The 4th

MicroArray Quality Control

Project Meeting

FINAL AGENDA

February 3-4, 2006

Boston, Massachusetts

http://edkb.fda.gov/MAQC/

Co-sponsored by:

The MAQC-4 Project Meeting Agenda, February 3-4, 2006
Room S3-127, Science Building, UMASS Boston Campus, Boston, MA

Day One: Friday, February 3, 2006 Chair: **Rick Jensen** (UMass Boston)		
8:00 am	Breakfast	
9:00 am	Chair's remarks	Rick Jensen
9:15 am	The Mysterious ***P***roblem in Microarray Technology: The Origin of Chaos in Data Analysis	Leming Shi (FDA/NCTR)
9:40 am	Standard Requirements in the Validation of Genomic Biomarkers	Federico Goodsaid (FDA/CDER)
Issue 1: QC Metrics and Thresholds: End User's Perspectives		
10:10 am	***Types of metrics:*** 1. Repeatability (Precision); 2. Reproducibility; 3. "Accuracy". ***Types of data comparison:*** 1. Intra-site; 2. Inter-site; 3. Cross-platform. ***Two measurement spaces:*** 1. Expressed genes: P/A; 2. Differential expressed genes: intensity, fold change, gene list, *etc*. ***Surrogates of "accuracy":*** 1. Use of titration mixtures; 2. False positives from self-self comparisons; 3. Concordance between platforms; 4. Concordance with alternative technologies. ***Platform-specific, single array QC metrics:*** 1. many … ***Determination of thresholds:*** 1. How good is good (normal) enough? 2. *A metric without corresponding acceptable thresholds does not help much.*	Discussion Team: **Federico Goodsaid** (Leader) Cecilie Boysen David Duewer Wendell Jones Yuling Luo Weida Tong Mike Wilson Russ Wolfinger Sheng Zhong ***"MAQC Guidance to Data Analysis", Draft V3, Nov-14-2005 (contact Leming Shi for more information)***
10:40 am	**Coffee break**	
Issue 2: Definition of Original, Non-normalized Data ("**Original Datasets**")		
11:00 am	***Background correction*** 1. Negative (missing) values; 2. Offset values; ***Transformation*** 1. log2 2. None? ***Replicating spots*** 1. Averaging? 2. As is? ***Outliers*** 1. Outlier arrays; 2. Outlier test sites; 3. Outlier platforms. ***AFX's "original" probe set data*** • Non-scaled MAS5 output?	Discussion Team: **Shawn Baker** (Leader) Jim Collins Francoise de Longueville Xu Guo Ernie Kawasaki Rich Shippy Yongming Sun Weida Tong

Issue 3: Expressed Genes ("**P/A**" calls / **Flags**)		
11:30 am	1. Summary of "P/A" detection calls; 2. Criteria for excluding flagged genes in analysis; 3. Impact of flagged genes on analysis results; 4. Call for unified flagging notation/definition: P, M, and A? ***Q**: What % of the genome is expressed in a single sample (or detectable by a platform)?*	Discussion Team: **Yongming Sun** (Leader) Shawn Baker; Jim Collins; Francoise de Longueville; Xu Guo; Ernie Kawasaki; Rich Shippy; Weida Tong
12:00 pm	**Lunch**	
Chair: **Uwe Scherf** (FDA/CDRH)		
Issue 4: Differentially Expressed Genes (**D.E.G.**)		
1:30 pm	***Normalization methods*** 1. Original; 2. Mean- (total-) scaling; 3. Median-scaling; 4. Quantile. *(AFX: MAS5, dCHIP, RMA, PLIER16 or apply norm. to non-scaled MAS5 output?)* ***Gene ranking (selection) methods*** 1. Fold change (FC); 2. *P*-value; 3. FC+*P*; 4. *P*+FC; 5. SAM; 6. FDR. • ***"What are we detecting with microarrays?"*** • ***Multi-array based normalization: Why? Practical implications?***	Discussion Team: **Leming Shi** (Leader) Cecilie Boysen Jim Chen Eugene Chudin Lisa Croner Lei Guo Xu Guo Rick Jensen Wendell Jones Walter Liggett (5 min. pres.) Weida Tong Sue-Jane Wang Russ Wolfinger
Issue 5: Cross-platform: Probe Sequence Mapping and Data Comparison		
2:00 pm	***Mapping to RefSeq transcripts*** 1. Database version 2. Mapping results ***Mapping to AceView genes*** 1. Database version 2. Mapping results ***Handling different mapping relations*** • 1-1; 1-n; n-1; n-n. ***Master mapping index*** ***Cross-platform data comparison under different mapping scenarios (3'-bias, probe proximity, etc.)*** • How much does it really matter? ***Call for public release of probe/primer sequences***	Discussion Team: **Damir Herman** (Leader) Shawn Baker Rick Jensen Scott Pine Zoltan Szallasi Jean Thierry-Mieg Chunlin Xiao
Issue 6: Applications of MAQC Outcomes		
2:30 pm	Applications of MAQC reference RNA samples and reference datasets for performance validation	Mike Wilson
2:50 pm	Assessing the performance of H25K platform	Mark Schena
3:10 pm	"Baseline practices document" for the submission and analysis of VGDS data sets	Federico Goodsaid
3:30 pm	**Coffee break**	

