



## Summary of the MAQC February-23-2006 Teleconference

### The MAQC Project: Calibrated RNA Samples, Reference Datasets, and QC Metrics/Thresholds for Microarray Quality Control

Teleconference Date: February 23, 2006 (9 am PST/11 am CST/12 pm EST/17:00 GMT)  
Summary Date: February 24, 2006  
Author: Leming Shi, +1-870-543-7387, [Leming.Shi@fda.hhs.gov](mailto:Leming.Shi@fda.hhs.gov)  
MAQC Website: <http://edkb.fda.gov/MAQC/>

- 1. Flags and Offset Values:** Manufacturers' latest (additional) instructions for handling flags and offset values have been posted for download ([FlagsOffsetHandling.zip](#)).
- 2. V2\_norm\_ILM:** Adding an extra offset value of 15 based on the originally distributed ILM data; has been posted for download.
- 3. V2\_norm\_GEH:** (1) Adding an offset value of 0.3 to each data point; (2) The request of GEH and GEH\_2 for replacing old GEH\_2 data with new data was extensively discussed and approved. According to Rick Jensen and Rich Shippy, the old GEH\_2 data were generated with protocols that significantly deviated from GEH's recommended SOP's for the MAQC project. Such deviations included the use of a wrong dye (SA-Alex instead of SA-Cy5) and a wrong washing solution (water instead of 0.1x SSC). In addition, Rick provided a PowerPoint slide that clearly demonstrated the dramatic decay of signal intensity during the scanning process while the old GEH\_2 data were collected in September 2005, a strong indication of the effect of ozone. GEH\_2 site repeated 20 arrays and collected new data on February 17, 2006 by exactly following the GEH recommended SOP's for the MAQC project and submitted the new data to FDA/NCTR on the same day. Because of the obvious problems with the old GEH\_2 data and the availability of the new GEH\_2 data that were generated by following the GEH SOP's and the fact that there will be no additional delay in final data analysis, the majority of the MAQC group on the TC approved the request for replacing new GEH\_2 data for MAQC data analysis. However, it was also agreed that in the main manuscript we should clearly state the importance of following manufacturer's SOP's and what exactly happened to the old GEH\_2 data, which will also be made publicly available along with other MAQC data. Therefore, the new GEH\_2 data should be used in all MAQC data analysis. The V2\_norm\_GEH data set has been posted for download.
- 4. One-One Sequence-mapping Table:** The MAQC group greatly appreciated the efforts of the NCBI team (Jean, Danielle, and Damir) for creating a comprehensive probe sequence mapping table with different categories of mapping quality (the file, [MAQC\\_mapping\\_2006\\_02\\_19.zip](#), can be downloaded from MAQC data download site). Such information will be very helpful for people to explain the discrepancies that may exist across different platforms. However, data analysis sites expressed strong desire of using a simple, one-one mapping table so that an overview of cross-platform comparability can be easily obtained. The one-one mapping table distributed by Rick before Palo Alto meeting was based on using 3'-closeness as the

major criterion to pick one probe (or probe set) from multiple perfect matches, a criterion previously agreed upon by all manufacturers and confirmed again during the TC. The MAQC group decided to continue to use a cleaned-up version (by removing 47 redundant records) of the one-one mapping table originally provided by Rick for any cross-platform data comparison across all manuscripts except for MS-5 (the mapping paper). The new table (`reference_mapping.zip`, with 14885 records for ABI, AFX, AG1 (AGL), GEH, ILM, NCI, and TAQ) is ready for download. Mapping tables for EPP, GEX, and QGN can be found in `EPP_GEX_QGN_reference_mapping.zip`.

5. **Manuscripts:** Each manuscript team leader was sent instructions (along with the current publication plan) to submit an initial draft (or very detailed outline with description on the types of figures/tables) to its team members (and copy to Leming Shi) by mid-night Tuesday, Feb-28-2006. A manuscript topic will be automatically eliminated from the publication plan for any further consideration for submission to *Nature Biotechnology* if the deadline is not met. The e-mail (and the plan) was also forwarded to the primary contact from official platform providers, RNA providers, and official data analysis sites.
6. **Others:** Summaries for TCs (Jan-12-06, Jan-26-06, and Feb-9-06) and Boston meeting have been posted at the MAQC web site.

Next Teleconference:

**Thursday, March 9, 2006** (9 am PST / 11 am CST / 12 pm EST / 17:00 GMT)

USA Toll Free Number: **888-566-5020**; International caller: +1-210-795-9594

PASSCODE: **79451**