Dynamics of CDKN1A in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit

Graphical Abstract

Highlights
- The eFlut toolset allows for affordable and flexible tagging of endogenous proteins
- Fused and cleaved reporters separate post- and pre-transcriptional regulation
- CDKN1A dynamics vary between individual proliferating cells and after DNA damage
- CDKN1A protein accumulation after DNA damage is cell-cycle dependent

Authors
Jacob Stewart-Ornstein, Galit Lahav

Correspondence
jacob_stewart-ornstein@hms.harvard.edu (J.S.-O.), galit@hms.harvard.edu (G.L.)

In Brief
Stewart-Ornstein and Lahav establish a low-cost systematic approach to generating fluorescent protein fusions in the mammalian genome, using a family of donor plasmids and Cas9 catalyzed DNA breaks. They use these tools to quantify the dynamics and regulation of the cell-cycle inhibitor CDKN1A in single cells.
Dynamics of CDKN1A in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit

Jacob Stewart-Ornstein1,2* and Galit Lahav1,2*

1Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA
2Correspondence: jacob_stewart-ornstein@hms.harvard.edu (J.S.-O.), galit@hms.harvard.edu (G.L.)

http://dx.doi.org/10.1016/j.celrep.2016.01.045

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

Observing the endogenous abundance, localization, and dynamics of proteins in mammalian cells is crucial to understanding their function and behavior. Currently, there is no systematic approach for the fluorescent tagging of endogenous loci. Here, we used Cas9-catalyzed DNA breaks, short homology arms, and a family of donor plasmids to establish endogenous Fluorescent tagging (eFlut): a low-cost and efficient approach to generating endogenous proteins with fluorescent labels. We validated this protocol on multiple proteins in several cell lines and species and applied our tools to study the cell-cycle inhibitor CDKN1A in single cells. We uncover heterogeneity in the timing and rate of CDKN1A induction post-DNA damage and show that this variability is post-transcriptionally regulated, depends on cell-cycle position, and has long-term consequences for cellular proliferation. The tools developed in this study should support widespread study of the dynamics and localization of diverse proteins in mammalian cells.

INTRODUCTION

Studying the endogenous localization, abundance, and behavior of proteins is crucial to understanding their regulation and function. Generation of endogenously tagged genes by random insertion of fluorescent proteins into the genome of mammalian cells has given important insights into cellular dynamics and signaling (Sigal et al., 2006; Cohen et al., 2008; Cohen-Saidon et al., 2009), as have targeted insertions with large homology regions in embryonic stem (ES) cells (Lengner et al., 2007) or using adenoviruses (Shaltiel et al., 2014). In budding yeast, systematic “tagging” of endogenous genes with fluorescent proteins has enabled proteome-wide surveys of protein localization (Huh et al., 2003), abundance (Schaemmaghami et al., 2003), and response to stimuli (Tkach et al., 2012). More generally, homologous recombination with short DNA homology regions (40–60 bp) and a set of template plasmids containing genetic markers for gene replacement, tagging, and modification give budding yeast part of its genetic power.

Similar techniques have not generally been applicable to mammalian genomes outside of mouse ES cells, mainly due to weaker homology-directed repair capacity. With the advent of CRISPR/Cas9 technology, which enables precise cutting of the genome, it may now be possible to develop efficient homology-directed tagging approaches for multicellular organisms, including mammalian cells. Indeed, groups have published targeting of specific endogenous proteins in Drosophila (Böttcher et al., 2014) and mammalian cells (Park et al., 2014), using CRISPR toolsets. However, there has not been a systematic approach to developing a common plasmid set that allows flexible tagging or modification of the genome with a range of fluorescent protein colors and variants. Here, we established such a systematic approach and used it to tag multiple fluorescent proteins to key signaling proteins in mammalian cells, including Erk2, beta-catenin, and RelA. Further, we take advantage of viral self-cleaving sequences to generate transcripational reporters that are transcribed and translated with the protein of interest but cleaved off to form a separate polypeptide, allowing separation of transcriptional and post-translational regulation. In yeast, we use PCR primers whose 5′ ends have ~40 nt of homology to the target gene sequences and 3′ ends that anneal to our plasmid cassettes (Baudin et al., 1993; Longtine et al., 1998). This minimal homology results in limited efficiency for tagging (~0.01%–1%), but selection with antibiotic markers allows for rapid enrichment of modified cells. Our endogenous Fluorescent tagging (eFlut) toolset allows for modification of loci with a range of markers and reporters, using a minimum of PCR primers.

Tagging of endogenous loci with fluorescent proteins, as opposed to adding exogenous reporters, minimizes the perturbation when tracking cellular components. This is particularly relevant for studying the cell cycle, where a delicate balance of cyclins, cyclin-dependent kinases (CDKs), and CDK inhibitors orchestrates cell-cycle entry, exit, and progression. A crucial component of this network is CDKN1A (p21), a DNA-damage-induced CDK inhibitor that regulates cell-cycle arrest after DNA damage (Dulić et al., 1994) and also plays a role in regulating quiescence and S-phase entry in the unperturbed cell cycle (Overton et al., 2014). Using eFlut, we endogenously tagged alleles of CDKN1A in a range of different cell lines and quantified the unperturbed and DNA-damage-responsive...
We decided to focus on the C terminus, to begin by constructing a series of plasmids that allow for eFlut: A Toolset for Mammalian Gene Tagging

