

TMA-Combiner, a simple software tool to permit analysis of replicate cores on tissue microarrays

Chih Long Liu¹, Kelli D Montgomery², Yasodha Natkunam², Robert B West², Torsten O Nielsen³, Maggie CU Cheang³, Dmitry A Turbin³, Robert J Marinelli⁴, Matt van de Rijn² and John PT Higgins²

¹Biological and Biomedical Sciences Program, Division of Medical Sciences, Harvard Medical School, Cambridge, MA, USA; ²Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA; ³Genetic Pathology Evaluation Centre, University of British Columbia, Vancouver Hospital and British Columbia Cancer Agency, Vancouver, British Columbia, Canada and ⁴Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA

We have previously published a suite of software tools that facilitates the reformulation of tissue microarray (TMA) data so that it may be analyzed using techniques originally devised for analysis of cDNA microarray data. However, current microarray data often feature multiple scores for a given tissue sample and antibody combination. Furthermore, an efficient and systematic method for combining scores that takes into account the differing staining properties of tissue epitopes has not been described. We thus present the TMA-Combiner, a new Microsoft Excel-based macro that permits analysis of data for which tissues may have two or more scores per antibody, and permits combination of data from multiple different tissue microarrays. It accomplishes this by rendering one score per tissue per antibody from two or more scores, using one of multiple user-selectable combination rules developed to account for the differing staining properties of tissue epitopes. This greatly facilitates analysis of tissue microarrays, particularly for users with large repositories of data, and may facilitate discovery of biological trends and help refine diagnostic accuracy of tissue markers in clinical samples. Modern Pathology (2005) 18, 1641–1648. doi:10.1038/modpathol.3800491; published online 16 September 2005

Keywords: immunohistochemistry; replicate; duplicate; bioinformatics

Tissue microarray data, like data derived from cDNA microarray studies, are characterized by a staggering degree of complexity that requires computer-based analysis. For each antibody-stained core, a score is rendered to reflect an assessment of the adequacy of the tissue, quality and degree of staining, and cell type of interest. Each array may contain several hundred cores and may be stained with dozens or even hundreds of antibodies. The process of identifying meaningful staining patterns in such an extensive data set would be extremely cumbersome if it required re-evaluation of cores using the traditional method of glass slides, a microscope, and a slide key. We have thus previously published a suite of software tools that facilitate the analysis of tissue microarray data¹ and which have since been

used in a number of subsequent studies.^{2–11} One component, the TMA-Deconvoluter, permits hand-scored tissue microarray data from multiple antibodies to be reformulated rapidly so that it can be hierarchically clustered and viewed with TreeView using software originally developed for cDNA microarray data analysis. For physical processing of TMAs, a high-throughput tissue microscope digital image acquisition and archival system permits rapid and efficient acquisition of a high-resolution digital image of each core. For on-line retrieval of digitized core images, Stainfinder fetches images of each stained core through hyperlinks embedded in a TreeView display.

While these tools greatly facilitate the analysis of TMA data, they do not address the problem presented by replicated cores. It has been suggested that as many as four cores of each sample may be necessary for accurate interpretation of protein expression by immunohistochemistry.¹² Using multiple cores will minimize the possibility of nonrepresentative tissue, poor-quality staining, or

Correspondence: CL Liu, PhD, Department of Chemistry and Chemical Biology, 12 Oxford Street, Cambridge, MA 02138, USA. E-mail: clliu@fas.harvard.edu
Received 25 May 2005; accepted 14 July 2005; published online 16 September 2005

focal staining in producing a false-negative result for a sample. However, this further increases the complexity of the TMA data set, since each tissue will now have up to four scores per antibody. More importantly, there will be many occasions where all four scores will not be in agreement, necessitating a system to be devised for addressing these cases. Furthermore, established laboratories will frequently have multiple TMAs generated for a given tissue class of interest (eg lymphoma), often with varying degrees of antibody and case overlap. And

on some occasions, there might be multiple reads done on the same slide by different pathologists. All of these situations point to the need for an ability to combine these data into a single data set for efficient analysis. We therefore present here the TMA-Combiner (Figure 1), a portable Excel-based macro that combines replicate cores into one score, and of multiple tissue arrays into one data set. The TMA-Combiner is ideally suited to using data in the format rendered by the TMA-Deconvoluter. Combined data sets may then be analyzed using Cluster

