

# Synonymous modification results in high-fidelity gene expression of repetitive protein and nucleotide sequences

Bin Wu,<sup>1,2,3</sup> Veronika Miskolci,<sup>1,3</sup> Hanae Sato,<sup>1</sup> Evelina Tutucci,<sup>1</sup> Charles A. Kenworthy,<sup>1</sup> Sara K. Donnelly,<sup>1,2</sup> Young J. Yoon,<sup>1</sup> Dianne Cox,<sup>1,2</sup> Robert H. Singer,<sup>1,2</sup> and Louis Hodgson<sup>1,2</sup>

<sup>1</sup>Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA;

<sup>2</sup>Gruss-Lipper Biophotonics Center, Albert Einstein College of Medicine, Bronx, New York 10461, USA

Repetitive nucleotide or amino acid sequences are often engineered into probes and biosensors to achieve functional readouts and robust signal amplification. However, these repeated sequences are notoriously prone to aberrant deletion and degradation, impacting the ability to correctly detect and interpret biological functions. Here, we introduce a facile and generalizable approach to solve this often unappreciated problem by modifying the nucleotide sequences of the target mRNA to make them nonrepetitive but still functional (“synonymous”). We first demonstrated the procedure by designing a cassette of synonymous MS2 RNA motifs and tandem coat proteins for RNA imaging and showed a dramatic improvement in signal and reproducibility in single-RNA detection in live cells. The same approach was extended to enhancing the stability of engineered fluorescent biosensors containing a fluorescent resonance energy transfer (FRET) pair of fluorescent proteins on which a great majority of systems thus far in the field are based. Using the synonymous modification to FRET biosensors, we achieved correct expression of full-length sensors, eliminating the aberrant truncation products that often were assumed to be due to nonspecific proteolytic cleavages. Importantly, the biological interpretations of the sensor are significantly different when a correct, full-length biosensor is expressed. Thus, we show here a useful and generally applicable method to maintain the integrity of expressed genes, critical for the correct interpretation of probe readouts.

[*Keywords:* MS2; biosensor; single molecule; genome integrity; retroviral reporter]

Supplemental material is available for this article.

Received January 26, 2015; revised version accepted March 18, 2015.

Repeated nucleotide or protein sequences occur frequently in nature or in engineered constructs. For example, the long terminal repeats of the human immunodeficiency virus (HIV) genome are indispensable for viral reproduction (Barre-Sinoussi et al. 2013). The repetitive telomeres at the end of chromosomes are crucial to maintain chromosome integrity (O’Sullivan and Karlseder 2010). In bioengineering, repeats are often employed for signal amplification. For instance, multiple MS2-binding sites (MBSs) are incorporated into RNA to visualize a single transcript (Bertrand et al. 1998). Tandem repeats of LacO are integrated into a genome to mark the genomic locus (Robinett et al. 1996). Recently, the target peptides of a single-chain variable fragment, GCN4, were multimer-

ized for imaging a single protein or attracting many activators to a single endogenous gene targeted by CRISPR (SunTag) (Tanenbaum et al. 2014). In the field of fluorescent biosensors, fluorescent proteins (FPs) with different wavelengths and optophysical characteristics (Campbell et al. 2002; Zhang et al. 2002; Shaner et al. 2004) are incorporated into a single peptide to achieve fluorescent resonance energy transfer (FRET). These FPs share substantial sequence homology, since they evolved from a few parental species, including the *Aequorea victoria* GFP (green FP) and *Discosoma* sp. dsRed. While the repeated nucleotide or peptide sequences are required components of these probes and biosensors, they are notoriously difficult to construct and maintain for stable

<sup>3</sup>These authors contributed equally to this work.

Corresponding authors: robert.singer@einstein.yu.edu, louis.hodgson@einstein.yu.edu

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.259358.115>.

© 2015 Wu et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

integration into the cellular genome. They are easily deleted or truncated, which occurs randomly and is not a priori predictable, resulting in anomalous gene expression. Therefore, a facile and generalizable method to correctly maintain and express these engineered constructs becomes absolutely crucial for these state-of-the-art sensor and probe approaches currently employed in the field.

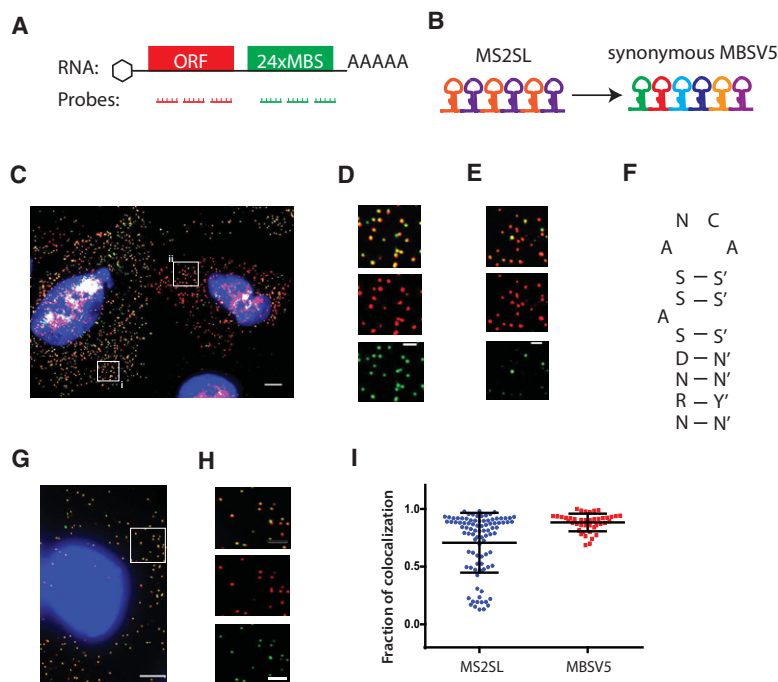
The deletion of repetitive sequences occurs frequently during bacterial amplification, homologous recombination, and viral transduction and in plasmid transfections. It happens most prominently during retroviral transduction, for which the mechanism is well documented (Delviks and Pathak 1999; An and Telesnitsky 2002). Depending on the length of and distance between the repeats, the deletion rate can reach 90% (Delviks and Pathak 1999). As the retroviral transductions can target with great efficiency both dividing and nondividing cells, including stem cells and neurons (Naldini et al. 1996; May et al. 2000), both in vitro and in vivo, it is critical to side-step this random deletion problem. Similarly, random deletion can manifest during other routine gene deliveries, including homologous recombination and transient transfections of plasmids followed by drug selection (Hocine et al. 2013). Previously, irregular repeated sequences have been developed to increase genomic stability in the bacteria (Lau et al. 2003). Here, we describe a simple and generally applicable method to solve the problem of recombinant deletion by modifying the repeats with nonrepetitive but functional (“synonymous”) sequences. By designing nonrepeating sequences and codons, the homology between the components is substantially reduced, which prevents the deletion through recombination. The key here is synonymy: the functional role of the “repeating” sequence is kept intact. We illustrate the broad appli-

cability of this approach by demonstrating it in several unrelated systems. First, we designed a set of nondegenerate MBSs and tandem coat proteins that significantly enhanced the signals and uniformity for single-RNA imaging in live cells. Next, we developed a general strategy for repetitive protein sequences by using synonymous codons. We applied the method to FRET biosensors. It drastically increases the fidelity of the expressed sensors and significantly influences the biological interpretation of experimental readouts. Thus, the method that we describe here should be routinely employed in the generation of probes and sensors containing multiple, repetitive motifs to achieve correct expression profiles.

