Outline of Topics to Discuss

- Overview of the FSS-i³ software
- How to create a new multiplex in multiplex manager, i.e. MiniFiler
- Work performed at NIST
  - Single source samples
  - Mixture samples
- Conclusions and future plans

Overview of the FSS-i³ Software

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.

Our publications and presentations are made available at: http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Overview of Software Components

- **i-STRess** 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.

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-STRess**
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**

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.

GeneMapper ID Raw Data Sorted

- **These bins were developed for FSS-i³ by Promega**

FSS-i³ Input Template

- **Custom assays not provided**

Controls what and how information will be brought into the FSS-i³ program.
User-Customizable Plate Layouts

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 Flagged Samples

Change of color

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

Files are created in html but can easily be imported into Excel for sorting and review.

---

**Audit Files Created by FSS-i3 i-STRess**

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

---

**How to Create a New Multiplex in Multiplex Manager**

Creating new multiplexes with Multiplex Manager.
Enter the Name of the New Multiplex

Designate the sequencer

Change parameters in Advanced Ladder Settings

Click Templates button and designate all names of ladders used

Adding new loci and allele designations

Entering first allele in ladder

Dye Colors:
1 = Blue
2 = Green
3 = Yellow
4 = Orange
5 = Red

Enter 1st allele position in ladder

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Add subsequent alleles from initial anchor position

Position from anchor depends on locus repeat number (i.e. D13S317 is a 4 bp repeat, so each allele is positioned accordingly.

Adding variant alleles

The deviation is different for variant alleles.

Complete ladder before designation of virtual alleles

Virtual allele designation, anchor point should be 1st REAL allele.

When the multiplex is complete, all anchor positions are displayed

Creation of a new ruleset for the multiplex.
Adding positive controls to the multiplex

Adding negative controls to the multiplex

Ladder for MiniFiler D13S317

Spikogram view of MiniFiler data

Ladder editor for Y-filer multiplex

Y-filer DYS456 ladder
Work Performed at NIST
Single Source Samples

Allele Concordance Studies at NIST with Single Source Samples

• 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)

Single Source Samples Examined with i-STRess

• We have previously examined 262 Identifiler samples with v4.0.1 and 656 PowerPlex16 samples with v4.1.3. Excellent concordance was found and the results can be found in past presentations:

• In this presentation I will present results from 982 MiniFiler 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 are currently available to the community

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>LOCN ID</th>
<th>Major Designation 1</th>
<th>Major Designation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201097</td>
<td>D38517</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201097</td>
<td>D38500</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201097</td>
<td>AMEL</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>12501</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Mf Flew RAW AA</td>
<td>D201138</td>
<td>D201138</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Each row is an individual locus

Data Transformation

Each row is an individual sample

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Becky Hill – Promega Tech Tour
Multiplexes, MiniSTRs, and Mixtures: Uses for the FSS-i³ Software at NIST

August 2007

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-integrity in looking for samples with closest genotypes through comparing each sample to all others

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

<table>
<thead>
<tr>
<th>Sample</th>
<th>Description</th>
<th>Penta D Penta E ETH1 TP05 VWA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unmatched</td>
<td>GT37019 all data</td>
<td>2.11 5.13 6.7 6.6 11</td>
</tr>
<tr>
<td>Unmatched</td>
<td>GT37019 all FP16 GM samples</td>
<td>2.11 5.13 6.7 6.6 11</td>
</tr>
<tr>
<td>ExactMatch</td>
<td>BC11352 all data, BC11352 all FP16 GM samples</td>
<td>10.11 7.12 6.83 6 14.17</td>
</tr>
<tr>
<td>ExactMatch</td>
<td>GA05070 all data, GA05070 all FP16 GM samples</td>
<td>12.14 7.17 7.9 8.12 14.19</td>
</tr>
<tr>
<td>ExactMatch</td>
<td>GA05071 all data, GA05071 all FP16 GM samples</td>
<td>10.11 11.12 7.03 8.11 19.17</td>
</tr>
<tr>
<td>ExactMatch</td>
<td>GC03594 all data, GC03594 all FP16 GM samples</td>
<td>10.11 12.15 6.7 6 17.18</td>
</tr>
</tbody>
</table>

