MULTIRESOLUTION IDENTIFICATION OF GERM LAYER COMPONENTS IN TERATOMAS DERIVED FROM HUMAN AND NONHUMAN PRIMATE EMBRYONIC STEM CELLS

Amina Chebira[1], John A. Ozolek[3], Carlos A. Castro[4], William G. Jenkinson[5], Mukta Gore[2], Ramamurthy Bhagavatula[2], Irina Khaimovich[1,2], Shauna E. Ormon[1,2], Christopher S. Navara[4], Meena Sukhwani[4], Kyle E. Orwig[4], Ahmi Ben-Yehudah[4], Gerald Schatten[4], Gustavo K. Rohde[1] and Jelena Kovačević[1,2]

[1] Dept. of BME and Center for Bioimage Informatics, [2] Dept. of ECE
Carnegie Mellon University, Pittsburgh, PA, USA
[3] Dept. of Pathology, Children's Hospital of Pittsburgh
[4] Dept. of Obstetrics and Gynecology, Magee-Womens Research Institute and Foundation, University of Pittsburgh, Pittsburgh, PA, USA
[5] Dept. of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

ABSTRACT

We propose a system for identification of germ layer components in teratomas derived from human and nonhuman primate embryonic stem cells. Tissue regeneration and repair, drug testing and discovery, the cure of genetic and developmental syndromes all may rest on the understanding of the biology and behavior of embryonic stem (ES) cells. Within the field of stem cell biology, an ES cell is not considered an ES cell until it can produce a teratoma tumor (the "gold" standard test); a seemingly disorganized mass of tissue derived from all three embryonic germ layers; ectoderm, mesoderm, and endoderm. Identification and quantification of tissue types within teratomas derived from ES cells may expand our knowledge of abnormal and normal developmental programming and the response of ES cells to genetic manipulation and/or toxic exposures. In addition, because of the tissue complexity, identifying and quantifying the tissue is tedious and time consuming, but in turn the teratoma provides an excellent biological platform to test robust image analysis algorithms. We use a multiresolution (MR) classification system with texture features, as well as develop novel nuclear texture features to recognize germ layer components. With redundant MR transform, we achieve a classification accuracy of approximately 88%.

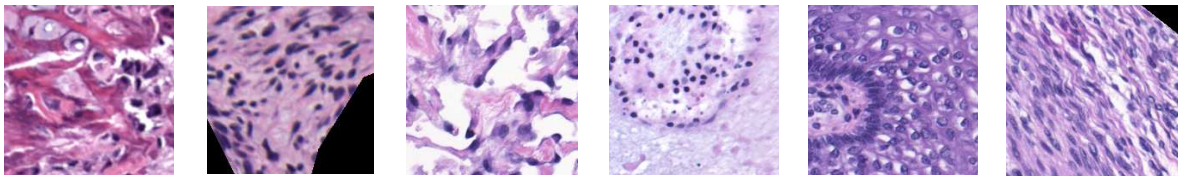
Index Terms— Stem cell biology, multiresolution, classification, feature extraction

1. INTRODUCTION

Embryonic stem cells (ES) and cells derived from them hold great promise both as therapeutic agents in clinical medicine as well as biological windows into the early stages of development. The range of therapeutic options includes repair of damaged or injured tissue (tissue regeneration after stroke, heart attack, cartilage renewal in arthritis), restoring defects in genetic, biochemical, and metabolic pathways, as well as drug testing and discovery [1, 2, 3]. Studying the sequence of genetic events within ES cells as they develop and differentiate into tissue, will have a significant impact on explaining and ultimately defining the therapy for a wide range of developmental syndromes. ES cells possess certain inherent characteristics that set them apart from any other cell type. They have the ability to self-renew, perpetuate indefinitely, and produce all three germ layers from which all tissue types are derived (pluripotency). Typically, in the laboratory, ES cells are defined by their expression of specific proteins and their behavior in cell culture. However, human and nonhuman primate cells isolated and cultured cannot be considered ES cells until they show the ability to produce a teratoma tumor when injected into immunocompromised mice. A teratoma is a tumor that is strictly defined by histological evidence of tissue types contributed by each of the original three germ layers. These include ectoderm (neuroepithelium, mature neuroglial tissue, skin), mesoderm (smooth and skeletal muscle, connective tissue, bone, and cartilage), and endoderm (lung and intestinal mucosa, pancreas, liver). While at first glance, most teratomas derived from ES cells appear as disorganized tissue masses with recognizable germ layer elements, little is known about the contribution of each germ layer to the lesion, and this information may hold important clues to normal and abnormal development.

