Agilent Technologies

Release Note

Agilent DNA Analytics 4.0.81

New for DNA Analytics 4.0.81

- **Users can create filters and templates in the workflow mode**
  In the 4.0 release, users were not able to create filters while in the workflow mode; they could only do so in the interactive mode. With this release, users can create filters and templates in the workflow mode.

- **Module specific license support**
  In the 4.0 release, concurrent licenses were not properly supported. With this release, different modules (such as the ChIP or CGH modules) can use different license types (server license or a text license) in the same program installation. This allows users to have a concurrent license for one application module, and a single-seat license for a different one.

- **Track visualization improvement**
  In 4.0, genomic tracks loaded in the viewer were placed on the same plane, making it difficult to view multiple tracks if they overlapped. In 4.0.76, the genomic tracks are spaced apart from each other in order to make them more distinguishable.

- **Gene name ON/OFF toggle**
  Users can now turn the display of gene names on/off in reports from the preferences dialog. This makes it easier to view data, and keeps the view from becoming overly confusing.

- **Configurable Interval Margin**
  To define an aberrant interval, it is sometimes desirable to add a "margin" to the interval – this helps to accommodate the fact that aberrant regions may extend beyond the first and last probe of an aberration call.

  In CGH 3.5 a default margin of 200 bp was added to all intervals, accompanied by an option to apply the margin symmetrically or asymmetrically. CGH 3.5 aberrant intervals started from 200 bp before the start position of first probe in interval and ended 200 bp from *start* position of last probe (i.e. typically 140 bp from *end* position of last probe).

  In DNA 4.0.76 there is an option, configurable within the *CGHAnalytics.properties* file, to control this behavior. Length of the margin to be added, if any; and the manner in which it is to be added can both be configured. Default settings are same as CGH 3.5.

- **Design 014693 added to the default data.**
  We have added the design file for the Human Genome CGH Microarray 244A design (014693) as part of the base data package in the application.

- **Track description added to all the default tracks**
  For all default tracks loaded into the application as part of the base data package, descriptors have been added

- **2 tracks (hg18_CNV_Track and hg18_miRNA_Track) removed from the default tracks shipped with application.**

  These genome tracks were duplicates within the application, and are represented by more up-to-date tracks in the base data package.

- **Support for genome builds with alphanumeric chromosome names.**

  As part of this release, we have enabled the application to support all chromosome names, this will allow for the creation of chromosomes for species other than Human, Mouse, and Rat, and will accommodate the creation of a ‘Mito chromosome’ to represent the mitochondrial genome.

- **.BED file format supported in genome build import function**

  The .BED file format can now be used to represent genome builds as part of the genome build import functionality.

Issues from DNA Analytics 4.0.xx, fixed in 4.0.81:

- **Log ratios vary for different genome build data of the same design**

  In DNA Analytics build 4.0.xx, if two design files with the same design file ID (AMADID) were loaded, but each file referenced a different genome build (e. hg17 and hg 18), a problem occurred in the mapping of probes from the Feature Extraction (FE) file to the correct genome. This incorrect mapping led to incorrect log ratios being associated with some genome regions, which in turn could result in the disappearance of aberration calls from one genome build to another.

  This problem occurred only when two or more versions of a design file are loaded. All hg17 only versions and hg18 only versions worked correctly.

  The problem occurred due to incorrect mapping of probe positions which have changed). Now the coordinates from the appropriate genome builds are mapped correctly. With the fix, two versions of a design file can be maintained in the application without creating the observed mapping problem.

  Note If you are importing experiments generated in previous versions of DNA Analytics into the 4.0.81 version(as opposed to importing the design files directly), and the conditions of different genome builds exist, the problem will still persist as no probe coordinate remapping occurs as part of experiment uploads.

- **Expression design gets added to CGH node for UDF.**

  With the fix, expression designs get added to the expression node.

- **Methylation: clicking on gene view resizes the window**

  With the fix, clicking on a gene will not resize the window.

