

Strategies for genetic inactivation of long noncoding RNAs in zebrafish

PERRINE LAVALOU, HELENE ECKERT, LOUISE DAMY, FLORIAN CONSTANTY, SARA MAJELLO, ANGELO BITETTI, ANTOINE GRAINDORGE, and ALENA SHKUMATAVA

Institut Curie, PSL Research University, CNRS UMR3215, INSERM U934, 75005 Paris, France

ABSTRACT

The number of annotated long noncoding RNAs (lncRNAs) continues to grow; however, their functional characterization in model organisms has been hampered by the lack of reliable genetic inactivation strategies. While partial or full deletions of lncRNA loci disrupt lncRNA expression, they do not permit the formal association of a phenotype with the encoded transcript. Here, we examined several alternative strategies for generating lncRNA null alleles in zebrafish and found that they often resulted in unpredicted changes to lncRNA expression. Removal of the transcription start sites (TSSs) of lncRNA genes resulted in hypomorphic mutants, due to the usage of either constitutive or tissue-specific alternative TSSs. Deletions of short, highly conserved lncRNA regions can also lead to overexpression of truncated transcripts. In contrast, knock-in of a polyadenylation signal enabled complete inactivation of *malat1*, the most abundant vertebrate lncRNA. In summary, lncRNA null alleles require extensive *in vivo* validation, and we propose insertion of transcription termination sequences as the most reliable approach to generate lncRNA-deficient zebrafish.

Keywords: CRISPR-Cas9; hypomorph; long noncoding RNAs; poly(A) signal; zebrafish

INTRODUCTION

Thousands of lncRNAs have been identified in multiple vertebrate species (Necsulea et al. 2014; Hezroni et al. 2015), but their biological functions remain mostly unknown. To study lncRNAs *in vivo*, genetic mutants have been generated in model animals, primarily using a mouse model (Leighton et al. 1995; Marahrens et al. 1997; Ripoche et al. 1997; Sado et al. 2001; Sleutels et al. 2002; Bond et al. 2009; Eissmann et al. 2012; Nakagawa et al. 2012, 2014; Zhang et al. 2012; Grote et al. 2013; Li et al. 2013; Sauvageau et al. 2013; Han et al. 2014, 2018; Goff and Rinn 2015; Lai et al. 2015; Amândio et al. 2016; Anderson et al. 2016; Ip et al. 2016; Kotzin et al. 2016; Isoda et al. 2017; Kleaveland et al. 2018), but have also more recently been reported in zebrafish (Kok et al. 2015; Hosono et al. 2017; Bitetti et al. 2018; Goudarzi et al. 2019).

Genetic inactivation of lncRNAs is less straightforward than for coding genes, where deletion of an exon or a point mutation in the open reading frame (ORF) often leads to stop codons or frame-shift mutations and subsequent loss of function. Several complementary strategies have been implemented to achieve genetic loss of lncRNA func-

tion, including full or partial deletion of the lncRNA locus, deletion and subsequent replacement of the lncRNA locus by a reporter gene (Nakagawa et al. 2012; Sauvageau et al. 2013), deletion of the lncRNA transcription start site (TSS) and upstream regulatory regions (Fitzpatrick et al. 2002; Zhang et al. 2012) and sequence inversions (Fig. 1; Bitetti et al. 2018). Although commonly used, these lncRNA inactivation strategies have several caveats and limitations. Full deletions of lncRNA loci, which often span several kilobases, or lncRNA replacement by a reporter gene are invasive and might lead to phenotypes that are caused by removal of regulatory DNA motifs. Deletions of lncRNA TSS and upstream promoter regions may result in usage of alternative TSSs or cryptic promoters and/or impact the expression of neighboring genes. A less invasive and more accurate approach is to inactivate lncRNAs by integrating a premature polyadenylation [poly(A)] cassette. This strategy has been successfully implemented in several recent mouse lncRNA studies (Fig. 1; Bond et al. 2009; Grote et al. 2013; Anderson et al. 2016; Ballarino et al. 2018). Whereas lncRNA locus deletion and partial lncRNA gene inversion strategies have been applied in zebrafish to genetically inactivate lncRNAs (Kok et al. 2015; Hosono et al. 2017; Bitetti et al. 2018; Goudarzi

Corresponding author: alena.shkumatava@curie.fr

Article is online at <http://www.rnajournal.org/cgi/doi/10.1261/rna.069484.118>. Freely available online through the RNA Open Access option.

