
TRANSBIG project & data analysis

Thierry Sengstag

Bioinformatics Core Facility (BCF)

Swiss Institute of Bioinformatics

Swiss Institute for Experimental Cancer Research

thierry.sengstag@isrec.unil.ch

Bioinformatics Core Facility



- Group started 2002
- Head: Mauro Delorenzi
- Equivalent 6 full-time positions (early 2005)
- Current missions:
 - Help researchers in the design of microarray experiments
 - Supervise/perform microarray experiment analyses (breast cancer, glioblastoma, Huntington disease, diabetes, etc...)
 - Develop statistical methods for microarray analysis
 - Provide bioinformatics resources for researchers
 - Address microarray annotation issues

Overview of TRANSBIG project



- EU Framework Programme VI Project: “**T**ranslating molecular knowledge into early breast cancer management: building on the Breast International Group (**BIG**) network for improved treatment tailoring”
- Started 2004
- Coordination: Jules Bordet Institute, Brussels
- 39 partners in 21 countries (Sept 2004), in UK:
 - Southwest Wales Cancer Institute
 - University of Glasgow
 - Oxford University
- Total funding: €24 millions (€7 millions EU)

Overview of TRANSBIG project



■ Goals:

- “To develop ways of individualizing breast cancer treatment, so that treatment is tailored to the person receiving it.”
- “To integrate, strengthen and facilitate translational and clinical breast cancer research ... by linking to an existing network for clinical breast cancer trials (BIG).”
- “To develop and run a major clinical trial aimed at validating the hypothesis that understanding the genetic make-up (signature) of a tumor can lead to better targeted treatment.”

Overview of TRANSBIG project



- TRANSBIG: family name for a series of subprojects
- 1st TRANSBIG subproject: MINDACT
Microarray for Node Negative Disease may Avoid Chemotherapy
Supervised by EORTC
- Goals:
 - Validation of 70-genes signature identified at Netherlands Cancer Institute (van 't Veer study)
 - Suggest new therapeutic orientations
- Large clinical study: 6'000 patients, 3 years
- Outcome:
 - Avoid unnecessary treatments (damageable and expensive) through better early cancer characterization

Current status of TRANSBIG



- Role of major institutions have been defined (sample collection, data analysis, ...)
- Current phase “*preliminary validation*”
- Recent milestones:
 - 70-gene hypothesis tested using samples different from those in original study...
 - ... and results shown in Dec 2004 at SABCS*
 - Jan 2005: Design of MINDACT study updated and refined
- Next phase:
 - Validation of new design and approval by EU

* San Antonio Breast Cancer Symposium

Issues in large microarray projects

- Center bias in multi-center studies
 - Confounding effects introduced by poor randomization
 - Missing or incomplete data in databases and tissue banks
 - Use of different pathological grading systems, inconsistent units (if unit recorded)
 - ...
-

Illustration: center effect

Color:
Protein marker
(Histochemistry)

