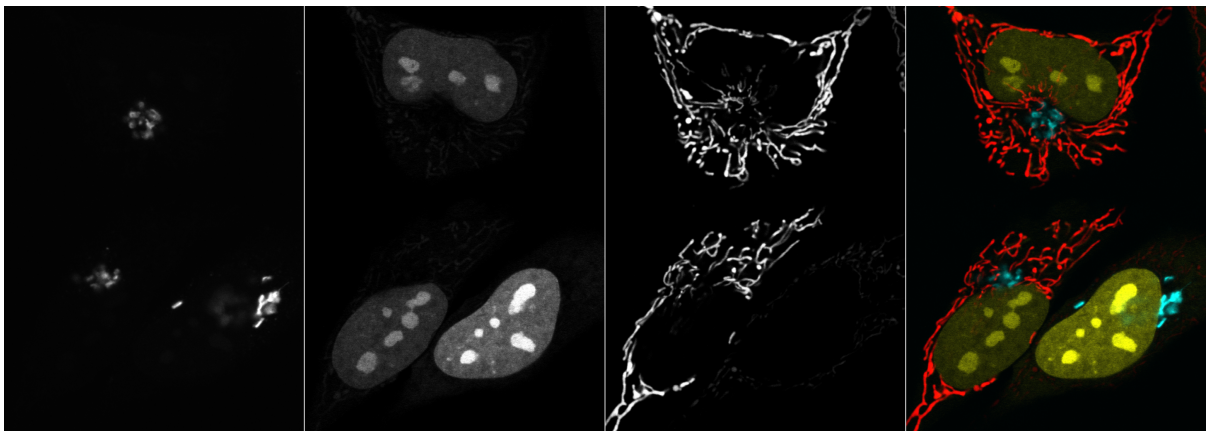
Robust and Bright Genetically Encoded Fluorescent Markers for Highlighting Structures and Compartments in Mammalian Cells.

Anna O. Chertkova[1], Marieke Mastop[1], Marten Postma[1], Nikki van Bommel[1], Sanne van der Niet[1], Kevin L.Batenburg[1], Linda Joosen[1], Theodorus W.J. Gadella Jr. [1], Yasushi Okada[2], Joachim Goedhart[1,#]

[1]Swammerdam Institute for Life Sciences, Section of Molecular Cytology, van Leeuwenhoek Centre for Advanced Microscopy, University of Amsterdam, P.O. Box 94215, NL-1090 GE Amsterdam, The Netherlands
[2]Laboratory for Cell Polarity Regulation, Quantitative Biology Center (QBiC), RIKEN, Osaka, Japan & Department of Physics and Universal Biology Institute, Graduate School of Science, the University of Tokyo, Tokyo, Japan



Abstract

To increase our understanding of cells, there is a need for specific markers to identify biomolecules, cellular structures and compartments. One type of markers comprises genetically encoded fluorescent probes that are linked with protein domains, peptides and/or signal sequences. These markers are encoded on a plasmid and they allow straightforward, convenient labeling of cultured mammalian cells by introducing the plasmid into the cells. Ideally, the fluorescent marker combines favorable spectroscopic properties (brightness, photostability) with specific labeling of the structure or compartment of interest. Here, we report on our ongoing efforts to generate robust and bright genetically encoded fluorescent markers for highlighting structures and compartments in living cells.

[#] corresponding author: j.goedhart@uva.nl, @joachimgoedhart

Introduction

It is difficult to overestimate the impact that Green Fluorescent Protein had and has on basic and applied research. The isolation of the cDNA encoding GFP from the jellyfish *Aequorea victoria* (Prasher et al., 1992) has enabled scientists to track molecules, organelles, protein complexes or cells at different time (milliseconds to days) and length scales (nanometer to centimeter). Fluorescent proteins (FPs) with different emission colors that have since been generated or isolated from different sea-resident animals can be used to highlight independent events simultaneously (Shaner et al., 2005).

By attaching the cDNA encoding a FP to the cDNA of a peptide or protein of interest, specific labeling is achieved. One particular application in cell biology is to use FP fusions as fluorescent markers that highlight structures and compartments in cells. These structures can be visualized with fluorescence microscopy. By employing markers with different emission colors (Rizzuto et al., 1996), the relative location of several structures can be studied (e.g. location of ER relative to mitochondria). Another application of markers is to identify the residency of a biomolecule with an uncharacterized location.

Regardless of the application, it is crucial to use markers that show specific, crisp labeling, without spurious, non-specific localization. In addition, it is practical to have a palette of bright fluorescent markers, to facilitate co-imaging. Here, we report on our ongoing efforts to generate robust and bright genetically encoded fluorescent markers for highlighting structures and compartments in living cells.

