% of all data pairs that are valid members of their joint bivariate normal distribution with a confidence level of 95%. This 95%/95% within-laboratory tolerance criterion is defined in the same manner as was the 99% among-laboratory criterion of Eq 5, using the parameters observed for the particular data rather than those expected for the given DNA fragments. The standardized distance from the center of a data-defined bivariate normal distribution is:

$$\hat{K}_{bi} = \frac{\left(\frac{x_1 - \bar{X}_1}{S_1}\right)^2 + \left(\frac{x_2 - \bar{X}_2}{S_2}\right)^2 - 2 \times R \times \left(\frac{x_1 - \bar{X}_1}{S_1}\right) \left(\frac{x_2 - \bar{X}_2}{S_2}\right)}{(1 - R^2)} \quad (12)$$

where S_1 and S_2 are the observed S for the respective data and R is the observed correlation between x_1 and x_2 over all measurement pairs. For these data, S_1 and S_2 are 38 bp and 17 bp, respectively, and R is about 0.54.

Given that the data pairs are from a single bivariate normal distribution, we can be 95% confident that 95% of the observed data pairs will have:

$$\hat{K}_{bi} \leq K_{bi95,95}^n \quad (13)$$

where the critical $K_{bi95,95}^n$ is a constant for a given number of observed data pairs, n . We know of no closed-form expression for $K_{bi95,95}^n$; however, values for $K_{bi95,95}^n$ are estimable using numerical simulations and are tabulated for n of 10 to 50 (16). The asymptotic limit of $K_{bi95,95}^n$ at very large n is the 95th percentile of a χ^2 distribution having 2 degrees of freedom. From this limiting value and the tabulated data, we have established an empirical estimate of $K_{bi95,95}^n$ for $n \geq 10$:

$$K_{bi95,95}^n \cong 5.99 + \frac{38.1}{n^{0.767} - 3.51} \quad (14)$$

The characteristic expanded uncertainty (95% confidence) of this estimation is ± 0.1 . The value of $K_{bi95,95}^{264}$ is about 6.54.

The 95%/95% bivariate tolerance ellipse is defined by solving Eq 12 for a set of $\{x_1, x_2\}$ values such that $\hat{K}_{bi} = K_{bi95,95}^n$:

$$\begin{aligned} x_2 &= \bar{X}_2 + [S_2 \times \omega] \\ x_1 &= \bar{X}_1 + \left[S_1 \times \left(\omega \times R \pm \sqrt{(1 - R^2)(K_{bi95,95}^n - \omega^2)} \right) \right] \end{aligned} \quad (15)$$

where ω has domain $-\sqrt{K_{bi95,95}^n} \leq \omega \leq +\sqrt{K_{bi95,95}^n}$.

Choice of Tolerance Criteria

The $\pm 2.5\%$ among-laboratory (Eq 3) and the $\pm 2.0\%$ within-laboratory (Eq 9) match windows are founded in forensic experience with bands mostly of size 1000 to about 8000 basepairs. For bands in this size range, the expected $S(\bar{X})$ is in fairly constant proportion to \bar{X} (4). Thus, the empirical and probabilistic tolerance criteria for the data of Figs. 1 and 2 are rather similar. As will be seen later, there is less similarity for very large (and very small) bands. In any case, display of some appropriate percent-of-band-size match window is valuable as it provides a familiar frame of reference for forensic analysis.

Positive correlation between x_1 and x_2 RFLP measurements is widely recognized, although with little agreement concerning quantitative properties (6,17,18). The bivariate probabilistic criteria (Eq 7 and 13) are thus more appropriate—if less familiar—than their univariate analogs (Eqs 4 and 10). As seen in Fig. 1, the univariate criteria are simultaneously too restrictive in the direction of the ellipse’s major axis (positive correlation, from lower left to upper right) and too accepting in the direction of the minor axis (negative correlation, upper left to lower right).

Examples

Within-Laboratory Tolerance Charts

While full-page SLC, such as Fig. 1, enable rapid identification of data-entry errors and other atypical data, the simultaneous display of data from different genetic loci helps to reveal system-wide phenomena. A simple way to do this is to plot several SLC on the same page. Each segment of such a “multiplot” is constructed in the manner described earlier. Since all axes for all loci are similarly standardized, the data occupy the same relative graphical space in each segment. The axis and tick-mark labels can be deleted for graphical clarity. We find that six loci can routinely be displayed on a single page.

