A Conditional System to Specifically Link Disruption of Protein-Coding Function with Reporter Expression in Mice

Shin-Heng Chiou,[1] Caroline Kim-Kiselak,[4] Viviana I. Risca,[1] Megan K. Heimann,[4] Chen-Hua Chuang,[1] Aurora A. Burds,[4] William J. Greenleaf,[1] Tyler E. Jacks,[4,5] David M. Feldser,[6,7] and Monte M. Winslow[1,2,3,*]

[1]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305-5120, USA
[2]Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305-5324, USA
[3]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA 94305-5456, USA
[4]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[5]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[6]Department of Cancer Biology, University of Pennsylvania, Philadelphia, PA 19104-6160, USA
[7]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, PA 19104-6160, USA
*Correspondence: mwinslow@stanford.edu
http://dx.doi.org/10.1016/j.celrep.2014.05.031

SUMMARY

Conditional gene deletion in mice has contributed immensely to our understanding of many biological and biomedical processes. Despite an increasing awareness of nonprotein-coding functional elements within protein-coding transcripts, current gene-targeting approaches typically involve simultaneous ablation of noncoding elements within targeted protein-coding genes. The potential for protein-coding genes to have additional noncoding functions necessitates the development of novel genetic tools capable of precisely interrogating individual functional elements. We present a strategy that couples Cre/*lox*P-mediated conditional gene disruption with faithful GFP reporter expression in mice in which Cre-mediated stable inversion of a splice acceptor-GFP-splice donor cassette concurrently disrupts protein production and creates a GFP fusion product. Importantly, cassette inversion maintains physiologic transcript structure, thereby ensuring proper microRNA-mediated regulation of the GFP reporter, as well as maintaining expression of nonprotein-coding elements. To test this potentially generalizable strategy, we generated and analyzed mice with this conditional knockin reporter targeted to the *Hmga2* locus.

INTRODUCTION

In recent years, methods allowing manipulation of genes of interest within living organisms have enabled a detailed understanding of numerous genes' function across diverse phyla. The ability to alter the mouse genome has been especially critical for the investigation of gene function in vivo and has contributed immensely to our understanding of many fundamental questions in developmental biology and biomedical sciences. In particular, the development of site-specific recombinase systems in mice, including the Cre/*lox*P and FLP/FRT systems, has allowed the inactivation of genes of interest with both spatial and temporal control. As a result, conditional gene disruption has become a critical tool for understanding gene function during development, homeostasis, as well as in disease states (Rajewsky, 2007; Schmidt-Supprian and Rajewsky, 2007). Fluorescent reporter alleles have proved to be another key resource in the dissection of gene function by allowing direct visualization and isolation of molecularly defined subsets of cells (Hadjantonakis et al., 2003).

The past decade has brought an increased awareness that protein-coding genes can also possess additional nonprotein-coding functions. Single transcriptional units with multiple functions suggest interesting mechanisms that influence the cellular state, but their very existence necessitates new genetic methods to specifically abrogate the protein-coding element without altering other functions of the transcript and to generate reporter alleles that remain subject to all endogenous forms of regulation. In developing an allele system to couple specific disruption of protein-coding function with reporter expression, we considered the potential noncoding functions of transcripts.

Approximately half of all microRNAs (miRNAs) are encoded within introns. Targeted deletion of several intronic miRNAs has resulted in severe phenotypes including embryonic lethality (Kuhnert et al., 2008; Miyaki et al., 2010; Nakamura et al., 2011; Wang et al., 2008; Zhao et al., 2007), whereas many other intronic miRNAs remain uncharacterized. Therefore, gene targeting that inadvertently disrupts the expression of intronic miRNAs could have unanticipated phenotypic consequences (Kuhnert et al., 2008; Osokine et al., 2008; Wang et al., 2008). Interestingly, rare exonic miRNAs have also been shown to have critical functions (Sundaram et al., 2013).

Further indicating the need for improved gene inactivation methods, recent work indicates that mRNAs may also directly influence miRNA function through competitive binding, suggesting that the relative abundance of one miRNA target may impact that miRNA's ability to repress other targets. Genetic deletion or truncation of transcripts that have this competitive endogenous RNA (ceRNA) activity could thus lead to inadvertent widespread alterations in the protein expression of conserved miRNA targets (Karreth et al., 2011; Kumar et al., 2014; Poliseno et al., 2010; Salmena et al., 2011; Tay et al., 2011). More generally, any titratable site-specific RNA-binding factor could be affected by the deletion of a highly expressed target transcript. Finally, recent reports of abundant circular RNA species comprised of exons of protein-coding genes with potential additional activity forewarn of further inadvertent changes in cellular state when transcripts are entirely ablated or truncated (Memczak et al., 2013; Salzman et al., 2012).

Additionally, the appropriate regulation of protein expression from reporter alleles can also be influenced by transcript structure. miRNAs, RNA-binding proteins, and RNA secondary structure regulate transcript stability and protein expression. Thus, reporters that are embedded within the native transcript will more faithfully recapitulate the expression of the gene of interest. Indeed, reporter alleles with endogenous versus exogenous 3′ UTRs can produce dramatically different expression patterns due to the presence or absence of appropriate miRNA-mediated regulation (Merritt et al., 2008; Yoo et al., 2009).

Collectively, the potential for transcripts to possess multiple functions indicates that more precise genome modification methods are required to avoid the potential phenotypic consequences of altering nonprotein-coding functions of genes. Based on these considerations, we developed a murine allele system in which conditional disruption of protein-coding function is directly coupled to fluorescent marker expression without impacting other aspects of transcript function.

RESULTS

A Strategy to Couple Conditional Gene Disruption with GFP Reporter Expression in Mice

The alteration of gene function with temporal and spatial control as well as the use of fluorescent reporter alleles has allowed the detailed investigation of a plethora of developmental processes. Coupling conditional Cre/loxP-mediated gene disruption with fluorescent reporter expression would not only generate two useful tools from one gene-targeting event but would also allow single-cell resolution of gene disruption, permit altered cells to be assessed relative to nontargeted cells, and enable the identification and isolation of cells that are attempting to express the disrupted gene.

Several genetic approaches in mice have been previously developed to inactivate a gene of interest and express a reporter gene, including the combination of Cre reporter alleles with floxed alleles (Alexander et al., 2009; Lao et al., 2012; Srinivas et al., 2001), mosaic analysis with double markers (Tasic et al., 2012; Zong et al., 2005), systems based on the inversion of gene-trap-like elements (Mandalos et al., 2012; Schnütgen et al., 2003, 2005; Schnütgen and Ghyselinck, 2007; Xin et al.,

2005), and floxing of the entire gene followed by a 3′ reporter (Moon et al., 2000; Potocnik et al., 2000; Theis et al., 2001; Wellershaus et al., 2008). Each of these methods has its unique strengths, but system-specific limitations have in all cases negated their widespread application (Table S1).

