Quantification of factors influencing fluorescent protein expression using RMCE to generate an allelic series in the ROSA26 locus in mice

Sara X. Chen1, Anna B. Osipovich1, Alessandro Ustione2, Leah A. Potter1,2, Susan Hipkens1, Rama Gangula1, Weiping Yuan1, David W. Piston3 and Mark A. Magnuson1,2*

SUMMARY

Fluorescent proteins (FPs) have great utility in identifying specific cell populations and in studying cellular dynamics in the mouse. To quantify the factors that determine both the expression and relative brightness of FPs in mouse embryonic stem cells (mESCs) and in mice, we generated eight different FP-expressing ROSA26 alleles using recombinase-mediated cassette exchange (RMCE). These alleles enabled us to analyze the effects on FP expression of a translational enhancer and different 3′-intrinsic and/or polyadenylation sequences, as well as the relative brightness of five different FPs, without the confounding position and copy number effects that are typically associated with randomly inserted transgenes. We found that the expression of a given FP can vary threefold or more depending on the genetic features present in the allele. The optimal FP expression cassette contained both a translational enhancer sequence in the 5′-untranslated region (UTR) and an intron-containing rabbit β-globin sequence within the 3′-UTR. The relative expressed brightness of individual FPs varied up to tenfold. Of the five different monomeric FPs tested, Citrine (YFP) was the brightest, followed by Apple, eGFP, Cerulean (CFP) and Cherry. Generation of a line of Cherry-expressing mice showed that there was a 30-fold variation of Cherry expression among different tissues and that there was a punctate expression pattern within cells of all tissues examined. This study should help investigators make better-informed design choices when expressing FPs in mESCs and mice.

INTRODUCTION

Fluorescent proteins (FPs) are widely used to monitor gene expression and study cellular dynamics in cells and tissues (Tsien, 1989; Heim et al., 1994; Tsien and Miyawaki, 1998; Shaner et al., 2004). However, the successful use of FPs requires that their relative brightness in a cell exceeds the experimental detection threshold (Niswender et al., 1995). Many variables, both intrinsic and extrinsic to the FP, influence the brightness of FPs when expressed in cells or mice. Factors intrinsic to the FP include quantum efficiency, the necessity for some FP variants to dimerize or fold in a specific manner, and the nature of the protein to which they are attached. Factors extrinsic to the FP include the transcriptional activity of the promoter being used and both the half-life and translational efficiency of the mRNA generated.

Although some of the intrinsic factors that influence FP brightness can readily be determined, extrinsic factors have been more difficult to quantify. This is largely owing to the confounding effects of transgene copy number and the genetic insertion site, which can cause large variations in gene expression. Although there have been prior attempts to compare different FP reporter constructs in vitro (Falcone and Andrews, 1991; Yew et al., 1997), the results obtained have been neither quantitative nor conclusive. Similarly, although a large amount of data has been collected on the properties of bacterially-expressed FPs (Olenych et al., 2007), there have been fewer studies to compare the relative brightness of FPs in mammalian cells. As a result, it is not uncommon for suboptimal reporter design to result in wasted time and effort when generating new lines of FP-expressing mice.

The method of recombinase-mediated cassette exchange (RMCE) enables the efficient generation of an allelic series in embryonic stem cells (ESCs) and mice (Araki et al., 2002). RMCE is a two-step procedure. The first step is the generation, by gene targeting, of an acceptor allele containing the appropriate recombinase recognition sites and chemically selectable markers. The second step is to exchange a cassette containing DNA sequences of interest into the acceptor allele by RMCE. An advantage of this method is that it enables multiple allelic variants to be generated at a defined genetic location with greater ease than can be achieved by repetitive gene targeting.

To quantify the effects of intrinsic and extrinsic variables that affect both the expression and brightness of FPs, we generated a loxed cassette acceptor (LCA) allele for the ubiquitously expressed Gt(ROSA)26Sor (ROSA26) gene (Zambrowicz et al., 1997; Kohlhepp et al., 2001; Safran et al., 2003). Utilizing the ROSA26LCA allele, we generated eight different FP reporter alleles by RMCE, which were used to assess the effects of several regulatory elements on FP expression and to compare the brightness of five different FP variants. As part of this effort, we also made several plasmids that can be used to generate both targeting vectors and exchange cassettes by BAC recombineering. The plasmids, mouse ESCs (mESCs), expression cassettes and findings that we describe should be of general use in generating other LCA alleles, basal exchange...
cassettes, and for optimizing the expression of FPs and other reporter alleles in mESCs and mice.

RESULTS
Generation of a ROSA26LCA allele
A targeting vector was made using a two-step BAC recombineering method, as shown in Fig. 1A. First, tandem lox71 and lox2272 sites, flanking both PGK-\(\text{pu\textendash}\Delta tk\) and EM7-neo\(^R\) selection cassettes, were inserted into a ROSA26 BAC clone (Fig. 1A) using kanamycin to select for recombinant clones. Second, the entire LCA selection cassette along with flanking upstream and downstream homology arms of 8.232 kb and 3.770 kb, respectively, was retrieved into pMCS.DT-A by recombineering-mediated gap repair. The targeting vector that was generated replaced a 5.165 kb region of the ROSA26 gene, containing both the promoter and first exon, with the dual PGK-\(\text{pu\textendash}\Delta tk\) and EM7-neo\(^R\) selection cassette. The heteromeric lox sites that were present in the targeting vector have been previously reported to allow efficient, unidirectional RMCE.

