applied biosystems

Addressing Current Data Interpretation Challenges with Applied Biosystems™ 3500 Data Collection Software v4.0 and GeneMapper™ *ID-X* Software v1.6

Pranapda Aumsuwan, Jaime Brachold, Xiaoyu Liu, Priti Sabadra. Thermo Fisher Scientific, 180 Oyster Point Blvd., South San Francisco, CA, USA, 94080

INTRODUCTION

3500 Data Collection Software v4.0 and GeneMapper ID-X v1.6 now offer an updated, integrated HID solution designed to significantly enhance productivity and confidence in the laboratory. The system addresses critical challenges faced in HID laboratories including data interpretation speed and reliability, user experience efficiency, and IT compliance.

Specifically addressing current data interpretation challenges and efficiency, this system offers the ability to reduce pull-up edits and capillary-to-capillary variation, improve first-pass success rate for database labs, and enable a more efficient data transfer for probabilistic genotyping (GMID-X export with stutter without the need to reanalyze).

The following summarizes highlights from the developmental validation performed at Thermo Fisher Scientific to demonstrate the functionality of pull-up reduction, signal optimization, off-scale data recovery, and streamlined data export.

MATERIALS AND METHODS

Instruments and computers

Three 3500xL and two 3500 instruments with data collection software v4.0 were used across the studies (not all instruments were used for each study; however, each instrument was used in at least, one study). Spatial based signal optimization was enabled on all instruments. GMID-X v1.6 was installed and tested on both Windows 7 and 10, 64-bit operating systems and used for data analysis.

Chemistry

All quantification was performed on a 7500 Real-Time PCR System for Human Identification using the Quantifiler™ Trio DNA Quantification kit. All STR amplifications were performed on GeneAmp™ PCR System 9700 thermal cyclers.

Representative STR kits were used in each study to demonstrate the ability of the instrument/software to achieve accurate results within the parameters of each test. The following kits were used for these studies: GlobalFiler™, GlobalFiler™ Express, NGM Detect™, Yfiler™ Plus, VeriFiler™ Express, Identifiler™ Plus, and Identifiler™ Direct (GF, GFE, NGMD, YFP, VFE, IDP, IDD).

All samples were amplified and run according to the standard protocols in the applicable User Guide, unless otherwise noted.

Summary of Studies

1. Pull-up reduction

Amplification was performed with 22 gDNA samples (1-2ng) with GF, NGMD, IDP, and YFP and 22 blood FTA card samples (1.2mm punch) with GFE and VFE. A 1:7 mixture of control DNA 007 and 9947A (1ng) was also amplified with GF. CE runs were performed using both the standard and off-scale recovery (OSR) dye sets. Samples were injected twice.

The same data set was processed with the pull-up reduction feature enabled and disabled in DC 4.0 using an internal Thermo Fisher Scientific software development tool. The total number of pull-up peaks were counted, the absolute percent value of each pull-up incidence was calculated, and results compared. The effect of pull-up reduction on the minor contributor peaks in the mixture data was also evaluated.

2. Signal optimization (SO)

Control 9947A DNA was run on three 3500xL instruments in 24 wells x 8 injections with spatial-dependent signal optimization enabled (SO factor calculated and applied) or disabled (no SO factor calculation).

Control 9947A DNA samples were run on three 3500xL instruments in 24 wells x 4 injections using the default injection position and an position higher in the well (z-offset).

The average peak height for each capillary in each run and %CV and the maximum-to-minimum (max:min) peak height ratio of the average peak height for all capillaries in each run was calculated.

3. Off-Scale Data Recovery (OSR)

Amplification was performed with kit positive controls (3-6ng) using GFE, VFE, and IDD. CE runs were performed using the OSR dye sets. Samples were injected three times with different injection voltages and times to generate peak heights (heterozygous alleles) of <32,000, 32,000-65,000, and >65,000 RFU.

Off-scale related flags and peak morphology was evaluated in GMID-X v1.5 and v1.6. Only GMID-X v1.6 can process the extended dynamic range; pull-up related to off-scale data should be minimized in both software versions.

