

## Image database analysis of Hodgkin lymphoma



Tim Schäfer<sup>a,1</sup>, Hendrik Schäfer<sup>a,1</sup>, Alexander Schmitz<sup>a,1</sup>, Jörg Ackermann<sup>a</sup>,  
Norbert Dichter<sup>a</sup>, Claudia Döring<sup>b</sup>, Sylvia Hartmann<sup>b</sup>,  
Martin-Leo Hansmann<sup>b</sup>, Ina Koch<sup>a,\*</sup>

<sup>a</sup> Molecular Bioinformatics, Institute of Computer Science, Cluster of Excellence "Macromolecular Complexes", Johann Wolfgang Goethe-University Frankfurt am Main, Robert-Mayer-Straße 11-15, 60325 Frankfurt am Main, Germany

<sup>b</sup> Senckenberg Institute of Pathology, Johann Wolfgang Goethe-University Frankfurt am Main, 60590 Frankfurt am Main, Germany

### ARTICLE INFO

#### Article history:

Received 21 March 2013

Received in revised form 26 April 2013

Accepted 26 April 2013

#### Keywords:

Hodgkin lymphoma

Image analysis

CD30<sup>+</sup> stain

Nodular sclerosis

Mixed cellularity

### ABSTRACT

Hodgkin lymphoma (HL) is a special type of B cell lymphoma, arising from germinal center B-cells. Morphological and immunohistochemical features of HL as well as the spatial distribution of malignant cells differ from other lymphoma and cancer types. Sophisticated protocols for immunostaining and the acquisition of high-resolution images become routine in pathological labs. Large and daily growing databases of high-resolution digital images are currently emerging. A systematic tissue image analysis and computer-aided exploration may provide new insights into HL pathology. The automated analysis of high resolution images, however, is a hard task in terms of required computing time and memory. Special concepts and pipelines for analyzing high-resolution images can boost the exploration of image databases.

In this paper, we report an analysis of digital color images recorded in high-resolution of HL tissue slides. Applying a protocol of CD30 immunostaining to identify malignant cells, we implement a pipeline to handle and explore image data of stained HL tissue images. To the best of our knowledge, this is the first systematic application of image analysis to HL tissue slides. To illustrate the concept and methods we analyze images of two different HL types, *nodular sclerosis* and *mixed cellularity* as the most common forms and reactive lymphoid tissue for comparison. We implemented a pipeline which is adapted to the special requirements of whole slide images of HL tissue and identifies relevant regions that contain malignant cells.

Using a preprocessing approach, we separate the relevant tissue region from the background. We assign pixels in the images to one of the six predefined classes: *Hematoxylin*<sup>+</sup>, *CD30*<sup>+</sup>, *Nonspecific red*, *Unstained*, *Background*, and *Low intensity*, applying a supervised recognition method. Local areas with pixels assigned to the class *CD30*<sup>+</sup> identify regions of interest. As expected, an increased amount of *CD30*<sup>+</sup> pixels is a characteristic feature of nodular sclerosis, and the non-lymphoma cases show a characteristically low amount of *CD30*<sup>+</sup> stain. Images of mixed cellularity samples include cases of high *CD30*<sup>+</sup> coloring as well as cases of low *CD30*<sup>+</sup> coloring.

© 2013 Elsevier Ltd. All rights reserved.