RESULTS

in mammalian cells. These results confirm that endogenous tagging of mammalian genes will enable high time resolution measurements of endogenous protein abundance and localization in many contexts, allowing for C-terminal tagging and selection, we obtained a mixed population of cells, roughly half of parental cells, the CDKN1A-YFP fusion in the knockin cell line, treatment led to the induction of CDKN1A in the wild-type (WT) parental cells, the CDKN1A-YFP fusion in the knockin cell line, and both the endogenous CDKN1A and the GFP-NLS in the region of CDKN1A, we transfected a mixture of Cas9/gRNA plasmid and PCR product from the YFP-P2A-NEO plasmid into MCF7 cells. In the case of CDKN1A, the gRNA stretched across the stop codon and was, therefore, destroyed when the final tagged recombination product was obtained. To tag loci without attaching a large fluorescent tag to the protein of interest. To accommodate this, we constructed a set of plasmids that contain a T2A cleavage tag between the loci of interest and the fluorescent protein and also attached an NLS-PeST cassette (Figure 1B). This causes an endogenous protein to be expressed and translated with a short additional peptide sequence from the T2A element and a separate polypeptide expressing the fluorescent protein with an NLS-PeST, which aids in quantification and time dynamics of the reporter. These constructs (termed 1B throughout the paper) allow for tracking the transcription and translation of an endogenous protein without the influence of post-translational regulation (Figure 1B).

Not all proteins are compatible with C-terminal tagging. For example, tail anchor proteins fail to correctly localize when C-terminal sequences are appended. To develop an N-terminal-tagging plasmid, we constructed an allele that contained a Neo marker followed by a T2A cleavage tag and an EGFP protein (termed 1C throughout the paper). This combination permits tagging of endogenous loci at the N terminus and selection with neomycin without disruption of the underlying open reading frame (ORF) (Figure 1C). In the future, we expect to expand the markers and colors available for translational tags and N-terminal fusions.

Initial Validation of the eFlut Toolset

To validate the set of C-terminal-tagging plasmids (1A), we first applied them to the CDK inhibitor, CDKN1A, as it was previously shown that C-terminal tags do not disrupt CDKN1A function (Overton et al., 2014). Using a Cas9 plasmid (Ran et al., 2013) and a guide RNA (gRNA) designed to target the stop codon region of CDKN1A, we transfected a mixture of Cas9/gRNA plasmid and PCR product from the YFP-P2A-NEO plasmid into MCF7 cells. In the case of CDKN1A, the gRNA stretched across the stop codon and was, therefore, destroyed when the final tagged recombination product was obtained. To tag loci without such a “self-inactivating” gRNA sequence, it is necessary to introduce additional mutations into the locus to destroy the gRNA binding site (see the Supplemental Experimental Procedures for detailed protocol). We obtained single-cell clones and used PCR to confirm an increase in product size, indicating that CDKN1A tagging was successful (Figures 2A and 2B). We also sequenced the PCR product for the CDKN1A-YFP junction and found that, as expected, the protocol led to a clean fusion of the YFP to the most C-terminal codon of CDKN1A (Figure 2B). To further verify CDKN1A tagging with the fusion (1A) and transcriptional (1B) constructs, we performed a western blot of these cells before and after treatment with Nutlin3A, a small molecule that activates p53, the upstream regulator of CDKN1A. Nutlin3A treatment led to the induction of CDKN1A in the wild-type (WT) parental cells, the CDKN1A-YFP fusion in the knockin cell line, and both the endogenous CDKN1A and the GFP-NLS in the line expressing the transcriptional reporter 1B (Figure 2C). These results suggest that the 1A and 1B constructs can successfully tag proteins with the expected molecular products.

Having validated the general short-homology-arm approach to gene tagging, we then systematically tested the plasmids from our collection. Testing the gene-fusion plasmids, after selection, we obtained a mixed population of cells, roughly half of

Figure 1. The eFlut Collection of Plasmids for PCR Amplification-Based Tagging of Endogenous Genes

(A) Diagram of the C-terminal fusion constructs (1A) with pairwise combination of three fluorescent proteins (FP) and selective markers (SM).

(B) Diagram of the transcriptional reporters (1B) containing a T2A viral cleavage and translated with a short additional peptide sequence from the T2A element and a separate polypeptide expressing the fluorescent protein.

(C) Diagram of the N-terminal tagging construct (1C). F1 and R1 indicate PCR primers for C-terminal tagging; F2 and R2 indicate PCR primers for N-terminal tagging.

See also Supplemental Experimental Procedures and Table S1.

kinetics of CDKN1A in single cells. Our analysis revealed that, in response to DNA damage, CDKN1A transcription is highly synchronized in a population, while CDKN1A-protein levels show distinct and complex dynamics linked to the cell-cycle phase. These results confirm that endogenous tagging of mammalian genes will enable high time resolution measurements of endogenous proteins’ abundance and localization in many contexts, which is essential for understanding their regulation and function in mammalian cells.

