PP082-S

Automated Detection, Quantification and Characterization of Osteoclasts in Cultures Using a Combined Image-Processing and Machine-Learning Strategy

A. Heindl[1,2], M. Schepelmann[1], R. Stumberger[1], A. Nussbaumer[1,3], P. Pietschmann[1], G. Bises[1], R. Rogojanu[1,4], T. Thalhammer[1], A. Seewald[2] and I. Ellinger[1]

[1]Department of Pathophysiology and Allergy Research, Center for Pathophysiology, Immunology and Infectiology, Medical University of Vienna, Austria; [2]Seewald Solutions, Vienna, Austria; [3]Institute of Specific Prophylaxis and Tropical Medicine, Center for Pathophysiology, Immunology and Infectiology, Medical University of Vienna; [4]TissueGnostics GmbH, Vienna, Austria

Introduction

Osteoclast (OC) cultures are used to study the effect of drugs and xenobotics on OC growth and function. Currently, human experts manually quantify OC in these cultures by counting tartrate-resistant acid phosphatase (TRAP) positive multi-nucleated cells. This time-consuming evaluation prevents the conduction of large-scale experiments. In addition, neither the amount of precursor cells in the culture nor cellular parameters like area or the number of nuclei can be determined. A major disadvantage of this method is a lack of combination with immunohistochemical or immunofluorescence staining methods in order to quantify protein-expression in OC identified by TRAP staining.

Applying image processing and machine-learning techniques to images derived from such biological samples can dramatically improve the speed, quality and reproducibility of cell identification. Automated analysis produces consistent quantitative measures of cell-associated parameters [1]. Using machine-learning techniques, tacit/intuitive knowledge, which can be applied but not explained by human experts (see also Algorithm B) might be incorporated in algorithms. Finally, an automated analysis is more efficient than a human and can operate 24 hours and 7 days a week.

Aim

Development and validation of a versatile image analysis system composed of several algorithms (A, B,..) that is capable of automated detection, quantification and characterization of cultured multinucleated mouse OC and their precursor cells on images derived from immunofluorescence microscopy.

Materials and Methods

Mouse OC in vitro culture and immunofluorescence staining protocol are given in detail on the corresponding Poster PP112-S (Schepelmann et al.). Table 1 summarizes the applied antibodies.

Primary antibody	Secondary antibody	Application
Anti calcitonin receptor	Alexa Fluor 750	Cell segmentation by Algorithm A
Anti ! -tubulin	Alexa Fluor 750	
Anti F4/80 macrophage marker	Alexa Fluor 568	OC identification by Algorithm A
DAPI (nucleic acid dye)		OC identification by Algorithm A and B

Table 1: Antibodies used in the immunofluorescence staining protocol. The anti-calcitonin receptor antibody and the anti-alpha-tubulin antibody are used to make all cells visible and detectable for algorithm 1. Lack of OC precursor-cell specific F4/80 macrophage marker is used for OC identification by algorithm A. DAPI-stained nuclei are identified by algorithm A and B.

Image (field of view, FOV) acquisition was done with an automated Axio Imager epifluorescence microscope (Zeiss) equipped with TissueFAXS hard- and software (TissueGnostics GmbH) using a 40x (oil) objective. The acquired FOVs were aligned and stitched to form one big image [2] for each fluorescence channel.

Validation of algorithms A and B was done by comparison of the output-masks to OC- or nuclei-markups done by human experts. Two human experts independently marked OCs in seven regions (90 FOVs) for algorithm A or marked and classified (OC/precursor cell) ~3500 nuclei for algorithm B. The results of the automated comparisons of the masks generated by algorithm A and B to the markups are show in Figure 5, respectively Figure 6. An extensive survey concerning the validation process can be found in [3].

Results

So far, two algorithms (A and B) have been developed for the image-processing system. Currently, they work independently. A combination of the algorithms to optimize performance is possible.

Algorithm A

Algorithm A employs image-processing techniques on cell cultures labeled by an immunofluorescence staining protocol (see Table 1). OC identification (green arrows in Figure 1A) is based on two biological criteria: OC are multinucleated cells (nuclei>=3) and F4/80 macrophage marker expression is low in OC.