![Fig. 1. BAC recombineering, gene targeting and RMCE strategies.](image)

(A) Steps for making a ROSA26 gene targeting vector by BAC recombineering. Short inner homology arms were amplified from a BAC containing the ROSA26 gene and cloned into the pLCA.71/2272 vector. In the first recombination step (1), the cassette containing a \(\text{pu\textendash}\Delta tk\) fusion gene surrounded by tandem lox71 and lox2272 sites was introduced into the BAC. After cloning short outer homology arms into pMCS.DT-A, a second recombination step (2) enabled retrieval of the targeting vector. DT, diptheria toxin A; MCS, multiple cloning site; HA1/2, short homology arm 1 or 2. (B) Generation of the ROSA26 exchange vector by BAC recombineering. pMCS.66/2272 was used to retrieve a 5.165-kb fragment of the ROSA26 gene, thereby generating a basal exchange vector. A variant of this plasmid was used to make all FP exchange vectors as shown in Fig. 2. (C) Gene targeting and RMCE strategy. A ROSA26LCA allele containing the \(\text{pu\textendash}\Delta tk\) fusion gene flanked by tandem lox71 and lox2272 sites was made by gene targeting in mESCs. Puromycin-resistant mESC clones were screened by Southern blot analysis. DNA was digested with \text{NsiI} and hybridized with the 3\textsubscript{\textprime} probe indicated in the figure. Correctly targeted clones containing the ROSA26LCA allele were used to perform RMCE to derive eight different FP-expressing alleles. PCR using primers 1-4 (red arrows) were used for the screening of clones that were resistant to puromycin and sensitive to gancyclovir. (D) Southern blot analysis. The presence of an 11 kb band indicates a correctly targeted mESC clone with the ROSA26LCA allele. (E) DNA PCR analysis. Representative PCR amplification results are shown, indicating correct exchange of the exchange cassette into the ROSA26LCA allele. Prior to RMCE, cells with the ROSA26LCA allele exhibit a 503 bp band after PCR using primers 1 and 2. Correctly exchanged clones exhibit both a 537 bp and 503 bp product using primers 1 and 2, and a 565 bp product with primers 3 and 4.
Expression of fluorescent proteins using ROSA26

(Araki et al., 2002). After electroporation, 270 puromycin-resistant mESC clones were obtained. Southern blot analysis identified three that had undergone the desired homologous recombination event (Fig. 1C,D). Clones 5B9 and 2A6, which demonstrated germline transmission, were used subsequently for all RMCE experiments.

**Design of an allelic series of RMCE vectors and insertion into the ROSA26 gene locus**

We next used pMCS.66/2272, a plasmid with tandem lox66 and lox2272 sites, to retrieve by BAC recombineering a 4.1 kb region of the ROSA26 gene that is absent in the ROSA26^{locreg} mESCs (Fig. 1B). The resulting plasmid, pROSA26.Ex, was used as the basis for making eight different RMCE vectors that contained a dual promoter (PGK and EM7)-driven neo^R cassette so as to enable G418 to be used as a positive selectable marker during RMCE (Fig. 2). The correct exchange of each vector into the ROSA26^{locreg} allele after RMCE was confirmed by PCR analysis (Fig. 1C,E). To simplify assembly, only a 1.065 kb region of the ROSA26 gene, containing the sequence downstream of the promoter, was replaced, thereby allowing expression of the FPs to occur without the need to include splice donor and acceptor sites (Nyabi et al., 2009). Other aspects of the cassette were varied in order to assess the impact of different transcriptional and translational components. First, we examined the impact of a putative 5′-untranslated region (UTR) translational enhancer sequence from the *Xenopus* β-globin gene, hereafter termed EN (Falcone and Andrews, 1991). Second, we examined three different 3′-UTR configurations: a Simian virus 40 polyadenylation sequence (SV40), a rabbit β-globin polyadenylation sequence (bpgA) and an intron-containing rabbit β-globin polyadenylation sequence. Lastly, we examined the relative brightness of five different FP variants: Apple, Cerulean (CFP), Cherry, Citrine (YFP) and enhanced GFP (GFP).