We devised a strategy based on the insertion of an inverted splice acceptor-eGFP-splice donor (SA-GFP-SD) cassette into an intron of the target gene. Pairs of heterologous loxP sites (2272 loxP and 5171 loxP) (Lee and Saito, 1998; Stern et al., 2008) flank the SA-GFP-SD cassette, thus allowing two sequential Cre-mediated recombination events to stably invert the SA-GFP-SD cassette. The inversion should disrupt gene function and concomitantly generate a GFP reporter (Figures 1A and S1). The first recombination event inverts the intervening sequence between either the 5171 loxP or the 2272 loxP sites and generates one of two intermediates in which alternate loxP variants are now in a head-to-tail orientation. The second recombination event then deletes the sequence between these sites leaving a final stably inverted cassette with one 2272 loxP site and one incompatible 5171 loxP site (Figures 1A and S1). Most importantly, we engineered a splice donor site following the stop codon of GFP to allow the transcript to continue through the remaining exons and 3′ UTR to the endogenous polyadenylation sites (Figures 1A and 1B). Because the reading frame at the end of the first exon of a gene of interest can be in any of the three reading frames, we generated base vectors with GFP in each reading frame to generate this style of conditional knockin (CK) allele (Figure S1).

Generation of an $Hmga2^{CK}$ Allele

To test our strategy to link Cre-mediated gene disruption with GFP reporter expression, we generated a CK allele of the chromatin-associated protein Hmga2 ($Hmga2^{CK}$). We chose this gene because of our interest in Hmga2 in cancer metastasis as well as its dramatic role in body size regulation and potential role in stem cell biology (Nishino et al., 2008; Winslow et al., 2011; Xiang et al., 1990; Yu et al., 2007; Zhou et al., 1995). The Hmga2 genomic locus is quite large (~116 kb), has all exons in the same frame, has an intronic miRNA (miR-763), and produces a transcript that is highly regulated by endogenous miRNAs. Despite the difficulty of accounting for these common elements with existing genetic-targeting approaches, these characteristics made Hmga2 an excellent candidate for our CK allele system. The Hmga2 locus was targeted with the SA-GFP-SD cassette in the reverse orientation (Figure 1A). Mice generated from correctly targeted embryonic stem (ES) cells were bred with mice expressing FLPe recombinase to delete the neomycin-resistance cassette, thereby generating the $Hmga2^{CK}$ allele (Farley et al., 2000) (Figure 1A). To characterize the Hmga2 allele after cassette inversion, $Hmga2^{CK}$ mice were then bred with mice expressing Cre recombinase to generate a germline $Hmga2^{GFP}$ allele (Wellershaus et al., 2008) (Figure 1A).

Cre-Mediated Stable Inversion of the SA-GFP-SD Cassette Disrupts Protein Expression

The $Hmga2^{GFP}$ allele generated a transcript in which exon 2 was replaced by GFP, resulting in a protein product containing the N terminus of Hmga2 fused to GFP (Figures 1B–1D and S1I). Although the SA within our cassette is well characterized
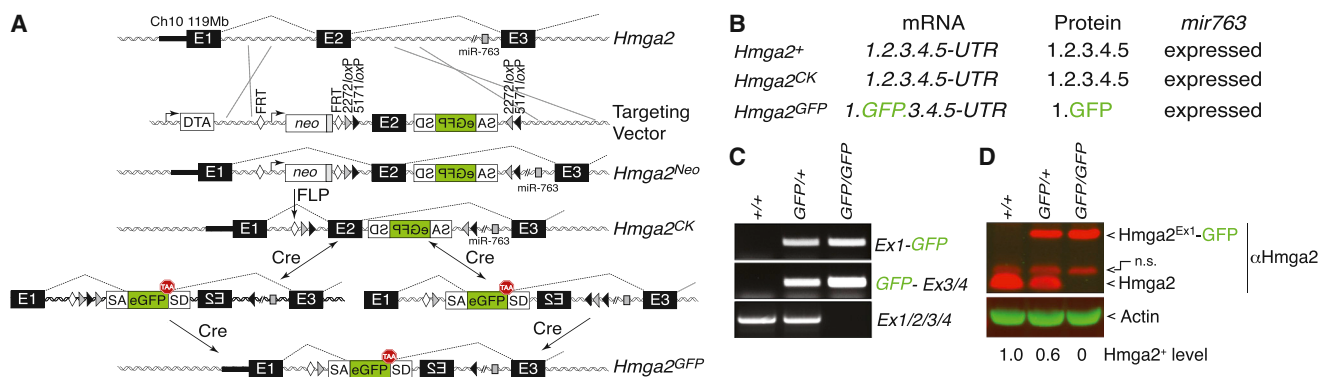
Figure 1. A Strategy to Integrate Conditional Gene Disruption with GFP Reporter Expression within an Otherwise Full-Length Transcript

(A) Schematic of the *Hmga2* targeting to create a Cre-regulated gene-trap allele that is both a conditional null allele and a conditional GFP reporter.

(B) Anticipated expression of the Hmga2 mRNA, protein, and intronic miRNA (miR-763) from the wild-type (*Hmga2⁺*), targeted (*Hmga2^CK^*), and inverted (*Hmga2^GFP^*) alleles.

(C) RT-PCR analysis of MEFs of the indicated *Hmga2* genotypes. The SA-GFP-splice donor in the *Hmga2^GFP^* allele generates the expected transcript containing exon 1 to GFP (*Ex1-GFP*) and GFP to exons 3 and 4 (*GFP-Ex3/Ex4*) while completely disrupting the full-length *Hmga2* transcript (*Ex1/Ex2/Ex3/Ex4*).

(D) Western blot analysis of MEFs shows that the *Hmga2^GFP^* allele generates an Hmga2^Ex1^-GFP fusion that is detected by the Hmga2 antibody, which recognizes an epitope in the N-terminal portion of Hmga2. No endogenous Hmga2 is detected in *Hmga2^GFP/GFP^* MEFs. Actin shows equal loading. Wild-type Hmga2⁺ protein quantification is shown. A nonspecific band (n.s.) is indicated. One of four independent experiments is shown.

and is capable of generating null-like alleles (Carette et al., 2009), it remained possible that some *Hmga2* transcripts from the *Hmga2^GFP^* allele could splice from exon 1 to exon 3. In *Hmga2^GFP/GFP^* embryos and murine embryonic fibroblasts (MEFs), no evidence of an mRNA generated from exon 1 splicing to exon 3 was detected by RT-PCR (data not shown). Additionally, whereas splicing from exon 1 to exon 3 would create an in-frame protein product, even very long exposure times gave no indication of any truncated Hmga2 protein product in *Hmga2^GFP/GFP^* embryos or MEFs (Figures 1D, S2A, S2B, and S3C; data not shown).

