Meeting Notes Kick-off Emerald

November 23, 2006 EBI, Hinxton, UK 09:00 – 15:00

Present: Martin Kuiper (VIB), Alvis Brazma (EBI), Wolfgang Huber (EBI), Helen Parkinson (EBI), Audrey Kauffmann (EBI), Carole Foy (LGC), Wolfgang Philipp (IRMM), Ulf Landegren (UU), Laslo Puskas (BRC), Arne Sandvik (NTNU) and Vidar Beisvag (NTNU)
Not present: Joaquin Dopazo (CIPF)

Agenda:

1) **Project introduction** – MK
2) **Introduction of participants**
3) **Project objectives** – what to accomplish? All
4) **Presentation of the work packages,** WP coordinators
5) **More extensive discussion on WP3**: the workshops.
6) **The consortium agreement** (MK)
7) **The scientific advisory board**: role and composition
8) **Relationships with MGED, MAQC, ERCC**, other (all)
9) **The Steering committees:** role and composition
10) **Conclusions, next meeting** (MK, all)

1. Project introduction (MK)

The project involves a 10 step process, with the development of Quality Control Metrics, processing ontologies and standards, and the dissemination of these.

2. Introduction of participants

3. Project objectives

Before going through the different workpackages and workshops there was a brief discussion on the bottom-line expectations that the different partners had of the project. Whereas it is obvious that quality metrics and quality assurance measures will be the major outcome of the project, and that dissemination of these will be crucial. However, concerning protocols and best laboratory practices we can do little more than making information about preferred protocols available, as we cannot tell people what to do.

4. Presentation of the work packages, by work package coordinators

WP1 Quality Metrics and Ontologies (Wolfgang Huber, Helen Parkinson; EBI)

Basic QM are already available, these have to be borne out in practice. How do we assess what are useful and what is meaningful? Suggestions were to use QM on large datasets form ArrayExpress.

QM works well for Affymetrix, but need to be extended to fit other commercial platforms like Illumina.

Discussion/open questions/action:

Should we also include QM for home grown arrays? This is harder but possible. NB: such a system has been established by The Norwegian Microarray Consortium in collaboration with Patrick Kemmeren (UMC Utrecht), we may take this approach for integration into AE. We believe that if the data are presented in a standard format, if these data are produced in substantial quantities and if we are able to develop a standard way of doing QM for these formats, it may be useful.

Should we work on QM for other microarray formats than expression arrays? This is for the time being an open question, we should first stick to expression arrays. One obvious QM that is applicable is reproducibility. In the future some of the established QM can possibly be converted to other array formats.

QM output needs to be inspected by humans because different sample types may give different quality of data, but some benchmark data can possibly be established to aid in this. Or, if the datasets stored in ArrayExpress are well annotated to ontologies, they can be used as source for such information. HP informed that MGED Ontology no longer will be maintained, and that they are now working with FuGO OBI. That work is already ongoing. HP is participating in the FuGO/OBI effort. A person will be hired and will work from this upper level common ontology into developing terms specifically for this project.

WP2 Standards (Carole Foy, LGC)

Many initiatives are established like; ERCC, MAQC MGED, ABRF, and MARG. However, we also may want to look at other projects which are more sector-specific like CLGGS, Eurogentest and SoGAT. With standard we mean reference materials and SOPS. From the CA IRMM? LGC hopes to learn what we (the microarray community) really need. As WP2 is closely related to WP5, LGC and IRMM need to work closely together to distribute the tasks.

Discussion/open questions/action:

It will already be very useful if links to these sources are displayed on a project website, thus making the microarray user aware about what efforts are ongoing. LGC will be responsible for the communication between our CA and the organizations behind these initiatives.

WP4 Data quality and systems biology (Martin Kuiper, VIB)

A first idea is to use the QM as a filter, to remove low quality data from total input data, when working with a compendium approach. We will then establish the impact of quality filtering on deriving co-regulation networks from such compendium data. We may also pursue other approaches if these seem to be more appropriate.

WP5 Standards and European Legislation (Wolfgang Philipp, IRMM)

Discussed together with WP2.

WP6 Implications of new technologies (Ulf Landegren, UU)

UL is involved in a centre of excellence connected to genetic analysis and MolTools, an Integrated Project to develop advanced molecular technologies. The contribution of UL will be to develop a database where information about technology developers and standard protocols are stored. This database should be made accessible through the Emerald website.

Discussion/open questions/action:

Should we feed this database with the microarray protocols currently uploaded to AE with the data? This may be worth while, and EBI is going to consider how this could be maintained after UU has transferred the database to EBI.

