J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Research Article

Feature-based analysis of mouse prostatic intraepithelial neoplasia in histological tissue sections

Pekka Ruusuvuori[1,2], Mira Valkonen[1], Matti Nykter[1], Tapio Visakorpi[1,3], Leena Latonen[1,3]

[1]Institute of Biosciences and Medical Technology - BioMediTech, University of Tampere, Tampere, [2]Tampere University of Technology, Pori, [3]Fimlab Laboratories, Tampere University Hospital, Tampere, Finland

E-mail: *Dr. Leena Latonen - leena.latonen@uta.fi
*Corresponding author

Received: 03 December 15 Accepted: 20 December 15 Published: 29 January 2016

Abstract

This paper describes work presented at the Nordic Symposium on Digital Pathology 2015, in Linköping, Sweden. Prostatic intraepithelial neoplasia (PIN) represents premalignant tissue involving epithelial growth confined in the lumen of prostatic acini. In the attempts to understand oncogenesis in the human prostate, early neoplastic changes can be modeled in the mouse with genetic manipulation of certain tumor suppressor genes or oncogenes. As with many early pathological changes, the PIN lesions in the mouse prostate are macroscopically small, but microscopically spanning areas often larger than single high magnification focus fields in microscopy. This poses a challenge to utilize full potential of the data acquired in histological specimens. We use whole prostates fixed in molecular fixative PAXgene™, embedded in paraffin, sectioned through and stained with H&E. To visualize and analyze the microscopic information spanning whole mouse PIN (mPIN) lesions, we utilize automated whole slide scanning and stacked sections through the tissue. The region of interests is masked, and the masked areas are processed using a cascade of automated image analysis steps. The images are normalized in color space, after which exclusion of secretion areas and feature extraction is performed. Machine learning is utilized to build a model of early PIN lesions for determining the probability for histological changes based on the calculated features. We performed a feature-based analysis to mPIN lesions. First, a quantitative representation of over 100 features was built, including several features representing pathological changes in PIN, especially describing the spatial growth pattern of lesions in the prostate tissue. Furthermore, we built a classification model, which is able to align PIN lesions corresponding to grading by visual inspection to more advanced and mild lesions. The classifier allowed both determining the probability of early histological changes for uncategorized tissue samples and interpretation of the model parameters. Here, we develop quantitative image analysis pipeline to describe morphological changes in histological images. Even subtle changes in mPIN lesion characteristics can be described with feature analysis and machine learning. Constructing and using multidimensional feature data to represent histological changes enables richer analysis and interpretation of early pathological lesions.

Key words: Histopathological image analysis, machine learning, prostatic intraepithelial neoplasia

Access this article online

Website: www.jpathinformatics.org

DOI: 10.4103/2153-3539.175378

Quick Response Code:

This article may be cited as: Ruusuvuori P, Valkonen M, Nykter M, Visakorpi T, Latonen L. Feature-based analysis of mouse prostatic intraepithelial neoplasia in histological tissue sections. J Pathol Inform 2016;7:5.
Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2016/7/1/5/175378

INTRODUCTION

This paper describes work presented at the Nordic Symposium on Digital Pathology 2015, in Linköping, Sweden.

Cancer is a major health care challenge and one of the most studied group of diseases worldwide.[1] In many cancers, e.g., with prostate cancer, differences between advanced cancer and normal tissue are relatively well known, yet the process of cancer development is still inadequately understood.[2] To gain further knowledge on the biological steps involved in oncogenesis, more understanding is needed for early changes leading to the development of cancer.

Prostatic intraepithelial neoplasia (PIN) represents a type of premalignant tissue with neoplastic epithelial growth confined in the lumen of prostatic acini.[3] To study oncogenesis in the prostate, early neoplastic changes in PIN are often modeled in the mouse by genetic manipulation of certain tumor suppressor genes or oncogenes.[4] E.g., it is well known that heterozygous deletion of tumor suppressor *Pten* induces PIN formation in the mouse prostate, in addition to hyperplasia and tumors in several other tissues.[5] Here, we analyze the histology of such mouse PIN (mPIN) lesions formed in the prostates of 10–11 months old *Pten* ± mice.