Figure 3 presents laboratory A14’s K562 cell line control data for loci D1S7, D2S44, and D10S28 in addition to D4S139. The small size and relatively consistent shape of the 95%/95% ellipses attest to excellent within-laboratory measurement precision. The nearly constant difference between the observed data median and the expected (NIST-certified) location for each band strongly suggests that this laboratory’s measurements are systematically biased one $S(\bar{X})$ low relative to the majority of other forensic laboratories.

The small cluster of potential outlier data at the low end of several of the ellipses may be from the same gels; however, this cannot be confirmed as no gel-identification information was supplied with these data. Gel codes can easily be displayed for any data exterior to one or more of the tolerance criteria. Examination of the gels from which these relatively low-size values were obtained could provide insight into the more subtle systematic bias.

Figure 4 is a modified multiplot, contrasting data from laboratories A14 and A15. While both laboratories provided data for four loci, only the three loci D1S7, D2S44, and D4S139 are in common. Only data exterior to the 95%/95% tolerance ellipse of each labo-

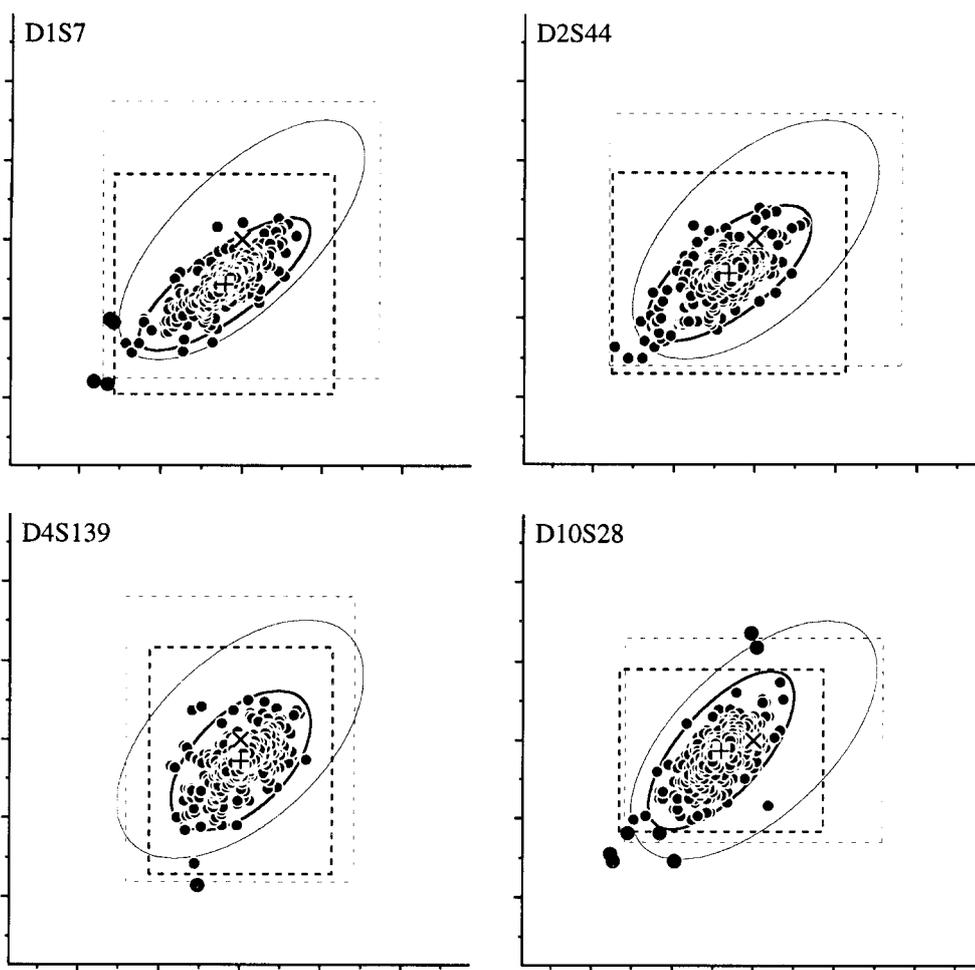


FIG. 3—Multiple SLCs for single-laboratory data. The four panels present D1S7, D2S44, D4S139, and D17S79 K562 cell line control data collected by laboratory A14 from 1990 to 1993.

ratory are individually plotted. This facilitates recognition of the 95%-of-the-time performance of the two laboratories without obscuring the “least usual” 5% of the data. These exterior data were identified by sorting in order of decreasing \hat{K} and plotting only those data with \hat{K} larger than $K_{bi,0.95}^n$.