## Results

### Developing a cassette of nondegenerate MBSs

In single-molecule RNA imaging with the MS2 system, repeated RNA motifs are used to amplify the signals on RNA (Bertrand et al. 1998). Briefly, a genetically encoded MBS derived from bacterial phage is inserted into the gene of interest, and a separate MS2 coat protein (MCP) fused to a FP is introduced into the same cell. MCP-FP binds to the MBS and labels the target mRNA. To increase the signal to noise ratio (SNR) of the mRNA, it is essential that multiple MBSs are incorporated into the mRNA such that its brightness exceeds that of the free MCP-FP. Typically, 24xMBSs are sufficient to image reporter mRNAs with single-molecule sensitivity. Lentiviral vectors are ideal tools to stably express the MS2 reporter mRNA in a variety of cell types. We constructed reporter mRNAs in a lentiviral backbone with 24xMBSs inserted in the 3′ untranslated region (Fig. 1A). The current version of the MBS



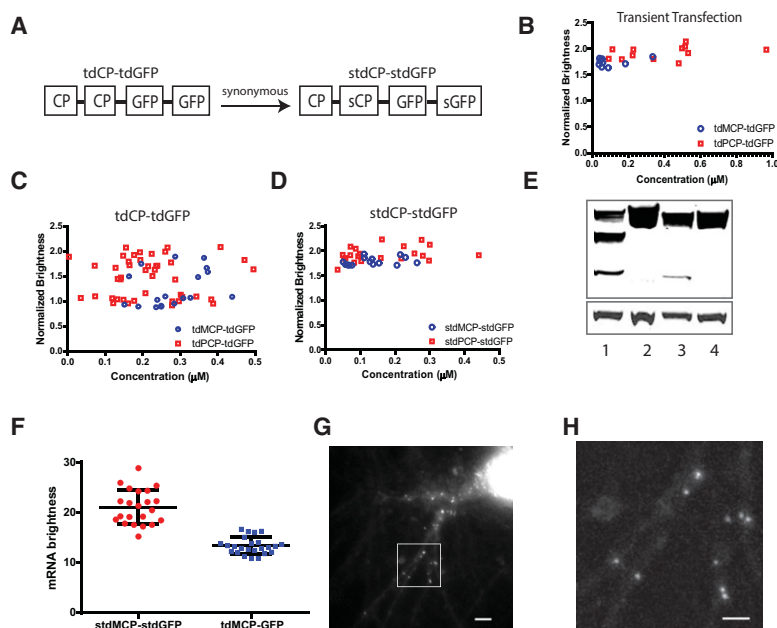
**Figure 1.** Synonymous transformation of MBSs eliminates deletion of repeats in viral transduction. (A) Schematic of the reporter RNA construct. 24xMBSs were inserted in the 3′ untranslated region. The FISH probes for the ORF (red) and MBS (green) are shown at the bottom. (B) MS2SL is the current version of MS2, which contains repeats of two alternating different MBSs. We constructed a new version of MBS, MBSV5, which contains 24 different MBSs with random linkers in between. (C) FISH image of ORF-24xMS2SL. (Red) ORF; (green) MS2SL. The RNA in the *left* cell had both green and red signals, while most RNAs in the *right* cell had only red signal. Boxes i and ii are enlarged in D and E, respectively. (D,E, top panel) Merged. (D,E, middle) MS2SL channel. (D,E, bottom) ORF channel. (F) The consensus sequence of MBS determined by SELEX experiment (Schneider et al. 1992). S represents a C or G; D represents an A, G, or U; R represents an A or G; and Y represents a C or U. (G) FISH image of ORF-24xMBSV5. (Red) ORF; (green) MBSV5. The box was enlarged in H. (H, top panel) Merged. (H, middle) MBSV5 channel. (H, bottom) ORF channel. (I) Quantification of FISH image. The fraction of red spots (ORF) that contained green spots (MBS) was measured in each cell. Bars: C,G, 5  $\mu$ m; D,E,H, 2  $\mu$ m.

(MS2SL; Addgene plasmid no. 31865) contains repeats of two alternating stem-loops (Fig. 1B). We applied single-molecule fluorescence in situ hybridization (FISH) (Femino et al. 1998; Raj et al. 2008) to evaluate the integrity of the mRNA by using probes against the ORF (Fig. 1A, red) and the MBS region (Fig. 1A, green), respectively. Full-length mRNA will have both the green and red FISH signals colocalized. Transiently transfected cells showed that the ORF and MBS were colocalized, indicating that the mRNA were full-length (Supplemental Fig. S1). However, when the reporters were stably expressed by lentiviral infection, the FISH images indicated cell heterogeneity. In Figure 1C, the mRNAs in the cell on the left side of the image are positive for both colors. An enlarged portion of the cell (Fig. 1C, box i) is shown in Figure 1D. On the contrary, the cell on the right in Figure 1C shows mostly red (ORF channel), indicating that the MS2SL had been deleted in this cell. A square in this cell (Fig. 1C, box ii) is enlarged in Figure 1E. To quantify the heterogeneity of the stably integrated reporters, we measured the fraction of green mRNAs (MBS) colocalized with red mRNAs (ORF) (fraction of colocalization in each cell). For MS2SL, the fraction among cells ranges from 10% to >90% (Fig. 1H). Furthermore, the intensities of MS2SL FISH spots in cells with low colocalization are less than in cells with high colocalization (Supplemental Fig. S2), indicating that MS2SL sequences had been partially deleted in the former. We reasoned that the deletion must happen during the viral transduction processes, since transiently transfected mRNA is full-length (Supplemental Fig. S1). In order to circumvent the deletion problem, we designed a new version of MBS to remove the repeated sequences (Materials and Methods). Briefly, we mutated the nonessential nucleotides in the consensus MBS sequence (Fig. 1F), determined from a systematic evolution of ligands by exponential enrichment (SELEX)

experiment (a tetranucleotide loop and a variable length stem with a bulged adenosine) (Schneider et al. 1992). We designed 24 different MBS sequences with random linkers between them. This new version of MBS was named MBSV5. To test this design, we constructed a lentiviral reporter mRNA similar to before (Fig. 1A). A two-color FISH experiment indicates that the majority of mRNAs have both ORF and MBS signals (Fig. 1G,H). Quantitative analysis confirms that all cells tested have fraction of colocalization >70% (the average is 88%) (Fig. 1I). Live-cell imaging shows that a stably expressed mRNA reporter labeled by MBSV5 can be visualized with single-molecule sensitivity (Supplemental Movie 1).

### Synonymous transformed tandemly repeated coat proteins

Repeated sequences also occur frequently in protein-coding regions. For example, MCP binds to MBS as a dimer. However, the dimerization affinity in the cell is low (Wu et al. 2012). We previously constructed tandem coat proteins (tdMCP and tdPCP) that greatly enhanced the SNR and uniformity of the RNA labeling by the MS2 and PP7 systems (Wu et al. 2012). However, one caveat was that the total number of FPs on mRNA was reduced by half, since only one FP was linked to a tdCP. One obvious strategy was to fuse a tandem FP to the tdCP (tdCP-tdFP) (Fig. 2A). Indeed, the brightness of the transiently transfected tdCP-tdGFP measured by fluorescence fluctuation spectroscopy (FFS) (Chen et al. 1999) was twice that of monomeric tdCP-GFP (Fig. 2B). However, the brightness of tdCP-tdGFP when stably expressed by lentiviral infection was heterogeneous (Fig. 2C). Some cells had brightness corresponding to a dimer, and some had monomer brightness, while still others had a mixture of both monomers and dimers. Western blots of the coat protein



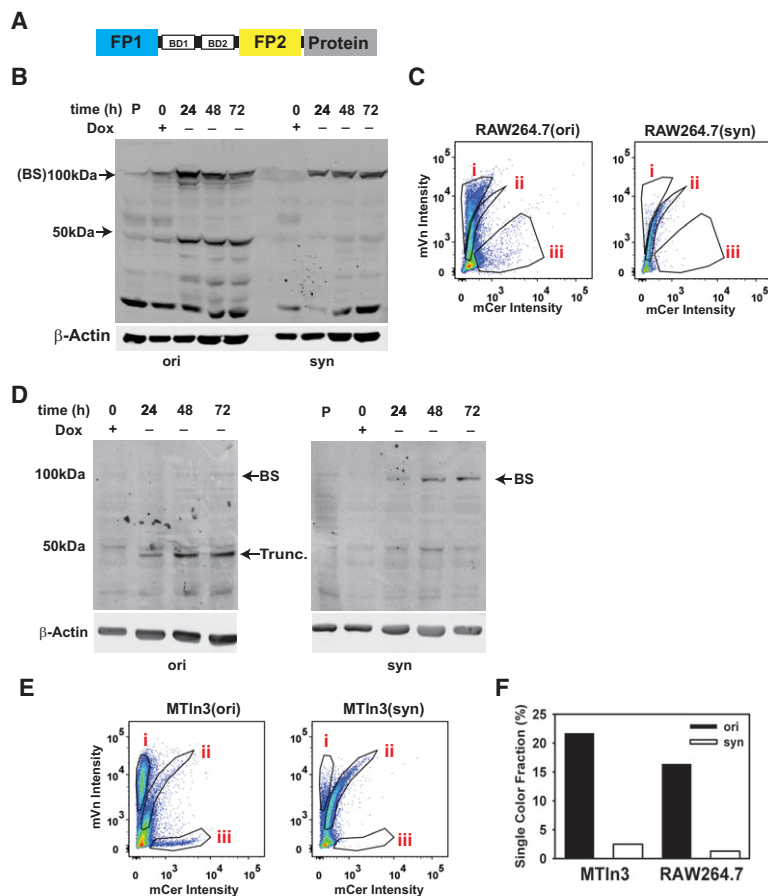
**Figure 2.** Synonymous transformation of CPs. (A) Schematic of the tandem coat protein constructs. (CP) Coat protein; (GFP) green FP; (sCP) synonymous coat protein; (sGFP) synonymous GFP. (B–D) The normalized fluorescence fluctuation brightness of tdPCP-tdGFP (red square, B,C), tdMCP-tdGFP (blue circle, B,C), stdPCP-stdGFP (red square, D), and stdMCP-stdGFP (blue circle, D) transiently transfected (B) or stably expressed by lentiviral infection (C,D) in U2OS cells. (E) Western blot of the stable cells shown in C and D. (Lane 1) tdPCP-tdGFP. (Lane 2) stdPCP-stdGFP. (Lane 3) tdMCP-tdGFP. (Lane 4) stdMCP-stdGFP. (Top panel) Anti-GFP antibody. (Bottom panel) Anti-Actb antibody. Lanes 1 and 3 show truncated protein products, while lanes 2 and 4 are mostly full-length proteins. (F) The stdMCP-stdGFP and mCherry-24xMBSV5 were stably expressed in U2OS, and the mRNA brightness measured by FFS (red circle) was significantly higher than transiently transfected CFP-24xMS2SL and tdMCP-GFP ( $P < 0.0001$ ). (G) Primary hippocampal neurons were infected with lentivirus expressing mCherry-24xMBSV5 and stdMCP-stdGFP. The box is enlarged in H. Bars: G, 5  $\mu$ m; H, 2  $\mu$ m.

