Automatic Analysis of Live Cell Image Sequences to determine Temporal Mitotic Phenotypes

Nathalie Harder[1], Felipe Mora-Bermúdez[2], William J. Godinez[1], Annelie Wünsche[2], Jan Ellenberg[2], Roland Eils[1], Karl Rohr[1]

[1]University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg, Dept. Bioinformatics and Functional Genomics, Biomedical Computer Vision Group
[2]European Molecular Biology Laboratory (EMBL) Heidelberg, Gene Expression and Cell Biology/Biophysics Programmes
n.harder@dkfz-heidelberg.de

Abstract. Automated screening platforms allow biologists to acquire large amounts of image data with high information content. However, reliable automatic methods for analyzing this data are often not available. Here, we present an approach for detailed cell cycle analysis based on live cell fluorescence microscopy image sequences. Our approach comprises segmentation and tracking of dividing cell nuclei, and classifies cells into seven cell cycle phases as well as five abnormal morphological phenotypes. Moreover, we automatically quantify cell cycle phase durations and perform a statistical analysis to determine temporal phenotypes. Our approach was successfully applied to images from gene knockdown experiments and experiments treated with small molecule drugs.

1 Introduction

Understanding gene regulation of the cell cycle is of high common interest since errors in this process may lead to serious diseases such as cancer. High-content image-based screening is a powerful technology for gene function studies, and comprises automated microscopy as well as computational analysis to automatically extract information in an unbiased way. In screening experiments for cell cycle analysis usually live cell images of multiple cells are acquired. Multi-cell image sequences can be either analyzed in a population-based manner, i.e. features are determined for all cells and changes are studied for the whole population over time. However, subtle effects, such as cell cycle phase prolongations of certain cells, cannot be detected in this way. Therefore, single cell-based analysis has to be performed, which requires to track the individual cells throughout an image sequence. Based on tracking, the temporal evolution of single cells can be investigated, in particular, to study cell cycle phase progression. Previously, this has been done based on phase contrast ([1]) and fluorescence ([2, 3]) microscopy images. However, these studies distinguished only up to five cell cycle phases and did not consider abnormal cellular morphologies. Also, none of these studies determined cell cycle phase durations which is an essential readout for gene

function studies. In [4] seven normal phases were distinguished, but abnormal morphologies were not considered. Here, we present an approach to automatically determine morphological and temporal phenotypes on a single cell basis.

2 Materials and Methods

In this work, we analyze fluorescence microscopy 3D image sequences with three slices per time step, including multiple cell nuclei (Fig. 1, left). We developed the following image analysis approach: First, segmentation and tracking is performed based on maximum intensity projection (MIP) images. Next, image features are extracted based on original images and MIP images. Then, we perform classification of each nucleus at each time step, and finally process the obtained phase sequences and quantify phase durations using a finite state machine.

2.1 Segmentation and Tracking

For segmentation we developed a region-adaptive thresholding approach based on Otsu's method. Local thresholds are computed in overlapping image regions and are applied to their non-overlapping center regions. For normal cell nuclei this method yields accurate results, however, for accurate segmentation of abnormal morphologies additional processing steps are necessary. In our application, abnormal morphologies are characterized by dim micronuclei attached to a bright main nucleus and small detached chromosomes. Proper segmentation of these structures is important to enable correct phenotype classification. For segmentation of dim micronuclei attached to a bright normal nucleus we apply the segmentation algorithm twice, where in the second run the bright normal nuclei are masked out. To merge single detached chromosomes to their corresponding main nucleus our approach automatically connects small objects (smaller than the smallest regular nuclei in anaphase) to the closest regular nucleus in their neighborhood by inserting a connecting line (Fig. 1, middle) .

