

YeastRGB: comparing the abundance and localization of yeast proteins across cells and libraries

Benjamin Dubreuil¹, Ehud Sass¹, Yotam Nadav¹, Meta Heidenreich¹, Joseph M. Georgeson¹, Uri Weill², Yuanqiang Duan³, Matthias Meurer³, Maya Schuldiner², Michael Knop^{3,4} and Emmanuel D. Levy^{1,*}

¹Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel, ²Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel, ³Zentrum für Molekulare Biologie der Universität Heidelberg (ZMBH), DKFZ-ZMBH Alliance, Heidelberg, Germany and ⁴Cell Morphogenesis and Signal Transduction, DKFZ-ZMBH Alliance and German Cancer Research Center (DKFZ), Heidelberg, Germany

Received August 14, 2018; Revised September 21, 2018; Editorial Decision September 25, 2018; Accepted October 18, 2018

ABSTRACT

The ability to measure the abundance and visualize the localization of proteins across the yeast proteome has stimulated hypotheses on gene function and fueled discoveries. While the classic C' tagged GFP yeast library has been the only resource for over a decade, the recent development of the SWAT technology has led to the creation of multiple novel yeast libraries where new-generation fluorescent reporters are fused at the N' and C' of open reading frames. Efficient access to these data requires a user interface to visualize and compare protein abundance, localization and co-localization across cells, strains, and libraries. YeastRGB (www.yeastRGB.org) was designed to address such a need, through a user-friendly interface that maximizes informative content. It employs a compact display where cells are cropped and tiled together into a 'cell-grid.' This representation enables viewing dozens of cells for a particular strain within a display unit, and up to 30 display units can be arrayed on a standard high-definition screen. Additionally, the display unit allows users to control zoom-level and overlay of images acquired using different color channels. Thus, YeastRGB makes comparing abundance and localization efficient, across thousands of cells from different strains and libraries.

INTRODUCTION

The systematic tagging of yeast open reading frames with GFP has been a landmark achievement enabling proteome-wide analyses of protein abundance and localization in living cells (1,2). Coupled with microscopy automation (3)

this library subsequently revealed how cells organize their proteome, and how the proteome reorganizes itself upon adaptation to environmental stresses (4–8). Fluorescence microscopy data gathered in such works represent a wealth of information that can be queried and mined via several databases (1,9–11), whose features are described in Table 1.

Browsing microscopy data of cells expressing fluorescently-tagged proteins entails two critical obstacles. First, a data bandwidth obstacle: an image acquired on a recent 16-bit camera with four million pixels is about 8 megabytes. Second, image dimensions exceed the resolution of an HD monitor, making it impossible to compare multiple images side-by-side without downscaling or cropping them.

These limitations are further exacerbated by recent developments in high-content microscopic imaging, which allow screening libraries in multiple conditions, and with multiple fluorescent tags. Moreover, the newly developed SWAp Tag (SWAT) technology enables creating arrayed yeast libraries with tags of choice at the N' or C' in about four weeks, thus drastically reducing the time and cost previously required to create such libraries (12–14). In fact, the N' and C' SWAT collections already led to genome-wide ORF tagging with latest-generation fluorescent proteins (12–14). These developments create a need for high-dimensional visualization and comparison of the same protein, i.e. imaged in different conditions, from multiple libraries, or tagged with multiple fluorescent proteins. For example, comparison of protein variants across libraries tagged at the N' versus C' can guide yeast biologists to choose a preferred fusion strategy. Indeed, interference between the fluorescent protein and regulatory regions could potentially be detected by a change of localization, as in the following examples [MST28|PMA1|PET54|YCK2|TOM70](#). YeastRGB has been designed with this goal in mind, which is to facilitate the comparison of dozens of strains imaged with multiple fluorescent tags on a single computer screen.

