Automated Estimation Parasitemia of Plasmodium berghei Infected Mice using CellProfiler

Otomatisasi Penghitungan Parasitemia Pada Mencit Terinfeksi Dengan *Plasmodium berghei* Menggunakan *CellProfiler*

Dwi Ramadhani *, Siti Nurhayati

Center for Technology of Radiation Safety and Metrology
National Nuclear Energy Agency of Indonesia
Jl. Lebak Bulus Raya No.49 Kotak Pos 7043 Jakarta Selatan 12070
Indonesia

Abstract

In this paper, we propose a technique for automatically recording parasitemia of mice infected with Plasmodium berghei by using Cellprofiler. Our purpose is to identify the difference number of parasitemia obtained by CellProfiler and manual assessment. We conducted a T-test analysis with $p < 0.05$. This value is considered to have a statistically significant different between automatic and manual process. Total of 50 thin blood smear images were analyzed for both automatically using CellProfiler and manual process. Results showed that there were insignificant difference between automatic and manual process ($p > 0.05$). It can be concluded that based on this research that automated quantification of parasitemia using CellProfiler was comparable but not better than manual.

keywords: Automation, CellProfiler, Parasitemia, Plasmodium berghei, Thin blood smear

Abstrak

Pada tulisan ini, penghitungan parasitemia pada mencit yang diinfeksi dengan *Plasmodium berghei* dengan *CellProfiler* diusulkan. Tujuan penelitian yang dilakukan adalah untuk mengetahui apakah terdapat perbedaan antara nilai parasitemia yang diperoleh dengan menggunakan *CellProfiler* dibandingkan dengan secara manual. Uji T digunakan untuk analisis statistik dengan $p < 0,05$ berarti terdapat perbedaan yang signifikan antara penghitungan otomatis dibandingkan dengan manual. Total sebanyak lima puluh citra apusan darah tipis mencit dianalisis secara otomatis menggunakan *CellProfiler* dan secara manual. Hasil penelitian menunjukkan bahwa tidak terdapat perbedaan signifikan antara nilai persentase yang diperoleh secara otomatis dibandingkan dengan manual ($p > 0,05$) . Dengan demikian dapat disimpulkan bahwa nilai persentase parasitemia yang diperoleh secara otomatis menggunakan *CellProfiler* sebanding tetapi tidak lebih baik dibandingkan dengan teknik manual.

kata kunci: CellProfiler, Otomatisasi, Parasitemia, *Plasmodium berghei*, Preparat Apusan Tipis

1. INTRODUCTION

Malaria is a serious global disease and a leading cause of morbidity and mortality in tropical and sub-tropical countries. It affects between 350 and 500 million people and causes more than one million deaths every year [1]. Malaria is caused by protozoan parasites of the genus Plasmodium. There are four species of Plasmodium that infect man and result in four kinds of malaria fever which were *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*. *Plasmodium falciparum* is most common in tropical and subtropical areas. It causes the most dangerous and malignant form of malaria without relapses and contributes to the majority of deaths associated with the disease [2].

Infection of mice with rodent Plasmodium species is routinely conducted to evaluate the efficacy of drugs and vaccines against malaria. Percentage of infected erythrocytes by parasitemia is used to monitor the progress of infection and recovery of infected mice. The most widely used technique for determining the development stage of the malaria disease is visual microscopical evaluation

*Corresponding Author. Tel: +6221-7513906
Email: dhani02@batan.go.id
Received: 22 Jun 2013; revised: 6 Jul 2013; accepted: 21 Aug 2013
Published online: 30 Nov 2013

of Giemsa stained blood smears. This process consists of manually counting the infected red blood cells against the number of red blood cells in a slide which is a laborious, time consuming and relies on the expertise of the experimenters with consequent person to-person variability [3]. Therefore, automation in the form of digital image analysis becomes obvious potential solution [4]. Infected red blood cell was indicated by the presence of dark spot inside it (red circle in Figure 1). When no dark spot is found, then is classified as a normal red blood cell (blue circle in Figure 1). If the dark spot was not inside the red blood cell then is simply an artifact (green circle in Figure 1).

