Affymetrix’ Tiling Informatics Strategy

Mike Lelivelt, Ph.D.
Sr. Manager, Informatics Applications

1st European ChIP-on-chip Scientific Forum
Presentation Overview

- Microarray Informatics Workflow
  - How do the pieces fit together?

- Command Console
  - Getting your data off the instrument

- TAS & Expression Console
  - Getting from probe intensities to molecular estimates

- GeneChip-compatible Program
  - Comparing Arrays & Putting Together the Biology

- Affymetrix Developers’ Network
  - Enabling Research Informatics

1st European ChIP-on-chip Scientific Forum
DAT to CEL to CHP

DAT

Scaling & Normalization

CEL

CHP

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Signal</th>
<th>Detection</th>
<th>Detection p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>200604_s_at</td>
<td>1929.0</td>
<td>P</td>
<td>0.000244</td>
</tr>
<tr>
<td>200605_s_at</td>
<td>2823.8</td>
<td>P</td>
<td>0.000244</td>
</tr>
<tr>
<td>200606_at</td>
<td>3270.9</td>
<td>P</td>
<td>0.000244</td>
</tr>
<tr>
<td>200607_s_at</td>
<td>1826.6</td>
<td>P</td>
<td>0.000244</td>
</tr>
</tbody>
</table>

Affymetrix Confidential
*.*ARR – Sample Attributes stored in xml format
*.*DAT – pixel level data
*.*CEL – feature/probe level data
*.*CHP – probe summarized/molecular level data
  – *.RPT now contained in CHP header

File lineage captured by embedded identifiers, which allows files to be moved and renamed.

AGCC brings new file formats so analysis application must be ready.
GCOS data handling

**Database**
- Sample Attributes
- UNC paths to files

**Files in One Folder**
- DAT, CEL, CHP

**Pain**
- No remote drives
- Can’t rename files

Affymetrix Confidential
Files as sole data source + Read-only slave index = AGCC
Use Off the shelf tools to automate transfers
What? New generation instrument control & data management software
  - Will replace GCOS – GCOS supported through 2008

When? Currently in third Beta release
  - Final Beta in April
  - Launch 1.0 in July (download)

How much? All licenses are included in system price
  - Support contracts are available for purchase

Forms/Versions? Available in Standalone & Workgroup versions

How is it different?
  - AGCC is file-based with slave search index
  - Focused on Core:
    ▪ Sample Reg. → Instrument Control → Image processing → CEL files
Ease of Use
- Drop and Scan, Quickly add additional storage, Data Management

Flexibility: Works the way our customers work
- Multiple Workflow and Configuration options

Data Mobility
- Drag & Drop flat files
  - No Import/Export, DTT or CAB files

Cross platform
- Files work on Mac/Unix/Linux/Windows
- Remote workflow access through Web pages

Rapid data searching & Filtering via Index

Separate Primary Analysis Tools
- Stable, Validated Core software
- Rapidly update CHP writers with new algorithms
## Command Console Benefits (2)