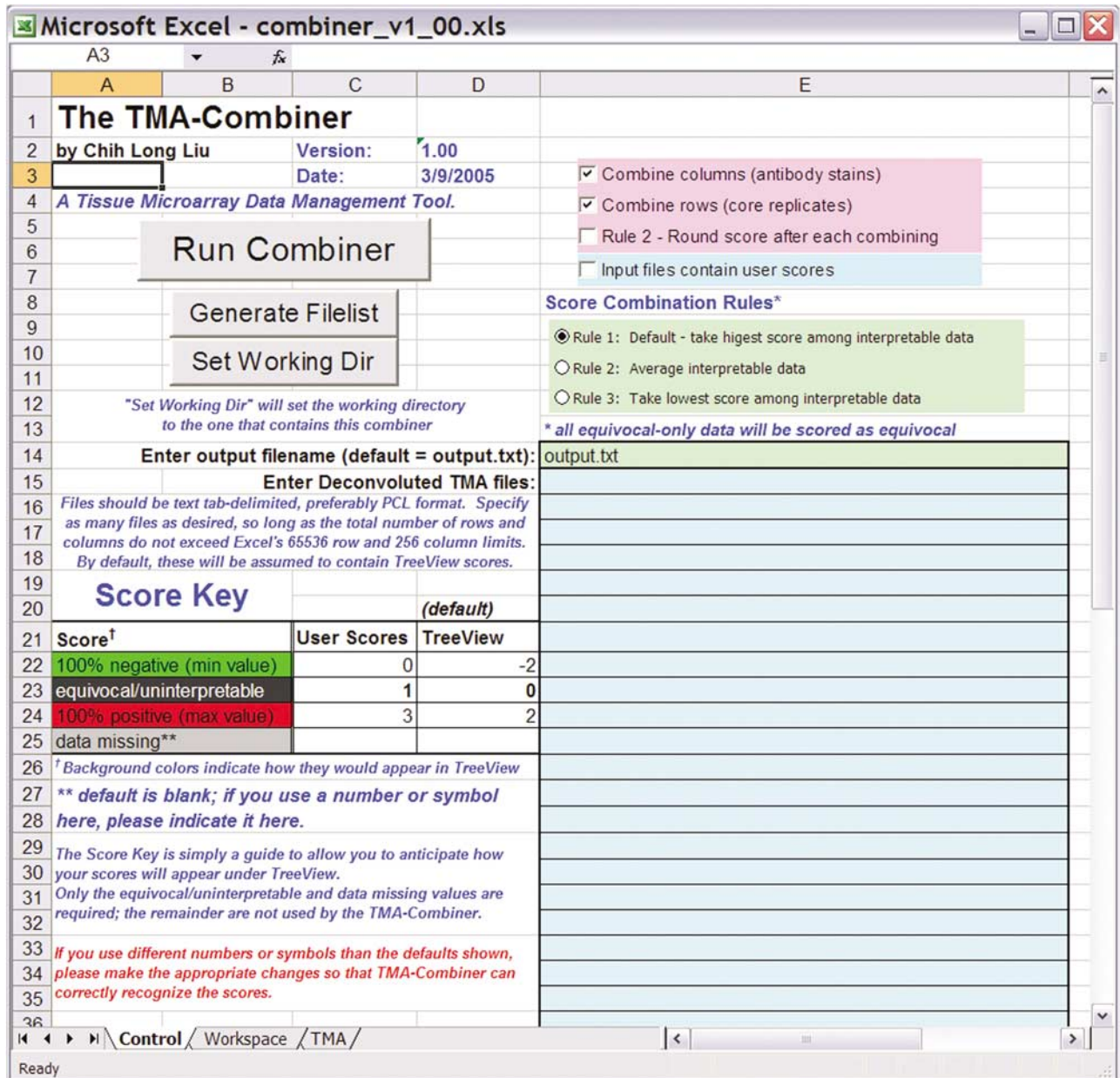


Figure 1 The TMA-Combiner User Interface. The TMA-Combiner is a macro-containing Microsoft Excel workbook, just like the TMA-Deconvoluter.¹ Users enter one or more text tab-delimited output files from the TMA-Deconvoluter, adjust the score key as needed, select a score combination rule, and then run the Combiner, producing an output file that can subsequently be hierarchically clustered. Refer to the online documentation in the download section of the supplemental website, <http://genome-www.stanford.edu/TMA/combiner/download.shtml>, or the on-line walkthrough, <http://genome-www.stanford.edu/TMA/combiner/walk-overview.shtml>, for specific details on operations and options.

and visualized using TreeView, as described previously.¹ The TMA-Combiner is freely available at <http://genome-www.stanford.edu/TMA/combiner>.