While sophisticated image analysis and bioinformatics are burgeoning fields within pathology, most of the imaging applications have focused on the automation and digitization of the tissue processed for histological examination. New techniques are in development that would allow the pathologist to manipulate high-resolution images and sign out cases at the computer rather than examine tissue under a microscope. Image analysis currently allows segmentation of tissue areas defined by specific immunohistochemical stains to highlight the tissue of interest. However, the ability to automate the recognition of specific and varied tissue types from the routine hematoxylin and eosin (H&E) stained tissue sections (used almost exclusively to make the vast majority of diagnoses) is not available. As this applies to ES cell biology and teratoma analysis, advanced digital imaging applications would help answer the following questions: How much of each specific tissue type is present? How are these specific tissue arranged in space with respect to one another? How are the tissues present affected in type and quantity when derived from

This work was supported in part by NSF grants CCF-0515152 and EF-0331657, the PA State Tobacco Settlement, Kamlet-Smith Bioinformatics Grant as well as NIH grant PO1 HD 47675.



(a) (b) (c) (d) (e) (f)

Fig. 1. Sample H&E-stained images from each tissue class: (a) bone, (b) mesenchyme (embryonic connective tissue), (c) myenteric plexus, (d) necrotic tissue, (e) skin, and (f) striated muscle.

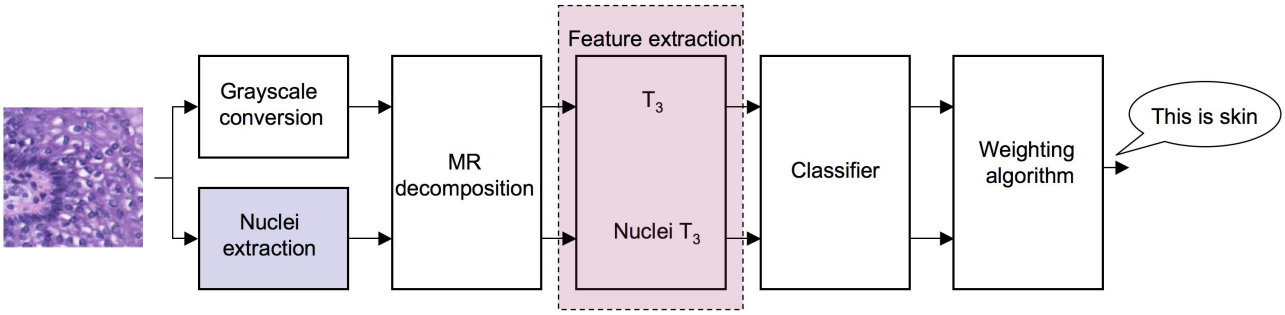


Fig. 2. Overview of the proposed H&E-stained tissue recognition system. The input is an H&E-stained image of one of the six tissue classes given in Fig. 1. The multiresolution (MR) nature of the system is accomplished through the MR decomposition block, after which all the processing is done in MR subspaces. We use T_3 features [4] and propose new nuclear texture features (see Fig. 4). The classifier is a simple neural network one. We use two versions of the weighting algorithm (open form and closed form). The output is the tissue class label.

ES cells that have been manipulated genetically, biochemically, and environmentally (that is, by drugs or toxins)? The ability to accurately detect and quantify specific tissue types will allow detection of species-specific differences in developmental programming and enable accurate three-dimensional reconstruction of teratomas and precise correlation with high-resolution MRI.

Since the teratoma contains many varied tissue types, the first step in this direction is to actually recognize these specific tissues types, a classification problem in image analysis. In our previous work [5, 4, 6, 7], we found that classifying in multiresolution (MR) subspaces adds to the discriminative power of a classification system. Moreover, based on our previous experience, modified Haralick texture features [4] (our improvement of the known Haralick texture features), were typically sufficient for high classification accuracies on those data sets (in the mid to upper 90s). The complex nature of the images in this work makes it unlikely that the texture features alone could produce such high accuracies.

In this work, we use our MR classification algorithm to design a system to recognize tissue types within teratomas derived from ES cells. We will show that MR classification again outperforms the nonMR one. We test various MR decomposition modules (bases and frames) and then develop a novel feature set for specific use in recognizing tissues in H&E stained images.

2. ADAPTIVE MULTIRESOLUTION CLASSIFICATION

This recognition problem can be posed as follows: Given an H&E stained image of a specific tissue type, output the label of that type. As in our previous work [4], we observe that these types of images (see Fig. 1) seem to possess a fair amount of features localized in space as well as frequency, leading us to hypothesize that classifying in MR subspaces would help discriminate tissues.