Results

Previously, we have published a number of markers that highlight a variety of structures and compartments (Bindels et al., 2017; Goedhart et al., 2012). However, we noted that some of these markers also showed spurious localization of the fluorescent protein, i.e. unrelated to the compartment of interest. Below we describe the results of (i) improvements of exisiting markers and (ii) construction of new combinations of existing markers with established bright fluorescent proteins.

Our choice of FPs is based on (i) high cellular brightness in a spectral class, (ii) good photostability and (iii) monomeric behavior in cells. Based on these criteria, we use the FPs mTurquoise2 (cyan), mNeonGreen (green) and mScarlet-I (red). In the future, this palette of fluorescent proteins may be complemented with other variants, including large Stokes-shift FPs (Shcherbakova et al., 2012) or IR emitting proteins (Shcherbakova et al., 2016) to allow for an even broader range of options.

Mitochondria

We observed that the mitochondrial marker, comprising mTurquoise2 and a signal peptide from COX8A (Goedhart et al., 2012), showed substantial cytoplasmic labeling. To increase mitochondrial targeting we replaced the single sequence by a 4 times repeated mitochondrial target sequence (mts) derived from 4mtD3cpv. Co-expression of the 1xmts and 4xmts variant show a crisp labeling of the mitochondria for the 4xmts variant, with hardly any cytoplasmic fluorescence (**Figure 1**).

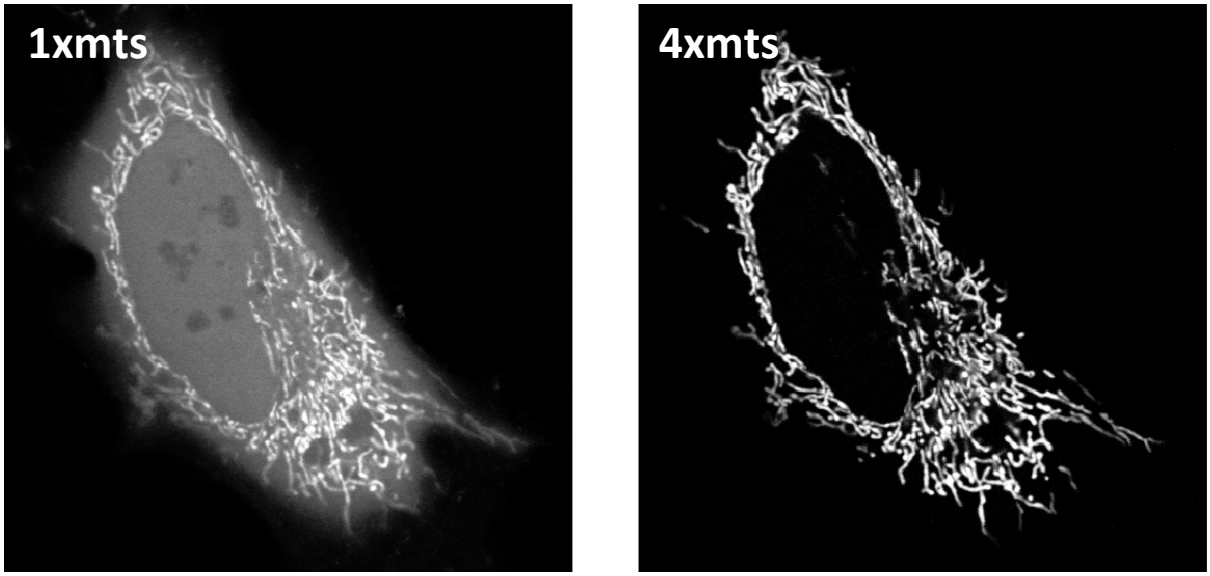


Figure 1: HeLa cell co-transfected with the plasmids encoding a fluorescent protein tagged with 1xmts and 4xmts to label mitochondria. The width of the images is 71 µm.
(Left: 1xmts-mTurquose2, right: 4xmts-mScarletI)

Nucleus

To target proteins to the nucleus, a strong nuclear localization signal (nls) is usually attached to the protein of interest. As fluorescent proteins are small enough to move out of the nucleus through the nuclear pores, substantial cytoplasmic fluorescence is observed for nls-mTurquoise. To improve nuclear localization, a triple nls, 3xnls (Joosen et al., 2014), sequence was employed. The variant with 3xnls shows increased nuclear labeling relative to cytoplasmic labeling when compared to a fluorescent protein with 1xnls (**Figure 2**). We also noted that the 3xnls variant shows enhanced labeling of nucleoli.

