Summary of the 4th Face-to-face MAQC Project Meeting
February 3-4, 2006, Boston, MA

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

Meeting Date/Place: February 3-4, 2006, Boston, MA
Summary Date: February 20, 2006
Author: Leming Shi, +1-870-543-7387, Leming.Shi@fda.hhs.gov
MAQC Website: http://edkb.fda.gov/MAQC/

Special thanks to Rick Jensen and his staff at UMass Boston for organizing and sponsoring this great event!

1. **Overview**: The 4th face-to-face MAQC project meeting, "by-invitation-only", was successfully held at a well-equipped computer laboratory in UMass Boston. Fifty-one participants attended the meeting, and each participant had access to the internet via a hardwired workstation or a laptop. Participants were able to share/present their analysis results and presentations in a live mode; such an arrangement was found to be very effective. Rick Jensen highlighted the specific tasks for the meeting. Thanks to session chairs for keeping the meeting on schedule: Rick Jensen, Uwe Scherf, Federico Goodsaid, Weida Tong, and Leming Shi. The most important thing is that our momentum has been maintained (and even accelerated); each manuscript team has been working very hard according to the firm timelines.
2. **Apparent Lack of Reproducibility of Reported Microarray Results**: Leming Shi gave a presentation on "*The enigmatic problem in microarray technology: The origin of chaos in data analysis*". Based on Leming's analyses, the apparent lack of reproducibility of microarray results (*i.e.*, gene lists) reported in the literature was mainly a result of inappropriate data analysis practices that put statistical significance (*P*) above scientific common sense (FC). Leming's presentation invited lively debates during and after the Boston meeting. Federico gave a presentation on "*Standard requirements in the validation of genomic biomarkers*".
3. **Specific Issues**: During the rest of the 1st day and part of the 2nd day, presentations and discussions were centered around specific issues that must be addressed before data analysis could be finalized. Such issues included QC metrics/thresholds (Federico Goodsaid, Wendell Jones, *et al.)*, definition of original (non-normalized) data set (Shawn Baker), handling absent calls (Andrew Sun) and offset values (Rich Shippy, Shawn, *et al.*), identification of differentially expressed genes including normalization and gene ranking methods (Leming Shi, James Chen, Russ Wolfinger, Sue-Jane Wang, *et al.*), cross-platform probe-sequence mapping and data comparison (Damir Herman and Jean Thierry-Mieg), applications of MAQC outcomes for performance validation (Mike Wilson and Janet Hager). Consensus and/or convergence was reached for some of these issues. Each platform provider was given a last chance to provide instructions to analysis sites on handling its data including

flags, control probes, offset values, *etc*.. These instructions will be distributed at the MAQC data download site as soon as possible.

4. **Manuscript Proposals**: The 2nd day of the meeting was focused on the discussion of preparing a set of manuscripts for submission to *Nature Biotechnology*. Each team of the proposals was given an opportunity to present the rationale and focus of each manuscript topic, followed by extensive discussions. Proposals MS-4 (Applications of MAQC outcomes) and MS-10 (Cross-hybridization) were dropped, and the topic of cross-hybridization will be merged into MS-5. Those manuscripts that do not meet the timelines will be automatically eliminated before submission to *NBT*.

No.	Topic	Team Leader
MS-0	Editorial	*Nature Biotechnology*
MS-1	Pharmacogenomics and the U.S. FDA's Critical Path Initiative	Janet Woodcock and Dan Casciano
MS-1A	Data quality in genomics	Ron Davis and Hanlee Ji
MS-2	U.S. FDA's VGDS and IPRG	Felix Frueh
MS-2A	U.S. EPA efforts to develop a framework for the use of genomics data in regulatory and risk assessment applications	David Dix
MS-3	MAQC main manuscript	Leming Shi
MS-5	Sequence mapping	Jean Thierry-Mieg
MS-6	The stability of gene lists from a microarray study	Leming Shi
MS-7	Quantitative technologies for validating microarray results	Federico Goodsaid
MS-8	Titration mixtures for assessing relative accuracy of microarrays	Rich Shippy
MS-9	Modeling technical variation	Walter Liggett
MS-11	One-color versus two-color	Tucker Patterson
MS-12	Spike-ins for array quality assessment	Weida Tong
MS-13	Identification of reproducible gene lists from real-world rat toxicogenomics data sets: a cross-platform study	Lei Guo
MS-14	Reproducibility analysis for microarray experiments	Sheng Zhong

5. **Publication Plan with *Nature Biotechnology***: Dr. Gaspar Taroncher-Oldenburg, representing *NBT*, attended the Boston meeting and addressed the MAQC team regarding publishing results in a supplemental issue. There should be a major "Article" that introduces the entire MAQC study and the data sets, followed by several "Analysis" manuscripts, each of which focuses on a specific research topic and gives a clear message (statements and recommendations) to readers. Each manuscript will be subject to *NBT*'s routine peer-review process, and *NBT* makes final decisions regarding the format and contents of the final publications. It is essential that each manuscript team strictly follow the tight timelines to ensure timely publication of MAQC results.

