Advances in cancer tissue microarray technology: Towards improved understanding and diagnostics

Wenjin Chena,∗, David J. Forana,b

a The Center for Biomedical Imaging & Informatics, University of Medicine and Dentistry of New Jersey, 675 Hoes Lane, Piscataway, NJ 08854, United States
b The Cancer Institute of New Jersey, University of Medicine and Dentistry of New Jersey, 195 Little Albany Street, New Brunswick, NJ 08903, United States

Received 1 September 2005; received in revised form 18 November 2005; accepted 21 November 2005

Available online 23 January 2006

Abstract

Over the past few years, tissue microarray (TMA) technology has been established as a standard method for assessing the expression of proteins or genes across large sets of tissue specimens. It is being adopted increasingly among leading research institutions around the world and utilized in cancer research in parallel with the cDNA microarray technology. This article summarizes various aspects of cancer understanding and diagnostics in which TMA has had great impact. Although tremendous advances continue to be made to facilitate imaging and archiving of TMA specimens, automatic evaluation and quantitative analysis of TMA still remains an important challenge for modern investigators.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Tissue microarray; Molecular profiling; Molecular marker quantification; Automatic tissue analysis

1. Introduction of tissue microarray technology

Tissue microarray (TMA) is a powerful new technology designed to assess the expression of proteins or genes across large sets of tissue specimens assembled on a single glass microscope slide, efficiently and economically [1–3].

Kononen’s original TMA creation method extracts tissue cylinders originating from either formalin-fixed paraffin-embedded tumor tissue blocks or freshly frozen tumors fixed in cold ethanol and embedded in paraffin (to reserve RNA or DNA for detection), and precisely arrays them into a “recipient” paraffin block [1], please see Fig. 1. About 200 consecutive sections 4–8 μm in thickness can be cut from each tissue microarray block, with each stained differently to establish a large-scale gene expression profile of cancers. It has been demonstrated that TMA sections can be used to detect DNA (fluorescent in situ hybridization, FISH), RNA (mRNA in situ hybridization, RNA-ISH), or protein (immunohistochemistry, IHC). This new technique should not be confused with DNA microarrays, in which each tiny spot on the grid represents a unique cloned complementary DNA (cDNA) or oligonucleotide.

Two of the most apparent advantages of TMA technology are that it allows amplification of limited tissue resources by providing the means for producing a large number of small core biopsies, rather than generating a single section, and that it saves time and antibody to analyze multiple specimens at once. Using this technology, a carefully planned array can be constructed with cases from pathology tissue block archives, such that a 20-year survival analysis can be performed on a cohort of 600 or more patients using only a few microliters of antibody. Another major advantage of the TMA technique is the fact that each specimen is treated in an identical manner. Consequently, reagent concentrations are consistent across discs within each TMA specimen, as are the incubation times, temperatures, and washing conditions. The staining results of these discs are hence comparable.

TMA technology continues to evolve and improve in order to simplify the process of constructing arrays [4–7], while accommodating a wider variety of specimens. Although the original TMA technique focused on profiling paraffin-embedded tissue specimens, Abbott et al. established a cell line microarray composed of cell pellets, which had been obtained from human tumor [8], and several groups reported success in creating TMAs from frozen needle biopsies or brain tissue [6,9–11]. Fergenbaum et al. reported their concern that the extended storage time for paraffin-embedded tissue might affect tissue’s
immunoreactivity and thereby giving rise to an increase in false-negative results in the corresponding TMAs [12]. Combined paraffin coating and nitrogen storage was reported by DiVito et al. [13] to be the best storage condition for preserving TMA antigenicity.