© 2019 Lavalou et al. This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at <http://creativecommons.org/licenses/by/4.0/>.

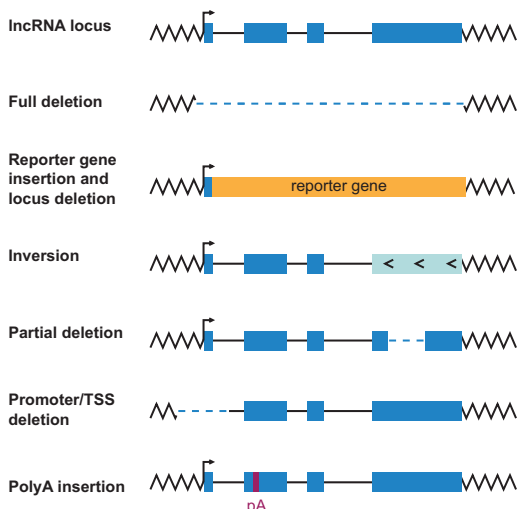


FIGURE 1. Strategies for genetic inactivation of lncRNAs in animals.

et al. 2019), analyses of complementary lncRNA silencing approaches including the minimally invasive insertion of the poly(A) sequences have not yet been carried out.

Here, we examined the efficiency of several strategies for CRISPR-Cas9-mediated inactivation of lncRNAs in zebrafish. Careful evaluation of lncRNA zebrafish mutants demonstrated that caution is required when analyzing each individual mutant allele. When genetically manipulating lncRNA loci, we found that usage of constitutive or tissue-specific alternative TSSs, overexpression or destabilization of truncated lncRNA transcripts commonly take place *in vivo*, minimizing or confounding the effect of the intended genetic intervention. In contrast, using our minimally invasive knock-in of a premature polyadenylation signal into the *malat1* locus diminished *malat1* transcripts to undetectable levels, effectively establishing a *malat1* null allele in zebrafish.

RESULTS

Deletion of the conserved region of the lncRNA *cyrano* leads to overexpression of the truncated transcript

A small fraction of zebrafish lncRNAs are conserved in mammals, representing a promising set of candidates for functional interrogation (Ulitsky et al. 2011; Hezroni et al. 2015). The conserved regions of lncRNAs are usually relatively short, ranging between 50–300 nucleotides (nt) (Ulitsky et al. 2011; Hezroni et al. 2015) and can be efficiently targeted for CRISPR-Cas9-mediated deletions in zebrafish, offering a minimally invasive strategy for functional inactivation (Fig. 1). To examine the effect of this strategy on lncRNA expression, we chose the deeply conserved lncRNA *cyrano* (Ulitsky et al. 2011) for genetic inter-

rogations in zebrafish. We generated a ~280 base pair (bp) deletion of the most conserved region of the 5.5 kb sequence, hereafter referred as *cyrano*^{ΔCR} (Fig. 2A,B; Ulitsky et al. 2011). Interestingly, we detected elevated levels of the residual truncated transcript in homozygous *cyrano*^{ΔCR} zebrafish embryos and across *cyrano*^{ΔCR} adult tissues apart from the brain (Fig. 2C,D; Supplemental Fig. 1A). These results suggest that removal of a relatively small region of a lncRNA may have an unexpected effect on the transcript levels, potentially leading to its unintended overexpression.

TSS deletion of the *cyrano* locus results in hypomorphic zebrafish mutants

Next, we tested if deleting the sequences surrounding and containing lncRNA TSS elements is a reliable alternative strategy for zebrafish lncRNA genetic inactivation. To this end, we generated a minimally invasive *cyrano*^{ΔTSS} mutant allele by removing sequences containing the *cyrano* TSS (0 to +84) (Fig. 2E). Although *cyrano* transcript levels were reduced in *cyrano*^{ΔTSS} fish, the transcript was still robustly detectable by RNA blot analysis and qRT-PCR, resulting in a hypomorphic *cyrano*^{ΔTSS} mutant (Fig. 2F,G). The 5' RACE (rapid amplification of cDNA ends) analysis demonstrated that in the absence of the two main TSSs usually used in WT animals, an alternative upstream TSS maintains *cyrano* expression in *cyrano*^{ΔTSS} mutant zebrafish (Supplemental Fig. 1B–D).

