

Optimizing Techniques for Tracking Transplanted Stem Cells In Vivo

TIMOTHY R. BRAZELTON,^{a,b} HELEN M. BLAU^a

^aBaxter Laboratory in Genetic Pharmacology, Department of Microbiology and Immunology and Department of Molecular Pharmacology, Stanford University School of Medicine, Stanford, California, USA; ^bDepartment of Surgery, The Children's Hospital of Philadelphia, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

Key Words. Adult bone marrow stem cells • Green fluorescent protein • Skeletal muscle • Tissue regeneration • In vivo tracking

ABSTRACT

The potential for bone marrow–derived cells (BMDCs) to contribute to nonhematopoietic tissues has generated considerable debate in recent years. Causes for the controversies include disparities in the techniques used to track engraftment of BMDCs, inappropriate tissue preparation, a lack of appropriate positive and negative controls, and basic misunderstandings about how to properly collect and interpret images from epifluorescent and confocal microscopes. Our laboratory was among the first to use bone marrow transplants from transgenic mice constitutively expressing enhanced green fluorescent protein (GFP) to study the ability of BMDCs to give rise to nonhematopoietic tissue types, a system that is now in widespread use. During our 6 years of experience using GFP,

as well as beta-galactosidase and the Y chromosome, to track BMDCs in vivo, we have identified many difficulties and have developed techniques to resolve them. We discuss several of these methods, and, in particular, we describe ratiometric analysis techniques for improving detection of transplanted cells derived from genetically modified bone marrow. Finally, to help resolve reported discrepancies regarding the frequency with which BMDCs contribute to skeletal myofibers, we demonstrate that the pattern of highly autofluorescent myofibers in skeletal muscle is clearly distinct from that of GFP-expressing myofibers and describe how unambiguous conclusions can be drawn from such data. *STEM CELLS* 2005;23:1251–1265

INTRODUCTION

The extent to which bone marrow–derived stem cells (BMDCs) contribute to nonhematopoietic tissues in vivo has been a topic of considerable debate over the past several years [1–6]. Different groups have reported disparate results, with some groups failing to observe the phenomena entirely and others reporting dramatically high contributions of BMDCs to various tissues. Some of the reports of adult stem cell plasticity seem, in retrospect, to be due to misinterpretation of microscopy results or flawed experimental design. In other cases, these failures were likely attributable to methodological difficulties that were unrecognized before publication due to a lack of positive controls. Because contradictory findings continue to be published, further experimentation is necessary to resolve the debates. To minimize additional confusion

in the field, care must be taken in designing future experiments so that they accommodate the distinct limitations that exist for each of the various methods available. Our hope is that through a frank discussion of the advantages and limitations of each technique and the methods for optimizing their use, the number of misleading reports will be minimized and the field will more quickly reach a consensus regarding these exciting biological phenomena.

BMDCs have been reported to contribute to a variety of non-hematopoietic cell types in adult mice, rats, and humans. A key feature of such contributions is nuclear reprogramming of the BMDC, whether it occurs in response to extracellular signals or in response to intracellular signals after fusion. Initial findings that BMDCs could contribute to skeletal myofibers and activate previously silent muscle-specific promoters, albeit at a low fre-

Correspondence: Helen M. Blau, Ph.D., Stanford University School of Medicine, 269 W. Campus Drive, Stanford, California 94305-5175, USA. Telephone: 650-723-6209; Fax: 650-736-0080; e-mail: hblau@stanford.edu Received April 4, 2005; accepted for publication July 19, 2005; first published online in *STEM CELLS EXPRESS* August 18, 2005. ©AlphaMed Press 1066-5099/2005/\$12.00/0 doi: 10.1634/stemcells.2005-0149

quency, challenged existing dogma about the restricted differentiation potential of BMDCs [7, 8]. Subsequent reports confirmed the low frequency of incorporation into most skeletal muscles but documented that this frequency varies 1,000-fold among muscles, reaches significant levels in specific muscles (5%), and is increased after tissue damage or an extended period of exercise (3.5%) [9, 10]. Moreover, this contribution was shown by parabiosis not to be associated with variables such as the irradiation inherent to the bone marrow transplant procedure used in most studies, suggesting that it was a response to local physiological or damage signals [11].

Several groups reported that BMDCs contribute to cardiomyocytes at a low frequency but that the frequency can increase after injury to the heart [12, 13]. Other groups have reported a very high frequency (up to 68%) of BMDCs in cardiac tissue [14]. Resolution of these major discrepancies may benefit from the techniques presented here.

BMDCs have been shown to contribute to liver hepatocytes in both animals and humans [15–18]. This ability was most clearly demonstrated when, after a bone marrow transplant with genetically healthy cells into mice lacking fumarylacetoacetate hydrolase, BMDCs contributed to 30%–50% of the liver and their nuclei were reprogrammed to express the missing enzyme, rescuing the animals from death [19].

BMDCs have also been reported to contribute to various cell populations in the central nervous system (CNS), including cells with protein expression patterns typical of microglia, astrocytes, and neurons [20, 21]. Subsequently, several studies have demonstrated that BMDCs contribute to Purkinje neurons in both humans and mice by a mechanism involving cell fusion. After fusion with a Purkinje neuron, it has been shown that the BMDC nucleus is reprogrammed to express a Purkinje gene. It has been suggested that such cell fusion events may be a means by which essential and morphologically complex cells are rescued from damage or injury [22, 23].

Our laboratory has more than 6 years of experience using transgenic mice that constitutively express enhanced green fluorescent protein (GFP) as bone marrow donors in experiments designed to track adult stem cells [21]. We demonstrated that the expression of GFP serves as a powerful tool for following the fate of transplanted BMDCs as they move into nonhematopoietic tissues and, occasionally, adopt new programs of nonhematopoietic gene expression [7, 9, 10, 21, 23]. This “fate change” can be either at the single-cell level or after fusion but always involves nuclear reprogramming to express genes that were previously silent and that are characteristic of the surrounding tissue. Before adopting GFP as our preferred marker for cell tracking, we spent longer than 1 year using ROSA26 mice, which constitutively express beta-galactosidase (B-gal), as bone marrow donors for studies of in vivo cell plasticity. We ultimately discarded the ROSA26 mice as unsuitable for detecting rare events because of both low levels

of endogenous B-gal activity in many cell types and rare occurrences of high endogenous B-gal activity in individual cells (see B-gal section below).

Exceptionally robust and unambiguous tracking systems are essential for monitoring rare events such as those reported for the contribution of circulating adult stem cells to various nonhematopoietic tissues. A low frequency of positive events needs to be balanced against the real possibility of an equally low frequency of false-positive events. For example, the upregulation of endogenous B-gal activity in rare cells needs to be characterized in a given tissue before using B-gal as a marker to evaluate endogenous cells. In the case of mice transplanted with ROSA bone marrow, given the variable endogenous expression of the recipient's own B-gal activity, this would require a labor-intensive evaluation of tissues from numerous mice transplanted with wild-type bone marrow and subjected to the same experimental protocol.

Because increasing numbers of scientists study the plasticity of adult stem cells in vivo using tracking markers such as GFP or B-gal, we think it is timely to share techniques developed in our laboratory as these tracking systems have been refined over the last 6 years. In particular, we demonstrate here how ratiometric imaging techniques allow the rapid and specific identification of GFP-expressing cells in tissues that have significant autofluorescence. Both GFP and B-gal have idiosyncrasies that require careful planning and appropriate experimental controls to avoid many of the common pitfalls that we and others have experienced while tracking BMDCs and evaluating the associated nuclear reprogramming.

MATERIALS AND METHODS

Bone Marrow Transplantation

Bone marrow was harvested from 8- to 10-week-old male transgenic mice that ubiquitously expressed GFP [24], B-gal (ROSA26 mice; Jackson Laboratory, Bar Harbor, ME, <http://www.jax.org>), or wild-type C57BL/6 mice (Stanford's in-house colony). The marrow of 8- to 10-week-old isogenic C57BL/6 recipient mice was ablated by lethal irradiation (9.5 Gy), after which each mouse received 6×10^6 nucleated unfractionated bone marrow cells by tail vein injection. By 8 weeks after transplant, $\geq 90\%$ of the circulating, nucleated cells were BMDCs in $>95\%$ of recipient mice. The care of all animals included in these studies was in accordance with Stanford University's institutional guidelines.

Intravascular Fixation

Recipient mice were anesthetized with 60 mg/kg Nembutol, and the heart was surgically exposed by incision with a small scissors. The right atrium was clipped to allow blood to exit, and the tip of a blunted 20-gauge needle was placed in the left ventricle. A perfusion pump was used to infuse 30 ml of 0°C 100-mM sodium phosphate buffer (PB) (pH 7.4) followed by 30 ml of 0°C 1.5% freshly

dissolved paraformaldehyde (PF) plus 0.1% glutaraldehyde in sodium PB at a rate of approximately 2.5 ml per minute. Lower concentrations of PF will decrease the amount of GFP held in the tissue, but concentrations down to 0.05% will often retain sufficient GFP to allow direct visualization. The addition of 0.1% glutaraldehyde results in a slight reduction in the background autofluorescence in several tissues, including skeletal muscle and the CNS, but is not necessary to hold in the GFP. Note that if additional fixation is desired, harvested organs or pieces of tissue can be placed in 0.5%–1.5% PF/0.1% glutaraldehyde/20% sucrose in PB (pH 7.4) at 4°C. The tissue is then fixed for 1 hour for every 2 mm of tissue penetration needed. It is essential to maintain the solution at or below 4°C for this entire step to prevent overfixation.