In response to questions posed by many investigators as to whether utilizing 0.6 mm diameter cores of tissue was sufficient in representing the heterogeneous morphology of any given tissue specimen, the TMA technology underwent tremendous scrutiny and was tested and validated repeatedly for use in cancer research over the period from 1998 to 2003. A number of research groups participated in this process by comparing TMA analysis with large tissue sections diagnosis and/or by validating its results with cDNA microarray findings in various cancer types: breast cancer [12,14–19]; prostate cancer [20–23]; gastric cancer [24]; colorectal cancer [25–27]; lymphomas [28–32]; multiple myeloma [33]; soft tissue sarcoma [34]; renal cell carcinoma [35]; bladder tumor [36]; glioma [37,38]; melanoma [39]; lung tumor [40,41]. It is now generally accepted that two to four samples taken from different regions of each donor tissue block provides a sufficient amount of morphologic information for reliably evaluating specimens [18,22,24,31,34,42]. In the case that the protein marker is expressed focally as opposed to homogeneously, the use of TMA is validated by Mucci et al., by increasing number of samples from each specimen [23]. TMA technology has also proven to be an excellent quality assurance tool that can be used to conduct inter-observer, intra-laboratory, and inter-laboratory studies [43–47].

2. TMA in cancer understanding and diagnosis

Although TMA is not currently used to render a primary diagnosis, it does provide unparalleled insight into cancer development, prognosis, and treatment.
Molecular profiling has been demonstrated to be a useful approach for providing new insight into cancer development and differentiation although the underlying mechanisms is not yet fully understood. Multivariate statistical analysis, including principal component analysis, hierarchical clustering analysis and other machine learning algorithms, were used in several of these efforts to elucidate the possible relationships between the expression profiles and biological or clinical parameters of the cases [81–88]. Using 13 protein markers, Callagy et al. revealed distinct tumor clusters which divided into two main groups correlating with tumor grade and nodal status, whereas none of the protein biomarkers that were tested could identify these groups individually [82]. Hierarchical clustering analysis applied to immunostaining results for 18 different markers showed that synovial sarcomas, leiomysarcomas, hemangiopericytomas, and solitary fibrous tumors clustered distinctly [84]. Using principle component analysis to characterize salivary gland tumors resulted in three components using a panel of 15 markers correlating with tumor grade and nodal status whereas none of the protein biomarkers that were tested could identify these groups individually [82].

It has also recently been discovered that, in addition to expression intensity, subcellular location of certain protein markers [90–94] is related to clinical outcome. Berger et al. conducted a survival study that included 555 tumors collected from the Yale University Department of Pathology archives [95]. The results of the study indicated that subcellular localization of activating transcription factor 2 (ATF2) could be used as an independent prognostic indicator for melanoma. Subcellular localization of 14-3-3sigma and Ski-related novel protein N (SnoN) are also identified as prognostic factor for breast cancer [92,93].

### 3. TMA imaging, archiving, and information management

Currently, the primary methods used to evaluate the arrays involve manual, interactive review of TMA samples while they are subjectively evaluated and scored. An alternative, but less utilized approach for evaluation is to digitize each specimen, sequentially, for subsequent semi-quantitative assessment [96]. Both procedures ultimately involve the interactive localization of TMA discs under the microscope, which is a slow, tedious process that is prone to error. For many researchers and technologists, simply navigating among the regularly arranged tissue cores while viewing them under a microscope makes it difficult to keep track the current disc position within the array. This is especially problematic at high magnifications. Research groups and manufacturers have implemented several automatic approaches to image TMA discs for storage and evaluation.

Beyond the algorithmic and software development that is required for analyzing tissue microarrays, reliable tools are also needed to facilitate large-scale, multi-site collaboration for a broad spectrum of research and clinical activities including tissue banking, proteomics, and outcome studies. Future progress in several key areas will rely upon the capacity of individuals to acquire, share, and assess microarrays and correlated data, dynamically.

Although some DNA microarray readers are capable of reading TMA samples, automatic imaging, and evaluation of TMA present a host of unique technical challenges. First, TMA samples often exhibit a higher level of morphological irregularities. For example, aside from the overall rotation of the grid on the slide, tissue discs are sometimes shifted from regular grid positions as a result of mechanical deformations, which can arise during construction of the physical TMA. Such artifacts are often introduced during the sectioning stage of specimen preparation. In addition, discs occasionally become detached and fall out during preparation. To address these issues our group developed a web-based prototype which features automated imaging, registration, and distributed archiving of tissue microarrays in multi-user network environments, please see Fig. 1. The imaging system features a robust alignment algorithm that could reliably recover sample grids and compensate for detached discs [97].

Several other groups of investigators have also developed their own relational databases to manage TMA related information [97–102]. To facilitate exchange of TMA data by researchers across all these different database implementations, the Technical Standards Committee of the Association for Pathology Informatics organized a TMA workshops in order to establish a TMA data exchange specification. The proposed XML-based TMA data exchange specification [103,104] is free and non-proprietary.