tdCP-tdGFP also show multiple bands when detected with GFP antibody (Fig. 2E, lanes 1,3). Notice that fluorescence-activated cell sorting (FACS) will not be able to sort for the “correct” cells because all measured cells are positive for GFP. Since the tdCP-tdGFP has significantly repeated sequences, we inferred that repeated sequences were deleted during the viral infection. The approach to remove the repetitiveness was to employ the redundancy in protein codons. For example, serine could be coded by any one of four codons: UCU, UCC, UCA, or UCG. Therefore, we could design a nonrepeating RNA sequence using synonymous codons while keeping the repetitive amino acid sequence intact (Fig. 2A). In practice, the homology between the synonymous RNA construct and the original sequence was kept below 75%. We directly synthesized the synonymous DNA sequences of both coat protein and GFP. We replaced the repeated coat protein and GFP in tdCP-tdGFP with the synonymous variants, respectively, to obtain stdCP-stdGFP (s indicates synonymous) (Fig. 2A). When stdCP-stdGFP was stably expressed using lentiviral infection in U2OS cells, all cells measured exhibited the brightness expected from a dimer (Fig. 2D). A Western blot also showed a single band of full-length proteins (Fig. 2D, lanes 2,4). When we stably expressed 24xMBSV5 reporter mRNA and stdMCP-stdGFP in U2OS cells, the single mRNA brightness was 1.6 times brighter than the transiently transfected 24xMBS

mRNA labeled with tdMCP-GFP (Fig. 2F), meaning an increase in the SNR. Because viruses are the predominant method for introducing genes into neurons, we stably expressed the MBSV5 reporter mRNA and stdMCP-stdGFP in primary hippocampal neurons. The single-reporter mRNAs could be readily visualized (Fig. 2G,H). Therefore, the MBSV5 and synonymous coat protein constructs make visualization of RNA dynamics in live neurons readily available.

### Synonymous transformation of FRET biosensors

Viral recombination deletion occurs for not only direct repeats but also partially overlapping sequences. FPs with many different wavelengths and optophysical characteristics (Campbell et al. 2002; Zhang et al. 2002; Shaner et al. 2004) were evolved from a few parental species, including the *A. victoria* GFP and *Discosoma* sp. dsRed. Thus, the FPs from the same family differ by only a few point mutations (e.g., the sequence homology between the popular FRET pair of FPs monomeric Cerulean1 [mCer1] and monomeric Venus [mVn] is up to 98%). When FPs of the same family were fused into one peptide, such as in fluorescent biosensors, the resulting RNA sequence homology made it highly susceptible to partial or complete deletion by viral recombination. Indeed, in the biosensor design shown in Figure 3A (Hanna et al. 2014; Moshfegh et al. 2014), two

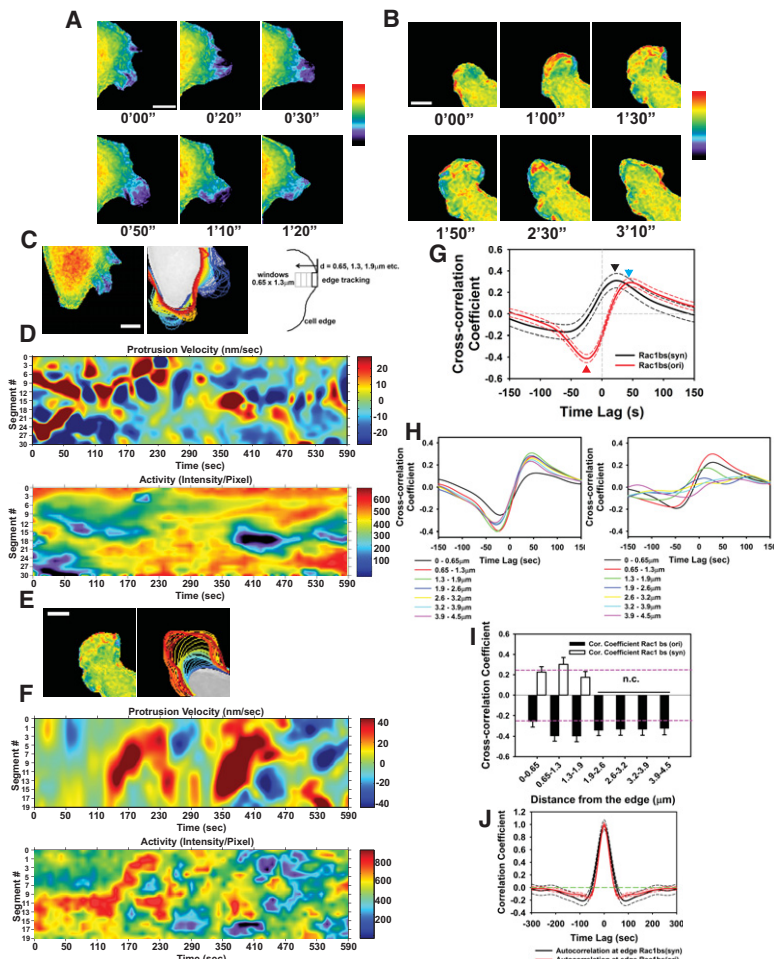


**Figure 3.** Synonymous codon transformation of the FRET biosensor for Rac1 GTPase stabilizes full-length expression in cells. (A) Schematic of the Rac1 GTPase biosensor based on FRET (Moshfegh et al. 2014). (FP1) Monomeric Cerulean 1 FP; (BD1) the wild-type p21-binding domain (PBD) of p21-activated kinase 1 (PAK1); (BD2) the GTPase binding-deficient mutant version of the PBD domain, incorporated for an autoinhibitory modulation of the BD1 domain affinity; (FP2) monomeric Venus FP; (Protein) wild-type Rac1 GTPase. (B) Rac1 biosensor induction and expression profiles in the RAW264.7 cell line. (Dox) Doxycycline at 2  $\mu$ g/mL; (P) untransduced, parental RAW264.7 lysate; (ori) original version of the biosensor; (syn) synonymous modified version of the biosensor. (C) FACS profiles of the RAW264.7 from B with and without synonymous codon modification (syn vs. ori). (D) Rac1 biosensor induction and expression profiles in the MTln3 cell line. (Dox) Doxycycline at 5  $\mu$ g/mL; (P) untransduced, parental MTln3 lysate. In the original version of the sensor (ori), truncated (Trunc.) product is seen at ~50 kDa, which is absent in the synonymous modified version (syn). (E) FACS profiles of the MTln3 cells from D with and without the synonymous codon modification (syn vs. ori). (F) Quantification of the indicated FACS gates from C and E showing changes in the single-positive fluorescent cell populations before and after the synonymous modification. The indicated FACS-gated populations are RAW264.7(ori) (i = 12.0%; ii = 18.4%; iii = 4.37%), RAW264.7(syn) (i = 0.35%; ii = 25.6%; iii = 0.93%), MTln3(ori) (i = 19.5%; ii = 11.9%; iii = 2.18%), and MTln3(syn) (i = 2.25%; ii = 22.8%; iii = 0.22%).