Early pathological changes in tissue tend to be small and/or subtle as far as histology is concerned. The neoplastic mPIN lesions used as a model here represent a group of samples with such early pathological changes with relatively narrow, but detectable, histological scope of alterations, and thus represent a challenge for histological quantitation. No previous studies to our knowledge exist providing computational models for assessment of mPIN histology. The mPIN lesions are macroscopically so small that they are undetectable by eye from the whole organ preparation. However, microscopically, the mPIN lesions can span areas larger that can be visualized in several high magnification focus fields in microscopy – both with an actual microscope or with digital pathology applications with standard computer screens. The generally accepted way to analyze such histological data is to snapshot single to several fields within the region of interest (ROI) to perform the analyses. In this way, only a fraction of the ROI is utilized often leaving most of the data unanalyzed. In addition, subjective errors are introduced through selection bias. There is a tendency to select areas not immediately adjacent to edges of the ROI, resulting in quantitation representing more the middle area of the lesion and leaving possible differences near the edges ignored.

More comprehensive ways to utilize the full potential of whole-slide scans and sequential sections are required. Visual inspection of only one to few slides and parts of the ROIs per lesion leaves a significant amount of information unobserved. Furthermore, to capture the potential spatial differences in the tissue, it is of importance to utilize information in three-dimensional (3D) to understand, e.g., cancer growth patterns better. By employing the full ROI areas and including 3D information, we can help basic scientific approaches to secure more comprehensive data. Eventually, using the enhanced methods for analysis, the goal is to ensure better clinical tools for the future.

Image analysis and machine learning provide powerful tools for processing microscope images and are increasingly applied in digital pathology.[6-8] Image analysis can be used for segmenting tissue areas and for extracting multidimensional feature data from full resolution whole slide images. Machine learning can be used for computationally learning a descriptive or discriminative model from the data either for segmentation or sample classification. The rationale behind using quantitative approach over the common use of few, simple readouts, such as size and volume, is that when making decisions of histology, also experts use several characteristics of the intensity, spatial distribution, and morphology of the tissue components. By quantifying such features, and using them for learning-based analysis, information about the properties common to tissue samples with histological changes can be acquired.

Our approach is to use a pipeline of automated image analysis methods for the exclusion of unwanted regions from further analysis and for extracting a large set of numerical descriptors for the studied tissue areas. Using this numerical representation of the ROIs, it is possible to visualize various properties and to observe similarities and differences within the population of samples. The features, along with the annotation obtained from ROIs segmented by an expert, are also used for training a classifier model. Machine learning enables incorporating expert knowledge in quantitative analysis in a sophisticated manner and is applicable in a small sample setting. Here, the classifier is used for determining the probability of histological changes in a given tissue sample. The contribution of individual features to classification provides information about differences in the histology of mPIN lesions.

METHODS

Tissue Material

We studied prostates of 10–11-months-old male FVB/N mice heterozygous for tumor suppressor *Pten* (n = 12; with 6 mPIN lesions/mouse on average). Half of the mice expressed ARR2PB-miR32 transgene (Latonen L, Visakorpi T, unpublished). Ethical approval for animal experimentation has been admitted by the Regional State Administrative Agency for Southern Finland (ESAVI/6271/04.10.03/2011).

Prostate tissues were fixed in PAXgene™ molecular fixative (PreAnalytiX GmbH, Hombrechtikon, Switzerland) according to manufacturer's recommendations and embedded in paraffin. The tissue blocks were sectioned through, and H&E staining was performed to three 5 μm sections every 50 μm apart to study prostate histology throughout the prostate. The slides were scanned with a Zeiss Axioskop 40 microscope (Carl Zeiss MicroImaging, NY, USA) with ×20 objective and a charge-coupled device color camera (QICAM Fast; QImaging, Canada) and a motorized specimen stage (Märzhäuser Wetzlar GmbH, Germany). The automated image acquisition was controlled by the Surveyor imaging system (Objective Imaging, UK). Uncompressed bitmap output was converted by JVSdicom Compressor application to JPEG2000 WSI format.[9]

Histological Assessment

Mouse prostate histology was assessed by an expert. These mice develop mPIN without any evidence of invasion through the basement membrane (Three, unpublished). Stages I–IV of mPIN were graded according to the guidelines introduced by Park *et al*.[10] The mPIN lesions which included areas with clear nuclear atypia and represented Grades II–IV, were included in the study. The presence of prominent nucleoli and increased atypia in nuclei were regarded as determinants in Grades III–IV. However, it must be noted that as the analysis takes into account the whole of the lesions in 3D, most of the lesions include areas representing several mPIN grades within them. Thus, it is not suited for the aims of the study to categorize lesions according to the grades. Rather, the terms "mild" and "advanced" phenotypes are used to describe the most variable nature of the alterations seen within individual lesions. "Mild" refers to the lesion with majority of Stage II–III alterations, as "advanced" refers to the lesion with majority of Stage III–IV alterations.