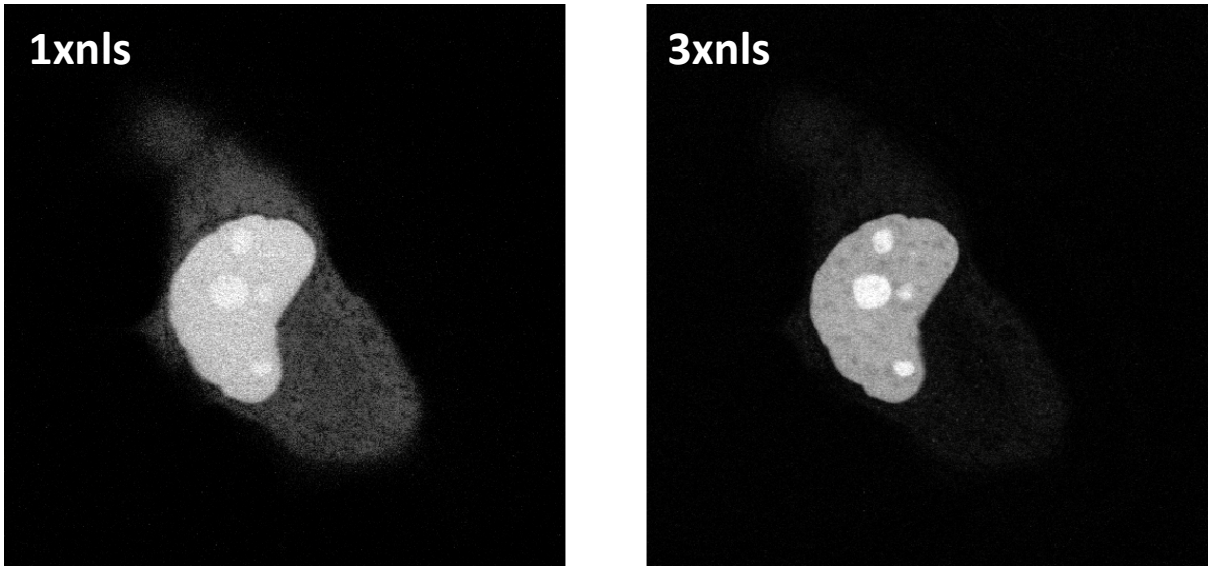


Figure 2: HeLa cell co-transfected with the plasmids encoding a fluorescent protein tagged with 1xnls and 3xnls to highlight the nucleus. The width of the images is 71 µm.
(Left: 1xnls-mTurquoise2, right: 3xnls-mScarletI).

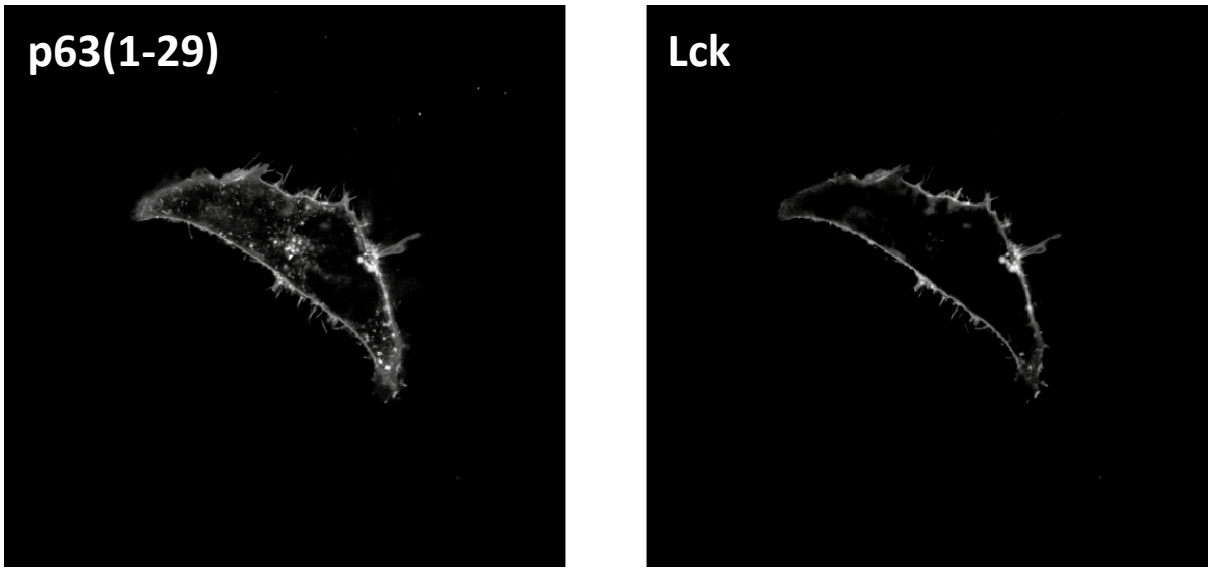


Figure 3: HeLa cell co-transfected with the plasmids encoding a fluorescent protein tagged with the plasma membrane targetin sequence derived from p63(1-29) and Lck. The width of the images is 108 µm.
(Left: p63(1-29)-mTurquoise2, right: Lck-mScarletI).

