The number of long noncoding RNAs (IncRNAs) has grown rapidly; however, our understanding of their function remains limited. Although cultured cells have facilitated investigations of IncRNA function at the molecular level, the use of animal models provides a rich context in which to investigate the phenotypic impact of these molecules. Promising initial studies using animal models demonstrated that IncRNAs influence a diverse number of phenotypes, ranging from subtle dysmorphia to viability. Here, we highlight the diversity of animal models and their unique advantages, discuss the use of animal models to profile IncRNA expression, evaluate experimental strategies to manipulate IncRNA function in vivo, and review the phenotypes attributable to IncRNAs. Despite a limited number of studies leveraging animal models, IncRNAs are already recognized as a notable class of molecules with important implications for health and disease.
Investigating long noncoding RNAs using animal models

Michael Feyder1,2 and Loyal A. Goff1,2,3

1McKusick-Nathans Institute for Genetic Medicine, 2Department of Neuroscience, and 3Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

The number of long noncoding RNAs (lncRNAs) has grown rapidly; however, our understanding of their function remains limited. Although cultured cells have facilitated investigations of lncRNA function at the molecular level, the use of animal models provides a richer context in which to investigate the phenotypic impact of these molecules. Promising initial studies using animal models demonstrated that lncRNAs influence a diverse number of phenotypes, ranging from subtle dysmorphia to viability. Here, we highlight the diversity of animal models and their unique advantages, discuss the use of animal models to profile lncRNA expression, evaluate experimental strategies to manipulate lncRNA function in vivo, and review the phenotypes attributable to lncRNAs. Despite a limited number of studies leveraging animal models, lncRNAs are already recognized as a notable class of molecules with important implications for health and disease.

Introduction

Approximately three-quarters of the mammalian genome is transcribed into RNA (1–3); however, only a fraction of this transcription produces mRNA, whose mature nucleotide sequence serves as a template for protein synthesis (3). The function of this nonprotein-coding RNA (or noncoding RNA) is mostly obscure, despite a larger number of noncoding genes than protein-coding genes. Given the broad functional repertoire derived from protein-coding RNA, it is perhaps not surprising that the relatively few known functions for noncoding RNAs also span diverse cellular processes, resulting in an array of noncoding RNA subclassifications (4).

One particular subclass, long noncoding RNAs (lncRNAs), represents a large family of noncoding RNA molecules with potentially broad implications for basic science, health, and disease. In this Review, we focus on the use of animal models to discover novel lncRNAs and to investigate their significance in vivo.

lncRNA definitions and mechanisms

lncRNAs represent a burgeoning class of molecules broadly defined as RNA transcripts longer than 200 nucleotides, with no protein-coding potential. This length is somewhat arbitrary, but it serves to distinguish them from shorter, biologically distinct noncoding RNAs, such as microRNAs. In humans, high-throughput experimental approaches have led to the rapid identification of approximately 16,000 lncRNA genes thus far, rivaling the approximately 20,000 protein-coding genes (5–10) (Figure 1A). These lncRNAs are informatically predicted to lack protein-coding potential, although notable cases of presumed lncRNAs encoding micropeptides (11, 12) or acting as precursors to microRNAs (13, 14) exist. Therefore, current definitions attempt to corral this emerging class of molecules, whose validation and function in vivo remain to be fully elucidated (Figure 1B).

lncRNAs are found in both the nucleus and the cytoplasm. The majority of lncRNAs reside in the nucleus (10), where they can act proximally or distally to their site of transcription, functioning in cis or trans, respectively. For example, during X chromosome inactivation in mice, the lncRNA Xist functions in cis to initiate silencing of genes across the same X chromosome from which it was originally transcribed (15). Conversely, the mouse lncRNA Trp53cor1 (also known as linRNA-p21) acts in trans to globally repress the expression of hundreds of genes distant from its site of transcription (16). Additionally, the transcription of a cis-acting lncRNA, per se, rather than its resulting RNA product, can also have a biological effect (17). As discussed below, this possibility raises important considerations when designing experiments to manipulate lncRNA expression and function.