After fixation, organs and tissues were placed in 20% sucrose in PB at 4°C for 1 hour for every 2 mm of tissue depth. The 20%-sucrose wash provides both cryoprotection and reduces the background, presumably by decreasing the concentration of PF in the tissue. Tissues were then blotted firmly on paper towels to remove excess fluid, embedded in TISSUE-TEK OCT compound (Sakura Finetek, Torrance, CA, <http://www.sakuraus.com>), and snap-frozen in a vial of isopentane suspended in liquid nitrogen.

NADH Activity Assay

The NADH activity in 12- μ m sections of fixed skeletal muscle was determined by incubation in 1.5 mM NADH (N-4505; Sigma-Aldrich, St. Louis, <http://www.sigmaaldrich.com>) and 1.5 mM Nitro-blue tetrazolium (N-46876; Sigma-Aldrich) in 37°C, 0.2 M Tris (Sigma-Aldrich) for 20 minutes.

Fluorescent In Situ Hybridization

The Y-chromosome fluorescent in situ hybridization (FISH) probe (a gift from E. Gussoni) was prepared by labeling 1 μ g of plasmid DNA with digoxigenin-11-dUTP as described [25, 26]. FISH was standardized on whole nuclei isolated from a male murine muscle cell line and on male muscle tissue sections. In situ hybridization was performed as described [26].

Microscopic Analysis

Each GFP⁺ cell was analyzed for GFP expression using epifluorescence and/or three-dimensional, confocal laser-scanning microscopy (Zeiss LSM 510; Carl Zeiss, Jena, Germany, <http://www.zeiss.com>). The filter sets used for the epifluorescent images contained the following excitation (EX), dichroic (D), and emission (EM) filters: GFP bandpass (BP) (EX 480/40, D 505LP, EM 535/50), rhodamine (EX 540/25, D 565LP, EM 605/55), GFP longpass (LP) (EX 480/40, D 505LP, EM 510LP), and dual fluorescein isothiocyanate/tetramethylrhodamine isothiocyanate BP (EX 482/20 and 545/30, D 525/55 and 570LP, EM 520/15 and 600/40). Epifluorescent images were captured in true color with a G5 digital camera (Canon, Tokyo, <http://www.canon.com>).

To capture the confocal images, each fluorophore was imaged individually with a single laser to prevent cross detection of fluorophores. The confocal settings used to image the GFP channel consisted of Argon laser excitation, EX 488, D 570, D 545, and EM 518/13 and for the rhodamine channel consisted of 543-nm HeNe laser excitation, EX 543, D 635, D 545, and EM 588/55.

RESULTS AND DISCUSSION

Demonstration of Colocalization Requires Confocal or Deconvolution Microscopy

Identification of a cell that has changed its fate relies on the identification of two kinds of markers: a tracking marker and an identifying marker. The tracking marker indicates the original identity of the cell. For example, a stem cell genetically engineered to express GFP constitutively can be transplanted into a wild-type mouse and its progeny can be detected by GFP expression, even years later. Other tracking markers include B-gal, membrane-binding dyes, and the Y chromosome that can be used to track male cells transplanted into female recipients.

In addition, one or more identifying markers are needed to determine that the nucleus of a given cell was reprogrammed to express new nonhematopoietic genes. For example, if a GFP-expressing BMDC is observed in the CNS, then expression of neuronal, astrocytic, or oligodendroglial proteins needs to be determined. Most rigorously, to claim that a cell has acquired a new identity, multiple proteins indicative of the new cell fate should be demonstrated as well as a loss of proteins indicative of the previous cell fate. The identifying markers are frequently proteins that are unique to specific cell types but can also include distinctive morphological features or functional characteristics of specific cell types. For example, Purkinje neurons and skeletal myofibers expressing GFP can be reasonably identified based on their distinctive morphological features (Figs. 1A, 1B). Similarly, the identity of BMDCs as neuron-like cells was bolstered by the detection of membrane depolarizations in response to specific neurostimuli [27].

When documenting a cell fate change, it is essential to demonstrate that the tracking marker and the identifying markers are expressed in the same cell. Cells can be arranged in morphologically complex ways, and a close apposition of two distinct cell types may represent a normal biological interaction. When tissue sections are analyzed in two dimensions, such as occurs with standard microscopy, a cell with a tracking marker may overlie a cell with an identifying marker, resulting in the mistaken appearance of a single cell that has changed its nuclear program or gene expression pattern. Although such events are rare, the reported frequencies of adult stem cells contributing to nonhematopoietic tissues are often equally rare. Thus, methods that allow the visualization of thin optical sections and three-dimensional reconstructions of cells in tissue sections are required. The most common methods to accomplish this are laser-scanning confocal microscopy and deconvolution microscopy.

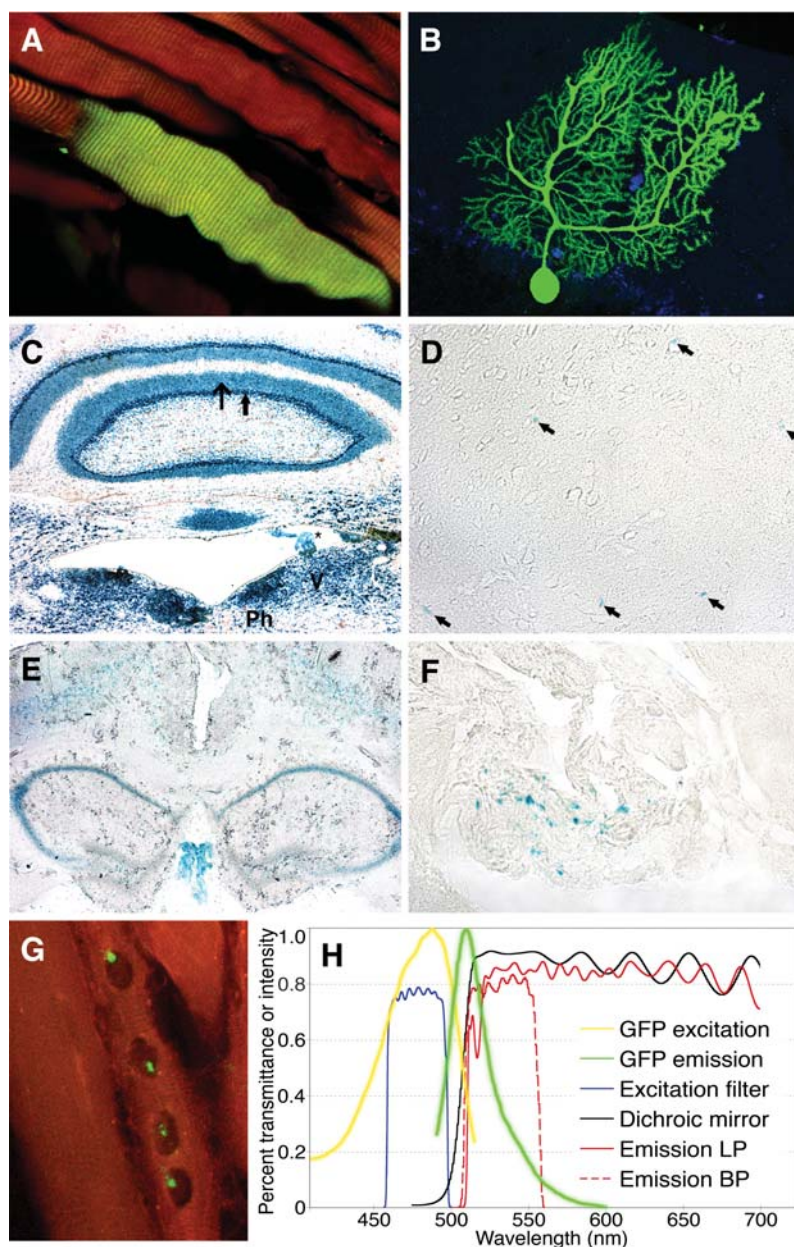


Figure 1. Several methods have been used to track transplanted cells, including (A, B) enhanced green fluorescent protein (GFP), (C–F) beta-galactosidase (B-gal), and (G) the Y chromosome. After a bone marrow transplant with GFP-expressing bone marrow into wild-type mice, GFP-expressing nuclei contribute to both (A) skeletal myofibers and (B) Purkinje neurons. Because GFP (green in A, B) is a small, soluble protein with a rapid diffusion rate, it fully distributes throughout the cell and highlights distinguishing morphological features such as (A) sarcomeric banding and (B) dendritic extensions. Compared with (C) ROSA26 mice, (D–F) wild-type mice have decreased but still significant B-gal (blue) activity. There is substantial heterogeneity in the amounts of B-gal activity among cell types in (C) the cerebellum of ROSA26 mice and in (E) the hippocampus of wild-type mice (both stained at a pH of 6.0). Cell groups with increased endogenous B-gal activity include the perihypoglossal nucleus (Ph), the vestibular nucleus (V), and the Purkinje (thick arrow) and granular (thin arrow) layers. The B-gal activity in the choroid plexus (*) is relatively resistant to increases in pH (data not shown). Even when stained at a pH of 7.4, substantial numbers of cells in wild-type mice continue to express significant B-gal activity in uninjured central nervous system (arrows; D) and at the site of a central nervous system stab injury (F). (G): Multiple male nuclei are identified by *in situ* hybridization against the Y chromosome (green) in a skeletal myofiber (red). The nuclei in the other myofibers are outside of the plane of the optical section. (H): Graph of the interaction of the excitation and emission spectra of GFP with a standard bandpass (BP) and longpass (LP) filter. The y-axis indicates the percent transmittance for the filter data and the relative excitation and emission intensities for the GFP spectral data.