4. GMID-X Data Export

With stutter filters enabled, export the genotypes from the plot view using the Export Table With Stutter option. Evaluate filtered stutter in the genotypes table and table export.

RESULTS

Pull-up Reduction

In the data collected with pull-up reduction enabled, the following results were observed:

- Fewer pull-up peaks above peak amplitude threshold were observed for all kits tested as shown in Figure 1.
- More samples with <1% were observed, and fewer samples with pull-up >5% were observed.
- The minor contributor genotype was not affected and the average peak height of the minor contributor was not significantly different as shown in Figure 2.

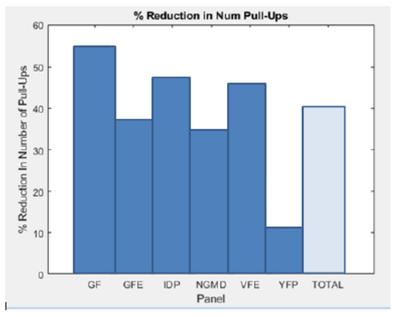

Figure 1: % reduction in pull-up with the pull-up reduction feature enabled in three dye sets (DS-33, DS-36, and DS-37). The range in reduction was ~10-55% depending on the sample set. The reduction of pull-up was greater in the samples that contained a higher incidence of pull-up data before the samples were reprocessed with the pull-up reduction feature enabled.

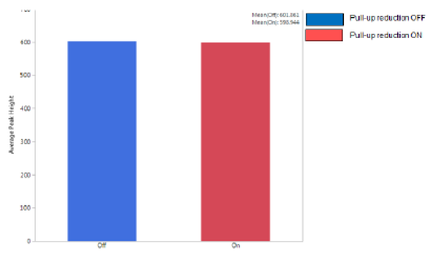

Figure 2: Average peak height of the unshared minor alleles with the pull-up reduction feature disabled (blue) is ~602 RFU and enabled (red) is ~599 RFU.

Signal Optimization (SO)

Using a spatial-dependent approach to signal optimization, an improvement in %CV and max:min ratio was seen across all 3500xL instruments. The average results across the instruments showed a delta %CV of 3.5% and delta max:min of 0.21. Results from a single instrument are shown in in Figure 3.

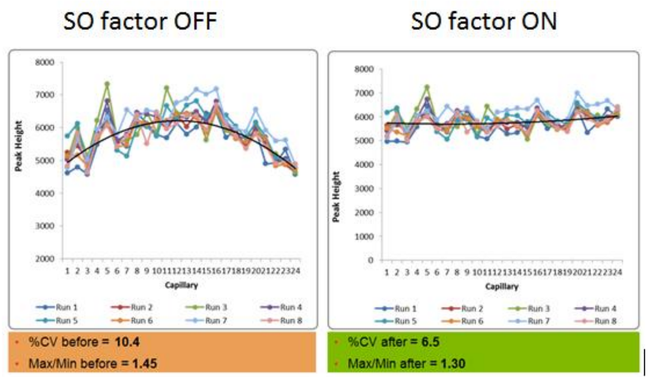

Figure 3: %CV and Max:Min ratios of the average sample peak heights across 24 capillaries on one instrument, from 8 injections of 9947A.

Using an injection position-dependent approach to signal optimization, an overall improvement in %CV and max:min ratio was seen across both 3500xL instruments. The level of improvement varied by instrument and was proportional to the initial variation observed; results are displayed in Figure 4.

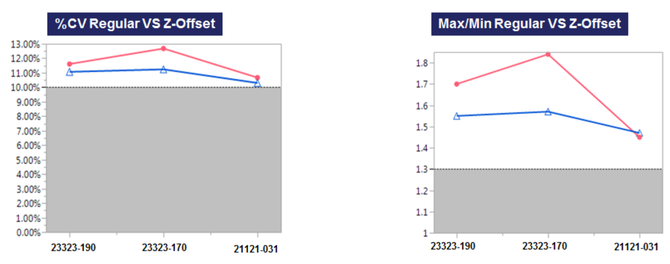

Figure 4: %CV and Max:Min ratios of the average peak heights, of the 400bp fragment of LIZ GS600, across 24 capillaries across three instruments, from 4 injections of 9947A.