Open Discussions/Presentations Chair: **Federico Goodsaid** (FDA/CDER)		
3:50 pm (1.5 hrs)	1. Surrounding the topics discussed so far 2. Terms for "Early Access" to MAQC data sets 3. Conf. presentations (IBC Chips to Hits, *etc*.) 4. Publication cost of supplemental issue 5. Additional discussion items	Federico Goodsaid Leming Shi Gaspar Taroncher
5:20 pm	Overview of updated manuscript proposals	Leming Shi
5:30 pm	MS-3: "Main" manuscript	Uwe Scherf
6:00 pm	Adjourn	

Day Two: Saturday, February 4, 2006 Chair: Weida Tong (FDA/NCTR)		
8:00 am	Breakfast	
9:00 am	Chair's remarks	Weida Tong
9:05 am	*Nature Biotechnology*'s plan for supplemental issue	Gaspar Taroncher
Editorial/Commentary/Perspective: MS-0, MS-1, MS-1A, MS-2, MS-2A		
No Presentations	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 (invited)	*Ron Davis and Hanlee Ji*
	MS-2: U.S. FDA's VGDS and IPRG	Felix Frueh and Federico Goodsaid
9:15 am	*MS-2A: U.S. EPA efforts to develop a framework for the use of genomics data in regulatory and risk assessment applications (proposed)*	*David Dix*
Research Articles: MS-3 to MS-14		
9:25 am	MS-5: Sequence mapping	Damir Herman
9:40 am	MS-6: Impact of normalization and gene selection	Leming Shi
9:55 am	MS-7: Alternative technologies	Federico Goodsaid
10:10 am	MS-8: Titration mixtures	Rich Shippy
10:25 am	MS-9: Modeling technical variation	Walter Liggett
10:40 am	**Coffee break**	
11:00 am	MS-10: Cross-hybridization	Zoltan Szallasi and Rick Jensen
11:15 am	MS-11: One-color versus two-color	Tucker Patterson
11:30 am	MS-12: Spike-ins for array quality assessment	Weida Tong
11:45 am	MS-4: Applications of MAQC RNA samples	TBD
12:00 pm	*MS-13: Validation of MAQC-recommended microarray data analysis methods using real-world toxicogenomics data sets (proposed)*	*Lei Guo* *David Dix* *Leming Shi*
12:15 pm	*MS-14: Reproducibility analysis for microarray experiments (proposed)*	*Sheng Zhong*
12:30 pm	**Lunch**	
1:30 pm (2 hrs 20 min.)	1. Resolving overlaps 2. Filling gaps 3. Task assignments 4. Co-authorship 5. Publication cost	Leming Shi Gaspar Taroncher ***Mar-15-2006***: 1st draft of full manuscripts due.
3:50 pm	Concluding remarks	Leming Shi
4:00 pm	Adjourn	

Many thanks to Dr. Rick Jensen and his staff for organizing the MAQC-4 project meeting.

Location:
The meeting will be held on February 3-4, 2006 at the UMASS Boston Campus (Room S3-127 in the Science Building).

Dear colleague,

You are invited to participate in the MAQC-4 "Data Analysis Jamboree" that will be held on Friday and Saturday, February 3-4, 2006 on the University of Massachusetts Boston Campus (www.umb.edu). Attendance is by invitation only and will be restricted to members of the MAQC data analysis groups and representatives of the participating platforms. In this 2 day workshop we should have ample time to present our (near) final analysis results of the MAQC data; sort out differences in methods, assumptions, and results; and work on the manuscripts for publication.

The workshop will run from 9:00-6:00 on Friday February 3 and 9:00-4:00 on Saturday February 4 in a new, large Computer Laboratory with 22 hardwired workstations and ample room for 20 more laptops with wireless internet access and full audio visual capabilities for presentations. (Room S3-127 in the Science building on the UMASS Boston Campus.) Breakfast, Coffee, Lunch and Afternoon Snacks will be provided both days.

The UMASS Boston campus is easily accessed by public transportation via the Red Line subway ("T"). It has its own stop on the Red Line, three stops south of South Station (the main train station) and is about 30 minutes by public transportation from the main Boston Logan Airport. From the UMASS/JFK "T" stop the free #1 campus shuttle bus runs continuously for the 5 minute ride to the main campus where the Science building is located.

Blocks of rooms have been reserved at the nearby Doubletree Bayside Hotel (15 minute walk, 5 minutes by the continuously running #1 campus shuttle bus) as well as the more upscale Seaport Hotel (25 minutes from UMB by subway). Please call the Hotels directly for reservations for the UMASS Boston MAQC conference.

Doubletree Club Hotel Boston Bayside
240 Mt. Vernon Street
617-822-3600
Rooms on hold for February 2nd, 3rd, 4th
20 rooms each night on hold until January 25th
Rate: $109/night

Seaport Hotel
1 Seaport Lane
617-385-5000
Rooms on hold for February 2nd, 3rd, 4th
10 rooms each night on hold until January 18th
Rate: $159/night
(The Seaport Hotel rooms are at a substantial discount.)
UMASS Boston is also about 20 minutes from the downtown Boston Hotels and 25 and 30 minutes from the Cambridge Hotels near MIT and Harvard on the Red Line subway.

Please confirm your plans to participate by return e-mail (so I can be sure there is plenty of coffee and food).

Also let me know if you have any questions or difficulties getting room reservations.

--- Rick

Roderick V. Jensen
Alton Brann Distinguished Professor of Physics, Biology, and Mathematics
and
Director of the Center for Environmental Health, Science, and Technology
University of Massachusetts Boston

Department of Physics
University of Massachusetts Boston
100 Morrissey Blvd
Boston, MA 02125

Office: 617-287-6032
Dept: 617-287-6050
FAX: 617-287-6053
www.biotech.umb.edu