Plasma membrane

The cysteine-rich sequence of p63RhoGEF was previously used to highlight the plasma membrane (Goedhart et al., 2012). In addition to plasma membrane located fluorescence, we noticed numerous intracellular puncta (**figure 3**, left panel). To create a marker for the plasma membrane with improved selectivity, we explored different peptides. Among the different sequences tested (derived from Lck, Gap43, Src, KRas), we observed the most efficient membrane labeling for the motif from Lck, with hardly any labeling of

endomembranes (**figure 3**). Therefore, we recommend to use the Lck-tagged fluorescent proteins as a robust marker for the plasma membrane.

Golgi apparatus

Initially, we used a sequence derived from GalT to make a Golgi marker (Goedhart et al., 2012). To evaluate whether this is the optimal marker for Golgi labeling we co-expressed this variant with fluorescent proteins tagged with sequences of Golgi resident proteins, i.e. Giantin (Komatsu et al., 2010) and Sialyltransferase. When introduced into HeLa cells, the fusions with giantin showed labeling that is confined to the Golgi apparatus, whereas the GalT marker showed labeling of punctuate structures, in addition to labeling of the Golgi apparatus (**figure 4**). Therefore, the giantin fusion is the marker of choice for the Golgi apparatus.

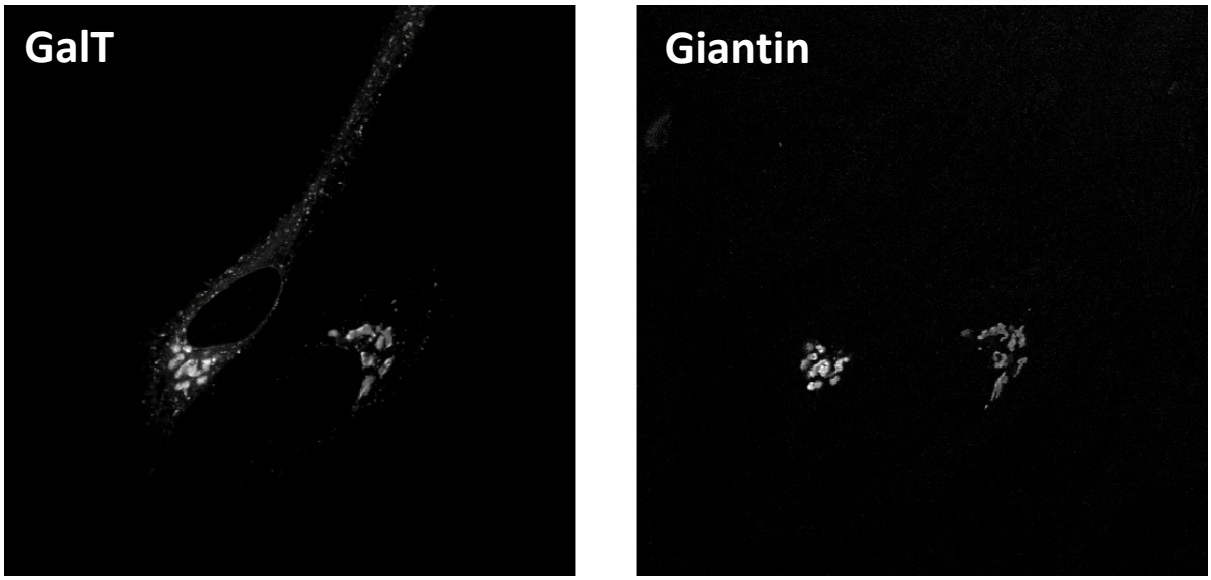


Figure 4: HeLa cell co-transfected with the plasmids encoding a fluorescent protein tagged with GalT and Giantin to mark the Golgi apparatus. The width of the images is 143 µm. (Left: GalT(1-61)-mTurquoise2, right: mScarletI-Giantin(3131-3259)).