For tracking dividing cells we developed a two-step approach. First, initial trajectories (*one-to-one* correspondences) are determined using a feature point tracking algorithm [5]. Second, mitosis events are detected and the respective trajectories are merged, establishing *one-to-many* correspondences. The detection of a mitosis event requires that (1) the daughter nuclei are smaller than the average nucleus, and (2) the Euclidean distance between daughter cells is below a threshold. If (1) and (2) are fulfilled, a measure for the likelihood of a mitosis event is computed, which is composed of three terms: The first term includes the ratio of the mean nucleus intensity and the intensities of the potential daughter nuclei, and yields values close to one if the daughter nuclei are very bright compared to the mean nucleus intensity. The second term considers the ratio of mother nucleus size and mean nucleus size, and yields values close to one if the mother nucleus is smaller than the average. The third term takes into account the analogous size ratio for the daughter nuclei, and consequently is close to one if both daughters are smaller than the average. The terms can be

individually weighted and sum up to a maximum value of one. If the likelihood measure is sufficiently high, the respective case is considered as mitosis (Fig. 1, right). Note that the number of possible daughter cells is not restricted to two, and thus, also abnormal splits into more than two daughter cells can be tracked.

2.2 Feature Extraction and Classification

We compute features for each cell nucleus based on the MIP images and based on single image slices of the original 3D images. In the latter case, the most informative slice is selected for each nucleus based on maximum total intensity. The reason for using the original image slices is that fine textures, which are important for the classification of certain phases (e.g., prophase), can be blurred in the MIP images leading to misclassifications. On the other hand, the selected slice not necessarily contains the entire object, e.g., detached chromosomes often were located in other slices. Consequently, we compute features related to texture, like Haralick texture features or wavelet features, based on the most informative slices, and features primarily related to object shape, e.g., size, circularity, or Zernike moments based on the MIP images. In addition, we use dynamic features representing the temporal changes of nuclei. These features are computed as differences of basic features (e.g., object size, mean and standard deviation of intensity, shape features) between subsequent frames. In total, we use 376 features. Our approach automatically classifies nuclei into 12 classes (inter-, pro-, prometa-, meta-, early ana-, late ana-, and telophase, abnormal early and late anaphase, abnormal telophase, abnormal interphase, and cell death) (Fig. 2, bottom, right) using support vector machines (SVMs) with a Gaussian radial basis function kernel. To obtain a more balanced data set we limited the number of interphase samples to 1000 per sequence, which yielded better results than using weighted SVMs.

2.3 Phase Sequence Analysis

To ensure consistency of the resulting phase sequences and to determine phase lengths we developed a finite state machine (FSM), which models cell cycle pro-

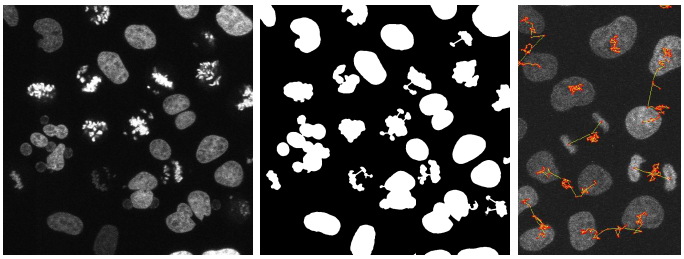

Fig. 1. MIP of a nocodazole treated experiment (left), segmentation with merged detached chromosomes (middle), enlarged tracking result with cell divisions (right).

gression and accepts only biologically plausible phase sequences. All 12 classes are represented by states of the FSM and all possible transitions between classes by state relations. Additionally, the FSM includes error states for all regular states that handle illegal phase transitions by correction or resetting. Note that using 12 instead of seven classes increases the model complexity significantly. The FSM processes phase sequences sequentially and each state includes a counter to determine phase durations.

3 Results

We applied our approach to 48 3D image sequences acquired on a confocal laser scanning microscope with an image acquisition interval of seven minutes and three slices per time step (1024×1024 pixels, 8 bit), and a total observation time of 15 h to 25 h. The imaged HeLa cells were fluorescently labeled with histone EGFP (visualizing the DNA), and treated with three different doses of the microtubule depolymerizing drug nocodazole (low, medium, and high). For each concentration we acquired six treated and six non-treated control sequences. Moreover, we performed siRNA knockdown experiments targeting the microtubule-associated gene *ch-TOG* and acquired six treated and six control sequences. Both treatments caused a delay in a prometaphase-like state, and afterwards chromosomal abnormalities such as lagging chromosomes, segregation defects, and appearance of micronuclei (Fig. 2, bottom, right).