RESULTS

eFlut: A Toolset for Mammalian Gene Tagging

We began by constructing a series of plasmids that allow for amplification with common PCR primers for C-terminal tagging (Figure 1A; Figure S1). We decided to focus on the C terminus, as 2 decades of experience in yeast suggest that the majority of proteins are unperturbed by such a fusion. These plasmids encoded a yellow fluorescent protein (YFP; mVenus), a CFP (mCerulean), or a red fluorescent protein (RFP; mKate2), followed by a viral cleavage tag and a selectable antibiotic marker (neomycin, blasticidin, or hygromycin). These plasmids (termed 1A throughout the paper) allow for C-terminal tagging and selection of in-frame integrations with antibiotics without introducing terminators or other RNA elements and, therefore, do not require additional use of recombinases to maintain near-endogenous loci without attaching a large fluorescent tag to the protein of interest. To accommodate this, we constructed a set of plasmids that contain a T2A cleavage tag between the loci of interest and the fluorescent protein and also attached an NLS-PeST cassette (Figure 1B). This causes an endogenous protein to be expressed and translated with a short additional peptide sequence from the T2A element and a separate polypeptide expressing the fluorescent protein with an NLS-PeST, which aids in quantification and time dynamics of the reporter. These constructs (termed 1B throughout the paper) allow for tracking the transcription and translation of an endogenous protein without the influence of post-translational regulation (Figure 1B).

Not all proteins are compatible with C-terminal tagging. For example, tail anchor proteins fail to correctly localize when C-terminal sequences are appended. To develop an N-terminal-tagging plasmid, we constructed an allele that contained a Neo marker followed by a T2A cleavage tag and an EGFP protein (termed 1C throughout the paper). This combination permits tagging of endogenous loci at the N terminus and selection with neomycin without disruption of the underlying open reading frame (ORF) (Figure 1C). In the future, we expect to expand the markers and colors available for translational tags and N-terminal fusions.

Initial Validation of the eFlut Toolset

To validate the set of C-terminal-tagging plasmids (1A), we first applied them to the CDK inhibitor, CDKN1A, as it was previously shown that C-terminal tags do not disrupt CDKN1A function (Overton et al., 2014). Using a Cas9 plasmid (Ran et al., 2013) and a guide RNA (gRNA) designed to target the stop codon region of CDKN1A, we transfected a mixture of Cas9/gRNA plasmid and PCR product from the YFP-P2A-NEO plasmid into MCF7 cells. In the case of CDKN1A, the gRNA stretched across the stop codon and was, therefore, destroyed when the final tagged recombination product was obtained. To tag loci without such a “self-inactivating” gRNA sequence, it is necessary to introduce additional mutations into the locus to destroy the gRNA binding site (see the Supplemental Experimental Procedures for detailed protocol). We obtained single-cell clones and used PCR to confirm an increase in product size, indicating that CDKN1A tagging was successful (Figures 2A and 2B). We also sequenced the PCR product for the CDKN1A-YFP junction and found that, as expected, the protocol led to a clean fusion of the YFP to the most C-terminal codon of CDKN1A (Figure 2B). To further verify CDKN1A tagging with the fusion (1A) and transcriptional (1B) constructs, we performed a western blot of these cells before and after treatment with Nutlin3A, a small molecule that activates p53, the upstream regulator of CDKN1A. Nutlin3A treatment led to the induction of CDKN1A in the wild-type (WT) parental cells, the CDKN1A-YFP fusion in the knockin cell line, and both the endogenous CDKN1A and the GFP-NLS in the line expressing the transcriptional reporter 1B (Figure 2C). These results suggest that the 1A and 1B constructs can successfully tag proteins with the expected molecular products.

Having validated the general short-homology-arm approach to gene tagging, we then systematically tested the plasmids from our collection. Testing the gene-fusion plasmids, after selection, we obtained a mixed population of cells, roughly half of...
clones from CDKN1A tagging of the aneuploidy MCF7 and knockins, we performed PCR checks on single-cell-derived cases. To compare the frequency of homozygous and heterozygous if, for example, the tagged protein is mildly hypomorphic. In other this cell line showed weak fluorescence in the absence of DNA damage. We observed a substantial increase in fluorescence, as anticipated for a tagged CDKN1A allele (el-Deiry et al., 1994; Figure 2D). Similar results were obtained with all nine 1A donors can contribute to gene tagging at the same locus, as did cells tagged with the RFP or CFP versions of the transcriptional reporter 1B (Figures 2E and S2B). Finally, to test the N-terminal cassette (1C plasmid), we tagged the tail anchor protein BBC3 (PUMA) with GFP. Expression of GFP-PUMA increased in response to stabilization of its transcriptional activator p53 by Nutlin3A and appeared mitochondrial, as expected (Figure 2F; Willing et al., 2012; Zhang et al., 2009).

In some circumstances, single-allele tagging may be desirable if, for example, the tagged protein is mildly hypomorphic. In other cases, following all molecules of a given protein may be critical. To compare the frequency of homozygous and heterozygous knockins, we performed PCR checks on single-cell-derived clones from CDKN1A tagging of the aneuploidy MCF7 and diploid NIH 3T3 cell line. We observed that 20%–30% of clones in each cell line show only the tagged species by PCR (4/18 for NIH 3T3 and 7/24 for MCF7; Figures S3A and S3B). Tagging of multiple alleles can occur by two ways; homologous recombination from a different donor oligo in each locus or sequential modification of a single locus followed by templated recombination onto the second locus. To determine whether different DNA donors can contribute to gene tagging at the same locus, we tagged CDKN1A in MCF7 cells using a mixed-donor population of YFP and RFP. Treating the polyclonal population, we observed that 20%–30% of the cells show both colors of CDKN1A (10/41) expressing a single color (RFP or YFP, respectively), compared to 12% (14/41) and 14% (10/41) expressing both colors, compared to 12% (14/41) and 14% (10/41) expressing both colors (RFP or YFP, respectively), showing that multiple donors can recombine in the same cell (Figures S3C–S3E). The approach to obtaining multicolor-tagged alleles is a potentially useful tool for studying the stochasticity of gene expression and allele specific expression patterns. For a clone expressing both a CDKN1A-YFP tag and a CDKN1A-RFP tag, using the 1A plasmids, we observed a strong correlation between expression of the two alleles after DNA damage (r = 0.7684; Figure 2G).
Multiple Genes Can Be Tagged Using eFlut in Various Cell Lines and Species

