Quantification of Subcellular Molecules in Tissue Microarray

Ali Can, Musodiq O. Bello, Michael J. Gerdes
GE Global Research Center, Niskayuna, NY, 12309, USA
{can, bellom, gerdes}@research.ge.com

Abstract

Quantifying expression levels of proteins with subcellular resolution is critical to many applications ranging from biomarker discovery to treatment planning. In this paper, we present a fully automated method and a new metric that quantifies the expression of target proteins in immunohisto-chemically stained tissue microarray (TMA) samples. The proposed metric is superior to existing intensity or ratio-based methods. We compared performance with the majority decision of a group of 19 observers scoring estrogen receptor (ER) status, achieving a detection rate of 96% with 90% specificity. The presented methods will accelerate the processes of biomarker discovery and transitioning of biomarkers from research bench to clinical utility.

1. Introduction

Quantification of target proteins at subcellular resolution enables direct association of the expression levels with the localization of specific biochemical activities. Frequently, the subcellular localization of a protein, namely in the membrane, cytoplasm, or nucleus, dictates the protein's function. Large-scale prediction and correlation studies can be designed based on automatically quantifying protein expression patterns in tissue microarrays (TMA) with known clinical outcomes [1]. In this paper, we present a new metric to quantify subcellular protein expression levels. Pathologists manually determine the expression levels in immuno-histochemically stained tissues to diagnose and grade cancer. For example, the expression level and percentage of Estrogen Receptor (ER) protein localized in the nuclei is visually evaluated to determine ER protein status in breast cancer patients [2-4]. The score assigned by the pathologist is based on a minimum percentage of epithelial nuclei having the marker present. Frequently, this score is expanded to also include the percentage of positive cells and the strength of the signal. Determination of ER protein expression is critical to ascertain the response of patients to drugs (Tamoxifen or other anti-estrogens); to predict survival time (ER+ is a favorable indicator); and to differentiate endocervical (ER-) from endometrial (ER+) adenocarcinomas. Different scoring methods are presented in the literature based on the percentage of tumor cells with positive expression with a range of 1-20% of cells with expression being considered ER+, or based on mixed percentage and strength expression levels [3, 5].

Chromogenic detection of antibodies bound to the target proteins, which is widely used in traditional pathology, relies on enzymatic amplification and can suffer from diffusion of the signal. On the other hand, fluorescence based detection is both linear and highly sensitive. A sample of fluorescence ER staining with a distinct nuclear localization pattern is illustrated in Figs. 1(a, c). The segmentation result is presented in Figs. 1(b, d).

We present a two-step automated subcellular quantification system. In the first step, the subcellular regions are segmented from a set of fluorescent images of compartmental markers. DAPI, a fluorescent dye that binds to DNA, is used to detect the nuclei and to generate the nuclear segmentation map. Fluorescently labeled antibodies conjugated to keratin protein are used to define the epithelia and cytoplasmic regions, while antibodies conjugated to pan-cadherin are used for defining epithelial extracellular membranes. In the second step, the distributions of a target protein (for example ER) in each of the epithelial compartments are calculated. In this paper we present a new quantification metric based on Kolmogorov Smirnov distance defined on these distributions. Our metric is robust to tissue auto-fluorescence, and non-specific background binding, and it is different from existing intensity or ratio based techniques [6-8]. Our fully automated quantification system can be used to score commonly used target markers ER, androgen receptor (AR), progesterone receptor (PR), tumor protein 53 (TP53) and human epidermal growth factor receptor 2 (Her2).

2. Segmentation of sub-cellular regions

The first step of our automated quantification system is to use a set of compartmental markers to segment the subcellular regions. We utilized a general likelihood function estimator to calculate the probability maps of membrane and nuclei-like structures in single channel images of membrane (pan-

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.624

cadherin) and nuclei (DAPI) markers. The probability maps encode the segmentation information of different shapes in images using probability values between zero and one. The algorithm iteratively estimates empirical likelihood functions of curvature and intensity based features. Geometric constraints are imposed on the curvature feature to detect nuclei or membrane structures in fluorescent images of tissues. Our method is non-parametric and can learn the distribution from the data. This is different from existing parametric approaches, because it can handle arbitrary mixtures of blob and ridge like structures. This is essential in applications such as in tissue imaging where a nuclei image in an epithelial tissue comprises both ridge-like and blob-like structures. The network of membrane structures in tissue images is another example where the intersection of ridges can form structures that are partially blobs. Further details can be found in [9].