While it is clear that much more research and development is needed to achieve reliable analysis of TMA, several commercial products that have already emerged for imaging and visualization of TMA including BLISS system by Bacus Laboratory (Lombard, IL); MedScan by Trestle (Irvine, CA); ScanScope by Aperio Technologies (Vista, CA); Arixl by Applied Imaging (San Jose, CA).
4. Quantitative analysis of TMA

Gene expression studies using cDNA microarrays have created an unprecedented need for the development and introduction of new analytical tools. Similarly, there now exists an urgent need for tools to analyze TMA specimens efficiently and effectively. However, despite the many commonalities between cDNA microarrays and TMA's, there is one important distinction which makes the analysis of tissue microarrays especially challenging. Unlike cDNA microarrays which can generally be considered homogeneous across a given spot, the discs comprising an array typically consist of a set of heterogeneous, stained tissues which essentially renders simple image analysis strategies completely inadequate in performing any meaningful assessment. The analysis is further complicated by the fact that, depending on the type of tumor or tissue section analyzed, the area of interest may represent nearly the entire disc or only a small percentage thereof (please see examples in Figs. 1 and 2). For example, a pancreatic carcinoma or lobular carcinoma of the breast with substantial dysplastic response may show stromal tissue representing a large percentage of the total area of the disc. If the goal of the assay is to determine epithelial cell expression of a given marker, a protocol must be used that evaluates only that component of the disc. The protocol must not only be able to identify the region of interest but it must also perform normalization operations so that the expression level for any given disc can be reliably compared with the level reported for others.

For these reasons, although a few groups made attempts to read TMA specimens using commercial cDNA microarray readers [105,106], developing reliable and effective methods and protocols specifically for quantitative analysis of tissue microarrays has become an extremely active research area since 2001 [94,107,108]. Since fluorescent TMA's and classic immunohistochemistry (IHC) TMA's each require specific hardware, most of the methods proposed to evaluate these specimens are only suitable for one type of processing.

4.1. Quantitative analysis of immunohistochemistry specimens

IHC is widely used for a broad number of applications throughout both the clinical and biomedical research communities. Research efforts focusing on automated quantification of IHC began with traditional biopsy specimens, but are now also being investigated for potential use in analyzing tissue microarrays. Similar to the TMA quantification, one of the most important tasks in quantification of IHC specimens is to generate tissue-specific or tumor-specific measurements. Another challenge of IHC quantification is how best to normalize the color differences that exist among specimens that have been acquired, stored, and processed under different conditions. Since TMA tissue cores are processed and stained on one single slide, tissue discs within a given slide can generally be considered as having already being normalized, making IHC TMA an ideal platform for developing and validating IHC quantification methods.

Several successful studies based on computerized IHC analysis have been reported including applications involving cell counting [109,110], microvessel counting [111], and cell morphology analysis [109]. There now also exist multiple...
commercial products which are aimed at quantitative IHC analysis such as the ChromaVision Automated Cellular Imaging System (ChromaVision Medical Systems Inc., San Juan Capistrano, CA) [112], the Image Tiler (Tripath Imaging, Burlington, NC) [113], and QPRODIT (Leica Imaging Systems Ltd., Cambridge, UK) [109]. TMAcore (Bacus Instruments, Lombard, IL), and TMAx (Breecher Instruments, Sun Prairie, WI). Most of these approaches rely upon manual delineation of the region of interest (ROI) before analysis can begin [109,114].

