IMAGE ANALYSIS OF NUCLEAR ENVELOPE BREAKDOWN EVENTS USING KNIME

Thorsten Rieß[1], Joseph Marino[2], Cornelia Wandke[2], Dorit Merhof [1], Oliver Deussen[1],
Gábor Csúcs[3], Ulrike Kutay[2], Péter Horváth[3]

[1]INCIDE, University of Konstanz, Germany
[2]Institute of Biochemistry, ETH Zürich, Switzerland
[3]Light Microscopy and Screening Centre, ETH Zürich, Switzerland
{thorsten.riess, dorit.merhof, oliver.deussen}@uni-konstanz.de, {jomarino, cwandke, csucsga, ulkutay, phorvath}@ethz.ch

ABSTRACT

Insights into the molecular mechanism underlying nuclear envelope breakdown (NEBD) are crucial to understand dynamic changes of the nuclear envelope (NE) that occur at mitotic entry of eukaryotic cells. In this paper, we present an image processing algorithm and its implementation in KNIME, which allows to automatically quantify image data obtained from the in vitro NEBD assay. The algorithm consists of image alignment via phase correlation, nucleus identification using adaptive thresholding, morphological operations, and intensity measurements. The results of these measurements are verified using images from a known nuclear envelope disassembly assay.

1. INTRODUCTION

Higher eukaryotic cells undergo an open mitosis which is characterized by the breakdown of the nuclear envelope (NE). The resulting loss of nucleo-cytoplasmic compartmentalization allows the formation of the mitotic spindle, which is essential for chromosome segregation during cell division [1]. In order to visualize morphological changes in the NE during nuclear disassembly at mitotic onset, we have previously established an in vitro nuclear disassembly assay [2]. In this assay semi-permeabilised HeLa cells expressing a fluorescently-tagged NE protein are incubated with mitotic HeLa cell extracts to induce NEBD. The gradual loss of the nuclear permeability barrier is followed by tracking fluorescent dextran influx into the nucleus using time-lapse confocal microscopy. At the same time, loss of the GFP-tagged NE marker from the nuclear rim is followed. This results in a (time-)sequence of microscopic images showing the nuclei as the envelope breakdown occurs. Typically, the nuclei are clearly visible as dark regions in the first image of the sequence, and are then subsequently flooded with the fluorescent marker in the course of the experiment.

From the image processing point of view, the problem is challenging for two main reasons: it is very hard to visibly detect the cell nuclei once the envelope broke down, and the cells move during the experiment. The movement of the cells is either due to the individual cell movement (which is, however, very little and thus negligible), or due

to the limited motorized microscopic stage location reproducibility during multi-location image acquisition process, which usually results in significant paraxial shifts.

We propose an algorithm and its implementation using KNIME [3] which solves the above mentioned problems and enables automated analysis of large data sets. The algorithm consists of the following steps: spatial image alignment, adaptive thresholding of the aligned image sequence, object identification, intensity measurements, and report generation. To our knowledge, there is no comparable approach published; alternative approaches using cell tracking instead of image alignment are almost certainly bound to fail due to the drastic change of the cell appearance in the course of the experiment.

The paper is organized as follows: Section 2 describes the biological background, the image processing pipeline is presented in Section 3. In Section 4 the implementation of the algorithm is presented, and in Section 5 results for some applications are shown. Finally, Section 6 gives a summary and outlook.

2. BIOLOGICAL BACKGROUND: NUCLEAR ENVELOPE BREAKDOWN

The genetic material of eukaryotic cells is enclosed by a double lipid bilayer termed NE. The NE consists of an outer and an inner nuclear membrane (ONM and INM, respectively) that are fused at sites of nuclear pores. The ONM is continuous with the endoplasmic reticulum (ER) and shares similar characteristics. In contrast, the INM contains a unique set of integral membrane proteins that are linked to chromatin and the nuclear lamina. ONM and INM are fused at numerous sites where nuclear pore complexes (NPCs) are inserted into the NE to allow for exchange of material between nucleus and cytoplasm. NPCs are composed of about 30 nucleoporins (nups) that together form assemblies of octagonal rotational symmetry mediating nucleo-cytoplasmic transport of macromolecules [1]. The molecular features of NPCs determine both the barrier and permeability characteristics of the NE.

