A simple image correction method for high-throughput microscopy

To the Editor: Advances in high-throughput fluorescence microscopy (HTFM) and image informatics have allowed single-cell phenotypic variability to be characterized and related to putative biological functions. However, estimates of single-cell variability may be skewed by inaccurate image correction. Furthermore, methods used in the HTFM literature do not generally take advantage of shared image properties, instead correcting each image independently. Here, we revisit the basic mathematical framework underlying image correction to show how the true distributions of commonly measured single-cell features are altered in uncorrected images. We also report an observation that allows the use of a simple method for accurate image correction in HTFM.

The intensity of a pixel, I, can be modeled by $I = S(F + B)$, where S represents shading due to uneven illumination, F is the biologically relevant foreground fluorescence and B is the background fluorescence[1–3] (**Supplementary Note**). Image correction is a data preprocessing step whose goal is to obtain F. (Subsequent normalization can address systematic experimental errors, such as row or column biases in microtiter plates[4,5].)

To investigate how the apparent cellular variability (for example, s.d. or coefficient of variation) is affected when S or B is not removed from images, we analyzed such effects for the distributions of commonly used single-cell features: average, total or ratiometric intensities of biomarkers (**Fig. 1a** and **Supplementary Note**). For these simple features, background and shading can alter apparent phenotypic variability in surprising ways. For example, the effects of either on the ratiometric feature are generally unpredictable, whereas background increases total feature variability owing to underlying differences in cell size. Such changes to distributions of single-cell features can be particularly important in screening, where variation is commonly used to determine whether experimental conditions differ statistically from control conditions.

In practice, the challenge in correction lies in the estimation of shading within an image, as subsequent background subtraction is relatively simple. How should shading in large HTFM image data sets be estimated? Every image from a microtiter plate could, in principle, have a unique shading pattern and background. Indeed, this is an implicit assumption of common correction methods in HTFM studies. However, we find that shading is not unique to every image but rather is predictable by image position within a well (**Fig. 1b**). This observation suggests a practical correction strategy for HTFM that estimates one reference shading pattern per in-well position (**Fig. 1b** and **Supplementary Note**). This empirical approach is simpler than commonly used methods that estimate correction parameters for each individual image and more accurate than applying one set of correction parameters to all images.

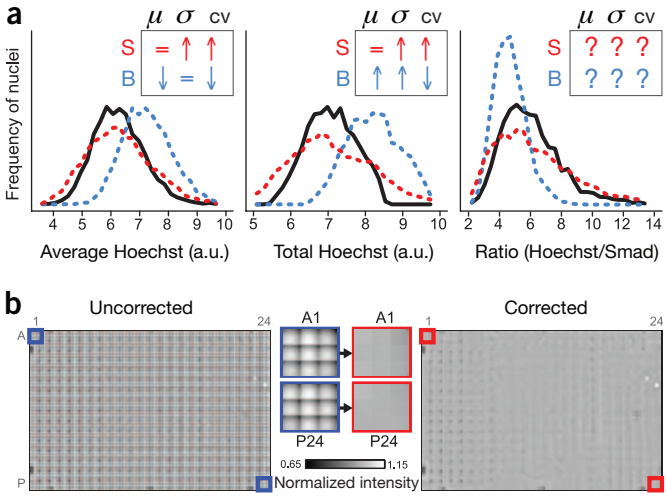

Figure 1 | A simple method for image correction in HTFM to improve estimates of single-cell phenotypes. (**a**) The effects of image background (B, blue) or shading (S, red) on distributions of three measured cellular features were mathematically analyzed (assumptions in **Supplementary Note**). Insets, predicted directions of change (↑, increase; ↓, decrease; '=', none; '?', indeterminate) of the mean (μ), s.d. (σ), or coefficient of variation (c.v. = σ/μ). For Hoechst- and Smad-stained cells (n > 3,700 cells), histograms illustrate changes to single-cell feature distributions before (black) or after the addition of synthetic background (blue) or shading (red). a.u., arbitrary fluorescence units. (**b**) Image grids from each well were normalized to the median within-well intensity and montaged across the entirety of a 384-well plate before (left) and after (right) positional image correction (details in **Supplementary Note**). Middle, close-up images of wells before and after image correction.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper (doi:10.1038/nmeth.2971).

ACKNOWLEDGMENTS

We thank J. Shay (University of Texas Southwestern Medical Center) for kindly providing human colonic epithelial cells. This work was supported by US National Cancer Institute R01CA133253 (S.J.A.); US National Institutes of Health R01 GM071794 (S.J.A.), R01 GM081549 (L.F.W.) and T32 GM008203 (A.D.C.); the Welch Foundation I-1619 (S.J.A.) and I-1644 (L.F.W.) and the Howard Hughes Medical Institute Med into Grad Initiative (A.D.C.).

AUTHOR CONTRIBUTIONS

A.D.C. performed the experiments, analyzed the data and prepared the figures. C.W. and S.R. provided critical interpretations of the data and revisions of the manuscript. A.D.C., L.F.W. and S.J.A. wrote the manuscript. S.R., L.F.W. and S.J.A. prepared the general mathematical formulations. All authors reviewed the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Adam D Coster, Chonlarat Wichaidit, Satwik Rajaram, Steven J Altschuler & Lani F Wu

Green Center for Systems Biology, University of Texas Southwestern Medical Center, Dallas, Texas, USA.
e-mail: steven.altschuler@utsouthwestern.edu or lani.wu@utsouthwestern.edu

1. Schultz, M.L., Lipkin, L.E., Wade, M.J., Lemkin, P.F. & Carman, G.M. *J. Histochem. Cytochem.* **22**, 751–754 (1974).
2. Likar, B., Maintz, J.B.A., Viergever, M.A. & Perus, F. *J. Microsc.* **197**, 285–295 (2000).
3. Madisetti, V. & Williams, D.B. *The Digital Signal Processing Handbook*. (CRC Press LLC, 1998).
4. Bray, M.-A. & Carpenter, A. in *Assay Guidance Manual* (eds. Sittampalam, G.S. *et al.*) (Eli Lilly and Company and the National Center for Advancing Translational Sciences, 2012).
5. Dragiev, P., Nadon, R. & Makarenkov, V. *BMC Bioinformatics* **12**, 25 (2011).