The molecular mechanisms for most lncRNAs remain largely unknown. They can bind to DNA, RNA, or proteins, and many techniques have been developed to assay these interactions (Figure 2A). Techniques exploring lncRNA-DNA interactions, such as capture hybridization analysis of RNA targets (CHART) (18) and chromatin isolation by RNA purification (ChiRIP) (19), utilize complementary oligonucleotides that hybridize to a lncRNA of interest and serve as an affinity handle to enrich for bound DNA. Similar hybridization approaches, such as radioimmunooassay sequencing (RIA-seq), can also be used to assay lncRNA-RNA interactions (20). Other techniques, such as RNA immunoprecipitation (RIP) or its variants, cross-linking immunoprecipitation (CLIP) (21, 22), and photoactivatable ribonucleoside–enhanced cross-linking immunoprecipitation (PAR-CLIP) (23), utilize antibodies to purify lncRNA-protein complexes. Additionally, high-throughput approaches to investigate RNA secondary structure, such as in vivo click selective 2′-hydroyl acylation and profiling experiment (icSHAPE) (24) or selective 2′-hydroxyl acylation analyzed...
Accumulating evidence suggests that IncRNA dysfunction promotes disease (26). Notably, approximately 40% of disease- or trait-associated SNPs are found within the noncoding regions flanking protein-coding genes, where a subclass of IncRNAs, termed long intergenic noncoding RNAs, reside (27). Although the functional consequences, if any, for many of these SNPs remain to be experimentally evaluated, they may promote disease by affecting IncRNA function and/or expression (28). For example, an SNP within the human IncRNA myocardial infarction–associated transcript (MIAT), which is associated with an increased risk for myocardial infarction, also increases the binding affinity of MIAT for nuclear proteins (29). This same SNP also increases expression levels of MIAT, although the precise molecular mechanisms remain unknown (29). Additionally, in a human cell line, the binding of IncRNAs to the protein complex Mediator promotes permissive chromatin states necessary for gene expression, and mutations in Mediator that result in intellectual disability have been shown to impair its interaction with IncRNAs (30). An alternative scenario, whereby an alteration of a IncRNA disrupts its interaction with Mediator, could conceivably produce similar, yet less pronounced, effects. These results contribute to a growing body of evidence implicating IncRNAs in disease.

Genetic linkage studies have uncovered new IncRNAs, providing additional evidence to suggest that IncRNA dysfunction promotes disease. For example, HELPP syndrome, which occurs in mothers during pregnancy and is characterized by hemolysis, elevated liver enzymes, and low platelet counts, is associated with a previously unknown placental IncRNA (31). Mutations within this IncRNA increase the in vitro proliferation of human placenta cells in a model of this syndrome (31). Additionally, brachydactyly type E can result from a chromosomal rearrangement that disrupts local interactions between the human IncRNA chondrogenesis-associated transcript (CISTR) and the spatially proximal protein-coding gene parathyroid hormone–like hormone (PTHLH), diminishing the expression of PTHLH mRNA (32). These findings highlight how IncRNA dysfunction impacts cellular processes with clinically relevant consequences.

Notably, IncRNA expression in human tumors is associated with clinical outcomes for a variety of cancers, and xenograft studies in mice have been instrumental when extending these findings in vivo. For example, more than 100 IncRNAs correlate with overall or progression-free survival for ovarian cancer, prostate cancer, glioblastomas, and lung squamous cell carcinomas (33). Nine of these IncRNAs consistently correlate across these cancer subtypes (33). Other examples include the IncRNA HOTAIR, which is overexpressed in primary breast tumors, and elevated expression of HOTAIR in these tumors correlates with an increased probability of metastasis and a decreased overall rate of survival (34). Additionally, xenografts overexpressing HOTAIR in a cell line derived from metastatic breast tissue have an increased propensity for metastasis in mice (34). Furthermore, approximately 1,900 IncRNAs are differentially expressed in T cell acute lymphoblastic leukemia (35). This finding is further supported in the same study by demonstrating in mice that one of these IncRNAs, LUNAR1, is necessary for xenograft tumor growth of a human T cell lymphoma cell line (35). Finally, xenograft studies can also support the in vivo significance of IncRNAs implicated in biological processes such as hypoxia that are known to promote cancer progression (36). For example, the IncRNAs NPTN-IT1 (also known as IncRNA-LET) (37), TP53COR1 (38), and LINC-ROR (39) all regulate hypoxia-induced signaling and affect xenograft growth. In summary, accumulating evidence implicates IncRNAs in cancer development, a finding supported by the use of laboratory animals for xenograft studies.

