Comparison of flow cytometry and image-based screening for cell cycle analysis

Damian J. Matuszewski[1,2,*], Ida-Maria Sintorn[1,2], Jordi Carreras Puigvert[1,3], Carolina Wählby[1,2]

[1] Science for Life Laboratory, Sweden,
[2] Centre for Image Analysis, Department of Information Technology, Uppsala University, Sweden
[3] Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden
* damian.matuszewski@it.uu.se (DJM)

Abstract. Quantitative cell state measurements can provide a wealth of information about mechanism of action of chemical compounds and gene functionality. Here we present a comparison of cell cycle disruption measurements from commonly used flow cytometry (generating one-dimensional signal data) and bioimaging (producing two-dimensional image data). Our results show high correlation between the two approaches indicating that image-based screening can be used as an alternative to flow cytometry. Furthermore, we discuss the benefits of image informatics over conventional single-signal flow cytometry.

Keywords: quantitative microscopy, DNA content histogram

1 INTRODUCTION

Optical microscopy is one of the most widely used techniques in cell and tissue research. As every form of cytologic instrumentation it represents a compromise between information content, fluorescence sensitivity, and acquisition speed. Conventional fluorescence microscopy provides high content information for the cost of low acquisition speed. Development of precise robotics has resulted in the creation of automated image-based high-content screening (IBHCS) systems. These systems are capable of imaging multiple stains of large numbers of cell populations in a short period of time [1]. Given that the samples are typically placed in microwell plates, containing up to 1536 wells per plate, this represents a perfect setting for high-throughput approaches. Consequently, the bottleneck has moved to analyzing the images: extracting information and interpreting the vast amount of generated data.

On the other hand, commonly used flow cytometry (FC) constitutes the opposite trade-off, with much higher single cell throughput (in the range of tens of thousands of analyzed cells per second) and information content reduced to a single or a handful of signals per cell. In a typical FC experiment hundreds of thousands of cells are analyzed. This allows for analysis of cell populations and

drawing robust conclusions about their distributions and dynamics. However, it does not allow for analyzing signal distributions or patterns within cells or how cells interact or spatially organize themselves. Additionally, despite the large analysis capacity per sample, numerous replicates require more preparation time, larger volumes, and higher initial cell numbers, thereby making FC an overall mid-throughput approach.

In this paper we present a comparative case study of two approaches for cell cycle analysis: FC and IBHCS. Similar image based approaches were presented in [2][3]. Here we quantitatively compare FC and IBHCS using automatic gate selection.

2 MATERIALS & METHODS

We used two different cell lines (lung cancer A549 and colon epithelial non-transformed CCD841) exposed to 5 different treatments: Dimethyl Sulfoxide (DMSO), Aphidicolin, Nocodazole, NaCl and Cisplatin. The cell lines were obtained from ATCC and maintained in Dulbecco's Modified Eagle Medium (Invitrogen) supplemented with 10% fetal bovine serum (Invitrogen) and 1% penicillin/streptomycin (Invitrogen), at 37°C and 5% CO2. NaCl and Cisplatin formulated in 0.9% NaCl were purchased from Hospira; Aphidicolin and Nocodazole were purchased from Sigma-Aldrich and dissolved in DMSO from Merck.

2.1 Image-based screening

An illustration of the workflow in the IBHCS is shown in Fig. 1A. A549 and CCD841 cells were seeded 24h prior to exposure to the compounds at a density of 1000 and 2500 cells per well respectively in imaging 384-well plates (Falcon). The cells were then exposed to the vehicle (DMSO or NaCl), 0.16µM-0.5µM of Aphidicolin, 0.16µM-0.5µM of Nocodazole, and 1.6µM-5µM of Cisplatin for 24h. Directly after, the cells were fixed in 4% paraformaldehyde (PFA) in PBS (Santa Cruz) for 15 minutes, and 2µg/ml Hoechst 33342 (Sigma-Aldrich) in Phosphate Buffered Saline (PBS) (Invitrogen) was added for 15 minutes to stain the DNA. Subsequently, the cells were imaged with an ImageXpress (Molecular Devices) high-throughput microscope. At this point, the sample preparation has typically taken approximately 3.5h, with minimal volumes used given the microwell plate format. Next, CellProfiler [4] was used to segment the cell nuclei by Gaussian smoothing followed by Otsu thresholding and watershed segmentation to split clustered nuclei based on intensity. Too small and too large objects were excluded. This commonly used segmentation approach provided good results (see Fig. 2), unaffected by small illumination variations between images. No background correction was necessary as the illumination field was even in this dataset. Finally, the total DNA content (integrated intensity of the DNA stain) was measured per nucleus. The CellProfiler processing pipelines (with all the parameters used) and sample images can be found at `http://www.cb.uu.se/~damian/IBS-FC_comparison/`.