Image Processing

In total, there were 72 mPIN lesions, with 387 ROIs, included in the analysis (1–25 ROIs/lesion; mean value 5.4). A multiresolution approach was used for segmentation of ROIs. The best quality section from three adjacent H&E-stained sections was selected from each whole slide scan. The segmentation was done using a freehand selection tool in ImageJ, (National Institutes of Health, Bethesda, MD, USA),[11] for a low-resolution version of the original image obtained using the reduction levels in wavelet decomposition. The resulting binary mask was then resized to match the original image size and used for extracting the ROI from the full resolution original H&E image for further processing. The image processing pipeline, including feature extraction and machine learning, was developed using Matlab (The MathWorks, Inc., Natick, MA, USA).

The colors of the extracted lesion areas were equalized using contrast-limited adaptive histogram equalization. This reduces the color variation between different lesion sections, and the colors through the whole stack of images become more consistent. Exclusion of secretion-filled and empty areas was performed to obtain effective tissue area within ROI. These extraction phases are based on the subtraction of different color channels, and the segmentation of these areas is done using basic thresholding method minimizing the intraclass intensity variance.

Feature Based Analysis

Information of ROIs in each section level was included in the calculation of features for each lesion. In this proof-of-principle study, we extracted features describing each lesion area from one color channel (red). For lesions spanning over more than one section, the feature values are averaged over the sections. Features are listed in Table 1. Morphological features refer to the quantitative analysis of form. Multiple features describing the shape and size, including features such as area, major axis length, and minor axis length and perimeter were extracted. Before subsequent analysis, the features were scaled to zero mean and unit variance.

Textures are complex visual patterns, properties of which can be quantified with features describing the frequency, regularity, roughness, linearity, or smoothness of the studied area. The extracted texture based features included, e.g., mean intensity value, contrast, correlation, and energy, calculated from gray level co-occurrence matrix (GLCM). Textural features were also extracted using local binary pattern (LBP).[12,13] Properties of the lesions were also extracted using histogram of

Table 1: Feature categories

Feature type	Description	Number of features
Morphological feature	Features describing the volume and shape of the lesion	18
Unwanted regions	The amount of secretion and "empty area" inside the lesion	2
Intensity features	Texture features describe the spatial arrangement of intensity values in an image region. These features can be structural or statistical	11
Local binary pattern	Describe the appearance of an image in a small neighborhood around a pixel	10
Histogram of oriented gradients	Counts occurrences if gradient orientations in localized image regions	81
Maximally stable extremal regions	A method for blob detection from an image	5
Total		127

oriented gradients (HOG) descriptor[14,15] as well as with maximally stable extremal regions (MSER).[16] MSER implementation by VLFeat[17] was used in this work.

Machine Learning for Detecting Histological Changes

Machine learning was applied for detection of histological changes in tissue as follows. We used the feature representations of lesions for building a model which determines the probability for a given lesion to belong to the group with more advanced histological changes when compared to the group with milder changes. The grouping of lesions to the training samples with advanced and mild histological changes was done by an expert for a subset of lesions (see above; mild, $n = 11$; advanced, $n = 9$). A logistic regression classifier with the lasso (least absolute selection and shrinkage operation) regularization[18,19] was constructed using the training data. The output from the classifier is the class conditional probability of the advanced group, i.e., tissues with more prominent histological changes.

RESULTS

We screened mPIN lesions in whole mouse prostates by sectioning through the tissue with 5 μm sections [Figure 1]. To utilize full data and omit the subjective errors made by human-directed field selection, we performed our analysis for lesion areas in a stack in z. H&E-stained histological images every 50 μm apart were processed and used in the analysis. All mPIN lesions were masked as ROIs, and unwanted empty and secretion-filled regions were selected out within the ROI areas [Figure 2]. Resulting tissue areas within ROIs were subjected to feature analysis.

