Tissue Microarrays for Rapid Linking of Molecular Changes to Clinical Endpoints

Joachim Torhorst,* Christoph Bucher,* Juha Kononen,† Philippe Haas,* Markus Zuber,‡ Ossi R. Köchli,§ Frank Mross,¶ Holger Dieterich,‖ Holger Moch,* Michael Mihatsch,* Olli-P. Kallioniemi,§ and Guido Sauter*†

From the Institute of Pathology,* the Department of Surgery,† Universitäts-Frauenklinik,‡ University of Basel, Basel, Switzerland; the Cancer Genetics Branch,§ National Human Genome Research Institute, National Institutes of Health,Bethesda, Maryland; Kreiskrankenhaus,¶ Lörrach, Germany; and Frauenklinik,‖ Rheinfelden, Germany

Advances in genomics and proteomics are dramatically increasing the need to evaluate large numbers of molecular targets for their diagnostic, predictive or prognostic value in clinical oncology. Conventional molecular pathology techniques are often tedious, time-consuming, and require a lot of tissue, thereby limiting both the number of targets that can be evaluated. Here, we demonstrate the power of our recently described tissue microarray (TMA) technology in analyzing prognostic markers in a series of 553 breast carcinomas. Four independent TMAs were constructed by acquiring 0.6 mm biopsies from one central and from three peripheral regions of each of the formalin-fixed paraffin embedded tumors. Immunostaining of TMA sections and conventional “large” sections were performed for two well-established prognostic markers, estrogen receptor (ER) and progesterone receptor (PR), as well as for p53, another frequently examined protein for which the data on prognostic utility in breast cancer are less unequivocal. Compared with conventional large section analysis, a single sample from each tumor identified about 95% of the information for ER, 75 to 81% for PR, and 70 to 74% for p53. However, all 12 TMA analyses (three antibodies on four different arrays) yielded as significant or more significant associations with tumor-specific survival than large section analyses (p < 0.0015 for each of the 12 comparisons). A single sample from each tumor was sufficient to identify associations between molecular alterations and clinical outcome. It is concluded that, contrary to expectations, tissue heterogeneity did not negatively influence the predictive power of the TMA results. TMA technology will be of substantial value in rapidly translating genomic and proteomics information to clinical applications. (Am J Pathol 2001, 159:2249–2256)

Analysis of prognostic and predictive markers in cancer has traditionally been accomplished by testing one marker at a time, starting from a relatively small sample size. However, before routine clinical application, large-scale studies of thousands of well-characterized tissue specimens with clinical follow-up information will need to be carried out to demonstrate the independent significance of the biomarker. During the past decades, only three well-characterized biomarkers have been implemented in the clinical routine in breast cancer: estrogen (ER) and progesterone receptors (PR), as well as the HER-2 oncogene.1–8

The translation of basic research findings to clinical applications is now becoming dramatically more challenging, with the introduction of high-throughput genomics and proteomics technologies.9 For example, in a single cDNA microarray experiment, one is able to determine the expression status of 50,000 human genes. These technologies often require fresh tissues, which makes it difficult to directly apply them in clinical studies. Formalin-fixed archival tissues provide a means to validate such genomic and proteomic screening in large sets of histologically well-characterized tumors with clinical endpoints. However, testing of even a small fraction of the human gene and protein targets is beyond the scope of traditional molecular pathology technologies. Not only are these techniques slow and tedious, but the availability of tissue is often rate-limiting. For example, one can only cut at most 300 sections from a typical archival tissue block. In the case of smaller tumors, or previously used precious research materials, the number of sections is often much smaller.