B-gal as a Tracking Marker

One of the earliest tracking markers used by many groups, including ours, was bacterial B-gal. The ROSA26 strain of mice, which constitutively express B-gal in most cell types, arose out of a promoter-trapping experiment in embryonic stem cells [28]. Since then, the ROSA26 promoter has been used to drive both B-gal and GFP expression [29]. Bacterial B-gal can be detected either by antibodies or by its ability to enzymatically cleave its substrate, a galactoside, thereby generating an amplified chromogenic or fluorescent product.

Antibodies specific for B-gal from *Escherichia coli* have been used with mixed success to detect ROSA26 cells in tissue sections. The bacterial B-gal antibodies work well to detect B-gal when it is partially purified by laboratory procedures such as

Western blots. Antibodies also effectively detect bacterial B-gal produced at high levels in tissue sections. However, the weak B-gal expression seen in ROSA26 mice has proven problematic for many investigators, often yielding relatively high and variable levels of background antibody staining in ROSA26 tissue sections, with differences among tissues.

Several substrates are available to detect B-gal based on its enzymatic activity. These include X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside), which creates a blue chromogenic substrate. A variety of substrates that generate a fluorescent signal after cleavage have also been developed, such as fluorescein digalactoside (fluorescein di-beta-D-galactopyranoside; Molecular Probes Inc., Eugene, OR, <http://probes.invitrogen.com>).

It has been known for years that a wide variety of mammalian cells possess significant B-gal activity, which is expressed at higher levels in the liver, kidney, pancreas, spleen, uterus, thyroid, and intestines than in other tissues [30–32]. Thus, the appearance of a chromogenic or fluorescent product of B-gal cleavage simply indicates that cleavage occurred. It does not indicate whether the cleavage event was mediated by *E. coli* B-gal (i.e., the marker in a ROSA26 mouse), endogenous mammalian B-gal, or another enzyme or pathway with this catalytic capacity. Moreover, it does not rule out the possibility that cleavage occurred simply by an unrestrained oxidative reaction.

Difficulties are encountered when using B-gal from ROSA26 mice as a tracking marker for two reasons. First, compared with most experimental systems using B-gal as a marker, the expression of B-gal in ROSA26 mice is relatively weak. Consequently, B-gal expression in ROSA26 mice can be difficult to distinguish from endogenous, mammalian B-gal activity.

Second, when rare events or unexpected results are observed using ROSA26 cells, it is necessary to rule out the possibility that the observed increase in B-gal activity is not due to an unanticipated upregulation of endogenous B-gal activity. Rare cells occasionally exhibit increased B-gal activity at a level similar to that seen in cells from ROSA26 mice (Figs. 1C–1F). Furthermore, some cell groups (e.g., hippocampal neurons, Purkinje cells, and the choroid plexus in the CNS) have substantially higher endogenous B-gal activity than adjacent cell types (Fig. 1E). Some groups were misled by these normal variations in endogenous B-gal activity and first reported and then subsequently retracted their conclusions that BMDCs had replaced large groups of CNS neurons [32]. Thus, a discussion of its optimal use is in order.

ROSA B-gal can be distinguished from endogenous B-gal by two characteristics. First, the B-gal activity in cells from ROSA26 mice frequently occurs within intracellular vesicles. Although the actual identity of these vesicles is not well characterized, the vesicular, punctate appearance of bacterial B-gal activity in ROSA26 cells is distinctive.

Second, although both the mammalian and bacterial B-gal exhibit increased enzymatic activity in acidic conditions, the activity of mammalian B-gal is less resistant to increases in pH. Thus, it is important for each laboratory to use both positive and negative controls at various pH levels to determine an optimal pH. Many laboratories find that a pH between 7 and 8 is reasonable, but even at a pH of 7.4, there can be significant endogenous B-gal activity depending on the tissue (Figs. 1D, 1F). Techniques to optimize B-gal–based tracking systems in the CNS have been reviewed recently [32].

In our personal experience, we ultimately decided that the ROSA26 system lacked sufficient robustness to make strong claims about rare and unexpected phenomena because the signal-to-noise ratio was so dependent on pH and was confounded by endogenous enzyme activity. As a consequence, we repeated all

of our ROSA26 experimental work using GFP-labeled BMDCs before our initial publication because GFP is advantageous in that it has no endogenous counterpart in mammals [21].

Regardless of the detection method used, proper fixation is essential to ensure that most of the enzyme remain within the tissue to maximize the signal strength. This typically involves intracardiac perfusion, fixation, or immersion of the target organ in fixative solution before sectioning. Attempts to fix ROSA tissue after sectioning typically result in most B-gal enzyme diffusing away from its tissue location as soon as the section is immersed in solution. Even very concentrated solutions of cross-linking fixatives such as PF or glutaraldehyde fail to hold most of the enzyme within the cell expressing it. This problem is analogous to, although not as severe as, that seen with the fixation of GFP in tissue sections (discussed below).

As we gain a better understanding of the biology and with careful experimental controls, the ROSA26 labeling system will again be a reasonable approach. The biggest risk, in our opinion, remains that of the rare cells that contain a level of endogenous B-gal activity as high as that seen in ROSA mice. If an experimental protocol, such as stress or ischemia, increases the frequency or strength of this endogenous B-gal expression, the results may be misinterpreted. For this reason, the documentation of novel cell fate transitions or reprogramming of nuclear gene expression using B-gal as the tracking marker will usually be suspect.

In Situ Hybridization to Track the Y Chromosome

The Y chromosome can serve as an excellent tracking marker. The Y chromosome is typically detected by in situ hybridization with fluorescent probes that result in single or multiple dots of fluorescence over a specific region of the Y chromosome (Fig. 1G). Under the proper conditions, both the specificity and sensitivity of in situ probes for the Y chromosome are very high. To minimize the incidence of false-positive Y-chromosome detection, it is essential that appropriate negative controls (i.e., female tissue) be processed and evaluated in parallel with the experimental tissue. Although laborious, this is the only method described here that allows the contribution of individual nuclei to be ascertained because both GFP and B-gal diffuse throughout the cytoplasm.

A greater problem with the Y chromosome is that because it lacks morphological characteristics, special care must be taken to demonstrate that the Y chromosome is actually contained within the cell of interest. This requires an approach that generates a high-resolution, three-dimensional representation of the boundaries of the cell in question to demonstrate that the Y chromosome is contained within that boundary.

Membrane Dyes

An additional class of tracking markers are the membrane dyes. A large number of membrane dyes are available, but one in particular, PKH26, has been used to track transplanted adult stem cells [33].

A unique feature of membrane dyes compared with the other tracking markers discussed here is that they are not encoded within the genome and, as a result, the amount of dye present in a cell diminishes by half with each cell division. Thus, if a labeled cell gives rise to a large number of progeny, none of the progeny may contain sufficient membrane dye to be recognized. On the other hand, if a cell containing a membrane dye is observed to have a new pattern of gene expression, it implies that a limited amount of proliferation occurred.

An important caveat of membrane dyes is that if an unlabeled phagocytic cell ingests a labeled cell, a process involving membrane fusion, membrane dye can be transferred to phagocyte membranes [34]. The frequency with which this occurs in vivo has not been established, and many of the studies using membrane dyes as a tracking marker have not adequately controlled for this possibility.

Strictly speaking, phagocytosis of a cell with any tracking marker could lead to a transient presence of that marker within the engulfing cell. However, although membrane components of a phagocyte and its target can merge, the remainder of the target cell is enclosed in a phagolysosome with a proteolytic environment that should rapidly destroy the fluorescent or enzymatic capacities of most tracking proteins [35]. This possibility has not been experimentally excluded in many published reports.

GFP

GFP is a unique protein with several positive features that make it an exceptional cell-tracking marker in eukaryotic cells and a few negative features that require special consideration. The green fluorescent protein that has made its way into common usage was originally derived from the jellyfish *Aequorea aequorea* and was subsequently modified for more rapid folding, improved stability, and greater brightness [36]. However, the initial green fluorescent protein variants suffered from limited expression in some organisms. To increase the expression in mammalian cells, an enhanced green fluorescent protein gene (referred to here as GFP) was synthesized that incorporated 190 silent base mutations, resulting in a gene with an open reading frame consisting only of preferred human codons. GFP is currently the most frequently used green fluorescent protein variant and has an excitation peak at 489 nm and an emission maxima at 508 nm. GFP is very stable and its fluorescent properties are preserved at temperatures up to 65°C and up to a pH of 11 [37, 38].

Importantly, GFP does not require an enzymatic substrate nor a cofactor for its fluorescent property, allowing it to be used as a reporter in both prokaryotic and eukaryotic cells. Furthermore, the GFP fluorophore is formed by the post-translational, autocyclization of Ser⁶⁵, Tyr⁶⁶, and Gly⁶⁷, a process requiring only the presence of oxygen [39]. This autocatalytic mechanism allows GFP to fold correctly in a wide diversity of organisms.