*To whom correspondence should be addressed. Tel: +972 89346715; Email: emmanuel.levy@weizmann.ac.il

Table 1. Comparison of features between YeastRGB and existing databases

	YeastGFP (1, 2)	YPL+	YeastRC (11)	LoQAtE (9)	CYCLoPs (10)	YeastRGB
Brightfield image	✓	✓	✓	✓		Download only
Microscopy type	Epi-fluorescence	Scanning confocal	Multiple types	Epi-fluorescence	Spinning disk confocal	Spinning disk confocal
Systematic contrast adjustment	✓	✓		✓		✓
Co-localization overlay	✓		✓		✓	✓
Batch download			✓	✓	✓ Upon request	✓
Cells displayed per strain	20–50	20–50	20–100	10–20	50–100	100–200
Maximal display dimensionality (strains $2 (1 \times 1 \times 2)$ \times screens \times colors)		$1 (1 \times 1 \times 1)$	$8 (2 \times 1 \times 4)$	$24 (4 \times 6 \times 1)$	$4 (1 \times 2 \times 2)$	$90 (10 \times 3 \times 3) (3 \times 10 \times 3)$
Freertext search to display protein groups			Search by GO id ^a	Search by pre-selected localization	Search by pre-selected keywords ^a	✓

^aResults of the search can yield several proteins but only one at a time can be displayed

The feature ‘Maximal display dimensionality’ is calculated as the combination of strains, conditions/libraries and color channels that can be displayed simultaneously on a screen. In YeastRGB, a 10×3 matrix of strains and conditions/libraries fits on a screen, where each strain can be rendered in up to three colors, yielding a maximal dimensionality equal to $10 \times 3 \times 3 = 90$.

Yeast libraries available on YeastRGB and their imaging

Currently, YeastRGB hosts image collections for three proteome-wide yeast libraries: The first is the classic C’ tagged GFP collection, (1,2) in which ORFs are expressed from their endogenous promoter fused to EGFP (15) at their C-terminus, and followed by the ADH1 terminator. The second library was derived from the C-SWAT strain collection (13). It contains ORFs under the expression of their native promoter, fused to mNeonGreen (16) at the C-terminus and followed by their endogenous terminator. This library was mated to a strain expressing NUP49 tagged with mCherry to highlight the nuclear membrane, and diploids with only one copy tagged with mNeonGreen were imaged. The third library was derived from the N-SWAT strain collection (12,14). It contains ORFs fused to a superfolder-GFP (17) at the N-terminus, under control of the constitutive *NOPI* promoter from *Saccharomyces paradoxus*. In this library, the native promoter controls the expression of mCherry (18) followed by the ADH1 terminator, and a blue fluorescent protein located in between (mTagBFP2 (19)) is expressed constitutively using the TDH3 promoter and URA3 terminator. All these libraries were imaged with the same microscopy system, which provided high-quality pictures (Yokogawa W1 confocal head coupled to an Olympus X83 microscope, dual Hamamatsu Flash 4 V2 cameras, 60X oil objective plan apo).

Additionally, we provide detailed instructions on the website for submitting microscopy screens to be added to yeastRGB. For such a submission, the main information that we require are images in TIFF format, which embed segmentation information. This information should be in the form of overlay objects highlighting the contour of individual cells, or highlighting their position (e.g. the coordinates of the nucleus).

The strain ‘display unit’ in YeastRGB

Existing databases show full-size images of microscopy pictures, which limits the number of strains and fluorescent channels that can be compared on a computer screen. Indeed, an HD screen has a resolution of 1920×1080 pixels, which is smaller than even a single image acquired by cur-

rent CMOS cameras (typically of $2k \times 2k$ pixels). An alternative strategy has been employed in the LoQAtE database (9), where a small cell-containing region is cropped for each strain. However, the cropped images contain ten to twenty cells at most (Table 1).

To address this limitation, we developed an image processing pipeline for our microscopy setup, which identifies, crops, and tiles 225 individual cells from multiple images into a 15×15 grid, as detailed in the section ‘Image pre-processing in YeastRGB’ below. This procedure eliminates uninformative regions where no cells are present (Figure 1A), while images of cells are embedded within a ‘display unit.’ Display units provide a user-friendly interface organized in five parts (Figure 1B):

- (1) A header provides gene identifier and name linked to SGD (20), the name of the screen and the multiwell plate position of the strain in the library.
- (2) A control panel allows users to customize image display (color-channel selection, zoom level) as well as download images of interest.
- (3) A cell-grid displays the tiled cells from individual color channel images.
- (4) A data table provides fluorescence intensity values of tagged proteins relating to absolute and relative cellular abundance as well as cell-to-cell variations in their abundance.
- (5) Collapsible panels inform on function and localization, as obtained from SGD (20) and GO.

Image data stored on YeastRGB are rendered through display units, which can be arrayed to compare multiple strains or screens (Figure 1C). Importantly, when multiple display units are shown, all of them can be controlled individually, using each unit’s controls, or simultaneously, using global control buttons on the dashboard (Figure 1D).

