

# MULTIPLEX\_QA

## An Exploratory Quality Assessment Tool for STR Multiplex Assays

David L. Duewer and John M. Butler

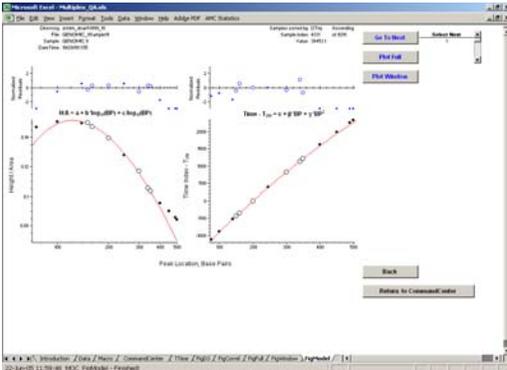
U.S. National Institute of Standards and Technology, Chemical Science and Technology Laboratory, Gaithersburg, MD 20899

The **Multiplex\_QA** system works in conjunction with the NIH-developed **BatchExtract** system to visualize short- and long-term changes in ABI 310 and 3100 electropherogram quality. **BatchExtract** converts ABI binary .fsa files into readily usable text data; **Multiplex\_QA** uses these files to estimate and display quality metrics that capture changes in electropherographic resolution and efficiency. These metrics are mostly based on the behavior of the internal sizing ladder included with every sample. Several different graphical displays enable identifying unusual events over time scales ranging from a single .fsa file to all .fsa files from a given instrument.

Multiplex\_QA, its User's Manual, and a demonstration dataset are available over the Internet through the NIST STRBase website: <http://www.cstl.nist.gov/biotech/strbase/software.htm>.

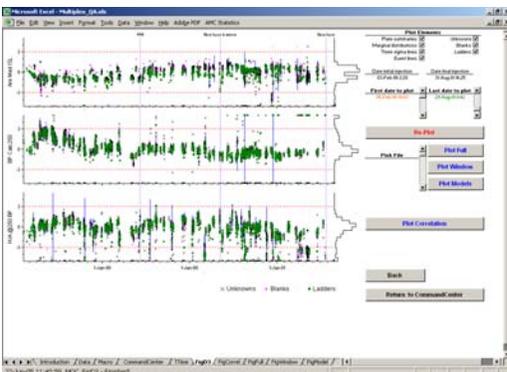
### Model Plots

Empirical regression models are used to characterize peak resolution and retention behavior. These models are defined for every sample using just the ABI-assigned values for height, area, bp size, and retention time of the internal sizing ladder peaks. All of these models can be visualized with **Model Plots**.



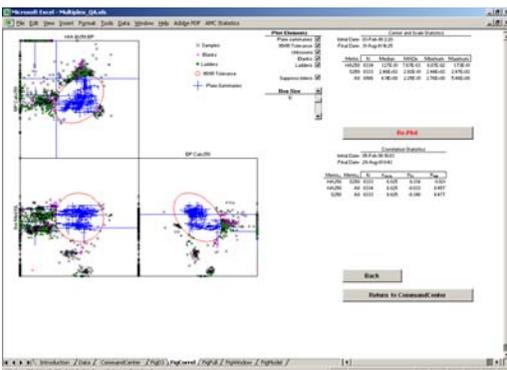
### D<sup>3</sup> Charts

The **Display, Document, and Discover** or "**D<sup>3</sup> Charts**" help you see how selected STR quality metrics change with time. In addition individual sample data, summary values for all samples amplified on the same plate can be displayed. Histograms and approximate 99% confidence intervals are options, along with markers for when particular events occurred. Up to five metrics can be stacked on the same Chart.



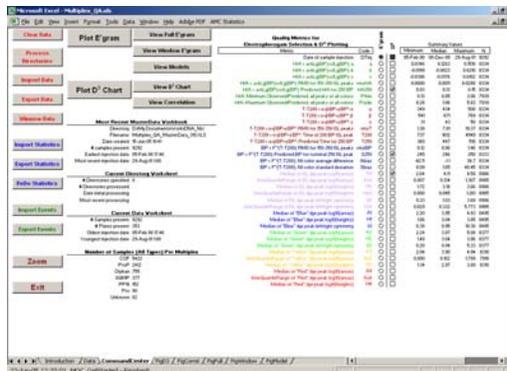
### Correlation Plots

**Correlation Plots** help you examine relations among the quality metric data displayed in the D<sup>3</sup> Charts. Summary values for all samples amplified on the same plate can again be displayed, along with tolerance ellipses that indicate the typical relationship among each pair of metrics.

