Frequently Asked Questions

affymetrix
Biology for a better world

Expression Console™ Software

1. **Why can't my favorite GeneChip-compatible™ software or third-party analysis package open CHP files created by Expression Console™ Software?**
   - Expression Console Software writes CHP files in GeneChip® Command Console® (AGCC) Software format by default. Some third-party analysis software packages may only be able to process CHP files in the legacy GeneChip Operating system (GCOS)-formatted CHP files. It is possible to save MAS5.0 CHP files in GCOS format using the ANALYSIS>Advanced Expression Configurations menu item. However, this is only available for CHP files analyzed with MAS5.0.

2. **Why don't I see any metrics or controls after I load my CHP files?**
   - If metric and control information is not included in the header of the CHP file, Expression Console Software will not display that information. This occurs for two reasons: (1) an application other than Expression Console Software created the CHP file and did not write the information to the header or (2) the control probes were not identified prior to the analysis run and were not included in the header. To solve this issue in either case, rerun the analysis within Expression Console Software.

3. **Why are the PLIER/RMA/MAS5 CHP files from Expression Console Software different than those produced by third-party software packages?**
   - Two reasons explain why the numbers may be different between the two software applications. First, the development environment and compilers were different between the two software applications, resulting in minutely different behavior with floating point values. Secondly, the parameters used to call the algorithm may be different.

4. **How do I see the Gene Names for the probe set IDs?**
   - To see the NetAffx® information associated with a particular probe set, view the probe level summaries under the REPORT menu item and double click on the probe ID of interest.

5. **Why doesn't the Zip study utility always include the library files when requested?**
   - The Zip utility only includes library files for exon arrays when CHP files are analyzed as part of the study. If the user creates a study and adds existing CHP files, then the library files will not be included in the Zip archive. To add the appropriate library files, identify them through the CHP file properties in Expression Console Software and then manually add them to the Zip archive.

6. **What is the difference between NetAffx library files and those downloaded from Affymetrix.com?**
   - There is no difference. The library files downloaded from NetAffx are not installed and registered with GCOS, but they are exactly the same files.

7. **I added my exon library files to my library folder, but the application still will not analyze exon arrays.**
   - Examine the messages in the status window to identify the missing library files in Expression Console Software. If the computer is connected to the internet, use the download library files functionality to add the necessary files. In addition to these files, Expression Console Software requires two additional files for exon analysis: the qcc file with control probe information and an exon_analysis_configuration file. These files are automatically included when Expression Console Software is used to download the library files; however, if you manually add files to Expression Console Software, you will need to obtain the qcc and exon_analysis_configuration files from affymetrix.com.

8. **What is the Signal Space Transformation-Robust Multi-Chip Analysis (SST-RMA) analysis algorithm?**
   - The SST-RMA analysis algorithm incorporates pre-processing steps to the CEL files before normalization and summarization with RMA. No changes have been made to RMA. RMA was the default analysis flow through Expression Console Software, version 1.3 for human and mouse transcriptome arrays. Beginning with Expression Console, version 1.4, the SST-RMA algorithm is the default algorithm for human and mouse transcriptome arrays.

9. **Why are we now using SST-RMA as the default algorithm?**
   - Historically, microarrays were perceived to underestimate fold change values when compared to other methods such as RT-PCR. For customers who are filtering using fold change cutoff, the new algorithm addresses this "fold change compression" by applying a GC correction and also by transforming the microarray data signal to a similar signal space of other methods.

10. **Are there any circumstances in which I would benefit from using RMA over SST-RMA with human and mouse transcriptome arrays?**
    - Users will benefit from using the SST-RMA when comparing data (significantly changed gene lists) with those processed using the RMA summarization only.

11. **What can I expect to see with SST-RMA compared to RMA?**
    - When filtering by fold change cutoffs, expect a larger number of differentially expressed genes compared to the standalone RMA. When comparing the number of differentially expressed genes (defined by fold change filters) to other methods, such as RT-PCR and RNA-Seq, expect this number to align more with these methods when using GCCN-SST. There is no impact on sensitivity or specificity of data. SST-RMA was designed to address comparability to other technologies. The same experimental design recommendations still apply when designing an expression study with microarrays.

12. **What effect does SST-RMA have on alternative splicing events as defined in Transcriptome Analysis Console (TAC) Software?**
    - With SST-RMA, more significant changes are observed in the number of expressed genes with alternatively spliced events compared to RMA.

13. **Will the QC controls look different in Expression Console Software with GCCN-SST RMA?**
    - The Relative Log Expression (RLE) boxplots are the only controls that might look different in Expression Console Software.