4. Extensive discussion on WP3

WP3 Organization and Dissemination (Sandvik, NTNU/Kuiper, VIB)

The organization of workshops: Coordination will be done by VB (NTNU) and someone in Martin's group. We have about 165,000 Euros available for organizing these workshops, to ensure that we should be able to get good speakers, and when needed at some WS we can also invite Sci Adv Board members. In addition, we can use some money for participants who need some travel support. Appropriate travel allowances will be reimbursed through VIB. A list of suggested WS 'champions' was presented:

Workshop 1: Organisation of the stakeholder community (NTNU, VIB, all)
Workshop 2: The development of QA/QC, pilot (EBI/LGC)
Workshop 3: Ontology Workshop (EBI)
Workshop 4: Dissemination of QA/QC results (EBI/LGC)
Workshop 5: Certification of standards (IRMM)
Workshop 6: Implications for new technologies (UU)
Workshop 7: Data quality and Systems Biology (VIB, CIPF)
Workshop 8: Dissemination to larger community (NTNU, VIB, all)

Each WP champion has the main responsibility for a draft agenda, including speakers and people to invite. This draft agenda will then be discussed with the Scientific Advisory Board and the Project Board. In general, these WS will be organized as much as possible back-to-back with other meetings relevant for the community. We briefly discussed the type of audience to invite for thee workshops, and came to the conclusion that some should specifically target a selected audience meaning that participants would come 'by invitation only' (closed), while other WS should be organised for all (open).

Discussion/open questions/action:

Workshop 1: Organization of the stakeholder community (NTNU, all): Open

Time: Suggested that this is held before the summer, possibly fitting in with the last MolTools meeting in Heidelberg (4 June could be a possible day), ISMB/ECCB Vienna July or Systems Biology Meeting in Paris. 7-8 June. Martin asks Joerg Hoheisel who will attend the meeting in Heidelberg.
We should decide pace and dates at latest mid December (this will be discussed further through email exchanges). We need to pinpoint this WS by mid-December. Once we have a first plan AB will enquire whether we can get this WS under the umbrella of MGED.
Who we should invite: Key Core Facilities for each EU country and people that need and want the QM. We should identify key representatives of commercial microarray manufacturers, users in the clinic, etc. Possible agenda: Presentation of EMERALD, Standards, QC, External keynote speakers (needed to be interesting), Invite relevant commercial interests/industry, presenters of use cases (pharma, clinical, biomarkers, systems biology); organize discussion on how work should proceed.

Workshop 2: The development of QA/QC, pilot (EBI/LGC)

Closed

Time: Autumn 2007, WS champions should make a survey and send suggestions to MK. Deadline is mid December. Likely this workshop will be held together with MGED10, in Brisbane, Australia. It was suggested to use part of the workshop funding for student travel grants. MK to pursue this with EU officer.

Workshop 3: Ontology Workshop (EBI)

Closed

Time: Spring 2008, WS champion should make a survey and send suggestions to MK. It may be best to organise this together with an MGED meeting as part of an ontology jamboree (Italy?).

Workshop 4: Dissemination of QA/QC results (EBI/LGC)

Open

Time: September 2008 Connected to MGED 11 in Italy.

Workshop 5: Certification of standards (IRMM)

Closed (may also invite other from policy-making area)

Time: Spring 2009, WS champions should make a survey and send suggestions to MK. Deadline mid December.

Workshop 6: Implications for new technologies (UU)

Open

Time: Autumn 2008 / Spring 2009, WS champions should make a survey and send suggestions to MK. Deadline mid December.

Workshop 7: Data quality and Systems Biology (VIB, CIPF)

Open

Time: WS champions should make a survey and send suggestions to MK. We do not need the exact time now.

Workshop 8: Dissemination to larger community (all)

Open

Time: Autumn 2009 WP coordinators should make a survey and send suggestions to MK. We do not need the exact time now.

Dissemination

A prototype of the web page may be up Dec 2006 and an official page will be up and running no later than Jan 2007 (EBI). A mailing list for EMERALD admin will be established as fast as possible (HP), but we will wait to set up a mailing list for the users community until we have delivered something (HP). We should consider dissemination via a journal, e.g. Nature Methods/Biotech, describing what the scope is. UL will contact Nature Methods editor and ask if they are interested in the project. MK will contact the editor of Nature Biotech. We reviewed the list of workshops and WH suggested that there was one event missing: a training session for QM tools. We should consider add some tutorial session at an MGED meeting.

5. The consortium agreement (MK)

Should be signed by all partners at the latest in December.

For the consortium agreement and legal questions: contact barbara.leyman@vib.be

6. The scientific advisory board (MK)

Current proposed members who have agreed to contribute:

Frank Holstege - Utrecht University / Netherlands
Helen Causton - Imperial College London / UK
Rafael Irrizarry - Johns Hopkins University / US
Joerg Hoheisel - DKFZ / Germany
Astrid Laegreid – Norwegian University of Science and Technology / Norway
Marc Salit - NIST/US

This list was approved. The SAB should be consulted every three months or so to review progress and planning of workshops.

Discussion/open questions/action:

Do we need to add some from journals or pharmaceutica? Some people were mentioned/suggested (Bertrand Jourdan, Janet Warrington, Stuart Cole and Charles Auffray). MK will contact one or two to probe their interest. Meanwhile we will see what the experiences are with the current SAB, to see if we need additional expertise.

7. The Steering committees (MK)

We discussed the role of the proposed Steering committees (one for each workpackage). As the many of the people in these steering committees are also project board members it was not clear what these committees could add beyond what the Project Board already contributed to the workpackage progress. We decided to dismiss this additional layer of project steering, and appropriate changes will be made in the Consortium Agreement to describe this (MK).

8. Conclusions

The project is now officially started.

First deliverables to focus on are the web page and WS1.

The next consortium meeting will be together with WS1.

VB 29.11.06