Several studies have been done to automate the counting of the malaria parasite [4–8], but only a few studies that used open source software for automated quantification the parasitemia in thin blood smears. Ma et al. [4] developed Plasmodium AutoCount, which can automatically generate parasitemia values from *Plasmodium yoelii* in mouse thin blood smears. Plasmodium AutoCount, was written in the Python Programming Language and makes use of the NumPy and SciPy packages for fast numerical calculation. Unfortunately the use of Plasmodium Auto-Count is not easy and it was difficult to change the source code for biologists without training in computer vision or programming. Therefore, it was needed to develop software that easy to use and downloadable as a freeware from internet for automatically generates parasitemia values from thin blood smears. Here the used of CellProfiler an open access cell image analysis software for automatically generate parasitemia values from thin blood smears of mice that infected with *P. berghei* was reported. The results were compared to the manual counting of parasitemia from thin blood smears.

CellProfiler is freely available modular image analysis software that capable of handling hundreds of thousands of images. The software contains already-developed methods for many cell types and assays and is also an open-source, flexible platform for the sharing, testing, and development of new methods by image analysis experts. CellProfiler uses the concept of a pipelines of individual modules. Each module processes the images in some methods. The modules are placed in sequential order to create a pipeline. The pipeline usually contains image processing, object identification and measurement. Most modules are automatic but CellProfiler also allows interactive modules (for example the user can outlinining a region of interest in each image). Modules are mixed and matched for a specific project and each module's settings are adjusted appropriately. Upon starting the analysis, each image (or group of images if multiple wavelengths are available) travels through the pipelines and is processed by each module in order [9].

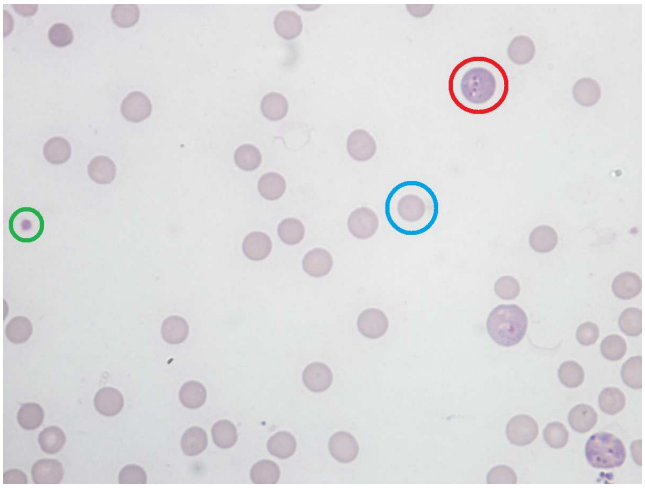

Figure 1. Infected red blood cell (red circle), normal red blood cell (blue circle) and artifact (green circle) in thin blood smear from infected mice.

2. METHODOLOGY

2.1 Giemsa-stained blood smears

Groups of mice were infected with *P. berghei* parasitized red blood cells. Three days post infection one drop of blood was taken from the tail tip of each mouse and used to make a thin blood smear. The smears were fixed with 100 % methanol for 2 minutes three times, stained with 10 % Giemsa for 15 minutes, and air-dried.

2.2 Image acquisition and standardization

A Nikon Eclipse E200 microscope with Nikon DS-Fi1 digital camera system was used to capture images of the smears. The smears were examined under oil immersion with a $100\times$ objective. Automated exposure of fixed light intensity through a fully opened iris with one push white balance was used. Images were captured at a resolution of 2560×1920 pixels using the NISH Element D software and saved as JPEG files with the NISH Element D software.

2.3 Manual counting of parasitemia

A manual counting tool, Cell Counting Aid that was developed by Ma et al. [4] used for counting parasitemia values from thin blood smears. Cell Counting Aid runs on the Microsoft Windows platform and was written in Visual Basic (The programme is free software released under the General Public License (GPL) version 2 license). After an image was opened with the software, the operator uses the mouse to point to each cell and clicks the left button if the cell is uninfected or