Cre-Mediated Stable Inversion of the SA-GFP-SD Cassette Generates a GFP Fusion Protein

In addition to functionally inactivating a targeted gene, a major power of our system is its ability to simultaneously generate a conditional GFP reporter allele. To assess GFP expression, we performed a series of experiments on embryos and MEFs from mice carrying the *Hmga2^CK^* or *Hmga2^GFP^* alleles. First, we examined embryos from several developmental stages and were able to easily distinguish *Hmga2^+/+^* from *Hmga2^GFP/+^* and *Hmga2^GFP/GFP^* mice by virtue of their GFP fluorescence (Figures 2A and S3A). Quantitative detection of the GFP signal confirmed that *Hmga2^GFP/GFP^* embryos are approximately twice as bright as *Hmga2^GFP/+^* embryos (Figure S3A). Analysis of MEFs by flow cytometry also confirmed robust GFP expression from the *Hmga2^GFP^* allele (Figure 2B). Immunohistochemical staining of embryonic day 14.5 (E14.5) *Hmga2^+/+^* embryos indicated that whereas a majority of cells within the embryos express Hmga2, only a fraction of the cells within the fetal liver express high levels of Hmga2 (Figure 2C). Therefore, we used flow cytometry to determine whether the *Hmga2^GFP^* allele could identify the Hmga2-expressing cell population within the fetal liver. We found that a subpopulation of fetal liver that was negative for hematopoietic lineage markers produced the highest level of Hmga2-GFP, with

most cells producing low levels of GFP and *Hmga2^GFP/GFP^* cells producing higher levels of GFP than *Hmga2^GFP/+^* cells (Figures 2D and 2E). Neither *Hmga2^CK/+^* nor *Hmga2^CK/CK^* embryos produced GFP as assessed by direct visualization under the fluorescence-dissecting scope, flow cytometry on fetal liver cells and MEFs, or western blot analysis of embryos and MEFs (Figures 2F, S2G, S2H, S3A, and S3C; data not shown). To verify the conditional nature of the allele, we infected *Hmga2^CK/+^* and *Hmga2^CK/CK^* MEFs with an adenoviral vector expressing Cre. Cre expression converted the CK alleles to their GFP conformation, with infected cells expressing high levels of GFP and lacking full-length Hmga2 protein (Figures 2F and S3C).

Cassette Inversion Maintains the Physiologic Expression of Intronic miRNAs

Approximately half of known miRNAs are encoded within the introns of protein-coding genes. Although some gene inactivation strategies disrupt transcript expression, thereby disrupting both the protein-coding mRNA and the intronic miRNA, our system maintains the production of the full-length transcript, suggesting that intronic miRNAs should not be affected (Figures 1A and 1B). miR-763 resides within intron 3 of the *Hmga2* locus; therefore, we quantified miR-763 expression in MEFs and embryos carrying the *Hmga2^CK^* and *Hmga2^GFP^* alleles. Importantly, neither the *Hmga2^CK^* nor the *Hmga2^GFP^* allele affected miR-763 expression, confirming that our gene-targeting strategy can specifically inactivate the protein-coding element without impacting other functional components within downstream introns (Figures S2I–S2L).

Cassette Inversion Maintains the Full-Length Transcript Allowing Physiologic miRNA-Mediated Regulation of the GFP Reporter

One theoretical advantage of generating a conditional allele using our general strategy is the ability to embed the GFP reporter
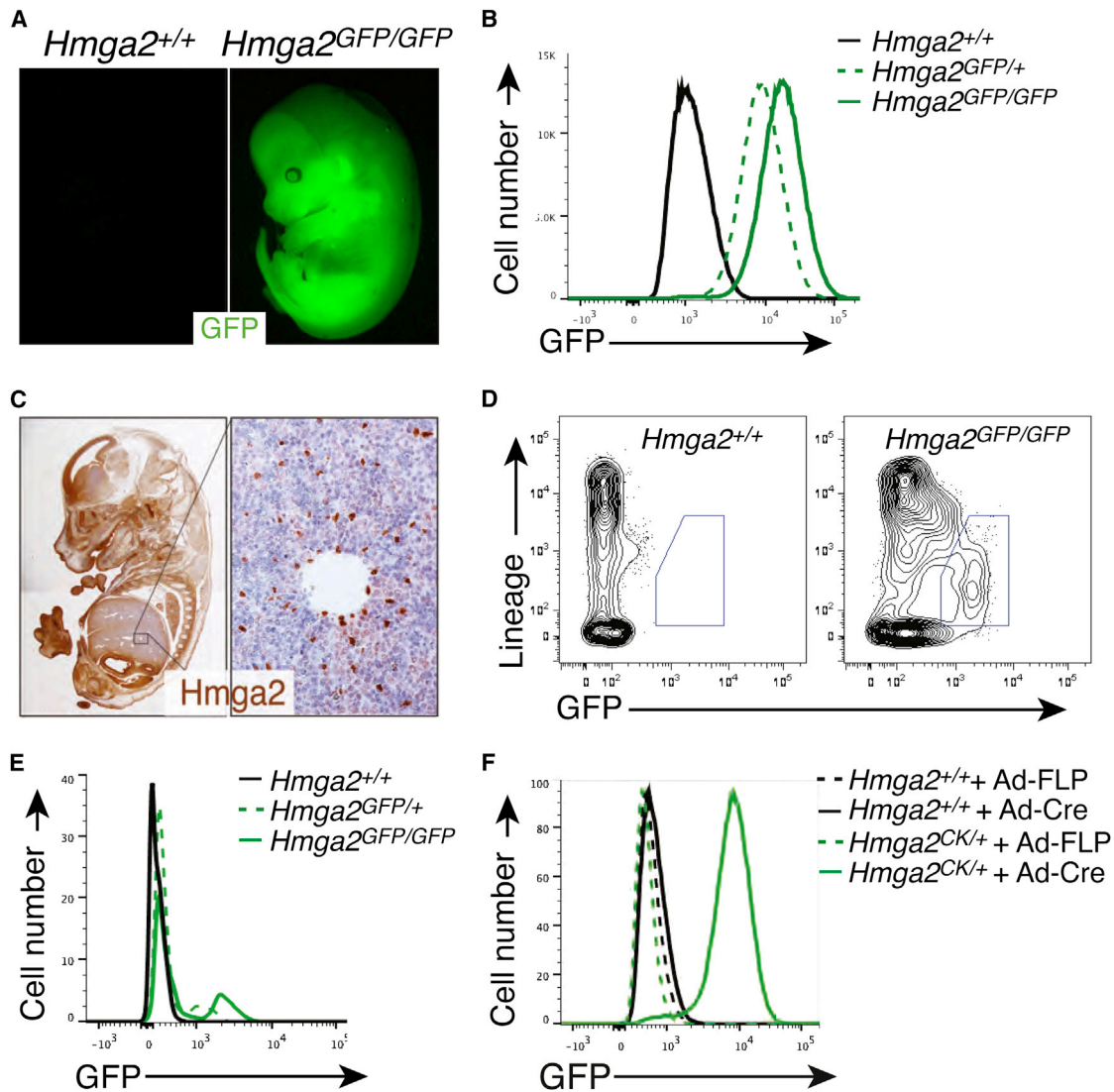
Figure 2. Gene-Trapped *Hmga2* Functions as a Conditional Reporter of Protein Expression

(A) Fluorescent images of *Hmga2*[+/+] and *Hmga2*[GFP/GFP] E14.5 embryos.