Laboratories A14 and A15 have the “extreme” low- and high-size systematic RFLP sizing biases present in the data published with Ref 2. The difference between laboratory medians for D2S44 and D4S139 bands is about two $S(\bar{X})$. The smaller difference for the D1S7 bands (4571 bp and 4231 bp) is thought to be attributable to both laboratories’ use of a sizing standard with an electrophoretically anomalous 4300 bp component. Largely as a result of the systematic biases, a few data from both A14 and A15 do not pass any of the among-laboratory tolerance criteria. Some of these data are separated by more than 5% of their average size; however, all $\pm 2.5\%$ “false exclusions” present in the 15,000 data published with Ref 2 are displayed in the locus D2S44 and D4S139 components of Fig. 4.

Among-Laboratory Tolerance Charts

Figure 5 displays locus D1S7 SLC for the five samples used in the Postgrad study. The 2.0% within-laboratory match window (Eq 10) is not relevant to these displays and has been suppressed;

however, the 95%/95% tolerance ellipse (Eq 12) provides an interesting summary of average performance and has been retained. Since no “true” band size values have been established for the four non-K562 samples, the data median is used as each plot’s center and axis tic labels have been retained. The data exterior to the 99% tolerance ellipse have been labeled with a participant code.

There are at least two potential D1S7 outlier data for Postgrad participants P10 and P14. To investigate whether this represents part of the analyst’s learning curve or a consistent bias, all replicate data for these two participants have been connected with dark and light lines, respectively. Neither participant appears to be consistently biased relative to the bulk of the data.

Treatment of Atypical Data

While the procedures described above suffice for most data, there are atypical situations that require special treatment: one-banded patterns, very large data sets, and data-entry blunders.

Completely or Partially Missing Band

Figure 6A displays multi-participant Postgrad data for sample SC, which is consistently reported as homozygous at locus