Open mitosis requires nuclear envelope breakdown (NEBD), which is accompanied by a gradual loss of the NE permeability barrier. Besides other mitotic kinases, the cyclin-dependent kinase 1 (CDK1) plays a key role in the

nuclear disassembly process. Furthermore, the disintegration of the NE membrane involves the dispersal of nups into the cytoplasm, the depolymerisation of the lamina and the absorption of NE membranes and their INM proteins into the ER [1]. When chromosome segregation is completed, the NE reforms around the decondensing chromatin masses of each daughter cell thereby re-establishing the NE permeability barrier [4]. The formerly established in vitro NE disassembly assay has proven to be a powerful tool to decipher the molecular requirements of NEBD [2, 5]. However, image processing of the data is still not fully automated and thus a time-consuming task. The KNIME-based application described here simplifies and accelerates the process of quantification of in vitro NEBD.

3. METHOD

The image processing algorithm consists of three steps. In the first step, the image sequence is aligned to account for the paraxial movement of the microscope. This registration is based on phase correlation [6] and it is crucial for the following steps. The paraxial nature of the alignment problem reduces the complexity to the detection of translation between consecutive frames. Let I_j and I_{j+1} denote two consecutive frames. For the phase correlation image of I_j and I_{j+1}, the Fourier transforms $\mathcal{F}\{I_j\}$ and $\mathcal{F}\{I_{j+1}\}$ are computed. Then the phase correlation image P is the inverse Fourier transform of the pixel-wise multiplication of $\mathcal{F}\{I_j\}$ and $\mathcal{F}\{I_{j+1}\}^c$. The position of the maximum value in the phase correlation image denotes the shift that is required from frame I_j to I_{j+1} that minimizes the difference between the frames. Note that it is also possible to only compute the phase correlation image of the first frame versus the remaining frames, but due to the significant change of the nucleus appearance in the course of the experiment, the use of consecutive frames is more robust. The second step is the segmentation of the cell nuclei in the first image of the sequence. At the beginning of the experiment, the cell nuclei are fully visible as dark areas (see Figure 2 upper right image), but after the NEBD event they cannot clearly be distinguished from the background. The segmentation in the first image is performed using the Otsu adaptive thresholding method [7]. Touching nuclei are separated with morphological operations, nuclei that cannot be separated are excluded by filtering out segments that are too large. To avoid measurement errors – caused by individual cell movements – the cell masks are reduced to a disk with significantly smaller size. Finally, in the third step, the location of the segmented nuclei is used to measure the mean intensity of each individual nucleus in each image of the registered sequence (note that nuclei that move too much or that have dextran influx already at the beginning of the experiment need to be excluded manually from the measurements). This intensity is divided by the mean intensity of the background to achieve an intensity ratio r (typically, $r \in [0,1]$), where $r = 0$ means that the nuclear envelope is intact and no dextran has entered, and $r = 1$ means that the background intensity is identical to the intensity inside

the nucleus. Nuclei displaying an intensity ratio $r = 0.3$ are defined to be dextran positive (this threshold is used in [5]). Algorithm 1 shows the detailed image processing pipeline in pseudo-code.

Algorithm 1 The image processing pipeline.

Require: Image sequence of N images:
$\quad \mathcal{I} = \{I_j\}_{j \in \mathcal{N}}, \mathcal{N} = \{1, 2, \ldots, N\}.$
Ensure: Ratio functions for the M observed nuclei:
$\quad r_k : \mathcal{N} \to [0,1], k = 1, \ldots, M.$

$\quad$**for** $j = 2 \to N$ **do**
$\quad\quad P \leftarrow$ phase correlation image of I_{j-1} and I_j.
$\quad\quad (\hat{x}_j, \hat{y}_j) \leftarrow \text{argmax}_{(x,y)} P(x, y).$
$\quad\quad I_j \leftarrow$ shift I_j by $\sum_{k=2}^{j}(\hat{x}_k, \hat{y}_k).$
$\quad$**end for**
$\quad \mathcal{I} \leftarrow$ cropped image sequence $\mathcal{I}$.
$\quad T \leftarrow$ thresholded binary image I_1.
$\quad \mathcal{C} \leftarrow$ connected components of T.
$\quad \mathcal{C} \leftarrow$ eroded components in $\mathcal{C}$.
$\quad M \leftarrow |\mathcal{C}|.$ $\{\mathcal{C} = \{C_1, \ldots, C_M\}\}$
$\quad B \leftarrow$ eroded T^c {background}
$\quad$**for** $j = 1 \to N$ **do**
$\quad\quad \{I|J$ means: image I restricted to the ROI $J\}$
$\quad\quad b \leftarrow$ mean intensity of $I_j|B$.
$\quad\quad$**for** $k = 1 \to M$ **do**
$\quad\quad\quad i \leftarrow$ mean intensity of $I_j|C_k$.
$\quad\quad\quad r_k(j) \leftarrow i/b$.
$\quad\quad$**end for**
$\quad$**end for**