Our recently developed tissue microarray (TMA) technology has the potential to significantly accelerate studies seeking for associations between molecular changes and clinical endpoints.10 In this technology, 0.6 mm tissue cylinders are punched from hundreds of different primary tumor blocks and subsequently brought into a recipient tissue microarray block. Sections from such array blocks can then be used for simultaneous in situ analysis of hundreds or thousands of primary tumors on DNA, RNA, and protein level. The high speed of arraying, the lack of a significant damage to donor blocks, and the
regular arrangement of arrayed specimens greatly facilitating automated analysis are the most significant advantages of the TMA technology over previous concepts of analyzing multiple different tissues in one paraffin block.11 To test the utility of our “tissue chip” approach for finding associations between molecular changes and clinical endpoints, we used breast cancer as a model system. The TMA analysis of the well-established prognostic markers ER and PR as well as of p53, another suggested prognostic parameter, suggests that tissue chips provide a means for rapid screening of the prognostic significance of molecular markers and may help to translate genomic and proteomics information to clinical applications.

**Materials and Methods**

**Patients**

Samples from 611 breast carcinomas had previously been included in a breast cancer TMA. The carcinomas of 553 patients of which follow-up data (tumor-specific survival and treatment information) could retrospectively be evaluated were included in this study. These patients had a median age of 61 (range, 33 to 97) years. They were treated for primary breast cancer at the University Hospital in Basel (Switzerland), Women’s Hospital Rheinfelden (Germany), and the Kreiskrankenhaus Lörach (Germany) between 1985 and 1994. The mean follow-up time was 65.8 months (range, 1 to 151). A systemic therapy had been performed in 273 patients including 172 with hormonal therapy alone, 52 with cytotoxic therapy alone, and 49 having both hormonal and cytotoxic treatment. Formalin-fixed, paraffin-embedded tumor material was available from the Institute of Pathology, University of Basel. The pathological stage, tumor diameter, and nodal status were obtained from the primary pathology reports. All slides from all tumors were reviewed by one pathologist (J.T.) to define the histological grade according to Elston and Ellis12 and the histological tumor type. The series included 405 ductal, 77 lobular, 16 medullary, 14 mucinous, 11 cribriform, 11 tubular, 7 papillary, 4 apocrine, 3 clear cell, 1 metaplastic, 1 atypical medullary, 1 large cell, 1 small cell, and 1 neuroendocrine cancer. Among 553 tumors, 27.8% were grade 1, 42.9% were grade 2, and 29.3% were grade 3. The local tumor stage was pT1 in 39.5%, pT2 in 46.3%, pT3 in 4.9%, and pT4 in 9.3%. The stage could not be unequivocally determined from the pathology reports in 6 tumors. Axillary lymph nodes had been examined in 519 patients (the nodal stage (pN) was: 52.4% pN0, 39.3% pN1, and 8.3% pN2). Stage, grade, and nodal status were strongly associated with tumor-specific survival of our patients (p < 0.0001 each).

**Tissue Microarray Construction**

Tumor samples were arrayed as previously described.10 Briefly, H & E-stained sections were made from each block to define representative tumor regions. Tissue cylinders with a diameter of 0.6 mm were then punched from selected areas of each “donor” block using a custom-made precision instrument (Beecher Instruments, Silver Spring, MD) and brought into a recipient paraffin block. The TMA blocks were constructed in four copies each containing one sample from a different region of all tumors. One sample was taken from the center (Figure 1) and three samples were taken from different peripheral areas of the tumors. After the TMA construction, large sections were cut from the “donor” blocks of 532 tumors having sufficient material available.

**Immunohistochemistry**

Three conventional “large” sections from all tumors and three sections from each of the four different replica TMA blocks were used for immunostaining. Standard indirect immunoperoxidase procedures (ABC-Elite, Vector Laboratories, Burlingame, CA) in combination with monoclonal antibodies were used for detection of p53 (DO-7, prediluted DAKO, Glostrup, Denmark), estrogen receptor (ER ID5, 1:1000, DAKO), and progesterone receptor (NCL-PGR, 1A6, 1:600, NOVOCASTRA Laboratories Ltd, Newcastle-upon-Tyne, UK). A microwave pretreatment was performed for p53 (30 minutes at 90°C) retrieval. Diaminobenzidine was used as a chromogen. Tumors with known positivity were used as positive controls. The primary antibody was omitted for negative controls. The same scoring criteria were applied in TMA and in large sections. All slides were manually read by one pathologist (J.T.). Tumors were considered positive for ER and PR if an unequivocal nuclear positivity was seen in at least 10% of tumor cells. To define a tumor as p53 positive, moderate staining intensity was requested in ≥ 20% of tumor cell nuclei. These cutoff values were arbitrarily selected before the beginning of the study based on previous suggestions.13,14 An only faint p53 staining was scored negative because such a staining can often be seen in non-neoplastic cells, for example, in the basal cell layer of squamous epithelium or urothelium. Examples of positive and negative tumors are given in Figure 2.