Diversity of animal models

Animal models vary in biological complexity, which can be leveraged depending on the area of investigation. Commonly used animal models more evolutionarily divergent from humans include the nematode (Caenorhabditis elegans), the fruit fly (Drosophila melanogaster), and the zebrafish (Danio rerio). The reduced com-
Elegans. C. dysfunction may contribute to human disease. For example, studies of protein-coding genes utilizing C. elegans and D. melanogaster have aided in understanding the molecular mechanisms of apoptosis, whose dysfunction contributes to cancer development and progression (40). Less evolutionarily divergent animal models include the rat (Rattus norvegicus), the mouse (Mus musculus), and the nonhuman primate. The greater biological complexity and physiological similarity of these organisms to humans may better model the complex biology of diseases, such as that underlying tumor growth and metastasis (41, 42). Additional considerations when choosing an appropriate animal model include financial cost, ethical considerations, potential throughput, and ease of experimental manipulation.

Evolutionary conservation of lncRNAs
In addition to humans, numerous lncRNAs are also found in the fruit fly (43–46), the nematode (47), the zebrafish (48, 49), and the mouse (1, 6). Cross-species comparisons have demonstrated that lncRNAs can contain short, conserved regions (9, 50) and also show evidence of purifying selection (51). In general, however, their primary sequence is weakly conserved across species (49, 52–56). For example, only 29 lncRNAs are conserved between zebrafish and humans (49). Strikingly, in the same study, the phenotype in zebrafish following functional inactivation of conserved lncRNAs can be rescued by an orthologous gene from either mice or humans, despite diverging primary sequences (49). These results suggest that the higher-order structure of a lncRNA, rather than its primary sequence, may be conserved. Given the approximately 500 computationally predicted, intergenic RNA secondary structures conserved across vertebrates (57), zebrafish may be a robust model for probing a conserved lncRNA function that is not readily apparent from the primary sequence.

Species-specific lncRNAs may broadly highlight the involvement of lncRNAs in certain conditions in which certain animal models have afforded advantages. For example, both the fruit fly and the nematode are preferentially used to model aging, as their life span is shorter than that of the mouse or the rat (58). Using a genetic model of aging in the nematode, it has been demonstrated that binding of the lncRNA tts-1 to the ribosome suppresses ribosomal levels and promotes longevity, implicating lncRNAs in this phenotype (59).

Other advantages of animal models more evolutionarily distant from humans include the ability to conduct genetic screens (60). For example, expanded nucleotide repeats within the human lncRNA ATXN8OS (also known as SCA8) result in the neurodegenerative disorder spinocerebellar ataxia (61). Using the retina of the fruit fly as a model system, the expression of ATXN8OS, either with or without expanded repeats, results in neurodegeneration (62). However, a genetic screen in this same study demonstrated that this phenotype is differentially modified by RNA-binding proteins (62), in agreement with accumulating evidence implicating aberrant RNA-protein interactions in neurodegenerative disorders (63). Therefore, the use of animal models more evolutionarily distant from humans to study lncRNA function may serve to uncover fundamental roles of lncRNAs and their dysfunction in disease.

Animal models and lncRNA profiling
Animal models generate a variety of tissue and cell types from which to profile lncRNA expression. This feature renders animal models well suited for the study of lncRNAs, given that lncRNA expression is more tissue and cell type specific than is protein-coding gene expression (5, 10, 64). When compared with cultured cells, material derived from animals is generally less abundant and more time-consuming to generate. Nonetheless, animal models provide direct, in vivo evidence of endogenous lncRNA expression and dramatically narrow the relevant search space when investigating phenotypes.