---

4. IMPLEMENTATION IN KNIME

KNIME [3] is a data mining and integration platform developed at the University of Konstanz. It is freely available for download at http://www.knime.org and runs on the major platforms. We choose KNIME as the main implementation framework because it is flexible, has an intuitive graphical user interface and has recently been enriched with a powerful image processing plug-in, which provides basic image processing algorithms that can be readily used. The concept of KNIME is that data is processed by a workflow consisting of nodes and connections between nodes. Each node can transform, input, output or visualize data. The intrinsic data structure that is passed from one node to another is a simple table structure, where each cell represents a value of a certain data type. The image processing plug-in is not part of the standard KNIME distribution, but can be downloaded at http://tech.knime.org. For the presented implementation KNIME version 2.5.1 was used.

The image processing pipeline presented in Algorithm 1 can readily be modeled as a KNIME-workflow using nodes from the image processing plug-in. After reading in the image sequence, the alignment-node aligns and crops the images according to the first part of the algorithm. Subsequent nodes handle the segmentation of the cell nuclei and background in the first image of the sequence and
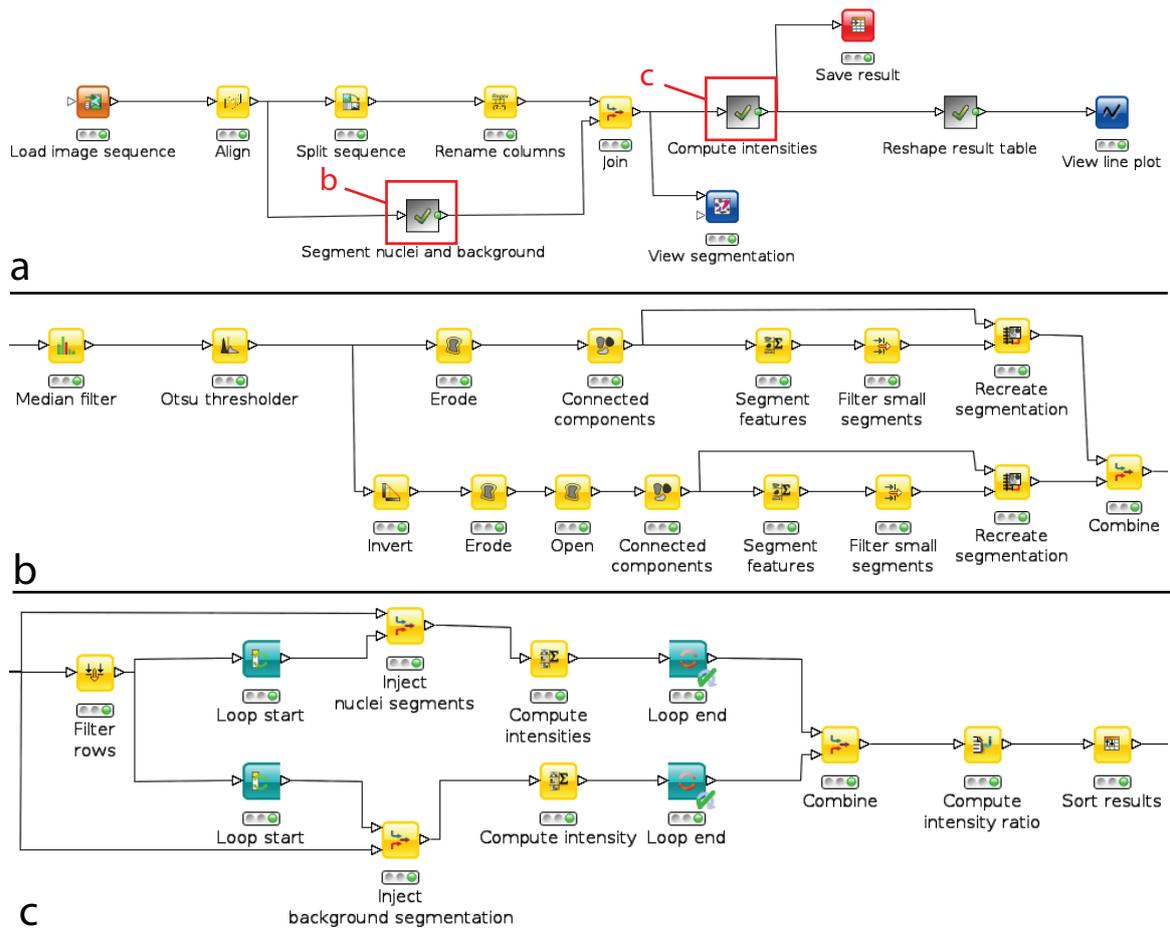
Figure 1. (a) the main KNIME-workflow. The segmentation of the nuclei and background as well as the intensity computation and the postprocessing are wrapped into sub-workflows to keep the main workflow as simple and clear as possible. (b) sub-workflow for the segmentation of the nuclei and the background in the first image of the sequence. (c) sub-workflow for the intensity measurements in each image of the sequence.