Lysosomes

Thus far, no lysosomal marker fused with either mTurquoise2 or mScarlet-I was available. To construct a lysosomal marker we used the lysosomal associated membrane protein 1 (LAMP1). This fusion protein is expected to decorate lysosomes and indeed, we observed labeling of punctuate structures (**figure 5**)

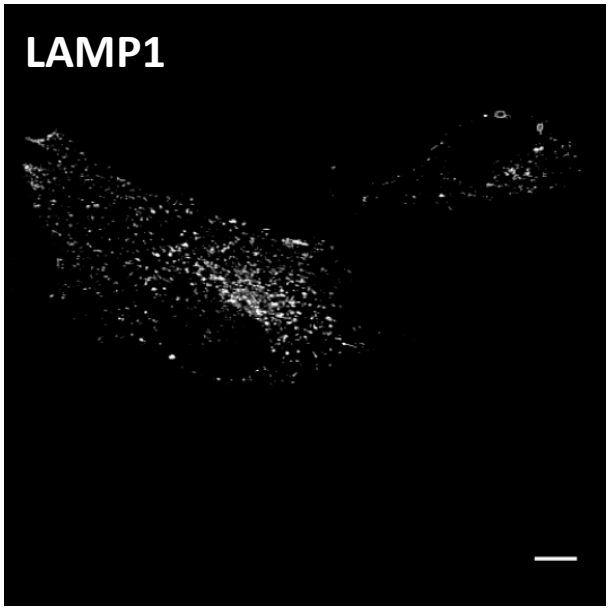


Figure 5: HeLa cell transfected with a plasmid encoding a fluorescent protein tagged with LAMP1 to mark the lysosomes. The width of the images is 143 µm.
(mScarletI-LAMP1)

Actin cytoskeleton

The actin cytoskeleton can be labeled by tagging actin itself with GFP. However, this results in labeling both the soluble, monomeric G-actin pool and the polymerized, filamentous actin (F-actin) pool. To study the cytoskeleton, it is often desired to selectively label actin filaments. To this end, peptide sequences or protein domains have been identified that preferentially bind F-actin, such as lifeact (Riedl et al., 2008), Utrophin-CH (Burkel et al., 2007) and ITPKA(9-52)/F-tractin (Johnson and Schell, 2009). Previously, we have described lifeact fusions with mTurquoise2 (Goedhart et al., 2012) or mScarlet (Bindels et al., 2017). To allow for a wider choice of markers, we made combinations of the bright fluorescent proteins mTurquoise2, mNeonGreen and mScarlet-I with the actin probes Lifeact, Utrophin-CH and ITPKA (residue 9-40). Note that we used a shorter version of Ftractin here (the original Ftractin probes corresponds with residue 9-52 from rat ITPKA, while here we use residue 9-40), without any noticeable difference. All three fusion constructs show pronounced labeling of the actin cytoskeleton (**figure 6**). It is of note that different markers may label different F-actin pools and may have different properties (Belin et al., 2014; Ganguly et al., 2015; Spracklen et al., 2014). Therefore, validation in the system of choice is warranted and the choice of the marker depends on the application.

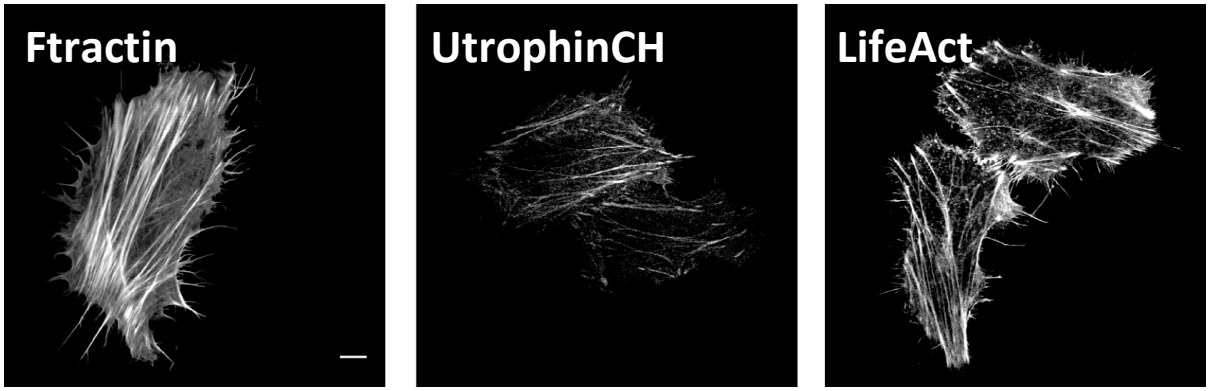


Figure 6: HeLa cell transfected with a plasmid encoding a fluorescent protein tagged with ITPKA or UtrophinCH or Lifeact to visualize the F-actin cytoskeleton. The width of the images from left to right is 116 µm, 143 µm and 143 µm.
(From left to right: ITPKA(9-40)-mTurquoise2, mNeonGreen-UtrCH, Lifeact-mScarletI).