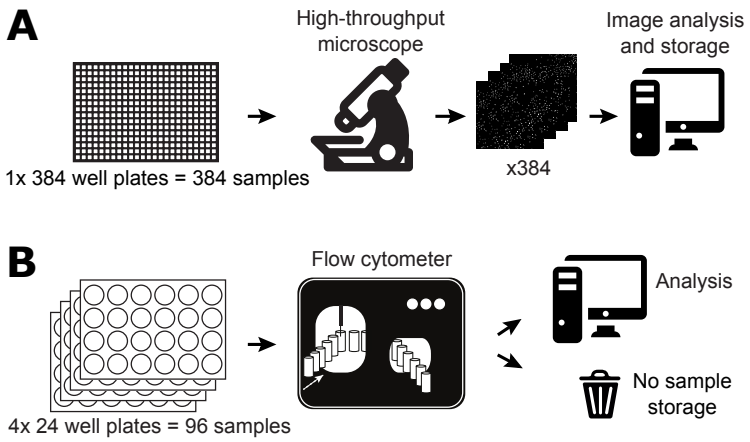

Fig. 1. Workflow diagram of the image-based screening (A) and flow cytometry (B).

The negative control histograms for DNA content analysis typically present two peaks that can be used to estimate the distribution of cells in different cell cycle phases, as shown in Fig. 3. The higher peak to the left (2N) corresponds to the normal amount of DNA, whereas the smaller peak to the right (4N) corresponds to the double amount of the DNA present in the nucleus after DNA replication during mitosis. The pooled histogram from all negative control wells for each cell line was analyzed to determine the integrated intensity values corresponding to the centers of the 2N and 4N sub-populations. These values were then applied as input parameters to define a search range for the exact 2N and 4N DNA peaks for each well and to normalize DNA intensity, such that the maximum of the 2N peak corresponds to 1 and the center of the 4N DNA peak corresponds to 2. If the histogram is normalized in this way, the individual cells can be categorized to one of the following five sub-populations according to DNA content as suggested in [3]:

- sub-2N all cells with DNA intensity below 0.75,
- 2N DNA intensity between 0.75 and 1.25,
- S DNA intensity between 1.25 and 1.75,
- 4N DNA intensity between 1.75 and 2.5,
- >4N DNA intensity above 2.5.

In order to avoid multiple peaks at 2N and 4N locations the histograms were smoothed with a Gaussian filter ($\sigma = 1.5$). The data analysis described here was performed with PopulationProfiler, a light-weight screening data analysis tool developed at Centre for Image Analysis, Uppsala University [5]. The DNA content measurements and the software are freely available [6].

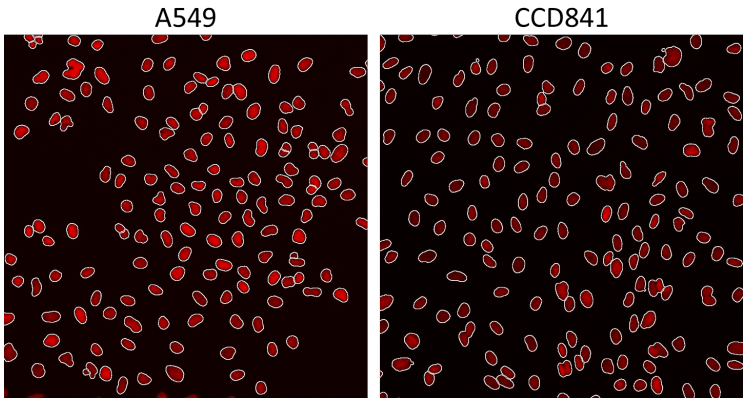

Fig. 2. Segmentation results for the negative controls of the two analyzed cell lines. White lines represent the outline of the cell segmentation results obtained with CellProfiler. The simple approach of Gaussian smoothing followed by Otsu thresholding and watershed segmentation gave satisfactory results.

