NIST Experience with FSS-i\(^3\) v4.1.3 Software Upgrade

Becky Hill
Amy E. Decker, Margaret C. Kline and John M. Butler
National Institute of Standards and Technology

17th International Symposium on Human Identification
October 11, 2006

Cost to the End User

- **Software**
  - Receive the v4.1.3 upgrade software
  - Single copy, single computer: $20,000

- **Maintenance agreement**
  - $4,000 per year (20\% of total software cost per year, max $15,000)
  - Software upgrades and patches are included

- **Training**
  - $2,000 if at Promega (plus your travel expenses)
  - $12,000 for up to 5 people if performed in your lab

- Requires GeneMapper ID or GeneScan/Genotyper software to already be in place in your lab

**Minimum starting cost of $26,000**

Disclaimers

**Funding:** Interagency Agreement 2003-IJ-R-029 between the National Institute of Justice and 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.

**Packaging for FSS i-Cubed Software**

**NIST Experience with Software Purchase**

- Attempted to purchase directly from FSS
  - No quote provided by Chris Macguire despite multiple attempts and email agreements to do so (Dec 2004, Jan, Feb, May, June 2005)

- Quote for software from Promega on Oct 18, 2005
  - Told that we had to purchase $4,000 maintenance agreement along with at least $2,000 training (plus travel expense to Madison, WI)
  - NIST contract officer signed off Dec 19, 2005

- Promega installed v4.0.1 software January 3, 2006

- Becky Hill went to Madison, WI Jan 9-13, 2006 for first training class held at Promega

- Received v4.1.2d August 3, 2006, installed by Bob McLaren

- Received and installed v4.1.3 patch September 9, 2006

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Overview of Software Components

**i-STReoss** quickly and accurately calls your allele types and objectively assesses the quality of your data. This allows analysts to reduce the time spent manually reviewing data and focus on “problem” samples.

**i-STReam** module evaluates two-person DNA mixtures and produces a best-fit major profile. This aids the reporting analyst in mixture deconvolution and unbiased interpretation.

**i-integrity** checks for potential sample-to-sample contamination within a batch by comparing all alleles called in a sample to the alleles in every other sample in the batch.

http://www.promega.com/applications/profile.asp?appname=Genetic+Identity&sku=FSSI3&spl=off

Introduction to FSS-i³ Software

**FSS-i³ Flow Chart**

1. Obtain data from 3100/3130xl
2. Process data with GeneMapper ID
3. Export data as a text file
4. Designate allele calls using i-STReoss
5. FSS-i³ checks ladders first
6. Input data into FSS-i³
7. Review flagged samples and edit as necessary
8. Generate output files of all data

**GeneMapper ID Raw Data Sorted**

Change GeneMapper ID Bin Settings

Min/Max Sizes were adjusted to be the same for all dye channels, and marker bin settings were lined up so there were no gaps within a dye channel.

These bins were developed for FSS-i³ by Promega

Text File Saved from GeneMapper ID “Raw Data”

Microsoft Excel used to examine data
FSS-i3 Input Template

Controls what and how information will be brought into the FSS-i3 program.

Import data as text file into FSS-i3

PowerPlex16 D3S1358 ladder

At least one ladder must pass for rules to be applied and alleles designated for each sample.

FSS-i3 Rule Sets and Thresholds

Each multiplex kit can have multiple rulesets.

Images from FSS-i3 i-STRess

Samples are displayed as "spikograms" rather than peaks using the peak size and height information gathered from GeneMapper ID or Genotyper "raw data".

Editing samples

Change of color.
FSS-i³ Output Template

Controls what and how information will be exported from the FSS-i³ program.

Output Files Created by FSS-i³ i-STRess

C:\Program Files\FSSi3 v412d\i-STRess\Files\PP16 Combo2 Output Files

Audit Files Created by FSS-i³ i-STRess

C:\Program Files\FSSi3 v412d\i-STRess\Files\PP16 Combo2 Audit Files

New Features of v4.1.3

FSS-i³ Upgrade

Creating new multiplexes with Multiplex Manager

Ladder editor for new multiplexes (Y-filer)

Y-filer loci were added manually.

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Becky Hill – NIST Experience with FSS-i3 v.4.1.3
17th Int Sym Hum Ident –Expert System breakout session

October 11, 2006

Y-filer DYS456 ladder

Spikogram view of Y-filer data

FSS-i3 data review can be saved as a “Batch”

User-Customizable Plate Layouts

Work Performed at NIST

Allele Concordance Studies at NIST

• Manual calls
  – with GeneScan/Genotyper v3.7
  – with GeneMapper ID v3.2

• Automated calls with GM/FSS-i3

• Comparison of output with Excel spreadsheets written by Dave Duewer (NIST)

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Single Source Samples Examined with i-STRess

- We have previously examined 262 Identifiler samples with v4.0.1. Excellent concordance was found and the results can be found in past presentations:

- In this presentation I will focus on results from 656 PowerPlex16 samples run with v4.1.3.

