



The Way Ahead.™

Affymetrix' Tiling Informatics Strategy

Mike Lelivelt, Ph.D.

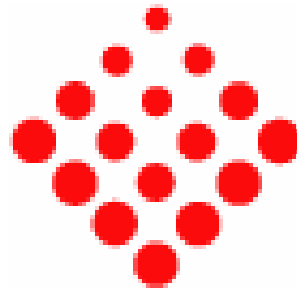
Sr. Manager, Informatics Applications

1st European ChIP-on-chip Scientific Forum



Our Sponsors

Genomatix
understanding gene regulation

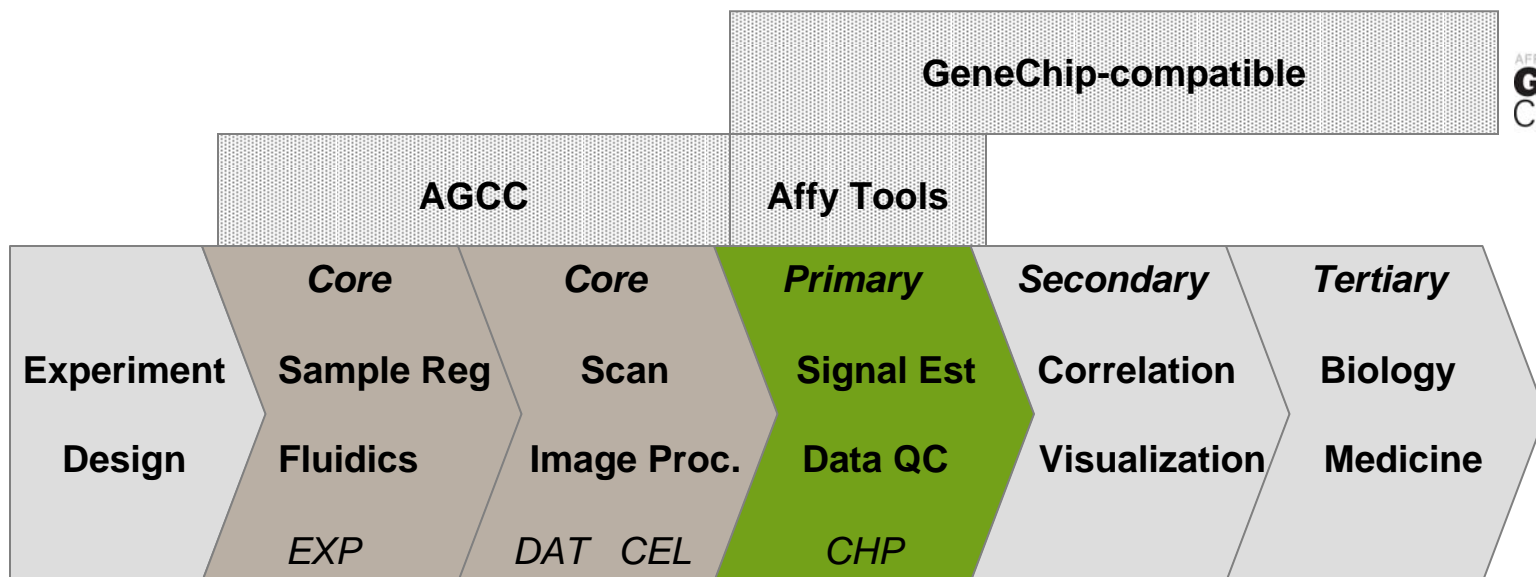


Partek[®]
turning data into **discovery**[®]


GENpathway[®]
The GMP assay experts.

abcam[®]