(B) Flow cytometric analysis of MEFs documents robust GFP expression and correspondingly higher expression in cells from *Hmga2*[GFP/GFP] mice.

(C) Hmga2 is widely expressed in embryos, but only a subset of fetal liver cells expresses high levels of Hmga2. Immunohistochemical staining of an *Hmga2*[+/+] embryo is shown. Enlarged area is indicated. Hematoxylin was used as counterstain.

(D) Flow cytometry analysis of fetal liver cells identifies a population of Hmga2[hi] Lineage[neg] cells. Forward-scatter/side-scatter-gated cells are shown.

(E) Overlay of the GFP intensity of fetal liver cells of the indicated genotypes. This histogram illustrates that most cells express a low level of Hmga2 and that GFP intensity increases in *Hmga2*[GFP/GFP] fetal liver cells.

(F) Conditional expression of the GFP reporter in MEFs after infection of *Hmga2*[CK/+] cells with Adenoviral-Cre (Ad-Cre). *Hmga2*[+/+] MEFs and Adenoviral-FLP (Ad-FLP) infection are negative controls.

within an otherwise full-length transcript, thereby ensuring that the GFP reporter remains under all the transcriptional and post-transcriptional regulatory control elements of the endogenous locus. In particular, the inclusion of a splice donor site after the stop codon of GFP allows transcription to continue all the way through the 3′ UTR, suggesting that the conditional GFP reporter should remain under physiologic miRNA regulation (Figures 1A, 1B, S1H, S3B, and S3D). To directly assess whether the *Hmga2*[GFP] allele maintains a physiologic and full-length

transcript structure, we performed northern blot analyses using *Hmga2*- and *GFP*-specific probes. The composite *Hmga2*[GFP] transcript was detected as a single predominant band that was similar in size to the *Hmga2*[wt] counterpart, indicating that cassette inversion embeds GFP within the full-length transcript (Figures S3E and S3F).

Because *Hmga2* is a well-characterized *let-7* miRNA target (Lee and Dutta, 2007; Mayr et al., 2007) and is regulated by *let-7* in MEFs (Viswanathan et al., 2008; Figures 3C and S4), we
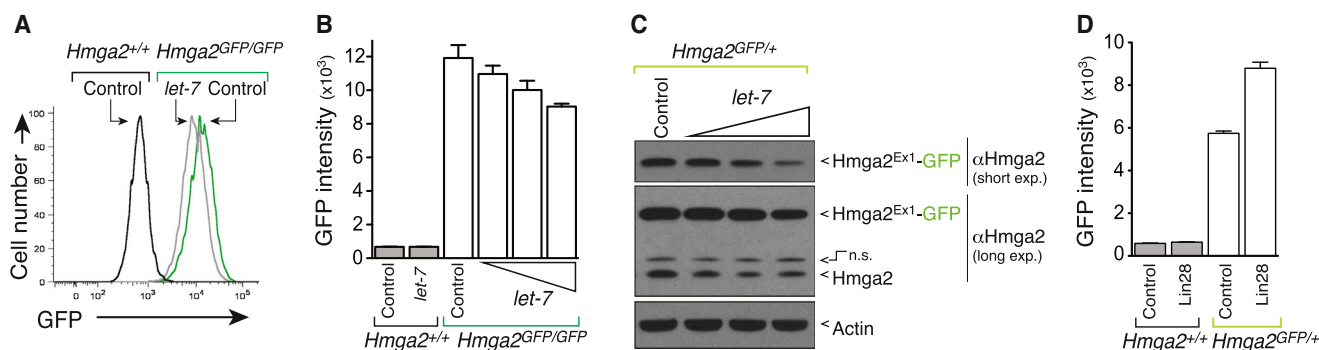
Figure 3. Hmga2^Ex1-GFP Remains under miRNA Control

(A) Flow cytometric analysis of *Hmga2^GFP/GFP* MEFs after transfection with a *let-7* mimic documents *let-7*-mediated reduction in GFP fluorescence.

(B) Quantification of the GFP mean fluorescence intensity (MFI) after transfection with a control miRNA mimic or increasing amounts of *let-7* mimic (2, 10, and 36 pmol; only 36 pmol *let-7* mimic is shown for *Hmga2^+/+* cells). Mean ± SD of triplicate wells is shown. Data are representative of four experiments.

(C) *Let-7*-mediated Hmga2^Ex1-GFP protein reduction parallels that of full-length Hmga2 in *Hmga2^GFP/+* MEFs. Cells were transfected with a control miRNA mimic (36 pmol) or increasing amounts of *let-7* mimic (2, 10, and 36 pmol). Both Hmga2 and Hmga2^Ex1-GFP expression decreases in a dose-dependent manner. A nonspecific band (n.s.) is marked. Actin shows equal loading. short exp., short exposure; long exp., long exposure.

(D) Quantification of GFP MFI after stable retroviral Lin28 expression. Mean ± SD of triplicate wells is shown.

performed a series of experiments to test whether the *Hmga2^GFP* allele remains under *let-7* control. Transfection of *Hmga2^GFP/+* or *Hmga2^GFP/GFP* MEFs with a synthetic *let-7* mimic reduced GFP expression in a dose-dependent manner (Figures 3A–3C, S4A, and S4B; data not shown). Furthermore, the reduction of Hmga2^Ex1-GFP protein mirrored that of endogenous Hmga2 after transfection of *Hmga2^GFP/+* MEFs with the *let-7* mimic (Figure 3C). The Hmga2-GFP reporter was also increased by stable expression of the *let-7* inhibitor Lin28 (Figures 3D and S4C). Collectively, these results indicate that the Hmga2^Ex1-GFP reporter generated after Cre-mediated cassette inversion remains under appropriate posttranscriptional 3′ UTR regulation.