Table I. Steps for the automated approaches used in this work using CellProfiler

No	Analysis steps	Modules use in CellProfiler and function of the module
1	Loading pipelines and image folder	Operator loads the pipelines with image folder in computer. Operator must check the Rescale Intensities setting so that saturated values are rescaled to 1.0 by dividing all pixels in the image by the maximum possible intensity value.
2	Resize image	Rezise module was used for resize an image due to memory constrains. Resizing method used is Resize by a fraction or multiple of the original size with the value of resizing factor is 0.5 and Interpolation method used is Bilinear.
3	Correct the illumination factor in image	ColorToGray,ImageMath,CorrectIlluminationCalculate, CorrectIlluminationApply, GrayToColor and ImageMath modules was used to corrected the illumination factor in the image. In CorrectIlluminationCalculate module the block size used is 50 smoothing method used is median filter and Smoothing filter size is 300.
4	Identify the cells that maybe infected by Plasmodium	UnmixColor module was used identify the infected cells based on their colors.
5	Identify all cells	IdentifyPrimaryObjects module was used to identify all cells in the image. Diameter of objects that categorized as cells was 50 to 200 (Min-Max). Otsu Global was selected as the thresholding method and Two-class thresholding was used. The threshold correction factor value was set to 1, and (0.0-1.0) was choose as a lower and upper bounds on threshold. Intensity was selected as a method to distinguish clumped objects and draw dividing lines between clumped objects. Size of smoothing filter was set to 30.
6	Identify Plasmodium inside the infected cells	Mask Image and IdentifyPrimaryObjects modules was used to identify Plasmodium inside the infected cells. Diameter objects that categorized as parasites was 15 to 200 (Min-Max). Otsu Global was selected as the thresholding method and Three-class thresholding was used. The threshold correction factor value was set to 2.8, and $(0.0 - 1.0)$ was choosed as a lower and upper bounds on threshold. None was selected as a method to distinguish clumped objects and chose continue at handling of objects if excessive number of objects identified.
7	Establish a parent-child relationship between cells and Plasmodium.	RelateObjects module was used to establish a parent-child relationship between cells as a parent and Plasmodium as a child in order to make sure that infected cells detected in step 4 contain Plasmodium inside it.
8	Identify infected cells contain Plasmodium inside it.	FilterObjects module was used to filter the infected cells contain Plasmodium that represented as a parent with minimally had one or more child
9	Calculated the percentage of infected cells.	CalculateMath module was used to calculate the percentage of infected cells.
10	Export data to spreadsheet	Export to spreadsheet module was used to exports all data at the end of the analysis to Microsoft Excell.

the right button if it is infected. Parasitemia values are recalculated after each mouse click. The total number of cells and the total number of infected cells are recorded and can be exported to Microsoft Excel for analysis.

2.4 Automated counting of parasitemia

2.4.1 Validation Stage

An open access cell image analysis software CellProfiler 2.0 r10997 that developed by Broad Institute was used for an automated counting of parasitemia. CellProfiler (CP) runs on Microsoft

Windows XP SP 2 32-bit platform. Processor type used inside the computer is AMD Athlon(tm) 64 X2 Dual Core 5000+ with memory (RAM) is 1.87 GB. Pipelines were developed to automatically count the parasitemia.. Before the pipelines was used to doing automatic count parasitemia, pleminary pipelines were developed for several testing images. The testing images was made using CorelDraw 13 software and the resolution of testing images was 2560×1920 pixels.

2.4.2 Running Stage

A pipelines that used in validation stage was modified to doing automatic count of parasitemia. Modification in the pipelines was done because the images files size that obtained from NISH Element D software was bigger compared to testing images in validation stage. Detail explanations about the pipelines can be found in Table I.

2.5 Statistical analysis

Data analysis was done using the Microsoft Excel and StatPlus. Statistical significance of the difference between group means was performed by T-test with $p < 0.05$ were considered to have a statistically significant different.

3. RESULTS AND DISCUSSION

Pleminary pipelines were successful to automatically count the parasitemia in testing images. A modified pipelines was used for total of 50 blood smear images, the results showed in Table II. Scatter plots graph show that analyzes using CellProfiler had quite good linear relation with manual counting ($r = 0.7549$; Figure 2). An average different value parasitemia between automatic counts using CellProfiler compare to manual count was 1.31%. Statistical analysis using T-test show that there were no significant different between manual counting and automated counting (P = 0.452).