Microarray Informatics Pipeline



Presentation Overview

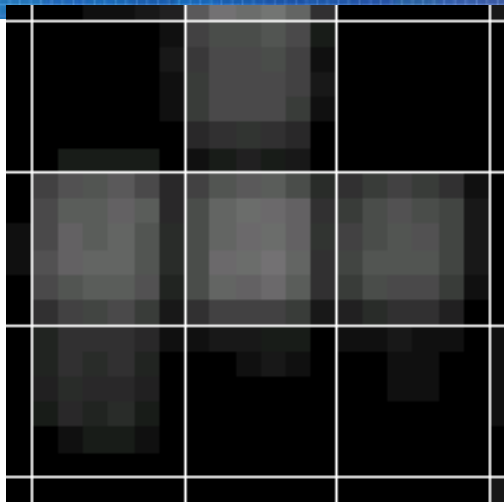
- Microarray Informatics Workflow
 - How do the pieces fit together?
- Command Console
 - Getting your data off the instrument
- TAS & Expression Console
 - Getting from probe intensities to molecular estimates
- GeneChip-compatible Program
 - Comparing Arrays & Putting Together the Biology
- Affymetrix Developers' Network
 - Enabling Research Informatics

1st European ChIP-on-chip Scientific Forum

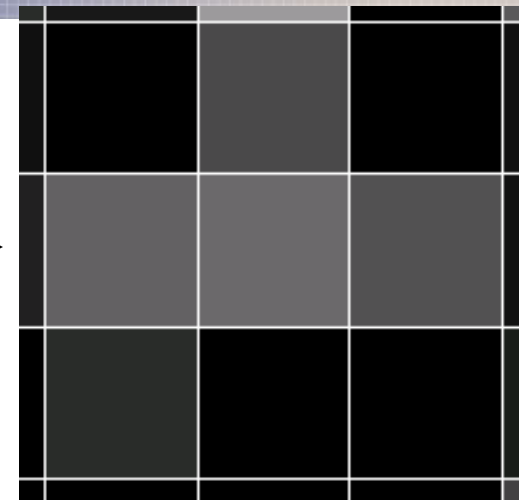


DAT to CEL to CHP

DAT



CEL



Scaling & Normalization



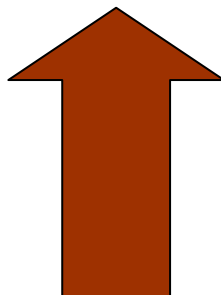
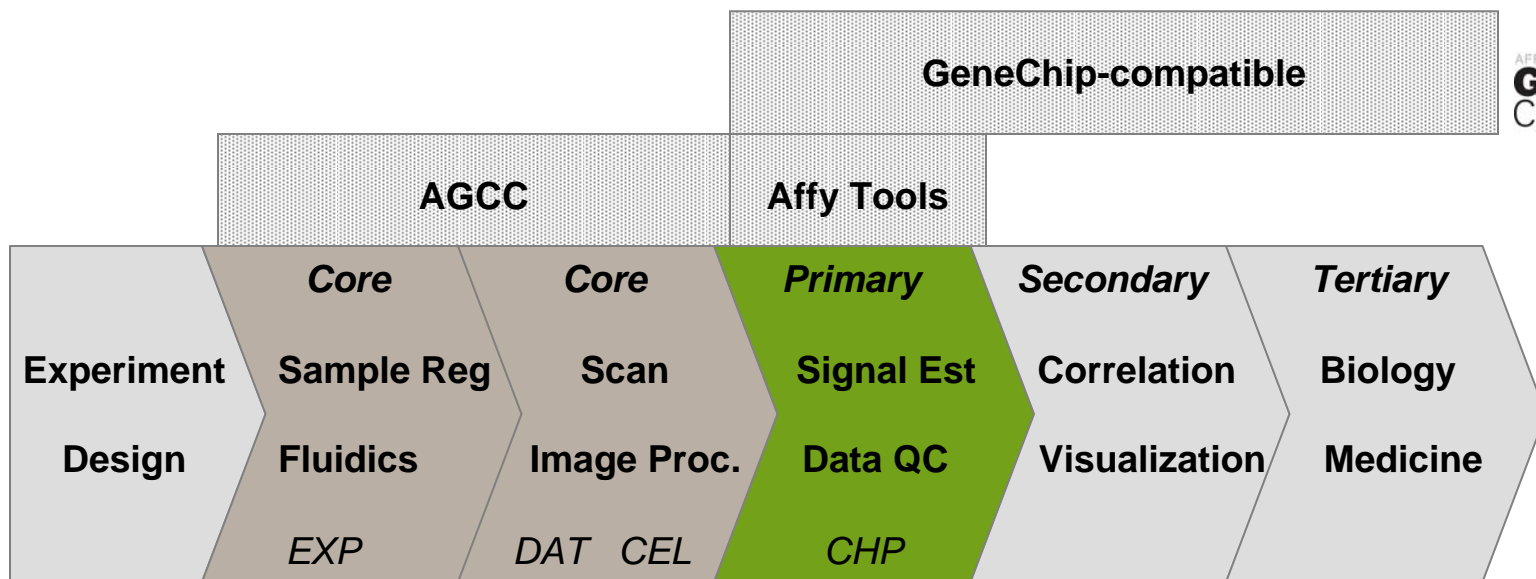
CHP

