METAcancer

Meeting Minutes

Technical Meeting

November 18th – 20th 2008
at VTT / Espoo, Fi

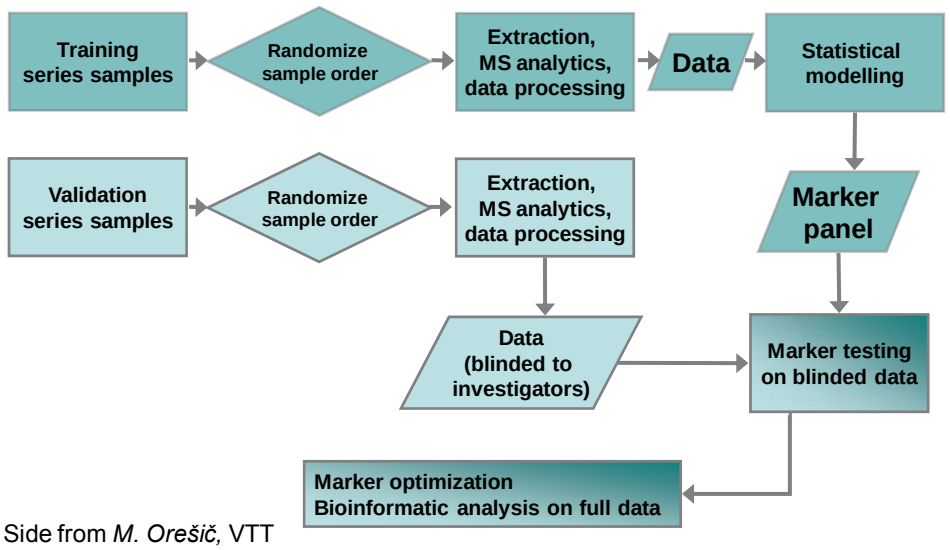
November 18th Technical Meeting

13:30 **Welcome address** by **Dr. Anu Kaukovirta-Norja, Vice President R&D Biotechnology, VTT**

13:40 **State of METAcancer** (C. DENKERT)
Overview on project progress To do list of Kick-off meeting – Follow up
➔ **Session shifted to afternoon**