We extracted in total 127 features assessing shape, texture and spatial arrangement of the lesions [Table 1]. The features included morphological, textural, LBP, scale-invariant feature transform, HOG, and MSER features. All sections through each lesion were processed and combined into a feature vector, which provides a numerical representation for each lesion. Principal component analysis of the lesions according to the features is shown in Figure 3a. The analysis shows the contribution of each feature to the two first principal components, and the features are shown as lines starting from the origo. The red dots represent mPIN lesions projected into the principal component space. The values for each lesion are calculated across all ROIs of that particular lesion and averaged per lesion. As several texture features (mainly HOG features) show separation roughly in both directions along the y-axis, emphasis on size- and shape-related features vs. composition and texture (e.g., solidity, Euler number, extent, homogeneity, convex area, area, axis lengths, bounding box, and equivalent diameter) is visible roughly along the x-axis [Figure 3a]. As an example of separation of lesions by composition and shape descriptors, a scatter plot of average solidity versus sum of the area of lesions is shown in Figure 3b. Lesions are separated by these descriptors to what can be interpreted as different growth patterns and/or stages of mPIN advancement, as is shown for representative lesions with example section masks [Figure 3c].

We applied machine learning for detection of the relatively small histological changes within the group of mPIN lesions. We used certain lesions as training samples for phenotypically mild ($n = 11$) and slightly advanced ($n = 9$) lesions and used the feature representations of lesions for building a model which determines

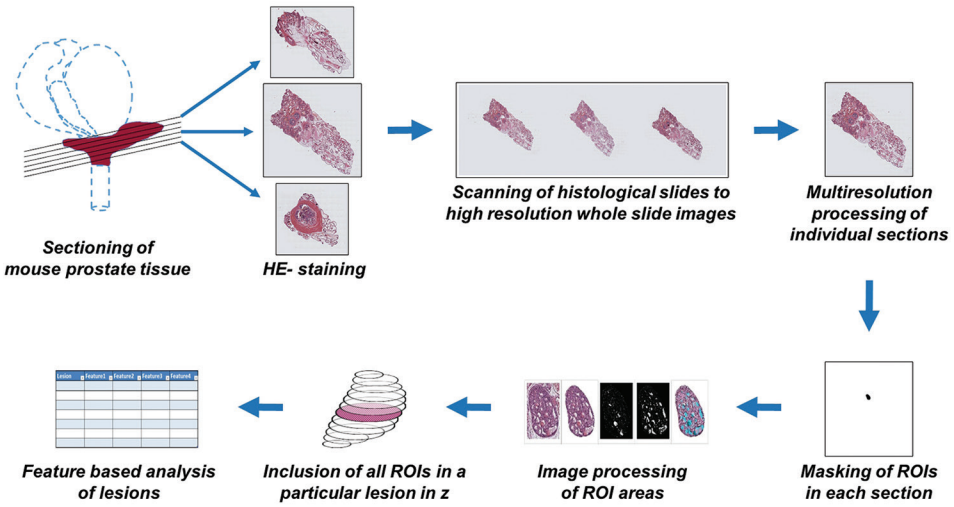

Figure 1: Processing of mouse prostate material for feature-based analysis. Whole mouse prostates were sectioned through with 5 μm sections. Sections were H&E stained, and the whole slide scanned to obtain high-resolution images. H&E-stained histological image every 50 μm apart was processed and used to mark mouse prostatic intraepithelial neoplasia lesions as a region of interests and subjected to image processing. All region of interests in a particular lesion obtained from a stack of histological images in z-direction were included and subjected to feature analysis