Even the statistical analysis showed that there was insignificant different between manual and automated counting, a significant variation between manual and automated counts was observed in several images that contain many dead parasites and white blood cells. The presence of dead parasites and white blood cells cause overestimation of parasitemia, which contributed to inaccuracies in automated counts. This problem commonly found in several researches on programme for determination parasitemia in thin blood smears. In our pipelines, the infected cells detection is based on parent-child relationship between the cells and the parasites. The red blood cell (parent) should count as the infected cell if it has at least a parasite (child)

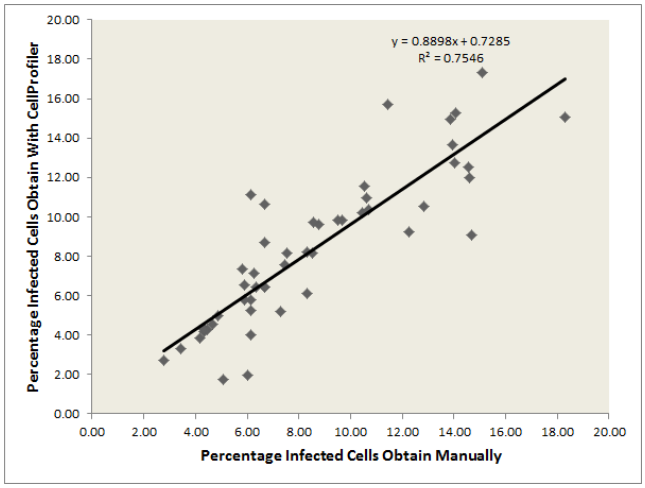

Figure 2. Scatter plots comparing parasitemia defined by automatic and manual counting.

in it. It will not count be counted as an infected cell if the parasite is not overlapping with the cell. Unfortunately this concept cannot be run entirely on a calculation using CellProfiler. Many dead parasites or artifacts from Giemsa stain are closely attached with the cells and as a consequence CellProfiler will counted cells as an infected cells and this make the CellProfiler overestimated the infected cells. Different with this, white blood cells that appear in image also considered as an infected cells because CellProfiler cannot differentiate the cell infected with parasitemia and white blood cells. In order to get an accurate automated counts using CellProfiler the quality of images is also important. Images must sharply focus and well illuminated, cell density in thin blood smear also must sparser than a typical smear to allow easy examination of the programme output. Interestingly CellProfiler have a CorrectIlluminationCalculate module to correct images with not having a good illumination. That was the reason after split the color images into Red, Green and Blue components and inverts each component a CorrectIlluminationCalculate module is used to correct the illumination images. After that we used a CorrectIlluminationApply module to correct all three images.

Another factor that may contribute to inaccurate determination of infected cells percentage was CellProfiler underestimated the red blood cells number. This because CellProfiler could not counted the red blood cells that appeared less than a half of the cells. For example, in Figure 3 two red blood cells were not counted with CellProfiler (in red circle). Overlapping red blood cells also can cause the inaccurate counted of the red blood cells because commonly it was counted as one cell. Overall, our pipelines worked very well for well stained and