	HG_U133A_replicate_1			
	Signal	Detection	Detection p-value	
200604_s_at	1929.0	P	0.000244	gt
200605_s_at	2823.8	P	0.000244	gt
200606_at	3270.9	P	0.000244	gt
200607_s_at	1826.6	P	0.000244	Cc

Primer on AFFX File Types

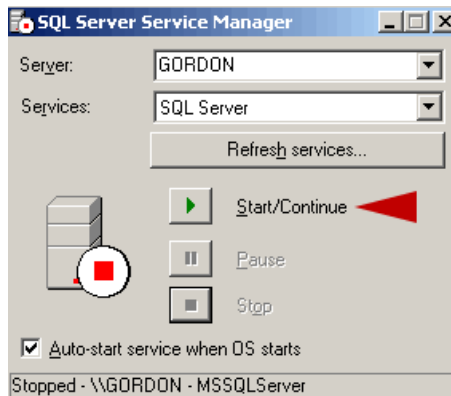
- *.ARR – Sample Attributes stored in xml format
- *.DAT – pixel level data
- *.CEL – feature/probe level data
- *.CHP – probe summarized/molecular level data
 - *.RPT now contained in CHP header
- File lineage captured by embedded identifiers, which allows files to be moved and renamed.
- AGCC brings new file formats so analysis application must be ready.