3. Protein expression quantification

The second step of our automated quantification is to quantify the relative distribution of target markers (such as ER, AR, PR, TP53, Her2) in the subcellular compartments. The intensity distribution of target protein in any compartment, *C*, is estimated using a Parzen window approach [10] with Gaussian kernels,

$$P_C(k) = \sum_{ij} w_{ij} \exp\left\{-(sk - I_{ij}^T)^2 / 2\sigma^2\right\}, \quad (1)$$

where weights, $w_{ij} = P(I_{ij} \in C)$, are the probability of a pixel belonging to compartment *C*, I_{ij}^T is the intensity of the target protein at pixel location *ij*. The scaling factor, *s*, is set to the ratio of the dynamic range of the image to the number of bins, *N*; and index *k* varies from 0 to *N*. The density function is normalized to give a sum of 1. The probability density function (PDF) of a target protein (Fig. 1(a)) on each of the epithelial regions (Fig. 1(b)) are illustrated in Fig. 1(e), where nuclear, membrane, and cytoplasmic distributions are plotted in blue, red and green respectively. The subcellular region of the target protein comprises of expressed and non-expressed (or non-specific, background) regions.

Signed Kolomogorov-Smirnov Distance: A well-known method to test if two distributions are different is to calculate the Kolmogorov-Smirnov (KS) distance [11] between the distributions. The associated test to measure statistical significance is commonly used in statistics and known as the Kolmogorov Significance test. We define a modified version of the KS distance and keep the sign of the distance to indicate which of the compartments is expressed more.

Given the PDF, the cumulative distribution function (CDF) is calculated as

$$F_C(k) = \sum_{\tau=0}^{k} P_C(\tau). \quad (2)$$

Figure 1(f) illustrates the CDF of the target distribution in Fig. 1(e), on each of the subcellular regions. The CDF clearly indicates over expression of the nuclear region (blue) where approximately 10% of the nuclear pixels express intensity values more than 50 on a [0,255] scale, as opposed to a very small percentage for the other compartments.

We introduce a signed version of the KS distance and call it signed KS (sKS) distance to determine which compartment is expressed. The positive one-sided KS test statistic between two distributions X and Y is given as

$$D_{XY}^{+} = \max_k \left(F_X(k) - F_Y(k)\right). \quad (3)$$

The expression level where the maximum positive cumulative differentiation is achieved is denoted by

$$k_{XY}^{+} = \arg\max_k \left(F_X(k) - F_Y(k)\right). \quad (4)$$

Similarly, the negative one-sided KS test and the associated expression level is calculated as,

$$D_{XY}^{-} = \max_k \left(F_Y(k) - F_X(k)\right) = D_{YX}^{+}, \quad (5)$$

$$k_{XY}^{-} = \arg\max_k \left(F_Y(k) - F_X(k)\right) = k_{YX}^{+}. \quad (6)$$

Note that the negative one-sided KS test is identical to positive one-sided KS test except that the order of the distributions is swapped. Since we are searching for the differences in the higher expression levels, we define a new signed- KS (sKS) distance as follows,

$$D_{XY} = \begin{cases} D_{XY}^{+} & \left|D_{XY}^{+}/D_{XY}^{-}\right| \geq \mathrm{T} \\ -D_{XY}^{-} & \left|D_{XY}^{+}/D_{XY}^{-}\right| \leq 1/\mathrm{T} \\ D_{XY}^{+} & 1/\mathrm{T} < \left|D_{XY}^{+}/D_{XY}^{-}\right| < \mathrm{T},\ k_{XY}^{+} > k_{XY}^{-} \\ -D_{XY}^{-} & 1/\mathrm{T} < \left|D_{XY}^{+}/D_{XY}^{-}\right| < \mathrm{T},\ k_{XY}^{+} \leq k_{XY}^{-} \end{cases} \quad (7)$$

where the threshold, T, was set to 10 in our experiments. Note that this is identical to the commonly used KS distance when the threshold is set to one. This new sKS distance is intentionally set to bias the distance that occurs at the high expression levels (the conditions in the lower two expressions in Eq. (7)). For example, if the CDF difference between

the membrane and the nuclei is a non-negative function, the smaller of the two one-sided KS distances is always zero and Eq. (7) gives the same results as the conventional KS distance. However, if the difference between the CDFs is not strictly non-negative or non-positive, the peak that is expressed at the higher values is taken as the true expression level. The metric D_{XY} is defined as a signed metric, where the positive values indicate the higher expression of the Y distribution, and the negative values indicate the higher expression of the X distribution. This is particularly important to identify which distribution is expressed higher in addition to the level of expression.