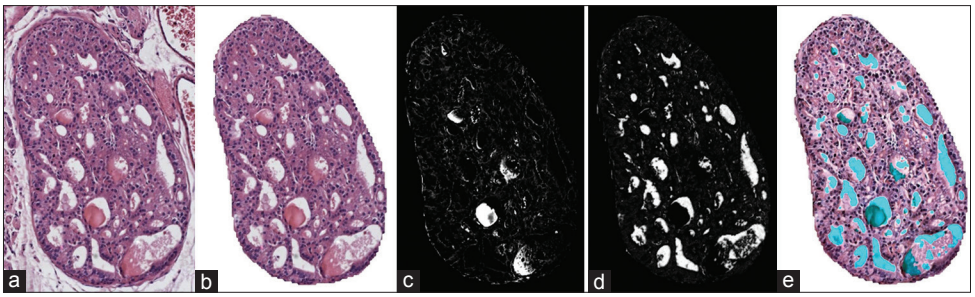

Figure 2: Image processing steps. An example of a PIN lesion area in a H&E-stained image (a) and its selection as a region of interest (b). Exclusion of secretion-filled (c) and empty (d) areas is performed to obtain effective tissue area within region of interest (e) excluded areas shown in turquoise

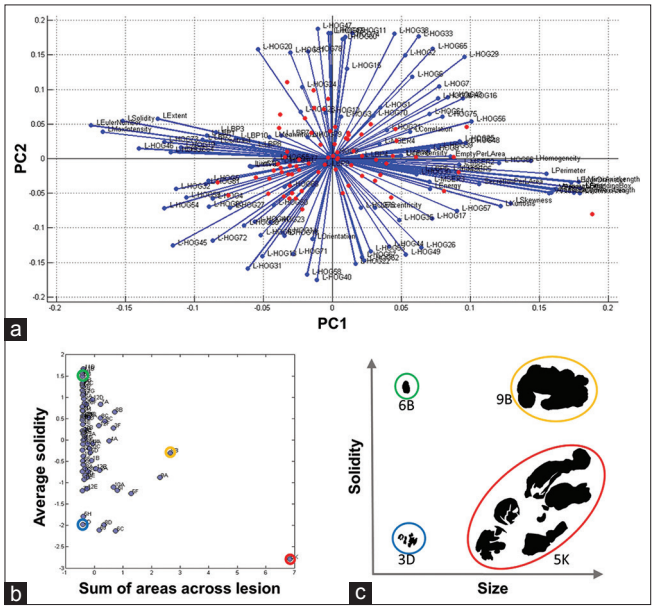

Figure 3: Feature analysis of mouse prostatic intraepithelial neoplasia lesions. (a) Separation of individual lesions (red dots, *n* = 72) according to principal component analysis and the relative weights of features (blue lines). The values for each lesion are calculated across all region of interests of that particular lesion. (b) Scatter plot of mouse prostatic intraepithelial neoplasia lesions with average solidity and sum of the area across lesion (i.e., the sum of areas of all lesion region of interests in all sections in z). (c) Examples of the type of lesion region of interest masks representing the morphological groups separated by lesion solidity and size (e.g., lesion area or volume). Colors in (b) indicate the same lesions of which an individual region of interest mask is shown in (c). Region masks in (c) are in the same scale

the probability for a given lesion to belong to either group [Figure 4a]. The plotted probabilities are averages from 1000 repetitions of leave-one-out experiments, where a randomly picked lesion was out of the training samples in Figure 4a, positive training samples are shown in red and negative in green. The hold-out experiment is a simulation of model stability in a small-sample setup. In addition to the left-out sample, the majority of all available lesions did not have an expert-defined class, and thus, were not used for training (blue markers in Figure 4a). The model successfully distinguished the remaining advanced lesions to the proper group with positive class conditional probabilities >0.5 and left the milder lesions below the threshold, examples of lesion ROIs are shown in Figure 4b. The models obtained during the 1000 repetitions composed of four features representing aspects of lesion size combined to shape (major axis length, eccentricity, equivalent diameter, and perimeter) and several texture features (LBPs, HOGs, and MSER2) [Figure 4c]. The histogram bins show the number of times each feature was selected in the model.

CONCLUSIONS

Studies of many diseases would benefit of enhanced quantitative assessment of early pathological changes. E.g., by understanding the alterations occurring early in cancer development, we could learn important aspects of neoplasia formation, which could help us find better ways to prevent, diagnose, and treat cancer. Early pathological changes leading to cancer formation are practically impossible to study in human samples, as early lesions go undetected and human samples lack the possibility for time series experimentation. Thus, as with many other diseases, cancer is often modeled in the mouse. Mice provide the possibility to follow neoplastic development and growth in time, as the genetically homogenous background of laboratory mice in combination with certain genetic manipulations provides relatively homogenous sample material over the specimens with a relatively constant development rate.