### 1. Introduction

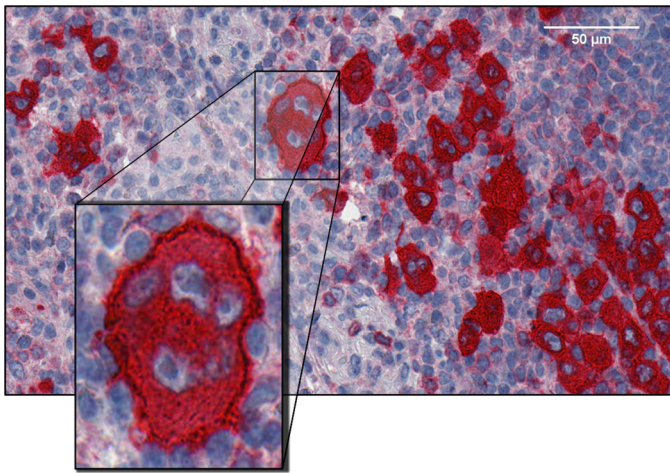
Hodgkin lymphoma (HL) is a type of lymphoma, originating from germinal center B-cells in most of the cases (Küppers et al., 1994). In HL, neoplastic cells circumvent immune surveillance and apoptotic control. Characteristically, in HL, the malignant cells make up only about 1% of the tumor mass, while being outnumbered by a microenvironment of reactive lymphoid cells (Steidl et al., 2011; Mathas et al., 2009). According to the WHO classification, two main subtypes can be described, the *classical HL (cHL)* and the *nodular lymphocyte-predominant HL (NLP HL)*.

The assignment to cHL is based on the identification of characteristic tumor cells termed *Hodgkin cells* and *Reed–Sternberg cells* (HRS cells) (Reed, 1902; Sternberg, 1898). Using immunohistochemical methods for immunophenotypic differentiation of HL, the *cluster-of-differentiation* (CD) protocols are of particular interest. They are used to identify cell surface proteins presented on leukocytes, e.g., CD30, CD20, and CD15. For example, CD30 belongs to the tumor necrosis receptor family (TNFR), and is usually expressed by HRS cells but can also be expressed by activated blasts in reactive lesions like lymphadenitis.

Fig. 1 illustrates a typical CD30<sup>+</sup> immunostained HRS cell within its microenvironment. Hematoxylin counterstaining is applied to visualize cell nuclei. Depending on immunophenotypic findings and the composition of the cellular microenvironment cHL is further divided into two subtypes, the most common one termed

\* Corresponding author.

<sup>1</sup> These authors contributed equally.



**Fig. 1.** A typical CD30<sup>+</sup> Reed-Sternberg cell. Malignant cells are marked with CD30 immunostain. Additional hematoxylin counterstain is provided to visualize nuclei. A giant RS cell can be seen in the center of the image outnumbered by a variety of other lymphoid cells with visible nuclei. Multiple irregularly shaped nuclei are visible within the RS cell, a distinguishing feature of this cell type.

*nodular sclerosis* (NS cHL) and the mixed cellularity (MC cHL) (Hansmann and Willenbrock, 2002). All subtypes of cHL share an immunophenotype of CD30<sup>+</sup> (Küppers, 2009). CD15 and CD20 can be variably expressed, with usual CD20-negativity and CD15-positivity. However, CD30 expression can also be observed in reactive lesions. Therefore a detailed investigation of the location of CD30 blasts is mandatory and image analysis may provide systematic information for this task.

Moreover, the expression of various cytokines and chemokines is thought to be responsible for the attraction of different cells, such as small lymphocytes, macrophages, mast cells, plasma cells, stromal cells, histiocytes, or fibroblasts (Aldinucci et al., 2010; Steidl et al., 2011). The extent of cellular infiltration, involving many different cell types as a response to inflammatory signals, represents a key characteristic of HL (Hartmann et al., 2013). The events that happen in the course of malignancy development as well as the complex interaction network of malignant cells in HL are not understood. Here, image analysis may provide new insights into the pathology of HL.

There is a broad range of literature on image analysis of non-lymphoid tumors. For example, Naik et al. (2008) used a Bayesian nuclear segmentation scheme in prostate and breast cancer histopathology. Bayesian models with additional contextual information from Markov random fields have been applied as a probabilistic approach for prostate cancer detection (Monaco et al., 2009). Lei et al. (2011) applied Gaussian mixture models to perform local and global clustering, extracting different tissue constituents in cervix histology images. Sertel et al. (2009) developed a system for segmentation of eosinophilic and basophilic structures in hematoxylin and eosin (H&E) stained tissue section images of neuroblastoma. Large-scale computations and machine learning algorithms have been applied to breast histology images (Petushi et al., 2006). Karacali and Tözeren (2007) proposed a high-throughput method of texture heterogeneity on breast tissue images to identify regions of interest. An automatic classification of three types of malignant lymphomas has been proposed by Orlov et al. (2010), taking into account chronic lymphocytic leukemia, follicular lymphoma, and mantle cell lymphoma. They achieved high accuracy in discriminating the different lymphoma types using various statistical and texture features. Sertel et al. developed a computer-aided method to detect follicles based on texture features in an immunostained image followed by the detection of centroblasts within the follicular regions (Sertel et al., 2008).