GFP is a small, soluble protein (236 to 265 amino acids) that distributes rapidly with diffusion rates at 37°C of 25 μm^2 per second in cytoplasm and 87 μm^2 per second in water [40]. The ability

of GFP to rapidly diffuse aids in the assessment of cell morphology, allowing it to fill even the thin dendritic and axonal extensions of neurons (Fig. 1B).

The expression of GFP does not seem to be overtly toxic to cells. The few reports of toxicity occurred only at very high GFP concentrations in transfected cells and seem to be the exception [41, 42]. Most GFP applications have not resulted in overt toxicity. Perhaps the best evidence for GFP's low toxicity is the large number of transgenic organisms that constitutively express GFP and yet are able to develop and breed normally. GFP has been transgenically expressed in *Dicytostelium* and *Arabidopsis thaliana* [43, 44], *Drosophila* [45], *Caenorhabditis elegans* [46], zebrafish [47, 48], mice [24], rabbits [49], and monkeys [50].

In our hands, the line of GFP-expressing transgenic mice that we use [24] work exceptionally well as cell donors for tracking studies (hereafter referred to as Okabe GFP mice). In addition to the Okabe GFP mice, several other lines of mice are available that ubiquitously express GFP or a GFP variant [51, 52]. The Okabe GFP mice seem to be physiologically healthy. They breed relatively well and reach an adult weight that is similar to wild-type C57BL/6 mice. The elements regulating GFP expression in these mice include a chicken β -actin promoter, a cytomegalovirus enhancer, a beta-actin intron, and a bovine globulin poly-adenylation signal [24]. Interestingly, when wild-type GFP or a variant similar to S65T was expressed using the same promoter, its expression was limited to specific tissues, such as skeletal muscle and pancreas, and it was never expressed in the brain or blood vessels [53]. Whether this was due to the GFP variant used or to the location of the transgenic insertion is not clear.

We have successfully used the Okabe GFP mouse line as donors to track the fate of transplanted bone marrow cells in several tissues. However, in our experience, this mouse line does have a few idiosyncrasies. First, a standard UV light source induces sufficient dermal fluorescence in newborn and adult mice to allow rapid characterization of GFP expression. However, caution should be used when illuminating GFP pups with the UV light source in the presence of their mother because, in our experience, this can increase the incidence of maternal infanticide.

Second, the level of GFP expression can vary among mice and the expression levels can be increased or decreased with selective breeding. Thus, based on the expression level of GFP, we routinely characterize our mice as negative, dim, moderate, or bright. Breeding brighter males to brighter females increases the expression levels over time but eventually seems to result in reduced breeding success (smaller newborns, decreased litter sizes, increased incidence of still births). We suspect, but have not formally tested, that too high of an intracellular GFP concentration eventually begins to compromise normal physiology. On the other hand, mice with lowest levels of dermal GFP expression occasionally fail to express detectable GFP in all tissue types. Thus, we suggest that colonies should be bred to maintain a moderate level

of GFP expression. This is best accomplished by breeding moderate males to moderate females because most pups resulting from such a mating will also express GFP at a moderate level.

Evaluation of GFP in Tissue Requires Fixation Before Sectioning

As mentioned above, GFP is a small, soluble protein with a rapid rate of diffusion in cytoplasm and water. If membrane integrity is compromised through cryopreservation or sectioning, most GFP will be lost within seconds of immersion of a tissue section in solution. This is true even when that solution is a concentrated fixative such as 37% PF. It is worth noting that GFP fluorescence can be observed directly in sectioned, unfixed tissue before immersion in solution or mounting, a feature that serves as a useful positive control when testing fixation protocols.

Thus, to quantitatively assess the expression of GFP, it is essential to fix the GFP into the tissue before cryopreservation and sectioning. Our laboratory uses standard intracardiac perfusion to intravascularly deliver dilute fixative throughout the mouse. Some investigators include an anticoagulant in the initial PB perfusion, but in our laboratory the addition of either heparin or EDTA does not seem to improve the degree or uniformity of fixation. Intravascular perfusion occasionally fails if the perfusion rate is too rapid because this causes fluid to back up into the lungs, where it compromises vascular integrity and allows the perfusate to exit through the oropharynx. Once this pathway has been initiated, most perfusate will travel via this route even if the perfusion rate is decreased. If a mouse has been adequately perfused with fixative, it should be cold and stiff. If poor perfusion is suspected, an additional fixation step can be performed after harvesting, as described in Materials and Methods.

The level of fixation required to fully hold GFP in some tissues can prevent the adequate penetration of antibody or *in situ* probes. If tissues cannot be fixed before sectioning, slightly higher amounts of GFP can be retained by heating the slide before depositing a tissue section on it to minimize condensation, followed by heat drying the cut tissue section onto the slide and then gently adding 37% PF that has been heated to 45°C. Alternatively, unfixed tissue containing GFP can sometimes be satisfactorily fixed with formaldehyde fumes. However, even if great care is taken to prevent condensation from forming on the tissue section, the results with the fumigation method are often variable.

The use of an antibody against GFP is the most sensitive method. It allows the identification of cells containing lower levels of GFP due to either insufficient fixation or a low level of expression. For example, in skeletal muscle in which each myofiber contains many nuclei, the use of an antibody to detect GFP increases the number of GFP-expressing myofibers observed by twofold to threefold (unpublished observations). In other words, anti-GFP antibody allows the detection of fibers that have very weak expression of GFP (i.e., probably contain only a few nuclei expressing GFP).

Even in tissue that was fixed only after cryosectioning and thus contains no visible GFP by direct visualization, GFP antibody will often detect some cells and myofibers. The frequency of detected cells or fibers is typically decreased relative to perfusion-fixed tissue, but this technique allows GFP-expressing cells or myofibers to be identified and further analyzed with fixation-sensitive techniques, such as immunohistochemistry with certain antibodies. Furthermore, a compromise can be made if the mouse is perfused with dilute fixative (0.005%–0.5% PF) and then GFP is detected with GFP antibody. Skeletal muscle fixed by perfusion with 0.05% PF and then washed in 20% sucrose is often very amenable to antibody staining that is difficult to achieve in heavily fixed tissue.

Thus, GFP is a robust tracking marker that is easily detected, is genetically transmitted, and clearly delineates the morphology of the cell in which it is expressed. Its primary limitation is its rapid diffusion rate out of unfixed cells, a feature requiring special consideration and methodology. In addition, even though GFP has no endogenous counterpart, concerns have been expressed that the excitation and emission wavelengths of GFP are similar to the autofluorescent emissions of a few endogenous molecules such as the flavins and NAD(P)H [36]. Because the levels of these molecules can vary substantially among cell types, this could result in the false identification of GFP within certain cell types. However, it seems likely that this debate is based largely on misunderstandings and misapplications of imaging technologies, and therefore a brief discussion of imaging techniques is in order.

Primer on Imaging and Autofluorescence

The wavelengths of light typically used in routine imaging range from approximately 350 nm (blue) to more than 750 nm (red). Each fluorophore has one or more specific wavelength ranges at which it absorbs or emits light energy (Fig. 1H). Although the overall intensity of emission varies with the excitation wavelength, the spectral distribution of emitted light is largely independent of the excitation wavelength [54]. Thus, fluorophores can be distinguished from each other based on their distinctive excitation and emission wavelengths.

Light can be described by three criteria: hue, intensity, and saturation. Any of these criteria can be used to distinguish emissions from different fluorophores. Hue refers to the color or, more specifically, the wavelength of the light. Intensity refers to the quantity of light. Saturation describes how pure the color of the light is (e.g., how red a red light is). For example, an emission consisting of a single wavelength is extremely saturated whereas an emission made up of a broad range of wavelengths has low saturation. Taken to the extreme, something with zero saturation is gray with the shade of gray described by the intensity, and if this intensity is high enough, the light appears white. In typical applications, most fluorophores emit over a narrow range of wavelengths and thus yield a signal that appears highly saturated.

In contrast to fluorophores, which have well-defined bands of excitation and emission, general autofluorescence usually has the potential to occur across the full spectrum of hues in visible light. In other words, general autofluorescence is not specific to any particular wavelength of light (Fig. 2A). The actual range of emitted wavelengths for general autofluorescence is determined by the specific excitation wavelength, in contrast to fluorophores, for which the emission profiles are mostly independent of the excitation wavelength. The maximum autofluorescent peak typically occurs at a wavelength 25–80 nm greater than the excitation wavelength (Fig. 2A). Even if a single wavelength from a laser is used as the energy source, the resulting emission profile of autofluorescence is typically much broader than the emission profiles of most fluorophores (Fig. 2A). Because autofluorescence occurs across a wide range of wavelengths, it has a low saturation. Thus, in general, fluorophores are distinguished from each other by the hue and from autofluorescence by the intensity and saturation.

There are a few molecules encountered in normal tissue sections that have the potential to generate autofluorescence with narrow emission profiles. Billinton and Knight [36] have written an excellent review on this topic. The distinctive hue and saturation of these autofluorescent molecules can resemble those of specific fluorophores, but the intensity of emissions from these molecules is considerably lower than that of typical fluorophores. Molecules with narrower emission profiles that could theoretically be confused with GFP include flavins (absorb 450–490 nm, emit 500–560 nm) and NADH and NADPH (absorb 360–370 nm, emit 440–470 nm) [36]. In general, the intensity of flavins and NAD(P)H is quite low relative to the wide-spectrum, fixation-induced autofluorescence seen in most tissues. However, various experimental conditions such as fixation, pH, and oxidative state can change the autofluorescent properties of these molecules. As a result, it is always essential that negative controls be carefully evaluated.