Notably, neither the *cyrano*^{ΔCR} mutant, with removal of the highly conserved miR-7 site (Ulitsky et al. 2011), nor the *cyrano*^{ΔTSS} mutant fish exhibited obvious morphological defects. This observation is consistent with recent zebrafish and mouse studies (Kleaveland et al. 2018; Goudarzi et al. 2019) and is in contrast to previous studies that used a morpholino-based knockdown approach to inactivate *cyrano* (Ulitsky et al. 2011; Sarangdhar et al. 2018).

lncRNA TSS removal leads to tissue-specific alternative TSS usage, maintaining lncRNA expression

To test if the usage of alternative TSSs is a prevalent cellular mechanism to maintain lncRNA gene expression, we examined the effect of TSS deletions on additional lncRNAs in zebrafish. We generated a lnc-*sox4a*^{ΔTSS} mutant allele by removing ~200 bp surrounding the lnc-*sox4a* TSS (–43 to +157) (Fig. 3A,B). lnc-*sox4a* (chr19:29,161,676–29,270,573; Zv9/danRer7) (Ulitsky et al. 2011) is highly expressed in the zebrafish ovary and was successfully abolished in lnc-*sox4a*^{ΔTSS} embryos and across lnc-*sox4a*^{ΔTSS} adult tissues (Fig. 3C,D). However, lnc-*sox4a* was robustly expressed in the adult lnc-*sox4a*^{ΔTSS} brain at levels comparable to WT (Fig. 3D). The 5' RACE analysis confirmed that a tissue-specific alternative TSS, located in an intron 70 kb