the transformation of this segmentation using standard binary morphological operations. Finally, a loop iterates through all images in the sequence and the mean intensities of the segmented nuclei and the background are computed for each image. The result is a table holding the ratio of these intensities for each nucleus at each image in the sequence. Standard KNIME nodes can be used to either save the result table or to visualize the result as a line plot. Figure 1 a. shows the main workflow as a skeleton for the registration, segmentation and measurement steps. The actual segmentation and measurement steps are implemented as sub-workflows shown in Figure 1 b. and c.

5. RESULTS

We applied the image processing pipeline described in Section 3 on real image data. In a typical experimental setup, the analyzed image sequences consist of 15 images taken in every 4 minutes between 16 and 68 minutes after addition of mitotic cell extracts to semi-permeabilized HeLa cells stably expressing GFP-Nup58. These image sequences are acquired at 4 different locations (each show-

ing 10-30 nuclei), and the whole experiment is repeated 3 times. The individual images usually have two channels, one for the fluorescent dextran and the other for the GFP-tagged nucleoporin. This results in 12 image sequences for each experimental condition. Due to slight shifts of the microscope alignment, the individual image sequences are usually shifted paraxially. (If the microscope delivers 3D image stacks, the best focal plain is manually selected for the computational analysis.)

The timing of NEBD is compared in cells incubated with either mitotic extracts (ME) or with extracts supplemented with the CDK1 inhibitor alsterpaullone (ME + AP). Figure 2 left shows ME images at 16, 36 and 56 minutes.

Paraxial microscope stage shifts are detected using the presented phase-correlation algorithm and images are cropped and aligned accordingly. On Figure 2 left red dashed rectangles show the detected common parts of the images. On the first image of every sequence cell nuclei are segmented. Results are shown in Figure 2 right. The in-