The search function of YeastRGB

The fact that YeastRGB is optimized to visualize and compare multiple strains simultaneously prompted the development of a versatile search function, which allows retrieving entire classes of proteins based on their name, function,

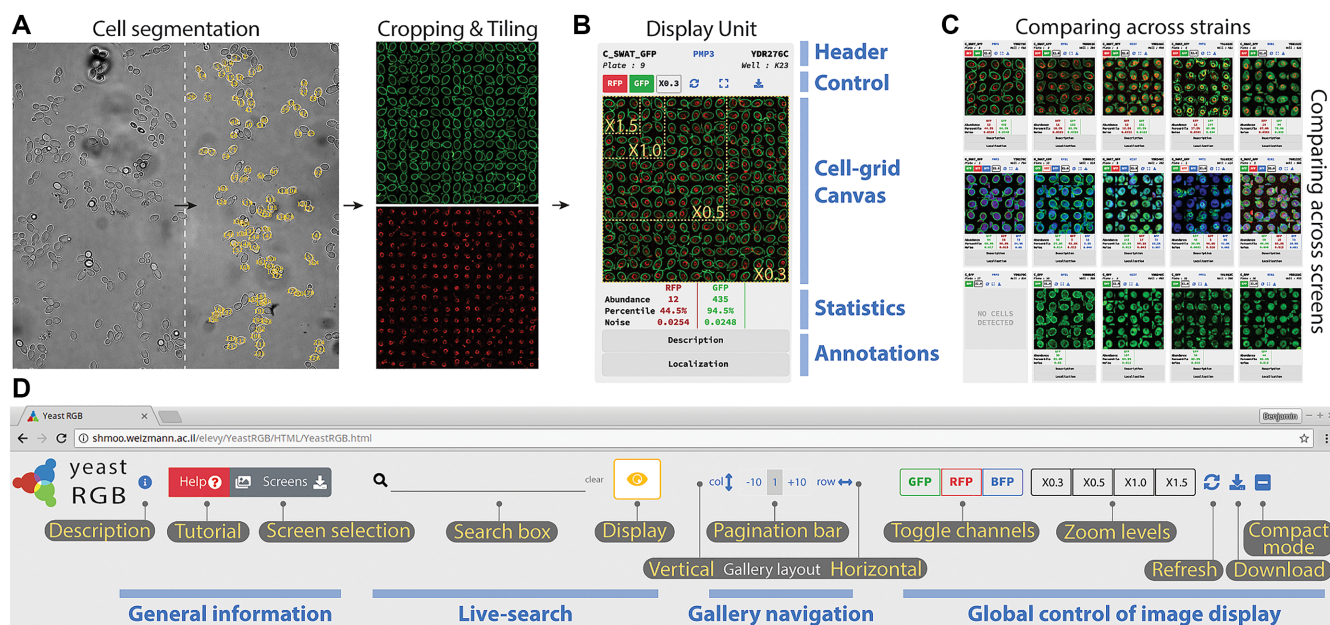


Figure 1. Cell tiling strategy of YeastRGB and its user interface. (A) Cells are identified in brightfield microscopy images, cropped and tiled before being displayed in the ‘display unit’. (B) Display units provide the user with information and control over images: The ‘header’ gives gene and screen information. The ‘Control’ part allows adjusting the zoom level, the color-channels overlay, and it contains a button to download images. The cell-grid section displays cropped and tiled cell images. Underneath, statistics provide the mean intensity and corresponding percentile across the screen for each fluorescent channel, as well as the normalized variance of intensities across cells. The annotation section can be expanded to reveal known information on function and localization. (C) Display units are compact elements, and up to 30 of them can be shown on an HD screen for comparing multiple strains, conditions or libraries. (D) The dashboard allows querying YeastRGB with live search and can be subsequently used to interact with display units globally. It also provides a help button triggering a short tutorial slideshow, and a ‘screens’ button to access details about the yeast libraries imaged.

localization, or keyword(s) present in their description or GO annotation. To guide new users through the interface, the dashboard includes a ‘Help’ button triggering a tutorial slideshow with explanations and screenshots of consecutive steps to a successful search.