FPs of *A. victoria* origin (mCer1 and mVn) were placed within a single-chain design as well as tandem binding domains BD1 and BD2 to facilitate an autoinhibitory regulation (Hanna et al. 2014; Moshfegh et al. 2014). When stably and inducibly transduced into the RAW264.7 mouse macrophage/monocyte cell line (Kheir et al. 2005) or the MTln3 rat mammary adenocarcinoma cell line (Neri and Nicolson 1981), the biosensors showed a significant number of products of intermediate sizes less than the molecular weight of the full-length peptide (Fig. 3B,D), which was not due to proteolytic cleavage, as protease inhibitor treatment failed to rescue this effect (Supplemental Fig. S3). Moreover, a significant presence of

the single-positive fractions for either the cyan or the yellow fluorescence was detected by FACS from the cell population expressing the original version of the biosensor (Fig. 3C,E, matching Western blots for these cells are shown in B,D). We introduced synonymous modifications to mCer1 and BD2 (Fig. 3A) to obtain ~66% homology with mVn and BD1, respectively (Supplemental Fig. S4). When the synonymous constructs were stably expressed in both cell lines, we achieved a significant improvement in expression of the full-length biosensor (Fig. 3B–F).

This truncation has a direct consequence on the measured protein activities in live-cell imaging (Fig. 4A, B; Supplemental Movies 2, 3). Using morphodynamics



**Figure 4.** Synonymous transformation affects the biological interpretation of the biosensors. (A) Time-lapse image panels of the RAW264.7 cells stably expressing the unmodified original Rac1 biosensor (Moshfegh et al. 2014). Pseudocolor ranges: black = 1.0; red = 1.44. Bar, 5  $\mu$ m. (B) Time-lapse image panels of the RAW264.7 cells stably expressing the synonymous codon modified Rac1 biosensor. Pseudocolor ranges: black = 1.0; red = 1.58. Bar, 5  $\mu$ m. (C) A representative example of a RAW264.7 cell stably expressing the Rac1(ori) biosensor showing a ratio image of an edge region (*left*), the computation edge tracking of the same region (*middle*), and a schematic of the measurement windows constructed at the edge of the protrusion. (D) Morphodynamics mapping of the protrusion edge velocities (*top*) and the corresponding biosensor activities (*bottom*) from the region of a cell edge shown in C. (E) A representative example of a RAW264.7 cell stably expressing the Rac1(syn) biosensor showing a ratio image of an edge region (*left*) and the computation edge tracking of the same region (*right*). (F) Morphodynamics mapping of the protrusion edge velocities (*top*) and the corresponding biosensor activities (*bottom*) from the region of a cell edge shown in E. (G) Morphodynamics mapping (Machacek et al. 2009) was used to obtain the cross-correlation coefficient distribution for the original (red line) versus the synonymous modified (black line) Rac1 biosensor in RAW264.7 macrophages during serum-stimulated random protrusions. Single traces at the distance of 0.65–1.3  $\mu$ m (location of the strongest cross-correlation) are shown to illustrate the difference in profiles. The corresponding dashed lines indicate 95% confidence intervals. The cross-correlation coefficients above and below 0.252 and –0.252, respectively, are significant at  $P < 0.05$ . The

primary peak of the Rac1bs(ori) is at –23.3 sec [–32.5 sec –14 sec] [cross-correlation coefficient value: –0.396 [–0.448 –0.345]; red arrowhead], whereas the Rac1bs(syn) is at +28 sec [7 sec 49 sec] [cross-correlation coefficient value: 0.302 [0.283 0.322]; black arrowhead],  $\pm$ SD limits. The secondary peak of Rac1bs(ori) at +47 sec [30 72] [cross-correlation coefficient value: 0.273 [0.231 0.320]; blue arrowhead] is not significantly different from the primary peak of Rac1bs(syn) at  $P = 0.261$ . Three-hundred-ninety-eight individual windows for Rac1bs(ori) (12 cells) and 362 individual windows for Rac1bs(syn) (13 cells) were analyzed. Statistical limits were calculated as in Machacek et al. (2009). (H) The spatial evolution of cross-correlation traces for the Rac1(ori) biosensor (*left*) and the Rac1(syn) biosensor (*right*), color coded to indicate spatial dependence. (I) Spatial dependence of the cross-correlation coefficient between the edge velocity and the biosensor ratio readouts. The original version of the biosensor consistently gives strong negative correlation, peaking at 0.65–1.3  $\mu$ m from the edge, whereas the synonymous modified biosensor gave more spatially limited coupling but with positive correlations peaking at the same spatial position. (n.c.) No correlation. Pink lines indicate  $P = 0.05$  limits, above and below which the measurements are significant. (J) The autocorrelation functions of the leading edge protrusion velocities in cells expressing the Rac1(ori) versus the Rac1(syn) biosensor. The zero-crossing times between the two conditions show  $P = 0.0626$  to indicate that the protrusion cycling periodicity between these two biosensor expressions are not distinguishable.

analysis (Machacek et al. 2009) to characterize the coupling of Rac1 activity to the leading edge protrusions in macrophages (Fig. 4C–F), we observed significant changes in the cross-correlation functions between the original versus the synonymous modified versions of the Rac1 biosensor (Fig. 4G,H). Rac1 activity patterns and its role in macrophage lamellipodia protrusions have not yet been fully elucidated using biosensors. Based on other studies of Rac1 activities in lamellipodia (Kraynov et al. 2000; Machacek et al. 2009) and the requirement for morphological cohesion of Rac1 in macrophages (Allen et al. 1997; Wheeler et al. 2006), we expected the protrusive coupling of Rac1 activity to depend strongly on the spatial position within the lamellipodia during serum-induced random protrusions. When the original version of the biosensor was used, the morphodynamic coupling was spatially homogenous, with the major peak at an approximately –23.3-sec time lag with significant negative cross-correlation (Fig. 4H,I; Supplemental Movie 2). Using the synonymous modified biosensor, the coupling of Rac1 activity to the protrusion velocities instead became highly dependent on the spatial position within the lamellipodia protrusion, with the maximum coupling occurring at 0.65–1.3  $\mu\text{m}$  close to the leading edge, with a positive 28-sec time lead ahead of the protrusion and a positive cross-correlation, and then falling off thereafter away from the edge (Fig. 4H,I; Supplemental Movie 3). The stable integration and the expression of the original versus the synonymous modified versions of the Rac1 biosensor did not change the basic protrusive dynamics (Fig. 4J). This striking difference in the dynamics of biosensor measurements in macrophage lamellipodia protrusions clearly indicated that synonymous modification of the biosensor had a significant effect on the readouts and thus the biological interpretation. This effect is not likely to be an isolated example simply in a proteolytically active cell type such as the macrophage cell line. Improvements in expression stability of the biosensor were similarly manifested in the rat mammary adenocarcinoma cell line MTLn3 (Fig. 3D,E), suggesting that it is cell type-independent and is essential to ascertain reliable biosensor readouts.

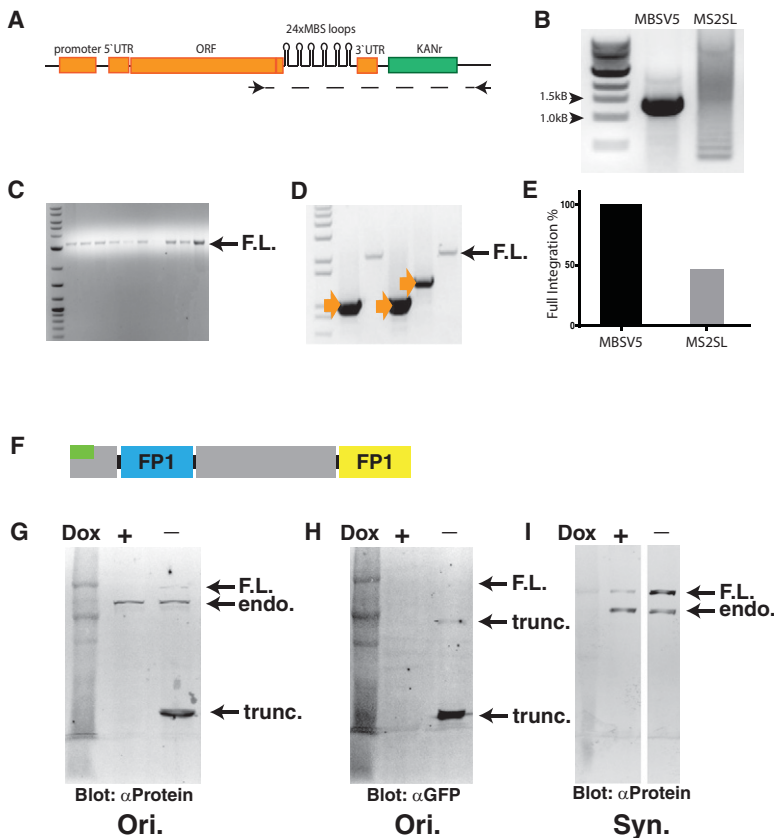
#### *Repetitive sequences in nonviral gene delivery*