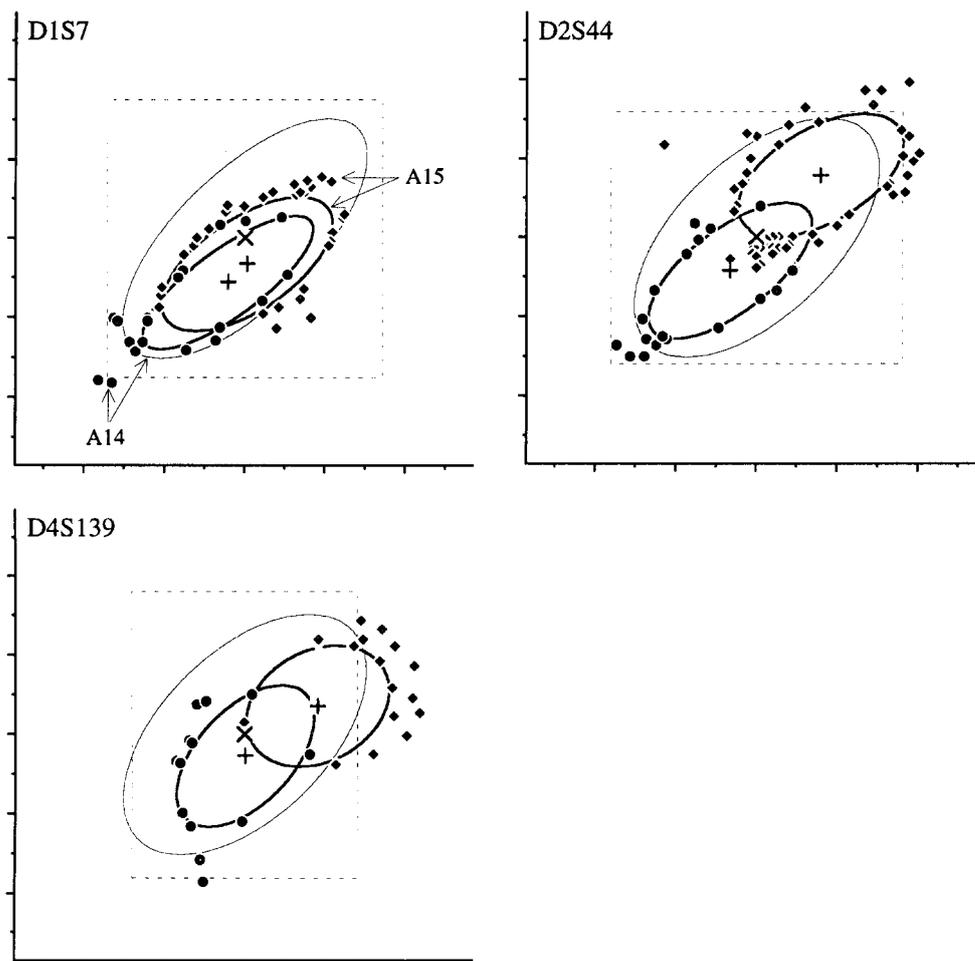


FIG. 4—Multiple SLCs contrasting data from two laboratories. The three panels present D1S7, D2S44, and D4S139 K562 cell line control data collected by laboratories A14 and A15 from 1990 to 1993.

D17S79. The single reported band size, x , is used for both x_1 and x_2 values. Although a univariate line plot or frequency histogram could portray the same information, maintaining the same graphical format for all loci facilitates data interpretation. The two-axis visualization of the same univariate among-laboratory match window criterion, while redundant, remains appropriate. Since the data are not bivariate, no tolerance ellipse is appropriate; however, the usually suppressed univariate tolerance boxes are appropriate and are now displayed.

Figure 6B displays multi-participant Postgrad data for sample JA, which is known to be heterozygous at locus D4S139 but gives bands of very similar size. When only a single value is reported for such a sample, we again use the value for both x_1 and x_2 . The SLC quite clearly presents the two data populations, apparent homozygote and closely spaced heterozygote. While the two populations could be analyzed separately, among-laboratory comparisons using a single composite model are less complicated than comparisons allowing two separate (measurement artifact) data populations.

Figure 6C presents an alternative treatment for partially missing data, redisplaying the D4S139 sample JA data but assigning a constant value, \bar{X}_2 , to all the missing x_2 . This approach is most appropriate when the x_2 band is small enough (for example, less than 900 bp) that it is sometimes completely eluted from the gel and when

some “experimental glitch” prevents measurement of one of the bands. A modification of this treatment is appropriate for bands that are too large or too small to be quantitatively measured, where the value of the upper or lower bound is the assigned constant.

Very Large Data Sets

Figure 6D displays an approach to handling very large amounts of data, such as provided here by laboratory A15 for locus D1S7, where most of the 607 $\{x_1, x_2\}$ values are “uninterestingly close” (from a data analyst’s viewpoint) to the data medians $\{\bar{X}_1, \bar{X}_2\}$. Printing an SLC that displays many hundreds to thousands of data pairs consumes considerable resources. As with the two-laboratory comparisons of Fig. 4, sorting the observed data in order of decreasing \hat{K} (Eq 10) and plotting just the “least usual” subset of the data retains most-to-all of the interesting information. Here, display of the first 150 data pairs is sufficient to unambiguously define the location of the remaining values.

Off-Scale Values

Figure 6E depicts our treatment for data that is beyond the one-in-a-million axis limits, using D1S7 data intentionally contaminated at one or both bands with values appropriate to other loci. This is one of several data-entry or data-retrieval blunders we have