Overview: using the TMA-Combiner

The TMA-Combiner is meant for researchers who have a number of tissue microarrays and want to do rapid analyses. It is freely available and can be rapidly implemented. Although the TMA-Combiner is not meant to take the place of a more elaborate, server-based database, the most important analyses

necessary for tissue microarray research such as frequency of positive staining can be performed using this simple program. Even correlation with outcome can be explored using commercially available statistical programs such as Winstat. The TMA-Combiner should be used when a TMA contains multiple cores taken from one or more samples on a single tissue microarray, or when two or more TMAs (which may or may not have replicate cores) are to be analyzed as a single data set. In either case, raw tissue microarray data should first be acquired and processed in a manner previously described.¹ Figure 2 shows an overview of the entire process.

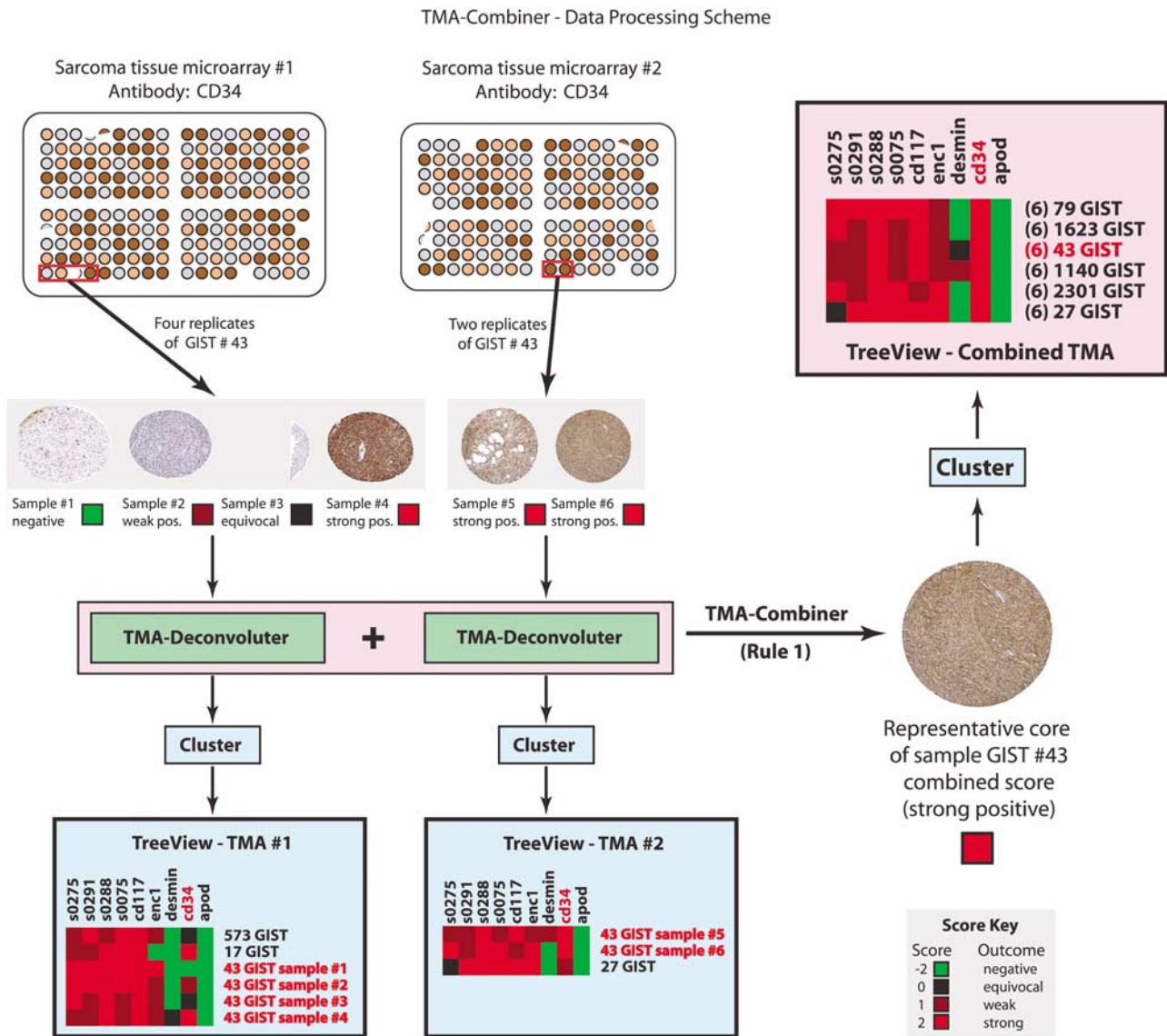


Figure 2 The TMA-combiner data processing scheme. This scheme is based on minor modifications to the data processing paradigm for our previously described software tools.¹ The process starts with one or more TMAs containing replicate cores, such as the two depicted here. Below these are the replicate core images, along with their scores and TreeView color box representations. Previously, each TMA would be separately processed, as shown. However, the TMA-Combiner combines the data from the two arrays immediately following deconvolution. The end result is a single TreeView window displaying combined data from both TMAs; the number in parentheses indicates the number of cores analyzed for this sample. While this figure depicts only a single representative core image for the combined GIST #43, the updated Stainfinder can retrieve and display all six cores used to produce the combined score.

Briefly, TMA data are stored in an Excel workbook in which each spreadsheet represents an antibody with which the array was stained. A master worksheet indicates the catalog number of each core, whose values can be retrieved with the investigator's catalog number look-up file. The TMA data set is then reformatted using TMA-Deconvolter into a two-dimensional tabulated data format. The TMA-Deconvolter software is freely available at: <http://genome-www.stanford.edu/TMA>.

Once the TMA data set has been deconvoluted, the data are ready for processing with the TMA-Combiner. First, for multiple TMAs to be combined together, the TMA-Combiner can accept multiple deconvoluted files, each corresponding to a different array. Next, combination can be limited to cases (via rows), antibodies (via columns), or both (in which case combination will start with cases, then proceed to antibodies). After this option has been selected, one of three combination rules (highest, average or lowest) can then be chosen. Which of the three rules should be selected will depend on the interests of the investigator or the characteristics of tissue staining of a given marker, as illustrated with examples in the next section.

Once the combination options have been selected, the TMA data are 'combined' into an output table that, like the deconvoluted input files, can be processed by the Cluster program. In the case of multiple TMA data sets having been combined together, there may be core/antibody combinations for which there are no data (eg the array has not been stained with those particular antibodies). As depicted in the score key (Figure 2), cores that lack data for an antibody will be rendered in gray on the TreeView hierarchical cluster (heatmap), as distinguished from cores that were stained but received equivocal scores, which will be rendered in black.