We demonstrated the deletion of repetitive sequences in retroviral transduction. However, the deletion due to repetitive sequences is not limited to gene delivery by retrovirus. We show here two examples of this process occurring in systems that are commonly used for stable gene expressions and recombinant genetic manipulations. The first example is during yeast homologous recombination. Yeast is an ideal model system to perform genetic manipulations, since its genome can be easily edited by homologous recombination. To label an endogenous RNA in *Saccharomyces cerevisiae* for live-cell imaging, we attempted to insert 24xMBSs into the 3' untranslated region (UTR) of the target gene (Fig. 5A). Normally, PCR is performed with primers spanning the homology sequence flanking the target genetic sequence. The

MBSV5 fragment was readily amplified by PCR. On the contrary, the PCR of MS2SL was inefficient and resulted in many lower-molecular-weight bands (Fig. 5B). With single-step PCR amplification, the 24xMBSV5 can be inserted into any target gene through homologous recombination. Importantly, 100% of positive clones have full-length insertion (Fig. 5C). Although it is not feasible to PCR-amplify the MS2SL segment directly, it is possible to clone the two homology sequences flanking the MS2SL, which is an extra two-step cloning for each gene to be targeted. Afterward, the MS2SL cassettes and the homology sequences are cut out directly from the plasmid and are used to transform the yeast. Nevertheless, significant loss of repeats was observed during the homologous recombination (Fig. 5D; Hocine et al. 2013). In contrast to 100% full-length insertion for MBSV5, only ~40% of MS2SL insertions maintain full-length (Fig. 5E). The second example is the production of stable cell lines from standard plasmid transfections followed by drug selection. Here, we constructed a 204-kDa molecule containing two FPs: mVn placed at the C terminus and mCer1 placed at amino acid position 260 (Fig. 5F). This construct was inserted into a tet-inducible backbone containing puromycin resistance for selection and was transfected into mouse embryonic fibroblasts (MEF/3T3) containing stable tet-OFF-tTA. Upon establishment of the stable cell population and induction of the protein expression, we saw a prominent, short-molecular-weight band at 55 kDa, detectable using an antibody against the N terminus of the protein (Fig. 5G), while a band of significantly reduced intensity is also visible at 204 kDa, indicating the presence of a full-length engineered protein. The presence of this short fragment at 55 kDa is recapitulated when the same molecule is stably integrated by viral transduction and induced for expression (Fig. 5H). The immunoblot against GFP (Fig. 5H) confirms the presence of FP in this band to suggest a recombinant truncation at the end of the FP1 (mCer1) (Fig. 5F). When the synonymous modified mCer1 is used, only the full-length band was observed (Fig. 5I). This was not an isolated example. We observed truncation in another stably transfected biosensor that contained two FPs flanking a protein domain (Supplemental Fig. S5; Cammer et al. 2009). Therefore, we caution that it is critical to carefully examine the expression profile using Western blots and FACS to characterize populations of cells that express the full-length protein when stably integrating gene cassettes containing repetitive sequences. The new approach presented here dramatically improves the fidelity of full-length gene expression and side-steps this uncertainty by eliminating the random expression profiles that can directly impact data interpretation.

#### **Discussion**

The recombination and deletion of repetitive sequences is a universal phenomenon in retroviral transduction, homologous recombination (Hocine et al. 2013), and stable transfection with drug selection. The synonymous modification introduced here is applicable to both repetitive RNA and peptide sequences to maintain the integrity of



**Figure 5.** Repetitive sequence in nonviral gene delivery. (A) Schematic of targeting yeast mRNA with 24xMBSs in the 3' UTR. (B) PCR amplification of 24xMBSV5 shows correct length on the gel, while amplification of 24xMS2SL exhibits many lower-molecular-weight bands. (C) PCR product of 24xMBSV5 with flanking sequence homology with the 3' UTR of ASH1 was directly used to transform *S. cerevisiae*. Screening of the genomic DNA by PCR shows that 100% of positive clones have full length (F.L.). (D) The 5' and 3' homology sequences of the GAL1 3' UTR were cloned to flanking 24xMS2SL, respectively. The whole construct was cut out by restriction enzyme digestion and gel-purified. Subsequently, it was used to transform the yeast *S. cerevisiae*. The positive clones were genotyped by PCR to measure the size of the insert on a gel. Only ~40% of the positive clones show the full-length insert. The arrows indicate the truncated inserts. Adapted with permission from Hocine et al. (2013). (E) Comparison of the percentage of full-length insertion for MBSV5 and MS2SL. (F) The schematic of an example 204-kDa peptide construct containing two FPs placed so as to flank a substantial portion of the protein. FP1 (mCer1) is placed at amino acid position 260, and FP2 (mVn) is placed at the C terminus at amino acid position 1200. The green segment indicates the location of the epitope for the primary antibody, which can be used to detect the protein. (G) The original (Ori.) example construct as shown in F was transfected into MEF3T3 fibroblasts under a tet-OFF-inducible system and was selected for stable integration using an antibiotic

selection. Removal of doxycycline from the medium resulted in a modest induction of the full-length biosensor (F.L.) and substantially more abundant expression of the truncated version (trunc.) at 55 kDa. The endogenous protein band (endo.) is also visible, detected using a primary antibody against the protein backbone. (H) The original (Ori.) example construct as shown in F was stably transduced using retrovirus into MEF3T3 fibroblasts under a tet-OFF-inducible system and selected for stable integration. Removal of doxycycline from the medium resulted in no detectable induction of the full-length biosensor (F.L.) and expressions of the truncated versions (trunc.) at ~95 kDa and at 55 kDa, detected using an antibody against FP. (I) The synonymous modified (Syn.) example construct as shown in F was stably transduced using retrovirus into MEF3T3 fibroblasts under a tet-OFF-inducible system and selected for stable integration. Removal of doxycycline from the medium resulted in an induction of the full-length biosensor (F.L.). No truncated products are visible. The endogenous protein band (endo.) is also visible, detected using a primary antibody against the protein backbone. MEF3T3 fibroblasts were cultured in 2  $\mu$ g/mL doxycycline to repress the biosensor induction for this data set. At higher levels of doxycycline concentration, the tet-OFF gene repression is more complete; however, it can result in minor toxicity and difficulty in achieving good induction. Thus, we used 2  $\mu$ g/mL doxycycline to illustrate the expression profile in this data set.

the reporter. As examples, we present two dissimilar areas of quantitative microscopy imaging in which this effect can significantly impact the biological readout and interpretations of data. For repetitive protein sequences, we developed a systematic approach that we termed "synonymous" codon transformation. For repeated nucleotides, it must be performed on a case-by-case basis. For MS2, the consensus sequence has been determined previously through SELEX experiments. A few key nucleotides essential for the binding were kept intact, and the rest could be modified. In addition, the linkers between the binding sites were fully randomized. We expect that this pattern of synonymous modification could be generalized to other RNA- or DNA-binding proteins as well.

In the biosensor field, it has traditionally been assumed that a certain amount of protein degradation is unavoidable due to the size and complexity of the engineered sensor constructs. Since there are many biosensors based on

two compatible FRET pairs of FPs (Miyawaki et al. 1997; Itoh et al. 2002; Kurokawa et al. 2005; Mitra et al. 2005; Kawase et al. 2006; Pertz et al. 2006; Cai et al. 2008; Ouyang et al. 2010; Zawistowski et al. 2013; Hanna et al. 2014; Moshfegh et al. 2014), their expression profiles in stably integrated cell systems must be carefully assessed by immunoblotting and FACS analysis. Indeed, in our previous works, several approaches were taken to minimize potential problems by carefully (1) optimizing the biosensor induction condition and duration, (2) analyzing the expression patterns of the biosensor on Western blots and FACS sorting for limited populations of cells that express both CFP and YFP, and (3) choosing, critically, cells expressing both FPs at near-expected intensities during single-cell microscopy imaging (Hodgson et al. 2010; Spiering et al. 2013). Especially in light of the recombinant deletion frequency being a function of the insert size between the two homologous components such as the FPs

(Delviks and Pathak 1999), those sensors containing long stretches of moieties between two homologous FPs are clearly at greater risk for this type of effect. Our new results presented here address these situations and thus could dramatically reduce the potential for misinterpretation of biosensor readouts due to incorrect expression patterns that are not a priori predictable or well characterized.