2.2 Flow cytometry

An illustration of the FC workflow is shown in Fig. 1B. FC collects single cell measurements in real time, as compared to IBHCS where image analysis has to be applied after the data has been collected. A549 and CCD841 cells were seeded in 24 well plates (Greiner) at a density of 50.000 and 75.000 cells per well, respectively. After 24h, the cells were exposed to the corresponding concentrations of the aforementioned compounds for 24h. Next, the cells were trypsinized, collected into 1.5ml Eppendorf tubes to be pelleted by centrifugation, and washed once with PBS. Subsequently the cells were lysed in Vindelv's PI solution containing propidium iodide (PI), Tris, NaCl, Tergitol-type NP-40 and RNase (all from Sigma-Aldrich). The cells were then incubated for 1h at 4°C in the dark, to allow for the staining of the DNA, and subsequently analyzed by FC using a Beckman Coulter Navios. At this point, the sample preparation and analysis has typically taken 5 to 6h. In the case of a FC capable of analyzing samples in 96 well plate format this time may be shorter. The analysis of the data was done with the Beckman Coulter Kaluza software. It is to be noted at this point that this procedure is intended to maintain the nuclei intact, and it is to be emphasized that after the described steps, there is a large loss of cells mainly due to the trypsinization and washing steps. Upon initial acquisition of the samples, a size exclusion gating was applied to ensure single cell population measurements by excluding cell debris and cell doublets. Next, gates corresponding to different cell cycle phases were set using the negative control (DMSO) as reference. The gates were left unaltered for the rest of the samples.

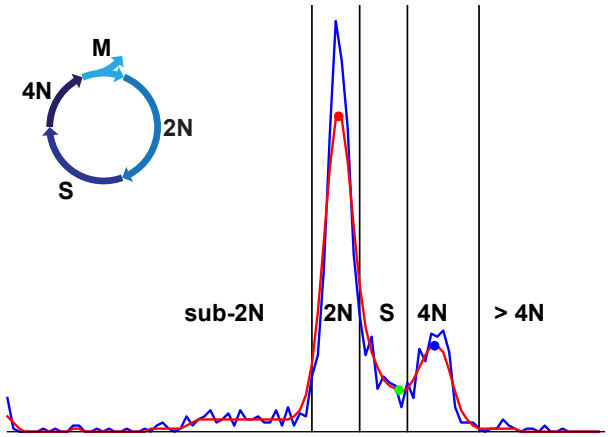

Fig. 3. Identification of the cell cycle sub-populations based on DNA content. The blue and red lines show data before and after smoothing, respectively. The red, blue and green dots on the histogram mark respectively the 2N peak, the 4N peak and the local minimum between them. The circular diagram in the top left corner illustrates transitions between the cell cycle phases: 2N (Gap 1), S (DNA synthesis), 4N (Gap 2) and M (Mitosis).

3 RESULTS

Figure 4A presents a table with Pearson's correlation coefficients between normalized cell cycle subpopulation distribution vectors found with IBHCS and FC. Each value corresponds to a crosswise pair of different drug-dose combinations. Figures 4B and 4C show tables with corresponding calculations but in these cases comparing results within IBHCS and FC respectively. For print clarity Fig. 4 shows results for pooled data from multiple runs of the same experiments. There were two replicates for each drug-dose combination in FC and three in IBHCS. Similar results were obtained when individual experiment runs were investigated. Figure 5 shows the mean contribution of the three main cell cycle subpopulations (2N, S and 4N) measured in % with the two methods.

For each table in Fig. 4, the background color is scaled so that white corresponds to low correlation and dark red to high. A characteristic cross pattern corresponding to low correlation caused by high response to Nocodazole (a drug affecting cell cycle by arresting cells in the 4N phase) is visible in all three tables. This demonstrates that both approaches provide similar results and can be successfully used for cell cycle analysis. The diagonal in Fig. 4A presents the Pearson's correlation coefficient between corresponding drug-dose combinations as outcomes from the two analysis approaches. In most cases it is above 90 %. The lowest correlation on the diagonal is observed for the two doses of Nocodazole. This corresponds to the biggest differences between IBHCS and FC percentages of 4N subpopulations in Fig. 5. We believe the cause for this is the very low cell count in IBHCS (below 400 for this particular treatment). For all

A

IMAGE-BASED SCREENING	No Cells	FLOW CYTOMETRY: DMSO	Noc 0.16	Noc 0.5	Aph 0.16	Aph 0.5	NaCl	Cisplatin 1.6	Cisplatin 5
DMSO-Noc	2449	0.97	0.01	0.07	0.90	0.62	0.97	0.93	0.56
DMSO-Aph	4108	0.98	-0.01	0.04	0.92	0.64	0.98	0.94	0.58
Noc 0.16	353	0.13	0.86	0.84	0.24	0.04	0.12	0.25	0.06
Noc 0.5	337	0.17	0.77	0.75	0.28	0.11	0.16	0.28	0.13
Aph 0.16	2332	0.72	0.44	0.49	0.90	0.86	0.72	0.87	0.86
Aph 0.5	2103	0.52	0.20	0.25	0.75	0.92	0.53	0.68	0.93
NaCl	2777	0.97	-0.01	0.05	0.91	0.64	0.98	0.94	0.58
Cisplatin 1.6	2162	0.90	0.29	0.35	0.97	0.80	0.90	0.97	0.77
Cisplatin 5	1671	0.63	0.53	0.59	0.87	0.89	0.64	0.82	0.90