**Effects of extrinsic factors on FP expression**

We first sought to determine the effects of the allele design by comparing the expression of the four alleles that express CFP (Cerulean). Cerulean has been shown to be a bright CFP variant with efficient folding at 37°C (Rizzo et al., 2004). Correctly exchanged clones were analyzed by fluorescence-activated cell sorting (FACS) for the absence of mosaicism (data not shown). Several that exhibited considerable non-fluorescent cell populations were not used in the subsequent studies. The mRNA expression of two different clones of each allele was determined by quantitative PCR and compared with that of the EN.CFP allele (which contains the *Xenopus* β-globin 5′-UTR translational enhancer sequence, CFP and an intron-containing rabbit β-globin polyadenylation sequence). As shown in Fig. 3B, the EN.CFP allele showed the highest level of expression, followed by nonEN.CFP (lacking the translational enhancer), EN.CFP.bpgA (containing a rabbit β-globin polyadenylation sequence) and finally EN.CFP.SV40 (containing a SV40 polyadenylation sequence). One-tailed Student's *t*-tests showed no significant difference between EN.CFP and nonEN.CFP mRNA levels. Similarly, there was no significant difference in the expression of EN.CFP.bpgA and EN.CFP.SV40. However, the presence of the β-globin intron resulted in a twofold higher level of expression as compared with the allele that only contains a polyadenylation sequence. Similarly, the mRNA expression level of both the EN.CFP and nonEN.CFP alleles was approximately 2.5-fold higher than that of EN.CFP.SV40.

Next, we explored the effects of regulatory elements on protein expression by using quantitative confocal microscopy. Relative FP concentrations were calculated on the basis of fluorescence intensities as compared with the ROSA26^{locreg} allele, which does not contain a FP and was used as a non-expressing control. Ten confocal images for each of the mESC clones were taken at 40× magnification with no saturation areas to eliminate non-linearity in intensity measurements (Fig. 3A). In parallel, we imaged serial dilutions of purified FP with known concentrations and measured the fluorescence intensity for each concentration (Patterson et al., 1997). A linear equation of fluorescence intensity against concentration was plotted and used to calculate average concentration of each sample, thereby allowing direct comparison between FPs. The percentage of photons emitted, which correlates to relative brightness, was calculated using the following equation: relative brightness = quantum yield × molar extinction coefficient × concentration. The concentration and relative brightness for each of the four different CFP-containing alleles are shown in Fig. 3C. The EN.CFP allele showed the highest level of protein expression, with 3.4% relative brightness, followed by nonEN.CFP, EN.CFP.bpgA and then EN.CFP.SV40 (Fig. 3C). The EN.CFP allele, which contains the translational enhancer sequence, had 1.26-fold higher protein expression levels than nonEN.CFP, which lacked the *Xenopus* translational enhancer sequence (*P*<3.8×10^-5). Mirroring the real-
time PCR data, the EN.CFP allele, which has an intron-containing rabbit $\beta$-globin polyadenylation sequence, exhibited 1.75-fold higher ($P<4.0 \times 10^{-19}$) protein expression than EN.CFP.bgpA, which only contained the polyadenylation signal sequence from the rabbit $\beta$-globin gene as the 3$^{\prime}$-UTR. Both the EN.CFP and nonEN.CFP alleles had 1.7-fold ($P<1.7 \times 10^{-16}$) and 1.4-fold ($P<5.62 \times 10^{-10}$) higher protein concentrations than the EN.CFP.SV40 allele, respectively. The EN.CFP.bgpA and EN.CFP.SV40 alleles did not show significant differences in protein expression when compared with each other.

To further assess the effects of the different regulatory elements on FP expression, we used FACS. Two clones for each allele were profiled, and the mean fluorescence intensity for each cell line was calculated (Fig. 3D). The results paralleled the confocal measurements, with the EN.CFP allele having the highest level of expression, followed by the nonEN.CFP, EN.CFP.bgpA and EN.CFP.SV40 alleles.

**Comparison of intrinsic brightness of five different FPs**

We next sought to examine the relative brightness of four other monomeric FPs (Cherry, Apple, Citrine and eGFP) with that of Cerulean. These FPs were chosen because they represent different parts of the visible light spectrum and have well-defined physical characteristics. For instance, eGFP is a widely used mutational variant of wild-type GFP owing to its superior brightness, photostability and relatively small sensitivity to pH (Cormack et al., 1996).

To limit variation in extrinsic variables that influence FP brightness, each allele was identical in design to the EN.CFP allele, with the only modification being the FP itself (Fig. 2A). As is well known, a number of factors influence apparent brightness of an FP beyond protein expression, including the efficiency and rate of maturation, the molar extinction coefficient values within the excitation wavelength range, and the quantum yield (Olenych et al., 2007; Rizzo et al., 2009). The known physical properties of the five FPs studied are summarized in Table 1. To directly compare the efficiency of different color variants of FPs (Fig. 2A,E-H), as represented by relative brightness, we employed the same quantitative confocal microscopy method described in the previous section. eGFP, Apple, Citrine and Cerulean (Fig. 3A and Fig. 4A) all exhibited uniform fluorescence within the mESC colonies. However,
Cherry (Fig. 4A) showed a distinctly punctate pattern that seemed to be confined to cell bodies or vesicles. To compare FP intensities, a maximum threshold was set for the intensity measurements of Cherry because the extreme aggregation of proteins within cells rendered saturation unavoidable. Calculations of average protein concentrations, on the basis of calibration intensities of purified corresponding FPs, allowed direct comparison between different color variants. Fig. 4B shows the average concentrations and relative brightness of each FP variant. Citrine showed the highest level of protein concentration (4 μM), followed by Cherry (1.4 μM), Cerulean (1.3 μM), eGFP (1.1 μM) and Apple (1.0 μM). Citrine was also, by far, the brightest FP variant, with a normalized actual brightness of 5.53 (defined as the product of concentration × extinction coefficient × quantum yield) compared with eGFP (1.00). Citrine was followed in brightness by Apple (0.91), Cerulean (0.60) and Cherry (0.52). Differences in relative brightness between all FP variants were statistically significant, with the exception of the difference between Apple and eGFP. It is interesting to note that Citrine is 10.4-fold brighter than the dimmest color variant, Cherry, and approximately sixfold brighter than the remaining color variants. An assessment of the monomeric red FPs revealed that Apple was 1.7-fold brighter than Cherry.