Cassette inversion introduced a premature stop codon after the GFP-coding sequence followed by several additional exons; therefore, it is possible that the composite transcript could be subject to nonsense-mediated decay (NMD) (Chang et al., 2007). To directly compare the expression of the *Hmga2^wt* and *Hmga2^GFP* transcripts in isogenic cells, we treated *Hmga2^CK/CK* cells with a Cre-expressing adenoviral vector to generate *Hmga2^GFP/GFP* cells. Quantitative PCR and northern blotting for total *Hmga2* indicate that the expression of *Hmga2^wt* in *Hmga2^CK/CK* cells is comparable to the expression of *Hmga2^GFP* in *Hmga2^GFP/GFP* cells. Additionally, inhibition of NMD induced comparable upregulation of *Hmga2^GFP* and *Hmga2^wt* transcripts, had no effect on *let-7*-mediated repression of either *Hmga2^wt* or *Hmga2^GFP* mRNA or protein, and did not alter Lin28-induced upregulation of Hmga2^GFP protein (Figure S4).

Hmga2 Protein Is Responsible for the Phenotypes Associated with *Hmga2* Deficiency

Previously, *Hmga2* null alleles have been generated through spontaneous deletion, fortuitous insertional mutagenesis, and conventional gene targeting. In all cases, *Hmga2* null mice displayed a dramatic dwarfism phenotype (Xiang et al., 1990; Zhou et al., 1995). To determine whether *Hmga2^GFP* mice phenocopy *Hmga2* null mice, we analyzed the growth of *Hmga2^+/+*,

Hmga2^GFP/+, and *Hmga2^GFP/GFP* mice. Consistent with the *Hmga2*-null phenotype, postnatal day 0 (P0) *Hmga2^GFP/GFP* mice were slightly smaller than *Hmga2^+/+* and *Hmga2^GFP/+* littermates (Figure 4A). Both male and female *Hmga2^GFP/GFP* mice recapitulated the expected dwarfism phenotype, with 12-week-old mice being shorter and ∼50% the weight of control mice (Figures 4B–4F). Additional aspects of the *Hmga2* null phenotype were also observed in the *Hmga2^GFP/GFP* mice, including a disproportionate reduction in fat tissue and pinnae size and a relatively mild reduction in brain size (Nishino et al., 2008) (data not shown). *Hmga2^CK/CK* mice were phenotypically indistinguishable from *Hmga2^+/+* control mice, indicating that the cassette itself does not dramatically affect the expression of the targeted allele prior to its inversion by Cre (Figures S2E and S2F).

DISCUSSION

The unequivocal coupling of gene disruption and reporter expression, unaltered expression of intronic miRNAs in the presence of disrupted protein-coding function, and the induction of a fluorescent reporter under the control of all endogenous genomic elements and 3′ UTR control regions all support the broad use of CK alleles to better understand gene function in vivo. Because miRNA-mediated control of protein expression and the role of ceRNA networks are almost completely unstudied in this context, combined approaches that employ both CK alleles and alleles that entirely abrogate transcript expression could represent an important future method to study the importance of ceRNA and other noncoding transcript functions in complex tissues and at the organismal level. This system also circumvents the major difficulties encountered with mosaic deletion resulting from ineffective Cre-ER or weak cell-type-specific Cre alleles because it allows cells that harbor the deleted target gene to be definitively distinguished from those that do not and, in fact, allows for neighboring control and null cells to be
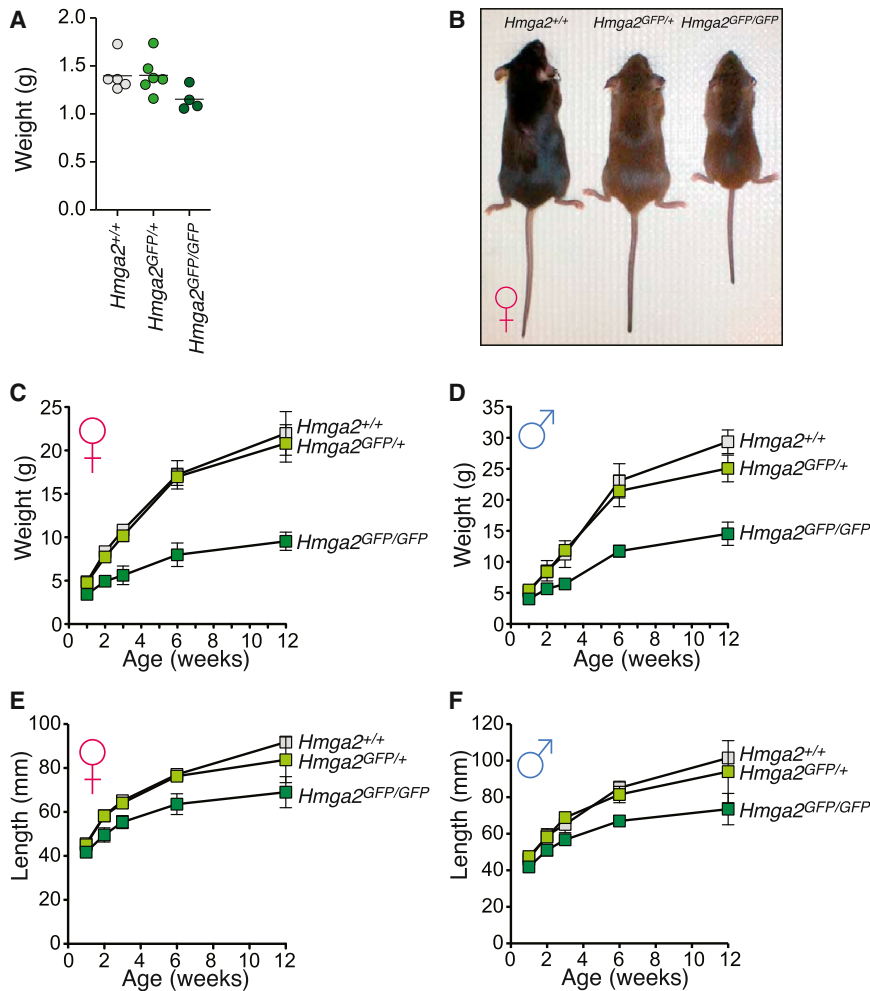
Figure 4. *Hmga2^GFP* Mice Recapitulate the Null Phenotype

(A) P0 *Hmga2^GFP/GFP* pups are ∼20% smaller than control littermates. Each dot represents a P0 mouse, and the bar represents the mean.

(B) Photo of 12-week-old *Hmga2^+/+*, *Hmga2^GFP/+*, and *Hmga2^GFP/GFP* female littermates.

(C and D) Growth curves of female (C) and male (D) *Hmga2^+/+*, *Hmga2^GFP/+*, and *Hmga2^GFP/GFP* mice. Mean ± SD is shown for each time point.