The accuracy of our segmentation approach was evaluated based on four sequences where ground truth was determined manually. We quantified the occurrence of under- and oversegmentations and obtained an overall accuracy of 98.1 %. The tracking accuracy was determined based on the same sequences, and we found that for a total number of about 16,900 matches, 40 mismatches occurred (mostly caused by segmentation errors), yielding an overall accuracy of 99.8 %. Mitosis detection was evaluated based on 22 sequences, yielding an accuracy of 95.4 % and a positive predictive value of 92.0 %. The false positives were caused by abnormal morphologies such as detaching micronuclei. To determine

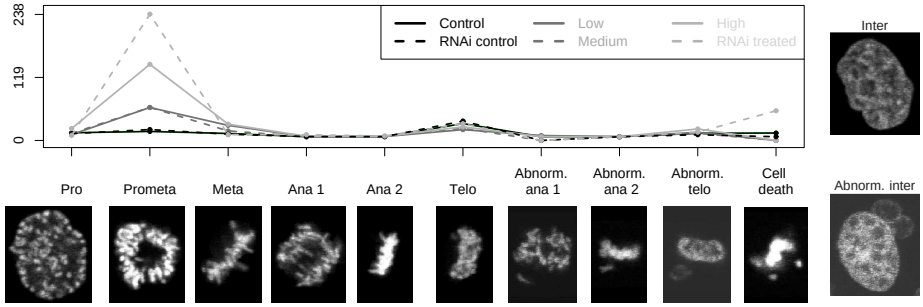

Fig. 2. Mean phase durations in minutes (y-axis) for the automatically analyzed image sequences, and example images for all 12 classes.

the classification accuracy we manually annotated 16 sequences from the nocodazole, and four sequences from the RNAi experiments. Using five-fold cross validation we yielded an overall classification accuracy of 93.9 % for the nocodazole, and 94.7 % for the RNAi data. Finally, we trained one SVM classifier with the annotated nocodazole and one with the annotated RNAi data, and applied them to the remaining, previously unseen test data. We determined the mitotic phase lengths based on the classification results using our finite state machine. To prove that our approach allows accurate determination of changes in mitotic progression, we analyzed the effects of nocodazole and siRNA treatment on the durations of mitotic phases. In particular, we tested whether prometaphase in perturbed cells was significantly longer than in controls, and compared the dose response of different nocodazole concentrations using Mann-Whitney tests (α=5 %). We found a highly significant prometaphase prolongation for all nocodazole concentrations (Fig. 2). The high concentration showed a stronger effect of higher significance (p=4.2·10^{-12}) compared to the medium and low concentrations (p=1.1·10^{-7} and p=4.1·10^{-7}). Cells treated with *ch-TOG* siRNA showed an even stronger increase of prometaphase duration (Fig. 2, p=2.3·10^{-12}).

4 Discussion

We presented an automatic approach for cell cycle analysis from live cell image sequences. Our approach allows accurate segmentation, tracking, and classification of normal as well as morphologically abnormal cell nuclei. We systematically evaluated the performance of the single steps of our approach based on real image sequences from different experiments. Our approach robustly quantifies the duration of mitotic phases and enables large-scale statistical analysis of phase durations to determine temporal phenotypes.

Acknowledgement. This work was supported by the EU projects MitoCheck and EuroDyna as well as the BMBF project NGFN-2.

References

1. Li K, Miller ED, Chen M, et al. Cell population tracking and lineage construction with spatiotemporal context. Med Image Anal. 2008;12(5):546–66.
2. Wang M, Zhou X, Li F, et al. Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. Bioinformatics. 2008;24(1):94–101.
3. Padfield DR, Rittscher J, Thomas N, et al. Spatio-temporal cell cycle phase analysis using level sets and fast marching methods. Med Image Anal. 2009;13:143–55.
4. Harder N, Mora-Bermúdez F, Godinez WJ, et al. Determination of mitotic delays in 3D fluorescence microscopy images of human cells using an error-correcting finite state machine. In: Proc. ISBI; 2007. p. 1044–7.
5. Chetverikov D, Verestoy J. Tracking feature points: a new algorithm. In: Proc ICIP; 1998. p. 1436–8.