The dashboard features a live search performed on a fixed-size table comprising 6692 rows for each gene from the complete SGD genome (21) (last update 23/04/2018) where columns are (i) the ORF identifier, (ii) the gene name if defined, (iii) the functional description, (iv) the full description, and (v) GO annotations (Cellular Component) of the gene of interest. Any input of at least three characters immediately triggers filtering of the proteome table, finding matching ORFs and highlighting the first occurrence of the matched query in their name and description. If an ORF is matched by its GO annotation, a checkmark appears in the fifth column, named ‘GO matched?’. This feature was introduced to keep a single line per matched entry. The live search will help users refine their selection of proteins before loading display units. Additionally, users can manually uncheck individual rows of the result table by clicking on their corresponding checkboxes.

Once the search and selection steps are completed, users can press the eye-shaped ‘Show’ button to generate the gallery of display units for all selected proteins and all collections picked in ‘Screen selection’ panel. For example, typing ‘YNL216W’ in the search box returns a single matching ORF. One display unit is then shown per imaged library, highlighting that this protein localizes to the nu-

cleus in the three libraries currently hosted in YeastRGB. To study cellular protein organization, users can also type more general words such as ‘cytoskeleton’, ‘stress granule’, or ‘nuclear pore complex’ to have access to images of proteins involved in a particular function, process, molecular phenotype or protein complex.

Importantly, searching on YeastRGB supports regular expressions. Users can use this feature to display groups of functionally related proteins. For example, ‘ \wedge RAD[0-9]+\$’ will search for names starting (\wedge) with ‘RAD’ followed by one or many (+) digits until the last (\$) character, returning a range of proteins that confer protection against radiation damage. It is also possible to retrieve lists of ORF-identifiers using parentheses and ‘|’ as a separator between each identifier, as in the following example: ‘(YBR160W|YPL004C|YOR101W|HSP31|ADE4)’. Also, a quick search can be performed by directly typing an address in the web browser, terminating the default URL with ‘?search=’ followed by the query, as illustrated in the hypertext links from this paragraph (e.g. <http://shmoo.weizmann.ac.il/levy/YeastRGB/HTML/YeastRGB.html?search=ATP>).

For requests involving the display of a single protein, the default layout is to show all the screened collections on a single row. However, when multiple proteins are to be displayed, the default layout is to show a matrix of display units, where each row corresponds to a particular collection, and each column corresponds to a protein. This layout can be toggled at any time by the ‘row’ and ‘col’ buttons from the dashboard (Figure 1D).

For queries involving more than ten proteins, a pagination bar appears in the dashboard and allows navigating between pages.

Image pre-processing in YeastRGB

High-content microscopy screens involve acquiring thousands of images from which single-cell data are extracted. Among the screens hosted on YeastRGB, five to eight picture sets (brightfield + fluorescent images) were acquired per strain, with each screen representing 5000 or more strains. A first step consisted in identifying and segmenting cells, which was achieved with custom algorithms (Matalon *et al.* 2018, [bioRxiv](#)) in Fiji/ImageJ (22) and facilitated by using a brightfield image taken out of focus, similarly to previously described (23). The cell masks of the segmented cells were used to define squares of 65 pixels centered on each cell. The squares were then cropped and assembled into a grid of 15×15 cells of dimension 975×975 pixels. The same masks were used in all channels, yielding one image per channel. To facilitate visualization, the contrast of each cell-grid image was adjusted such that 0.35% of pixels above a signal threshold were saturated (where the signal threshold was defined as 10 arbitrary units of fluorescence in all channels). Importantly, this process preserves relative fluorescence intensities across cells, thus allowing to assess for noisy protein expression (e.g. high variance across cells), as in the case of [HSP26](#) or [HXK1](#).

By default, 25 cells (5×5) are shown in the cell-grid of a single display unit. However, the display unit provides a zoom feature with four different magnifications (3×3 , 5×5 , 10×10 and 15×15 cells) showing up to 225 cells (15×15). On a computer screen with HD resolution, ~ 30 display units (or 40 in ‘compact mode,’ Figure 1D) can be viewed simultaneously enabling comparison for over 6000 cells at once.

Data download

The ‘Screens’ button (dashboard, Figure 1D) opens a panel providing access to the list of yeast collections available. A table summarizes information about the strains contained in each collection and references the publication describing each collection construction. Links allow downloading the tiled images in bulk for each screen, and also point to a tabular file with details of the proteins present in each collection.

The content of each display unit can be retrieved as a compressed JPG image through the download button of the display unit. The dashboard also has a ‘download’ button that creates a zipped folder with images from all display units present on the page.