Early pathological changes in tissue tend to be macroscopically small and histologically subtle. However, they can span areas larger than single high magnification focus field in microscopy. To utilize the full potential of the data acquired in histological specimens with modern whole slide scanning systems, entire lesions should be used in quantitative assessments. This challenges the traditional way of analyzing few single focus fields with microscopes or snapshots from virtual slide viewers to obtain quantitative data. In this study, we tackled this problem by using sections with certain intervals (50 μm; every 10th of 5 μm sections) through the whole mouse prostate tissue and analyzing all ROIs of each lesion.

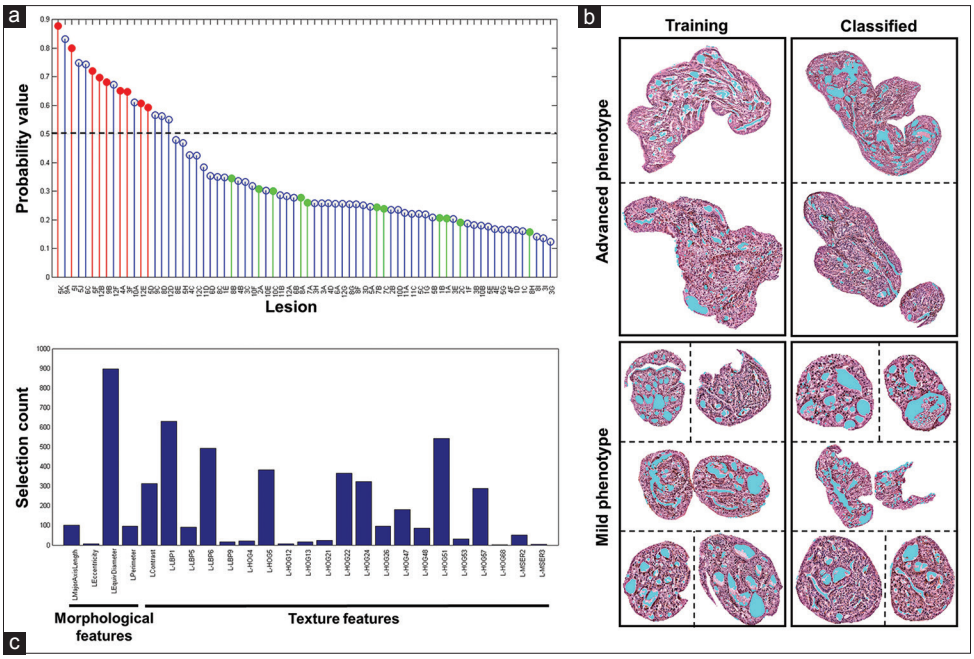

Figure 4: Mouse prostatic intraepithelial neoplasia lesion classifier model based on features obtained by machine learning. (a) Lesions sorted according to positive class conditional probabilities obtained during the 1000 repetitions of the classifier design by random hold-out of a training sample. Training samples are marked with colors (red, advanced phenotype; green, mild phenotype). A probability threshold of 0.5 is marked with a dashed line. (b) Examples of lesion region of interests representing the two phenotypes in the classification model. Examples of both training samples and classified samples are shown. It must be noted that the lesions are not in the same scale. (c) The histogram bins showing the number of times the features have been selected in the classifier model

The mPIN lesions, depending on their size and growth direction relative to cut orientation, spanned from 1 to 25 images.

As expected, the size- and shape-related morphological features were good descriptors of the mPIN lesion advancement. Size and visual complexity of growth pattern roughly correlate to growth rate and potency of neoplastic lesions, and characterization of such lesions by eye is partly dependent on these characters. As these parameters correspond well to what is observed by visual inspection of the lesion histology, it may not, at first, seem like their quantitation could bring novel information compared to human-inspected pathological screening. However, human eye has difficulties keeping track of growth patterns over several sections, especially if the growth curls and intertwines with other structures in the tissue and/or several lesions are included in the same tissue. Thus, computer-aided quantification enables inspection of lesion growth patterns in an additional dimension compared to two-dimensional (2D) visual inspection.