(E and F) Length curves of female (E) and male (F) *Hmga2^+/+*, *Hmga2^GFP/+*, and *Hmga2^GFP/GFP* mice. Mean ± SD is shown for each time point.

protein-coding function by introducing a premature stop codon after GFP could induce NMD (Brogna and Wen, 2009). Although we found that switching of the *Hmga2^wt* transcript to *Hmga2^GFP* did not reduce steady-state mRNA abundance, the effect of cassette inversion on transcript stability on additional targets will have to be determined empirically.

Hmga2 is an important regulator of body size across a wide range of species, including humans (Makvandi-Nejad et al., 2012; Weedon et al., 2007). Previous murine *Hmga2* null alleles have all resulted from the abrogation of the entire *Hmga2* transcript (Xiang et al., 1990; Zhou et al., 1995). That our *Hmga2^GFP* allele recapitulates the null phenotypes confirms that most, if not all, of the developmental phenotypic consequences observed in these earlier models are driven by the loss of Hmga2 protein function rather

analyzed within the same tissue. Although other strategies to link gene disruption with reporter expression have been developed, none allows the generalizable and streamlined generation of alleles that specifically disrupt protein-coding function while leaving all nonprotein-coding functions intact (Table S1). The CK allele system will be widely applicable for creating Cre-regulated alleles to better understand the protein-coding element of genes during many stages of development, normal homeostasis, and disease states.

The CK allele system should be amenable for targeting most genes because it employs a well-characterized SA that has been used for multiple genome-wide loss-of-function screens in haploid cells (Bürckstümmer et al., 2013; Carette et al., 2011). However, a few situations do exist that preclude the use of the CK technology. Introns flanked by noncanonical SA-donor pairs are rare, constituting less than 1% of introns, but would be insensitive to an inverted CK allele. Additionally, our system is also limited to genes that contain at least one intron and in which an exon 1-GFP fusion would lack function.

mRNA stability is influenced by RNA sequence and structure, and both are altered in the CK allele after cassette inversion and incorporation of the GFP sequence. Additionally, interrupting the

than the function of miR-763 or disruption of a *let-7-Hmga2* ceRNA network. Our results provide a loss-of-function complement to the increased body size in transgenic mice expressing increased levels of Hmga2 protein (Arlotta et al., 2000; Battista et al., 1999). Given that the molecular functions of Hmga2 during development and disease remain poorly characterized, our allele should aid in dissecting the role of this important chromatin-regulating protein.

EXPERIMENTAL PROCEDURES

Creation of Base-Targeting Vectors

The allele design (termed the CK system) is based on the creation of an in-frame GFP fusion protein; therefore, CK plasmids were generated for each of the three reading frames (Figure S1; pCK [+0], pCK [+1], and pCK [+2]). To generate these plasmids, eGFP-SA-2272*lox*P-5171*lox*P fragments in each of the three reading frames were amplified with the addition of a 5′ multiple cloning site, a consensus splice donor site (in reverse, actcacctt), and a 3′ PacI site. These fragments were ligated into a plasmid backbone that contained a diphtheria toxin A expression cassette, a multiple cloning site, an FRT-flanked neomycin-resistance gene (*neo*), 2272*lox*P-5171*lox*P sites, and a final multiple cloning site (Figures 1A and S1; Supplemental Experimental Procedures).

Generation of CK Alleles

Construction of a CK allele for any gene of interest requires the amplification and ligation of the targeting arms into the appropriate pCK plasmid. Having an exon within the region of the genome to be inverted is not required for the system to work, but an exon can be included if desired (as in the $Hmga2^{CK}$ allele). Reasons to incorporate an exon may include the ATG start site lies within the second exon, a potential alternate ATG start lies within the second exon, or if one simply desires to further guarantee gene disruption. If no exon is included, it would likely be preferable to insert the cassette in the first intron of the target gene to limit the length of the endogenous gene produced prior to splicing into the SA-GFP-splice donor cassette after Cre-mediated inversion. In cases where an exon is included within the region that will be inverted, the overall conformation would likely be very similar to the $Hmga2^{CK}$ allele that we have generated. If an exon is to be included, it will be necessary to clone three regions into the pCK vector in any order; otherwise only two regions will be required (Figure S1A; Supplemental Experimental Procedures).

Hmga2 Targeting and Generation of *Hmga2^CK^* and *Hmga2^GFP^* Mice

The *Hmga2*-targeting construct was linearized and electroporated into v6.5 ES cells (a kind gift from Rudolph Jaenisch) using standard conditions. Neomycin (300 μg/ml G418)-resistant colonies were picked, expanded, and screened by long-range PCR for correct targeting of the *Hmga2* locus. Of 294 clones, 5 (1.7%) were correctly targeted. C57BL/6J blastocysts were injected with targeted ES cells resulting in several high-percentage chimeras. Targeted $Hmga2^{Neo}$ mice were bred to $Rosa26^{FLPe}$ mice (The Jackson Laboratory; stock number 003946) to delete the *neo* cassette and generate the $Hmga2^{CK}$ allele. A subset of $Hmga2^{CK}$ mice was bred to β-actin-Cre mice to create mice with a germline $Hmga2^{GFP}$ allele (Figure 1A). The MIT and Stanford Institutional Animal Care and Use Committees approved all animal studies and procedures.

Generation of MEFs, Adenoviral Infection, and Flow Cytometry

MEFs were generated from embryos between E12.5 and E16.5. To express Cre recombinase in MEFs, cells were infected with either Adeno-Cre or Ad-eno-FLPo (control) purchased from the Gene Transfer Vector Core at the University of Iowa. Analysis of the cells was performed 3 days after infection. Data were collected on a BD LSR II analyzer (BD Biosciences) at the Stanford Shared FACS Facility. For the quantification of GFP in MEFs on a LSR II flow cytometer, cells were thoroughly trypsinized and stained with DAPI to exclude dead cells.

Western Blot Analysis

MEFs were trypsinized and pelleted before being lysed. Denatured samples were separated by SDS-PAGE, followed by transfer to polyvinylidene fluoride membranes. Membranes were stained with the indicated primary antibodies rabbit anti-Hmga2-P1 (BioCheck; 59170AP), rabbit anti-GFP (Cell Signaling Technology; 2956), and β-actin antibody (Sigma-Aldrich; A5441). For LI-COR imaging, membranes were first stained with the indicated primary antibodies followed by staining with either the goat anti-mouse-IRDye 800CW (926-32210; LI-COR Biosciences) or the goat anti-rabbit-Alexa Fluor 680 (A-21109; Life Technologies) for 1 hr at room temperature in the dark. After three washes (1× PBS, 0.05% Tween 20), the membranes were imaged with the LI-COR imager LI-COR Odyssey.