155 kDa TRITC Dextran

t=16m

t=36m

t=56m
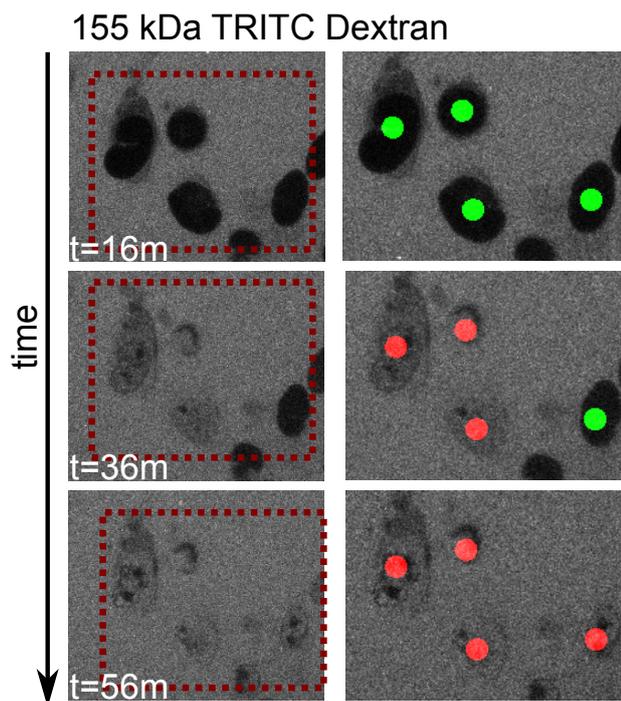
time

Figure 2. The influx of TRITC-labelled 155 kDa dextran into nuclei is shown for three different time points (left) and segmented nuclei (right). The discs overlayed on the right show the area of the actual intensity measurement and are green for intact nulcei and red for nuclei classified as dextran positive (intensity ratio larger than 0.3).

tensity measurements are taken using a disk mask significantly smaller than the nuclear size placed on the centroid of segmented regions. Figure 3 shows the intensity ratio measurements over time for the two experimental conditions analyzing image series taken in 3 independent experiments. It can be observed that CDK1 inhibition by alsterpaullone strongly delays the time that is required for 50% nuclei to achieve NEBD.

6. DISCUSSION

In this paper we present an image processing pipeline for the automated image analysis of NEBD events. The implementation of the pipeline is based on the image processing plug-in of the data analysis framework KNIME, which allows for a very flexible prototype workflow and can readily be used by biologists. The resulting intensity ratios can either be saved and used externally, or they can be further processed or visualized using the tools already built into KNIME. Moreover, the workflow-based approach allows batch-processing of huge amounts of experimental data, also in parallel. The presented tools allow the biologist to use the workflow as-is, but also allows tuning all parameters such that the workflow can be used for a wide range of experiments and does not require a computer science expert to adjust to different image data. The measurement results allow an accurate estimation of the nuclear breakdown events for individual nuclei, which is
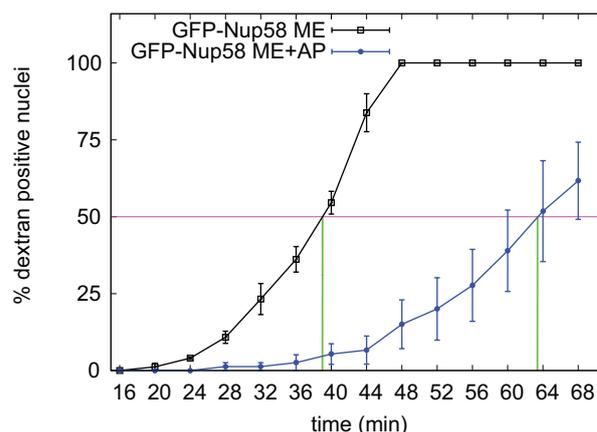
Figure 3. Quantification of dextran positive nuclei over time. The purple line marks the time point at which 50% of the nuclei are dextran positive. Error bars represent the standard error of the mean.

neither biased nor subjective as a manual evaluation would be.

7. REFERENCES

[1] S. Guttinger, E. Laurell, and U. Kutay, "Orchestrating nuclear envelope disassembly and reassembly during mitosis," *Nat Rev Mol Cell Biol*, vol. 10, no. 3, pp. 178–91, 2009.

[2] P. Muhlhausser and U. Kutay, "An in vitro nuclear disassembly system reveals a role for the RanGTPase system and microtubule-dependent steps in nuclear envelope breakdown," *J Cell Biol*, vol. 178, no. 4, pp. 595–610, 2007.

[3] M. Berthold, N. Cebron, F. Dill, T. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, and B. Wiswedel, "KNIME: The Konstanz Information Miner," in *Proc. Data Analysis, Machine Learning and Applications*, 2008, pp. 319–326.

[4] B. Burke and J. Ellenberg, "Remodelling the walls of the nucleus," *Nat Rev Mol Cell Biol*, vol. 3, no. 7, pp. 487–97, 2002.

[5] E. Laurell, K. Beck, K. Krupina, G. Theerthagiri, B. Bodenmiller, P. Horvath, R. Aebersold, W. Antonin, and U. Kutay, "Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry," *Cell*, vol. 144, no. 4, pp. 539–50, 2011.

[6] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite fourier transforms.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 700–703, 1987.

[7] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.