B

IMAGE-BASED SCREENING	DMSO-Noc	DMSO-Aph	Noc 0.16	Noc 0.5	Aph 0.16	Aph 0.5	NaCl	Cisplatin 1.6	Cisplatin 5
DMSO-Noc	1.00	1.00	0.28	0.35	0.79	0.59	1.00	0.93	0.68
DMSO-Aph	1.00	1.00	0.23	0.31	0.78	0.60	1.00	0.93	0.68
Noc 0.16	0.28	0.23	1.00	0.98	0.49	0.20	0.25	0.42	0.49
Noc 0.5	0.35	0.31	0.98	1.00	0.57	0.33	0.32	0.48	0.54
Aph 0.16	0.79	0.78	0.49	0.57	1.00	0.89	0.79	0.95	0.98
Aph 0.5	0.59	0.60	0.20	0.33	0.89	1.00	0.60	0.77	0.89
NaCl	1.00	1.00	0.25	0.32	0.79	0.60	1.00	0.93	0.68
Cisplatin 1.6	0.93	0.93	0.42	0.48	0.95	0.77	0.93	1.00	0.90
Cisplatin 5	0.68	0.68	0.49	0.54	0.98	0.89	0.68	0.90	1.00

C

FLOW CYTOMETRY	No Cells	DMSO	Noc 0.16	Noc 0.5	Aph 0.16	Aph 0.5	NaCl	Cisplatin 1.6	Cisplatin 5
DMSO	19142	1.00	-0.05	0.01	0.93	0.65	1.00	0.96	0.58
Noc 0.16	11784	-0.05	1.00	1.00	0.20	0.17	-0.06	0.16	0.21
Noc 0.5	10518	0.01	1.00	1.00	0.27	0.25	0.00	0.23	0.29
Aph 0.16	9997	0.93	0.20	0.27	1.00	0.86	0.93	0.99	0.82
Aph 0.5	10519	0.65	0.17	0.25	0.86	1.00	0.65	0.81	1.00
NaCl	17480	1.00	-0.06	0.00	0.93	0.65	1.00	0.96	0.59
Cisplatin 1.6	14115	0.96	0.16	0.23	0.99	0.81	0.96	1.00	0.76
Cisplatin 5	9196	0.58	0.21	0.29	0.82	1.00	0.59	0.76	1.00

Fig. 4. Pearson's correlation coefficients of normalized cell cycle subpopulation vectors image-based screening vs. flow cytometry (A), image-based screening vs. image-based screening (B), and flow cytometry vs. flow cytometry (C). Various treatments: Aphidicolin (Aph), Nocodazole (Noc), NaCl and Cisplatin were applied to cell line A549. The drug dose is stated by the name (in µM). Dark background indicates high correlation.

analyzed drug-dose combinations the cell count in IBHCS was much lower than in FC (see Fig. 4). Nevertheless, even in the least populous case of Nocodazole, it was still sufficient to observe significant drug response (low correlation between Nocodazole and negative controls, as seen in Fig. 4B). This shows that less cells suffice to perform cell cycle analysis using IBHCS.

Comparing Fig. 4B and 4C it can be observed that while the overall correlation pattern is very similar, the individual values can be quite different, especially in the case of correlations between Nocodazole and other treatments. In this case FC always shows lower correlation value than IBHCS. Again we believe that the reason for this is the low cell count and the fact that the gating is not exactly the same in the two approaches and also that the strong effect of Nocodazole on the cell cycle manifests in different ways.

4 CONCLUSIONS

FC is frequently chosen for cell cycle analysis. However, it is often more time consuming than IBHCS, especially if many treatments are to be tested, and leads to irretrievable loss of the analyzed sample. This makes discrimination between