Upf1 Knockdown

The small interfering RNA (siRNA) specific for murine *Upf1* transcripts was purchased from Life Technologies (siRNA ID, s72879). The procedure to deliver the control and *Upf1* siRNA into MEFs is described below. Transfectants were harvested 3 days posttransfection and lysed for quantification of mRNA. For functional validation of *Upf1* knockdown, *Gas5* transcript was used as described by Keeling et al. (2013).

Northern Blot Analysis

Northern blots were performed using standard methods. In brief, radiolabeled probes were generated through random hexamer-primed synthesis by using a PCR-amplified cDNA sequence as the template and DNA polymerase I Klenow fragment and α-32P-dCTP. Probes were purified using a spin column kit

(Nucleotide Removal Kit; QIAGEN). Membranes were prewashed, hybridized with radiolabeled probe for 16–20 hr at 42°C, washed twice with 2× saline sodium citrate (SSC) and 0.1% SDS at 42°C for 20 min, and twice with 0.1× SSC and 0.1% SDS at 55°C for 20 min. Membranes were imaged on a Molecular Dynamics Typhoon 9400 Imager (Amersham/GE Healthcare).

Let-7 Transfection and Retroviral Lin28 Expression

To deliver control (mirVana-negative control mimic) and/or *let-7* (mirVana mmu-let-7a-5p miRNA mimic) miRNA mimics into MEFs, Lipofectamine RNAiMAX was used according to the manufacturer's description (Life Technologies). To normalize total input mimics, compensatory levels of control mimic were added with titrated *let-7* mimic. Overexpression of the *let-7* inhibitor Lin28 (Viswanathan et al., 2008) was achieved by transducing MEFs with a retroviral vector (murine stem cell virus [MSCV]) harboring the *Lin28* cDNA (Addgene; pMSCV-mLin28A/*neo*, plasmid No. 26357) (Viswanathan et al., 2009). An MSCV-*neo* vector without Lin28 was used as a negative control. Transduced MEFs were cultured in the presence of G418 (0.5 mg/ml) for 3 days to allow complete enrichment of transductants. Selected MEFs were assessed for GFP and Hmga2 levels using flow cytometry and western blot analysis.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.05.031.

ACKNOWLEDGMENTS

REFERENCES

Alexander, C.M., Puchalski, J., Klos, K.S., Badders, N., Ailles, L., Kim, C.F., Dirks, P., and Smalley, M.J. (2009). Separating stem cells by flow cytometry: reducing variability for solid tissues. Cell Stem Cell *5*, 579–583.

Arlotta, P., Tai, A.K., Manfioletti, G., Clifford, C., Jay, G., and Ono, S.J. (2000). Transgenic mice expressing a truncated form of the high mobility group I-C protein develop adiposity and an abnormally high prevalence of lipomas. J. Biol. Chem. *275*, 14394–14400.

Battista, S., Fidanza, V., Fedele, M., Klein-Szanto, A.J., Outwater, E., Brunner, H., Santoro, M., Croce, C.M., and Fusco, A. (1999). The expression of a

truncated HMGI-C gene induces gigantism associated with lipomatosis. Cancer Res. *59*, 4793–4797.

Brogna, S., and Wen, J. (2009). Nonsense-mediated mRNA decay (NMD) mechanisms. Nat. Struct. Mol. Biol. *16*, 107–113.

Bürckstümmer, T., Banning, C., Hainzl, P., Schobesberger, R., Kerzendorfer, C., Pauler, F.M., Chen, D., Them, N., Schischlik, F., Rebsamen, M., et al. (2013). A reversible gene trap collection empowers haploid genetics in human cells. Nat. Methods *10*, 965–971.

Carette, J.E., Guimaraes, C.P., Varadarajan, M., Park, A.S., Wuethrich, I., Godarova, A., Kotecki, M., Cochran, B.H., Spooner, E., Ploegh, H.L., and Brummelkamp, T.R. (2009). Haploid genetic screens in human cells identify host factors used by pathogens. Science *326*, 1231–1235.

Carette, J.E., Guimaraes, C.P., Wuethrich, I., Blomen, V.A., Varadarajan, M., Sun, C., Bell, G., Yuan, B., Muellner, M.K., Nijman, S.M., et al. (2011). Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. Nat. Biotechnol. *29*, 542–546.

Chang, Y.F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. Annu. Rev. Biochem. *76*, 51–74.

Farley, F.W., Soriano, P., Steffen, L.S., and Dymecki, S.M. (2000). Widespread recombinase expression using FLPeR (flipper) mice. Genesis *28*, 106–110.

Hadjantonakis, A.K., Dickinson, M.E., Fraser, S.E., and Papaioannou, V.E. (2003). Technicolour transgenics: imaging tools for functional genomics in the mouse. Nat. Rev. Genet. *4*, 613–625.

Karreth, F.A., Tay, Y., Perna, D., Ala, U., Tan, S.M., Rust, A.G., DeNicola, G., Webster, K.A., Weiss, D., Perez-Mancera, P.A., et al. (2011). In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell *147*, 382–395.

Keeling, K.M., Wang, D., Dai, Y., Murugesan, S., Chenna, B., Clark, J., Belakhov, V., Kandasamy, J., Velu, S.E., Baasov, T., and Bedwell, D.M. (2013). Attenuation of nonsense-mediated mRNA decay enhances in vivo nonsense suppression. PLoS One *8*, e60478.

Kuhnert, F., Mancuso, M.R., Hampton, J., Stankunas, K., Asano, T., Chen, C.Z., and Kuo, C.J. (2008). Attribution of vascular phenotypes of the murine Egfl7 locus to the microRNA miR-126. Development *135*, 3989–3993.

Kumar, M.S., Armenteros-Monterroso, E., East, P., Chakravorty, P., Matthews, N., Winslow, M.M., and Downward, J. (2014). HMGA2 functions as a competing endogenous RNA to promote lung cancer progression. Nature *505*, 212–217.

Lao, Z., Raju, G.P., Bai, C.B., and Joyner, A.L. (2012). MASTR: a technique for mosaic mutant analysis with spatial and temporal control of recombination using conditional floxed alleles in mice. Cell Reports *2*, 386–396.

Lee, G., and Saito, I. (1998). Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. Gene *216*, 55–65.

Lee, Y.S., and Dutta, A. (2007). The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev. *21*, 1025–1030.

Makvandi-Nejad, S., Hoffman, G.E., Allen, J.J., Chu, E., Gu, E., Chandler, A.M., Loredo, A.I., Bellone, R.R., Mezey, J.G., Brooks, S.A., and Sutter, N.B. (2012). Four loci explain 83% of size variation in the horse. PLoS One *7*, e39929.