Filters

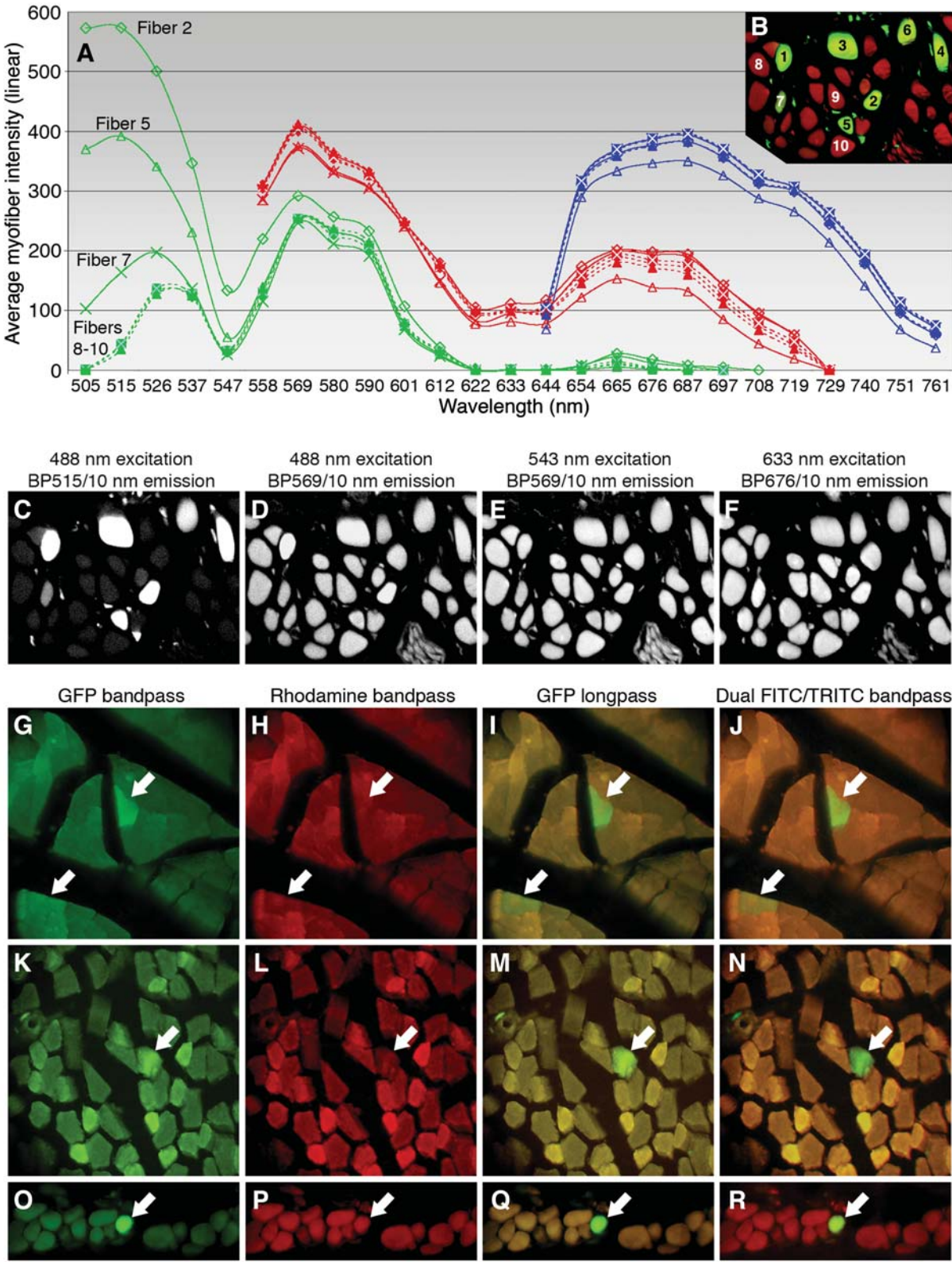
Two primary kinds of optical filters, each using a distinctive strategy, are used to view the emissions of fluorophores selectively. Bandpass (BP) filters allow only a selected range of wavelengths to pass. They are described by their center wavelength and bandwidth. For example, a BP filter commonly used for GFP emission is a BP540/40 that allows only wavelengths of 520–560 nm to pass. In contrast, longpass (LP) and shortpass (SP) filters selectively allow only wavelengths to pass through that are, respectively, longer than or shorter than a specified wavelength (Fig. 1H).

BP filters attempt to distinguish fluorophore signals from autofluorescence based solely on intensity differences. By selectively allowing only the peak emission wavelengths from a given fluorophore to pass, they increase the intensity of the fluorophore signal relative to the background. On the other hand, because they allow only a selected range of wavelengths to pass, most of the

Figure 2. Two methods of distinguishing green fluorescent protein (GFP)–expressing myofibers from background autofluorescence are (A–F) spectral analysis and (G–R) hue-based discrimination. (A–F): Myofibers in the panniculus carnosus muscle were imaged with a Zeiss 510 Meta, three-laser, confocal microscope, which has a dispersive grating that separates the fluorescence spectrum into 32 10.5-nm-wide channels. (B): The image in the upper right shows the seven GFP-positive fibers (green, 1–7) and three GFP-negative fibers (red, 8–10) that were analyzed. The GFP fibers are ordered by GFP intensity, with 1 having the brightest GFP signal and 7 having the lowest. (A): The graph shows the relative intensities (y-axis) for each myofiber for every 10.5-nm-wide wavelength range (x-axis) after excitation with a laser generating a single wavelength of light. Each fiber was analyzed with a single Argon laser (green lines; 488 nm), 543-nm HeNe laser (red lines), or 633-nm HeNe laser (blue lines). In other words, each individual myofiber is represented by a single green, blue, and red line. Fibers 1, 3, 4, and 6 were excluded for clarity but had similar curves with the expected relative intensities. The peak intensity of myofiber 1 was greater than 900 in the 515-nm wavelength range after excitation with 488-nm light (data not shown). (C–F): The bottom row displays representative monochrome images resulting from the excitation and emission criteria shown. There is a clear difference in the emission intensities observed for GFP-positive myofibers compared with GFP-negative myofibers at the spectral bands centered around 505–526 nm for all myofibers, and this difference extends out to the 550-nm range for the brightest GFP-positive myofibers. However, at emission wavelength ranges beyond approximately 550 nm, there is no difference in the emission intensity because this is the result of general autofluorescence. Furthermore, the autofluorescent emission potential of all myofibers exists across all visible wavelengths of light and occurs irrespective of the excitation wavelength. GFP-expressing myofibers (arrows) in the (G–N) tibialis anterior and the (O–R) panniculus carnosus can be rapidly distinguished under the epifluorescent microscope using the following distinct filter sets: (G, K, O) GFP bandpass, (H, L, P) rhodamine, (I, M, Q) GFP longpass, or (J, N, R) dual fluorescein isothiocyanate/tetramethylrhodamine isothiocyanate (FITC/TRITC) bandpass. (G–R): Captured in true color using a Canon G5 digital camera. The tissues shown in G–R all have unusually high autofluorescence or weak GFP expression and demonstrate how difficult it can be to identify GFP-expressing myofibers in these conditions using only a GFP bandpass filter. In contrast, both the GFP longpass and dual bandpass filters readily allow the identification of GFP-expressing myofibers (arrows) based on their unique hue and saturation, even in the context of very high nonspecific autofluorescence. The filter sets used for the epifluorescent images (G–R) contained the following excitation (EX), dichroic (D), and emission (EM) filters: GFP bandpass (EX 480/40, D 505LP, EM 535/50), rhodamine (EX 540/25, D 565LP, EM 605/55), GFP longpass (EX 480/40, D 505LP, EM 510LP), and dual FITC/TRITC bandpass (EX 482/20 and 545/30, D 525/55 and 570LP, EM 520/15 and 600/40).

saturation and hue data are lost. In other words, for narrow BP filters, the hue of light originating from both the fluorophore and autofluorescence is indistinguishable because the observed hue is determined by the BP filter. For example, white light viewed through a BP filter for GFP appears green.

LP or SP filters take the opposite approach, relying primarily on hue and saturation data to distinguish fluorophores from auto-



fluorescence. Because materials emit at a higher wavelength than they absorb, a phenomenon termed the Stokes shift, LP filters are primarily used for emissions and SP filters are primarily used for excitation. The autofluorescence emissions, although of low intensity for a given wavelength, when summed over a wide range of wavelengths, such as those passed by LP filters, can generate a signal of relatively high intensity. This can be advantageous because it allows an investigator to clearly see the tissue surrounding a cell of interest. On the other hand, because the use of data from LP filters requires the ability to discriminate signals based on the hue and saturation, this limits their utility as this capability is lacking in many laboratory imaging devices.

Because LP filters require hue and saturation data to discriminate signal from background, they are less effective for direct visualization at higher wavelengths since the end of the visible spectrum effectively turns them into a BP filter. Thus, LP filters work well for blue or green fluorophores but become ineffective for imaging red wavelengths of light.