To that end, we developed an automated and adaptive MR classification algorithm [4]. This system adds MR decomposition in front of a generic classification consisting of feature extraction and classification in each MR subspace, yielding local decisions which are then combined into a global decision using an adaptive weighting procedure (the white blocks in Fig. 2). This system has proven accurate for various classification tasks [5, 4, 6, 7]; we now briefly describe it.

Multiresolution Block. In our classification system, any MR transform can be used, which we divide into two classes: nonredundant MR transforms (MR bases, MRBs) and redundant MR transforms (MR frames, MRFs). In particular, for MRBs, we used the discrete wavelet transform (DWT), while for MRFs we used the most redundant one, the stationary wavelet transform (SWT). (Note that here, we use all the subbands of the decomposition tree, not only the leaves, and thus, it might be abuse of language to call a transform a DWT. For example, for 2 levels, we have a total of $S = 21$ subbands (original image + 4 subbands at level 1 + 16 subbands at level 2).

Feature Extraction. Features are numerical descriptors that characterize the input data, usually in a lower-dimensional space. We focus on the following feature sets:

Haralick Texture Features (T_1). These features are calculated using four co-occurrence matrices [8]: 1) P_H (horizontal nearest neighbors), 2) P_V (vertical nearest neighbors), 3) P_{LD} (left diagonal nearest neighbors), and 4) P_{RD} (right diagonal nearest neighbors). Haralick calculates 13 measures on each of these four matrices. A T_2 set was also available but not found to be useful.

New Texture Feature Set T_3. We found that the Haralick texture features seem to possess the most discriminative power, so we examined these more closely. We changed the way that Haralick combines the initial four sets of features as P_H and P_V are fundamentally different from P_{LD} and P_{HD}, because adjacent neighboring pixels are

spatially closer than diagonal neighboring pixels. Therefore, instead of averaging the features from all four sets, we create our first set of 13 features by averaging horizontal and vertical measures, and a second set of 13 features by averaging diagonal measures, resulting in a new feature set T_3 with 26 features [4].

Neural Networks (NN). We used a two-layer NN classifier: The first contains a node for each of the input features, each node using the Tan-Sigmoid transfer function, while the second contains a node for each output and uses a linear transfer function (no hidden layers are used). When training, each output from the second layer corresponds to a class, and each training image will have an output of 1 for the class of which it is a member and a 0 for all other classes.

Weighting Procedure. Here we combine all of the subband decisions into one (see Fig. 2). We use weights for each subband to adjust the importance a particular subband has on the overall decision made by the classification system. As the combination with no-decomposition weight equal to 1 and all other weights equal to 0 is equivalent to not using the adaptive MR system, we know that there exists a weight combination that will do at least as well as the generic classifier (when no MR is involved) in the training phase. Our goal is to decide how to find the weight vector that achieves the highest overall classification accuracy on a given data set. We developed two versions of the weighting algorithm: open form (OF) and closed form (CF).

The NN block outputs a series of decision vectors for each subband of each training image. Each decision vector $d_s^{(r)}$ contains C numbers (where C denotes number of classes) that correspond to the "local" decisions made by the subband s for a specific image r. The classifier is evaluated using nested cross validations (five-fold cross validation in the NN block and ten-fold during the weighting process). To avoid the issue of the initial ordering of the images determining which images are grouped together for training and testing in each fold of the cross validation, we run multiple trials, each with a random initial ordering of the images.

Open-Form Algorithm. This is an iterative algorithm where we initialize all the weights, and for each image, a global decision vector is computed using a weighted sum of the local decisions, yielding an initial class label. If that class label is correct, we go to the next image. If it is incorrect, we look at the local decisions of each subband and adjust the weights of each subband s in a reward/punishment fashion; Those subbands making correct decisions are rewarded with an increased weight, whereas the ones making wrong decisions are punished with a decreased weight. We continue cycling through the images until there is no increase in classification accuracy on the training set for a given number of iterations.

Closed-Form Algorithm. The CF solution finds the weight vector w by solving a minimization problem in the least-squares sense. Assuming that d is the desired target decision vector and D the matrix grouping the local decisions of all of the subbands, we can write the minimization problem as:

$$w_{win} = \arg\min_{w} \|d - Dw\|,$$

which possesses a closed-form solution and can be computed efficiently. Then, for a testing image t, the winning class corresponds to the index of the highest coefficient in the decision vector of t.