Modern techniques available for genomics research, such as RNA-seq, provide an unbiased, high-throughput approach to investigating endogenous lncRNA expression and to discovering and assembling novel lncRNAs de novo (64). LncRNA expres-
development of the cerebral cortex (64, 68). Profiling endogenous lncRNA expression may therefore highlight the diversity of genetically defined cell types and their potential to promote disease. Additionally, environmental stimuli may affect lncRNA expression, making animal models an attractive resource for certain areas of research, such as neuroscience. Interestingly, lncRNAs are overrepresented in the human brain (10), where they may be fundamental to synaptic transmission (56). Additionally, both neuronal activation (69) and drugs of abuse (70, 71) promote lncRNA expression. The diversity of cell types within the brain, combined with myriad environmental stimuli, has the potential to reveal numerous unrecognized lncRNAs. In sum, animal models are a rich source of material from which lncRNA expression can be profiled across tissues or cell types, providing insight into their potential function in vivo.

Animal models and lncRNA function

The function of a gene is commonly inferred by attenuating or ablating its expression. Strategies to do so include physically interacting with its RNA (RNA-targeted approaches) or altering the underlying genetic locus (DNA-targeted approaches) (Figure 3, A and B). In animal models, these approaches may be applied at the zygotic stage, thereby modeling congenital conditions. They may also be applied to a specific cell population or during a defined time during development. Precise spatial and/or temporal control can refine the conclusions of a study or can circumvent deleterious developmental effects, such as embryonic lethality. Although these approaches have been successfully applied to protein-coding genes, their use requires caution. We discuss the advantages and limitations of these approaches below, highlighting examples from the literature to illustrate the diverse functions attributed to lncRNAs in vivo.

**RNA-targeted approaches and animal models.** RNA-targeted strategies introduce an exogenous RNA or an RNA analog that specifically binds to and functionally inactivates an endogenous lncRNA through complementary base pairing. Commonly used RNA-targeted approaches include RNAi and the use of morpholinos. These two approaches differ mechanistically. RNAi promotes lncRNA degradation, while morpholinos sterically hinder cellular processes, such as splicing, and prevent the formation or function of a mature lncRNA. Steric interference can also disrupt the binding of a lncRNA to other macromolecules, preventing the formation of functional complexes (49).
Zebrafish are an attractive animal model for investigating lncRNAs, because morpholinos can be readily microinjected at the embryonic, one-cell stage to attenuate lncRNA function throughout development and into adulthood (72). In these animals, the lncRNAs terminator, alien, and punisher are necessary for cardiovascular development, complementing similar inferences derived from cultured human and mouse cells (73). Additionally, attenuation of megamind (also known as tuna) in zebrafish impairs locomotion and disrupts brain development, while attenuation of cyrano in zebrafish results in neural tube defects and dysmorphic head and eyes (49). These results highlight the pronounced roles of lncRNAs in zebrafish development.

Mouse models are another valuable resource for understanding lncRNA function in vivo. Transgenic mice engineered to constitutively overexpress an antisense RNA targeted to Sfia3 (also known as Nanci) have abnormal epithelial morphogenesis in their developing lungs (74). A more localized, tissue-specific interference of lncRNAs using RNAi can also have phenotypic consequences. shRNA-mediated downregulation of Miat (also known as Gomafu) in the medial prefrontal cortex of the brain promotes anxiety-like behaviors (75), while local siRNA-mediated downregulation of Munc, which is specifically expressed in skeletal muscle, impairs myogenesis (76). Additionally, the use of an shRNA to revert the upregulation of Arid2-IR in the mouse kidney following renal inflammation also reverts the biochemical signatures of this condition (77). Future work using mutant- and virus-mediated strategies will likely reveal additional phenotypes directly attributable to lncRNAs.

Limitations of RNA-targeted approaches. While informative, these RNA-targeted approaches have important caveats. RNAi is effective for RNA exported to the cytoplasm but is relatively inefficient for RNAs residing in the nucleus (78), where many lncRNAs are found (10). This inefficiency may be particularly apparent for cis-acting lncRNAs, which act near their site of origin and therefore may necessitate rapid binding for inactivation. This limitation may be circumvented by using RNA analogs, such as locked nucleic acids, which are better suited for nuclear targets because of their faster kinetics (79). However, sterically interfering RNA analogs, including both locked nucleic acids and morpholinos, cannot be genetically encoded, precluding the generation of transgenic animals. Furthermore, both strategies need sufficient expression levels to be effective, potentially inducing off-target effects and general toxicity within the cell. Conversely, insufficient levels for either strategy may only attenuate, rather than eliminate, lncRNA function and may not be sufficient to produce a phenotypic effect (80). Finally, binding of an exogenous RNA after transcription may not affect lncRNA function, because its transcription, per se, may have a biological effect. Despite these limitations, successfully applied RNA-targeted approaches have revealed significant phenotypes following lncRNA dysfunction.