4.2. Quantification of fluorescent TMA

Rao et al. developed a TMA-based quantitative fluorescence image analysis (QFIA) method and validated its use by analyzing distinctive BRCA1 expression patterns in high-grade and low-grade sporadic epithelial ovarian cancers and their adjacent dysplastic fields [106]. However, his system could not automatically distinguish epithelial tissue from surroundings, instead it depended on manual segmentation to locate regions of interest. Haedicke et al. [105] used a cDNA microarray two-color laser scanner to evaluate fluorescent TMA specimens. In his unique approach using double-label indirect immunofluorescence, the target protein expression signal was normalized by a second fluorescent label, cytokeratin, on the same TMA specimen. Since the cDNA microarray scanner only measures total signal strength of a specific antigen, the staining intensity can be normalized according to the amount of tissue which underlies the underlying staining characteristics of even those cells that suffer from visual masking due to the counterstaining. Particularly, this counterstaining can be visualized using DAB and hematoxylin distributes along a hyperplane in the $L^*a^*b^*$ color space. By performing a polar transformation of the $L^*a^*b^*$ color space plot and identifying the two principal peak colors in this $L^*a^*b^*$ representation, the principal color vectors are determined. Therefore, the staining characteristics of each tissue disc appear as a continuous representation of staining intensity for each of the dyes. It is interesting to note that the protocol that we have reported is able to unveil and quantify the underlying staining characteristics of even those cells that suffer from visual masking due to the counterstaining, please see Fig. 2. We also explored the use of the color decomposition algorithms for a broader set of immunohistochemical staining agents. Recent feasibility studies showed that this approach was also effective in analyzing histological specimens stained with Nova Red and Vector Red against hematoxylin.

5. Conclusion

The power and efficiency of TMA technology rests in its capacity to provide investigators with a means for evaluating gene expression in tissue or cell lines while processing large cohorts of specimens simultaneously. Due to the heterogeneous and complex nature of histological samples, the staining characteristics of TMA specimens can be described and quantified at multiple levels of detail. The most common approaches for assessment have focused primarily on determining the overall staining intensity of the individual tissue discs comprising the arrays. At the tissue level, however, this overall staining intensity can be normalized according to the amount of tissue which is present each disc. The next two levels of detail are at the cellular and subcellular resolution where the assessment can provide even greater insight related to the marker distribution in vivo. A tremendous amount of research efforts are targeting these levels of detail in order to promote better understanding of the underlying mechanisms of disease progression. This level of detailed analyses accentuates the need for higher standards for TMA imaging resolution and data management solutions. It is anticipated that the advances in TMA image analysis will also serve to impact the manner in which routine diagnostic IHC specimens are imaged, archived, and analyzed. Although it is not anticipated that TMA’s will have direct clinical application in the near future, incremental steps are necessary to support TMA analysis, the majority cannot distinguish between tumor cells and non-malignant tissue components such as stroma or inflammation, in their analysis.

Elie et al. attempted automatic tissue specific nuclei quantification [115]. Their method for delineation of epithelial from stromal compartments was based on intensity difference in the green color channel, which is not necessarily applicable to other IHC specimens. Lacking the technology to un-mix multiple stains on a given IHC specimen, they generated their measurements by manipulating RGB color channels while approximately isolating DAB stained colors. We have developed the principal color decomposition method to detect and characterize the staining signatures for each dye within the specimen based upon vector decomposition in $L^*a^*b^*$ color space [97]. The method explored the fact that color vectors that are generated for a representative TMA specimen stained with DAB and hematoxylin distributes along a hyperplane in the $L^*a^*b^*$ color space. By performing a polar transformation of the $L^*a^*b^*$ color space plot and identifying the two principal peak colors in this $L^*a^*b^*$ representation, the principal color vectors are determined. Therefore, the staining characteristics of each tissue disc appear as a continuous representation of staining intensity for each of the dyes. It is interesting to note that the protocol that we have reported is able to unveil and quantify the underlying staining characteristics of even those cells that suffer from visual masking due to the counterstaining, please see Fig. 2. We also explored the use of the color decomposition algorithms for a broader set of immunohistochemical staining agents. Recent feasibility studies showed that this approach was also effective in analyzing histological specimens stained with Nova Red and Vector Red against hematoxylin.
being taken by the clinical and research communities to design, develop, and evaluate the computational tools which are necessary for performing reliable, reproducible comparative analysis of expression patterns in arrayed tissues. As these tools become increasingly reliable they will have tremendous potential in the areas of drug discovery, clinical outcomes, and therapy planning.

Acknowledgements

The research of the authors mentioned in this review is supported, in part, by a grant from The Cancer Institute of New Jersey and NIH contract 5R01LM007455-03 from the National Library of Medicine and 1R01EB003587-01A2 from the National Institute on Biomedical Imaging and Bioengineering.

References