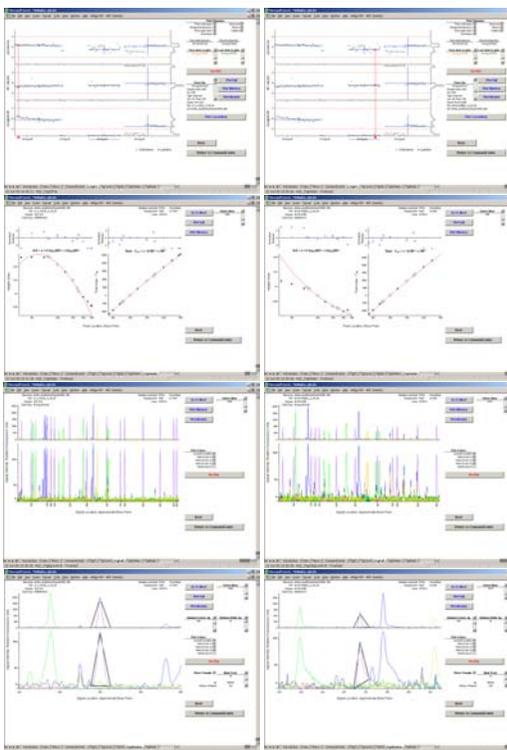


### The CommandCenter

**Multiplex\_QA** is a button-driven system intended to help you monitor different performance characteristics of multiplex STR analysis. The best way to determine if this system can help you is to **start clicking!**

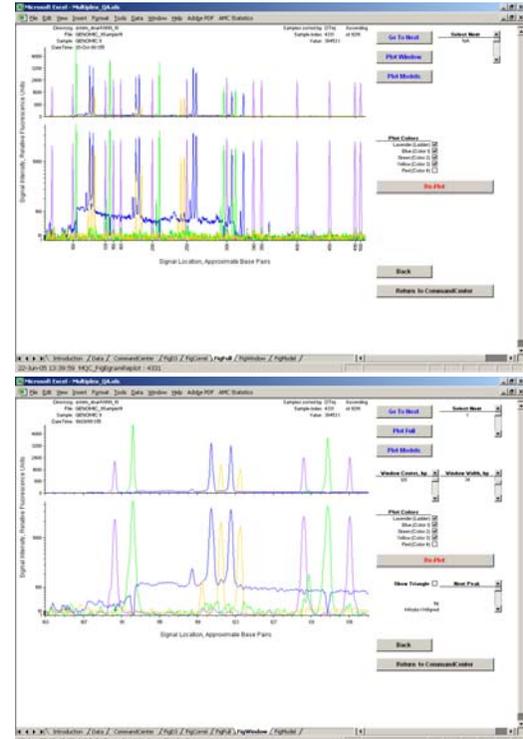


### Interconnecting Graphics



### Electropherograms

Full-scale and high-resolution electropherograms (**Full** and **Window E'grams**) help you see the quality of your analyses. In addition to the typical linear scale of fluorescence intensity, both E'grams also display the data along a modified logarithmic scale to reveal "baseline" structure. E'grams can be used to investigate samples with unusually thin or wide peaks (e.g., spikes or dye blobs), as well as to visualize potential performance issues identified using particular quality metrics.



### Multiplex\_QA is an exploratory analysis system.

It doesn't provide magic answers. Rather, it is intended to provide the tools you need to ask questions of your data:

- **when** did measurement performance change?
- **how much** did performance change?
- **which** samples are affected?, and
- **what** is going on in the electropherograms?

The **D<sup>3</sup> Chart** is the heart of the system, visualizing different quality metrics over time. Once you identify "something funny", you can then identify the individual samples, check the data used in the regression models with the **Model Plot**, view the complete electropherogram as a **Full E'gram**, and view selected regions of the electropherogram with **Window E'gram**. Hopefully, this information will help you understand **why** things changed and **whether** you need to take corrective action.

We do anticipate that particular patterns of changes can be associated with particular events. The sudden change in the height/area ratio (the metric "H/A@250 BP" at the bottom of the D<sup>3</sup> Charts) from a sudden onset of peak tailing. Unfortunately, the instrument used to take these data was used by many researchers and only an incomplete use log is available. We can not now connect all the evidence to a single common cause.

*We are therefore looking for a few good datasets!*

Do you have a few 1000 samples worth of ABI data from a single instrument, a pretty complete use log (when things happened, like capillary changes and maintenance), and some hours to spend helping answer a few of the **whys**?

### Acknowledgments and Disclaimer

These projects were funded by the National Institute of Justice through interagency agreement 2003-IJ-R-029 to the NIST Office of Law Enforcement Standards. Points of view are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice. Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.