Mandalos, N., Saridaki, M., Harper, J.L., Kotsoni, A., Yang, P., Economides, A.N., and Remboutsika, E. (2012). Application of a novel strategy of engineering conditional alleles to a single exon gene, Sox2. PLoS One *7*, e45768.

Mayr, C., Hemann, M.T., and Bartel, D.P. (2007). Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science *315*, 1576–1579.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature *495*, 333–338.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3′ UTRs are the primary regulators of gene expression in the C. elegans germline. Curr. Biol. *18*, 1476–1482.

Miyaki, S., Sato, T., Inoue, A., Otsuki, S., Ito, Y., Yokoyama, S., Kato, Y., Takemoto, F., Nakasa, T., Yamashita, S., et al. (2010). MicroRNA-140 plays dual roles in both cartilage development and homeostasis. Genes Dev. *24*, 1173–1185.

Moon, A.M., Boulet, A.M., and Capecchi, M.R. (2000). Normal limb development in conditional mutants of Fgf4. Development *127*, 989–996.

Nakamura, Y., Inloes, J.B., Katagiri, T., and Kobayashi, T. (2011). Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling. Mol. Cell. Biol. *31*, 3019–3028.

Nishino, J., Kim, I., Chada, K., and Morrison, S.J. (2008). Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell *135*, 227–239.

Osokine, I., Hsu, R., Loeb, G.B., and McManus, M.T. (2008). Unintentional miRNA ablation is a risk factor in gene knockout studies: a short report. PLoS Genet. *4*, e34.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033–1038.

Potocnik, A.J., Brakebusch, C., and Fässler, R. (2000). Fetal and adult hematopoietic stem cells require beta1 integrin function for colonizing fetal liver, spleen, and bone marrow. Immunity *12*, 653–663.

Rajewsky, K. (2007). From a dream to reality. Eur. J. Immunol. *37* (*Suppl 1*), S134–S137.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell *146*, 353–358.

Salzman, J., Gawad, C., Wang, P.L., Lacayo, N., and Brown, P.O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS One *7*, e30733.

Schmidt-Supprian, M., and Rajewsky, K. (2007). Vagaries of conditional gene targeting. Nat. Immunol. *8*, 665–668.

Schnütgen, F., and Ghyselinck, N.B. (2007). Adopting the good reFLEXes when generating conditional alterations in the mouse genome. Transgenic Res. *16*, 405–413.

Schnütgen, F., Doerflinger, N., Calléja, C., Wendling, O., Chambon, P., and Ghyselinck, N.B. (2003). A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. Nat. Biotechnol. *21*, 562–565.

Schnütgen, F., De-Zolt, S., Van Sloun, P., Hollatz, M., Floss, T., Hansen, J., Altschmied, J., Seisenberger, C., Ghyselinck, N.B., Ruiz, P., et al. (2005). Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. Proc. Natl. Acad. Sci. USA *102*, 7221–7226.

Srinivas, S., Watanabe, T., Lin, C.S., William, C.M., Tanabe, Y., Jessell, T.M., and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev. Biol. *1*, 4.

Stern, P., Astrof, S., Erkeland, S.J., Schustak, J., Sharp, P.A., and Hynes, R.O. (2008). A system for Cre-regulated RNA interference in vivo. Proc. Natl. Acad. Sci. USA *105*, 13895–13900.

Sundaram, G.M., Common, J.E., Gopal, F.E., Srikanta, S., Lakshman, K., Lunny, D.P., Lim, T.C., Tanavde, V., Lane, E.B., and Sampath, P. (2013). 'See-saw' expression of microRNA-198 and FSTL1 from a single transcript in wound healing. Nature *495*, 103–106.

Tasic, B., Miyamichi, K., Hippenmeyer, S., Dani, V.S., Zeng, H., Joo, W., Zong, H., Chen-Tsai, Y., and Luo, L. (2012). Extensions of MADM (mosaic analysis with double markers) in mice. PLoS One *7*, e33332.

Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S.M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F., et al. (2011). Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell *147*, 344–357.

Theis, M., de Wit, C., Schlaeger, T.M., Eckardt, D., Krüger, O., Döring, B., Risau, W., Deutsch, U., Pohl, U., and Willecke, K. (2001). Endothelium-specific replacement of the connexin43 coding region by a lacZ reporter gene. Genesis *29*, 1–13.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97–100.

Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L., et al. (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat. Genet. *41*, 843–848.

Wang, S., Aurora, A.B., Johnson, B.A., Qi, X., McAnally, J., Hill, J.A., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2008). The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. Dev. Cell *15*, 261–271.

Weedon, M.N., Lettre, G., Freathy, R.M., Lindgren, C.M., Voight, B.F., Perry, J.R., Elliott, K.S., Hackett, R., Guiducci, C., Shields, B., et al.; Diabetes Genetics Initiative; Wellcome Trust Case Control Consortium (2007). A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat. Genet. *39*, 1245–1250.

Wellershaus, K., Degen, J., Deuchars, J., Theis, M., Charollais, A., Caille, D., Gauthier, B., Janssen-Bienhold, U., Sonntag, S., Herrera, P., et al. (2008). A new conditional mouse mutant reveals specific expression and functions of connexin36 in neurons and pancreatic beta-cells. Exp. Cell Res. *314*, 997–1012.

Winslow, M.M., Dayton, T.L., Verhaak, R.G., Kim-Kiselak, C., Snyder, E.L., Feldser, D.M., Hubbard, D.D., DuPage, M.J., Whittaker, C.A., Hoersch, S., et al. (2011). Suppression of lung adenocarcinoma progression by Nkx2-1. Nature *473*, 101–104.

Xiang, X., Benson, K.F., and Chada, K. (1990). Mini-mouse: disruption of the pygmy locus in a transgenic insertional mutant. Science *247*, 967–969.

Xin, H.B., Deng, K.Y., Shui, B., Qu, S., Sun, Q., Lee, J., Greene, K.S., Wilson, J., Yu, Y., Feldman, M., and Kotlikoff, M.I. (2005). Gene trap and gene inversion methods for conditional gene inactivation in the mouse. Nucleic Acids Res. *33*, e14.

Yoo, A.S., Staahl, B.T., Chen, L., and Crabtree, G.R. (2009). MicroRNA-mediated switching of chromatin-remodelling complexes in neural development. Nature *460*, 642–646.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., and Song, E. (2007). let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell *131*, 1109–1123.

Zhao, Y., Ransom, J.F., Li, A., Vedantham, V., von Drehle, M., Muth, A.N., Tsuchihashi, T., McManus, M.T., Schwartz, R.J., and Srivastava, D. (2007). Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. Cell *129*, 303–317.

Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. (1995). Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature *376*, 771–774.

Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D., and Luo, L. (2005). Mosaic analysis with double markers in mice. Cell *121*, 479–492.