Next, we tested how general the eFlut toolset is by fluorescently tagging nine additional proteins in their endogenous locus (MDM2, TP53, FOXO1, RelA, ERK2, CTNNB1, GMNN, PCNA, and KI67) with the C-terminal fusion YFP-NEO plasmid (1A). We found that each tagged protein retained its expected localization and response to stimulus or cell-cycle progression (Figures 3A and 3B; Figure S4; Lahav et al., 2004; Tay et al., 2010; Calnan and Brunet, 2008; Costa et al., 2006; Chen et al., 2014; Sakaue-Sawano et al., 2008; Celis and Celis, 1985; van Dieren-Croes et al., 1989). We further validated a subset of lines by western blot and genomic PCR and observed an expected GFP-protein fusion weight and PCR products (Figures S3F and S3G). Two other proteins (ATM and P38) did not generate successful tags. In each case, we used only a single gRNA, making it possible that testing of further gRNA, or slightly modified or longer homology regions, would allow for tagging of even these refractory genes. In addition, we applied our toolset to six additional cell lines, tagging CDKN1A or p53 in three human lines (HCT116, A549, and UACC257), two mouse lines (NIH 3T3 and HEPA1C1C7), and one dog line (MDCK) (Figures 3C and 3D).
These results suggest that the eFlut toolset has wide applicability across cell lines, species, and proteins.

To further validate that the endogenous fusions generated by eFlut behave similarly to previously described constructs, we quantified the dynamics of the transcription factors RelA and TP53. The dynamics of these proteins have been well studied; they both exhibit highly stereotyped oscillations that require transcriptional functionality and precise degradation kinetics and, therefore, represent a stringent test of the tagged proteins functionality. Both TP53-YFP and RelA-YFP (using the C-terminal fusion plasmids 1A) exhibited normal response to DNA damage and tumor necrosis factor α (TNFα) treatment, showing oscillations with periods of 5 hr and 90 min, respectively (Figures 3E and 3F; Lahav et al., 2004; Tay et al., 2010). We also established a cell line expressing the cycle-dependent protein Geminin (GMNN) fused to YFP (using the C-terminal fusion plasmids 1A) and also expressing the widely used GMNN (1-110)-CFP “FUCI” cell-cycle-tracking construct (Sakaue-Sawano et al., 2008). We observed near-identical dynamics of the endogenous GMNN-YFP and FUCI transgene, with accumulation in the S/G2 phase and depletion after cytokinesis (Figure 3G). Note that the endogenous tag that we established exhibited less variability in expression level between cell cycles. We concluded that endogenous tagging using our eFlut toolset does not disrupt the dynamics of p53, RelA, and Geminin.

Another useful purpose for tagging endogenous genes is to use the tag as a handle for manipulating protein levels. We obtained an anti-GFP nanobody fused to an E3-ligase component (deGradFP, NSlmb-vhhGFP4) that has been used for depletion of GFP-tagged proteins in Drosophila and mammalian cells (Caussinus et al., 2012), placed it under a doxycycline-inducible promoter, and expressed it in cells with the PCNA-YFP knockin gene (Figure 3H). The addition of doxycycline resulted in the depletion of PCNA-YFP over 48 hr, validating that this approach is applicable to manipulating tagged genes (Figure 3I).

A Tagged CDKN1A Is Regulated by the Cell Cycle and p53

Next, we investigated the utility of eFlut technology for studying cell-cycle regulation of CDKN1A. We quantified CDKN1A-YFP, as cells progressed through the cell cycle, by time-lapse microscopy over 48 hrs (roughly, two cell cycles). In both MCF7 and NIH 3T3 cells, CDKN1A oscillated with the cell cycle, rising in the G2-M-G1 phase and rapidly degrading as cells entered the S phase, as estimated by cell division time and in silico alignment of single cells (Figures 4A and 4B; Figure S5; Movies S1 and S2; Loewer et al., 2010). Note that the CDKN1A behavior and cell-cycle timing varies substantially between individual cells (Figure 4B). The source of this heterogeneity is still unknown and may represent intrinsic random variations or the presence of an unmeasured underlying factor. To further verify the relationship between CDKN1A degradation and the cell cycle, we compared the accumulation and degradation of GMNN-YFP and CDKN1A-CFP in a cell line expressing both tags. As expected, we observed that degradation of CDKN1A-YFP preceded GMNN-YFP accumulation, which is known to occur in the S phase (Figure 4C; Bornstein et al., 2003).