The main steps of the algorithm depicted in Figure 2 are: The image processing (1), preparing the image for further analysis. Classification (2) of all identified cells in OC and precursor cell according to three criteria computed in the algorithm. Result (3) generation of output cell mask (Fiugre 1B) and subsequent measurements such as cell area, mean/median and standard deviation of staining intensity. The obtained data can be further processed in any spreadsheet program.

An evaluation of algorithm A is shown in Figure 5. Manual identification of OCs in various regions exhibits inter human variabilities. The green bars show the amount of cells that were classified as OC by both experts. The mean detection rate of these cells in all seven regions was 86%. A first biological application of algorithm A can be found on Poster PP112-S (Schepelmann et al). The effect of the pineal hormone melatonin on OC growth and expression of Mel-1A receptor in OC was quantified.

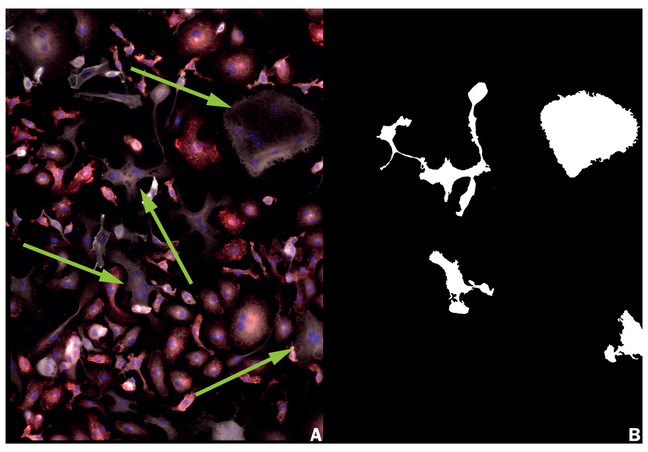

Figure 1: Image 1A is a subregion (2x2 FOVs) of a mouse bone marrow culture. All cells are stained with anti-alpha-tubulin and anti-calcitonin receptor antibodies (white). Nuclei are stained with DAPI (blue). F4/80 macrophage marker (red) predominantly labels OC precursor cells. OC are indicated with green arrows. Image 1B is a binary image output mask of algorithm A where detected OC are indicated as white spots. This mask is later used to quantify cell-associated features (expression of target proteins, area, ...)

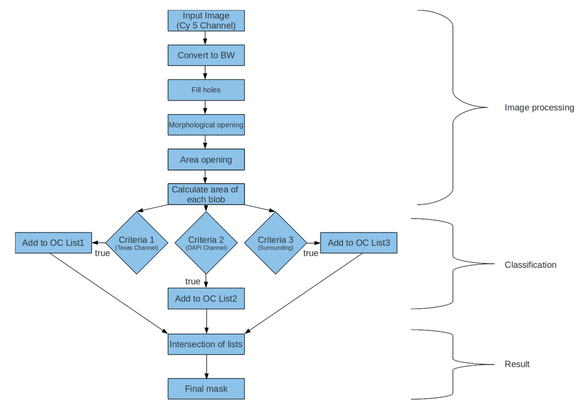

Figure 2: This flowchart illustrates Algorithm A. It contains three parts. The image processing steps that prepare the image for further processing, and eliminate small artifacts and disturbances. A classification step that measures three criteria unique to OC. Finally the result that leads to a binary mask (Figure 1B) used to compute various statistical moments.

Algorithm B

This cell classification is based on the intuitive, not clearly specified human observation that OC nuclei and precursor cell nuclei appear “somehow” different. Algorithm B applies machine learning to classify OCs and precursors by parameters of their nuclei shape to model the intuitive human knowledge.

The main steps of the algorithm B are the following: 1/ Two human experts marked up ~3500 nuclei in stitched images of the DAPI channel by drawing perimeters (white line in Figure 3). In the cellular context, the experts also classified OCs and precursor cells as well as their nuclei (i.e. Ground truth data in Figure 6). The perimeters were analyzed by a machine-learning system. To map the free-hand perimeters to mathematically easier describable shapes, we employed ellipse fitting (orange line in Figure 3). 2/ Various features (n=9) were computed from these ellipses such as mean of intensity, standard deviation of intensity, area, eccentricity, ... No single parameter was able to differentiate OC and precursor nuclei. However, a comparison of all parameters illustrated by a sammon-mapping [4], a procedure that maps this high dimensional space into a low dimensional space, reveals that indeed two different clusters of nuclei exist (Figure 4).