One of the broad implications of this study is that it offers a generalizable, system- and platform-independent solution to stabilize the expression of exogenous reporter systems. This includes but is not limited to most genetically encoded single-chain biosensors based on the FP-FRET published thus far, many reporter systems using the tandem signal amplification motifs, and many polycistronic plasmids with repeated motifs. The current strategy is straightforward to implement, with easily affordable gene synthesis solutions. Practically, the homology of the synonymous transformed RNA sequences versus the original should be kept at minimum. We found that the sequence homology of 75% or less is ideal, while taking care to choose the species-specific codons with relative abundance of  $\geq 20\%$ . We expect that facile algorithms can be designed to automate this synonymization procedure in the future for many species and organisms, taking into consideration specific codon optimization strategies that are now available from a wide variety of vendors.

In conclusion, we believe that the approach will find its application in many areas of biology for stable expression of target genes containing repetitive sequences.

## Materials and methods

### *Design and synthesis of the synonymous constructs*

The consensus MBS sequences that bind to the MCP include a tetraloop of the primary sequence of ANCA and a variable-length stem with a bulged adenosine residue at a specific stem position (Schneider et al. 1992). We designed 24 different MBS sequences with random linkers between them. The secondary structure of the whole sequence was predicted using Mfold software (<http://mfold.rna.albany.edu/?q=mfold>). Individual nucleotides were changed until the predicted secondary structure had 24 correct MBSs. We also ensured that there were no stop codons or repeated sequences  $>15$  nucleotides (nt).

For repeated proteins, we exploited the redundancy of the codon usage to remove the repeats in RNA sequence. We first used the Integrated DNA Technology online codon optimization software (<https://www.idtdna.com/CodonOpt>) to choose the codon of a protein in a different organism such as yeast *S. cerevisiae*. The sequence was further modified to keep the homology to  $<75\%$  with the original.

The designed sequences were synthesized by GenScript and subsequently cloned into retroviral vectors.

### *Mammalian cell lines*

Human U2OS cells (American Type Culture Collection [ATCC], HTB-96) were grown at  $37^{\circ}\text{C}$  and  $5\%$   $\text{CO}_2$  in DMEM containing  $4.5\text{ g/L}$  glucose,  $10\%$  FBS, and  $1\%$  penicillin–streptomycin.

RAW264.7 mouse monocyte/macrophage cells (ATCC, TIB-71) (Kheir et al. 2005) were maintained at  $37^{\circ}\text{C}$  and  $5\%$   $\text{CO}_2$  in RPMI containing  $10\%$  newborn calf serum and  $1\%$  penicillin–streptomycin.

MTln3 rat mammary adenocarcinoma cells (Neri and Nicolson 1981) were maintained at  $37^{\circ}\text{C}$  and  $5\%$   $\text{CO}_2$  in  $\alpha$ -MEM containing  $5\%$  FBS and  $1\%$  penicillin–streptomycin.

Mouse embryonic fibroblasts with stably integrated tet-OFF were obtained from Clontech and were grown at  $37^{\circ}\text{C}$  and  $5\%$   $\text{CO}_2$  in DMEM containing  $4.5\text{ g/L}$  glucose,  $10\%$  FBS, and  $1\%$  penicillin–streptomycin.

### *Culture of primary postnatal hippocampal neurons from mice*

Postnatal day 1 mouse hippocampal tissue was isolated from C57BL6 wild-type pups. Hippocampi were placed in  $0.25\%$  trypsin for  $15\text{ min}$  at  $37^{\circ}\text{C}$ . Tissue was triturated and plated onto poly-D-lysine (Sigma)-coated glass-bottom dishes (Matek) at  $75,000$  cells per dish and cultured in Neurobasal A medium supplemented with B-27, GlutaMax, and primocin (InvivoGen). Neurons were imaged between 2 and 3 wk in culture. For live imaging of neurons, cells were washed in prewarmed HBS medium and imaged at  $35^{\circ}\text{C}$ – $37^{\circ}\text{C}$ .

### *Plasmids, transfection, and viral transduction*

All lentiviral constructs for MBS and MCP were cloned into the phage-*ubc*-RIG lentiviral vector (Mostoslavsky et al. 2006) from which the DSRed-IRES-GFP fragment had been excised. For mRNA reporters,  $24\times\text{MS2SL}$  or  $24\times\text{MBSV5}$  was inserted after the stop codon and before the Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE). For coat proteins, tdMCP-GFP, tdMCP-tdGFP, std-stdGFP, tdPCP-GFP, tdPCP-tdGFP, and stdPCP-stdGFP were all cloned to replace DSRed-IRES-GFP, respectively. All of the coat proteins had nuclear localization signals at the N terminus.

The lentiviral plasmids were cotransfected with helper plasmids in 293T cells as described previously (Naldini et al. 1996). Collected lentiviral particles were purified with lenti-X concentrator (Clontech). U2OS cells were infected with the purified lentiviral particles. The positive cells were sorted with FACS.

Tet-inducible RAW264.7 and MTln3 cells were produced using the pRetro-X-tet-OFF-Advanced system following the manufacturer's protocols (Clontech). A fluorescent biosensor for Rac1 was transduced into the tet-OFF RAW264.7 and MTln3 cell lines using the pRetro-X-Zeo and pRetro-X-Puro backbones (Clontech), respectively, and selected for stable transductants as previously described (Hanna et al. 2014; Moshfegh et al. 2014). Cells were kept under  $2\text{ }\mu\text{g/mL}$  doxycycline in the culture medium to repress the biosensor expression under normal culture and propagation.

For plasmid transfection-based stable cell line production, mouse embryonic fibroblast cells were transfected using the Jet-Prime (Polyplus) reagent following the manufacturer's protocols using biosensor constructs in a plasmid system containing the tet-inducible promoter driving the biosensor cassette and a puromycin-resistance gene expression cassette driven by a PGK promoter (pRetro-X; Clontech). Forty-eight hours after transfection, cells were checked for fluorescence to confirm the expression and were cultured in growth medium containing  $2\text{ }\mu\text{g/mL}$  doxycycline and  $1\text{ }\mu\text{g/mL}$  puromycin. Puromycin concentration was gradually increased to a final concentration of  $10\text{ }\mu\text{g/mL}$  to achieve stable selection. For the published Wiskott-Aldrich syndrome protein biosensor, the biosensor design, plasmid backbone (Clontech, pECFP-C1), and transfection protocols were described previously (Cammer et al. 2009).

### *Yeast strain construction*

All *S. cerevisiae* strains were derived from BY4741 (MATa, his3 $\Delta$ 1; leu2 $\Delta$ 0; met15 $\Delta$ 0; ura3 $\Delta$ 0).



In order to insert the 24xMBSV5 at the 3' UTR of the ASH1 gene by homologous recombination, we generated the plasmid p415-24xMBSV5-LoxP-KANr-LoxP (pET194). PCR was performed using pET194 as a template and ASH1-specific primers (Fw, TGCGAAATTGAAGGGTACCGTTGCTTATTTGTAA TTACATAACTGAGACAGTAGAGAATTGAAACCTACAAAC GGGTGGAGGATCA; Rev, TGTACAATTGTTTCGTGATAA TGCTCTTATTAGTTGAAAGAGATTACAGTTATCCATGT AGATATCACCTAATAACTTCGTATAG). The product fragment containing the homology sequences was purified on an agarose gel and transformed in *S. cerevisiae* using standard procedures. Transformed cells were spread on YPD-G418 plates and incubated for 3 d at 30°C and replica-plated on YPD-G418 plates. To identify positive transformants, genomic DNA was prepared and used as a template for PCR with primers that annealed outside the cassette (Fw, AACACATACAAGATGTTTGAACG; Rev, AGCAGGTTCCGCTATTTCAGTG).

#### Single-molecule FISH and analysis

Mammalian cell single-RNA FISH was performed as previously described (Femino et al. 1998; Raj et al. 2008). FISH probes, each ~20 nt long, contained one or two end-labeled dyes. The MS2SL probes were dually labeled with Cy3. The MBSV5 was singly labeled with Quasar 570. The GFP probes (used for CFP) were singly labeled with Quasar 670. Images were acquired using a microscope setup previously described (Zenklusen et al. 2008). Cells were imaged with 200-nm z steps, spanning a depth of 6 µm.

FISH images were analyzed using FISH Quant (Mueller et al. 2013). Briefly, after background subtraction, the FISH spots in the cytoplasm were fit to a three-dimensional (3D) Gaussian to determine the coordinates of the mRNA in each color. The intensity and width of the 3D Gaussian were thresholded to exclude autofluorescent particles and nonspecific signals. For colocalization analysis, the RNA positions in each color were first determined by FISH Quant independently. To find green RNA (MBS) colocalized with red RNA (ORF), we applied a mathematical framework: linear assignment problem (LAP) (Jaqaman et al. 2008). The distances between red and green particles were used as the cost. A threshold of 200 nm was used as the cost if the particle was not linked. Costs for particles that were 2 µm further apart were set to an inhibitory large number, so they were not considered for linking. The assignment of green particles to red ones that minimize the total cost was chosen. The fractions of red particles (ORF) that had a green partner (MBS) are reported here.