CONCLUSION

The cell-tiling strategy employed in YeastRGB allows for a compact representation to mine high-content imaging data of yeast collections, which we hope will facilitate data reuse and re-exploration. Indeed, storage of the raw data for our proteome-wide screens decreased up to 50 folds, from 0.63TB (C’ GFP), 1.38TB (C’ SWAT) and 1.97TB

(N’ SWAT) to 17GB, 28GB, and 39GB respectively (values before compression). Moreover, it will be simpler to apply classification algorithms on the segmented cells than it would be on raw, unsegmented images (6,24–27), e.g. for the automated identification of protein localization.

The display design presents new possibilities for visually comparing protein abundance and localization across multiple strains and infer hypotheses on protein function. For example, localization discrepancies among subunits of a particular protein complex could suggest moonlighting (28), whereas discrepancies across N’ and C’ tagging could suggest interference of the tag (Davidi *et al.* 2018, [bioRxiv](#)).

The URL-embedded search function, which we use in this manuscript to link to specific queries, could be used more broadly by the yeast community. For example, to create dynamic links to groups of proteins of interest (e.g. all the physical or genetic interactors of a particular protein), or could serve an educational purpose (e.g. to illustrate the main subcellular localizations). Last, scientific hypotheses and advances often stem from serendipitous observations, and we hope that the user-friendly interface of yeastRGB will spark numerous such observations.

DATA AVAILABILITY

YeastRGB is an open source initiative available at www.yeastRGB.org. All images contained in the database are available for bulk download on the website, except the N’ SWAT library, for which the green fluorescent channel is available, while the Red and Blue fluorescent channels will be made available upon publication of the work associated to them.

Open reading frames identifiers cited here use the nomenclature defined by SGD to refer to a specific gene locus onto *Saccharomyces cerevisiae* chromosomes. Gene names correspond to conventional names, as published by the [SGD on 23 April 2018](#).

ACKNOWLEDGEMENTS

We thank members of the Levy, Knop and Schuldiner labs, Michal Breker and Suzi Aleksander for feedback on the web interface.

FUNDING

Israel Science Foundation and the Israeli Centers of Research Excellence (I-CORE) Program of the Planning and Budgeting Committee [2179/14 to E.D.L., 1452/18 to E.D.L.]; China Scholarship Council (to Y.D.); German Research Foundation (DFG) [KN498/11-1 to M.K., SFB grant 1190 to M.S.]; European Research Council [CoG 646604/Peroxisystem to M.S.]; Estelle Funk Foundation; Estate of Fannie Sherr; Estate of Albert Deligher; Merle S. Cahn Foundation; Mrs. Mildred S. Gosden; Estate of Elizabeth Wachsman; Arnold Bortman Family Foundation (to E.D.L.); M.S. is an incumbent of the Dr Gilbert Omenn and Martha Darling Professorial Chair in Molecular Genetics; E.L.D. is an incumbent of the Recanati Career Development Chair of Cancer Research. Funding for open access charge: Israel Science Foundation.

Conflict of interest statement. None declared.