- Same platform (Affy)
- Data normalized together
- Gene-specific behavior

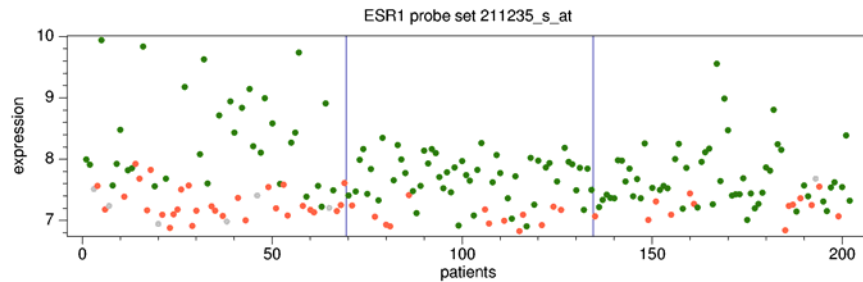
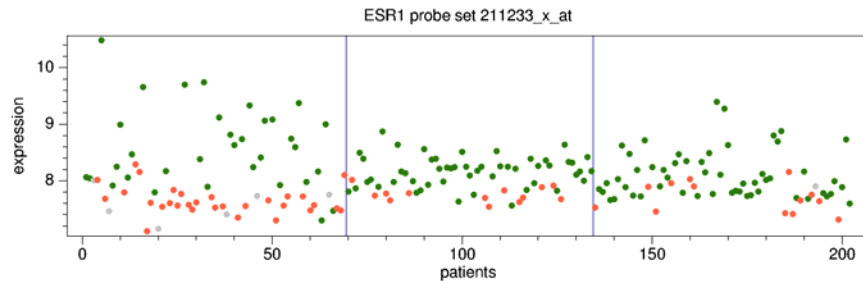
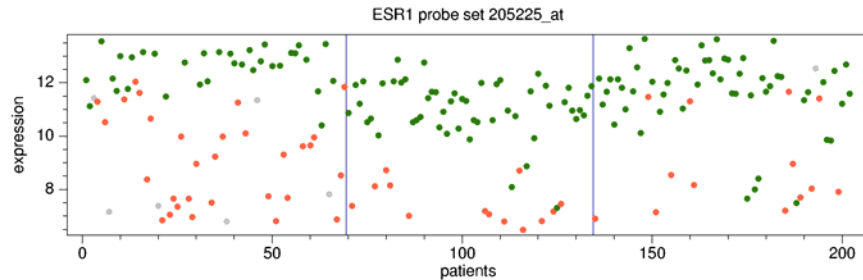
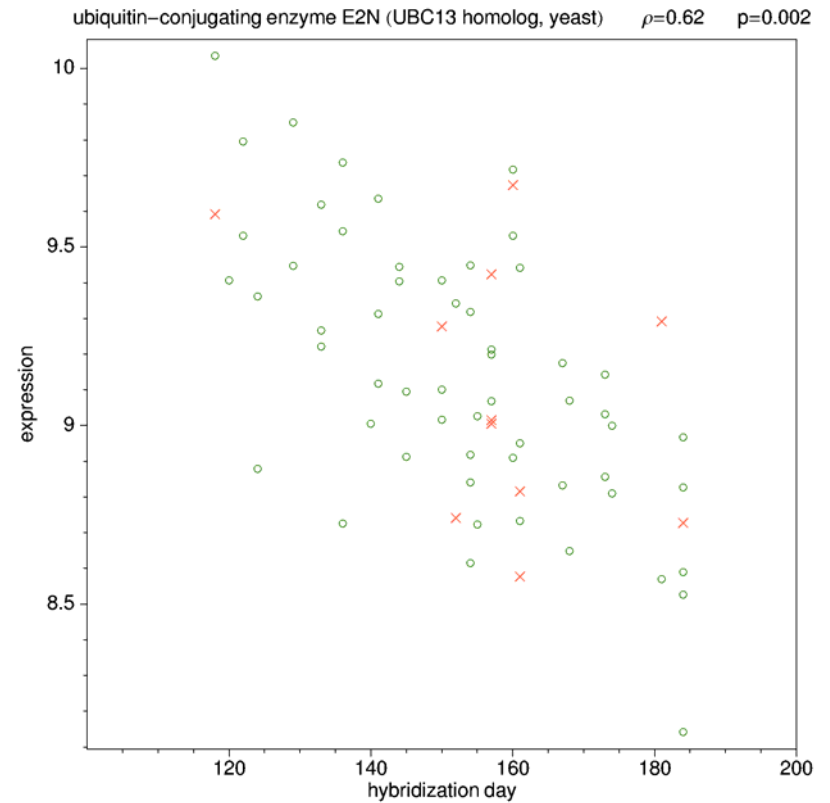
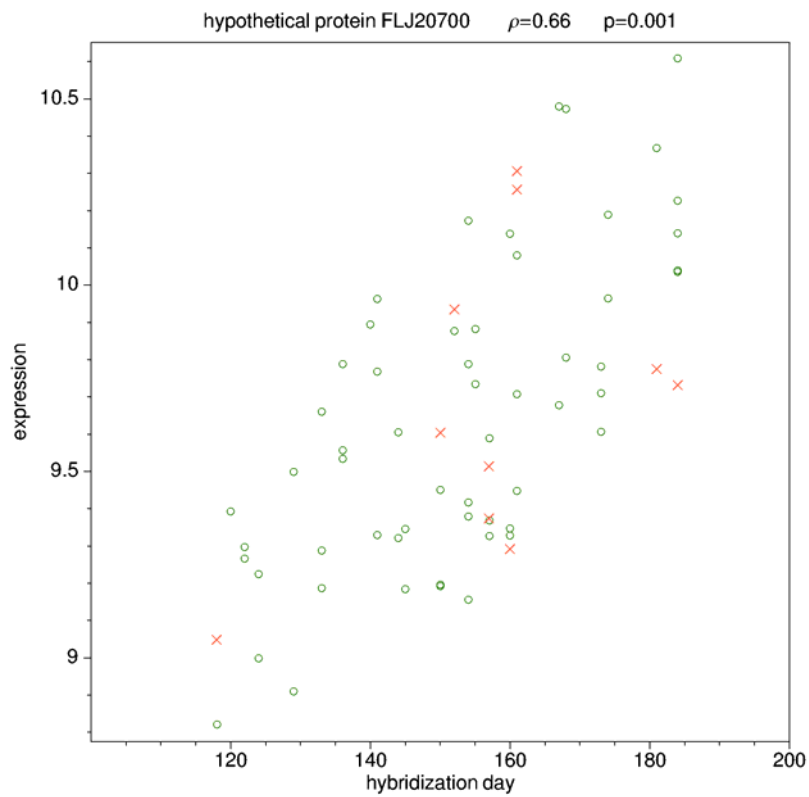


Illustration: date effect



Information required for analysis (I)

- Need for detailed information that allows identification of systematic errors
 - “Low level” sample description (MIAME)
 - “Higher level” (clinical) sample description
 - MIAME can help:
 - Identifying technical confounding factors...
 - Technician effect
 - Protocol effect
 - Date effect
 - ... to mitigate them
-

Information required for analysis (II)

- Detailed record of patient history
 - Sex, age, etc...
 - Date and conditions of illness diagnosis
 - History of previous treatments, relapse
 - ...
- Sample description parameters
 - Pathological grading (~semi-objective)
 - Marker status (ER, Erbb2, ...)
 - Units !!! (log2/log10/ln and other “real” units)
 - ...

Experimental design

- Early involvement of statistician can avoid systematic errors
 - Randomization
 - Which technician will conduct which hybridization
 - When should experiments take place
 - Which pathologist will grade which sample
 - etc...
 - Define minimal number of experiments/conditions to isolate at best parameters under scrutiny
-

In TRANSBIG context



- Organizational issues addressed
 - Main analysts take care of uniformity of data
Marc Buyse (IDDI), Patrick Therasse (EORTC), ...
 - BCF role: microarray analysis quality assurance
(test reproducibility of results by independent analysis)
- Standardization of clinical information
 - *ad hoc*, overviewed by main analysts
 - probably not transposable “as is” to other studies

Global picture

- Based on our experience:
 - **Study specific standardization:**
Organizational recommendations may be more important than defining detailed standards.
 - **Towards an universal database:**
Manual curation of clinical information might be unavoidable if uniformity required.
-