The use of BP or LP filters is often most effective when carried to the extremes. Just as an ideal BP filter is narrowly focused over the peak emission wavelengths, an ideal LP filter needs to capture a wide enough range of hues so that they can be adequately distinguished. The worst case is the use of wide BP filters of approximately 100 nm in width. They do not capture a sufficient width of the spectrum to easily distinguish hues but they do greatly increase the intensity of background autofluorescence, resulting in an almost monochrome image with high background. However, in some cases, such wide BP filters may be necessary to capture enough light to detect a signal. In this case, special techniques [36] must be used to reduce background signal.

Instrumentation

Image-capture devices can be categorized into instruments that collect either monochrome or color images. Most electronic instruments used for fluorescent imaging such as laser-scanning confocal microscopes and low-light digital cameras capture only a monochromatic image. Although such devices do generate a color image, it is actually a falsely colored image. Consequently, the hue and saturation are arbitrarily determined and constant throughout the image. The only wavelength or color information available comes from knowledge of the filter sets used to select the wavelengths of light reaching the CCD or PMT sensor.

Because only intensity values are recorded, the strategy used for monochrome image-capture devices is to use BP filters to select a wavelength range that maximizes the fluorophore signal relative to the combination of background autofluorescence and detector noise. However, because only monochrome intensity values are collected, negative controls lacking the fluorophore signal are required to ensure that the monochrome image is due to the desired signal instead of an aberrant background artifact. Because narrow BP filters discard hue and saturation data, the cautions

and controls required for any type of monochrome imaging apply. This key point is ignored by many investigators. To repeat, the use of narrow BP filters is a form of monochrome imaging. This is true even when viewed with the best hue discriminator available, the human eye.

The most advanced approach of using BP filters with monochrome sensors involves spectral analysis. Spectral analysis uses a dispersive grating to separate an emission signal into discrete wavelength ranges (typically ~10 nm wide). Each spectral component is then individually quantified by a detector and the monochrome intensities are plotted for each spectral component (Fig. 2A). The generated curves can be compared directly, as was done in Figure 2A, or spectral signatures can be calculated for each fluorophore and the fluorescence emissions can be separated computationally.

In contrast to monochrome imaging, traditional color film and some newer digital cameras capture accurate hue and saturation as well as intensity data from fluorescent samples. This distinction is important to consider when documenting fluorescence signals. The example that follows illustrates how an underappreciation of this distinction has resulted in confusion regarding the detection of GFP.

A Subset of Highly Autofluorescent Skeletal Myofibers Can Mistakenly Be Interpreted to Express GFP

Over the past several years, numerous investigators have reported that BMDCs are able to contribute to a wide variety of nonhematopoietic tissues *in vivo*. Many other investigators have failed to observe such events, and, thus, these findings have been the subject of considerable debate.

In addition to several reports documenting the contribution of BMDCs to the CNS, we recently reported that after a bone marrow transplant, GFP-expressing BMDC incorporates into skeletal myofibers throughout the body [10]. We find that most skeletal muscles have a low frequency of GFP-expressing skeletal myofibers (0.0022%–0.26%), in agreement with previous reports. However, in one muscle, the panniculus carnosus, up to 5% of skeletal myofibers expressed GFP without any perturbation to the muscle. The panniculus carnosus is a thin, subcutaneous muscle that surrounds the trunk of hairy mammals and may be involved in thermoregulation. Thus, we observed, remarkably, that the frequency with which BMDC incorporated into skeletal muscle differed over 1,000-fold among different muscles, suggesting that there are physiologic differences in the uptake of BMDCs.

Using a similar bone marrow transplant model, our laboratory demonstrated that the BMDC contribution to skeletal myofibers occurs via a satellite cell intermediate [9], a finding recently replicated [55]. In the course of this study we also found that after 6 months of exercise, 3.5% of myofibers in the tibialis anterior expressed GFP, a 20-fold increase compared with non-exercised controls.

The recent article by Jackson et al. [56] that highlights the risks of interpreting the autofluorescent properties of certain types of skeletal muscle fibers as indicative of the presence of GFP is worthy of discussion. Skeletal myofibers contain both flavins and NADH, and thus concerns that the emissions from these molecules could be mistaken for GFP emission are theoretically sound. Furthermore, compared with glycolytic myofibers, oxidative myofibers have higher levels of flavins and NAD(P)H and, after fixation in PF, demonstrate greater autofluorescent emissions compared with glycolytic myofibers (Fig. 3). These characteristics result in considerable heterogeneity in the autofluorescent intensities of skeletal myofibers in some muscle groups (Figs. 2, 3). However, in fixed tissue, this autofluorescence signal is broad, covering the entire visible spectrum, and is more than two to three orders of magnitude brighter than the flavin-based, green hue-specific autofluorescence that can mimic GFP (Fig. 2A). Furthermore, the pattern of highly autofluorescent fibers is distinct from that of GFP-expressing myofibers in muscle with mixed fiber types (Fig. 3), which makes it unlikely that the highly autofluorescent myofibers were mistaken for GFP-expressing myofibers in previous studies [9, 10, 57].

However, if fixed skeletal muscle sections are imaged using improper techniques, the subset of oxidative myofibers can appear to be GFP-positive. This risk is particularly significant when narrow BP filters are used to evaluate fluorescent signals because, as discussed above, this form of monochromatic imaging discards critical spectral data. This common mistake was illustrated by Jackson et al. [56] when they reported that “normal GFP-negative mice . . . displayed a distinct subset of muscle fibers that were similarly bright green, with the clear appearance of GFP expression.” In fact, this autofluorescence is only green when viewed through a green BP filter (Figs. 2G, 2K, 2O). For example, if Jackson et al. [56] had viewed the same tissue through a red or blue filter, they would have discovered red or blue autofluorescence, respectively (Figs. 2H, 2L, 2P). In fact, the autofluorescent emissions of the fixed skeletal myofibers in question, like most autofluorescence, occur over the full range of hues in the visible light spectrum (Fig. 2A).

Jackson et al. [56] conclude that BP filters that pass emission wavelengths from approximately 510 to 530 nm are the best approach to distinguish GFP from autofluorescence. They apply this observation to both confocal imaging (a monochrome device) and epifluorescence imaging (a color device). However, Jackson et al. [56] failed to make a critical distinction between these two different types of imaging modalities.

Because the autofluorescence from fixed skeletal muscle can be considerable (Figs. 2, 3), narrow BP filters, in our experience, often cannot discriminate conclusively between GFP and autofluorescence. Thus, when epifluorescence can be evaluated in true color, LP filters will more unequivocally distinguish GFP

from background based on its distinctive hue and color saturation. In our own studies of skeletal muscle, we have selected our LP filter (510LP; Chroma set 41012) over several BP filters to which we have access. The images resulting from an LP compared with a BP filter are shown in Figure 2. With the LP filter, the muscle tissue is yellow because it is comprised roughly equally of all visible wavelengths greater than 510 nm.

A major hidden challenge in monochrome imaging techniques is to capture an image of appropriate intensity. Monochrome imaging relies solely on differences in intensity to distinguish signal from background, and, conversely, the intensity of an image is determined by user-controlled settings. However, many experienced scientists routinely collect data with overexposed images because they look brighter. The primary reason for this is that even when an area of the monochrome image is fully exposed (i.e., every pixel is at maximum intensity, which is 255 for an 8-bit image), the application of a false color to it by the instrument results in an image in which the area of maximum intensity is represented by a fully saturated color. In contrast, if traditional film is overexposed, it becomes progressively less saturated until all color is lost and that part of the image turns white. Thus, most investigators associate a bright saturated color with correct exposure. In reality, all monochrome imaging devices should aim to collect data that average middle gray in tone (i.e., of moderate intensity) and that may look somewhat grayish and dull when combined with a false color. In general, it is much easier to determine the correct exposure of grayscale images, so many experienced users do not apply the false color until after images have been captured.

Overexposed images are the result of a too-lengthy detection period, too much excitation energy, or having the gain or amplifier on the detector turned up too high. Improperly adjusted electronic imaging devices can obtain bright images from any normal nonfluorescent tissue, and that image will display complex variances in intensity that to the untrained eye resemble true fluorophore emissions (Fig. 3B; supplemental online Fig. 1). For example, Jackson et al. [56] clearly demonstrated how images of normal skeletal muscle could appear similar to those containing GFP when overexposure is combined with false coloring of monochrome images.

As always, the most important aspect of the detection of any specific signal in a tissue is the use of appropriate negative controls to evaluate potential confounding factors. For example, when documenting GFP-positive myofibers after a transplant with GFP-expressing bone marrow, an analysis of the skeletal myofibers of mice transplanted with wild-type bone marrow is an essential negative control. We have never seen any autofluorescence in our fixed negative controls that had the hue characteristics of GFP. The fixation necessary to hold GFP into skeletal muscle increases the broad-spectrum autofluorescence sufficiently to overwhelm any flavin- or NADH-based fluorescence.