In MCF7 cells, we noted that CDKN1A levels often rose sharply in the G1 phase. This sharp increase resembled the spontaneous pulses of p53 that we have previously observed in the G1 phase (Loewer et al., 2010). To test whether p53 contributed to CDKN1A accumulation in the G1 phase, we compared CDKN1A-YFP accumulation in WT cells to that in cells in which p53 was depleted by a retroviral short hairpin RNA (shRNA) construct. We found a CDKN1A pulse in the G1 phase in WT p53 cells, but not in shp53 cells, supporting a role for p53 in triggering CDKN1A pulses in the G1 phase (Figure 4D). To further quantify the relationship between p53 and CDKN1A pulses in the G1 phase, we constructed a cell line expressing both a p53-YFP reporter and a CFP-tagged CDKN1A (plasmid 1A). Imaging these cells as they passed through the cell cycle, we observed that the sharp increases in CDKN1A-CFP signal were typically preceded by a pulse of p53-YFP (Figures 4E and 4F; Movie S3). Note that the small simultaneous peaks in p53 and CDKN1A at mitosis result from cell rounding and autofluorescence during division.

Interestingly, not all p53 pulses led to CDKN1A pulses, and even sister cells, which show identical bursts of p53 activity, often show divergent CDKN1A accumulation (Figure 4G). To determine whether specific aspects of the p53 pulse predict the CDKN1A response, we quantified p53 and CDKN1A in >100 cells and computationally identified all spontaneous p53 peaks and subsequent CDKN1A dynamics. In silico synchronization of CDKN1A traces to the peak of the p53 pulse revealed that p53 induction is, indeed, followed by an increase in CDKN1A level on average (Figure 4H). In addition, we have observed a moderate (R = 0.36) but significant (p < 0.05) association between p53-YFP fold change and CDKN1A-T2A-CFP transcriptional reporter (Figure 4I). We tested other quantitative features of p53 dynamics, including its integral, amplitude, and pulse width, but did not observe increased predictive power (Figure S9). Collectively, these experiments suggest that the sharp accumulation of CDKN1A observed in the G1 phase results from p53 induction. The observation that not all p53 pulses are productive in activating CDKN1A suggest additional mechanisms that buffer against p53 activity in non-damaged proliferating cells.

The Timing of CDKN1A Induction Post-DNA Damage Depends on the Cell-Cycle Phase

In response to DNA damage, CDKN1A is induced by p53 and leads to cell-cycle arrest. To quantify the dynamics of CDKN1A in response to DNA damage, we imaged CDKN1A-RFP cells after DNA double-strand breaks induced by the radiomimetic drug necrozinostatin (NCS). Interestingly, although CDKN1A-RFP levels increased in all cells, the timing and rate of the increase were heterogeneous (Figure 5A; Movie S4). Specifically, some cells immediately induced CDKN1A, whereas in others, we observed CDKN1A protein expression only >10 hr post-DNA damage (Figure 5A). To verify the heterogeneous behavior of CDKN1A after DNA damage, we performed immunofluorescence on the endogenous CDKN1A protein in fixed cells, which allows rapid measurements and computational analysis of thousands of cells. Endogenous CDKN1A protein showed a long-tailed distribution of expression levels after DNA damage (Figure 5B), consistent with the large variation in the dynamics...
Figure 4. CDKN1A Oscillation during the Cell Cycle Are Driven by Rapid Degradation in S Phase and p53

(A) MCF7 cells expressing CDKN1A tagged with YFP at the endogenous locus were tracked over 48 hr. Divisions are indicated by a dashed line.

(B) A histogram showing the distribution of times between mitosis and the loss of CDKN1A signal (n = 58) in MCF7 cells. The peak at 8 hr is consistent with entry into the S phase in these cells (Loewer et al., 2010).

(C) Cells expressing GMNN-YFP and CDKN1A-CFP from the endogenous loci were imaged over 36 hr. Depletion of CDKN1A occurs shortly before GMNN-YFP accumulation, consistent with CDKN1A degradation in the early S phase. norm., normalized.

(D) Cells expressing CDKN1A-YFP were analyzed for YFP accumulation in the G1 phase measured in WT (n = 88) or p53 depleted (n = 54) cells. Only WT cells show an accumulation of CDKN1A in the G1 phase. Faint traces represent examples from single cells. Bold traces represent the averaged behavior.

(E) MCF7 cells expressing p53-YFP and CDKN1A-CFP were imaged for 48 hr. Single-cell traces of two cells are shown with pulses of p53 (blue) preceding CDKN1A (yellow). Dashed lines represent division times.

(F) Quantification of the distributions of times between the peak of a p53 pulse and the subsequent CDKN1A expression (n = 42 cells).

(G) Not all p53 pulses result in CDKN1A expression. A pair of sister cells is shown; both show p53 pulses after mitosis (blue), but only one of which subsequently activates CDKN1A (yellow). Asterisks indicate the sister cells that are followed in the traces below.

(legend continued on next page)
of the CDKN1A-RFP tagged in live cells (Figure 5A). To test whether the variation in the timing of CDKN1A induction post damage is due to variation in p53, we examined p53 levels in response to DNA damage. We observed a strong and homogenous induction of p53 after DNA damage (Figure 5C), indicating that other cellular factors are responsible for the variation in CDKN1A levels post-damage.