4. Results

We demonstrated the effectiveness of the sKS metric on 123 TMA images from 55 patients (some patients are represented as multiple tissue cores) stained with DAPI (nuclei), pan-cadherin (membrane), Keratin (tumor/epithelial mask), and ER markers. DAPI and pan-cadherin were used to segment the subcellular compartments, and keratin to segment the epithelial mask. The computed PDF and CDF distributions of ER in each of the subcellular compartments are presented in Figs. 1(e, f). The sKS metric computed between the membrane and nuclei for the ER pattern in Fig. 1(a) is -0.24, where the negative sign is an indication of nuclear staining.

The distribution of the ER protein was calculated on each of the epithelial subcellular regions. Biologically, the ER protein is partially or fully localized only in the nuclear regions. The nuclear ER distribution comprises a mixture of partial expression, non-specific expression, and autofluorescence (AF). Using the sKS distance, the nuclear ER distribution is compared to the membrane ER distribution comprising only non-specific binding and AF. Therefore, the membrane compartment serves as a normalization factor.

The sKS distance for each image in the TMA was automatically calculated and compared to the independent assessment of 19 observers. The observers (non-pathologist) were trained to score the collected data set of images as ER positive if they visually identified more than 10% of the nuclei with ER expression. Then the majority of the votes determined the most likely score for each image. In addition to the majority score, we also recorded the percentage of observers voting positive. The observer cutoff was set to 50% to determine the ER positive with the majority rule, while the sKS cutoff is set to -3%. The negative sign represents that this is a nuclear expression relative to the membrane. The estimated

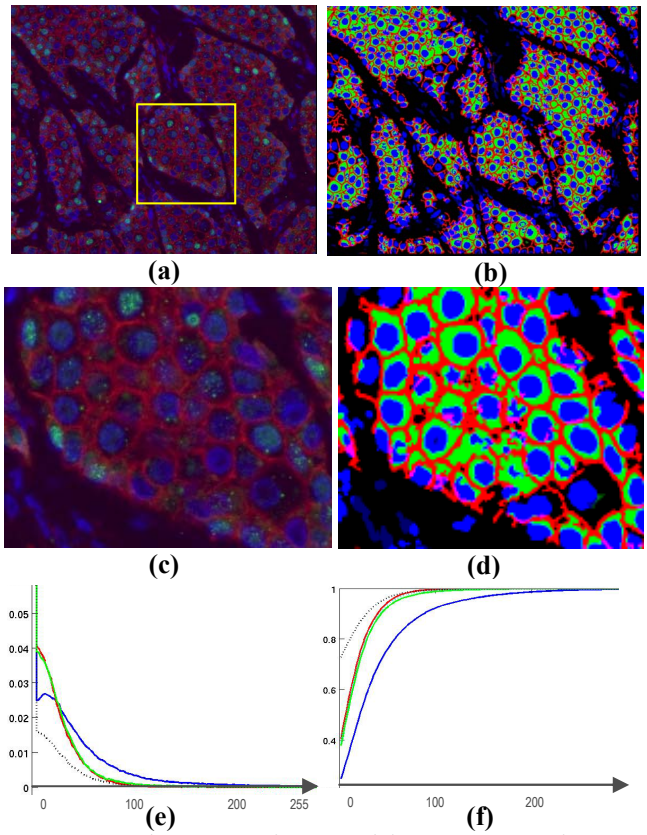
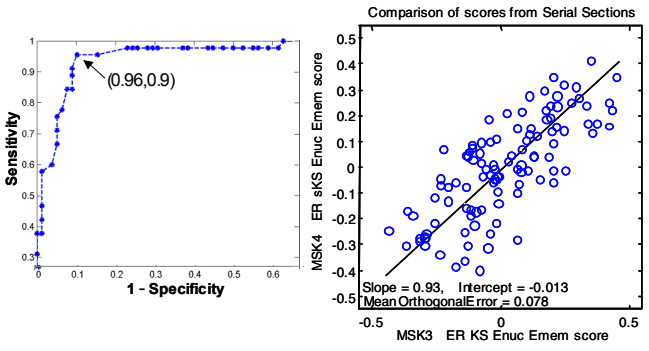


Figure 1: Fluorescent image of breast cancer tissue samples that are immunofluorescently stained. **a&c)** Immunofluorescent staining of ER (green) with partial expression; pan-cadherin stains membrane (red), DAPI stains nuclei (blue). **b&d)** Automatically segmented subcellular regions; membrane (red), nuclei (blue), cytoplasm (green). **e)** PDF of the ER on each of the subcellular compartments; membrane (red), nuclei (blue), cytoplasm (green). **f)** CDF of the ER distribution. The yellow box in (a) shows the region that is zoomed in the middle panels (c&d).

expressed percentages represent the percentage of nuclear area expressed rather than the percentage of number of nuclei. The automated score was very well correlated with the manual score, yielding only 8 false positives, and 2 false negatives compared to the majority of observers. Fig. 2(a) illustrates the ROC curve of the sKS score when the majority of the human observers is considered as the ground truth. At –3% cutoff, 96% sensitivity and 90% specificity was achieved.