Again, FACS analysis was used to confirm the differences between different fluorescent color variants. It was not possible to quantitatively compare the FPs by FACS using the same excitation wavelengths used in confocal microscopy. However, the same FACS excitation wavelengths were used for YFP and eGFP, as well as for Cherry and Apple. Thus, it was possible to quantitatively compare fluorescence mean intensity between YFP and eGFP (Fig. 4C), and

<table>
<thead>
<tr>
<th>Protein (color)</th>
<th>Excitation peak (nm)</th>
<th>Emission peak (nm)</th>
<th>Brightness*</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerulean (cyan)</td>
<td>433/445</td>
<td>475/503</td>
<td>27/24</td>
<td>Rizzo et al., 2004</td>
</tr>
<tr>
<td>EGFP (green)</td>
<td>488</td>
<td>507</td>
<td>34</td>
<td>Cormack et al., 1996</td>
</tr>
<tr>
<td>Citrine (yellow)</td>
<td>516</td>
<td>529</td>
<td>59</td>
<td>Griesbeck et al., 2001</td>
</tr>
<tr>
<td>Cherry (red)</td>
<td>587</td>
<td>610</td>
<td>16</td>
<td>Shaner et al., 2008</td>
</tr>
<tr>
<td>Apple (red)</td>
<td>568</td>
<td>592</td>
<td>37</td>
<td>Shaner et al., 2008</td>
</tr>
</tbody>
</table>

Protein names and property values are based on literature. *Product of molar extinction coefficient and quantum yield (mM × cm)–1.

![Fig. 4. Comparison of intrinsic brightness of different color variants of FPs.](image)

(A) Confocal images of ESC clones with alleles that express GFP (EN.GFP), Citrine (EN.YFP), Apple (EN.Apple) and Cherry (EN.Cherry). Calibration bars indicate the fluorescence intensity level that corresponds to one color shade in the image; intensity is expressed in arbitrary units. (B) Quantitative confocal microscopy measurements of FP concentration (black line) and relative actual expressed brightness (concentration × extinction coefficient × quantum yield; colored bars). The mean intensities for two clones, with ten images acquired for each clone, were averaged together for each data point. **P<0.01 for all comparisons except between EN.GFP and EN.Apple. Bars indicate s.e.m. (n=20 each).

(C,D) Quantification of mean fluorescence intensities of EN.GFP and EN.YFP by FACS using a 488 nm excitation laser (C), and EN.Apple and EN.Cherry using a 532 nm excitation laser (D). Bars indicate s.e.m. (n=2 each).
between Cherry and Apple (Fig. 4D). In this setting, Apple exhibited 1.4-fold greater fluorescence intensity than Cherry. By contrast, Citrine and eGFP showed no significant difference.

**Generation of polychrome fluorescent embryos and confirmation of germline competency of the mESCs**

To confirm the pluripotency of mutant mESC lines after gene targeting and RMCE, and to validate the expression of three of the FPs in vivo, we performed blastocyst injections of mESCs expressing cyan (EN.CFP), green (EN.GFP) and red (EN.Cherry) FPs. Four ESCs of each clone were simultaneously injected per blastocyst then transferred into pseudopregnant females. Seven out of nine transfers resulted in pregnancies. Embryos were harvested from two of the pregnant animals at E15.5 and two at E17.5 for fluorescent imaging. In total, 21 of 34 embryos exhibited varying degrees of fluorescence, with the majority of the embryos showing one or two colors (Fig. 5A-D). However, one high-percentage chimeric embryo displayed polychrome fluorescence, indicating that it was derived from all of the different mESC clones (Fig. 5A,B). The remaining pregnant females were brought to term, and several high-degree chimeric males were bred to C57Bl/6j females to test for germline transmission. Although it was not our goal to generate a mouse for every different FP tested, one agouti mouse from these matings that exhibited bright red fluorescence (Fig. 5E,F) was used to establish a ROSA26Cherry mouse line.

**Analysis of Cherry expression in ROSA26Cherry mice**

To assess Cherry expression from the ROSA26 locus in vivo, we performed quantitative confocal microscopy using whole-mount tissues. The punctate pattern of Cherry fluorescence observed in mESCs was also present in all of the examined tissues (Fig. 6A). Interestingly, the amount of Cherry in different organs was highly variable, with approximately a 30-fold difference between skin and muscle, which were the highest and lowest expressing tissues, respectively (Fig. 6B).