Clinical application of RNA-targeted approaches. Findings from animals models serve as a translational proof of concept for therapeutic interventions targeting lncRNAs, though significant hurdles exist (81). In mice, systemic administration of an antisense oligonucleotide can downregulate its complementary RNA in muscle tissue (82). For example, systemic administration of an antisense oligonucleotide complementary to the lncRNA Malat1 (also known as Neat2) can attenuate its expression by 80% (82). Although this approach was principally used to attenuate a pathological protein-coding RNA in a mouse model of myotonic dystrophy type 1, the attenuation of Malat1 suggests that lncRNAs are potential targets for this condition and possibly other muscle disorders. However, only RNAs retained in the nucleus are sensitive to this approach (82), a feature well suited for targeting the abundance of nuclear lncRNAs (10). This work not only highlights the value of animal models in advancing novel therapeutic approaches but also the feasibility of correcting pathological gene expression, including that of lncRNAs.

DNA-targeted approaches and animal models. Strategies to manipulate the genetic locus encoding a lncRNA, many of which have recently been reviewed (83), represent an alternative to RNA-targeted approaches. Transcription of both cis- and trans-acting lncRNAs can be eliminated, in contrast to the dose-dependent, posttranscriptional effects of RNA-targeted approaches. Furthermore, if this ablation occurs in the germline, it is possible to generate transgenic animals whose offspring constitutively lack a given lncRNA. These aspects of DNA-targeted approaches confer several advantages compared with RNA-targeted approaches.

DNA-targeted approaches can facilitate expression profiling of a lncRNA if its locus is replaced with a reporter gene. This strategy was used in mice for 18 different lncRNA gene loci (84). For example, the lncRNA Peril was found to be expressed in discrete regions of the mouse brain and spinal cord, while Mdgt expression in mice appears restricted to the testes, brain, thymus, and colon (84). This methodology may also detect lncRNAs that are found only in a minority of cells. For example, the same study revealed that mouse Pantr2 (also known as linc-Brn1b) is expressed in select brain regions throughout cortical development and is selectively expressed in upper layers of the cortex in adulthood (84). This reporter strategy, therefore, provides sufficient resolution to assess lncRNA expression, not only between tissues, but also within a tissue.

A spectrum of phenotypes has been reported following the ablation of a lncRNA by DNA-targeted approaches. For example, mice lacking Peril or Mdgt die perinatally with varying degrees of penetrance, demonstrating an essential, life-supporting role for both genes (84). Surviving Peril- and Mdgt-deficient mice are developmentally stunted, having smaller body sizes and reduced body weight compared with WT animals (84). A similar developmental phenotype was observed in mice lacking the lncRNA Pint (84). lncRNA ablation can also result in more subtle phenotypes. For example, deletion of Srat in mice improves obesity-related measures when animals are fed a high-fat diet (85), and deletion of Hotair in mice results in skeletal abnormalities of the vertebrae and wrist (86). Thus, DNA-targeted approaches have demonstrated that lncRNAs may affect not only dramatic phenotypes but also more nuanced ones.

DNA-targeted approaches and genomic imprinting. Animal models of lncRNA function using DNA-targeted approaches have been instrumental in investigating the phenotypic effects of genomic imprinting, whereby gene expression is derived from a parent-specific allele, while transcription from the other allele is epigenetically repressed. Both viability and body weight are influenced by imprinted lncRNA expression. For example, deletion of maternal...
Meg-3 (also known as Gtl2) in mice results in perinatal death, an effect not observed following paternal deletion (87). Similarly, ablation of maternal H19 in mice results in greater offspring body mass compared with that of offspring with paternal deletion (88, 89). Additionally, female mice lacking a single Tsix allele, a lncRNA implicated in X chromosome inactivation, produce fewer surviving offspring than do males lacking the same gene (90, 91). Finally, deletion of the X-linked Tsix in male mice results in smaller testes, reduced fear-related behaviors, and enhanced short-term memory (92). Collectively, these studies demonstrate diverse in vivo functions for imprinted lncRNAs.