REFERENCES

- Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S. (2003) Global analysis of protein expression in yeast. *Nature*, **425**, 737–741.
- Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S. and O'Shea, E.K. (2003) Global analysis of protein localization in budding yeast. *Nature*, **425**, 686–691.
- Mattiazzi Usaj, M., Styles, E.B., Verster, A.J., Friesen, H., Boone, C. and Andrews, B.J. (2016) High-Content screening for quantitative cell biology. *Trends Cell Biol.*, **26**, 598–611.
- Breker, M., Gymrek, M. and Schuldiner, M. (2013) A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *J. Cell Biol.*, **200**, 839–850.
- Tkach, J.M., Yimit, A., Lee, A.Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J.A., Ou, J., Moffat, J., Boone, C. *et al.* (2012) Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat. Cell Biol.*, **14**, 966–976.
- Lu, A.X., Chong, Y.T., Hsu, I.S., Strome, B., Handfield, L.-F., Kraus, O., Andrews, B.J. and Moses, A.M. (2018) Integrating images from multiple microscopy screens reveals diverse patterns of change in the subcellular localization of proteins. *Elife*, **7**, e31872.
- Chong, Y.T., Koh, J.L.Y., Friesen, H., Duffy, S.K., Duffy, K., Cox, M.J., Moses, A., Moffat, J., Boone, C. and Andrews, B.J. (2015) Yeast proteome dynamics from single cell imaging and automated analysis. *Cell*, **161**, 1413–1424.
- Dénervaud, N., Becker, J., Delgado-Gonzalo, R., Damay, P., Rajkumar, A.S., Unser, M., Shore, D., Naef, F. and Maerkl, S.J. (2013) A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 15842–15847.
- Breker, M., Gymrek, M., Moldavski, O. and Schuldiner, M. (2013) LoQAtE—Localization and quantitation ATlas of the yeast proteome. A new tool for multiparametric dissection of single-protein behavior in response to biological perturbations in yeast. *Nucleic Acids Res.*, **42**, D726–D730.
- Koh, J.L.Y., Chong, Y.T., Friesen, H., Moses, A., Boone, C., Andrews, B.J. and Moffat, J. (2015) CYCLOPs: A comprehensive database constructed from automated analysis of protein abundance and subcellular localization patterns in *Saccharomyces cerevisiae*. *G3*, **5**, 1223–1232.
- Riffle, M. and Davis, T.N. (2010) The yeast resource center public image repository: A large database of fluorescence microscopy images. *BMC Bioinformatics*, **11**, 263.
- Yofe, I., Weill, U., Meurer, M., Chuartzman, S., Zalckvar, E., Goldman, O., Ben-Dor, S., Schütze, C., Wiedemann, N., Knop, M. *et al.* (2016) One library to make them all: streamlining the creation of yeast libraries via a SWAp-Tag strategy. *Nat. Methods*, **13**, 371–378.
- Meurer, M., Duan, Y., Sass, E., Kats, I., Herbst, K., Buchmuller, B.C., Dederer, V., Huber, F., Kirrmaier, D., Štefl, M. *et al.* (2018) Genome-wide C-SWAT library for high-throughput yeast genome tagging. *Nat. Methods*, **15**, 598–600.
- Weill, U., Yofe, I., Sass, E., Stynen, B., Davidi, D., Natarajan, J., Ben-Menachem, R., Avihou, Z., Goldman, O., Harpaz, N. *et al.* (2018) Genome-wide SWAp-Tag yeast libraries for proteome exploration. *Nat. Methods*, **15**, 617–622.
- Heim, R., Cubitt, A.B. and Tsien, R.Y. (1995) Improved green fluorescence. *Nature*, **373**, 663–664.
- Hostettler, L., Grundy, L., Käser-Pébernard, S., Wicky, C., Schafer, W.R. and Glauser, D.A. (2017) The bright fluorescent protein mNeonGreen facilitates protein expression analysis in vivo. *G3*, **7**, 607–615.
- Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C. and Waldo, G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.*, **24**, 79–88.
- Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E. and Tsien, R.Y. (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.*, **22**, 1567–1572.
- Subach, O.M., Gundorov, I.S., Yoshimura, M., Subach, F.V., Zhang, J., Grünwald, D., Souslova, E.A., Chudakov, D.M. and Verkhusha, V.V. (2008) Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.*, **15**, 1116–1124.
- Cherry, J.M., Hong, E.L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E.T., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R. *et al.* (2012) *Saccharomyces* Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.*, **40**, D700–D705.
- Engel, S.R., Dietrich, F.S., Fisk, D.G., Binkley, G., Balakrishnan, R., Costanzo, M.C., Dwight, S.S., Hitz, B.C., Karra, K., Nash, R.S. *et al.* (2014) The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda)*, **4**, 389–398.
- Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
- Gordon, A., Colman-Lerner, A., Chin, T.E., Benjamin, K.R., Yu, R.C. and Brent, R. (2007) Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat. Methods*, **4**, 175–181.
- Pärnamäa, T. and Parts, L. (2017) Accurate classification of protein subcellular localization from High-Throughput microscopy images using deep learning. *G3*, **7**, 1385–1392.
- Kraus, O.Z., Grysb, B.T., Ba, J., Chong, Y., Frey, B.J., Boone, C. and Andrews, B.J. (2017) Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.*, **13**, 924.
- Handfield, L.-F., Strome, B., Chong, Y.T. and Moses, A.M. (2015) Local statistics allow quantification of cell-to-cell variability from high-throughput microscope images. *Bioinformatics*, **31**, 940–947.
- McQuin, C., Goodman, A., Chernyshev, V., Kamensky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D. *et al.* (2018) CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.*, **16**, e2005970.
- Matalon, O., Horovitz, A. and Levy, E.D. (2014) Different subunits belonging to the same protein complex often exhibit discordant expression levels and evolutionary properties. *Curr. Opin. Struct. Biol.*, **26**, 113–120.