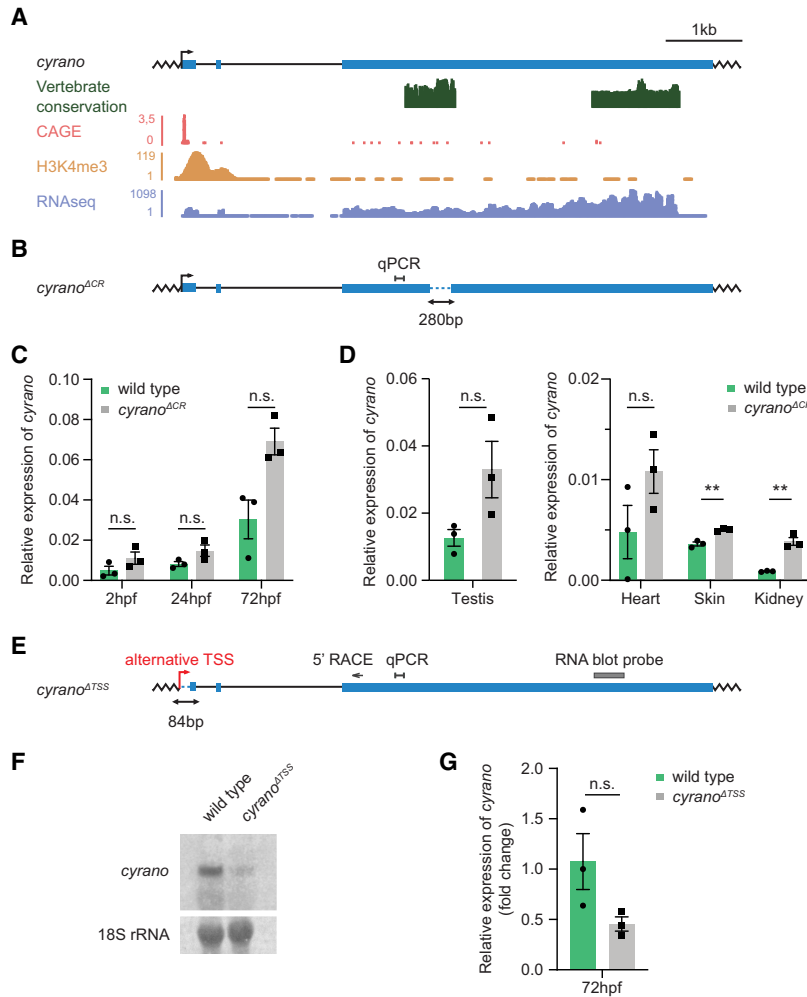


FIGURE 2. Genetic perturbations of the lncRNA *cyrano* in zebrafish result in overexpression and hypomorphic alleles. (A) Gene architecture of the lncRNA *cyrano*. Shown are the corresponding CAGE (Nepal et al. 2013; Haberle et al. 2014), H3K4me3 ChIP-Seq (Ulitsky et al. 2011), and RNA-seq tracks from wild-type (WT) zebrafish. Vertebrate conservation plots based on the eight-genome alignment indicate the location of conserved sequences. (B) The *cyrano*^{ΔCR} mutant allele showing the deletion of the most conserved region of the transcript (dotted, blue line) in zebrafish. Position of the qRT-PCR product is indicated. (C) *cyrano* expression in WT and homozygous *cyrano*^{ΔCR} embryos detected by qRT-PCR at 2 h postfertilization (hpf), 24 and 72 hpf. (D) *cyrano* expression across WT and homozygous *cyrano*^{ΔCR} adult tissues detected by qRT-PCR. (E) The *cyrano*^{ΔTSS} zebrafish allele showing deletion of the sequence around the TSS (dotted, blue line). Indicated are positions of the 5' RACE primer, qPCR product, RNA blot probe and alternative TSS. (F) *cyrano* expression in 72 hpf WT and homozygous *cyrano*^{ΔTSS} embryos detected by an RNA blot. 18S rRNA was used as a reference gene. (G) *cyrano* expression in 72 hpf WT and homozygous *cyrano*^{ΔTSS} embryos detected by qRT-PCR. *eef1a111* was used as a reference gene in all qRT-PCR experiments. Each dot represents an individual biological replicate. Data are presented as mean ± S.E.M.; (*) *P* < 0.05, n.s., not significant, unpaired t-tests.

downstream from the main TSS (Fig. 3B; Supplemental Fig. 2A,B), was used only in the *lnc-sox4a*^{ΔTSS} animals and maintained lncRNA expression specifically in the adult brain (Fig. 3D). While homozygous *lnc-sox4a*^{ΔTSS} fish were viable and fertile, our alternative strategy to eliminate *lnc-sox4a* expression by deleting the last exon failed to generate homozygous fish (Supplemental Fig. 2C,D).

We generated an additional lncRNA mutant by removing ~390 bp surrounding the *lnc-pou2af1* TSS (-74 to +315) (Fig. 4A,B). Similar to the *lnc-sox4a*^{ΔTSS} allele, the level of *lnc-pou2af1* (chr15:16770170-16773603; Zv9/danRer7) was abolished in *lnc-pou2af1*^{ΔTSS} embryos and in a subset of tested *lnc-pou2af1*^{ΔTSS} adult tissues (Fig. 4C; Supplemental Fig. 3A). However, in skin, kidney, intestine and testis, expression of *lnc-pou2af1* was robustly detected in *lnc-pou2af1*^{ΔTSS} fish (Fig. 4D,E). The 5' RACE analysis showed that several alternative TSSs, located ~1 kb upstream of the main TSS, were used in the *lnc-pou2af1*^{ΔTSS} animals in a tissue-specific manner (Fig. 4B; Supplemental Fig. 3B,C). Expression of *lnc-pou2af1* from alternative TSSs generated new tissue-specific lncRNA exons at the 5' of the transcript (Fig. 4B; Supplemental Fig. 3C,D).

Together, our data showed that in the absence of the main TSS, alternative TSSs can be used in a tissue-specific manner, generating hypomorphic mutants, and minimizing the effect of the intended gene inactivation.

Insertion of a polyadenylation signal resulted in a *malat1* null allele in zebrafish

Given the evidence that usage of alternative TSSs may be a common cellular mechanism to confer lncRNA expression, we tested if knock-in of a poly(A) signal into a lncRNA locus can be applied in zebrafish as a minimally invasive alternative to generate lncRNA null alleles. This approach has been successfully used to inactivate lncRNAs in mice (Grote et al. 2013; Anderson et al. 2016; Isoda et al. 2017; Ballarino et al. 2018).