**Statistics**

Contingency table analysis and χ² tests were used to study the relationship between immunohistochemical re-
results on large section and on TMAs. Survival curves were plotted according to Kaplan-Meier. A log-rank test was applied to examine the relationship between ER, PR, or p53 positivity and tumor-specific survival. Patients were censored at the latest date when they were seen alive or at the date of their non-tumor-related death.

**Results**

The frequency of ER positivity ranged from 78.9% to 80.8% in sections from the four replicate TMAs. Almost the same frequency was observed on large tissue sections (79.8%) where ER staining was usually homogeneous. Therefore, combining the results of multiple TMAs did not significantly increase the number of ER-positive cases (Figure 3). Loss of ER expression was strongly associated with poor prognosis, regardless of whether the statistical analysis was done on data obtained from individual TMAs, combinations of multiple TMAs, or on conventional large sections from each tumor (Figure 4).

The frequency of PR positivity ranged from 41.1% to 53.1% in the four TMAs. A slightly higher frequency was found on large tissue sections (60.3%). The combination of the results of multiple arrays led to a gradual increase of the fraction of positive tumors (Figure 3), so that 4 different TMAs provided essentially the same frequency as observed on large sections. Loss of PR expression was also strongly associated with poor prognosis both when evaluated from individual TMAs, combinations of multiple TMAs, and on large sections from each tumor (Figure 5).

**Figure 2.** Examples of positive and negative tumors. Single punches (0.6 mm diameter) positive for ER (A), PR (B), and positive and negative for p53 respectively (C and D).

**Figure 3.** Immunohistochemical analyses of ER, PR, and p53 on TMAs and on large sections. Only tumors that were interpretable on all four TMAs and on large sections were included in this analysis. The bars on the left of each group reflect the positivity found in the individual TMA from the tumor center (C) and the three peripheral areas (P1, P2, P3). The bars marked as 2A, 3A, and 4A give the frequency of positivity that was obtained by combining the data from 2 TMAs (2A: center + periphery 1), 3 TMAs (3A: center + periphery 1 + periphery 2), or from all four TMAs (4A). Tumors are considered positive in this calculation if at least one sample was considered positive. The bar on the right gives the frequency detected on large sections (LS).
The prevalence of p53 positivity among the four TMAs ranged from 15.2% to 20.9% (Figure 3). If the results from multiple arrays were combined, there was a slight increase in the overall positivity rate, up to 24.1%. However, the frequency of p53 positivity on large sections (42.8%) was almost twice as high as on TMAs. Unexpectedly, the results obtained on individual TMAs or by combining results from multiple TMAs were much more strongly linked to poor prognosis than the results obtained on large sections (Figure 6). Further analysis showed that tumors that were positive both on TMAs and large sections had a poor prognosis. However, the prognosis of these 111 tumors that were only positive on large sections but not on arrays was equally good as in tumors that were p53 negative on large sections (data not shown). A review of the discrepant cases showed that differences between TMA and large section data were not due to tissue heterogeneity in most cases. The vast majority of discrepant cases had a faint-to-moderate p53 staining in 15 to 30% of tumor cells in large sections. These tumors were often considered negative on array sections because the criteria of positivity (at least 20% of cells with unequivocal positivity) were usually not reached in the small arrayed samples. The same tissues were often considered positive on large sections, where the criteria for positivity were frequently met at least in small tumor areas. Reanalysis of the prognostic significance of the large section p53 staining revealed that a large percentage of positive cells (Figure 6G) and a strong staining intensity (Figure 6H) were strongly linked to poor prognosis. A moderate staining intensity or fraction of positive cells between 20% and 50% was not indicative of poor prognosis, although such tumors were considered p53 positive in our previously selected definition.