Microtubules

To label the microtubules, FPs can be fused either to the N-terminus of alpha-tubulin (Bindels et al., 2017; Goedhart et al., 2012) or to the C-terminus of beta-tubulin. The protein level of fusion tubulin should be carefully controlled for optimal results. Too much expression leads to higher cytoplasmic signals from unpolymerized tubulin, and also affects the microtubule dynamics. Detailed analyses on the effects by tubulin fusion protein expression are reported with fission yeast (Snaith et al., 2010).

The microtubule binding domain of microtubule associated proteins (MAPs) such as EMTB can also used as marker for microtubules, but these have a tendency to facilitate stabilization and bundling of microtubules.

Another class of microtubule binding proteins, +-TIPs that bind preferentially to the growing plus-ends of the microtubules can be used as the marker for microtubule growth (Stepanova et al., 2003). Among them, EB1 or EB3 fused with an FP is widely used. We observed "comet"-like staining of the growing ends of microtubules with both EB1 (not shown) and EB3 based fusion constructs of mTurquoise2 (**Figure 7**), mNeonGreen, and mScarlet-I. Their brightness and the photostability along with the rapid exchange of the bound populations with the cytoplasmic pool enabled live imaging of microtubule growth at sufficiently low expression levels. A low concentration of the marker is essential for correct measurement of the microtubule dynamics. Higher expression levels can cause hyperstabilization of the microtubules, lower contrast to the background signals from the cytoplasmic pool. At even higher expression levels, their signals extend from the tips to the body of the microtubules.

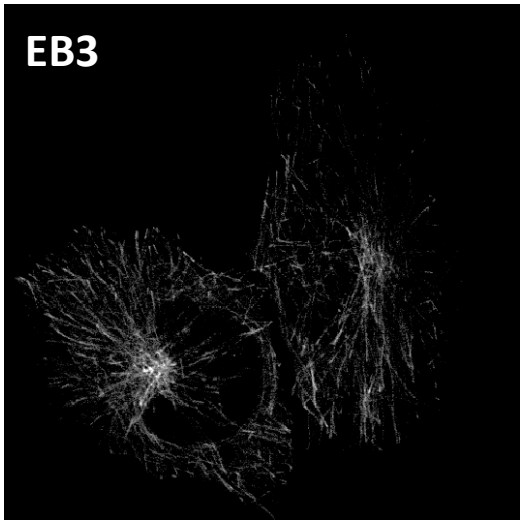


Figure 7: HeLa cell transfected with a plasmid encoding a fluorescent protein tagged with EB3 to detect the growing tips of microtubuli (EB3-mTurquoise2). The width of the image is 86 µm.

Conclusion

Fluorescent markers are essential components of the cell biologist's toolkit. Several of the previously reported markers show non-specific, suboptimal labeling of structures or compartments. Here, we report markers with improved labeling of the structure or compartment of interest with reduced non-specifc labeling. In addition, we have generated new combinations of markers with the brightest monomeric fluorescent proteins currently available. Our analysis of the labeling patterns and improvements is limited to observations after transient transfection in HeLa cells. It is highly recommended that independent verification of the labeling quality is performed. This is of particular importance, if the markers are intended for different cell lines or organisms.

To conclude, we anticipate that the improved markers will be valuable for future live cell imaging studies.

Materials and Methods

Plasmids

The bright monomeric fluorescent proteins used in this study, mTurquoise2 (Goedhart et al., 2012), mNeonGreen (Shaner et al., 2013) and mScarlet-I (Bindels et al., 2017) have been described previously. Plasmids with Lck (van Unen et al., 2016) and Giantin (Bindels et al., 2017; van Unen et al., 2015) have been described before.
Plasmids with mCherry-LaminB (Bas van Steensel), EB3-EGFP-N1 (Niels Galjart) EGFP-LAMP1 (Eric Reits) were kind gifts.
Plasmids with pcDNA-4mtD3cpv (addgene Plasmid #36324), GFP-UtrCH (addgene plasmid #26737), 3xnls (addgene plasmid #60492) were from www.addgene.org

In many cases, existing plasmids were used to generate new combinations by exchanging the fluorescent protein in the clontech-style plasmid that encodes the marker (usually using the restriction enzymes AgeI/BsrGI), with the following exceptions:

The 4mts-mTurquoise2 is made by a partial digest of pcDNA-4mtD3cpv (Addgene Plasmid #36324) with HindIII and the insert is ligated in mTurquoise2 cut with HindIII.