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

Concordance Evaluation

• MiniFiler collected on ABI 3130xl; 982 samples processed in GeneMapper ID and FSS-i³
• Typed manually with GeneMapper ID
• Same data processed through GeneMapper ID/FSS-i³
• When rules were fired, profiles were reviewed
• Results from 982 samples compared with STR_MatchSamples.xls
• 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 review

Example 1- Microvariant

Unedited FSS-i³ Data:
D21S11 Allele 33.1

No virtual bin for this microvariant was entered into software when MiniFiler was created therefore this allele was not being called in FSS-i³

Example 2- MiniFiler Issue

Unedited FSS-i³ Data:
D16S539 Alleles not called by software

Pref Amp Rule firing due to peak imbalance.

Concordance Studies Reveal Potential Primer Binding Site Mutations with Different Primer Sets

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

- Each genotype was carefully re-reviewed with STR_MatchSamples.xls
- 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...

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

Work Performed at NIST
Mixture Samples

Mixture Interpretation Results

Once the mixture interpretation calculations are generated, the data is imported as major allele designations

i-STReam Mixture Analysis

i-STReam Summary Sheet

The summary sheet displays mixture calculations and final allele designations

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Experiment 1 – MIX05 Data
Mixture Deconvolution

- Several STR kits were used in the study:
  - SGM+, Profiler Plus, Identifier, COifiler, Powerplex 16
- 3100 Genetic Analyzer was used to process the samples
- Data was previously collected and profiles analyzed in GMID v3.2
- Mixture deconvolution:
  - FSS-i³ i-STReam, stand-alone version

Experiment 2 – Replicates and Ratios
Mixture Deconvolution

- Identifier, COifiler, Profiler Plus kits were used
- 1:2, 1:3, 1:5, and 1:8 mixture ratios were prepared
- 6-7 amplification replicates for each mix ratio
  - To determine PCR variation
- How does i-STReam handle this variation?
  - Different results for the same mixture?
  - Incorrect calls?

Total i-STReam Results

- 1:2 ratio worst results
  - 56% Correct Allele Calls
- 1:3 ratio best results
  - 78% Correct Allele Calls
- Drop-out observed in 1:5 and 1:8 Ratios

External Link: [http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm](http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm)
Conclusions and Future Plans

Some i-STReam Observations…

- GeneMapper ID minus A and stutter filters set at zero to allow all alleles into FSS-i3
  - Some minor alleles filtered out as stutter and not called
- Some incorrect calls
  - Incorrect calls can be explained by variation in peak height ratios
  - 26 / 4080 alleles (0.64%)
- Very conservative
  - F designations allow the program to not make a call

In Summary

- FSS-i3 has the capability to create new multiplex kits (Y-filer, PowerPlex Y, MiniFiler and custom assays)
- Dave Duewer software programs are currently available on STRBase: http://www.cstl.nist.gov/biotech/strbase/software.htm
- A total of 2162 profiles have been analyzed using the FSS-i3 software with concordance checks performed. Full concordance has been achieved after careful review.
- In general, FSS-i3 i-STReam is conservative in its mixture deconvolution; however, only 26 out of 4080 allele calls were called incorrectly (0.64%).

Future Plans

- We plan to explore i-STReam capabilities further
- We will run more data sets that are available at NIST
- Publish recommendations on approaches for validation of expert system software

Acknowledgments

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

*Special thanks to Angie Dolph, our summer intern from Marshall U. for the mixture study data

Software Installation at NIST: Bob McLaren

Training at Promega:
Bob McLaren
Kim Huston
Curtis Knox

http://www.promega.com/products/STCR/STCR01/STCR01_TG.pdf