Between 70.3% and 76.5% of the arrayed samples were interpretable for ER, PR, and p53 immunostaining on each individual TMA section. The combination of information from multiple TMAs resulted in a significant increase of interpretable tumors. The fraction of interpretable tumors was 86.4 to 93.1% if two, 92.2 to 95.5% if three, and 94.9 to 96.9% if information from all four replicate TMAs were combined. For most tumors, identical results were obtained from all four arrayed samples (Figure 7A). The fraction of tumors with heterogeneous findings was 8.8% for ER, 28.9% for PR, and 11.3% for p53. These heterogeneous tumors tended to have an intermediate clinical outcome as compared with a homogeneously positive or negative finding (Figure 7, B–D).

**Discussion**

Prognostic significance of ER, PR, and p53 was assessed in quadruplicate experiments for over 500 pa-
patients using the TMA technology. The results demonstrate that minute tissue samples in an array format can be sufficiently representative of their donor tumors to establish associations between molecular alterations and clinical endpoints. The possibility to miniaturize tissue analyses will substantially facilitate translational and clinical cancer research in a number of ways. In a single experiment, up to 500 to 1000 tissues can be evaluated on the same microscope slide. This increases the speed of analysis of very large clinical datasets, and will also facilitate the standardization and interpretation of the results. The TMA format also greatly increases the number of targets that can be analyzed from the same set of tumors. Calculations indicate that tens of thousands of TMA sections can be generated from one paraffin block containing 10 × 10 mm of tumor area with a depth of 3 mm. This is hundreds of times more than could be accomplished using traditional techniques based on sectioning entire tumor blocks, where less than 200 sections can be generated before the blocks are exhausted. Investigators using genomics and proteomics technologies are now identifying literally thousands of candidate molecular markers that appear to be involved in cancer and could represent candidate diagnostic and therapeutic possibilities. The tissue microarray strategy will almost be the only means of comprehensively evaluating and prioritizing such large numbers of markers based on their clinical significance in large clinical materials composed of hundreds or thousands of patient specimens.

This study was specifically designed to address the most obvious limitation of the TMA technique: the sampling of large, potentially heterogeneous tumors. The substantial heterogeneity of tumors is often evident both at the morphological and genetic level. This heterogeneity is thought to represent the genetic instability of tumors and forms the basis for the current concepts of progression and clonal evolution of cancer. We observed that the frequency of ER positivity in the entire material was virtually the same when measured from a single TMA section (each tumor represented by a 0.6 mm diameter tissue spot), as compared to entire sections of breast cancers. For PR, the concordance was slightly lower (88%) and three samples from each tumor were required to achieve the same level of positivity as large section analyses. However, it must be kept in mind that large sections often represent a small fraction (ie, 0.004 × 10 × 10 mm) of large tumors (ie, 30 × 30 × 30 mm). The question to what extent TMA data can reproduce large section data are therefore much less important than whether clinicopathological associations can be reproduced or newly detected on TMAs. In this study, all TMA analyses provided highly significant association

Figure 5. PR immunostaining and prognosis. The association with tumor-specific survival is shown for PR data obtained on the TMA containing tissue from the tumor center (A), the three TMAs containing different samples from the tumor periphery (B–D), the combination of the data from all four TMAs counting every tumor as positive if at least one sample was scored positive (E) and on large sections (F).
with prognosis for ER, PR, and p53. Possible heterogeneity or fixation differences between central or peripheral tumor regions did not affect these results even though the fraction of positivity was generally lower in samples from tumor center and periphery 1 than in periphery 2 and 3. It appears that the high number of tumors that can be included in a TMA study compensates for some false negative results which may be equally frequent in all subgroups of one arrayed tumor set. Therefore, in the case of large study materials, a single sample from each tumor may often be sufficient to derive information on clinical associations. This is also supported by our previous observations, where TMA analyses made it possible to reproduce numerous clinicopathological associations that were previously reported in the literature using conventional techniques based on large tissue specimens.10,23,24