**DISCUSSION**

In this study, we generated a ROSA26LCA allele and used it to create an allelic series of mESCs to quantitatively assess both intrinsic and extrinsic variables that affect the relative brightness of FPs when expressed in mammalian cells and tissues. The ROSA26 locus was chosen for this study because it is widely thought to be expressed at a consistent level in mouse cells, and its use has been proven in numerous studies examining the effects of protein misexpression and as a Cre-inducible genetic tag (Zambrowicz et al., 1997; Kohlhepp et al., 2001; Safran et al., 2003).

Several features of how the ROSA26LCA allele was generated merit discussion. First, the plasmids used to generate this allele can be used to generate other LCA alleles. These plasmids are designed to enable the use of BAC recombineering, thereby simplifying the assembly of both gene targeting and RMCE vectors. Second, the allele contains puA3k, a fusion protein that can be used to perform both positive and negative selection. This fusion protein also enables mESCs containing the LCA allele to be tested for germline competency without the problematic Sertoli cell toxicity of the wild-type HSV-TK (Chen and Bradley, 2000). Third, the combination of mutant loxP sites present in the LCA allele and exchange vector enable efficient Cre-driven cassette exchange (Araki et al., 2002). Indeed, we found that the use of the tandem lox66/71-lox2272 site strategy is much more efficient than the loxP/inverted-loxP combination that we previously reported for both the Gck and Ptf1a alleles (Long et al., 2004; Burlison et al., 2008). Moreover, because recombination between the lox66 and lox71 sites results in a doubly mutant lox site that can no longer efficiently bind Cre or undergo recombination, the post-RMCE allele is predicted to remain stable in the presence of Cre (Araki et al., 1997).

**Fig. 5. Polychrome mouse embryos.** Four mESCs expressing cyan (EN.CFP), green (EN.GFP) and red (EN.Cherry) FPs were microinjected into 3-day-old blastocysts and implanted into pseudopregnant female mice. Embryos were harvested at E15.5 and E17.5, and fluorescent images obtained. (A,B) This high-degree chimeric embryo displayed cyan, green and red fluorescence in different tissues, indicating that it was derived from the three different mESC types. (A,B) Right and left perspectives, respectively. (C,D) In most embryos, one FP dominated expression, e.g. green (C) or blue (D). (E,F) Offspring of a polychrome chimera show germline transmission of the EN.Cherry allele. Images of two littersmates showing one with the EN.Cherry allele and one with a wild-type ROSA26 allele under white (E) and fluorescence excitation (F). Calibration bars indicate the fluorescence intensity level that corresponds to one color shade in the image; intensity is expressed in arbitrary units.
Using the ROSA26\textsuperscript{LCA} allele, we generated a series of FP-expressing alleles by RMCE, thereby enabling us to perform a side-by-side comparison of specific design variations and to accurately assess the impact of both extrinsic and intrinsic variables on FP expression and relative brightness. This approach, which results in the insertion of a single copy transgene into a defined genetic location, circumvents confounding issues, such as gene copy number and location, that have previously limited the quantitative assessment of different transgenes or design features that influence FP expression or, for that matter, expression of any other gene. As a result, we were able to accurately assess the impact of several different allele design features.

First, we examined the effect of a translational enhancer on FP expression and efficiency. Previously, it was shown that the 5'\textsuperscript{UTR} of the \textit{Xenopus} \(\beta\)-globin gene, along with an optimized translational initiation site (ACCAUGG), increased translational efficiency by 10- to 300-fold (Falcone and Andrews, 1991). When this sequence was inserted in the 5'\textsuperscript{UTR}, we observed only a modest 1.26-fold increase in FP concentration and/or brightness, with no change in mRNA expression. Both EN.CFP and nonEN.CFP contain the optimized initiation site (ACCAUGG), but only EN.CFP contains the \textit{Xenopus} \(\beta\)-globin 5'\textsuperscript{UTR} leader sequence. Therefore, although the effect is modest, inclusion of the 5'\textsuperscript{UTR} of the \textit{Xenopus} \(\beta\)-globin gene improves FP expression in mammalian cells by enhancing translational efficiency.

Second, we examined three different 3'\textsuperscript{UTR} configurations. Several prior studies have reported that intervening intron sequences can have a major role on gene expression, both in vitro (Buchman and Berg, 1988; Nott et al., 2004) and in vivo (Choi et al., 1991). For instance, Buchman et al. showed that including an
intron in the transcription unit of a recombinant SV40 virus carrying a rabbit β-globin gene sequence resulted in up to a 400-fold increase in mRNA expression. Similarly, Choi et al. showed that insertion of intervening sequences in a transgene caused a 5- to 300-fold increase in expression (Choi et al., 1991). Other studies have shown that a β-globin polyadenylation sequence results in greater mRNA accumulation than a SV40 polyadenylation site (Westwood et al., 1993). However, the effect on expression of placing these different 3′-UTR sequences into a given transgene or allele remains poorly quantitated. In this study, we observed at least a twofold increase in FP expression and a 1.75-fold increase in translation when using a β-globin splice and polyadenylation sequence compared with an allele that only contains a β-globin or SV40 polyadenylation signal sequence. Thus, the cassette containing both the Xenopus 5′-UTR translational enhancer sequence and the intron-containing rabbit β-globin polyadenylation sequence provided the highest FP expression among the design variants that we tested.