6. **Timeline for MAQC Manuscripts**:
   Feb-28-06: Initial draft manuscripts (for circulation within each manuscript team);
   Mar-15-06: 1st draft of manuscripts (for circulation among manuscript teams);
   Mar-22-06: Discuss 1st version of manuscripts in San Diego (after the MAQC panel discussion at the CHI QPCR meeting) (tentative);
   Apr-10-06: 2nd version for internal review (many organizations);
   May-10-06: 3rd version for internal review (many organizations);

May-31-06: Submission of manuscripts;
Jun/Aug-06: Peer review and revision;
Sep-06: Publication

7. **No "Early Access" to MAQC Data Sets**: Considering the fact that about 40 organizations have already been given full access to the MAQC data sets and in order to avoid any potential compromises for the MAQC group to have the results published in a timely fashion, the MAQC group decided not to grant additional organizations access to the data sets until the data are formally deposited in public repositories. Leming Shi acknowledged the interests of the scientific community in the MAQC data sets and felt sorry for being unable to convince the MAQC group to grant permission to those who requested "early access" to the data.
8. **Presentation at IBC's Chips to Hits Conference**: The MAQC group tentatively agreed to present MAQC results at IBC's Chips to Hits conference, to be held on Sept. 25-28, 2006, Boston, MA, in the format of a presentation and a panel discussion. Details will be arranged with IBC.

MAQC-4 Meeting Participants, February 3-4, 2006, Boston, MA

No.	Name	Organization	No.	Name	Organization
1	Baker, Shawn C.	Illumina	27	Mei, Nan	FDA/NCTR
2	Barbacioru, Catalin	Applied Biosystems	28	Papallo, Adam	UMass Boston
3	Bertholet, Vincent	Eppendorf	29	Patterson, Tucker A.	FDA/NCTR
4	Boysen, Cecilie	ViaLogy	30	Peterson, Ron	Novartis
5	Canales, Roger	Applied Biosystems	31	Rowley, Patricia	UMass Boston
6	Chen, James J.	FDA/NCTR	32	Ruppel, Patty	Ianalytics
7	Chudin, Eugene	Illumina	33	Scherf, Uwe	FDA/CDRH
8	Collins, Jim	Agilent	34	Setterquist, Robert A.	Ambion
9	Croner, Lisa J.	Biogen Idec	35	Shchegrova, Svetlana	Agilent
10	de Longueville, Francoise	Eppendorf	36	Shi, Leming	FDA/NCTR
11	Dix, David J.	EPA	37	Shippy, Richard	GE Healthcare
12	Eklund, Aron	UMass Boston	38	Sun, Yongming	Applied Biosystems
13	Goodsaid, Federico	FDA/CDER	39	Szallasi, Zoltan	Harvard/CHIP
14	Guo, Lei	FDA/NCTR	40	Taroncher-Oldenburg, Gaspar	*Nature Biotechnology*
15	Guo, Xu	Affymetrix	41	Thierry-Mieg, Danielle	NIH/NCBI
16	Hager, Janet	Yale University	42	Thierry-Mieg, Jean	NIH/NCBI
17	Haje, Paul K.	TeleChem ArrayIt	43	Tong, Weida	FDA/NCTR
18	Herman, Damir	NIH/NCBI	44	Walker, Stephen J.	Wake Forest University
19	Hueter, Irene	CUNY	45	Wang, Sue Jane	FDA/CDER
20	Jensen, Roderick	UMass Boston	46	Willey, James C.	Gene Express (Ohio Medical University)
21	Jones, Wendell	Expression Analysis	47	Wilson, Mike	Ambion
22	Kawasaki, Ernest	NIH/NCI	48	Wolber, Paul K.	Agilent
23	Lee, Kathy Y.	Applied Biosystems	49	Wolfinger, Russ	SAS
24	Liggett, Walter	NIST	50	Xiao, Chunlin	Applied Biosystems
25	Luo, Yuling	Genospectra	51	Zhong, Sheng	UIUC
26	Ma, Yunqing	Genospectra			