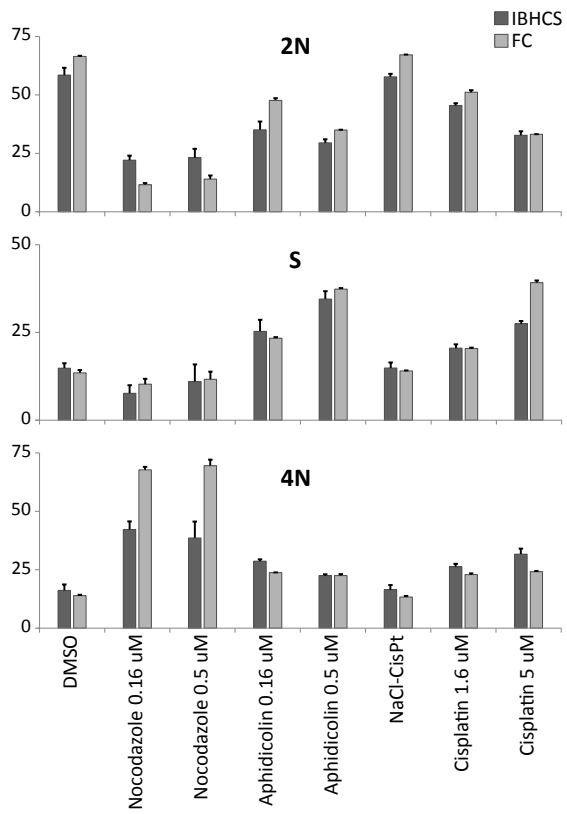

Fig. 5. Comparison between IBHCS and FC measures of the contribution (in %) of the three main cell cycle subpopulations (2N, S and 4N) for various tested treatments. The standard deviation was calculated for three and two replicates of IBHCS and FC respectively.

true signals and artifacts, and tracing errors in the analysis very difficult. On the other hand, IBHCS preserves raw data in the form of images, so that potential artifacts can be revealed by visual analysis. Furthermore, original biological samples can be re-imaged at a higher resolution, potentially revealing additional information. This contributes to reproducible measurements and possibility of further morphological analysis of intensity distribution in the nucleus or of other stains/compartments for interesting treatments. That is, IBHCS allows measuring a multitude of morphological features as well as comprehensive texture analysis of cells directly from images.

The sample preparation procedures for FC often lead to loss of cultured cells due to trypsinization, a treatment that may also have variable effect on cells in different phases of the cell cycle. As a consequence, this approach requires much larger amounts of cells and results may be skewed. IBHCS utilizes the biological sample more efficiently. However, new FCs that require a lower cell number

are becoming available [7]. In the near future we will observe more systems that merge the strengths of microscopy with those of FC. In fact the first multispectral image FCs, instruments that combine the speed and sample size of FC with the resolution and sensitivity of microscopy, are already commercially available [8].

Comparison of the results from the presented image-based DNA content analysis with those obtained using FC shows high correlation between the two approaches. The Pearson's correlation coefficient for corresponding results is above 75 % for all tested drug-dose combinations and above 90 % in more than 66 % of cases. The lowest correlation is observed for the two doses of Nocodazole (86 % and 75 % for 1.6 and 5 microM respectively). This, we believe, is caused by the low number of cells in the IBHCS analysis (these are the two least populated samples), as well as by the fact that Nocodazole has a strong effect on the cell cycle. Since the gating in the two approaches is not done in exactly the same manner, the effect manifests in different ways (some cells classified as 4N in FC would be considered as > 4N in IBHCS) which also affects the correlation.

References

1. Zanella, F., Lorens, J.B., and Link, W.: High content screening: seeing is believing. Trends in biotechnology 28, 237–245 (2010)
2. Gasparri, F., Cappella, P., and Galvani, A.: Multiparametric cell cycle analysis by automated microscopy. J. Biomol. Screen. 11, 586–598 (2006)
3. Chan, G.K.Y., Kleinheinz, T.L., Peterson, D., and Moffat J.G.: A simple high-content cell cycle assay reveals frequent discrepancies between cell number and ATP and MTS proliferation assays. PLoS ONE 8, doi: 10.1371/journal.pone.0063583 (2013)
4. Carpenter, A.E., Jones. T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O. et al.: CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. 7, doi: 10.1186/gb-2006-7-10-r100 (2006)
5. Matuszewski, D.J., Wählby, C, Puigvert, J.C., Sintorn, I.M.: PopulationProfiler: a tool for population analysis and visualization of image-based cell screening data. PLOS ONE 11, doi:10.1371/journal.pone.0151554 (2016)
6. Matuszewski, D.J., Wählby, C, Puigvert, J.C., Sintorn, I.M.: PopulationProfiler site `http://cb.uu.se/~damian/PopulationProfiler.html`
7. Black, C.B., Duensing, T.D., Trinkle, L.S., and Dunlay, R.T.: Cell-Based Screening Using High-Throughput Flow Cytometry. Assay and Drug Development Technologies 9, 13–20 (2011)
8. Amnis Corporation. `https://amnis.com/flowsight.html` [last visited on 25/02/2016]