In this study, 83 tumors with a borderline p53 staining (15% to 30% positive cells) on large sections were called negative on TMAs but considered positive in the large section analysis. The significant associations with clinical outcome detected on all four replica TMAs but not in the large section analysis prompted a reanalysis of large sections revealing that the prognosis of breast cancer

Figure 6. p53 immunostaining and prognosis. The association with tumor-specific survival is shown for p53 data obtained on the TMA containing tissue from the tumor center (A), the three TMAs containing different samples from the tumor periphery (B–D), the combination of the data from all four TMAs counting every tumor as positive if at least one sample was scored positive (E), and on large sections (F). G and H: Results of a quantitative analysis of the large sections. The percentage of positive cells (G) and the staining intensity (H) were linked to prognosis.

Figure 7. Heterogeneity of results in individual tumors. Only tumors that were interpretable in all four TMAs were included in these analyses. A shows the percentage of individual tumors with 0 to 4 positive TMA results. The survival rates for tumors with homogenous and heterogeneous IHC results are shown for ER (B), PR (C), and p53 (D).
patients was dependent on the fraction of p53-positive cells and their staining intensity. The obvious difficulties in the quantitation of p53 staining may be a reason for the observed discrepancies in the literature. While the majority of previous studies had reported an association between nuclear p53 accumulation and poor prognosis in breast cancer, several other studies had not confirmed this observation. Our p53 data also showed that data obtained on TMAs can be superior to large section data. Several technical issues apparently compensate for some loss of information due to the small tissue size. The staining of a single TMA slide provides a much greater degree of consistency and standardization than the immunostaining of hundreds of individual slides. Furthermore, quantitation of immunostainings is markedly easier on arrayed samples than on large sections. For example, it is possible to directly compare staining intensities of the different specimens on the same TMA slides, thereby improving the subjective interpretation of staining intensities. Most of all, the interpretation is, by default, limited to a small predefined area in arrayed samples. This facilitates a reproducible application of the selected scoring criteria because the entire tissue is always used for interpretation and the subjective selection of one tumor area for decision making is avoided. In the future, the TMA technology may help to optimize and standardize the interpretation of immunostainings, which is currently subjective and poorly reproducible and often leads to major discrepancies in studies investigating clinical associations for novel biomarkers. An exchange of stained or unstained TMA slides between laboratories reporting controversial data would help to rapidly unmask technical or interpretational reasons for conflicting study results.

Our data suggest that taking multiple punches from each tumor not only increases the number of interpretable tumors but also allows the distinction of three subgroups (positive, negative, and heterogeneous). The similar prognostic value for ER heterogeneous tumors as found for ER positive tumors suggests that false negative staining in one or several arrayed samples due to regional fixation problems may be the reason for heterogeneous findings in some of these tumors. The similarly poor prognosis of ER heterogenous tumors as found for ER negative tumors could theoretically be explained by an insufficient response of heterogeneous tumors to hormonal therapy. In this study, the use of four different samples per tumor did also result in a marked increase of interpretable tumors if data from multiple replica arrays were combined. However, analysis failure in up to 30% of arrayed samples was caused by technical problems related to the early generation of tissue arrays. Improved array making will reduce the need of multiple samples per tumor to increase the number of cases available for evaluation. The current success rate in our laboratory for arraying breast cancers is now greater than 90%.

In summary, our data suggest that TMAs can be successfully used to establish associations between molecular changes and clinical endpoints. Array-based tissue analysis is a rapid, cost-effective, and tissue-saving method for high-throughput clinicopathological studies.

Molecular markers that appear most promising based on TMAs constructed from retrospectively collected sets of tumors will then need to be prospectively validated using molecular analyses of larger tissue specimens available from prospective clinical trials.

References