Score combination rules

There are several possible methods for combining scores from multiple different cores taken from the same tissue. One obvious choice would be to average the scores from each replicate (Figure 3, Rule 2). However, not all tissue epitopes stain in a quantitative fashion. For example, in the case of whole sections of a poorly differentiated neoplasm stained for keratin, focal staining is taken as presumptive evidence for epithelial differentiation. This may correspond to only one of four cores staining on a tissue microarray. Therefore, even if only one core shows convincing staining, it is probably sufficient to infer differentiation and the positive core should be taken to represent the most appropriate interpretation for that tumor. On the other hand, focal staining on whole sections of a breast carcinoma stained for Her2neu would be considered negative for overexpression. In this case, if only one of four cores shows strong Her2neu

expression, one of the negative cores should be taken as representing the tumor. Other markers such as hormone receptors or Ki67 would normally be averaged on whole sections. For these markers, negative, weak positive, and strong positive cores should all be taken into consideration in order to derive an estimate of the average level of expression of these markers.

Thus, the following rules we describe represent our logical and straight-forward approach to score combination, based on our current scoring system.¹ In constructing the rules, we first reasoned that equivocal or otherwise uninterpretable scores should first be removed from the analysis, since they are tantamount to missing data points. Since simple averaging (Rule 2) is not appropriate for all situations, we created Rules 1 and 3, which allow one to take focal staining as significant or to require diffuse strong staining for significance. We therefore anticipate that most users of the TMA-Combiner will want to combine their tissue microarrays according to the three rule options built into the combiner interface, since these options reflect the current interpretive standards for different antibodies in clinical use.

The rules themselves, illustrated in Figure 3, consist of the following:

1. (default rule) *Take highest interpretable score.*
Mechanics: This rule will automatically assign the highest score (usually *strong positive*) among the data points to the combined data point. However, if the data points consist only of negative (and equivocal) scores, the combined data point will be assigned a score of 'negative'. If the data points consist only of equivocal scores, the combined data point will be assigned a score of 'equivocal'.
Justification: this rule corresponds to the standard clinical practice of diagnostic immunohistochemistry used for many antibodies, particularly in the tissue microarray setting, hence its default status. For example, any cytokeratin expression is taken as evidence of epithelial differentiation. Even focal staining with anticytokeratin antibodies in a poorly differentiated neoplasm may be scored as positive when deciding between melanoma, lymphoma and poorly differentiated carcinoma, and taken to provide support for epithelial differentiation. A similar approach to diagnosis is taken for most antibodies, including those directed against HMB45, LCA, CD20, muscle actin, desmin, and others. We have previously described this method of combining discrepant scores from duplicate cores.^{11,13}
2. *Average interpretable data.* *Mechanics:* this rule takes the data points to be combined, removes equivocal scores, and calculates an arithmetic mean based on the remaining data points. If there are no good data points remaining, the combined data point will be assigned an equivocal score.