Table II. Percentage Infected Cells Obtain Manually compared with CellProfiler

No	Image	Analysis Using CellProfiler			Manual Analysis			Difference
		TCN	ICN	PIC (%)	TCN	ICN	PIC (%)	
1	P101	49	3	6.12	52	3	5.76	0.36
2	P103	86	9	10.46	88	9	10.22	0.24
3	P104	48	2	4.16	52	2	3.84	0.32
4	P105	76	8	10.52	78	9	11.53	1.01
5	P107	66	7	10.60	64	7	10.93	0.33
6	P109	43	2	4.65	44	2	4.54	0.11
7	P110	49	3	6.12	36	4	11.11	4.99
8	P111	45	3	6.66	47	5	10.63	3.97
9	P112	49	3	6.12	38	2	5.26	0.86
10	P115	46	2	4.34	48	2	4.16	0.18
11	P116	46	2	4.34	46	2	4.34	0
12	P117	48	3	6.25	42	3	7.14	0.89
13	P118	57	5	8.77	52	5	9.61	0.84
14	P119	45	2	4.44	47	2	4.25	0.19
15	P120	46	2	4.34	47	2	4.25	0.09
16	P121	57	8	14.03	55	7	12.72	1.31
17	P122	62	6	9.67	61	6	9.83	0.16
18	P123	68	4	5.88	69	4	5.76	0.09
19	P125	47	4	8.51	49	4	8.16	0.35
20	P126	51	3	5.88	46	3	6.52	0.64
21	P127	53	4	7.54	49	4	8.16	0.62
22	P129	75	5	6.66	78	5	6.41	0.25
23	P130	70	6	8.57	72	7	9.72	1.15
24	P131	65	9	13.84	67	10	14.92	1.08
25	P132	71	10	14.08	72	11	15.27	1.19
27	P134	55	13	18.30	73	11	15.09	3.21
28	P136	63	8	15.09	52	9	17.30	2.21
29	P137	58	4	7.27	58	3	5.17	2.1
30	P139	63	6	9.52	61	6	9.83	0.31
31	P140	58	2	3.44	60	2	3.33	0.11
32	P141	49	3	6.12	50	2	4	2.12
33	P142	50	3	6	51	1	1.96	4.04
34	P143	48	7	14.58	50	6	12	2.58
35	P144	48	4	8.33	49	3	6.12	2.21
36	P146	30	2	6.66	31	2	6.45	0.21
37	P148	41	2	4.87	40	2	5	0.13
38	P149	45	3	6.66	46	4	8.69	2.03
39	P151	36	1	2.77	37	1	2.70	0.07
40	P152	59	3	5.08	57	1	1.75	3.33
41	P153	69	4	5.79	68	5	7.35	1.56
42	P154	67	5	7.46	66	5	7.57	0.11
43	P155	72	6	8.33	73	6	8.21	0.12
44	P156	78	10	12.82	76	8	10.52	2.3
45	P157	75	11	14.66	77	7	9.09	5.57
46	P158	63	4	6.34	62	4	6.45	0.11
47	P159	56	6	10.71	58	6	10.34	0.37
48	P160	57	7	12.28	54	5	9.25	3.03
49	P161	55	8	14.54	56	7	12.5	2.04
50	P162	43	6	13.95	44	6	13.63	0.32

Mean: 1.314 ± 1.44

TCN = Total Cells Number, ICN = Infected Cells Number, PIC = Percentage Infected Cells

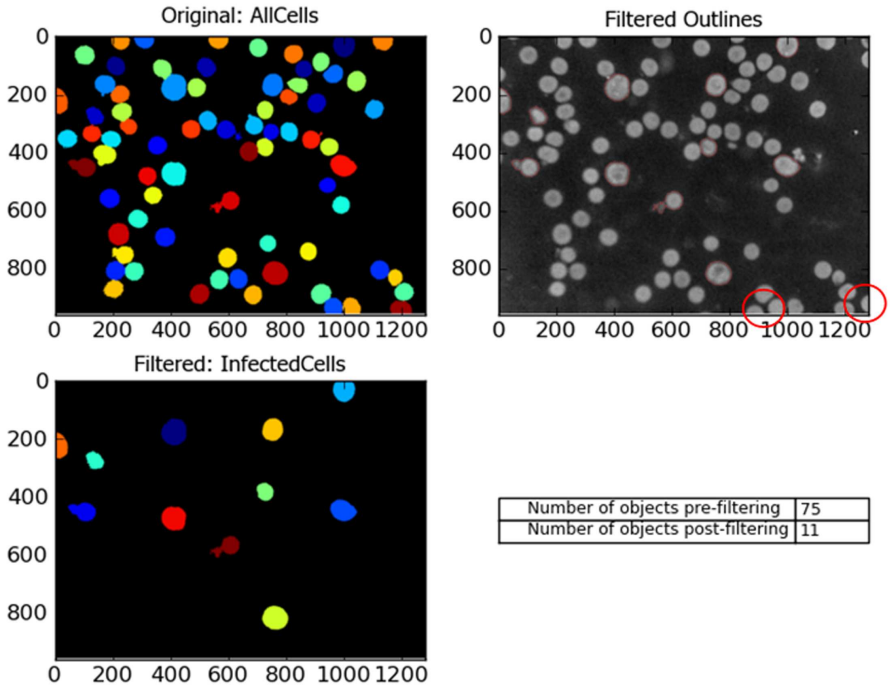

Figure 3. Underestimated RBC counting with CellProfiler.