The Utrophin CH domain is excised from the plasmid GFP-UtrCH (addgene plasmid #26737) with BsrGI and SacI and ligated in mTurquoise2-C1 cut with BsrGI and SacI.

LAMP1-mTurquoise2 was generated by a PCR on LAMP1-GFP (primer sequences: forward, 5'-AGCGTCGACATGGCGGCCCCCG-3' and reverse, 5' AGCGGATCCTTGATAGTCTGGTAGCCTGCGT-3'), digestion SalI and BamHI and ligation in the pmTurquoise2-N1 plasmid, cut with the same enzymes.

Detailed sequence information can be accessed at www.addgene.org or requested from the corresponding author.

Cell culture & Sample Preparation

HeLa cells (American Tissue Culture Collection: Manassas, VA, USA) were cultured using Dulbecco's Modified Eagle Medium (DMEM) supplied with Glutamax, 10% FBS (without antibiotics). Cell culture, transfection and live cell microscopy conditions were previously described (REF). Transfection was

performed with lipofectamine 2000, using 1-2 µl of lipofectamine, 50 µl of OptiMEM and ~500 ng of plasmid DNA per 24 mm coverslip. Alternatively, we used PEI, with 4.5 µl of PEI, 200 µl of OptiMEM and ~ 500 ng of plasmid DNA per 24 mm coverslip.

Microscopy

Images were acquired on a Zeiss LSM META confocal microscope with settings as follows. For mTurquoise2 we used 458 nm excitation and a BP480-520 nm emission filter. For mNeonGreen we used 488 nm excitation and a BP500-530 nm emission filter. For mScarletI we used 543 nm excitation and a LP560 nm.
Alternatively we use a Nikon A1 confocal microscope with settings as follows.
For mTurquoise2 we used 458 nm excitation and a 482/35 nm emission filter.
For mScarlet-I we used 561 nm excitation and a BP 595/50 nm emission filter.
All images were acquired with the pinhole set to 1 airy unit.

Availability of materials

Plasmid DNA encoding the improved markers described in this study will be distributed by Addgene if possible, together with full plasmid sequences and additional information at https://www.addgene.org/Dorus_Gadella/

Acknowledgments

The cDNA encoding mCherry-LaminB (Bas van Steensel), EB3-EGFP (Niels Galjart) EGFP-LAMP1 (Eric Reits) were kind gifts from colleagues.

Competing interests

The authors declare no competing or financial interests.

Author contributions

All authors participated in this study, either by generating plasmids with markers, fluorescence microscopy or writing (or a combination thereof).

References

Belin, B.J., Goins, L.M., and Mullins, R.D. (2014). Comparative analysis of tools for live cell imaging of actin network architecture. Bioarchitecture *4*, 189–202.

Bindels, D.S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K.E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M.A., et al. (2017). mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat. Methods *14*, 53–56.

Burkel, B.M., von Dassow, G., and Bement, W.M. (2007). Versatile fluorescent probes for actin filaments based on the actin-binding domain of utrophin. Cell Motil. Cytoskeleton *64*, 822–832.

Ganguly, A., Tang, Y., Wang, L., Ladt, K., Loi, J., Dargent, B., Leterrier, C., and Roy, S. (2015). A dynamic formin-dependent deep F-actin network in axons. J. Cell Biol. *210*, 401 LP-417.

Goedhart, J., von Stetten, D., Noirclerc-Savoye, M., Lelimousin, M., Joosen, L., Hink, M.A., van Weeren, L., Gadella, T.W.J., and Royant, A. (2012). Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. Nat. Commun. *3*, 751.

Johnson, H.W., and Schell, M.J. (2009). Neuronal IP3 3-Kinase is an F-actin–bundling Protein: Role in Dendritic Targeting and Regulation of Spine Morphology. Mol. Biol. Cell *20*, 5166–5180.

Joosen, L., Hink, M.A., Gadella, T.W.J., and Goedhart, J. (2014). Effect of fixation procedures on the fluorescence lifetimes of Aequorea victoria derived fluorescent proteins. J. Microsc. *256*, 166–176.

Komatsu, T., Kukelyansky, I., McCaffery, J.M., Ueno, T., Varela, L.C., and Inoue, T. (2010). Organelle-specific, rapid induction of molecular activities and membrane tethering. Nat Meth *7*, 206–208.

Prasher, D.C., Eckenrode, V.K., Ward, W.W., Prendergast, F.G., and Cormier, M.J. (1992). Primary structure of the Aequorea victoria green-fluorescent protein. Gene *111*, 229–233.