The *malat1* locus produces one of the most abundant lncRNAs in vertebrate genomes (Ulitsky et al. 2011; Hezroni et al. 2015). Because *malat1* is a mono-exonic lncRNA of ~7.5 kb and its locus contains multiple TSSs and clustered enhancers forming a so-called super-enhancer (Pérez-Rico et al. 2017), any deletion strategy of the locus, including TSS removal, has a strong potential

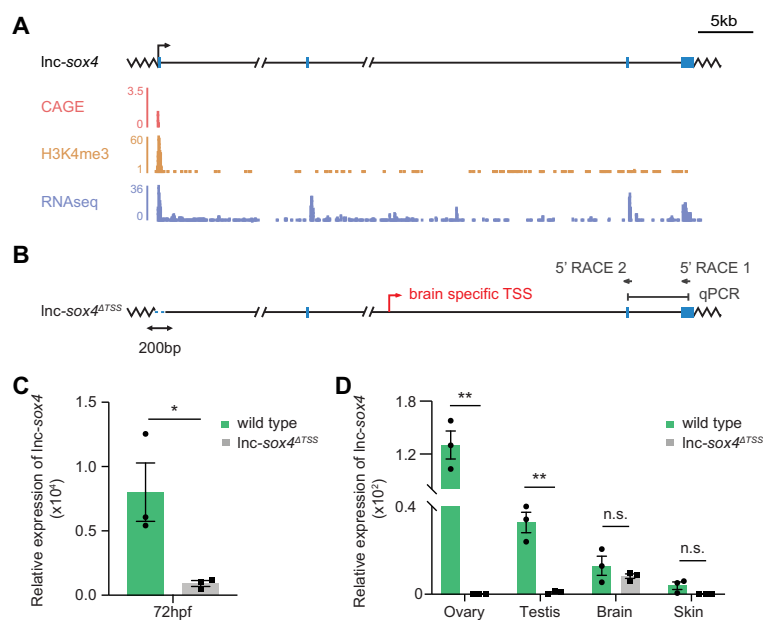


FIGURE 3. Presence of a tissue-specific alternative TSS leads to a brain-specific rescue of *Inc-sox4a* expression. (A) The *Inc-sox4a* locus in zebrafish (chr19:29,161,676–29,270,573). Shown are the corresponding CAGE (Nepal et al. 2013; Haberle et al. 2014), H3K4me3 ChIP-Seq (Ulitsky et al. 2011), and RNA-seq tracks from WT zebrafish. (B) The *Inc-sox4a*^{ΔTSS} mutant allele showing deletion of the sequence around the TSS (dotted, blue line). Indicated are positions of the 5' RACE primers, qPCR primers, and alternative TSS. (C) *Inc-sox4a* expression in 72 h post-fertilization (hpf) WT and homozygous *Inc-sox4a*^{ΔTSS} embryos detected by qRT-PCR. (D) *Inc-sox4a* expression across adult WT and homozygous *Inc-sox4a*^{ΔTSS} zebrafish tissues detected by qRT-PCR. *eef1a11* was used as a reference gene in all qRT-PCR experiments. Each dot represents an individual biological replicate. Data are presented as mean ± S.E.M.; (*) $P < 0.05$, (**) $P < 0.01$, n.s., not significant, unpaired t-tests.

to affect *cis* regulatory elements (Fig. 5A). Therefore, we applied our improved protocol for the efficient targeted knock-in to insert a 131 bp SV40 poly(A) signal into the *malat1* locus in zebrafish (Fig. 5B; see Materials and Methods; Supplemental Fig. 4A). The targeted knock-in of the poly(A) sequence completely abolished *malat1* expression in zebrafish embryos and in all examined adult tissues (Fig. 5C,D; Supplemental Fig. 4B,C). Despite effective inactivation of *malat1*, *malat1*^{poly(A)} zebrafish were viable and fertile and displayed no obvious morphological defects. The lack of overall morphological abnormalities is consistent with previously reported *Malat1*^{-/-} mice (Eissmann et al. 2012; Nakagawa et al. 2012; Zhang et al. 2012) and is in contrast to morpholino-based *malat1* inactivation in zebrafish (Wu et al. 2018). Taken together, compared to lncRNA deletion strategies, poly(A) signal insertion was the most efficient and least invasive approach in zebrafish.