Third, we assessed the relative brightness of five different FP variants when placed into the identical genetic context of a mammalian cell. Quantitative microscopic imaging enabled us to determine that Citrine was both the brightest and highest expressed FP, followed by Apple, Cerulean, eGFP and, finally, Cherry. These results differ significantly from prior measurements of relative brightness of purified FPs. However, the differences can be understood when one considers that our in vivo measurements also reflect differences in protein accumulation within living mammalian cells. Our measurements also reflect the excitation wavelength used, as well as the portion of the emission spectra that was measured. For instance, in our confocal measurements of FP brightness, we observed that Citrine that was excited at 514 nm was over fivefold brighter than eGFP that was excited at 488 nm, but there was no difference when brightness was assessed using a 488 nm laser in FACS. This difference reflects the attenuation of Citrine absorbance caused by the off-peak excitation by the 488 nm laser used in the FACS analysis.

We also directly compared Apple with Cherry and found that Apple is 1.7-fold brighter by confocal microscopy analysis and 1.4-fold brighter by FACS analysis. The slight differences observed are probably due to the different excitation wavelengths and filter sets used for the two analyses. Interestingly, Cherry displayed a punctate cellular labeling pattern, whereas Apple, like eGFP, Cerulean and Citrine, was uniformly expressed throughout the cells. The observed punctate Cherry pattern might be due to protein misfolding that results in cellular aggregates. Similar aggregates were observed when Cherry was overexpressed in the Phoenix Eco cell line (Schepmo et al., 2007). Although our findings indicate that Cherry is not as bright as Apple, the farther-red-shifted emission of Cherry might offer an advantage for experiments that involve multiple FPs. Because cellular autofluorescence is low in the red region, the minor expression differences between Cherry and Apple are unlikely to limit either for in vivo use. Indeed, even though Cherry was the least bright FP when assessed in cells, Cherry fluorescence from the ROSA26Cherry allele was easily visualized in 12 different tissues.

The ROSA26Cherry mouse also allowed us to assess the pattern and levels of Cherry expression in different tissues. We found that Cherry displayed punctate labeling in all tissues tested, similar to that observed in mESCs. Interestingly, although Cherry was detected in all tissues examined, expression varied up to 30-fold, with the highest expression being in skin and the lowest in muscle. The reason for these expression differences was not explored but could either reflect differences in ROSA26 promoter activity in different tissues in adult mice, or different FP turnover rates.

Although our studies were performed using only a single gene locus, they have implications for the design and use of other FP-containing transgenes and alleles. Although we did not design nor test alleles that were deliberately engineered for poor expression, on the basis of these data we can extrapolate that a poorly designed allele might be more than threefold lower in expression than the most highly expressed allele used in this study. The design of transgenes and/or alleles is highly variable in terms of sequences placed on the 5′ and 3′ sides of a coding sequence, such as those for an FP. Moreover, splice and polyadenylation sequences from endogenously expressed genes might intentionally be retained when gene targeting is used to insert an FP into a gene, even though the impact of these sequences on RNA stability is almost never known. Although a threefold difference in expression is seemingly small, in some settings it might mean the difference between experimental success and failure, especially if the excitation wavelengths available for use result in off-peak excitation or if the FP emission spectra overlaps with other proteins that exhibit significant autofluorescence.

METHODS

Plasmids

Four general use plasmids were made using standard cloning techniques and confirmed by restriction enzyme analysis and sequencing. The following four plasmids were made: (1) pLCA.71/2272, an insertion vector for gene targeting, contains tandemlox71 and lox2272 sites (Araki et al., 2002) flanking a fusion protein between puromycin N-acetyltransferase and a truncated form of HSV1 thymidine kinase (pu₄tk), under control of a mouse phosphoglycerol kinase (PGK) promoter, as well as an EM7-driven neomycin resistance (neo₄) cassette [from PL452 (Liu et al., 2003)]. The pu₄tk fusion gene does not cause male infertility in transgenic animals (Chen and Bradley, 2000), thereby allowing LCA-containing mESCs to be tested for germline competence. (2) pMCS.DT-A, a retrieval vector for gene targeting, contains an MCS flanked by tandem FLP recognition target (FRT) sites. These vectors have shown that a 5′-UTR translational enhancer sequence and the intron-containing rabbit β-globin splice and polyadenylation sequence might be retained when gene targeting is used to insert an FP into a gene, even though the impact of these sequences on RNA stability is almost never known. Although a threefold difference in expression is seemingly small, in some settings it might mean the difference between experimental success and failure, especially if the excitation wavelengths available for use result in off-peak excitation or if the FP emission spectra overlaps with other proteins that exhibit significant autofluorescence.