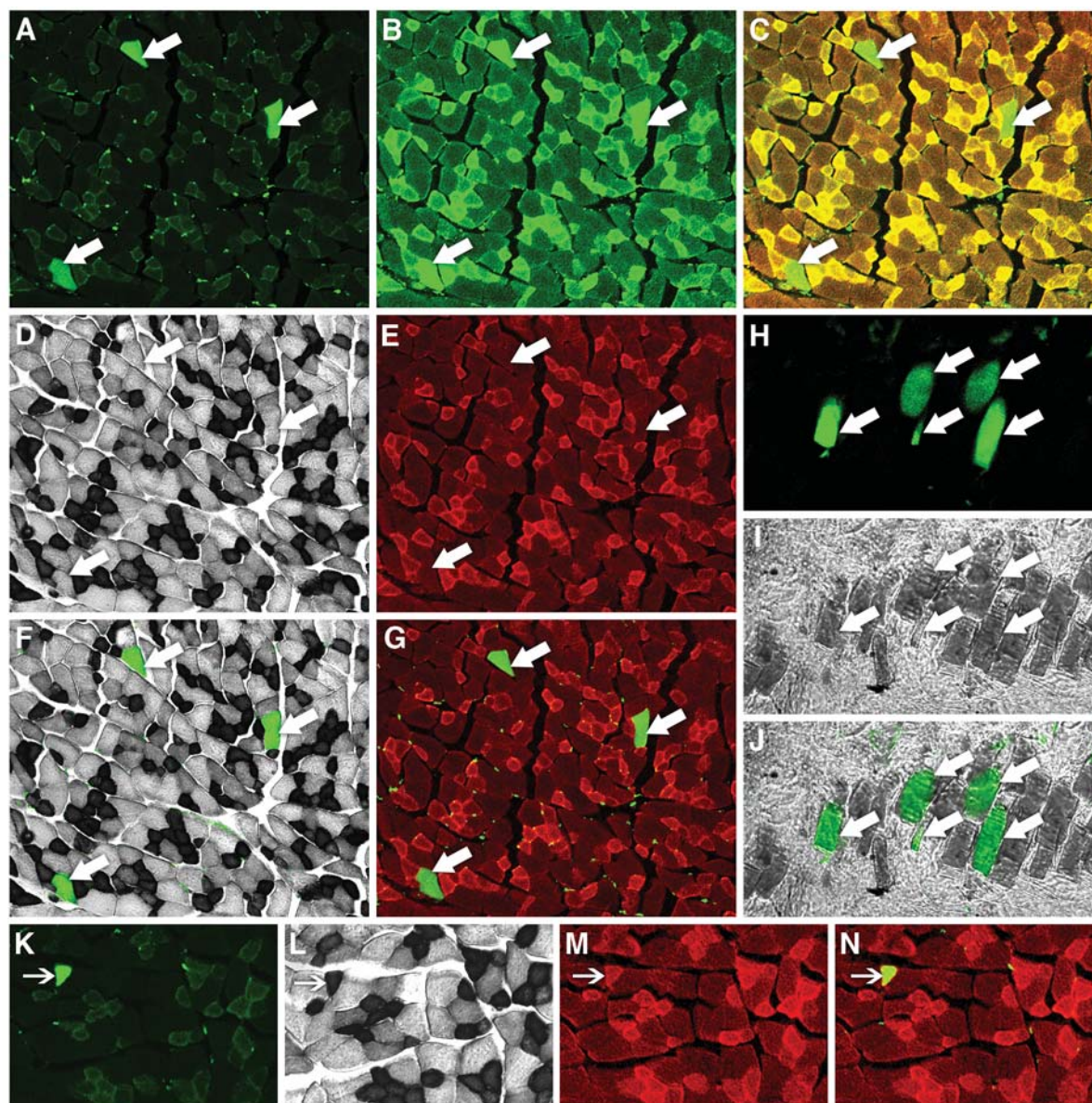


Figure 3. Images of the (A–G, K–N) tibialis anterior and (H–J) panniculus carnosus captured using a laser-scanning confocal microscope with channels optimized for green fluorescent protein (GFP) (green in A–C, F, G, H, J, K, N), rhodamine (red in C, E, G, M, N) or transmitted light (D, F, I, J, L). (C, F, G, J, N): The result of computationally merging single channels. (B): A severely overexposed image of the same field as (A), which is correctly exposed. The tissues shown in (B, E, M) demonstrate the heterogeneity that exists in the autofluorescent intensities among different types of skeletal myofibers. (B): This heterogeneity can confuse analysis when only single bandpass filters are used, particularly when an image is overexposed. (B): Because all confocal and most digital cameras used for fluorescent imaging capture a monochrome image and then apply a false color, overexposure can result in the appearance of GFP-expressing myofibers. (C): However, even if an image is severely overexposed, as in image B, it can be computationally combined with a red colored image of the tissue autofluorescence for the unequivocal identification of the GFP-expressing myofibers. The fibers with the highest autofluorescence are oxidative myofibers expressing higher levels of NADH dehydrogenase activity (darker myofibers in D, F, I, J, L). (K–N): Outside of the panniculus carnosus, most GFP-expressing myofibers are glycolytic fibers (>95%) that lack high NADH expression, although exceptions are seen. Interestingly, in the panniculus carnosus, where a higher proportion of GFP-expressing myofibers is observed, all of the myofibers have homogeneously high levels of both autofluorescence and NADH activity.

Ratiometric Analysis Techniques

Jackson et al. [56] propose that single-laser spectral analysis is the best method to distinguish GFP from autofluorescence. However, spectral imaging requires specialized and expensive equipment, is time consuming, and generates poorer quality images than stan-

dard confocal microscopy. Alternatively, GFP can be readily distinguished from autofluorescence with a simple ratiometric analysis comparing emissions in the GFP range to emissions at a higher wavelength that represents general autofluorescence (Figs. 2G–2R, 3A–3N). This can be achieved by two distinct approaches.

In the first approach, ratiometric analysis can be achieved visually and in real time by using either an LP filter or a dual BP filter on an epifluorescent microscope (Figs. 2G–2R). This is a simple and rapid technique to visually identify GFP-expressing cells based on hue discrimination. The caveat, however, is that documentation of the visual field created by these filters requires a true color-imaging instrument.

In the second approach, ratiometric analysis can be achieved computationally by merging two (or more) monochromatic images, each of which is falsely colored with a single hue representing a primary color (Figs. 3A–3N). If these monochromatic images are combined, then more than two hues can be created by the interactions of different intensity levels for each hue. For example, monochromatic digital images are frequently represented by 8 bits, which allows them to distinguish 256 intensity levels. If two 8-bit monochromatic images are combined and each is a different primary color (red, green, or blue), then more than 65,000 theoretical colors result. This approach can be routinely performed on many monochromatic imaging devices and often clearly identifies the presence of a fluorophore that was not apparent in a single monochromatic image.

For example, we routinely image GFP-expressing myofibers by collecting and merging two channels of monochromatic data on our confocal microscope. The first channel is designed to maximize GFP emission and minimize autofluorescence (Argon laser excitation, 488-nm excitation filter, emissions selected by 570 dichroic, 545 dichroic, and 505- to 530-BP filter). The second channel is one that is typically used to image red fluorochromes but is used here to evaluate the relative autofluorescence of individual skeletal myofibers (543-nm HeNe laser excitation, 543-nm excitation filter, emissions selected by 635 dichroic, 545 dichroic, and 560- to 615-BP filter).

In this example of the second approach, channel 1 is falsely colored green and channel 2 is falsely colored red. When the two images are combined, the autofluorescence in each channel is closely proportional and thus all of the non-GFP-expressing tissue yields a homogeneous hue. However, wherever GFP is present, the green intensity is proportionally higher than the red intensity and a hue shift results. Thus, this method serves as a means to rapidly distinguish the presence of a fluorophore in the presence of substantial background autofluorescence, requires only a stan-

dard confocal, and simultaneously generates high-quality images of the tissue.

Thus, dual-band, ratiometric techniques are based on a comparison of two wavelength bands, one that includes autofluorescence and fluorophore signals and one that includes only autofluorescence. Such techniques are rapid, visually obvious, do not require expensive equipment, and unequivocally distinguish fluorophores from background autofluorescence.

CONCLUSIONS

The field of adult stem cell plasticity has generated numerous conflicting reports and debates in recent years. Many of these debates have an origin in inadequate preparation of tissues, inappropriate choice of a tracking marker, or a lack of appropriate positive or negative controls or occurred as a result of poor methods for imaging experimental tissue. As demonstrated here, one of the areas of debate has centered on the frequency with which GFP-expressing BMDCs contribute to skeletal myofibers. For example, due to the heterogeneous pattern of autofluorescence in some skeletal muscles, particular care is required to distinguish GFP from autofluorescence emissions. A powerful and efficient technique to do this is ratiometric analysis performed either directly with filters or by computationally merging multiple images of distinct colors.

Many advances in imaging technology have occurred in the past decade, and, consequently, image data of improved quality can be collected faster and more easily. However, an understanding of the limitations of these new devices is essential to ensure that the image on the screen or printed in a publication be representative of the biology being documented.

ACKNOWLEDGMENTS

This research was supported by a Life and Health Insurance Medical Research Fund fellowship and NIH predoctoral training grant GM07149 to T.R.B. and NIH grants AG009521, HD018179, AG020961, AG024987, HL065572, a Senior Ellison Medical Foundation award AG-SS-0817, and support from the Baxter Foundation to H.M.B.

DISCLOSURES

The authors indicate no potential conflicts of interest.