Microarray Informatics Pipeline



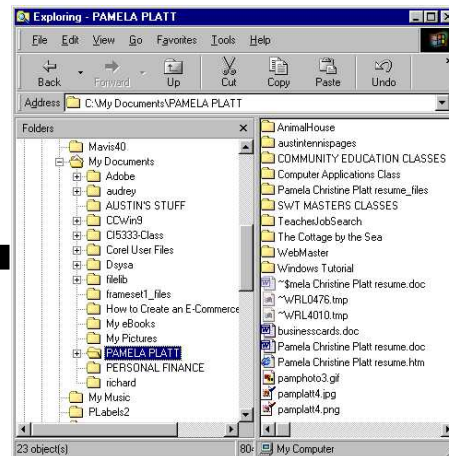
GCOS data handling

Database



- Sample Attributes
- UNC paths to files

Files in One Folder



DAT, CEL, CHP

Pain



DTT & CAB files

- No remote drives
- Can't rename files

Command Console workflows

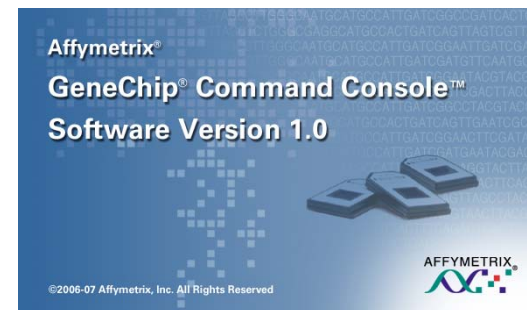
Files as sole data source



Read-only slave index



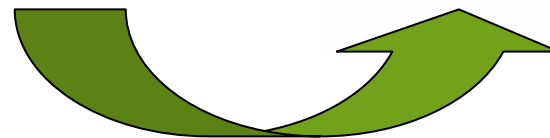
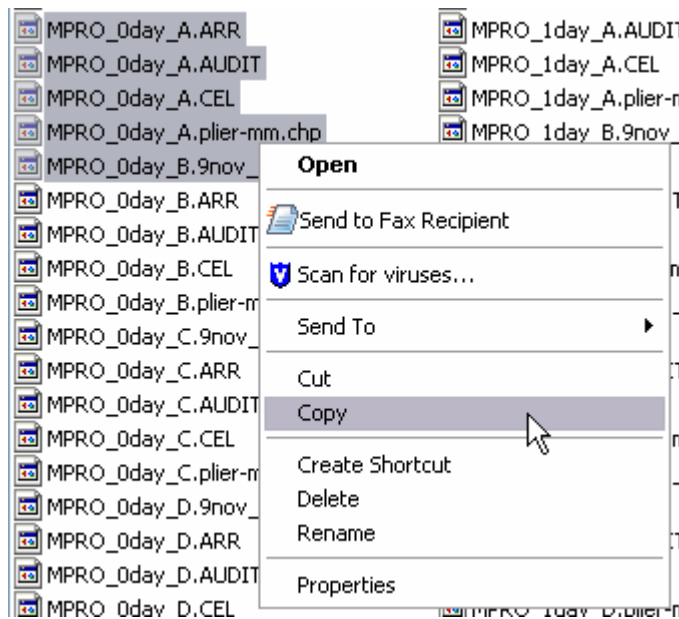
AGCC





The Way Ahead.™

AGCC Data Transfers





The Way Ahead.™

Use *Off the shelf* tools to automate t'fers

The screenshot displays the Allway Sync software interface, version 5.5.8, which is described as "file synchronization that works". The interface is set to a task named "fototime_backup". It shows two source folders: "C:\Users\mike\Pictures" and "\\HB479\FotoTimeStore". A list of 25 files is shown, with columns for file name, size, date, and status. The status column indicates that all files are "Updated". At the bottom, there are buttons for "Analyze", "Synchronize", and "Clear", along with a status indicator that says "Automatic synchronization is on". A summary bar at the very bottom provides statistics: "[3/5/2007 4:40 AM] Summary: Files processed: 21,048; Files copied: 48; Bytes processed: 2,045,011,568; Bytes copied: 418,198,838."

File Name	Size	Date	Status
C:\Users\mike\Pictures\Pics_2007\IMG_1022.JPG	4,137,940	3/4/2007 9:42 AM	Updated
C:\Users\mike\Pictures\Pics_2007\IMG_1021.JPG	4,571,921	3/4/2007 9:42 AM	Updated
C:\Users\mike\Pictures\Pics_2007\IMG_1023.JPG	4,311,321	3/4/2007 9:42 AM	Updated
C:\Users\mike\Pictures\Pics_2007\IMG_1024.JPG	4,366,515	3/4/2007 9:42 AM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0836.JPG.syn	4,073,393	2/24/2007 12:03 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0777.JPG.syn	3,056,211	2/23/2007 2:39 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0848.JPG.syn	4,192,362	2/24/2007 1:04 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_1003.JPG.syn	4,139,836	2/27/2007 4:56 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0905.JPG.syn	4,624,124	2/25/2007 5:37 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0954.JPG.syn	3,764,834	2/27/2007 8:04 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0715.JPG.syn	1,463,662	2/28/2007 10:52 AM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0847.JPG.syn	4,179,094	2/24/2007 12:53 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0745.JPG.syn	6,243,238	2/22/2007 9:16 AM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_1014.JPG.syn	3,889,332	2/27/2007 8:07 PM	Updated
C:\Users\mike\Pictures\Ndx\SyncBackups\IMG_0941.JPG.syn	3,830,000	2/26/2007 10:36 AM	Updated