Additional testing and results showed that on the 3500xL instrument, a greater reduction in variation was observed when using a combination of the spatial optimization and the z-offset. Further, results indicated there was not significant impact on the 3500 8-capillary instrument when spatial-optimization was applied.

Off-Scale Data Recovery (OSR)

Figure 5 outlines the results of peak morphology and off-scale related quality flags in GMID-X v1.5 and v1.6 for the three data sets. Additionally, when comparing the same sample from the 32,000-65,000 RFU data set and <32,000 RFU data set, there was 100% genotyping concordance and intra-locus peak balance within 10%. Although the 32,000-65,000 RFU data set was saturated in GMID-X v1.5, with OSR enabled data quality improved and fewer pull-up peaks under the off-scale peaks were observed; the same was true for off-scale related pull-up in GMID-X v1.6.

Data set/signal (RFU)	GeneMapper™ *ID-X* Software	
	v1.6	v1.5
≥65,000	Peaks are flat-topped and flagged as off-scale (OS).	• Peaks are off-scale. • Signal maximum is 32,000 RFU.
32,000–65,000	Peaks are on-scale with expected peak morphology.	
<32,000	Peaks are on-scale with expected peak morphology.	Peaks are on-scale with expected peak morphology.

Figure 5: Observations for the peak height of the heterozygous allele.

GMID-X Data Export

With stutter filtered, when the Genotypes Table from the Plot View is exported with stutter using the new Export Table With Stutter option the following is observed:

- Filtered stutter alleles are exported for Samples
- Exporting with stutter does not affect the peaks that are labeled in the plot or listed in the Genotypes table
- Stutter alleles are not exported for Allelic Ladder samples
- Only columns displayed in the Genotypes Table at the time of export are included in the export file
- PQVs in the exported file are based on the original analysis, no on all exported/unfiltered alleles
- Peaks with artifact or custom artifact labels are no exported
- Peaks with custom allele labels are export with the custom allele label name

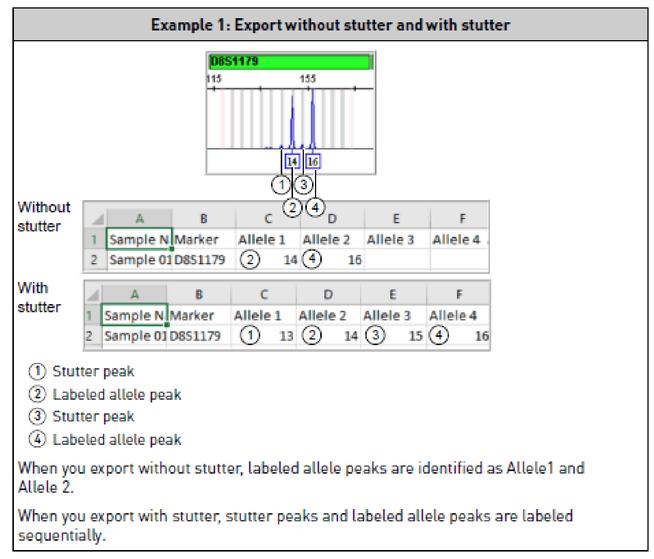

Figure 6: Example of export files without and with stutter.

CONCLUSIONS

The pull-up reduction and off-scale recovery features significantly reduced the number of pull-up peaks seen in the data and the number of samples/markers flagged as off-scale with using Applied Biosystems 5 or 6-dye chemistry. Both of these features should reduce the time spent on manual edits and/or review during data analysis. Peak height variation and signal uniformity across capillaries in an injection is improved by the signal optimization features. The ability to export with stutter, without the need to reanalyze, streamlines the process for downstream use of probabilistic genotyping software.

The updates included in 3500 Data Collection Software v4.0 and GeneMapper ID-X v1.6 address recent customer needs for improving data interpretation and efficiency. For the complete set of developmental validation studies and results refer to 355 Series Data Collection Software 4 for HID User Bulletin (Publication Number 100075298) and GeneMapper *ID-X* Software v1.6 User Bulletin (Publication Number 100073905).

ThermoFisher SCIENTIFIC