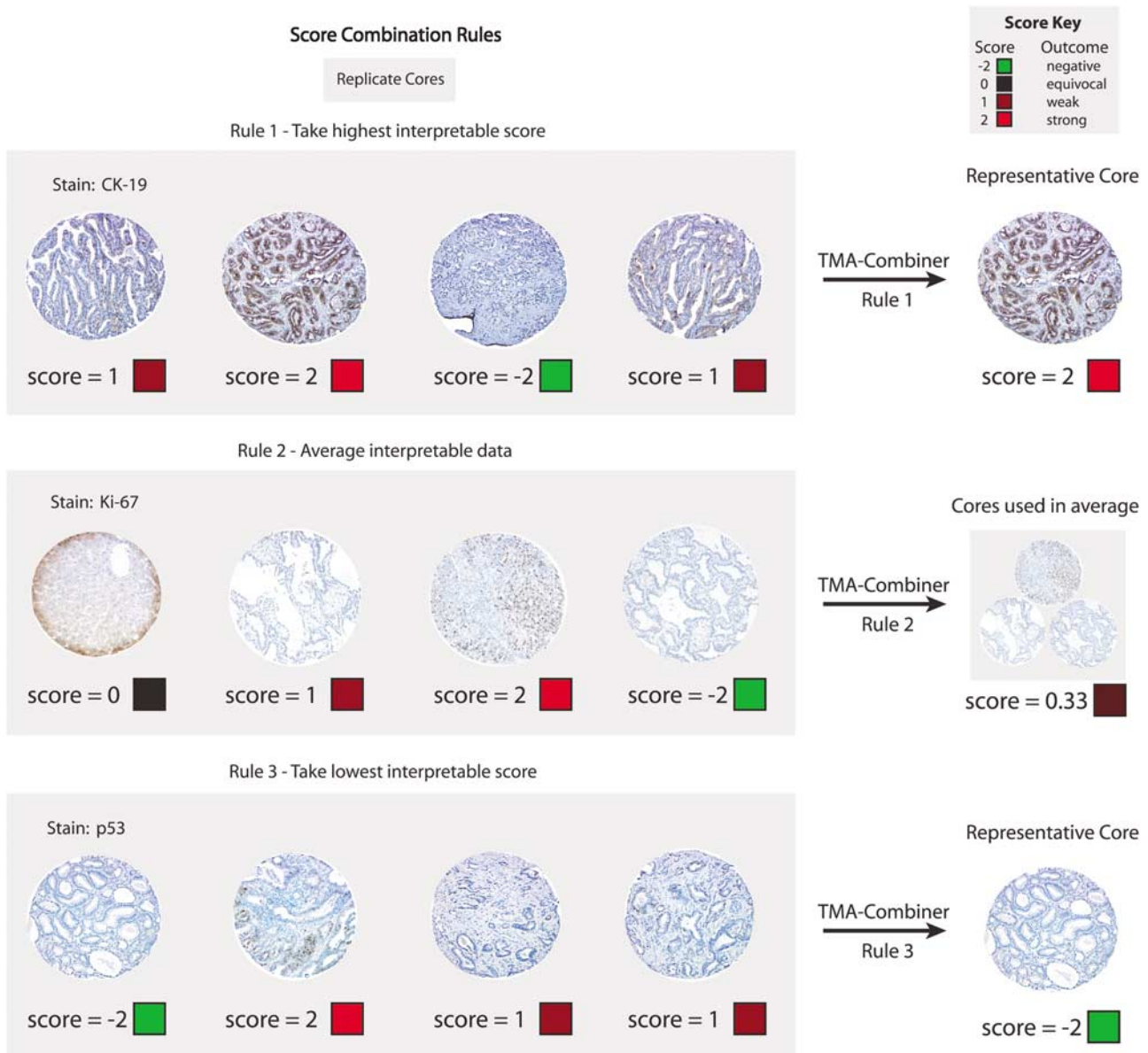


Figure 3 Score combination rules. Three different score combination rules are available for combining scores from multiple replicate cores into a single value. For each rule, four replicate cores are shown, along with their TreeView-compatible scores and representative TreeView color boxes. In Rule 1, the highest score from interpretable cores is taken, as indicated by the representative core exhibiting strong positive staining. In Rule 2, all interpretable scores are used to calculate an average score. Note that the core with equivocal staining on the far left was not included in the average. For the purposes of this figure, only representative cores are shown for the combined score. However, a combined data set configured with Stainfinder will retrieve all the replicate core images used to produce the combined score shown in the CaseExplorer or TreeView data set browser. A number of previously published data sets^{2,4,8} exhibiting this property are available for online browsing here: <http://genome-www.stanford.edu/TMA/combiner/explore.shtml>.