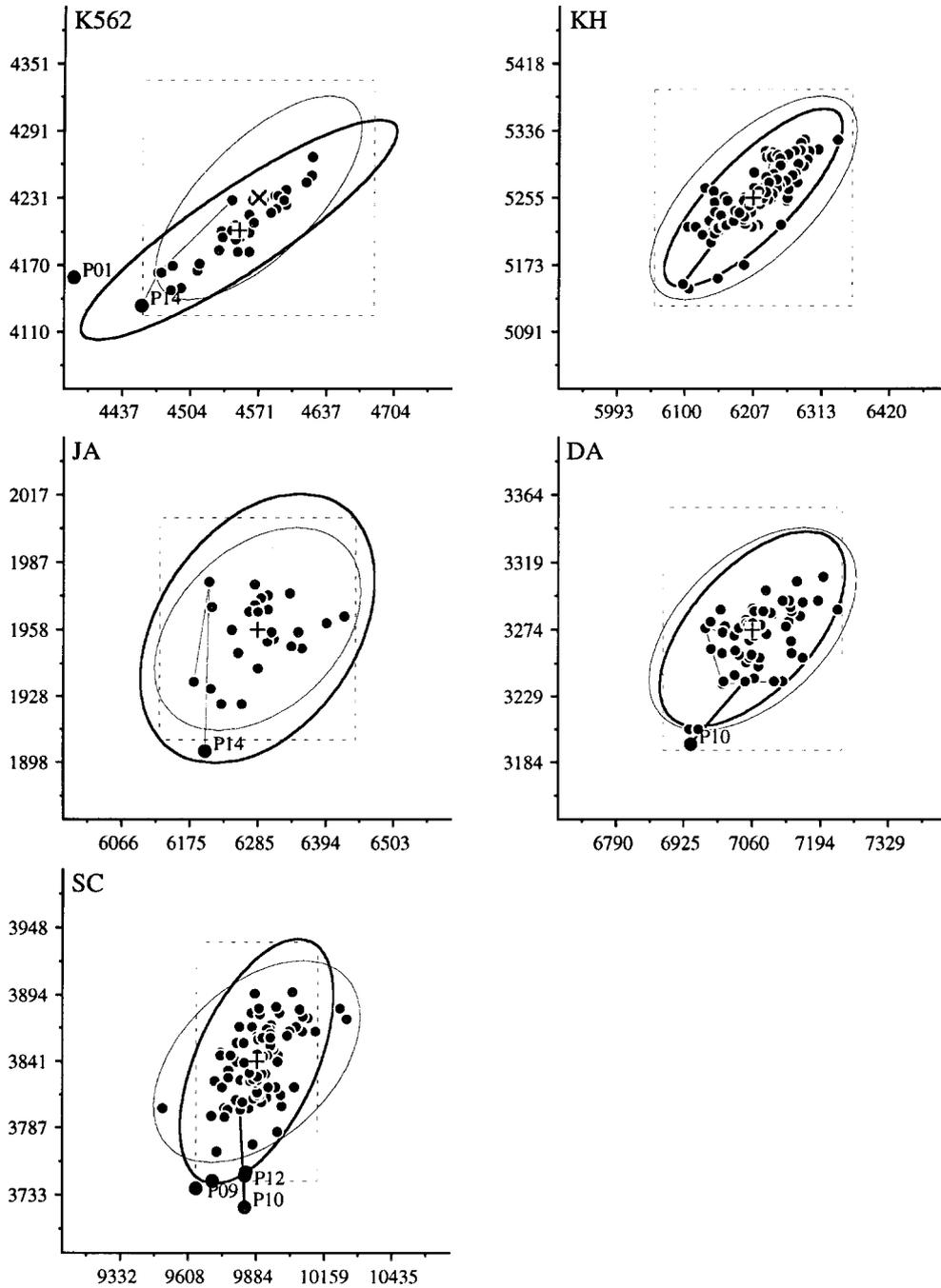


FIG. 5—Multiple SLCs for analysis of multi-laboratory data. Each panel presents locus DIS7 data for one of five different FBI postgraduate training samples, provided by 18 different analysts, collected from 1990 to 1992. The light lines connect replicate data submitted by analyst P14; the dark lines connect replicate data from analyst P10.