State of METAcancer, summary on project progress by C. DENKERT

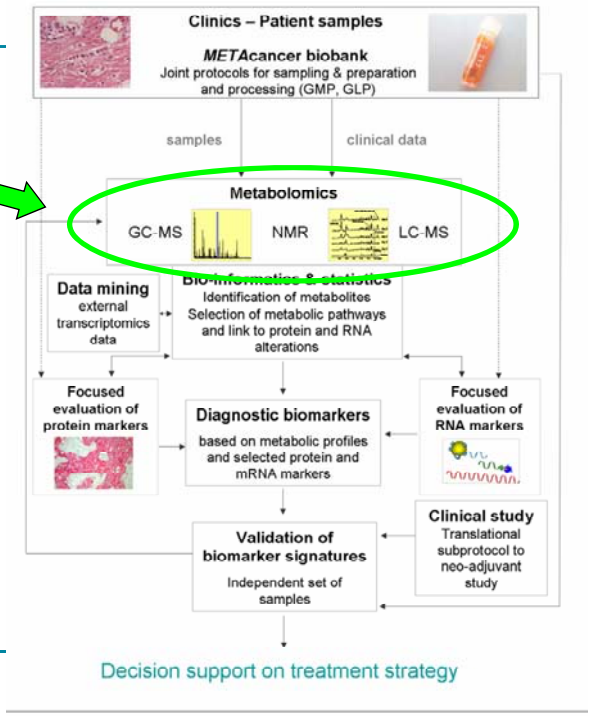
Information workflow



Side from *M. Orešič,* VTT

METAcancer

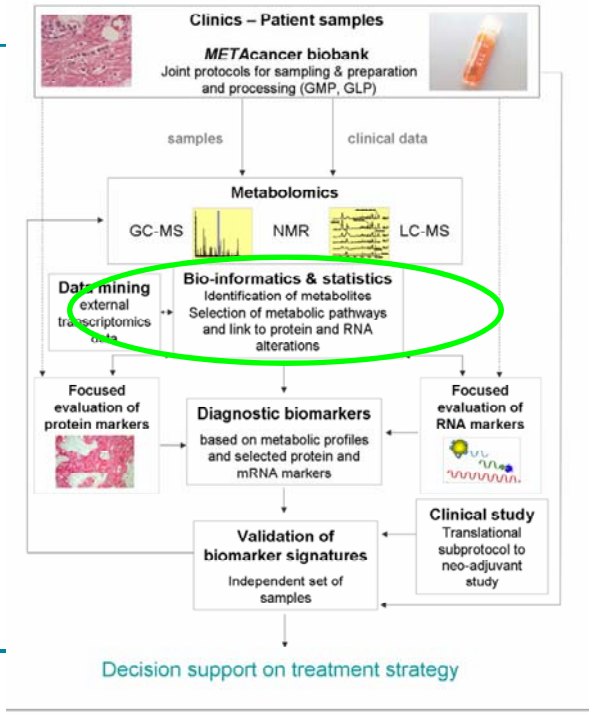
We are here (month 6)



FP7 HEALTH-2007-2.4.1-2

METAcancer

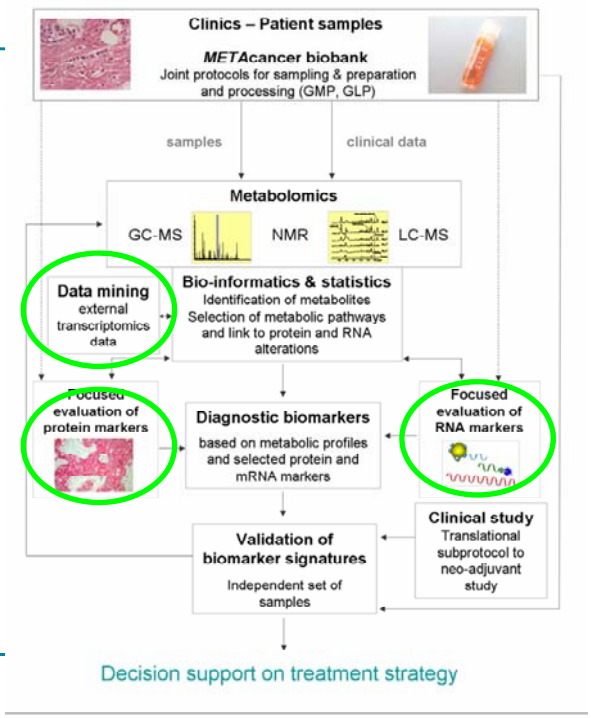
Focus 1



FP7 HEALTH-2007-2.4.1-2

METAcancer

Focus 2



FP7 HEALTH-2007-2.4.1-2

To do list from Kickoff meeting

1. Sample transfer
- Charité: Sent samples within 4-6 weeks (as soon as possible) **done July15th / 8 weeks**
- LSOC samples will be prepared at Charité (transfer to Charité)**done**
- * Cambridge needs 2 samples of frozen tissue **done**
- Consider to sent some Aliquots for extracts from VTT to UCD (for quality control at a later stage, 6 samples)

2. Metabolomics
- Metabolomics data acquisition should be done until September/October, 2008 (if possible) **in progress**
- Consider weighting of samples / protein content determination as additonal normalization strategies
- Consider to exchange protocols (O. Fiehn / RM Salek) for grinding
- Use Pubchem (unique metabolite code)

3. Prepare evaluation **in progress**

Working groups for statistical analysis should be established
- The data evaluation should start approximately October, 2008. Strategies will be discussed at the next Workshop in November 2008

To do list from Kickoff meeting

4. link between genomics and metabolomics **in progress**

Linking the metabolomic data to specific pathways (first step); gene sets (second step)

- * First focus on breast cancer samples (metabolites) -> go back to the general cancer tissues
- VTT group should exchange with Jan Budczies to add more breast cancer datasets with follow up, categorize by treatment if possible