Justification: for several antibodies in clinical use, the assessment is made by the degree of positive staining. Estrogen and progesterone receptor staining is scored for both intensity of staining and proportion of the tumor cell population that stains. Staining for the proliferation marker, Ki-67, is generally reported as a percentage of cells that stain. This rule is thus designed to take degree of staining into consideration. If a protein is focally expressed, this will likely manifest itself

in some replicate cores showing staining, with other replicates negative. With this rule, such tissues will be assigned accordingly lower values. We anticipate that this rule will be used next most frequently after Rule 1, and would probably be the most appropriate rule for quantitative or semi-quantitative scoring systems, such as the Allred scoring system^{14,15} or the H-score method.¹⁶ In fact, many genes that show promise as diagnostic markers by cDNA expression studies may also

show differential expression across samples without being completely switched off in any of them. In this case, degree of expression of the cognate protein may be critical in distinguishing between different classes of tissues.

3. *Take lowest interpretable score. Mechanics:* this rule will automatically assign the lowest score (usually negative) among the data points to the combined data point. Thus, even if weak or strong positives are present, if a negative is present, the combined score will be designated as negative. If the data points consist only of equivocal scores, the combined data point will be assigned an equivocal score. This rule is appropriate for antibodies and/or tissue types that are susceptible to a substantial incidence of false positives or when only strong, diffuse expression of an oncogene may correlate with a biologic parameter.

Justification: This rule represents the inverse of Rule 1 in which the bias is given to a lack of staining. A potential utility for this rule has been illustrated by Torhorst *et al*,¹⁷ using p53 staining of breast carcinomas on tissue microarrays. These authors found a strong correlation of TMA results with prognosis while the correlation for whole sections was much weaker. Close inspection revealed that this resulted from the authors inadvertently using more stringent criteria for 'positive' for the TMA than for the whole sections. In other words, in order to correlate with a poor prognosis, p53 expression needs to be diffusely and strongly positive. A similar approach would apply to Her2neu. The FDA approved HerceptTest from DAKO is scored as negative, 0 (no staining or staining in less than 10% of tumor cells), negative, 1+ (faint, barely perceptible staining in more than 10% of tumor cells), positive, 2+ (weak overexpression, partial or weak membrane staining in 10% of tumor cells) and positive, 3+ (strong overexpression, strong, complete membrane staining in over 10% of cells).¹⁸ This scoring scheme was devised because only diffuse strong positive staining correlates with amplification of the Her2neu gene. On tissue microarrays, one may wish to require that every core shows extensive strong staining, before suggesting that Her2neu, or a similar protein, is over-expressed.

Note: In all of these rules, if there are no scores assigned to all of the data points to be combined, the resultant combined score will be left blank. Furthermore, these rules assume the scoring system as shown in Figure 1, the TMA-Combiner user interface. For a more detailed discussion on how these rules work, please refer to the score combination section of the online walkthrough for the TMA-Combiner here: <http://genome-www.stanford.edu/TMA/combiner/walk-comb2.shtml>.

Display of combined TMA files in TreeView

Combined TMA data sets that have been hierarchically clustered will have a number of minor differences when compared with hierarchically clustered files of uncombined TMA files (Figure 2). In combined TMA data sets, a number in parentheses appears in front of the case catalog number of the tissue sample and represents the number of rows in the original uncombined deconvoluted files that have been combined into the current row. For example '(6)' means that the current row was derived from 6 rows that were designated with that case catalog number. Furthermore, in the case of Rule 2, when used with a scoring system consisting of discrete integral values such as that of our own scoring system, combined TMA data sets may produce a heatmap with additional gradations of red and green intensity not normally seen in Rule 1 or 3 combination or in uncombined data sets. Under such circumstances, these represent averaged scores taken from the multiple cores scored for each antibody and may result in nonintegral values. For example, as depicted in Rule 2 of Figure 3, if a tissue was represented by three good cores with scores of 1, 2, and -2, then the resultant value would be 0.33, and the color depicting the result would be a very dim red, dimmer than the weak-positive score as depicted on the score key (Figure 3). In addition, some gray blocks will be present on the heatmap, which will usually result from combination of multiple TMAs. This occurs because some of the samples combined were never stained with some of the antibodies present in the data set, or if some of the constituent TMAs included cases not represented on other TMAs.