Command Console (AGCC)

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

Command Console Benefits

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



The Way Ahead.™

Command Console Benefits (2)

Strong Ties to Console and Downstream Applications

Select Probe Cell Intensity Files

Current Directory: c:\temp\U133A panel Up One Level

Name	Biological System	De...	Organ	SampleType	Sex	Tissue
3AMH02082316.CEL	Hemic and Immune	adult	Blood	frozen sample	male	white blood cells
3AMH02082801.CEL	Hemic and Immune	adult	Blood	frozen sample	female	white blood cells
3AMH02082802.CEL	Hemic and Immune	adult	Blood	frozen sample	female	white blood cells
3AMH02082803.CEL	Hemic and Immune	adult	Blood	frozen sample	male	white blood cells
3AMH02082804.CEL	Hemic and Immune	adult	Blood	frozen sample	male	white blood cells
3AMH02082805.CEL	Hemic and Immune	adult	Blood	frozen sample	female	white blood cells
3AMH02082806.CEL	Hemic and Immune	adult	Blood	frozen sample	female	white blood cells
3AMH02082809.CEL	Hemic and Immune	adult	Blood	frozen sample	unknown	white blood cells
3AMH02082810.CEL	Hemic and Immune	adult	Blood	frozen sample	unknown	white blood cells
3ARS0207253IA.CEL	Reproductive System	adult	Placenta	frozen sample	female	placenta
3ARS0207263HB.CEL	Reproductive System	adult	Placenta	frozen sample	female	placenta
3ARS02080736e.CEL	Nervous System	adult	Dorsal ...	frozen sample	male	Dorsal Root Ganglion
3ARS02080736f.CEL	Nervous System	adult	Dorsal ...	frozen sample	male	Dorsal Root Ganglion
3ARS02080772a.CEL	Cardiovascular System	adult	Heart	frozen sample	male	atrioventricular node
3ARS02080772b.CEL	Cardiovascular System	adult	Heart	frozen sample	male	atrioventricular node
3ARS02080773a.CEL	Nervous System	adult	Cranial ...	frozen sample	male	ciliary ganglion
3ARS02080773b.CEL	Nervous System	adult	Cranial ...	frozen sample	male	ciliary ganglion
3ARS02080774a.CEL	Nervous System	adult	Brain	frozen sample	male	globus pallidus
3ARS02080774b.CEL	Nervous System	adult	Brain	frozen sample	male	globus pallidus
3ARS02080776a.CEL	Integumentary System	adult	Skin	frozen sample	female	Skin
3ARS02080776b.CEL	Integumentary System	adult	Skin	frozen sample	female	Skin
3ARS02080777a.CEL	Nervous System	adult	Brain	frozen sample	male	subthalamic nucleus
3ARS02080777b.CEL	Nervous System	adult	Brain	frozen sample	male	subthalamic nucleus