**Strong Ties to Console and Downstream Applications**

### Select Probe Cell Intensity Files

<table>
<thead>
<tr>
<th>Name</th>
<th>Biological System</th>
<th>Developmental Stage</th>
<th>Organ</th>
<th>Sample Type</th>
<th>Sex</th>
<th>Tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>3AMH02082316.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>male</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082801.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>female</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082802.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>female</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082803.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>male</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082804.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>female</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082805.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>female</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082806.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>female</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082807.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>unknown</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082808.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>unknown</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082809.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>unknown</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3AMH02082810.CEL</td>
<td>Hemic and Immune</td>
<td>adult</td>
<td>Blood</td>
<td>frozen sample</td>
<td>unknown</td>
<td>white blood cells</td>
</tr>
<tr>
<td>3ARS0207253A.CEL</td>
<td>Reproductive System</td>
<td>adult</td>
<td>Placenta</td>
<td>frozen sample</td>
<td>female</td>
<td>placenta</td>
</tr>
<tr>
<td>3ARS0207263A.CEL</td>
<td>Reproductive System</td>
<td>adult</td>
<td>Placenta</td>
<td>frozen sample</td>
<td>female</td>
<td>placenta</td>
</tr>
<tr>
<td>3ARS02080736a.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Dorsal</td>
<td>frozen sample</td>
<td>male</td>
<td>Dorsal Root Ganglion</td>
</tr>
<tr>
<td>3ARS02080736f.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Dorsal</td>
<td>frozen sample</td>
<td>male</td>
<td>Dorsal Root Ganglion</td>
</tr>
<tr>
<td>3ARS02080772a.CEL</td>
<td>Cardiovascular System</td>
<td>adult</td>
<td>Heart</td>
<td>frozen sample</td>
<td>male</td>
<td>atrioventricular node</td>
</tr>
<tr>
<td>3ARS02080772b.CEL</td>
<td>Cardiovascular System</td>
<td>adult</td>
<td>Heart</td>
<td>frozen sample</td>
<td>male</td>
<td>atrioventricular node</td>
</tr>
<tr>
<td>3ARS02080773a.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Cranial</td>
<td>frozen sample</td>
<td>male</td>
<td>ciliary ganglion</td>
</tr>
<tr>
<td>3ARS02080773b.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Cranial</td>
<td>frozen sample</td>
<td>male</td>
<td>ciliary ganglion</td>
</tr>
<tr>
<td>3ARS02080774a.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Brain</td>
<td>frozen sample</td>
<td>male</td>
<td>globus pallidus</td>
</tr>
<tr>
<td>3ARS02080774b.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Brain</td>
<td>frozen sample</td>
<td>male</td>
<td>globus pallidus</td>
</tr>
<tr>
<td>3ARS02080776a.CEL</td>
<td>Integumentary System</td>
<td>adult</td>
<td>Skin</td>
<td>frozen sample</td>
<td>female</td>
<td>Skin</td>
</tr>
<tr>
<td>3ARS02080776b.CEL</td>
<td>Integumentary System</td>
<td>adult</td>
<td>Skin</td>
<td>frozen sample</td>
<td>female</td>
<td>Skin</td>
</tr>
<tr>
<td>3ARS02080777a.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Brain</td>
<td>frozen sample</td>
<td>male</td>
<td>subthalamic nucleus</td>
</tr>
<tr>
<td>3ARS02080777b.CEL</td>
<td>Nervous System</td>
<td>adult</td>
<td>Brain</td>
<td>frozen sample</td>
<td>male</td>
<td>subthalamic nucleus</td>
</tr>
</tbody>
</table>

- **Current Directory**: `c:\temp\U133A panel`
- **Select Probe Cell Intensity Files**
- **File name**: 
- **Files of type**: Probe Cell Intensity Files
- **Open**
- **Cancel**
AGCC Ready Applications

AGCC Ready Logo for 2007

GeneChip-compatible vendors will support the new Console data formats

- GeneChip-compatible software will be AGCC Ready

GeneChip-compatible data management vendors will support integrated sample registration into AGCC
Expression Software from AFFX

- Expression Console launched Oct ’06
- Creates signal estimates (probe set data)
- Enables quality control monitoring
- Freely available at affymetrix.com
- Fully supported by Affymetrix
- Tiling Analysis Software (TAS) v1.1.02
- Freely available at affymetrix.com
- Writes BPMAP files in GCOS formats
- Writes CHP files in AGCC formats
- Used for both CHiP on chip and transcript mapping applications
- Source code is available
Console Compatible Applications

Command Console
- Replaces GCOS

Expression Console
- 3'IVT, Exon, WT Gene
- Replaces GCOS, ExACT

Genotyping Console
- SNP 5.0 and 6.0 Genotyping Support in Summer
- Copy Number + Genotyping Support in Fall
- Replaces CNAT, BAT and GTYPE

TAS
- Tiling Array Software
- AGCC Ready now

More Updates Coming
- GSEQ
- GTGS
Software bearing this logo will work smoothly with the Affymetrix platform, both today and tomorrow

www.affymetrix.com/genechip/compatible
Compatible Applications in Tiling

- Partek Genomics Suite
- Genomatix ChipInspector
Affymetrix Developers’ Network

- One of the largest collection of software developers’ in life science
  - Thousands and thousands of members

- [www.affymetrix.com/genechip/developer](http://www.affymetrix.com/genechip/developer)

- Monthly newsletters

- White papers, documentation, webcasts

- Sample code & Affymetrix Power Tools

Affymetrix Confidential
Fostering Algorithm Development

- Expression Console
  - RMA, MASv5, PLIER, but no GC-RMA or MBEI
- TAS algorithm, but not MAT
- All available as open source code
- Why isn’t *my favorite new algorithm* in *my favorite AFFX application*?

**AFFX fully discloses its methods by releasing public source code and enables researchers to enhance these methods.**
Tiling Informatics Strategy

- Create bulletproof, easy to use, core software
  - Get the data off the instrument & to the analysis workstation

- Create free primary analysis software
  - One number per biological unit per array

- Compare arrays using GeneChip-compatible (or other) software solutions
  - Analysis isn’t the core competency of AFFX

- Enable others to innovate and develop through the Affymetrix Developers’ Network