14. **Will the SST-RMA algorithm be available for other array types, such as GeneChip® Human Genome U133 Plus 2.0 Array?**
    - No, the SST-RMA analysis algorithm will not be available in Expression Console Software for legacy array designs, due to vast amounts of publically available data. Customers who wish to take advantage of this analysis method for other array types should use SST-RMA with Affymetrix® Power Tools (APT) software package.

15. **Will SST-RMA be available in APT?**
    - Yes, the SST-RMA analysis algorithm will be available in APT.

16. **When should data analyzed with RMA be re-analyzed using SST-RMA?**
    - You should re-analyze your data if comparing with RT-PCR or RNA-Seq data using fold change values for filtering.

17. **What are the positive and negative controls?**
    - Positive controls are probe sets designed against putative exons of about 100 housekeeping genes shown to be expressed at detectable levels across a variety of tissues. Since the extent of alternative splicing and transcript expression is not known for all tissues, not all exons are expected to be expressed in all tissues.

- Negative controls are putative introns of the same 100 housekeeping genes chosen for positive controls. These probe sets may be expressed in certain tissues through intron retention. They are not true negative controls. Overall, the positive and negative control probe sets provide a medium-size dataset with expected high and low signal values, respectively. This data set is useful to estimate overall data quality though the Pos_vs_neg_auc value.

18. What quality metrics can I examine to determine if dim or bright hybridizations have impacted my data?

- Use either All_mean or PM mean to assay for hybridization intensity. All_mean is a probe-set metric. PM_mean is a probe-level metric, and is the mean of perfect match raw intensities prior to any transformations, such as normalization or probe summarization. PM_mean and All_mean can be compared to understand the effect that data processing steps have on the average intensity of an array because All_mean has been subject to any data transformations that have been performed during signal estimation and normalization. Apparent outliers only based on PM_mean can be ignored when corrected through data normalization in All_mean.

19. Why is the bgrd_mean (background) value sometimes higher than that of PM_mean (signal)?

- The mean of the background probe signal values is based on background probes defined in the background probe file, which are by default the anti-genomic probes. Anti-genomic probes consist of about 1,000 probes for each level of GC content (0 to 25) with no homology to most studied organisms. This set has a higher GC content than the average probe on the array, and therefore can have relatively higher signal values than the mean of all probes (PM_mean).

20. Can I measure the quality of a single hybridization without the rest of the experiment?

- Affymetrix does not recommend measuring the quality of a single-array performance without consideration for the rest of the experiment. In large-scale expression experiments using similar sample types, researchers are likely to develop their own single-array guidelines on what metric values are predictive of high- or poor-quality samples. However, these guidelines are likely to be dependent on sample type, and Affymetrix is unable to recommend such guidelines for all possible situations. Note that the trend toward favoring model-based signal estimation algorithms (for all microarray experiments even beyond the Affymetrix platform) makes single-array quality determination very difficult due to the necessity of simultaneously analyzing multiple arrays to calculate signal estimates.

21. Which quality metric is most predictive of the quality of hybridization?

- Pos_vs_Neg_AUC is a good first-pass metric. A value below 0.8 is a good indicator that sample problems exist. However, a value above 0.8 does not guarantee that the sample is good.

22. One of my hybridizations results in skewed values for several of the metrics, but the values aren't grossly out of line. I could make a case to keep it in or discard it. What should I do?

- If a sample appears "on the line," it is best to leave the sample in the experiment for analysis and simply flag it as questionable. During the initial analysis, treat all samples uniformly. Once candidate genes have been identified, review how the questionable sample changes data relative to the other replicates within the sample. It is always possible to remove a sample later in the analysis workflows.

Affymetrix, Inc: (US) +1-888-362-2447, +1-408-731-5000 ▪ **(EU)** +44-(0)1628-552550 ▪ **(JP)** +81-(0)3-6430-4020 ▪ **(CN)** +86-21-63915511
eBioscience Products: (US) +1-888-999-1371, +1-858-642-2058 ▪ **(EU)** +43 1 796 40 40 305 ▪ **(JP)** +81-(0)3-6430-4020
USB Products: (US) +1-800-321-9322, +1-216-765-5000 ▪ **(EU)** +44-(0)1628-552600

www.affymetrix.com Please visit our website for international distributor contact information.

For Research Use Only. Not for use in diagnostic procedures.

P/N EMI04825 Rev. 2

© 2016 Affymetrix, Inc. All rights reserved. Please see affymetrix.com/trademarks for a complete list of Affymetrix trademarks. All other trademarks are the property of their respective owners.