5. Serum samples: **in progress**

- needed: 3 aliquots of 20 (UCD & VTT), 3 aliquots of 150 (Cambridge)
- 100 samples are available at the moment, pre- and postoperative
- ship samples of 400 µl, aliquot at metabolomic site;
- stratify for the Her-2-expression / pCR / hormon receptor (check for clinical data)

6. further meetings

- 3 days November 2008: meeting in Espoo Nov 18-20, 2008 at VTT / Espoo
- 25.-27. May (3 days) 2009: California UC Davis (Oliver Fiehn)

14:00 **WP1: METAcancer biobank and histopathological analysis** (C. DENKERT)

Session shifted to afternoon

T 1.5 Distribution of frozen tissue for metabolic analysis as well as for selected mRNA analysis (Mo 1-3 and Mo 12-18)

T 1.6: Construction of tissue microarrays and distribution of slides for immunohistochemical analysis (Mo 6-18)

D 1.1	SOPs for asservation of samples available	Mo 3
D 1.2	Report on histopathological quality control of existing samples for metabolic profiling	Mo 3
D 1.3	Central database for annotation of samples available	Mo 6

14:30 **WP 2: Metabolite profiling of breast cancer tissues by GC-MS for investigation of predictive signatures and connection to in-situ proteomics** (G. Wohlgemuth)

Presentation of progress related to:

T 2.3: Unsupervised statistical analysis and supervised statistic analysis and link to tumor characteristics and outcome data (Mo 4-10)

T2.4: Integration of GC-MS data with the other metabolic approaches and identification of key regulatory proteins by integration of metabolic data with pathway information from KEGG databases (Mo 10-18)

T2.5: Evaluate protein markers by immunohistochemistry using tissue-microarrays, production of virtual slides and image analysis (Mo 10-24)

D 2.1	Processed GC-MS data set with identified metabolites available	Mo 6

Discussion notes:
- the whole training set will be tested in December (at this time: only "poor quality samples" were tested)
- metabolites have to be present in 80% samples of one class to get a new BinBase entry
- probe table: poor quality / training set / test set -> should we all use the same splitting? Decision: yes
- different content of tumor percentage (Davis / VTT) (weight of probes)
- general normalisations are necessary!

15:00 **WP 3: High resolution 1H NMR spectroscopic analysis of breast tumours to investigate metabolomics in intact tumours** (J. GRIFFIN)

Presentation of progress related to:

T 3.2: Further characterise tumours of sufficient size to also allow solution state 1H, 13C and 31P NMR spectroscopy (Mo 6-12)

T 3.3: Develop a HRMAS 1H NMR spectroscopy database of the tumours with clinicopathological data (Mo 6-18)

Discussion notes:

- *analyzed ca. 40 samples in duplicate so far*
- *amount of tissue: 20 mg tumor -> Different weights of samples between VTT / Davis?*
- *wish list: 25 mg in one sample; only a subset to be run by MAS if large numbers; run samples at RT for higher throughout*

Questions:

- *- How many samples?*
  - What temperature do we run HRMAS?
- *– Which two dimensional spectra are needed?*
- *Databasing results: We have a schema, but how will it interface with the rest of the project?*

- *Jan: Is it possible to correlate the whole spectra with the clinicopathological parameters? -> amino acids ok, fatty acids not so good*
- *Oliver: protein degradation at high temperature? -> how much time between surgery and frozen? (1/2 hour?)*
- *How many probes do we have in the whole project? (300 or more? Second validation cohort?)*

15:20 **Coffee Break**

15:40 **WP 4: LC/MS analysis and validation of metabolic biomarkers in breast cancer tissue samples** (M. ORESIC)

Presentation of progress related to:

T 4.1: UPLC/MS analysis (ESI+ and ESI-) of selected breast cancer tissue lipid extracts for detection of major lipids (e.g. phospholipids, sphingolipids, acylglycerols, cholesterol esters). UPLC/MS analysis (ESI-) of cardiolipins and glycophingolipids from breast cancer tissue lipid extracts. UPLC/MS analysis (ESI+ and ESI-) of hydrophilic components breast cancer tissue samples
(Mo 1-12)

T4.2: Advanced chemometric analysis on dataset to obtain most relevant metabolite features, followed by bioinformatics analysis using megNet system. (Mo 10-18)