#### FFS

The FFS experiments were performed on a home-built two-photon fluorescence fluctuation microscope described previously (Wu et al. 2012). A 60× Plan-Apo 1.4 NA oil immersion objective (Olympus) was used to focus the laser (Chameleon Ultra) and collect the fluorescence, which was detected with avalanche photodiodes (APDs) (SPCM-AQR-14; PerkinElmer). The normalized brightness of an EGFP-labeled fluorescent molecule was defined as  $b = \lambda_{app} / \lambda_{EGFP}$ , where  $\lambda_{app}$  was the sample brightness (Chen et al. 2003), and  $\lambda_{EGFP}$  was monomeric EGFP brightness obtained in a calibration experiment by measuring cells transfected with EGFP only.

#### Live-cell fluorescence imaging

Fluorescence images were taken with an automated IX-81 microscope using a 150× 1.45 NA objective (UAPO, Olympus). The excitation was achieved with a 491-nm (Cobolt) diode laser. The image was recorded with an EMCCD camera (Andor iXon3). The microscope was controlled with MetaMorph imaging soft-

ware, and the images were analyzed with ImageJ (National Institutes of Health).

#### FRET imaging of the biosensor

The single-chain, FRET-based Rac1 biosensor was imaged in the RAW264.7 monocyte/macrophage cell line as previously described (Hanna et al. 2014). Briefly, cells were induced for biosensor expression 48 h prior to imaging by a brief trypsinization followed by removal of 1 µg/mL doxycycline from the medium. Twenty-four hours prior to imaging, cells were trypsinized and replated onto 25-mm round glass #1.5 coverslips (Warner Instruments) at  $1 \times 10^5$  cells per coverslip. On the day of the assay, coverslips containing cells were transferred to the live-cell imaging chamber (Bravo-Cordero et al. 2013) with imaging medium consisting of BWD solution (20 mM HEPES, 125 mM NaCl, 5 mM KCl, 5 mM glucose, 10 mM NaHCO<sub>3</sub>, 1 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub> at pH 7.4) (Greenberg et al. 1991; Gevrey et al. 2005) with 5% FBS. Cells were imaged using a custom, optimized multichannel epifluorescence microscope (Spiering and Hodgson 2012) under a 60× 1.45 NA DIC oil immersion objective lens. The FRET and mCER channels were acquired simultaneously using two Coolsnap ES2 cameras (Photometrics) mounted on the side of the Olympus IX81 microscope via an optimized beam splitter, allowing for simultaneous image acquisition to eliminate motion artifacts, as previously described (Spiering and Hodgson 2012; Spiering et al. 2013). The third camera, mounted on the bottom port of the same microscope, acquired the matching DIC image set. The optical setup, filters, and mirror combinations were described previously (Spiering and Hodgson 2012). Acquired image sets were flat-field-corrected and noise- and background-subtracted as previously described. The FRET and mCER channels were pixel-by-pixel-matched and aligned using nonlinear coordinate transformation and morphing (Spiering et al. 2013) prior to image masking and the ratiometric calculation. The linear pseudocolor lookup table was applied where warmer colors corresponded to regions of high Rac1 activity and the colder colors corresponded to lower activity. For quantification of the Rac1 activity levels, active protrusions were measured for average Rac1 activity per unit area within a small region of interest and directly normalized against the same measurement of the nucleus within the same image frame in the same cell.

#### Morphodynamics mapping and cross-correlation analysis

Morphodynamics and cross-correlation analysis were performed as previously described (Machacek et al. 2009). Briefly, the edge motion was tracked from the cropped image stack using the prPanel.m protrusion tracking software (Machacek et al. 2009). The measurement windows of  $3 \times 6$  pixels, which translated to 0.648 µm by 1.296 µm at 60× magnification  $2 \times 2$  binning, were constructed as previously described (Machacek et al. 2009). The leading edge segment typically contained 20–40 measurement windows, depending on the overall length of the segment. Window positions were successively adjusted back and away from the leading edge in 3-pixel units in order to characterize the spatial dependence of the cross-correlation functions. The correlational coupling was measured up to 4.536 µm away from the leading edge. The normalized cross-correlation coefficient at each window was calculated between the normal direction protrusion velocity at the edge and the changes in Rac1 activity at the corresponding window using the Matlab function `xcov`. The individual cross-correlation coefficient distribution at each window was considered as an independent measurement entity, smooth spline-fitted, and pooled between all cells imaged, and

the average maximal cross-correlation coefficient time lag location and the 95% confidence interval were calculated by a non-parametric bootstrap method (Efron 1979). Total individual sampling windows of 362 (Syn;  $n = 13$  cells) and 398 (Ori;  $n = 12$  cells) were measured for this analysis. Briefly, morphodynamics analysis was based on an assumption that when cells are randomly responding to serum-containing medium with no gradient of chemoattractant present, the constitutive edge fluctuations are assumed to be coupled to the associated Rho GTPase activities in the same manner within each individual sampling window along the edge. Furthermore, it is also assumed that such associations are conserved between all cells of a given population (Machacek et al. 2009). These specific assumptions form the key basis for the morphodynamics analysis, and thus we treated the individual windows as an independent measurement entity, and the characteristic coupling is significant if the averaged peak value of the cross-correlation coefficient function is greater (above or below) than the 95% confidence levels of Pearson's correlation coefficient. Furthermore, in a previous work (Machacek et al. 2009), we performed careful simulations in which up to 70% of the windows were replaced by those that did not contain significant coupling between the GTPase activity and the edge motion. This was the threshold level above which the resulting peak location began to register a measurable temporal deviation. Thus, this method is robust enough to predict with confidence the coupling dynamics of protein activities and the edge motion, as long as the averaged cross-correlation coefficient function registers greater than the 95% confidence levels.

#### Western blotting

We used the following primary antibodies: Roche mouse monoclonal anti-GFP (clones 7.1 and 13.1), Santa Cruz Biotechnology mouse monoclonal anti- $\beta$ -actin (clone AC-15), and goat anti-mouse IRDye800CW secondary antibody (Li-Cor). The images were acquired on an Odyssey infrared imaging system.

#### Acknowledgments

We thank Xiuhua Meng for cloning the plasmids used in the study, and Erik L. Snapp for sharing the BFP plasmid. This work is in partial fulfillment of the Ph.D. degree requirement for V.M. This work was supported by grants GM057071, EB013571, and NS083085 from the National Institutes of Health to R.H.S.; FNSNF Fellowship P2GEP3\_155692 to E.T.; GM093121 and CA181838 to L.H.; T32GM007491 to V.M.; GM071828 to D.C.; and the Integrated Imaging Program of the Gruss Lipper Biophotonics Center.