Parameters related to lesion size are sensitive to tissue cut orientation. E.g., in our case, mPIN in prostate tissue grow within the prostatic acini, creating more or less cylinder-shaped lesions. It needs to be taken into account in which orientation acini are cut in relation to how many sections there are. As in many other parameters, mean of ROIs (across sections) describes well the general result of each feature, in area-related features the entity of the lesion needs to be taken into account using, e.g., sums of all ROIs of a particular lesion or relation to volumetric estimations.

Qualities such as shape, texture, and spatial arrangement of lesions can expand our understanding of tumor biology and eventually could have the potential to improve the accuracy of clinical prognosis. Histopathologic classification based on visual inspection of an expert is subjective even with experienced pathologists, a challenge to, e.g. Gleason scoring for prostate cancer.[20] Features such as texture might be even impossible to accurately characterize by visually inspecting the images. Computer-aided quantitative analysis of histological tissue images is an effective tool to detect specific and small alterations too subtle or small to be separated by eye. For example, here, the histological changes described in features such as solidity or convex area may be spotted by eye when prominent enough, however, computer-aided quantitative analysis is far more accurate in positioning a particular sample within the range of variation within each parameter. Furthermore, the more complex and especially upto pixel-fidelity features, such as LBP and HOG texture features, compose of information not necessarily detected by human eye, and thus can provide an additional level of resolution to histological subtyping.

Machine learning and statistical pattern recognition typically require large sample sizes in order to provide

reliable models for prediction purposes. However, machine learning may also be seen as a way of using the expert labeled training for learning more about the available data even when the number of samples is low, as is often the case with the costly mouse model studies. Here, we used a classifier for determining the likelihood for histological changes, enabling ranking of uncategorized samples. Furthermore, machine learning is potentially useful in providing information about the importance of individual features in grouping of samples. For the purpose of interpreting the data using the classifier model parameters, we chose regularized logistic regression classifier. The motivation for this selection is that logistic regression, where model complexity is penalized with regularization, tend to produce sparse models,[18] which can be used for providing a selection of an informative subset of features in various applications, e.g., in the analysis of growth and morphology.[21] Here, the features selected by the classifier are potentially interesting parameters related to mPIN histology.

The ROI size-dependent parameters selected during the repeated training rounds, namely the sum of convex areas, eccentricity, equivalent diameter, and perimeter, can all be considered numerical descriptors of regularity versus irregularity in lesion shape. Early neoplastic lesions are small and regular in shape. As they grow, the space limitations of the tissue environment result in penetration to neighboring areas by growth pressure squeezing the surrounding tissue or, e.g., growth along lumen of an acinus. The more pronounced the growth, the stronger the pressure against the surrounding tissue, resulting in increasing acini diameters and the introduction of irregularity to formerly near-spherical or -cylindrical objects visible as round or elliptical shapes in 2D. Further progress to malignancy induces invasion to surrounding tissue. This generalized view of evolution in neoplastic growth patterns comply with the idea that increasing irregularity is a decent measure of histological stage of the lesion.

The texture features selected to the classifier are particularly interesting and potentially provide novel indicators of histological changes in early neoplastic tissue. Contrast, calculated from the GLCM, describes the difference of intensities between neighboring pixels. HOGs are edge and gradient based descriptors. The appearance and shape of an object can be characterized with the distribution of local gradients and edge directions without knowing the positions of the gradients and edges. The processed image is divided into smaller windows, and the distribution of local gradients and edge directions is calculated for each window.[15] MSER is an image element detector method, which generates features that are invariant to affine transformations.[16] The method is useful in baseline matching and in all, very robust and fast feature detector. LBP method in texture analysis is computationally simple and robust in terms of grayscale variations. These both are very important factors when considering histological image analysis, as the grayscale varies in these tissue section images due to lack of standard staining and scanning systems.

In summary, we report the development of quantitative image analysis pipeline to describe morphological changes in histological images using mPIN in mouse prostate tissue as a model for early neoplastic changes in the prostate. Our approach is to introduce quantitative rigor through image analysis and machine learning regardless of the sample size. To the best of our knowledge, such thorough computational pipeline for quantitative analysis of mouse PIN lesions using machine learning has not been presented elsewhere in the literature. Importantly, we show how these computational methods are powerful even in small sample settings, typical in mouse model studies. Representing histological changes by constructing and using multidimensional feature data can significantly contribute to interpretation and research of early pathological lesions. The methods described here, used in combination with potentially automated digital pathology applications, could provide improved and faster pipelines to grade early neoplastic lesions and to screen for potentially more malignant lesions from large sets of data. These methods could both enhance research of tumor models in, e.g., genetically manipulated model organism tissues, as well as aid in future attempts to develop better tools for digital pathology.