File name: Open

Files of type: Probe Cell Intensity Files Cancel

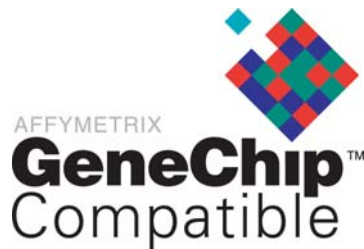


The Way Ahead.™

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

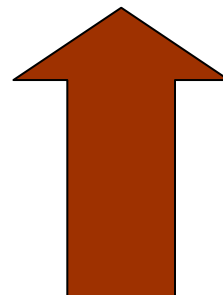
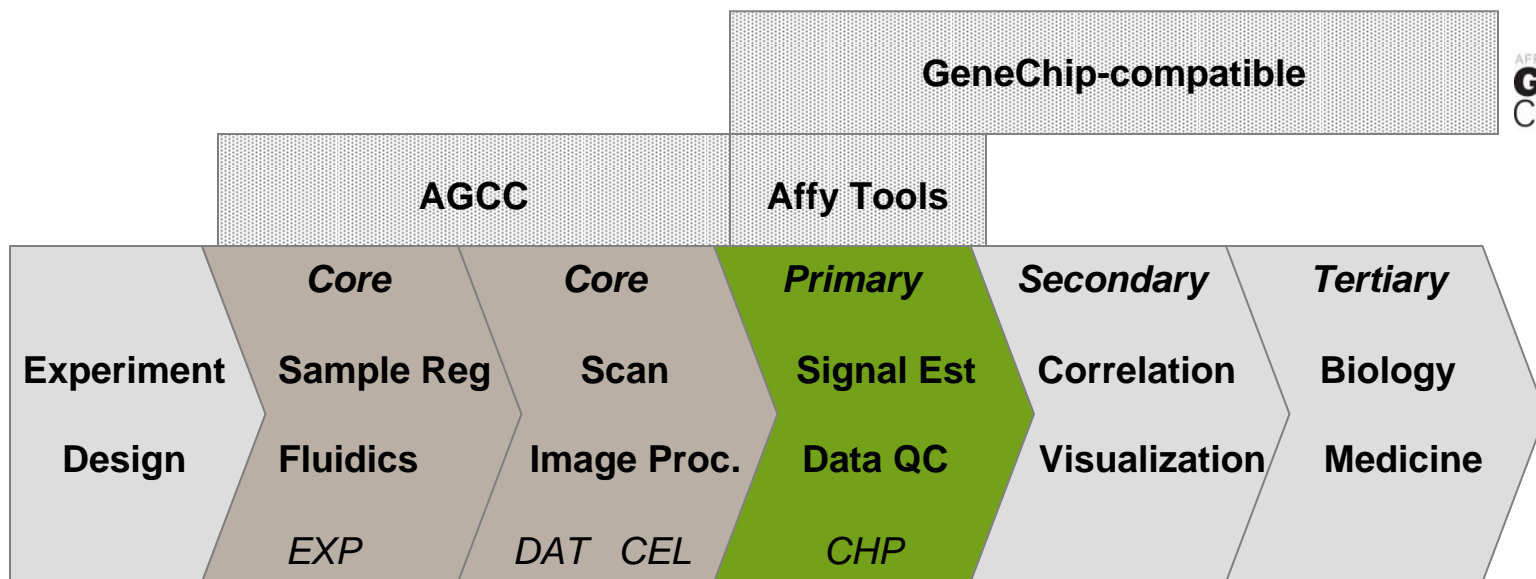
Applied Maths' GeneMaths XT

[GeneMaths XT](#) | [Technical Specs](#) | [Support](#) | [Sales](#) | [Applied Maths](#)

GeneMaths XT is perhaps the most versatile software for microarray expression analysis available today. Its advanced concept of layers and subsets literally gives an extra dimension to expression analysis, by making it possible to handle multiple data outputs, for example repeats, standard deviations, error values, etc., as well as multiple subsets of



Microarray Informatics Pipeline



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

Regulation / Tiling Arrays



- Tiling Analysis Software (TAS) v1.1.02
- Freely available at affymetrix.com
- Writes BMAP files in GCOS formats
- Writes CHP files in AGCC formats
- Used for both CHiP on chip and transcript mapping applications
- Source code is available