DISCUSSION

The identification of lncRNAs in model vertebrates, their comparative genomics analyses and recent progress in genome editing technologies has led to the generation of

multiple mutant lncRNA alleles. Because common strategies for genetic inactivation of lncRNAs often do not allow distinguishing between functions mediated by the lncRNA transcript and those mediated by overlapping DNA regulatory motifs, the generation and interpretation of lncRNA null alleles can be challenging. Here, we compared zebrafish lncRNA mutant alleles generated using several alternative and commonly applied CRISPR-Cas9 strategies for lncRNA inactivation.

We demonstrated that relatively small deletions of conserved regions of lncRNAs, which represent attractive target sequences to eliminate or diminish lncRNA functions (Bitetti et al. 2018; Kleaveland et al. 2018), might result in unexpected changes in lncRNA levels, such as overexpression of the remaining transcript, as demonstrated for *cyrano*. One possibility is that deletion of the conserved region of *cyrano*, which removed a highly conserved and extensively paired site to miR-7 (Ulitsky et al. 2011), stabilized the *cyrano* transcript in zebrafish. Alternatively, deletion of this region of *cyrano* in zebrafish might

have caused transcriptional up-regulation. For example, if deletion of this region abrogated *cyrano* function, cells might have boosted transcription of the locus in an attempt to restore *cyrano* activity. Deletion of the conserved region of mouse *cyrano* does not lead to increased lncRNA levels (Kleaveland et al. 2018), which suggests that *cyrano* regulation has diverged between fish and mammals. A better understanding of *cyrano* regulation and function will help identify the source of this ectopic effect on the remaining lncRNA transcript observed in fish and how this effect might complicate interpretation of the deletion results.

Moreover, we showed that the removal of TSS and upstream regulatory regions, a commonly used approach considered to be straightforward to interpret, can result in the presence of either constitutive or tissue-specific alternative TSSs that preclude efficient inactivation of lncRNAs and result in hypomorph mutant animals. Although not shown in this study, usage of temporal-specific alternative TSSs might also contribute to the maintenance of lncRNA expression at specific developmental stages, complicating the analysis and interpretation of TSS mutant alleles in animal models. Interestingly, a recent study reported that a 326 bp deletion removing *cyrano*'s TSS leads to loss of the lncRNA expression (Goudarzi et al. 2019). The

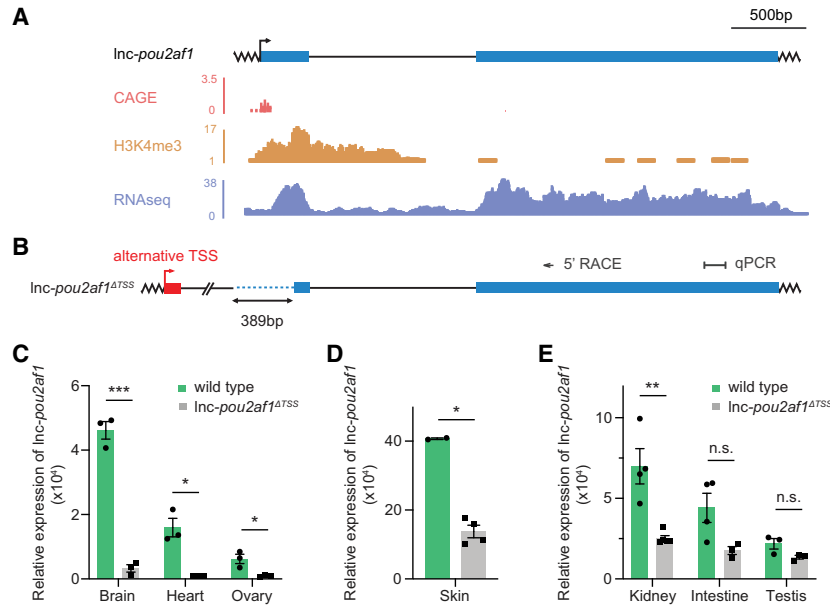


FIGURE 4. Usage of tissue-specific alternative TSSs maintains *lnc-pou2af1* expression in a subset of adult tissues. (A) The *lnc-pou2af1* locus in zebrafish. Shown are the corresponding CAGE (Nepal et al. 2013; Haberle et al. 2014), H3K4me3 ChIP-Seq (Ulitsky et al. 2011), and RNA-seq tracks from WT zebrafish. (B) The *lnc-pou2af1*^{ΔTSS} mutant allele showing deletion of the sequence around the TSS (dotted, blue line). Indicated are positions of the 5' RACE and qPCR primers and alternative TSS. Red box represents a new exon generated from the alternative TSS. (C–E) *lnc-pou2af1* expression across a subset of adult WT and homozygous *lnc-pou2af1*^{ΔTSS} zebrafish tissues detected by qRT-PCR. *eef1a111* was used as a reference gene in all qRT-PCR experiments. Each dot represents an individual biological replicate. Data are presented as mean ± S.E.M.; (*) $P < 0.05$, (**) $P < 0.01$, (***) $P < 0.001$, n.s., not significant, unpaired *t*-tests.