Gene targeting

The pROSA26LCA targeting vector was made by a two-step BAC recombineering procedure (Copeland et al., 2001). First, two inner homology regions from the ROSA26 gene locus were PCR-amplified from a mouse ROSA26 BAC (clone number: 58-D17, RPCI-22 library) and cloned into pLCA.71/2272. The resulting construct was used to insert the lox71-, PGK-pu₄tk-, EM7-neo₄- and lox2272-containing cassette into the BAC. Second, two outer homology regions were cloned into pMCS.DT-A, and the resulting
plasmid was used to retrieve a longer segment of DNA containing the loxped PGK-puAStk and EM7-neo<sup>R</sup> cassette, thereby generating the pROSA26.LCA gene targeting vector. 200 μg of this plasmid was linearized with NotI and used to electroporate 3.5×10⁵ TL-1 ESCs (Labosky et al., 1994) at 240 V and a capacitance setting of 500 μFD. After selection using 1.5 μg/ml puromycin (Sigma) for 7-8 days, surviving colonies were isolated for DNA analysis by Southern blot.

**FP exchange vectors**

pROSA26.Ex, which was made from pMCS.66/2272 and contains a 4.081 kb fragment of the mouse ROSA26 promoter that was removed from the gene during the generation of the ROSA26<sup>LC</sup>A allele, served as the starting point for assembling the FP-containing exchange vectors. CFP (Cerulean) was obtained from pCerulean-C1 vector (Rizzo et al., 2004), eGFP from pEGFP-N1 vector (Clontech), YFP (Citrine) from pmCIT-N3 vector (Rizzo et al., 2004), Cherry from pmCherry-C1 vector (Clontech), and Apple from pmApple-N1 vector (Clontech). Five of the exchange cassettes, EN.CFP, EN.GFP, EN.YFP, EN.Cherry and EN.Apple, contain the 5′-UTR from the Xenopus β-globin gene, which has been reported to function as a translational enhancer (Falcone and Andrews, 1991). In addition, these exchange plasmids contain an 886 bp DNA fragment from pB1-LG (Clontech) containing the 3′ end of exon 2, intron 2, exon 3 and polyadenylation sequences of the rabbit β-globin gene. The nonEN.CFP exchange vector lacks the translational enhancer sequence but is identical in other respects. In EN.CFP,bpA, CFP is followed by a 540 bp intronless polyadenylation sequence from the rabbit β-globin gene, also cloned from pB1-LG. EN.CFPSV40 contains an SV40 polyadenylation sequence cloned from pTet-on-Advance (Clontech). Each exchange vector contains a PGK/EM7-neo<sup>R</sup> cassette flanked by FRT sites. Annotated sequences for these FP-negative containing exchange vectors were deposited into GenBank (accession numbers from HM771696 to HM771703). The vectors are available upon request.

**Recombinase-mediated cassette exchange**

RMCE was performed using a positive-negative selection strategy previously described (Long et al., 2004). In brief, 5.6×10⁶ mESCs containing the ROSA26<sup>LC</sup>A allele were co-electroporated with 40 μg of an exchange plasmid and 40 μg of pB185, a Cre-expression plasmid (Sauer and Henderson, 1989). 200 μg/ml neomycin (Invitrogen) was used for positive selection and 8 μM gancyclovir (Sigma, St Louis, MO) for negative selection. Surviving clones were screened by DNA PCR using the following primers: (1) 5′-AGACT-TATCTAATCATGGTAGT-3′, (2) 5′-GTGAGAACAGATAGTACATCTG-3′, (3) 5′-GGAGATCATAACTCCAGCAT-3′, (4) 5′-TCACAAAGCATAATACCTGTAGT-3′. Primers 1 and 2 amplify a 537 bp DNA fragment from an exchanged allele and 503 bp DNA fragment from the wild-type ROSA26 locus. Primers 3 and 4 detect a 565 bp sequence only in an exchanged allele. Properly exchanging clones were expanded from the master plate and stocked for future use.

**Blastocyst microinjections**

All procedures performed on mice were approved by the Vanderbilt University Animal Care and Use Committee. Chimeric mice were generated by the microinjection of three mutant mESCs: R26.EN.CFP clone 5B9:1C3, R26.EN.Cherry clone 5B9:1C11 and R26.EN.GFP clone 5B9:1D10. Four cells for each clone were injected into each blastocyst then transferred into pseudopregnant females. After 15.5 and 17.5 days of gestation, pregnant females were euthanized and chimeric embryos were imaged by fluorescent microscopy. Several post-injection litters were brought to term, and three chimeras were mated with C57BL/6 females to test for germline transmission. Confirmation of germline transmission of the EN.Cherry allele and genotyping of established ROSA26<sup>Cherry</sup> mice were carried out by PCR.