An evaluation of the OC nuclei classification-performance is shown in Figure 6. Ground truth data were obtained from human expert markups and classification of nuclei in the cellular context. The algorithm achieved an accuracy of ~75% correctly classified nuclei. When the experts tried to classify nuclei without cellular context, the best expert only got ~55% percent correct.

Figure 3: In this image, white represents the human expert markup (perimeter) of a nucleus. As this freehand drawn line is hardly describable in mathematical terms, an orange ellipse is fitted that estimates this outline. Based on this well-defined mathematical shape, various features can be computed within reasonable time.

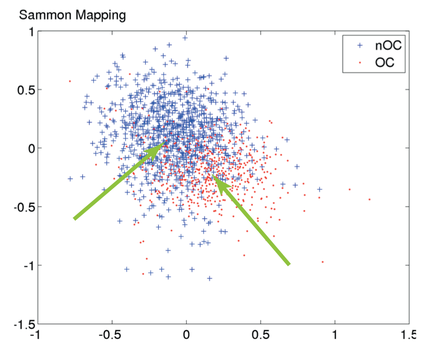

Figure 4: Sammon mapping projects the high dimensional space to a space of lower dimensionality by trying to preserve the structure of inter-point distances in high dimensional space in the lower dimensional space. A 2D plot of all computed features of the ellipses fitted to OC and precursor nuclei is shown in the image. The blue crosses represent nuclei from OC precursors whereas the red dots represent OC nuclei. Green arrows indicate the centers of the two clusters.

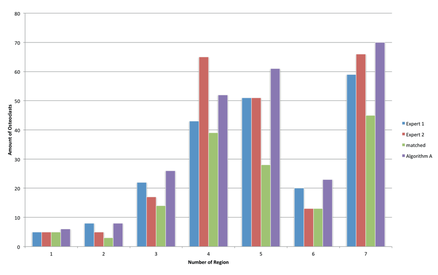

Figure 5: Evaluation of Algorithm A. The blue bars show the amount of OC identified by expert 1 respectively red those by expert 2. Green represents the amount of cells that were classified as OC by both experts. The purple bar is the amount of OC detected by algorithm A.

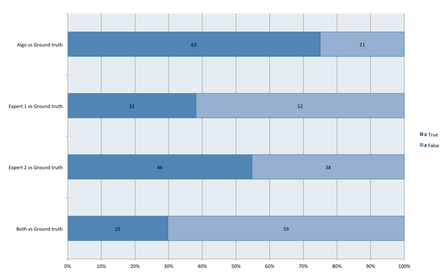

Figure 6: Evaluation of Algorithm B. The algorithm achieved an accuracy of ~75% correctly classified nuclei compared to ground truth. Without seeing a cellular context, the best expert classified only ~55% percent of nuclei correct.

Literature

[1] Murphy R. F., Boland M. V., Velliste M. (2000): Towards a Systematics for Protein Subcellular Location: Quantitative Description of Protein Localization Patterns and Automated Analysis of Fluorescence Microscope Images. ISMB-00 Proceedings, AAAI, 2000.

[2] Heindl A., Ecker R., Steiner G., Bises G., Ellinger I., Thalhammer T., Fuchs R., Uhrova H., Seewald A.: Automated cell-detection technologies for science and diagnostics. Poster presentation at the 5th PhD Symposium of the Young Scientist Association of the Medical University of Vienna, June 2009, Vienna, Austria.

[3] Heindl A., Schepelmann M., Ecker R., Pietschmann P., Ellinger I., Seewald A., Thalhammer T. (2011): Towards the automated detection and characterization of osteoclasts in microscopic images. Book chapter in Principles of Osteoimmunology: Molecular Mechanisms and Clinical Applications, Ed. P.Pietschmann, Wien. Springer-Verlag 2011.

[4] Sammon J.W., Jr. (1969): A Nonlinear Mapping for Data Structure Analysis. Transaction on Computers, Issue 5, Pages 401-409, ISSN: 0018-9340.

Summary

We have developed two algorithms (A and B) for automated identification of OC and precursor cells in cultures. These algorithms achieve high accuracies when validated against human experts (86% and 75%, respectively) and might be combined to further increase performance. First biological applications are also presented (PP112-S).

Acknowledgement

Supported by the FFG (Bridge-Project [818094])