GeneMapper® ID-X Software v1.0: Expert System and next generation forensic data analysis software for casework samples

Lisa Calandro, M.P.H.*, Jaime Handelsman, Nicola Oldroyd, Thomas McElroy, Ravi Gupta
Disclosure

- This presentation references commercial products from
  - Applied Biosystems
Outline

• Introduction
• Casework sample analysis workflow
• Technical review features
• Quality control tools
• Mixture analysis
• Conclusions
Introduction

• GeneMapper® ID-X v1.0 software design goals
  • Conform to forensic casework and databasing analysis workflows
  • Provide expert system functionality for single source samples
  • Incorporate auditing and security features provides control of technical records
  • Introduce quality control functionality to aid in data interpretation
  • Allow documentation and tracking of edits for streamlined technical review
Casework Data Analysis Workflow

Analysis of Sample Files
- Analysis Requirements Check
- Analysis Summary

Primary Analyst Interpretation
- New PQVs
- Enhanced peak labeling
- Colored marker header bars
- Profile concordance tool

Second Reader Interpretation
- Auditing functionality
- Label Edit Viewer

Data Output
Software Operations and Workflow

Add Samples & Press Analyze

Analysis Requirement Check → Peak Detection & Sizing → Allelic Ladder Quality Check → Allele Calling → Quality Value Checks → Analysis Summary

View & Interpret Data
Analysis Requirements Check & Summary (Automatic)

Highlights samples with missing information before proceeding with analysis.

Indicates samples for which Analysis Requirements Not Met (ARNM)
Allelic Ladder Analysis and Quality Assessment (automatic)

- Allelic ladder samples are sized and genotyped prior to sample analysis
- Allelic ladder samples are assessed based on specific allelic ladder requirements
- Only those ladders meeting all requirements are used to generate bin offsets
- Users may override quality assessment as needed
- If at least one allelic ladder has met requirements, the software may automatically proceed with analysis

Eliminates the need to manually inspect allelic ladder samples
Allelic Ladder Analysis Summary

Can elect to continue analysis (saves time) or view low quality allelic ladder(s)
Allelic ladder not labeled due to peaks below threshold

Click on marker header bars

Quality Value Details
- GQ = 0.0
- One or more allelic ladder requirements have not been met. See below:
  - One or more peaks were not detected

Due to the above reason(s), this ladder has not been used to generate bin offsets.

To display reason for failure
Sample Level Quality Assessment

- Sample Offscale (SOS)
- Sizing Quality (SQ)
- Sample Spikes (SSPK)
- Potential Mixture (MIX)
- Outside Marker Range (OMR)
- Composite Genotype Quality (CGQ)
  - Factors in marker-specific GQ values
Analysis Summary Generation (automatic)

- Provides snapshot of allelic ladder, control and sample quality
- Interactive window allowing users to link to a desired group of samples
  - Samples that met all thresholds
  - Samples that did not meet one or more thresholds
### Sample Status Summary

<table>
<thead>
<tr>
<th>Sample Status</th>
<th>Total # of Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unanalyzed</td>
<td>0</td>
</tr>
<tr>
<td>Analyzed</td>
<td>13</td>
</tr>
<tr>
<td>Analysis Setting Changed</td>
<td>0</td>
</tr>
</tbody>
</table>

### Allelic Ladder Quality Summary*

*Eliminates need to manually review allelic ladder samples

### Control Quality Summary*

*Eliminates need to manually review passing controls

### Sample Quality Summary

<table>
<thead>
<tr>
<th></th>
<th>Total # of Samples</th>
<th>All thresholds met</th>
<th>One or more thresholds not met</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Click on link to view only those samples in a specific category
Review and Interpret Samples

• One of the most time-consuming steps in the casework data analysis workflow
• Manual review tools specifically designed to improve efficiency of manual review and interpretation of artifacts
Tools Designed to Minimize Manual Review Time

• Quality value flags
  • Sample level
  • Marker level

• Plot display options
  • Colored marker header bars correspond to Genotype Quality (GQ)
  • Quality Values Details window
  • Quality value trigger labels

• Labeling scheme
Evaluate Sample Level Quality Values

- **SSPK flag is yellow**: one or more spikes were detected
- **CGQ is yellow**: either the spike is within a marker size range and/or there is another anomaly in a marker size range
Display Sample Plot

Electropherogram

Genotypes Table

Quality Value Details
View Problematic Markers

Scan colored marker header bars to quickly find markers with yellow or red Genotype Qualities.
View Quality Value Details

- Peak height ratio of 19 and 25 allele peaks is 0.5747
  - The threshold is 0.7 or greater, thus PHR flag is fired
- Peak detected that met the rule for spike

<table>
<thead>
<tr>
<th>PQV</th>
<th>Size</th>
<th>Observed Value</th>
<th>Threshold</th>
<th>GQ Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHR</td>
<td>323.82, 347.37</td>
<td>0.5747</td>
<td>&lt;= 0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>SPK</td>
<td>349.54</td>
<td>0.2089</td>
<td>&lt;= 0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Confirm Artifacts using Raw Data

View peak in raw data with direct link from Sample Plot
Edit (if required)

No editing required for spike peak

- Automatic “artifact” labeling for spikes
- “Artifact” labels do not get entered in Genotypes Table
Manually Accept Marker Genotype

Override GQ

- GQ flag turns to green square
- All other PQV flags turn grey, but retain original shape

Provides evidence anomalies within the marker were visually inspected and accepted
Manually Accept the Sample Genotype