3. NOVEL NUCLEAR TEXTURE FEATURES

The classification system described in the previous section works with gray-scale intensity images. The images obtained through histological techniques have the advantage of being highly detailed and showing distinctive features of the tissues at different resolutions. One of these important features is color. Therefore, it is important that the classification system takes advantage of all the information available from the histological images and exploits a feature as important as color. In particular, when using H&E staining, the nuclei turn a marked purplish/blue color from hematoxylin while the cytoplasm becomes varying shades of red due to exposure to eosin.

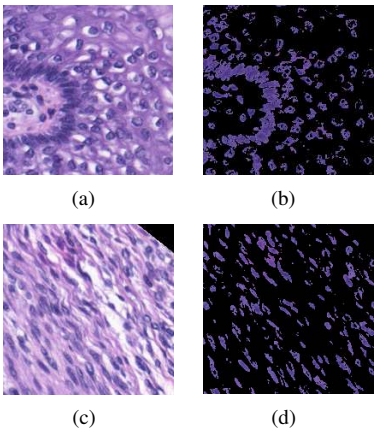


Fig. 3. Examples of tissue and nuclear-only images. On the left, are the original images of the tissue, while on the right, the corresponding extracted images of the nuclei are shown. The top row represents an image of skin and the bottom row is striated muscle.

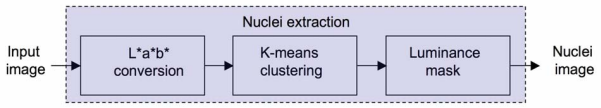


Fig. 4. Nuclear image extraction.

We observed that the cell nuclei have a distinctive distribution and texture depending on the tissue type. For example, nuclei in striated muscle image are mostly elongated and have the same orientation (Fig. 3(d)), whereas the nuclei in skin images are irregular with a jagged pattern (Fig. 3(b)). This prompted us to extract nuclei images from the original histological images and then compute texture features (NT_3) on these, to incorporate them in our classification system. Nuclear features in conjunction with other morphological features have also been used to classify breast cancer tumors in [9].

Fig. 4 depicts a block diagram of the nuclei extraction method; We first convert the original images from RGB to the perceptually uniform $L^*a^*b^*$ color space. The L^* channel is the luminance channel and a^* and b^* are the chrominance (color) channels that indicate where the color falls along the red-green and blue-yellow axes, respectively. Given that the images depict mainly three colors: white, blue and pink, we then use the K-means clustering algorithm on the a^* and b^* channels to derive three clusters corresponding to those three colors. We observe experimentally that the centroid with the largest difference between its red and blue channel pixel values corresponds to the blue cluster. Finally, since the nuclei take on the dark shades of blue, we use an adaptive threshold on the luminance channel to mask the lighter shades of blue from the blue cluster image, thus obtaining the nuclei images [10].

4. RESULTS AND DISCUSSION

Derivation of ES Cells and Teratomas. Human ES cells (H7 line) and putative nonhuman primate ES cells (derived as outlined by Navara [11]) were introduced into the testes of immune deficient SCID mice by modified efferent duct injection [12]. Cells were injected using an Eppendorf Femtojet pressure injector into the interstitial space of the testis. Tumors typically developed between two and four months after injection. Tumors were removed and processed by routine histological methods. Multiple serial sections were examined for evidence of tissue derived from the three germ layers.
Data Set. We have 45 images of size 200×200 for each of the six tissue classes contained in the teratomas: mesenchyme (embryonic connective tissue), skin, myenteric plexus, bone, necrotic (dying or dead tissue), and striated muscle. The images have been taken at 100x magnification (see Fig. 1) and have been labeled by a pathologist, thus we have access to ground truth. Note that the test images were never seen by the system during the training phase.
Experimental Setup. We used the following in our algorithm: For the MR block, we tested both MRBs in the form of the DWT as well as MRFs in the form of SWT. In the feature extraction block, we tested our new NT_3 features alone, the T_3 features alone, as well as the combination of both: T_3, NT_3. The classifier was the NN classifier as explained earlier, while we tested both the OF as well as the CF versions of the weighting algorithm.

Classification Accuracy [%]				
	Weight	NT_3	T_3	T_3, NT_3
NMR	NW	54.44	65.74	71.74
MRB	OF	63.22	70.02	77.29
	CF	64.10	71.38	78.20
MRF	OF	71.51	82.37	86.56
	CF	73.05	84.40	**87.72**

Table 1. Classification accuracy for tissue types in teratomas derived from ES cells. Along each row, feature sets are arranged by increased accuracy (with (T_3, NT_3) being the best). Along each column, MR blocks as well as weighting algorithms are arranged by increased accuracy as well; the MRF gives the best results.