To further validate the new sKS metric, two serial sections for the same patients acquired at different times were used to assess the robustness of the quantification to staining and tissue variations. Figure 2(b) illustrates the scores from the two serial sections. The slope of the orthogonal regression is close to 1

(0.926), and the intercept is close to 0. The mean orthogonal distance (0.078) to the fitted line is an indication of the robustness of the segmentation and quantification methods over time. We also compared the performance of the sKS metric with normalized mean difference (NMD), percentage and log ratio of expressions in the two compartments tested against the nuclei. As summarized in Table 1, sKS performs as good as or better than other metrics using Area Under the Curve (AUC), and Kappa Statistics as performance measures.



(a) (b)

Figure 2: a) ROC curve of the sKS score when the majority of the human observers are considered as the ground truth. **b)** Comparison of the automated scores between serial tissue sections stained and imaged six months apart.

5. Discussion and Further Research

We have shown a new fully automated method to determine the subcellular localization and relative abundance of partially expressed molecules in immunofluorescently labeled TMAs. Our methods are fully automated and designed to eliminate observer bias from the scoring. Using the intrinsic compartments as a normalization factor, our methods simplify the quality control process. In addition to clinical use, it can be used as a quality control tool to evaluate specimen preparation conditions as well. Extensive TMA analysis with clinical outcome data (survival times, drug resistance, recurrence, etc.) is needed to evaluate systems level performance.

Table 1: Comparison of performance

	Nuclei/Membrane		Nuclei/Cytoplasm	
	AUC(%)	Kappa	AUC(%)	Kappa
sKS	98.14	0.84	93.51	0.71
NMD	98.01	0.84	91.51	0.69
Percent	97.77	0.82	90.91	0.71
LogRatio	97.72	0.85	93.04	0.73

Acknowledgements

This work is supported by GE Global Research MI&D AT Program. We would like to thank F. Ginty, M. Montalto, M. Seel, S. Adak, X. Tao, Z. Pang, N. Barnhardt, A. Bhaduri, A. Sood, T. Treynor, S. Dinn, F. Pavan-Woolfe, S. Abbot, and J. Klimash, for discussions and participating in the observer study. The tissue sections were provided by Dr. W. Gerald from MSKCC.

References

[1] Ross, J., J. Fletcher, G. Linette, J. Stec, E. Clark, M. Ayers, W. Symmans, L. Pusztai, and K. Bloom, *The HER-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy*. The Oncologist, 2003. **8**(4) p. 307-25.

[2] Wasielewski, R., M. Mengel, B. Wiese, T. Rüdiger, H.K. Müller-Hermelink, and H. Kreipe, *Tissue array technology for testing interlaboratory and interobserver reproducibility of immunohistochemical estrogen receptor analysis in a large multicenter trial.* Am. J. Clinl Pathol. 2002. **118**(5): p. 675-82.

[3] Diaz, L.K. and N. Sneige, *Estrogen receptor analysis for breast cancer: current issues and keys to increasing testing accuracy.* Adv Anat Pathol., 2005. **12**(1): p. 10-19.

[4] Remmele, W. and H.E. Stegner, *Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection in breast cancer tissue.* Pathologe, 1987. **8**(3): p. 138-40.

[5] Layfield, L.J., D. Gupta and E.E. Mooney, *Assessment of tissue estrogen and progesterone receptor levels: a survey of current practice, techniques, and quantitation methods.* The Breast J. 2000. **6**(3): p. 189-96.

[6] Camp, R.L., G.G. Chung, and D.L. Rimm, *Automated subcellular localization and quantification of protein expression in tissue microarrays.* Nat Med. 2002. **8**(11): p. 1323-8.

[7] McClelland, R.A., *Automated quantitation of immuno-cytochemically localized estrogen receptors in human breast cancer.* Cancer Res., 1990. **50**(12): p. 3545-50.

[8] Said, M.V. and E.R. Barrack, *Image analysis of androgen receptor immunostaining in metastatic prostate cancer.* Cancer, 1993. **71**(11): p. 2574-2583.

[9] Can, A., M. Bello, H. Cline, X. Tao, P. Mendonca and M. Gerdes, *A unified segmentation method for detecting subcellular compartments in immuno-fluorescently labeled tissue images. MIAAB, Bethseda, MD, Sept. 2009.*

[10] Hastie, T., R. Tibshirani and J. Friedman, The elements of statistical learning: data mining, inference, and prediction. Springer, 2009.

[11] Devroye, L., L. Györfi and G. Lugosi, A probabilistic theory of pattern recognition. Springer, 1996.