#### References

- Allen WE, Jones GE, Pollard JW, Ridley AJ. 1997. Rho, Rac and Cdc42 regulate actin organization and cell adhesion in macrophages. *J Cell Sci* **110**: 707–720.
- An W, Telesnitsky A. 2002. Effects of varying sequence similarity on the frequency of repeat deletion during reverse transcription of a human immunodeficiency virus type 1 vector. *J Virol* **76**: 7897–7902.
- Barre-Sinoussi F, Ross AL, Delfraissy JF. 2013. Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* **11**: 877–883.
- Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. 1998. Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **2**: 437–445.
- Bravo-Cordero JJ, Moshfegh Y, Condeelis J, Hodgson L. 2013. Live cell imaging of RhoGTPase biosensors in tumor cells. *Methods Mol Biol* **1046**: 359–370.
- Cai X, Lietha D, Ceccarelli DF, Karginov AV, Rajfur Z, Jacobson K, Hahn KM, Eck MJ, Schaller MD. 2008. Spatial and temporal regulation of focal adhesion kinase activity in living cells. *Mol Cell Biol* **28**: 201–214.
- Cammer M, Gevrey JC, Lorenz M, Dovas A, Condeelis J, Cox D. 2009. The mechanism of CSF-1-induced Wiskott-Aldrich syndrome protein activation in vivo: a role for phosphatidylinositol 3-kinase and Cdc42. *J Biol Chem* **284**: 23302–23311.
- Campbell RE, Tour O, Palmer AE, Steinbach PA, Baird GS, Zacharias DA, Tsien RY. 2002. A monomeric red fluorescent protein. *Proc Natl Acad Sci* **99**: 7877–7882.
- Chen Y, Muller JD, So PT, Gratton E. 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophys J* **77**: 553–567.
- Chen Y, Wei LN, Muller JD. 2003. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proc Natl Acad Sci* **100**: 15492–15497.
- Delviks KA, Pathak VK. 1999. Effect of distance between homologous sequences and 3' homology on the frequency of retroviral reverse transcriptase template switching. *J Virol* **73**: 7923–7932.
- Efron B. 1979. Bootstrap methods: another look at the Jackknife. *Ann Stat* **7**: 1–26.
- Femino AM, Fay FS, Fogarty K, Singer RH. 1998. Visualization of single RNA transcripts in situ. *Science* **280**: 585–590.
- Gevrey JC, Isaac BM, Cox D. 2005. Syk is required for monocyte/macrophage chemotaxis to CX3CL1 (Fractalkine). *J Immunol* **175**: 3737–3745.
- Greenberg S, el Khoury J, di Virgilio F, Kaplan EM, Silverstein SC. 1991. Ca<sup>2+</sup>-independent F-actin assembly and disassembly during Fc receptor-mediated phagocytosis in mouse macrophages. *J Cell Biol* **113**: 757–767.
- Hanna S, Miskolci V, Cox D, Hodgson L. 2014. A new genetically encoded single-chain biosensor for Cdc42 based on FRET, useful for live-cell imaging. *PLoS One* **9**: e96469.
- Hocine S, Raymond P, Zenklusen D, Chao JA, Singer RH. 2013. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods* **10**: 119–121.
- Hodgson L, Shen F, Hahn K. 2010. Biosensors for characterizing the dynamics of rho family GTPases in living cells. *Curr Protoc Cell Biol* **46**: 4.11.1–4.11.26.
- Itoh RE, Kurokawa K, Ohba Y, Yoshizaki H, Mochizuki N, Matsuda M. 2002. Activation of rac and cdc42 video imaged by fluorescent resonance energy transfer-based single-molecule probes in the membrane of living cells. *Mol Cell Biol* **22**: 6582–6591.
- Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G. 2008. Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods* **5**: 695–702.
- Kawase K, Nakamura T, Takaya A, Aoki K, Namikawa K, Kiyama H, Inagaki S, Takemoto H, Saltiel AR, Matsuda M. 2006. GTP hydrolysis by the Rho family GTPase TC10 promotes exocytic vesicle fusion. *Dev Cell* **11**: 411–421.
- Kheir WA, Gevrey JC, Yamaguchi H, Isaac B, Cox D. 2005. A WAVE2-Abi1 complex mediates CSF-1-induced F-actin-rich membrane protrusions and migration in macrophages. *J Cell Sci* **118**: 5369–5379.
- Kraynov VS, Chamberlain C, Bokoch GM, Schwartz MA, Slaugh S, Hahn KM. 2000. Localized Rac activation dynamics visualized in living cells. *Science* **290**: 333–337.
- Kurokawa K, Nakamura T, Aoki K, Matsuda M. 2005. Mechanism and role of localized activation of Rho-family GTPases

- in growth factor-stimulated fibroblasts and neuronal cells. *Biochem Soc Trans* **33**: 631–634.
- Lau IF, Filipe SR, Soballe B, Okstad OA, Barre FX, Sherratt DJ. 2003. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol* **49**: 731–743.
- Machacek M, Hodgson L, Welch C, Elliott H, Pertz O, Nalbant P, Abell A, Johnson GL, Hahn KM, Danuser G. 2009. Coordination of Rho GTPase activities during cell protrusion. *Nature* **461**: 99–103.
- May C, Rivella S, Callegari J, Heller G, Gaensler KM, Luzzatto L, Sadelain M. 2000. Therapeutic haemoglobin synthesis in  $\beta$ -thalassaemic mice expressing lentivirus-encoded human  $\beta$ -globin. *Nature* **406**: 82–86.
- Mitra SK, Hanson DA, Schlaepfer DD. 2005. Focal adhesion kinase: in command and control of cell motility. *Nat Rev Mol Cell Biol* **6**: 56–68.
- Miyawaki A, Llopis J, Heim R, McCaffery JM, Adams JA, Ikura M, Tsien RY. 1997. Fluorescent indicators for Ca<sup>2+</sup> based on green fluorescent proteins and calmodulin. *Nature* **388**: 882–887.
- Moshfegh Y, Bravo-Cordero JJ, Miskolci V, Condeelis J, Hodgson L. 2014. A Trio-Rac1-Pak1 signalling axis drives invadopodia disassembly. *Nat Cell Biol* **16**: 574–586.
- Mostoslavsky G, Fabian AJ, Rooney S, Alt FW, Mulligan RC. 2006. Complete correction of murine Artemis immunodeficiency by lentiviral vector-mediated gene transfer. *Proc Natl Acad Sci* **103**: 16406–16411.
- Mueller F, Senecal A, Tantale K, Marie-Nelly H, Ly N, Collin O, Basyuk E, Bertrand E, Darzacq X, Zimmer C. 2013. FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat Methods* **10**: 277–278.
- Naldini L, Blomer U, Gallay P, Ory D, Mulligan R, Gage FH, Verma IM, Trono D. 1996. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**: 263–267.
- Neri A, Nicolson GL. 1981. Phenotypic drift of metastatic and cell-surface properties of mammary adenocarcinoma cell clones during growth in vitro. *Int J Cancer* **28**: 731–738.
- O'Sullivan RJ, Karlseder J. 2010. Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* **11**: 171–181.
- Ouyang M, Huang H, Shaner NC, Remacle AG, Shiryayev SA, Strongin AY, Tsien RY, Wang Y. 2010. Simultaneous visualization of protumorigenic Src and MT1-MMP activities with fluorescence resonance energy transfer. *Cancer Res* **70**: 2204–2212.
- Pertz O, Hodgson L, Klemke RL, Hahn KM. 2006. Spatiotemporal dynamics of RhoA activity in migrating cells. *Nature* **440**: 1069–1072.
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. 2008. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**: 877–879.
- Robinett CC, Straight A, Li G, Willhelm C, Sudlow G, Murray A, Belmont AS. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* **135**: 1685–1700.
- Schneider D, Tuerk C, Gold L. 1992. Selection of high affinity RNA ligands to the bacteriophage R17 coat protein. *J Mol Biol* **228**: 862–869.
- Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* **22**: 1567–1572.
- Spiering D, Hodgson L. 2012. Multiplex imaging of Rho family GTPase activities in living cells. *Methods Mol Biol* **827**: 215–234.
- Spiering D, Bravo-Cordero JJ, Moshfegh Y, Miskolci V, Hodgson L. 2013. Quantitative ratiometric imaging of FRET-biosensors in living cells. *Methods Cell Biol* **114**: 593–609.
- Tanenbaum ME, Gilbert LA, Qi LS, Weissman JS, Vale RD. 2014. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**: 635–646.
- Wheeler AP, Wells CM, Smith SD, Vega FM, Henderson RB, Tybulewicz VL, Ridley AJ. 2006. Rac1 and Rac2 regulate macrophage morphology but are not essential for migration. *J Cell Sci* **119**: 2749–2757.
- Wu B, Chao JA, Singer RH. 2012. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys J* **102**: 2936–2944.
- Zawistowski JS, Sabouri-Ghomi M, Danuser G, Hahn KM, Hodgson L. 2013. A RhoC biosensor reveals differences in the activation kinetics of RhoA and RhoC in migrating cells. *PLoS One* **8**: e79877.
- Zenkus D, Larson DR, Singer RH. 2008. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat Struct Mol Biol* **15**: 1263–1271.
- Zhang J, Campbell RE, Ting AY, Tsien RY. 2002. Creating new fluorescent probes for cell biology. *Nat Rev Mol Cell Biol* **3**: 906–918.



## Synonymous modification results in high-fidelity gene expression of repetitive protein and nucleotide sequences

Bin Wu, Veronika Miskolci, Hanae Sato, et al.

*Genes Dev.* 2015 29: 876-886

Access the most recent version at doi:[10.1101/gad.259358.115](https://doi.org/10.1101/gad.259358.115)

---

**Supplemental  
Material**

<http://genesdev.cshlp.org/content/suppl/2015/04/15/29.8.876.DC1.html>

**References**

This article cites 49 articles, 18 of which can be accessed free at:  
<http://genesdev.cshlp.org/content/29/8/876.full.html#ref-list-1>

**Creative  
Commons  
License**

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting  
Service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

Finally! Antibodies you can trust.  
Unparalleled ChIP & ChIP-seq with rigorously validated antibodies

★★★★★

[Learn more](#)

**diagenode**  
Innovating Epigenetic Solutions

---

To subscribe to *Genes & Development* go to:  
<http://genesdev.cshlp.org/subscriptions>

---