Linking a combined TMA sample through TreeView to Stainfinder also required modifications to Stainfinder. The TMA-Combiner has been designed such that it will pass all the image filenames from all the combined spot images to an updated version of Stainfinder (available at <http://genome-www.stanford.edu/TMA/combiner/download.shtml>). In a properly configured system, when a user clicks on an embedded link in the heatmap of a combined TMA data set, the new Stainfinder will retrieve cores from all of the samples represented by the single combined row. A number of previously published data sets,^{2,4,8} using the TMA-Combiner and configured with the updated Stainfinder, are available for on-line browsing here: <http://genome-www.stanford.edu/TMA/combiner/explore.shtml>.

Discussion

A central precept of tissue microarray technology is that a small core of tissue will show similar staining characteristics to the entire block of tissue from which it was taken. This appears to be true to a great degree. However, tissue microarrays inevitably show

cores with imperfections that complicate the analysis. Such cores result, for instance, from samples that completely or partially fall off the slide during processing (Figure 2), do not contain the cells of interest, or have an uneven staining pattern that is difficult or impossible to interpret. Even among the clearly positively stained cores, there are many that present interpretive challenges in assessing the degree of staining. Furthermore, experience with immunohistochemistry on whole sections clearly suggests that some proteins are expressed only focally within a neoplasm.

These problems can be mitigated by the use of two or more cores per sample. One formal study determined that the optimal number of 0.6 mm cores is up to four per sample;¹² reviewed in Packeisen *et al.*¹⁹ However, this manifold representation of each sample presents a further problem with data analysis: how to render a single, definitive score for each sample. It is impractical, at least in high-throughput fashion, to do this manually by inspecting all of the stained cores and picking the one that contains the greatest volume of the tissue of interest and the clearest pattern of staining. Furthermore, such a labor intensive, manual method runs the risk of introducing user-specific biases and inconsistencies in the selection process. Thus, TMA-Combiner makes it possible to rapidly select the best cores available for each tissue sample and to select a method by which the staining interpretation on multiple cores from the same source can be combined into a single value. The software works equally well whether the replicated cores are present on one array or are distributed over multiple arrays. While we anticipate that the current, greatest utility for this software will be combining cores together from multiple arrays for a single tissue, we believe that combining multiple replicates of a single antibody stain into a single score will become very important as the evolving process of constructing TMAs increasingly involve replicate cores.

Indeed, the process of making and using TMAs has evolved significantly within the past 3 years. At the time of our earlier publication on tools used in tissue microarray analysis,¹ each tissue was represented by a single core on our arrays. However, due to the aforementioned reasons, most of our current arrays now use two or four cores. Furthermore, for most pathologic categories, we are now on our second or third generation of TMAs. This has entailed further need for the ability to combine multiple TMA data sets together.

For example, an initial lymphoma array in our laboratory was produced and used to characterize a novel immunohistochemical marker. Subsequently, the tissue in that array was exhausted, while additional cases of interest were discovered. A new array was thus produced that included many of the cases from the first array but also these additional cases. Furthermore, this second array was then used to characterize novel markers, some but

not all of which were used on the first array. Our existing tools were inadequate for managing these additional complexities, so we created the TMA-Combiner, which presents a rapid means of combining the data from these arrays into an integrated data set.

Our current scoring system has remained unchanged since publication of our TMA data management system.¹ As a result of this, we have presented the TMA-Combiner accordingly, illustrating how we use it to fulfill our own TMA data set combination needs. However, we have enabled the TMA-Combiner to handle alternative and quantitative scoring systems, and we have updated the TMA-Deconvoluter with this added feature. We have designed the software to be applicable to the general pathology community, as we have done with the TMA-Deconvoluter and the rest of our TMA data management system. Although a large-scale laboratory analyzing data from hundreds or thousands of different arrays may require a server-based database, such as that provided by Oracle, we anticipate that the vast majority of tissue microarray researchers will find the TMA-Combiner well suited to their needs.