Riedl, J., Crevenna, A.H., Kessenbrock, K., Yu, J.H., Neukirchen, D., Bista, M., Bradke, F., Jenne, D., Holak, T.A., Werb, Z., et al. (2008). Lifeact: a versatile marker to visualize F-actin. Nat Methods *5*, 605–607.

Rizzuto, R., Brini, M., Giorgi, F. De, Rossi, R., Heim, R., Tsien, R.Y., and Pozzan, T. (1996). Double labelling of subcellular structures with organelle-targeted GFP mutants in vivo. Curr. Biol. *6*, 183–188.

Shaner, N.C., Steinbach, P.A., and Tsien, R.Y. (2005). A guide to choosing fluorescent proteins. Nat Methods *2*, 905–909.

Shaner, N.C., Lambert, G.G., Chammas, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M., et al. (2013). A bright monomeric green fluorescent protein derived from Branchiostoma lanceolatum. Nat. Methods *10*, 407–409.

Shcherbakova, D.M., Hink, M.A., Joosen, L., Gadella, T.W., and Verkhusha, V. V (2012). An orange fluorescent protein with a large Stokes shift for single-excitation multicolor FCCS and FRET imaging. J Am Chem Soc *134*, 7913–7923.

Shcherbakova, D.M., Baloban, M., Emelyanov, A. V, Brenowitz, M., Guo, P., and Verkhusha, V. V (2016). Bright monomeric near-infrared fluorescent proteins as tags and biosensors for multiscale imaging. Nat. Commun. *7*, 12405.

Snaith, H.A., Anders, A., Samejima, I., and Sawin, K.E. (2010). New and Old Reagents for Fluorescent Protein Tagging of Microtubules in Fission Yeast. Methods Cell Biol. *97*, 147–172.

Spracklen, A.J., Fagan, T.N., Lovander, K.E., and Tootle, T.L. (2014). The pros and cons of common actin labeling tools for visualizing actin dynamics during Drosophila oogenesis. Dev. Biol. *393*, 209–226.

Stepanova, T., Slemmer, J., Hoogenraad, C.C., Lansbergen, G., Dortland, B., De Zeeuw, C.I., Grosveld, F., van Cappellen, G., Akhmanova, A., and Galjart, N. (2003). Visualization of Microtubule Growth in Cultured Neurons via the Use of EB3-GFP (End-Binding Protein 3-Green Fluorescent Protein). J. Neurosci. *23*, 2655 LP-2664.

van Unen, J., Reinhard, N., Yin, T., Wu, Y., Postma, M., Gadella, T., and Goedhart, J. (2015). Plasma membrane restricted RhoGEF activity is sufficient for RhoA-mediated actin polymerization. Sci. Rep. *5*, 14693.

van Unen, J., Yin, T., Wu, Y.I., Mastop, M., Gadella, T.W.J., and Goedhart, J. (2016). Kinetics of recruitment and allosteric activation of ARHGEF25 isoforms by the heterotrimeric G-protein Gαq. Sci. Rep. *6*, 36825.

Supplemental information

Amino acid sequences of the targeting peptides or protein domains

1xmts is residue 1-29 of HsCOX8A:
MSVLTPLLLRGLTGSARRLPVPRAKIHSL

4xmts:
MSVLTPLLLRGLTGSARRLPVPRAKIHSLGDPMSVLTPLLLRGLTGSARR
LPVPRAKIHSLGKLATMSVLTPLLLRGLTGSARRLPVPRAKIHSLGDPMS
VLTPLLLRGLTGSARRLPVPRAKIHSLG.

1xnls is residue 126-132 of the SV40 large T antigen:
PKKKRKV

3xnls is:
MGLRSRADPKKKRKVDPKKKRKVDPKKKRKV

Lck is residue 1-10 of MmLck:
MGCVCSSNPE

p63(1-29) from Hsp63RhoGEF is:
MRGGHKGGRCACPRVIRKVLAKCGCCFAR

GalT has residue 1-61 of HsB4GALT:
MRLREPLLSGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLSRL
PQLVGVSTPLQ

Giantin is residue 3131-3259 of HsGiantin:
EPQQSFSEAQQQLCNTRQEVNELRKLLEEERDQRVAAENALSVAEEQIRR
LEHSEWDSSRTPIIGSCGTQEQALLIDLTSNSCRRTRSGVGWKRVLRSLC
HSRTRVPLLAAIYFLMIHVLLILCFTGHL

IPTKA also known as **Ftractin** is residue 9-40 of Rat IPTKA:
MGMARPRGAGPCSPGLERAPRRSVGELRLLFEA

UtrCH is residue 1-261 from HsUtrophin

Lifeact is residue 1-17 from Abp140 (Saccharomyces cerevisiae):
MGVADLIKKFESISKEE