Discussion notes:

- *training cohort: "no tumor" – n=7, normal n=14 ??? (definition of subgroups?)*
- *Lipidomics runs will be ready before Christmas*
- *Oliver: Normalization (internal standards, total protein content) -> does it make sense?*
- *Combination of the data!*
- *Target analyzes (cardio lipid acids), relevance of hydroxylated acids?*

16:00 **WP 5: Transcriptomic data mining and pathway analysis** (O. KALLIONIEMI)

Presentation of progress related to:

T 5.1: Definition of the expression pattern of all key metabolic enzymes relevant for the interpretation of cancer metabolomic data by In silico Transcriptomics database (IST) - analysis of 8000 samples (Mo 1-12)

T 5.2: Explore the coexpression profiles and de-regulation of genes along defined metabolic pathways in breast cancer and other cancers (Mo 1-12)

T5.3: Cluster a series of breast cancer samples using pre-defined gene sets based on key metabolic enzymes (Mo 8-24)

Discussion notes:

- *IST- database: "virtual biobank with 43 normal tissue types, 227 cell types, 68 different cancer types, 105 cancer subtypes, ~900 breast cancer samples*
- *Does anybody of us will make gene expression analyses? (budget??)*
- *Sample table: % of adipose tissue, inflammation cells, … is necessary for*

our analyses (Oliver) -> it is in the table! (Check the "partner-version"!)

- *Example 1: genes involved in lipid (fatty acid), example 2: analysis of PPAR-gamma correlating genes in breast cancer (Cancer 2008, may)*
- *Characterization of adipose-tissue specific genes expressed in breast cancer*

16:20 **WP 6: Advanced strategies for identification of metabolites using mass spectrometry** (O. FIEHN)

Presentation of progress related to:

T 6.1: Improve metabolite annotation by GCxGC-TOF. Delivering samples, evaluate results, store and disseminate data, statistical analysis. Sample preparation, data acquisition, raw data processing. Comparison to identified compounds by LC-MS, pathway mapping, NMR analysis of isolated fractions
(Mo 1-12)

T 6.2: Reduce technical noise by novel injectors: Technical implementation; data acquisition; data processing - delivering samples - statistical analysis - pathway analysis (Mo 6-24)

Discussion notes:

- *Presentation of the "setupX-database" (metacancer; password: carstendenkert)*

16:40 **WP 7: Implementation of metabolic profiling into translational research concepts in neoadjuvant clinical studies** (M. KOMOR)

Presentation of progress related to:

T 7.1: Collection of serum samples from GeparQuattro patients in a central serum bank and analysis using different metabolomics technologies (NMR -GC-MS - LC-MS)

D 7.1	Collection and transfer of at least 100 patient serum samples to participating partners. Central documentation of clinical data.	Mo 6

Discussion notes:

- *100 Her-2 neg. samples are ready for shipping to UCAM and will be distributed to the partners from there (17 µl for VTT, 30 µl for Davis)*
- *the clinical data were prepared next week*
- *we are currently identifying approx. 70 pts to build up the HER2 positive group*
- *aims: pCR (for GeparQuattro); What are you looking for in the serum?*

17:00 **WP 8: Management** (P. ZALUD)

Presentation of progress related to:

D8.2h	Set-up of a restricted communication and knowledge sharing platform	Mo 3

Discussion notes:

- *platform set up, presentation of functionality and document repository*
- *presentation of WebEx conference system, will be established on the METAcancer site, to be used to discuss on WP or deliverable level*
- *Meeting planning: next meeting will take place in Cambridge: clarifying final questions related to the 1st contractual reporting for the report; around October 2009, Reza will clarify the date.*

17:10 **WP Training & Dissemination** (P. ZALUD)

Presentation of progress related to:

D 9.1	Interdisciplinary Workshop	Mo 6
D 9.3	PR kit Public Website, Press Release	Mo 2
D 9.5	Internal Progress Survey	Mo 6
D 9.6	Consolidated Literature and Patent Survey	Mo 2