Next, we asked whether the different CDKN1A kinetics in different cells result from stochastic activation of the CDKN1A promoter. We measured the DNA-damage response of a cell line with the T2A-GFP-NLS transcriptional reporter (plasmid 1B) knocked into the endogenous locus. In this construct, the GFP is cleaved off CDKN1A, eliminating post-translational regulation. Consistent with the uniform expression of p53 (Figure 5C), the cleaved fluorophore was induced homogenously across cells in response to DNA damage (Figure 5D), implicating that post-translational regulation of CDKN1A introduces population heterogeneity (Figures 5A and 5B). We verified this conclusion with single-molecule fluorescence in situ hybridization (smFISH) for the CDKN1A transcript before and after DNA damage, noting a strong and uniform increase in CDKN1A mRNA levels after DNA damage (Figure 5E; Raj et al., 2008). Comparing the behaviors of the protein fusion (using plasmid 1A) and transcriptional CDKN1A (using plasmid 1B) reporters in time-lapse microscopy, we note a rapid and fairly uniform increase in the transcriptional reporter compared to a much more heterogeneous response of the protein fusion (Figure 5F; Movies S4 and S5). Taken together, these results indicate that the variation in CDKN1A expression emerges from post-translational regulation and not from transcriptional heterogeneity.

Next, we asked whether cell-cycle position can explain the variation in accumulation of CDKN1A after DNA damage. To test this, we compared the time elapsed since cell division to the CDKN1A-RFP expression 4 hr after DNA damage. We observed that cells that have recently divided (within 8 hr before damage), or divided long before damage (>16 hr), accumulate CDKN1A, whereas cells that divided between 8 hr and 16 hr before damage (which are likely in S phase at the moment of damage) show negligible accumulation of CDKN1A protein (Figures 5G and 5H). These results suggest that the cell-cycle phase in which damage occurs plays a major role in determining the kinetics of CDKN1A protein accumulation. As CDKN1A accumulation has been implicated in resistance to apoptosis (Wang and Walsh, 1996; Maddocks et al., 2013), regulation of CDKN1A accumulation is one potential mechanism by which cell-cycle position could influence response to a range of cytotoxic treatments.

**CDKN1A Expression after DNA Damage Predicts Reduced Proliferation**

To test the consequences of CDKN1A induction and, specifically, whether the diversity of CDKN1A expression after DNA damage results in long-term proliferative defects, we developed a cell line expressing the proliferation marker ki67 tagged with YFP and CDKN1A tagged with RFP (plasmids 1A). We found that, as predicted, in the absence of DNA damage, most cells are ki67-YFP high and CDKN1A low. After DNA damage, CDKN1A was induced in a fraction of cells, and ki67 expression was lost in those same cells (Figure 6A; Movie S6). We quantified the transition from ki67-YFP high/CDKN1A low to a ki67-YFP low/CDKN1A high and verified that after DNA damage, cells lose proliferative markers and increase CDKN1A expression (Figures 6B and 6C). Then, we asked whether early CDKN1A expression could predict subsequent loss of ki67 and proliferative potential by quantifying CDKN1A-RFP expression 5 hr after DNA damage and ki67 levels 24 hr after damage (Figure 6D). We observed that early expression of CDKN1A does, indeed, predict loss of proliferative potential, suggesting that early, cell-cycle-regulated accumulation of CDKN1A (Figures 5G and 5H) results in long-term fate-determining consequences for the cell.

**DISCUSSION**

The recent availability of Cas9 reagents has allowed many groups to perform low- to high-throughput gene-knockout experiments. Here, we developed a novel plasmid set and used PCR with short DNA oligos to introduce fluorescent tags into various proteins in the genome of multiple cell lines. The flexibility of our eFlut toolset will complement the existing Cas9 technology and make gene tagging in mammalian cells a more standardized, transparent, and general approach.

We suggest that short-homology-mediated recombination, which we apply here to tag genes, will be extended to other modifications, such as promoter swaps and gene replacement. Further, we expect that, as with the yeast-tagging plasmid collections, there will be substantial room to optimize and extend these reagents with improved linkers, fluorophores, and markers (Sheff and Thorn, 2004). We show that the toolset described here enables rapid modification and observation of proteins in several cell lines and expect that incorporating additional approaches to improve homologous recombination efficiency or delivery of the Cas9/gRNA complex will extend the flexibility of these tools further (Yu et al., 2015). Finally, the low cost of this approach and relatively limited labor involved may enable large collections of tagged proteins to be assembled.

With regard to CDKN1A, we observe cell-cycle regulation that likely relies on the regulation of protein stability (Bornstein et al., 2003; Overton et al., 2014) and also transcriptional regulation by p53, the activity of which, we find, often precedes an increase in CDKN1A expression. Consistent with previous work, we observed that not all p53 pulses induce CDKN1A expression (Loewer et al., 2010), and defining the characteristics of the p53 kinetics or modifications that result in productive CDKN1A expression is an unanswered question.
Figure 5. The Timing and Rate of CDKN1A Induction in Response to DNA Damage Depend on the Cell Cycle

(A) Schematic of signal transduction in response to DNA damage, showing CDKN1A induction through transcriptional activation by p53. MCF7 cells expressing CDKN1A-YFP were exposed to DNA damage and traced over time. Some cells immediately accumulated CDKN1A, whereas others show a delay.

(B) Expression of the CDKN1A endogenous protein and reporter show heterogeneous abundance across single cells. Images: MCF7 cells expressing CDKN1A tagged with RFP (1A) were imaged before and after (4 hr) DNA damage. Histograms: distribution of endogenous CDKN1A quantified by immuofluorescence. The red lines indicate the mean.

(C) Expression of endogenous p53 or the p53 reporter is uniformly increased after damage. Images: MCF7 cells expressing p53-YFP were imaged before and after (2 hr) DNA damage. Histograms: distribution of endogenous p53 quantified by immunofluorescence. The red lines indicate the mean.