Results. The results are given in Table 1 (note that NMR denotes the version of the algorithm where no MR transform is used). We note the following trends (the first three are consistent with the trends observed in all our previous work [5, 4, 6, 7]): (1) For all feature combinations, MR transforms (both MRBs and MRFs) significantly outperform NMR, thus showing that classifying in MR subspaces indeed improves the classification accuracy. (2) MRFs considerably outperforms MRBs and give the best classification accuracy of 87.72%. (3) The CF version of the algorithm outperforms the OF one. (4) Incorporating nuclear texture features significantly improves the accuracy using any MR block and any weighting algorithm. In particular, classification accuracy increases from 73.05% using NT_3 features only, to 84.40% using T_3 features only, and to 87.72% using the combined feature set T_3, NT_3.

5. CONCLUSIONS

We presented an automated and accurate MR identification algorithm of germ layer components in teratomas derived from human and nonhuman primate embryonic stem cells. The algorithm achieves an accuracy of 87.72% using MR frames.

For tissue type identification, we hope to improve the performance of the classification system by adding morphological features, as the shape of the cell and nucleus, which as noted demonstrates distinctive variations across tissue types. We will also incorporated 3D features as well as obtain more images for the training set so the classifier can see a larger variation during training. Our larger goal is to build an automated toolbox for extraction, recognition and quantification of the varied tissue types present in teratomas derived from ES cells and other pathological specimens using only routine hematoxylin and eosin stained tissue sections.

6. REFERENCES

[1] R. P. Lanza, *Essentials of Stem-Cell Biology*, Elsevier, 2006.

[2] C. W. Pouton and J. M. Haynes, "Embryonic stem cells as a source of models for drug discovery," *Nat. Rev. Drug. Discov.*, vol. 6, pp. 605–616, 2007.

[3] H. Thomson, "Bioprocessing of embryonic stem cells for drug discovery," *Trends in Biotechnol.*, vol. 25, pp. 224–230, 2007.

[4] A. Chebira, Y. Barbotin, C. Jackson, T. Merryman, G. Srinivasa, R. F. Murphy, and J. Kovačević, "A multiresolution approach to automated classification of protein subcellular location images," *BMC Bioinformatics*, vol. 8, no. 210, 2007, http://www.andrew.cmu.edu/user/jelenak/Repository/-07_ChebiraBJMSMK/07_ChebiraBJMSMK.html.

[5] R. A. Kellogg, A. Chebira, A. Goyal, P. A. Cuadra, S. F. Zappe, J. S. Minden, and J. Kovačević, "Towards an image analysis toolbox for high-throughput Drosophila embryo RNAi screens," in *Proc. IEEE Int. Symp. Biomed. Imaging*, Arlington, VA, Apr. 2007, pp. 288–291.

[6] A. Chebira, L. P. Coelho, A. Sandryhaila, S. Lin, G. W. Jenkinson, J. MacSleyne, C. Hoffman, P. Cuadra, C. Jackson, M. Püschel, and J. Kovačević, "An adaptive multiresolution approach to fingerprint recognition," in *Proc. IEEE Int. Conf. Image Proc.*, San Antonio, TX, Sep. 2007, vol. 1, pp. 457 – 460.

[7] A. Chebira and J. Kovačević, "Adaptive multiresolution frame classification of biological and biometric images," in *Proc. SPIE Conf. Wavelet Appl. in Signal and Image Proc.*, 6701:16, Ed., San Diego, CA, Aug. 2007, vol. 1-15.

[8] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, 1979.

[9] S. Petushi, C. Katsinis, M. M. Haber, F. U. Garcia, and A. Tozeren, "Large scale computations on histology images revels grade differentiating parameters for breast cancer," *BMC Medical Imaging*, vol. 6, no. 14, 2006.

[10] A.R. Weeks and G. E. Hague, "Color segmentation in the hsi color space using the k-means algorithm," *Proc. SPIE*, vol. 3026, pp. 143–154, Apr. 1997.

[11] C. S. Navara, J. D. Mich-Basso, C. J. Redinger, A. Ben-Yehudah, E. Jacoby, E. Kovkarova-Naumovski, and et al., "Pedigreed primate embryonic stem cells express homogeneous familial gene profiles," *Stem Cells*, vol. 25, pp. 2695–2704, 2007.

[12] T. Ogawa, J. M. Aréchaga, M. R. Avarbock, and R. L. Brinster, "Transplantation of testis germinal cells into mouse seminiferous tubules," *Int. J. Dev. Biol.*, vol. 41, pp. 111–122, 1997.