- *concept and rational of internal progress audits*
- *Public Relation: Communication of project's background, results and impact to the public for awareness creation: METAcancer flyer will be designed by tp21*
- *Literature and patents will be put on the internal project website only with abstracts due to copyrights of the editors*

General discussion (C. Denkert)

METAcancer

Aims for discussion

- Discuss the project progress
- Identify and discuss problems
- Establish database structure for combined bioinformatic evaluation
- Refine project structure (train-validation strategy)
- Define questions and hypotheses for scientific publications
- Develop and collect new ideas

METAcancer

Work in progress: Complete data aquisition

- Charité: update of clinical data –until December 15th
- VTT - metabolomic data aquisition completed until christmas
- UCD
  - 290 ‘training set’ Dec 01-Dec15
  - 140 ‘test set’ (=Validation set) Jan 01-Jan 10. Are the samples suitable for analysis?
- UCAM – analysis is ongoing, focus on more defined subgroups

FP7 HEALTH-2007-2.4.1-2 Technical Meeting November 18th-20th, Espoo

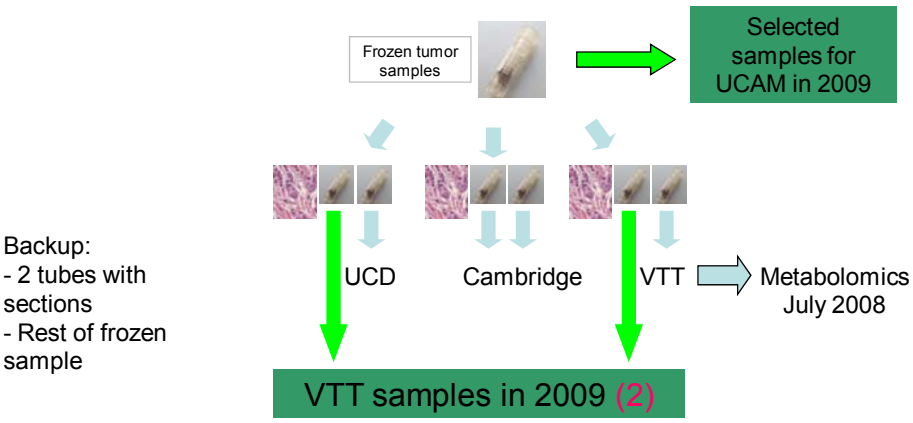
METAcancer

Discussion – sample transfer

- GBG will send serum samples to all partners.
- Decision on additional samples will be made after data acquisition.
- Current status on tumor samples:
  - UCD might not need additional samples
  - VTT needs more samples
  - UCAM needs larger samples in one tube (but not for the complete cohort) (one thicker sample)

METAcancer

Suggestion for sample transfer



Backup:
- 2 tubes with sections
- Rest of frozen sample

METAcancer

Task 1.3 Central database for annotation of samples

- Discussion needed:
  - System for exchange of data/update of data, all partners should have access
  - On METAcancer website (no preliminary data)
  - UCD data should be uploaded by tp21 from SetupX (simply data files)
  - Categories:
    - Clinical data
    - Raw metabolomic data after quality control (before normalization)
    - Normalized metabolomic data
    - Statistical analysis
    - Use pubchem ID
  - What is the best file type? (Excel; large data: spreadsheet)

METAcancer

Why two separate cohorts?

Publication strategy

- Standard reviewer comment in major journals:
  - "results should be validated in an independent set of samples"
  - Problem: independent cohorts are different in clinical data (different times of collection/centers of collection/therapy changes over time)
- Options:
  - Best: two completely separate cohorts
  - Samples should be randomized to each cohorts (Kullback-Leibler divergence)
  - If not possible: multiple random validation approach
  - Project structure should make clear that the cohorts are truly separate
- This is a major change of the workflow compared to the proposal, but quality is much improved.
- There should be a formal steering committee decision on this change. -> ok
- (VTT) replication: a third group would be necessary

FP7 HEALTH-2007-2.4.1-2 Technical Meeting November 18th-20th, Espoo

Please confirm the SSC decision via agreement of these minutes.