Acknowledgments

Funding for this work has been obtained from the Finnish Funding Agency for Innovation, project "three dimensional histogenomic modeling of whole prostate," Academy of Finland (grant no. 279270), the Cancer Society of Finland, the Sigrid Juselius Foundation, and the Medical Research Fund of Tampere University Hospital. We thank Marika Vähä-Jaakkola and Katja Liljeström for skillful technical assistance.

Financial Support and Sponsorship

Nil.

Conflicts of Interest

There are no conflicts of interest.

REFERENCES

1. Stewart BW, Wild CP, editors. IACR World Cancer Report 2014. Lyon, France: WHO Press; 2014.
2. Tomlins SA, Rubin MA, Chinnaiyan AM. Integrative biology of prostate cancer progression. Annu Rev Pathol 2006;1:243-71.
3. Merrimen JL, Evans AJ, Srigley JR. Preneoplasia in the prostate gland with emphasis on high grade prostatic intraepithelial neoplasia. Pathology 2013;45:251-63.
4. Shappell SB, Thomas GV, Roberts RL, Herbert R, Ittmann MM, Rubin MA, *et al.* Prostate pathology of genetically engineered mice: Definitions and classification. The consensus report from the Bar Harbor meeting of

the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. Cancer Res 2004;64:2270-305.
5. Di Cristofano A, Pesce B, Cordon-Cardo C, Pandolfi PP. Pten is essential for embryonic development and tumour suppression. Nat Genet 1998;19:348-55.
6. Ghaznavi F, Evans A, Madabhushi A, Feldman M. Digital imaging in pathology: Whole-slide imaging and beyond. Annu Rev Pathol 2013;8:331-59.
7. Hamilton PW, Bankhead P, Wang Y, Hutchinson R, Kieran D, McArt DG, *et al.* Digital pathology and image analysis in tissue biomarker research. Methods 2014;70:59-73.
8. Veta M, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: A review. IEEE Trans Biomed Eng 2014;61:1400-11.
9. Tuominen VJ, Isola J. Linking whole-slide microscope images with DICOM by using JPEG2000 interactive protocol. J Digit Imaging 2010;23:454-62.
10. Park JH, Walls JE, Galvez JJ, Kim M, Abate-Shen C, Shen MM, *et al.* Prostatic intraepithelial neoplasia in genetically engineered mice. Am J Pathol 2002;161:727-35.
11. Schneider CA, Rasband WS, Eliceiri KW. NIH image to imagej: 25 years of image analysis. Nat Methods 2012;9:671-5.
12. Ojala T, Pietikäinen M, Mäenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 2002;24:971-87.
13. Pietikäinen M, Ojala T, Xu Z. Rotation-invariant texture classification using feature distributions. Pattern Recognit 2000;33:43-52.
14. Ludwig O, Delgado D, Goncalves V, Nunes U. Trainable Classifier-fusion Schemes: An Application to Pedestrian Detection. Vol. 1. In: 12th International IEEE Conference on Intelligent Transportation Systems. 2009. p. 432-7.
15. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol. 1. 2005. p. 886-93.
16. Matas J, Chum O, Urban M, Pajdla T. Robust wide-baseline stereo from maximally stable extremal regions. Image Vis Comput Puting 2004;22:761-7.
17. Vedaldi A, Fulkerson B. VLFeat: An Open and Portable Library of Computer Vision Algorithms; 2008. Available from: http://www.vlfeat.org/. [Last accessed on 2016 Jan 16].
18. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc 1996;58:267-88.
19. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1-22.
20. Iczkowski KA, Lucia MS. Current perspectives on Gleason grading of prostate cancer. Curr Urol Rep 2011;12:216-22.
21. Ruusuvuori P, Lin J, Scott AC, Tan Z, Sorsa S, Kallio A, *et al.* Quantitative analysis of colony morphology in yeast. Biotechniques 2014;56:18-27.