Animal models of lncRNA function have also been instrumental when investigating the complementary molecular mechanisms underlying genomic imprinting (93). For example, deletion of paternal Kcnq1ot1 (also known as KvDMRI) in mice results in offspring with reduced body mass and de-repression of proximal genes, an effect not observed after maternal deletion (94). These results were further refined following a more targeted ablation strategy in mice, in which a premature termination sequence was inserted downstream of the Kcnq1ot1 promoter (95). This alteration also resulted in gene de-repression, demonstrating that Kcnq1ot1 transcription is necessary for gene silencing (95). Finally, complementary molecular studies have demonstrated that mouse Kcnq1ot1 interacts with chromatin and epigenetic modifiers with tissue specificity (96). This complementary use of animal models to investigate lncRNA function highlights the utility of animal models when exploring lncRNA function in vivo.

Limitations of DNA-targeted approaches. DNA-targeted approaches also have important limitations. Oftentimes, a large region encompassing the majority, if not all, of a lncRNA gene is removed, although smaller domain-specific (97) and promoter-specific (90, 94, 95) deletions may be possible. Insertion of a premature termination sequence in the gene body may also ablate lncRNA function (98). These strategies differ from those used for protein-coding genes, whereby a single nucleotide deletion or insertion is often sufficient to abolish a protein product. These larger genomic alterations may introduce unintended and confounding consequences by removing regulatory elements within the deleted region to affect the expression of neighboring genes (83, 99). For example, during embryonic stem cell differentiation in mice, overexpression of the lncRNA Haunt diminishes the expression of the neighboring HOXA gene (100). Conversely, enhancer elements within the Haunt locus facilitate HOXA expression (100). These opposing influences within a locus may complicate efforts to alter a lncRNA locus without also perturbing regulatory elements embedded within that locus. Because of the extensive number of regulatory elements within both the human (101) and mouse (102) genomes, these concerns may be particularly acute and represent the norm rather than the exception. Rescue strategies in which the disrupted gene is experimentally reintroduced may control for these potentially confounding effects. However, this experimental design can only apply to trans-acting lncRNAs for which the integration site of the transgene is independent of its function. These limitations are important caveats to consider when interpreting data generated by DNA-targeted approaches.

Another important consideration is that different DNA-targeted approaches may result in different phenotypes. In mice, replacing the lncRNA gene Fendrr with a reporter gene results in lung defects and perinatal death (84), a phenotype in agreement with clinical studies demonstrating that deletions within this locus in humans results in abnormal lung development and neonatal death (103). In contrast, a second study demonstrated that insertion of a premature transcriptional termination sequence within the same mouse locus results in perinatal death, body wall abnormalities, and heart malfunction (80). These contrasting phenotypes exemplify how different DNA-targeting strategies may produce inconsistent phenotypes and highlight the fact that caution should be exercised in making premature functional conclusions based on a single approach. Notably, this second study (80) also reported an attempted RNA-targeted strategy, in which a constitutively expressed antisense oligonucleotide resulted in a 60% reduction of Fendrr expression levels in mice that lacked any abnormal phenotype. Thus, three different strategies to manipulate lncRNA expression produced three different phenotypes, highlighting the challenges encountered when designing and interpreting lncRNA studies.

Interpretation of an absent phenotype. Loss of a lncRNA can result in no discernible phenotype. Despite being highly conserved throughout mammalian evolution, ablation of the brain-expressed Linc0046 (also known as Visc-2) in mice does not result in any overt anatomical or behavioral phenotype (104). Similarly, three independent mouse strains lacking Malat1, which is highly expressed in the brain and liver, appear to develop normally (105–107). Finally, the loss of Neat1 in mice does not result in any overt phenotype, except for the loss of mammalian-specific nuclear subregions termed paraspeckles, where Neat1 is usually localized (108). Because paraspeckles are thought to reflect higher-order compartmentalization within the nucleus that is necessary for intricate regulation of mammalian gene expression, the absence of any obvious phenotype following the loss of Neat1 prompted a critical reassessment of their function (109). Paraspeckles are induced following exposure to infectious diseases or cellular stressors, and it is possible that a phenotype in mice lacking Neat1 may only be apparent following an environmental manipulation, such as viral infection or exposure to microbes, that would normally induce paraspeckles (108).