METAcancer

Discussion

- Class labels in UCD data
  - Background: UCD uses class labels for randomisation of samples during measurement – ok: use **Sample_Type**
  - UCD uses class labels to exclude rare metabolites from primary data
    - Problem: primary data should be as unbiased as possible
    - Biological subgroups of breast cancer might consist of only 5-10% of samples
    - Some exclusion necessary
    - (most rare metabolites will not be relevant)

    - CD: so many possible classes in breast cancer -> it is impossible to define the classes now before statistical analyses; O.Fiehn: not agree; CD: acquire the data without including the hypothesis for an independent analyses; more strictly define for the first part, the other for part number five

METAcancer

Discussion

- Normalization:
  - UCD: use sum of all known metabolites; US: internal standards are accepted; 30 external standards are used; 2009: mixture of internal and external standards; internal standards could be used
  - VTT: normalization of the proteins; prefer 5-10 internal standards (including weight and proteins), class specific
  - UCAM: used total sum (like UCD)
- Can we agree on a common normalization strategy?
  - VTT & UCD: not possible; O.Fiehn: new available normalization (ordered)

FP7 HEALTH-2007-2.4.1-2 Technical Meeting November 18th-20th, Espoo

METAcancer

Discussion: Integration of metabolomic data to existing expression databases (O. Kall.)

- Sample size for the transcriptomic analyses?
- (CD) Transcriptomic is not planned, but it should be done during the project.
- Analyses can be done here (O. Kall.)
- M.O. (VTT) identify typical samples -> do transcriptomics on these samples, if there is enough material left.
- “Normal tissue” is a mixed tissue -> comparing tumors with an upregulating profile with other tumors.
- Combine IHC (Charité) with transcriptome -> use these results for GBG-trials -> stain these markers in neoadjuvant cohorts
- (OK) Mammaglobine: submitted paper (chemoresponse)
- (CD) BMI is a strong independent predictor for chemoresponse -> biological background?

METAcancer

Discussion: Integration of metabolomic data – the VTT experience (Sandra/ Jarkko, Jing Tang)

- Difference to Binbase: MZmine uses rawdata
- VTT method: elasticnet regression

FP7 HEALTH-2007-2.4.1-2 Technical Meeting November 18th-20th, Espoo

METAcancer

Discussion: Data analysis of metabolomic data – the Cambridge experience (Reza S.)

- 32 components could be identify at the moment (-> for further analyses in VTT / UC Davis on the METAcancer website)
- More precise regarding to quantification than the other methods
- Methanol / ethanol peak -> knife (sample preparation)?
- Typical NMR breast cancer peak: Choline and O-Phospho-Colonin could be detected in poor quality samples

METAcancer

Next meetings

- UC Davis: All cooperative partners and the Breast Cancer Group (UC Davis) -> open forum
- Cambridge: prepare everything for the report; around October 2009
- Metabolomic-meeting: July 5-8 in UK (abstract deadline: 13.02.2009), organized by J. Griffin; 2010: Amsterdam

FP7 HEALTH-2007-2.4.1-2 Technical Meeting November 18th-20th, Espoo

End of the meeting

METAcancer

Technical Meeting & 1st Interdisciplinary Workshop

November 18th – 20th 2008
at VTT / Espoo, Fi

Attendance Record

Name	Institution	Signature
REZA SALEK	U. CAM	
Matej Orešič	VTT	
Mark Rolfe	GSK	
Oliver Fiehn	UCD	
Berit Müller	charité	
JAN BUDCZIES	charité	
GERT WOHLGEMUTH	UCD	
JING TANG	VTT	
GOPAL PEDDINTI	VTT	
HENRIK EDGREN	VTT	

METAcancer

Attendance Record

Name	Institution	Signature
Sirkku Pollari	VTT	
Elmar Bucher	VTT	
Juha Kokkonen	VTT	
Tuulikki Seppänen-Laakso	VTT	
Laxmana Rao Yetukuri	VTT	
Petra Jakob	tp21	
Katja Haug	tp21	
Sandra Castillo	VTT	
Marko Sysi-Aho	VTT	
Catherine Boursaythip	VTT	
Carola Denkert	Charité	
Olli Kallioniemi	VTT	
ZHAO HAN	VTT	