REFERENCES

- Pomerantz J, Blau HM. Nuclear reprogramming: a key to stem cell function in regenerative medicine. *Nat Cell Biol* 2004;6:810–816.
- Horwitz EM. Stem cell plasticity: the growing potential of cellular therapy. *Arch Med Res* 2003;34:600–606.
- Prockop DJ. Further proof of the plasticity of adult stem cells and their role in tissue repair. *J Cell Biol* 2003;160:807–809.
- Alison MR, Poulosom R, Otto WR et al. Recipes for adult stem cell plasticity: fusion cuisine or readymade? *J Clin Pathol* 2004;57:113–120.
- Wagers AJ, Weissman IL. Plasticity of adult stem cells. *Cell* 2004;116:639–648.
- Grove JE, Bruscia E, Krause DS. Plasticity of bone marrow-derived stem cells. *STEM CELLS* 2004;22:487–500.
- Ferrari G, Cusella-De Angelis G, Coletta M et al. Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* 1998;279:1528–1530.
- Gussoni E, Soneoka Y, Strickland CD et al. Dystrophin expression in the *mdx* mouse restored by stem cell transplantation. *Nature* 1999;401:390–394.

- 9 LaBarge MA, Blau HM. Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. *Cell* 2002;111:589–601.
- 10 Brazelton TR, Nystrom M, Blau HM. Significant differences among skeletal muscles in the incorporation of bone marrow-derived cells. *Dev Biol* 2003;262:64–74.
- 11 Palermo AT, Labarge MA, Doyonnas R et al. Bone marrow contribution to skeletal muscle: a physiological response to stress. *Dev Biol* 2005;279:336–344.
- 12 Laflamme MA, Myerson D, Saffitz JE et al. Evidence for cardiomyocyte repopulation by extracardiac progenitors in transplanted human hearts. *Circ Res* 2002;90:634–640.
- 13 Bittner RE, Schofer C, Weipoltshammer K et al. Recruitment of bone-marrow-derived cells by skeletal and cardiac muscle in adult dystrophic *mdx* mice. *Anat Embryol (Berl)* 1999;199:391–396.
- 14 Orlic D, Kajstura J, Chimenti S et al. Bone marrow cells regenerate infarcted myocardium. *Nature* 2001;401:701–705.
- 15 Vassilopoulos G, Wang PR, Russell DW. Transplanted bone marrow regenerates liver by cell fusion. *Nature* 2003;422:901–904.
- 16 Wang X, Willenbring H, Akkari Y et al. Cell fusion is the principal source of bone-marrow-derived hepatocytes. *Nature* 2003;422:897–901.
- 17 Petersen BE, Bowen WC, Patrene KD et al. Bone marrow as a potential source of hepatic oval cells. *Science* 1999;284:1168–1170.
- 18 Theise ND, Badve S, Saxena R et al. Derivation of hepatocytes from bone marrow cells in mice after radiation-induced myeloablation. *Hepatology* 2000;31:235–240.
- 19 Lagasse E, Connors H, Al-Dhalimy M et al. Purified hematopoietic stem cells can differentiate into hepatocytes in vivo. *Nat Med* 2000;6:1229–1234.
- 20 Mezey E, Chandross KJ, Harta G et al. Turning blood into brain: cells bearing neuronal antigens generated in vivo from bone marrow. *Science* 2000;290:1779–1782.
- 21 Brazelton TR, Rossi FM, Keshet GI et al. From marrow to brain: expression of neuronal phenotypes in adult mice. *Science* 2000;290:1775–1779.
- 22 Weimann JM, Charlton CA, Brazelton TR et al. Contribution of transplanted bone marrow cells to Purkinje neurons in human adult brains. *Proc Natl Acad Sci U S A* 2003;100:2088–2093.
- 23 Weimann JM, Johansson CB, Trejo A et al. Stable reprogrammed heterokaryons form spontaneously in Purkinje neurons after bone marrow transplant. *Nat Cell Biol* 2003;5:959–966.
- 24 Okabe M, Ikawa M, Kominami K et al. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett* 1997;407:313–319.
- 25 Gussoni E, Wang Y, Fraefel C et al. A method to codetect introduced genes and their products in gene therapy protocols. *Nat Biotechnol* 1996;14:1012–1016.
- 26 Lichter P, Cremer T, Tang CJ et al. Rapid detection of human chromosome 21 aberrations by in situ hybridization. *Proc Natl Acad Sci U S A* 1988;85:9664–9668.
- 27 Jiang Y, Henderson D, Blackstad M et al. Neuroectodermal differentiation from mouse multipotent adult progenitor cells. *Proc Natl Acad Sci U S A* 2003;100(suppl 1):11854–11860.
- 28 Zambrowicz BP, Imamoto A, Fiering S et al. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* 1997;94:3789–3794.
- 29 Kisseberth WC, Brettingen NT, Lohse JK et al. Ubiquitous expression of marker transgenes in mice and rats. *Dev Biol* 1999;214:128–138.
- 30 Cohen RB, Rutenburg SH, Tsou KC et al. The colorimetric estimation and histochemical demonstration of beta-d-galactosidase. *J Biol Chem* 1952;195:239–249.
- 31 Hatton JD, Lin L. Demonstration of specific neuronal cell groups in rat brain by beta-galactosidase enzyme histochemistry. *J Neurosci Methods* 1992;45:147–153.
- 32 Sanchez-Ramos J, Song S, Dailey M et al. The X-gal caution in neural transplantation studies. *Cell Transplant* 2000;9:657–667.
- 33 Krause DS, Theise ND, Collector MI et al. Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* 2001;105:369–377.
- 34 Spotl L, Sarti A, Dierich MP et al. Cell membrane labeling with fluorescent dyes for the demonstration of cytokine-induced fusion between monocytes and tumor cells. *Cytometry* 1995;21:160–169.
- 35 Belkhir A, Lytvyn V, Guilbault C et al. A noninvasive cell-based assay for monitoring proteolytic activity within a specific subcellular compartment. *Anal Biochem* 2002;306:237–246.
- 36 Billinton N, Knight AW. Seeing the wood through the trees: a review of techniques for distinguishing green fluorescent protein from endogenous autofluorescence. *Anal Biochem* 2001;291:175–197.
- 37 Prasher DC. Using GFP to see the light. *Trends Genet* 1995;11:320–323.
- 38 Cubitt AB, Heim R, Adams SR et al. Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* 1995;20:448–455.
- 39 Heim R, Prasher DC, Tsien RY. Wavelength mutations and posttranslational autooxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* 1994;91:12501–12504.
- 40 Lippincott-Schwartz J, Snapp E, Kenworthy A. Studying protein dynamics in living cells. *Nat Rev Mol Cell Biol* 2001;2:444–456.
- 41 Liu HS, Jan MS, Chou CK et al. Is green fluorescent protein toxic to the living cells? *Biochem Biophys Res Commun* 1999;260:712–717.
- 42 Hanazono Y, Yu JM, Dunbar CE et al. Green fluorescent protein retroviral vectors: low titer and high recombination frequency suggest a selective disadvantage. *Hum Gene Ther* 1997;8:1313–1319.
- 43 Hu W, Cheng CL. Expression of Aequorea green fluorescent protein in plant cells. *FEBS Lett* 1995;369:331–334.
- 44 Sheen J, Hwang S, Niwa Y et al. Green-fluorescent protein as a new vital marker in plant cells. *Plant J* 1995;8:777–784.
- 45 Wang S, Hazelrigg T. Implications for bcd mRNA localization from spatial distribution of exu protein in *Drosophila* oogenesis. *Nature* 1994;369:400–403.
- 46 Chalfie M, Tu Y, Euskirchen G et al. Green fluorescent protein as a marker for gene expression. *Science* 1994;263:802–805.
- 47 Amsterdam A, Lin S, Hopkins N. The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos. *Dev Biol* 1995;171:123–129.
- 48 Peters KG, Rao PS, Bell BS et al. Green fluorescent fusion proteins: powerful tools for monitoring protein expression in live zebrafish embryos. *Dev Biol* 1995;171:252–257.
- 49 Madoff SH. The wonders of genetics breed a new art. *New York Times*. May 26, 2002; Arts & Leisure section: 1, 30.
- 50 Chan AW, Chong KY, Martinovich C et al. Transgenic monkeys produced by retroviral gene transfer into mature oocytes. *Science* 2001;291:309–312.

- 51 Hadjantonakis AK, Gertsenstein M, Ikawa M et al. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech Dev* 1998;76:79–90.
- 52 Hadjantonakis AK, Macmaster S, Nagy A. Embryonic stem cells and mice expressing different GFP variants for multiple non-invasive reporter usage within a single animal. *BMC Biotechnol* 2002;2:11.
- 53 Ikawa M, Yamada S, Nakanishi T et al. 'Green mice' and their potential usage in biological research. *FEBS Lett* 1998;430:83–87.
- 54 Lakowicz J. *Principles of Fluorescence Spectroscopy*. New York: Plenum Press, 1983.
- 55 Sherwood RI, Christensen JL, Conboy IM et al. Isolation of adult mouse myogenic progenitors: functional heterogeneity of cells within and engrafting skeletal muscle. *Cell* 2004;119:543–554.
- 56 Jackson KA, Snyder DS, Goodell MA. Skeletal muscle fiber-specific green autofluorescence: potential for stem cell engraftment artifacts. *STEM CELLS* 2004;22:180–187.
- 57 Corbel SY, Lee A, Yi L et al. Contribution of hematopoietic stem cells to skeletal muscle. *Nat Med* 2003;9:1528–1532.



See www.StemCells.com for supplemental material available online.