Data Comparison Between Methods

- Dave Duewer (NIST Analytical Chemistry Division) has written several computer programs to convert and compare FSS-i3 data that utilize Excel macros
  - DNA_FSSi3_Convert.xls (converts data format)
  - STR_MatchSamples.xls (compares samples)

- These programs will be made available to the community after additional testing and refinement

DNA_FSSi3_Convert.xls

First five columns in FSS-i3 output are converted to be like GeneMapper ID allele designation table

<table>
<thead>
<tr>
<th>Batch ID</th>
<th>Sample ID</th>
<th>Locus ID</th>
<th>Major Designation 1</th>
<th>Major Designation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP16</td>
<td>GT37465</td>
<td>D551568</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Mpp16</td>
<td>GT37465</td>
<td>THO1</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>PP15</td>
<td>GT37465</td>
<td>D21511</td>
<td>29</td>
<td>9</td>
</tr>
<tr>
<td>Mpp15</td>
<td>GT37465</td>
<td>D21511</td>
<td>29</td>
<td>9</td>
</tr>
<tr>
<td>PP16</td>
<td>GT37465</td>
<td>D19551</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Mpp16</td>
<td>GT37465</td>
<td>D19551</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>PP16</td>
<td>GT37465</td>
<td>Penta E</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Mpp16</td>
<td>GT37465</td>
<td>D551610</td>
<td>0</td>
<td>16</td>
</tr>
</tbody>
</table>

Each row is an individual locus

Data Transformation

- Each row is an individual sample

STR_MatchSamples.xls

Two or more data sets can be compared to one another

- Creates a list of all samples that are fully concordant at all loci between the samples being compared

- Similar to i-ntegrity in looking for samples with closest genotypes through comparing each sample to all others

Exact Matches (Full Concordance) Observed with STR_MatchSamples.xls Program

- Unmatched sample type flags discordant calls
- ExactMatch sample type indicates full concordance between FSS-i3 and GeneMapper ID samples

Concordance Evaluation

- PowerPlex16 data collected on ABI 3130xl; 656 samples processed in GeneMapper ID and FSS-i3
- Typed manually with GeneMapper ID
- Same data processed through GeneMapper ID/FSS-i3

- Results from 656 samples compared
  - 613 samples matched with no data review
  - 43 pairs exhibited a mismatch with unedited FSS-i3 results

- Examination of mismatches to determine which rules were fired and if user would be able to make correct calls following editing: All calls were concordant after careful review

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
If No Rules Fired, Were There Any Mistakes?

• Each genotype was carefully re-reviewed manually

• No discrepancies (discordance) were noted in calls based on rules set

• This observation provides confidence that when no rules are fired, data quality is acceptable in the data sets reviewed thus far...

Microvariant alleles are not found in GeneMapper ID bins

There is no "6" allele in software provided bins, but the user can manually add this microvariant bin

Reviewing a Large Data Set

Nice Features
• Rapid check of all allelic ladders and generation of composite allelic ladders
• Rapid processing of data

Cumbersome Features
• Having to click through every sample in order to review rule firings

i-integrity Contamination Evaluation

i-integrity Contamination Analysis

These 2 samples are showing a high match % of 84.38, so we want to examine these samples further

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
These 2 different samples are showing the same profile and thus should be 100% match rather than 84.38%. This discrepancy is due to the presence of homozygous peaks.

i-STReam Summary Sheet


The summary sheet displays mixture calculations and final allele designations

In Summary

The FSS-i3 v4.1.3 upgrade is a vast improvement over v4.0.1

- Opens up capabilities for creating new multiplex kits (Y-filer, PowerPlex Y and custom assays)
- Projects can be saved during data analysis sessions
- i-STReam provides mixture results much more quickly
- It is now possible to alter the plate layouts
- Can edit the input and output templates

Future Plans

- We plan to explore i-STReam capabilities further
- We will run more data sets that are available at NIST
- Release additional software tools on STRBase (Dave Duewer programs) soon
- Publish recommendations on approaches for validation of expert system software

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Acknowledgments

Funding from interagency agreement 2003-IJ-R-029 between NIJ and the NIST Office of Law Enforcement Standards

Software Installation at NIST:
Bob McLaren
Pete Vallone
Jan Redman
Amy Decker
Becky Hill
Dave Duewer

Training at Promega:
Bob McLaren
Kim Huston
Curtis Knox