Override CGQ

CGQ flag becomes green hexagon

Dialog appears automatically when all GQs overridden
Peer/Technical Review

- Multiple ways to visualize edits made by primary analyst
- GQ and CGQ overrides are obvious to technical reviewer
  - Pre-defined table settings allow overrides to be selected
- Electronic records of manual edits easy to generate for a given project
- Audit history records when each edit was made and user name of analyst who made the edit
Gray marker header bars and PQVs indicate edits
Electronic Peer/Technical Review: Label Edit Viewer

- Lists all edits for samples displayed in the samples plot
- Click on a row in the table to highlight the corresponding peak
Select View Edited Samples table to display edits

- Checkmark in SE column indicates sample edit
- Sample-level PQVs grayed out
GQ Override

GQ Yellow due to MPH flag
PQVs grey out and GQ turns to green upon override
Sample Data Quality Control

- Ability to use unlimited custom positive controls
- Three types of profile comparisons performed:
  - Project sample to sample
  - Project sample to user imported lab reference profiles
  - Project sample to QC sample/custom control profiles
- Designed to:
  - Group concordant samples
  - Quickly identify potential contributors to samples
  - Allow automatic blind QC checks
Profile Comparison Tool: Sample Comparison

User defined percent match threshold

Matching samples

Percent Match

Matching alleles highlighted in blue
Contamination Check

Low level peaks observed in negative control sample
Comparison to Lab Reference Database

<table>
<thead>
<tr>
<th>Profile Comparison</th>
<th>Sample Concordance</th>
<th>Sample Comparison</th>
<th>Lab Reference Comparison</th>
<th>Control/QC Comparison</th>
</tr>
</thead>
</table>

Percent Match Threshold (Percent of reference profile alleles detected in the comparison profile)

80%  

Lab Reference Comparison

Project sample/source-group used as reference to calculate percent match

<table>
<thead>
<tr>
<th>Gene</th>
<th>ID</th>
<th>NegControl.fsa</th>
<th>NegControl (Reference)</th>
<th>% Match</th>
</tr>
</thead>
<tbody>
<tr>
<td>G11</td>
<td></td>
<td></td>
<td></td>
<td>100.0%</td>
</tr>
<tr>
<td>ANALYST A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6S1179</td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>D21S11</td>
<td></td>
<td></td>
<td></td>
<td>30, 33.2</td>
</tr>
<tr>
<td>D7S820</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>CSF1PO</td>
<td></td>
<td></td>
<td></td>
<td>9, 12</td>
</tr>
<tr>
<td>D8S1179</td>
<td></td>
<td></td>
<td></td>
<td>9, 9.3</td>
</tr>
<tr>
<td>THO1</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>D13S317</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>

100% of alleles detected in negative control sample match Analyst A
GeneMapper® ID-X Software v 1.1

Development of a Mixture Analysis Tool

- Integrated tool to assist forensic casework scientists with the analysis and interpretation of mixed DNA samples
- Analysed and edited data moves smoothly and seamlessly from the GeneMapper® ID-X project into the mixture analysis tool
- Performs three major functions
Functions Performed in the Mixture Analysis Module

- Categorizes all samples within a project as Single Source, minimum 2 contributors or 3 or more contributors
- Extraction of individual contributors from 2 person mixtures
  - Separates the alleles in a mixed sample into two distinct contributors
  - Allows access to all possible genotype combinations but narrows the selections automatically into a most likely set for inclusion and subsequent analysis
  - Manages instances where one contributor to a mixture is known
- Evaluating multiple statistical calculations based on possible contributors
  - Provides different options for the different sample mixture categories
  - Allows import of User defined population databases
Summary

- GeneMapper® ID-X Software streamlines the data analysis workflow utilized by the casework laboratory
- The software follows a logical and efficient workflow that may eliminate the need for several manual steps
- Manual review tools have been designed to decrease the amount of time spent interpreting anomalies
- Quality control functionality is incorporated to easily assess controls and identify potential contributors in unknown samples
- The validated software offers security, auditing and multi-user database functionality that may be customized to meet the needs of each individual laboratory
- A tool is in development to assist the casework analyst with the interpretation of mixtures
Acknowledgements

- **Applied Markets Software Group**: Chris Vinson, Bobby Soni, Yelena Ximin, Manjula Aliminati, Gloria Lam, Ravi Gupta, Sharada Vijaychander, Xi Hau, Carol Kosman, Bruce deSimas
- **Applied Markets Product Assurance Group**: April Baugher, Liwei Qi, Eric Tam
- **Applied Markets Product Management/Marketing**: Thomas McElroy, Lisa Schade, Lisa Calandro, Nicola Oldroyd, Jonathan Tabak, Peter Christey
- **Applied Markets Field Support Group**: Melissa Kotkin, Michelle Shepherd, Yogesh Prasad, Andrew Masel
Legal Statement

For Research, Forensic or Paternity Use Only. Not for use in diagnostic procedures.

Notice to Purchaser:

Purchase of this software product alone does not imply any license under any process, instrument or other apparatus, system, composition, reagent or kit rights under patent claims owned or otherwise controlled by Applera Corporation, either expressly, or by estoppel.

GeneMapper ID-X Software version 1.0 and GeneMapper ID Software version 3.1 and version 3.2 have undergone a verification process defined by Applied Biosystems. However, human identification laboratories analyzing forensic, paternity, databasing, and single-source samples that choose to use GeneMapper ID-X Software for data analysis should perform their own appropriate validation studies.

Applera, Applied Biosystems, AB (Design), AmpFLSTR, COfiler, GeneMapper, Identifiler, Profiler Plus, SGM Plus and Yfiler are registered trademarks and MiniFiler and SEfiler are trademarks of Applera Corporation or its subsidiaries in the US and/or certain other countries.

© 2008 Applied Biosystems. All rights reserved.
Thank You!

Questions?
Lisa.M.Calandro@appliedbiosystems.com