(legend continued on next page)
In response to DNA damage, there is substantial heterogeneity in the accumulation of CDKN1A, which we find to be a consequence of post-translational regulation, as is largely explained by cell-cycle position at the moment of damage. The delayed CDKN1A accumulation in S-phase damaged cells is intriguing, as cells in the S phase have long been known to be more resistant to DNA damage (Griffith and Tolmach, 1976). This has largely been presumed to be due to pre-adapted repair machinery. Low CDKN1A expression—and, thus, no firm block of cell-cycle progression when the cell eventually exits the S phase—could also provide a proliferative advantage. Consistent with this hypothesis, we observed that cells inducing CDKN1A within 5 hr of DNA damage show decreased proliferation 24 hr later. The eFlut toolset we present here will enable the construction of additional reporters in diverse cell lines to further explore the relationship between cell-cycle phase and resistance to DNA damage.

**EXPERIMENTAL PROCEDURES**

**Cell Culture**

Parental cell lines were generally obtained from ATCC, thawed, and propagated in RPMI (GIBCO) with 5% fetal bovine serum (FBS) or DMEM with 10% fetal calf serum (FCS) (NIH 3T3). For microscopy, RPMI lacking phenol red and riboflavin was used. MDCK cells were obtained from the Harvard Digestive Disease Center. HEPA1C1C7 was a gift from Dr. Charles Weitz, Harvard Medical School. Cells were transfected using LT1 reagent (Mirus) according to the manufacturer’s instructions. Single-cell clones were obtained by limiting dilution.

**Gene Tagging**

See the Supplemental Experimental Procedures for detailed protocol. Briefly, cells were transfected with a cocktail of PCR-produced homology donor and a Cas9 plasmid expressing a gRNA directed at the locus of interest. In the case of multi-color tagging (Figure 2G), two separate donor sequences (different colors with the same antibiotic) were mixed at a 1:1 ratio and transfected. Cells were allowed to recover from the transfection and recombine before selection with the appropriate antibiotic. Selected populations were then used as a polyclonal mixture or selected for single clones by limiting dilution.

**Chemical Treatments**

DNA damage was inflicted by addition of the radiomimetic drug (NCS (Sigma) at 100 ng/ml. Nutlin3A was purchased from Sigma, dissolved in DMSO, and applied at a final concentration of 5 μM for the indicated times. Other small-molecule inhibitors used in Figure 3 were applied and assayed at the following concentrations and times: TNFx (10 ng/ml; 30 min; Sigma), AZD8055 (100 mM, 10 min; Biovision), and CH99021 (5 μM, 30 min; Sigma). Doxycycline (Sigma) was applied at 1 μg/ml to induce the Tet-NSLMB construct.

**Antibodies and Reagents**

Primary antibodies for CDKN1A (Calbiochem), p53 (FL393, Santa Cruz Biotechnology), and GFP (Invitrogen) were purchased and used at 1:400–800. Secondary goat anti-mouse or anti-rabbit antibodies conjugated to AF555 or AF647 were purchased from Invitrogen. DAPI was purchased from Life Technologies.

**Plasmids and Cloning**

Standard molecular biology techniques using restriction-enzyme-based cloning were applied to construct template plasmids. Homology donors were amplified from templates with Phusion DNA Polymerase (New England Bio-labs) and gel purified. Briefly, fluorescent proteins with a 10AA QA linker were amplified and cloned into a kanamycin-marked shuttle vector’s multiple cloning site with Xho1/EcoR1. Subsequently, P2A-selection marker sequences were amplified and cloned into EcoR1 and Sac1 of the same vector. For XFP-NLS-PEST, the sequence was amplified and cloned into the same vector with a Saal/EcoR1 digestion (destroying the Xho1 site; this was necessary as the PEST sequence contained a Xho1 site). The N-terminal-tagging plasmid vector was constructed similarly, with Neo cloned between Xho1/EcoR1 and T2A-EGFP cloned between EcoR1 and Sac1.

Our Cas9 expression plasmid (pSpCas9(BB)-2A-Puro (PX459)) was a gift from Dr. Feng Zhang (Addgene plasmid #48139). The GFP degradation construct was constructed by amplifying NSlmb-rrhGFPS4 from pcDNA3_NSlmb-rrhGFPS4 plasmid, which was a gift from Dr. Markus Affolter (Addgene plasmid #39579), and cloning it into a doxycycline-inducible vector. The pRetroSuper-p53th vector was a gift from Dr. Reuven Agami (Netherlands Cancer Institute; NKI). Oligos were ordered from Integrated DNA Technologies (sequences are in Table S1).

**Microscopy**

Cells were plated in glass-bottom 35-mm dishes (MatTek) 24–48 hr before imaging; 1–2 hr before imaging, cells were switched to transparent media (RPMI lacking riboflavin and phenol red; Invitrogen). Live-cell imaging was performed with a Nikon Eclipse Ti-E microscope equipped with a heating chamber (37°C) and CO2 control (5%); an epi-fluorescent source (either mercury arc lamp [Prior] or LED system [Lumenco]), automated stage (Prior); YFP, mCherry, or GFP filter set (Chroma); and a charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) camera (Hamamatsu). All live-cell imaging was performed with a 20× PA objective (Nikon).