This possibility extends to all studies using genetic ablation in animal models. Gene-environment interactions may unmask unrecognized phenotypes that are important in understanding complex diseases such as psychiatric disorders with known environmental risk factors (110, 111). The mechanism of risk conferral for many of these environmental factors, such as exposure to environmental pathogens, maternal stress during pregnancy, or maternal substance abuse during pregnancy, may be investigated by using animal models in a controlled laboratory setting to reveal latent phenotypes and resolve gene-environment interactions (Figure 3C) (112–116).

Moreover, highly exploratory research involving lncRNAs of unknown function and/or expression profile may constitute a “fishing expedition,” where only a fraction of the potential phenotypes are explored. A possible phenotype could span a large spectrum consisting of biochemical, anatomical, physiological, or behavioral effects, and thus the lack of a reported phenotype may simply represent a limited selection of unaffected pheno-
types. Compensatory effects may further mitigate the appearance of a phenotype.

**Alternative approaches.** Overexpression of a lncRNA is an alternative strategy to assess its phenotypic effects. This approach has been successfully applied to animal models using protein-coding genes to study neuropsychiatric disorders that may result from a failure to maintain homeostasis after a gain or loss of gene expression (117). When applied to lncRNAs, overexpression experiments may produce a phenotype opposite that occurring after lncRNA ablation, indicating bidirectional effects of lncRNA expression levels. Alternatively, they may induce a phenotype that is absent or seemingly unrelated to those observed following lncRNA ablation. However, this approach is only applicable for trans-acting lncRNAs, whose function is independent of their site of genomic integration.

**Genome-editing technology.** Major technological advances in genome editing (118) open the possibility of altering the genomes not only of conventional model organisms, but also of less widely used animal models that have historically been less amenable to alterations. TALENs or clustered regularly interspaced palindromic repeat (CRISPR/Cas9) methods have been widely used to excise or invert regions of a lncRNA locus in zebrafish (119, 120), mice (121, 122), and rats (123). These technologies may be extended to other animal models, adopted for specialized areas of investigation, such as the pig used to model cystic fibrosis (124), the songbird used to model language acquisition (125), and the nonhuman primate used to model the neurodegenerative disorder Huntington’s disease (126, 127). However, caveats and concerns similar to those previously discussed for standard DNA-targeted approaches still apply, given the comparatively large segments of the genome that may need to be altered in order to affect lncRNA expression and function.

**Conclusions**

Animal models are promising tools for aiding in the discovery of novel lncRNAs and investigating the phenotypic significance of these molecules. However, lncRNAs inherently differ from their protein-coding counterparts, and hence their investigation requires overcoming a host of important new challenges and addressing new considerations with respect to experimental design and data interpretation. Despite these caveats, it has been possible to convincingly demonstrate that lncRNA dysfunction in animal models results in diverse phenotypes, ranging from lethality to subtle dysmophia. So far, only a fraction of lncRNAs have been assessed, and this substantial gap in knowledge highlights a pressing need to progress beyond the initial cataloging of lncRNAs and to understand their impact in vivo. We anticipate that these initial findings represent a promising beginning to the diverse functions of lncRNAs in vivo and in understanding their relevance to disease.

**Acknowledgments**

We thank Andrew Holmes, Caroline Siebald, and members of the Goff laboratory for their helpful comments during the preparation of this manuscript.

**Address correspondence to:** Loyal A. Goff, Institute of Genetic Medicine, Miller Research Building, Room 449, 733 N. Broadway Ave., Baltimore, Maryland 21205, USA. Phone: 443.287.0251; E-mail: loyalgoff@jhmi.edu.


70. Sauvageau M, et al. Multiple knockout mouse models reveal lincRNAs are required for life and review.


