Large scale TMA experiments: automation and data management

Andrea Sboner [(1)], Mattia Barbareschi [(2)], Rossana Dell'Anna [(1)] , and Francesca Demichelis [(1)]

[(1)] Bioinformatics, SRA Division, ITC-irst, Via Sommarive 18, 38050 Povo, Trento-Italy
{sboner, dellanna, michelis}@itc.it
[(2)] Istituto di Anatomia Patologica, Ospedale Santa Chiara, L.go Medaglie d'Oro 8, 38100 Trento-Italy
barbareschi@tn.aziendasanitaria.trentino.it

Keywords. Tissue Microarray (TMA), Gene expression, Automation.

Introduction

Characterization of gene-expression profiles with DNA microarrays provides a powerful mean to discover disease-related genes, particularly in cancer. It is well known that clinical validation of disease-relates genes, through standard molecular analysis on individual tissue sections needs enormous effort in terms of time and costs. To overcome this problem, the Tissue Microarray (TMA) methodology has been recently developed [1]: a high-throughput technology enabling "genome-scale" molecular pathology studies.
In this paper we briefly present our technological platform [2] designed and optimized for the complete management of Tissue Microarrays experiments. Our comprehensive system is very flexible regarding the management of data and it allows a wide range of microarray experiments on different diseases. We also obtained promising new results of biomarker expressions on ovarian and breast cancer, in terms of discrimination of patients' overall survival and relapse free survival.

Motivation

The integration of cDNA microarray, high-density Tissue Microarray, and linked clinical and pathology data is emerging as powerful approach to molecular profiling of human cancer. Beside pathological and clinical aspects TMA technique poses significant challenges with data collection, data sharing and data interpretation, which bioinformaticians are only now beginning to address [3, 4]. Moreover, to allow high-throughput screening of TMAs in large-scale applications, a comprehensive system for TMA data management should include procedures to automatically acquire digital images of single core sections, therefore enabling digital pathology, for eventually quantifying biomarkers expression in automated way. Finally a system that exalts the full potentiality of TMA must be very flexible and must allow "open" experiments. We decided to address these issues by developing a web system with relational database (RDBMS) for data management together with TMA *ad hoc* virtual cases [5]. The choice of developing a web-based system is driven by different motivations. First and foremost, a web-based system allows collecting data from several institutions, providing access form virtually everywhere with an Internet connection. Secondly, the RDBMS ensures data consistency, correctness and completeness. Finally, TMA *ad hoc* virtual cases allow the evaluation on monitor of gene and/or protein expression by pathologists of different institutions and even more important allows quantifying biomarkers expression in an inherently objective and replicable way.

System Overview

The whole system (Fig 1.) has been designed following the standard web-based 3-tiers architecture, and consists of 1. a web based application to collect all the data involved in TMA studies: block array information, patients

data, array slide staining, core section evaluations, etc., 2. a relational database and 3. an environment (TMAenv) to automatically perform digital image acquisition of TMA glass slides and biomarker expression measurements. Particular attention has been given to the user interface design, addressing different needs of different users who operate in distinct work phases (biologist, pathologist, technicians, clinicians and bioinformaticians). Workflow needs, differentiated for input and retrieval phases, have been considered. The relational database univocally links records between different tables and, together with the middle layer, reduces possible errors in linking information: biological/genetic, pathological, and clinical. TMAenv is based on the usage of a robotic microscope (to automatically acquire digital image of single core section) and image processing algorithms to properly locate each detected core section in its block array and thus linking it to corresponding bio-clinical data. Digital images are also stored in the relational database.

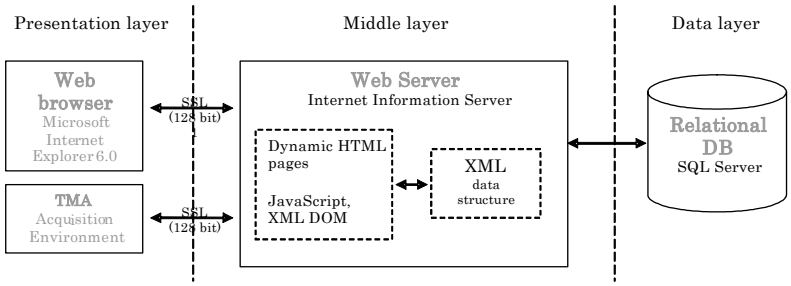

Fig 1 – System components and architecture

Results and Conclusions

The described system has been already implemented and deployed and is currently in use. Up to now (February 2004) about 1900 breast, ovarian and lung tumors have been collected together with the information of about 140 block arrays. About 700 glass slides have been cut, stained and stored. Moreover, biomarker evaluations of about 480 glass slides (about 34,500 cores sections) were input by pathologists from different institution. Automatic expression measurements have also been performed and tested for two biomarkers on single core section digital images (about 2,300 core sections). This kind of large-scale and distributed experiments is enormously promoted by this comprehensive system in terms of time costs. Moreover, system controls on data consistency reduce error-propagation risks, which affect high-throughput experiments.
Preliminary results of TMA data analysis confirmed known biomedical knowledge. We also obtained promising new results of biomarker expressions on ovarian and breast cancer, in terms of discrimination of patients' overall survival and relapse free survival. In conclusion, this system fosters the *translational research,* i.e. the clinical application of biopathological research findings, which is the aim it has been created for.

Acknowledgements

This work is partially funded by Ministry of Health of Italian Government.

References

[1] J. Kononen,. L. Bubendorf, A. Kallioniemi, M. Barlund, P. Schraml, S. Leighton, J. Torhorst, M.J. Mihatsch, G. Sauter, O.P. Kallioniemi, Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med,* 4(7):844-7C, 1998.
[2] https://bioinfo.itc.it/TMA/
[3] S. Manley, N.R. Mucci, A.M. De Marzo, M.A. Rubin: Relational database structure to manage high-density tissue microarray data and images for pathology studies focusing on clinical outcome: the prostate specialized program of research excellence model. *Am J. Pathol*. 2001; 62159(3):837-843.
[4] R. Simon, M. Mirlacher, G. Sauter. Tissue Microarray *Biotechniques.* 2004 Jan;36(1):98-105.
[5] F. Demichelis, M. Barbareschi, P. Dalla Palma, S. Forti. The virtual case: a new method to completely digitize cytological and histological slides. *Virchows Arch*, 441(2):159-64, 2002.