- **Graphical common aberration results different from common aberration text report for ADM2 algorithm**

  With the fix, common aberration results are the same in the graphical and text versions for the ADM2 Algorithm.

- **Intervals are not rendered in UI but reported in the interval summary report and also rendered in the table view.**

  With the fix, intervals are rendered in the UI, in addition to the summary report and table view.

- **Display of overlapping annotations of tracks is like Gene View.**

  With the fix, overlapping annotations are easier to read.

- **Wrong no of probes reported in interval penetrance report.**

  With the fix, the correct number of probes is reported in the interval penetrance report.

- **Removal of Duplicate Gene Names from ‘create genelist’ functionality.**

  Duplicate gene names where generated with the ‘create genelist’ functionality in previous releases, this has been fixed.

- **Handling of “dye flipped” arrays with ADM-2.**

  Since ADM2 utilizes the log ratio and log ratio error information, the log ratio columns for ‘flipped” arrays needs to be flipped. However, no such flip is necessary for log ratio error column due to structure of error model calculation. In DNA Analytics 4.0, the flipping was inadvertently applied to log ratio error column as well. Due to this, the log ratio error was not being taken into consideration and the aberration results were calculated using ADM -1 error model (i.e. the results were correct as per ADM -1 model but the additional refinements using log ratio error present in FE files was not being taken into account) This has been rectified in DNA 4.0.76.

- **CBS “Chromosome 1” Issue**

  When using genomic boundaries in defining regions to search for aberrations, Aberrations would be calculated only for chromosome 1, regardless of the genomic boundary placement. This defect has been fixed so that aberrations will be called appropriately in the defined boundaries.

Known Issues in DNA 4.0.81

1. DNA Analytics 4.0 only renders scatter plots in log scale with a fixed range of 0.25-fold to 4-fold.

Symptom	1. Ratios that are higher than 4-fold, which are very common in ChIP-chip data, are rendered as 4-fold. During visual inspection of the data, this inadvertently under-reports the shape of peaks of some true binding events. 2. The visualization of the noise level is exaggerated. The baseline for "good" ChIP-chip experiments can often vary between 1 to 2 log units, which look relatively clean in dynamic raw- (non-log) - scale, but noisy in fixed log-scale.
Fix/Workaround	None
Notes	Note that this issue is entirely limited to visualization, and that the algorithmic results are not affected. It will be addressed in future versions of the software.

2. In ChIP Module, cannot generate a proper probe report for experiments with replicate arrays

Symptom	Within the ChIP module of analytics, a probe report can be generated for an experiment, however the replicate arrays are not combined, instead are shown together in the report.
Fix/Workaround	None
Notes	None

3. On MAC, QC reports are not getting generated through the interactive mode.

Symptom	On MAC, QC reports are not getting generated through the interactive mode.
Fix/Workaround	QC reports on MAC do get generated in the workflow mode
Notes	None

4. Common Aberration Defect

Symptom	Context Corrected common aberration and aberration filter. The algorithm for context corrected common aberration expects the candidate intervals to be consecutive. When one applies aberration filters, it is not guaranteed that the filtered intervals would be consecutive. Thus. In such cases, there is an message shown (“Invalid Intervals”)
Fix/Workaround	Perform context corrected common aberration without aberration filters. The report generated contains the aberration summary (score, number of probes for the COMMON interval and p value) . Further selection would have to be done on this report.
Notes	None

5. Log ratios vary for different genome build data for uploaded experiments

Symptom	If you are importing experiments generated in previous versions of DNA Analytics into the 4.0.81 version (as opposed to importing the design files directly), different genome builds are referenced in each of the experiments, and all other conditions are the same, a difference in log ratio may be observed.
Fix/Workaround	Import the design files directly into the application, and re-create the experiment(s).
Notes	None

Preloaded Data

These are the datasets that are preloaded into the DNA Analytics 4.0 application:

Type	AMADID	Genome Build	# Probes (Size)	FE Files
CH3	014791	Hg18	199399 (200K)	No
CGH 3.0	Design113568592115 2101915_hg17	Hg17	6044 (6K)	41 3.0 files
CGH	014698 (Both hg17 and hg18)	Hg17 Hg18	Hg17 = 99785 (100K) Hg18 = 99805 (100K)	1. US22502705_251469814934_S01_CGH-v4_95_Feb07_1_1 2. US22502705_251469814934_S01_CGH-v4_95_Feb07_1_2 3. US22502705_251469814935_S01_CGH-v4_95_Feb07_1_1 4. US22502705_251469814935_S01_CGH-v4_95_Feb07_1_2
CGH	014950	Hg18	43016 (44K)	No
CGH	018897	Hg18	230499 (244K)	No
CGH	018898	Hg18	227599 (244K)	No
CGH	014693	Hg18	237834 (244K)	No
Exp 3.0	Design113568922571 319439176_hg17	Hg17	6044 (6K)	41 3.0 files
Exp	014850	Hg18	42816 (44K)	No
Exp	014868	Mm8	42993 (44K)	No
ChIP	014706	Hg18	237198 (244K)	No
ChIP	014707	Hg18	237194 (244K)	No

System Requirements

- Windows XP or Server 2003 with Pentium 4 or later *or* Mac OS X 10.4.x with JVM 1.5 *or* Redhat Linux 9
- 2GHz CPU or greater, dual core CPU recommended
- 4GB RAM minimum (need more when working with 244K or higher microarrays)
- 40GB Free Disk Space
- 1280x768 Display minimum

Installation

Part 1. Install the software (Windows XP)

1. Start the installation.
    - *If you are installing from a CD, put the CD into the CD drive. The setup program will start automatically.*
    - *If you are installing from another location, run the installation program on the computer on which you want to install the program.*

2. Read the Introduction screen and click **Next**.
3. Read the License Agreement, click **I accept the terms of the License Agreement**, then click **Next**.
4. In the Choose Install Folder dialog box, click **Next** to accept the default location for installed files, or click **Choose** for a different destination.
5. In the Choose Shortcut Folder dialog box, select where you want the shortcut for accessing the application to be displayed, then click **Next**.
6. In the Pre-Installation Summary dialog box, review the list of installation parameters, then click **Install**.
7. When the Install Complete screen appears, click **Done**.

Part 2. Change memory settings, if necessary

If you want to maximize the speed of processing, you can change the memory setting for the "heap size" of several processes. The heap size is controlled by two flags, –Xms<size> and –Xmx<size>. JVM starts with -Xms amount of memory and can grow to a maximum of -Xmx amount of memory. The 32-bit machine JVM does not support over 1400MB.

To change memory settings for running DNA Analytics 4.0

1. In Notepad, open the file ".../Program Files/Agilent/DNA Analytics 4.0/DNA Analytics 4.0.lax".
2. Find the line "lax.nl.java.option.additional=-Xms1024m -Xmx1400m -Dsun.java2d.noddraw=true".
3. Change the flags to your preferable memory setting. For example, if you have 2 GB RAM, change the line to read "lax.nl.java.option.additional=-Xms1400m -Xmx1400m -Dsun.java2d.noddraw=true".

   Note: *Make sure the letter m is present at end of size and there is no space between the number and m.*

To change memory settings for Workflow mode

1. In Notepad, open the file ".../Program Files/Agilent/DNA Analytics 4.0/config/config_workflow.properties".
2. Find the property, HEAP_SIZE=-Xmx1200m.
3. Change "1200" to your preferable memory setting. For example, if you have 2 GB RAM, change the property to HEAP_SIZE=-Xmx1400m.

To change memory settings for the background process for importing FE data files

1. In Notepad, open the file ".../Program Files/Agilent/DNA Analytics 4.0/config/config_FEImport.properties".
2. Find the property, HEAP_SIZE=-Xmx512m.
3. Change "512" to your preferable memory setting as you did for Workflow mode.