Thus, in such a rapidly evolving technology platform as tissue microarrays, not only does the TMA-Combiner facilitate the analysis of larger collections of TMA data, it offers opportunity in discovery of new biological trends and increased diagnostic accuracy, opportunities that might not be available through the more limited analysis of smaller, individual data sets. The freely available TMA-Combiner therefore enhances the suite of tools that facilitate analysis of tissue microarray data.

Acknowledgements

This work was supported by a Graduate Research Fellowship from the National Science Foundation. TON is a scholar of the Michael Smith Foundation for Health Research. MCUC and DAT were supported in part by an unrestricted educational grant from Aventis, Inc.

References

- 1 Liu CL, Prapong W, Natkunam Y, *et al.* Software tools for high-throughput analysis and archiving of immunohistochemistry staining data obtained with tissue microarrays. *Am J Pathol* 2002;161:1557–1565.
- 2 Nielsen TO, Hsu FD, O'Connell JX, *et al.* Tissue microarray validation of epidermal growth factor receptor and SALL2 in synovial sarcoma with comparison to tumors of similar histology. *Am J Pathol* 2003; 163:1449–1456.
- 3 Makretsov NA, Huntsman DG, Nielsen TO, *et al.* Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically signifi-

- cant groups of breast carcinoma. *Clin Cancer Res* 2004;10:6143–6151.
- 4 West RB, Harvell J, Linn SC, *et al*. Apo D in soft tissue tumors: a novel marker for dermatofibrosarcoma protuberans. *Am J Surg Pathol* 2004;28:1063–1069.
 - 5 Natkunam Y, Lossos IS, Taidi B, *et al*. Expression of the human germinal center-associated lymphoma (HGAL) protein, A new marker of germinal center B cell derivation. *Blood* 2005;105:3979–3986.
 - 6 Subramanian S, West RB, Corless CL, *et al*. Gastrointestinal stromal tumors (GISTs) with KIT and PDGFRA mutations have distinct gene expression profiles. *Oncogene* 2004;23:7780–7790.
 - 7 Kazanjian A, Wallis D, Au N, *et al*. Growth factor independence-1 is expressed in primary human neuroendocrine lung carcinomas and mediates the differentiation of murine pulmonary neuroendocrine cells. *Cancer Res* 2004;64:6874–6882.
 - 8 West RB, Corless CL, Chen X, *et al*. The novel marker, DOG1, is expressed ubiquitously in gastrointestinal stromal tumors irrespective of KIT or PDGFRA mutation status. *Am J Pathol* 2004;165:107–113.
 - 9 Nielsen TO, Hsu FD, Jensen K, *et al*. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 2004;10:5367–5374.
 - 10 Somasiri A, Nielsen JS, Makretsov N, *et al*. Overexpression of the anti-adhesin podocalyxin is an independent predictor of breast cancer progression. *Cancer Res* 2004;64:5068–5073.
 - 11 Alkushi A, Irving J, Hsu F, *et al*. Immunoprofile of cervical and endometrial adenocarcinomas using a tissue microarray. *Virchows Arch* 2003;442:271–277.
 - 12 Rubin MA, Dunn R, Strawderman M, *et al*. Tissue microarray sampling strategy for prostate cancer biomarker analysis. *Am J Surg Pathol* 2002;26:312–319.
 - 13 Hsu FD, Nielsen TO, Alkushi A, *et al*. Tissue microarrays are an effective quality assurance tool for diagnostic immunohistochemistry. *Mod Pathol* 2002;15:1374–1380.
 - 14 Allred DC, Clark GM, Elledge R, *et al*. Association of p53 protein expression with tumor cell proliferation rate and clinical outcome in node-negative breast cancer. *J Natl Cancer Inst* 1993;85:200–206.
 - 15 Allred DC, Harvey JM, Berardo M, *et al*. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. *Mod Pathol* 1998;11:155–168.
 - 16 Umemura S, Itoh J, Itoh H, *et al*. Immunohistochemical evaluation of hormone receptors in breast cancer: which scoring system is suitable for highly sensitive procedures? *Appl Immunohistochem Mol Morphol* 2004;12:8–13.
 - 17 Torhorst J, Bucher C, Kononen J, *et al*. Tissue microarrays for rapid linking of molecular changes to clinical endpoints. *Am J Pathol* 2001;159:2249–2256.
 - 18 Birner P, Oberhuber G, Stani J, *et al*. Evaluation of the United States Food and Drug Administration-approved scoring and test system of HER-2 protein expression in breast cancer. *Clin Cancer Res* 2001;7:1669–1675.
 - 19 Packeisen J, Korsching E, Herbst H, *et al*. Demystified tissue microarray technology. *Mol Pathol* 2003;56:198–204.