Today, a number of software solutions for analyzing images are available. Commercial software includes general tools like MATLAB (Matlab, 2013) as well as software geared especially toward digital pathology like Aperio ImageScope (Aperio, 2013), Definiens Tissue Studio (Definiens), and many other. There exist also freely available tools, the most prominent of which are CellProfiler (Lamprecht et al., 2007), ImageJ (Abramoff et al., 2004), GNU Octave (Eaton et al., 2008), and Scilab (Enterprises and Consortium, 2013). Despite all these approaches, no special work exists for the analysis of HL images that takes into account the specific characteristics of that lymphoma type. That is important, because NS cHL and MC cHL represent two different types of HL which often have distinct patterns of tumor cell distribution. In NS cHL, the tumor cells grow in a nodular pattern with broad collagen bands separating the nodules. In contrast, MC cHL tumor cells are more isolated and often distributed over the whole tissue section. NS cHL accounts for about 64%, whereas MC cHL is found in about 30% of the cHL cases in our database. While the visual differentiation between the two types of cHL is possible in many cases, non-lymphoma NL images sometimes exhibit features similar to cHL subtypes as for example, lymphadenitis which is an infection of the lymph nodes, resulting from certain bacterial or viral infections and often leading to inflammation and swelling of lymph nodes.

Our work aims to analyze the considered types based on CD30 stained tissue sections, using pixel based image processing techniques. To be flexible in handling and exploring HL whole slide tissue-images and to easily extend the software in future projects we implemented a new pipeline. Our software is intended to evolve into a tool that extracts statistical information from a large database of Hodgkin lymphoma images. We aim to combine our software solutions with existing imaging software solutions like CellProfiler (Lamprecht et al., 2007).

## 2. Material

The image sources used in this study are CD30 stained tissue slides of lymph node sections. The tissue slides are pretreated, and immunostained for CD30 as previously described (Hartmann et al., 2011). The reddish coloring is basically produced by the enzyme alkaline phosphatase that is bound via a biotin-streptavidin linker to a CD30 antibody and hydrolyzes the magenta dye new fuchsin. Hematoxylin counterstain was applied to stain cell nuclei. The digital slides are captured using an Aperio ScanScope XT scanning device with a 40× objective lens. Typical dimensions of tissue samples are on the average 15 mm<sup>2</sup>, and the images reach resolutions of up to 100,000 × 100,000 pixels (0.23 μm per pixel). The images are recorded in the specific Aperio SVS format, a single-file pyramidal TIFF with non-standard metadata and compression. Each picture in SVS format stores the digitized images in high resolution (0.23 μm per pixel) and simultaneously in various levels of lower resolutions. If not otherwise mentioned, we use the second level that is down-scaled by a factor of 16 (3.68 μm per pixel). The slides were manually assigned by pathologists to NS cHL (nodular sclerosis classical Hodgkin lymphoma), MC cHL (mixed cellularity classical Hodgkin lymphoma), and NL (non-lymphoma, lymphadenitis with and without follicular hyperplasia). Samples are randomly selected by pathologists who eliminated untypical cases from an anonymized data set from the Senckenberg Institute of Pathology Frankfurt am Main. Tissue sections of small size are ignored, because they are not representative for the complete lymph node. Furthermore, only tissue sections with distinct immunostaining are considered. In the selected images, CD30<sup>+</sup> cells can be visually distinguished from the background. The selection yields a representative set of 62 images of NS cHL, 57 images of MC cHL, and 51 images of NL.

Download English Version:

<https://daneshyari.com/en/article/6487187>

Download Persian Version:

<https://daneshyari.com/article/6487187>

[Daneshyari.com](https://daneshyari.com)