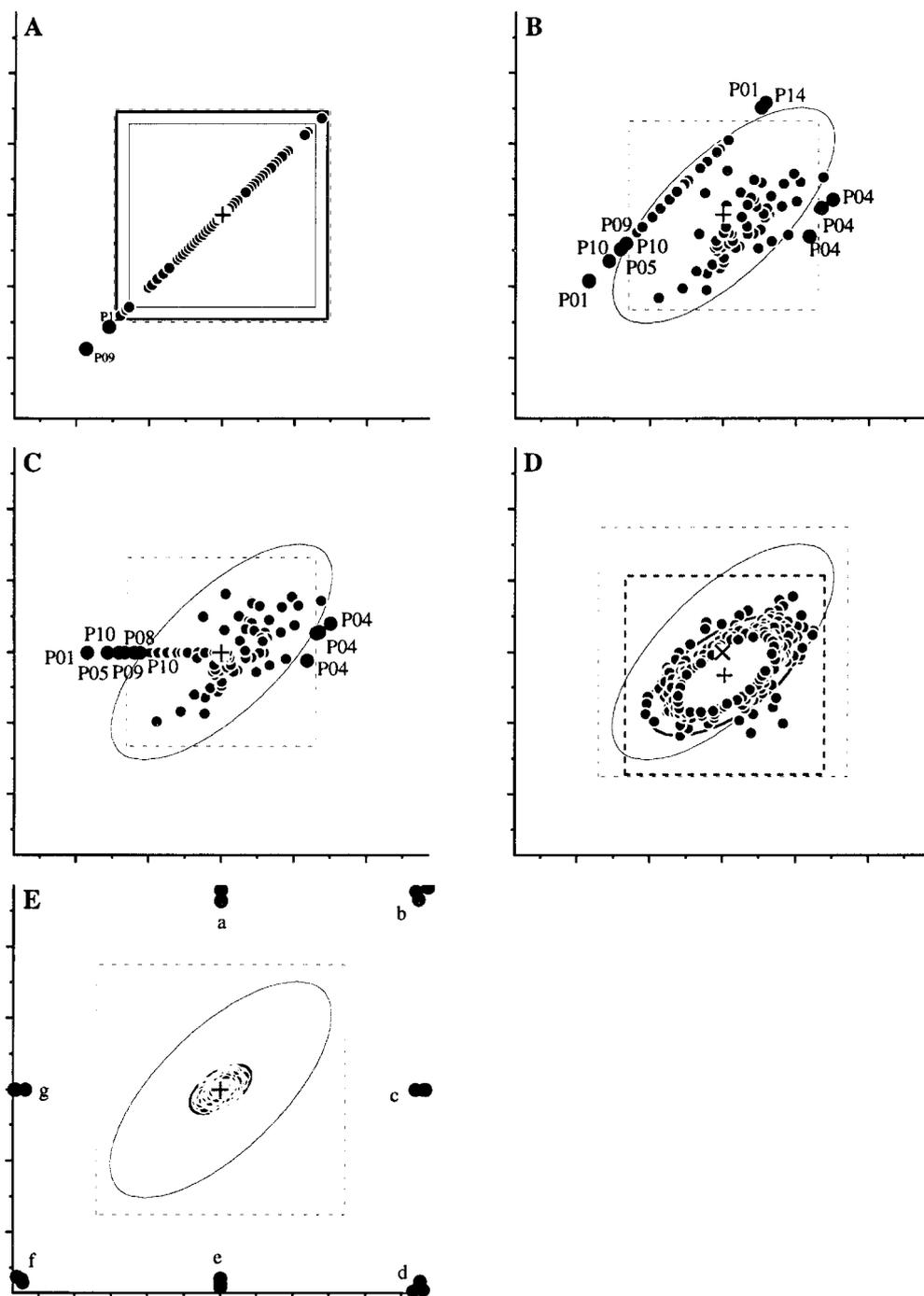


FIG. 6—Atypical data. Figure 6A is a pseudo-bivariate display for homozygous data, displaying FBI postgraduate training data for sample SC at locus D17S79. Figures 6B and 6C display two treatments for mixed-population resolved heterozygous and apparent-homozygous data for the postgraduate training sample JA at locus D4S139. Figure 6D describes the 607 K562 at locus D1S7 data pairs provided by laboratory A15 while actually plotting just the “least usual” 150. Figure 6E displays treatment of extreme outlier data, characteristic of some data-entry or data-retrieval errors.

encountered that give rise to such extreme outlier values. These “off the chart” outliers are assigned a graphical location along the edge of the chart that is closest to the true location of the data pair. Whenever the calculated location of an outlier is the same as that for another outlier, the location is slightly adjusted (dithered) so that the number of outlier data can be easily determined. Since there is extremely low probability that such outlier values are in fact members of the primary bivariate normal distribution, all ex-

treme outliers are excluded from the data-based location, dispersion, and correlation estimates.