The Way Ahead.™

Console Compatible Applications

Affymetrix®

GeneChip® Command Console™
Software Version 1.0



©2006-07 Affymetrix, Inc. All Rights Reserved

Command Console

- Replaces GCOS

Affymetrix®

Expression Console™
Version 1.1



©2006-07 Affymetrix, Inc. All Rights Reserved

Expression Console

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

Affymetrix®

Genotyping Console™
Version 1.0



©2006-07 Affymetrix, Inc. All Rights Reserved

Genotyping Console

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

TAS

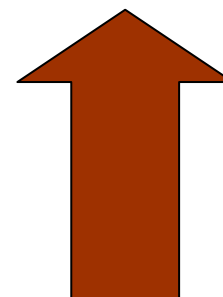
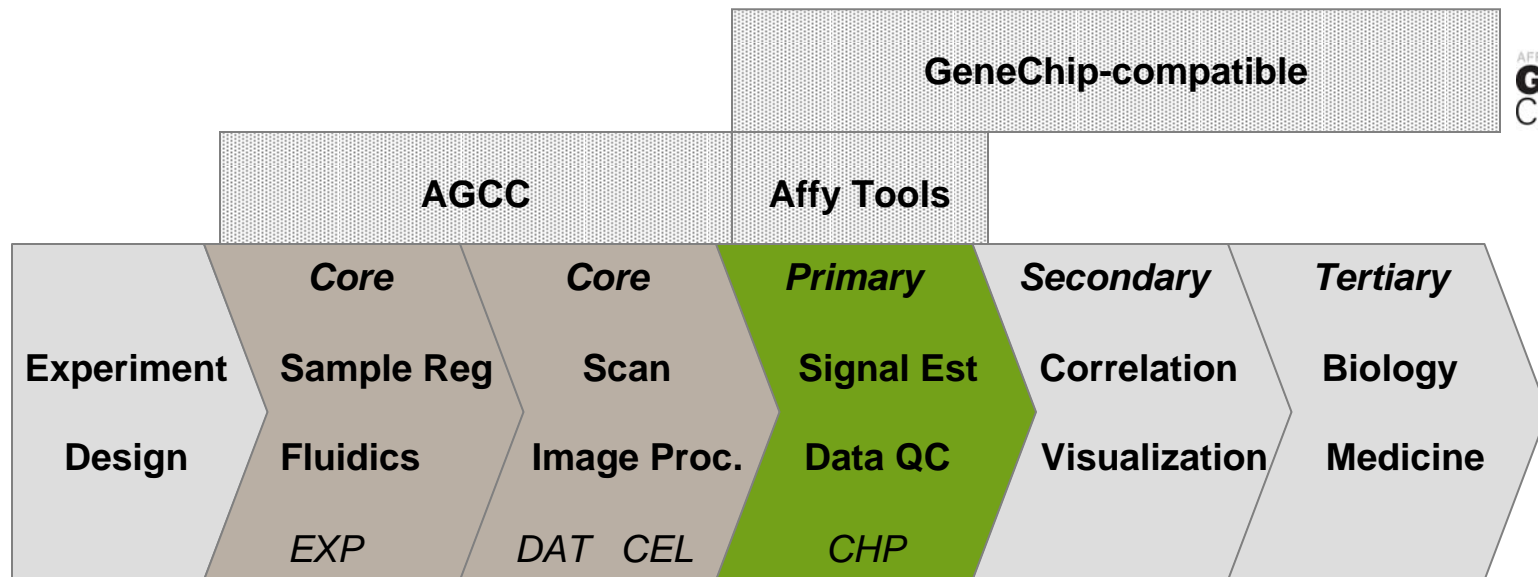
- Tiling Array Software
- AGCC Ready now

More Updates Coming

- GSEQ
- GTGS

Affymetrix Confidential

Microarray Informatics Pipeline





The Way Ahead.™

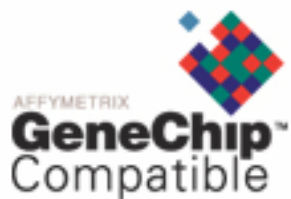
GeneChip-compatible™



AFFYMETRIX

GeneChip™ Compatible

- Software bearing this logo will work smoothly with the Affymetrix platform, both today and tomorrow
- www.affymetrix.com/genechip/compatible



GeneChip-compatible Partners



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
- Monthly newsletters
- White papers, documentation, webcasts
- Sample code & Affymetrix Power Tools



**Developers'
Network**

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