**Immunofluorescence**

Cells were plated in 35-mm dishes on glass coverslips; at the appropriate time after treatment, they were fixed with 2% formaldehyde (Alfa Aesar) for 10 min at room temperature, followed by permeabilization with 0.1% Triton (EMD). Cells were sequentially stained with antibodies for CDKN1A, TP53, or GFP, and secondary antibodies conjugated to Alexa Fluor 647 or Alexa Fluor 555. Cells were imaged on a Nikon Eclipse Ti-E microscope equipped with an epi-fluorescent source (Prior); an automated stage (Prior); YFP, Cy3, or Cy5 filter sets (Chroma); and a CCD camera (Hamamatsu).
expression after DNA damage (24 hr after NCS).

**Response to DNA Damage**

Figure 6. CDKN1A Accumulation Predicts Loss of Proliferation in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit, Cell Reports (2016), http://dx.doi.org/10.1016/j.celrep.2016.01.045

Please cite this article in press as: Stewart-Ornstein and Lahav, Dynamics of CDKN1A in Single Cells Defined by an Endogenous Fluorescent Tagging Toolkit, Cell Reports (2016), http://dx.doi.org/10.1016/j.celrep.2016.01.045

**Immunoblotting**

Cells were harvested and protein extracts obtained by lysis in the presence of protease and phosphatase inhibitors. Total protein amount was quantified using the Bradford assay (InVitrogen) and transferred to nitrocellulose membranes by electoblotting. Membranes were blocked with 5% non-fat dried milk, incubated with primary antibody, washed, and incubated with secondary antibody coupled to IR800 or IR680 dyes (LI-COR Biosciences). Protein levels were detected using an Odyssey scanner (LI-COR Biosciences).

**smFISH**

Cells were fixed with 2% formaldehyde for 10 min, followed by treatment with 75% ethanol at 4°C. smFISH was performed using a previously described protocol (Raj et al., 2008; Purvis et al., 2012). Probes for CDKN1A were purchased from Biosearch Technologies. Hybridization was at 37°C for 4 hr, and washing conditions were 2 x 30 min at 37°C, 10% formamide. We acquired 15–25 focal planes at 0.2-µm intervals using MetaMorph acquisition software (Molecular Devices). Typical exposure times per individual focal plane were 0.2 s. Gene expression was quantified by counting cytoplasmic intensity above a given threshold at 60× magnification.

**Selection of gRNAs**

We downloaded the human genome sequence from NCBI and used Reference Sequence (RefSeq) transcript data to define the C terminus of each ORF. Potential gRNA sequences with cut sites within 20 bp of the stop codon were examined for GC content, uniqueness in the genome, nearness to the stop codon, and lack of polyN sequences. A gRNA was selected for each gene and cloned into the Cas9 vector (see Table S1 for sequences). Note that gRNAs may cut the genome in multiple locations, leading to additional non-targeted integrations.

**Selection of Homology Regions**

We selected 40 bp immediately before and after the stop codon as our homology regions. In situations in which the gRNA cut site would be retained in the recombined sequence, we introduced silent mutations to eliminate the gRNA protospacer adjacent motif (PAM) site or one of the first five nucleotides in its recognition region (see Table S1 for sequences).

**Microscopy Data Analysis**

Microscopy data (live and fixed) were processed with custom MATLAB code. Single cells were tracked manually, using the phase images with a MATLAB interface. Single-cell tracks were projected onto the fluorescent images, which were then background corrected (by median filtering and subsequent Top-hat background subtraction), and nuclear signal (estimated as the average of the top ten pixels in the nuclear area) was then computed for the applicable channels. Fixed images were segmented using the DAPI channel with a watershed interface. Single-cell tracks were projected onto the fluorescent images, which were then background corrected (by median filtering and subsequent Top-hat background subtraction), and nuclear signal (estimated as the average of the top ten pixels in the nuclear area) was then computed for the applicable channels.

**Cell-Cycle Analysis and In Silico Alignment**

To study the influence of the cell cycle on CDKN1A accumulation, we measured single-cell trajectories of CDKN1A (or p53) and also noted the timing of mitosis in each cell. Then, we computationally aligned each trajectory to the first mitosis. This analysis results in a computationally synchronized population to study cell-cycle-regulated protein signaling.

(B) Single-cell traces of CDKN1A-RFP and ki67-YFP after DNA damage show gain of CDKN1A-RFP and loss of ki67-YFP.

(C) The levels of CDKN1A-RFP and ki67-YFP are plotted at the indicated time points after DNA damage, showing largely exclusive staining patterns with a gradual accumulation of CDKN1A-RFP and loss of ki67-YFP cells over time (n > 1,000 cells per time point).

(D) CDKN1A expression predicts subsequent loss of ki67-YFP. Cells that show accumulation of CDKN1A-RFP relative to time 0 (CDKN1A+ cells) show substantially lower ki67-YFP expression at 24 hr. n = 295.
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, three tables, and six movies and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.01.045.

AUTHOR CONTRIBUTIONS

J.S.-O. conceived of the project, performed the experiments, and analyzed the data; G.L. and J.S.-O. designed the experiments and wrote the paper.

ACKNOWLEDGMENTS

We thank Feng Zhang for the Cas9 plasmid and Charles Weitz for the HEP-A1C1C7 cell line. We thank the Nikon Imaging Center at Harvard Medical School for help with light microscopy, members of the G.L. lab for helpful comments and suggestions throughout this work, Jose Reyes for graphical design, and Caroline Mock for advice and support on tissue culture and cell line construction. This work was supported by NIH grant GM083303 and the Harvard Ludwig Cancer Research Center. Received: June 16, 2015 Revised: December 22, 2015 Accepted: January 12, 2016 Published: February 11, 2016

REFERENCES