Acknowledgments

We thank Stephen Niezgoda for his encouragement and suggestions, and for enabling CODIS system users to have ready access to their K562 quality assurance data. We thank Kenneth Gary for his

diligent validation and gentle correction of our arithmetic and algebraic notation. We again thank all the individuals and institutions who have made their data available for our studies. We are particularly grateful to those who have struggled with our software; many excellent suggestions have been made, intermixed with screams of agony.

References

1. National Institute of Standards and Technology. Certificate of Analysis, Standard Reference Material 2390, DNA Profiling Standard. Gaithersburg, MD: NIST Standard Reference Materials Program 1992.
2. Mudd JL, Baechtel FS, Duewer DL, Currie LA, Reeder DJ, Leigh SD, et al. Interlaboratory comparison of autoradiographic DNA profiling measurements. 1. Data and summary statistics. *Anal Chem* 1994;66:3303–17.
3. Duewer DL, Currie LA, Reeder DJ, Leigh SD, Liu H-K, Mudd JL. Interlaboratory comparison of autoradiographic DNA profiling measurements. 2. Measurement uncertainty and its propagation. *Anal Chem* 1995;67:1220–31.
4. Stolorow AM, Duewer DL, Reeder DJ, Buel E, Herrin G. Interlaboratory comparison of autoradiographic DNA profiling measurements. Part 3. Repeatability and reproducibility of RFLP band sizing, particularly bands of molecular size >10k base pairs. *Anal Chem* 1996;68:1941–7.
5. Duewer DL, Currie LA, Reeder DJ, Leigh SD, Liu H-K, Mudd JL. Interlaboratory comparison of autoradiographic DNA profiling measurements. 4. Protocol effects. *Anal Chem* 1997;69:1882–92.
6. Duewer DL, Benzinger EA. Products of partial digestion with *Hae* III. Part 2. Quantification. *J Forensic Sci* 1997;42(5):864–72.
7. Duewer DL, Lalonde SA, Aubin R, Fournery R, Reeder DJ. Interlaboratory comparison of autoradiographic DNA profiling measurements: precision and concordance. *J Forensic Sci* 1998;43(3):465–71.
8. Federal Bureau of Investigation. Draft, National DNA Index System (NDIS) Standards for CODIS Tolerance of DNA RFLP Data at NDIS. Washington: FBI Laboratory Room 3658, 20-May-1996.
9. College of American Pathologists. Anatomic Pathology Excellence Programs: Forensic Identity Program 1995 Set FID-A Final Critique. Chicago: CAP 1996.
10. Collaborative Testing Services, Inc. Paper & Paperboard Collaborative Reference Program Report No. 167S. Herndon, VA: CTS. 1997.
11. Leete CG. Personal communication 1998.
12. Laber TL, Iverson JT, Liberty JA, Giese SA. The evaluation and implementation of match criteria for forensic analysis of DNA. *J Forensic Sci* 1995;40:1058–64.
13. Tracy ND, Young JC, Mason RL. A bivariate control chart for paired measurements. *J Quality Tech* 1995;27:370–6.
14. Beyer WH, Ed. *CRC Standard Probability and Statistics: Tables and Formulae*. Boca Raton, FL: CRC Press 1991.
15. Hahn GJ, Meeker WQ. *Statistical Intervals: a guide for practitioners*. New York, NY: Wiley 1991.
16. Hall IJ, Sheldon DD. Improved bivariate normal tolerance regions with some applications. *J Quality Tech* 1979;11:13–9.
17. Evett IW, Scrance JK, Pinchin R. An efficient procedure for interpreting DNA single locus profiling data in crime cases. *J Forensic Sci Soc* 1992;32:307–24.
18. Eriksen B, Bertelsen A, Svensmark O. Statistical analysis of the measurement errors in the determination of fragment length in DNA-RFLP analysis. *Forensic Sci Int* 1992;52:181–91.

Additional information and reprint requests:
 Dennis J. Reeder
 NIST
 100 Bureau Drive, STOP 8311
 Gaithersburg, MD 20899-8311
 email: dennis.reeder@NIST.gov