**Quantitative reverse-transcription PCR**

mESC clones were grown to 70% confluency and total RNA isolated using the RNeasy mini kit (Qiagen). cDNA was synthesized from 2 μg RNA with Omniscript reverse transcriptase and oligo(dT) primers (Applied Biosystems). Quantitative reverse-transcription (RT)-PCR was performed using a 7900HT Real-time PCR System (Applied Biosystems) with 2 ng of template and Power SYBR Green PCR Mastermix (Applied Biosystems). The primers used to detect CFP sequences were 5′-TGCTTGGCGCAACC- ACTAC-3′ and 5′-CGGTACGACATCCAGCAGG-3′. β-actin cDNA was detected with 5′-ACATATGCTCCCCGGGCTGTAGTCTTC-3′ and 5′-TCTCTTTGCTCTGGCCTGTCACC-3′. The ΔΔCt method was used to determine relative expression. Data was normalized to β-actin and expressed relative to the ROSA26<sup>LC</sup>A allele.

**Quantitative confocal microscopy**

For analysis of mESCs, cells were grown on glass-bottom dishes (MatTek) on a monolayer of irradiated MEFs. For analysis of mouse tissues, different organs from two mice (6-weeks old and 20-weeks old) were dissected in PBS, cut into thin slices with a razor blade and mounted under coverslips in Aqua Poly/Mount (Polysciences) on MatTek dishes. Confocal microscopy was performed on a LSM710 (Zeiss) using a Fluor 40× 1.3 NA oil-immersion objective. The following excitation wavelengths and emission bandpass filters were used for FP detection: eGFP: 488 nm and 494-650 nm bandpass; Citrine: 514 nm and 520-640 nm bandpass; Cerulean: 561 nm and 520-640 nm bandpass; Cherry and Apple: 561 nm and 641-665 nm bandpass. All images were acquired using 512×512, 0.415 μm diameter pixels and 12-bit gray levels. DIC images were acquired simultaneously with the fluorescence images. For each clone, ten fields of view per dish were measured, and detector gain was set to avoid saturation. After each imaging experiment, serial dilutions of the respective purified protein solution were quantified using the same settings to establish a calibration curve for the conversion of gray level values into protein concentration values. Image analysis was performed using ImageJ 1.43s software (http://rsb.info.nih.gov/ij/). DIC images were used to create regions of interest (ROI) encompassing each cell. The ROI was used to determine mean fluorescence intensity of each cell, and the area outside the ROI was used to determine background to be subtracted from the fluorescence intensity. Intensity values were converted to concentration values by interpolation with the calibration curve.

**FACSs**

Cells were grown to confluency on gelatin-coated plates without MEFs, harvested with 0.25% trypsin/EDTA and resuspended in

Expression of fluorescent proteins using ROSA26

[545]
expression of fluorescent proteins using ROSA26

BACKGROUND

Fluorescent proteins (FPs) are widely used to monitor gene expression and study cellular dynamics in cells and tissues. However, many variables can significantly affect experimental outcomes when FPs are expressed in the context of transgenes or gene targeting vectors in cell lines and mice. Variables that are intrinsic to the FP itself include the excitation and emission wavelengths, as well as the quantum yield of the FP. Variables extrinsic to the FP include the transcriptional activity of the specific promoter used to drive expression of the FP, the translational efficiency of the mRNA generated and the stability of the mRNA. Although many of the intrinsic factors have been precisely characterized, the extrinsic factors have been more difficult to accurately assess. Specifically, the variability that occurs owing to transgene copy number and insertion site has been a major impediment in quantifying the genetic factors that influence FP reporter gene expression in cells and mice.

RESULTS

The authors of this study overcome this difficulty by using recombinase-mediated cassette exchange (RMCE) to generate an allelic series in mouse embryonic stem cells (mESCs), thereby allowing the quantitative assessment of intrinsic and extrinsic variables that influence FP expression in cells and mice. Eight different ROSA26 alleles are generated and compared to one another to assess the effects of intrinsic and extrinsic variables on FP protein expression in cells, as well as the relative brightness of different color variants of FPs. The authors show that the combination of genetic elements in an expression cassette influences FP expression up to threefold, and that relative expressed brightness of individual FPs varied up to tenfold in relation to one another. Among the five FPs examined, Citrine was the brightest, followed by Apple, eGFP, Cerulean and finally Cherry. In vivo expression of Cherry in ROSA26Cherry mice was observed in all organs, yet varied up to 30-fold between different tissues.

IMPPLICATIONS AND FUTURE DIRECTIONS

These results should help investigators to make more informed choices regarding the design of transgenes or vectors for expressing FPs and other proteins in cells and mice. In addition, the basal plasmids used in this study have design features that enable their use in BAC recombineering strategies and that simplify other aspects related to the design of gene targeting and RMCE vectors. Overall, these resources will simplify the generation of new colored cassette acceptor (LCA) alleles, gene targeting and RMCE vectors, and mouse models, especially those containing FPs.

REFERENCES

Labosky, P. A., Barlow, D. P. and Hogan, B. L. (1994). Mouse embryonic germ (EG) cell lines: transmission through the germine and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines. Development 120, 3197-3204.
Expression of fluorescent proteins using ROSA26