Table III. Percentage Infected Cells Obtain Manually and with CellProfiler of *P. falciparum*

No	Image Number	Analysis Using CellProfiler			Manual Analysis			Difference
		TCN	ICN	PIC(%)	TCN	ICN	PIC(%)	
1	M168	126	3	2.38	126	1	0.79	1.58
2	M170	151	2	1.32	152	1	0.65	0.66
3	M132	99	11	11.11	100	10	10	1.11
4	M138	122	9	7.38	121	9	7.43	0.05
5	M139	122	6	4.92	121	5	4.13	0.78
								Mean: 0.83 ± 0.56

well separated cells and free from dead parasites or artifacts from Giemsa staining that closely or attached to the cells. We also had tried in five images of thin bloods smear of P. falciparum that well-stained, well-separated cells and free from dead parasited or artifacts from Giemsa stain. The results showed that our pipelines also can be used for images that came from *P. falciparum* (Table III).

Several automated image-processing approaches for blood smear analysis have been attempted with some reported success. For example, an automated image processing programme has been developed by Ross et al. [6] for the diagnosis and classification of Plasmodium species, which reported a sensitivity of 85 % and a positive predictive value of 81 %. Another programme that successfully determined parasitemia in thin blood smears is Plasmodium AutoCount. Plasmodium AutoCount which was developed by Sio et al. [7]. It was designed using the MATLAB platform involved the detection of the red blood cells using edge detection, binary morphology and clump splitting routines. Inaccuracies parasitemia detection arose mainly due to fields that contained overlapping or lysed cells and poor cell separation which resulted in Plasmodium AutoCount are overestimated the parasitemia. Plasmodium AutoCount needs approximately 30 seconds to process a single image.

Our pipelines need a longer time to process one single image compared to Plasmodium AutoCount (approximately 1.49 minutes). We are now in the process of developing another pipelines than can be used in smaller images, because at capturing images using NISH Element D we can change to smaller resolution for example to 640×480 pixels. We hope this can increase speeds of processing image. The advantages using CellProfiler was it could be instructed to process images in

batches of several hundred to automatically generate parasitemia values without the need for supervision. This also eliminates factors such as user fatigue and lack of standardization that are often associated with manual enumeration.

4. CONCLUSION

We have developed pipelines for CellProfiler software that can be used to determination of parasitemia from infected mouse blood. The parasitemia values obtain from CellProfiler show a quite good correlation with those determined by manual counting, and the variations between them are small (1.314 %). The pipelines can be used for *P. berghei and P. falciparum* and also can expand to estimate parasitemia more quickly. Overall it appears that in this research parasitemia defined with CellProfiler is comparable. However, the performance is not better than manual.

Daftar Pustaka

[1] W. UNICEF, "World malaria report," WMR and UNICEF, Tech. Rep., 2005.

[2] H. Hisaeda, K. Yasutomo, and K. Himeno, "Malaria: immune evasion by parasite," *International Journal of Biochemistry and Cell Biology*, vol. 37, no. 4, pp. 700–706, 2005.

[3] J. D. Smyth, *Introduction to Animal Parasitology*. Cambridge, New York, USA: Cambridge University Press, 1994.

[4] C. Ma, P. Harrison, L. Wang, and R. Coppel, "Automated estimation of parasitemia of plasmodium yoelii-infected mice by digital image analysis of giemsa stained thin blood smears," *Malaria Journal*, vol. 9, pp. 1–9, 2010.

[5] J. Frean, "Improving quantitation of malaria burden with digital image analysis," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 102, no. 11, pp. 1062–1063, 2008.

[6] N. Ross, C. Pritchard, D. Rubin, and A. Duse, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears," *Medical & Biological Engineering & Computing*, vol. 44, no. 5, pp. 427–436, 2006.

[7] S. Sio, W. Sun, S. Kumar, W. Bin, S. Tan, S. Ong, H. Kikuchi, Y. Oshima, and K. Tan, "Malariacount: an image analysis-based program for the accurate determination of parasitemia," *Journal of Microbiological Methods*, vol. 68, no. 1, pp. 11–18, 2007.

[8] M. Le, T. Bretschneider, C. Kuss, and P. Preiser, "A novel semi-automatic image processing approach to determine plasmodium falciparum parasitemia in giemsa-stained thin blood smears," *BMC Cell Biology*, vol. 9, pp. 1–12, 2008.

[9] A. Carpenter, T. Jones, M. Lamprecht, C. Clarke, I. Kang, O. Friman, D. Guertin, J. Chang, R. Lindquist, J. Moffat, P. Golland, and D. Sabatini, "Cellprofiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, pp. 1–11, 2006.