difference observed between the *cyrano*^{ΔTSS} alleles may be a consequence of the larger deletion used by Goudarzi et al. potentially leading to a more effective down-regulation of *cyrano*. In addition, the choice of the lncRNA detection method as well as the developmental timing of detection are important. Our data show that in TSS deletion alleles, lncRNA expression is often abolished at early embryonic stages and robustly reestablished later during development by tissue-specific alternative TSSs. These collective observations underscore the necessity to carefully validate TSS deletion alleles.

Importantly, our improved protocol for efficient targeted knock-in in zebrafish enabled examination of the effect of a poly(A) signal insertion into the most abundant and enhancer-dense lncRNA locus. We demonstrate that this minimally invasive genome editing strategy, previously shown to be successful for lncRNA inactivation in mice (Grote et al. 2013; Anderson et al. 2016; Isoda et al. 2017; Ballarino et al. 2018), is a highly effective strategy in zebrafish. Given the ease of our knock-in approach, which combines the use of a single-strand oligo as a template for homologous recombination and inhibition of nonhomologous end joining, we anticipate that the insertion of a poly(A) sequence will become a widespread strategy for generating lncRNA mutant alleles in zebrafish. Furthermore, the

knock-in strategy can be used for genetic tagging of lncRNAs with self-cleaving ribozymes, which has been demonstrated to perturb lncRNA expression in mouse embryonic stem cells (Tuck et al. 2018) but has not been tested yet in model organisms.

Taken together, evaluation of several independent lncRNA mutant alleles in zebrafish indicates that a combination of complementary lncRNA inactivation approaches and their careful analyses are required for robust and accurate lncRNA functional interrogation.

MATERIALS AND METHODS

Generation of lncRNA mutant alleles in zebrafish

All lncRNA mutant alleles were generated using CRISPR/Cas9-mediated genome editing. To generate *lnc-sox4a*^{ΔTSS}, *lnc-pou2af1*^{ΔTSS}, *cyrano*^{ΔTSS}, and *cyrano*^{ΔCR} alleles, two sgRNAs (9 ng each, Supplemental Table 1) and 150 ng in vitro transcribed Cas9 mRNA were coinjected into the one-cell stage AB zebrafish embryos (Hwang et al. 2013). To generate *lnc-sox4a*^{Δ3'exon} allele, two sgRNAs (100 ng each, Supplemental Table 1) and Cas9 protein (50 ng/μL, a gift of the Concordet Lab, Muséum d'Histoire Naturelle, Paris) were coinjected into the one-cell stage AB zebrafish embryos (Hwang et al. 2013). sgRNAs and Cas9 mRNA were generated as described previously (Hwang et al. 2013), using the codon-optimized plasmid JDS246 for the Cas9 mRNA synthesis (Addgene #43861), purified with RNeasyMini Kit (Qiagen). Genomic DNA was extracted as described previously (Bitetti et al. 2018) and used for genotyping by PCR, DNA sequencing and mapping of genetic amplification product. The genotyping primers are listed in Supplemental Table 2.

All zebrafish were bred and maintained at Institut Curie, Paris. Animal care and use for this study were performed in accordance with the recommendations of the European Community (2010/63/UE) for the care and use of laboratory animals. Experimental procedures were specifically approved by the ethics committee of Institut Curie CEEA-IC #118 (project CEEA-IC 2017-017) in compliance with the international guidelines. Zebrafish were staged using standard procedures (Kimmel et al. 1995).

Generation of the *malat1*^{poly(A)} allele by CRISPR/Cas9-mediated homologous recombination in zebrafish

The CRISPR/Cas9-mediated knock-in protocol was optimized as described in Supplemental Figure 4A. Zebrafish *malat1*^{poly(A)} mutant was generated by insertion of a single SV40 poly(A) signal

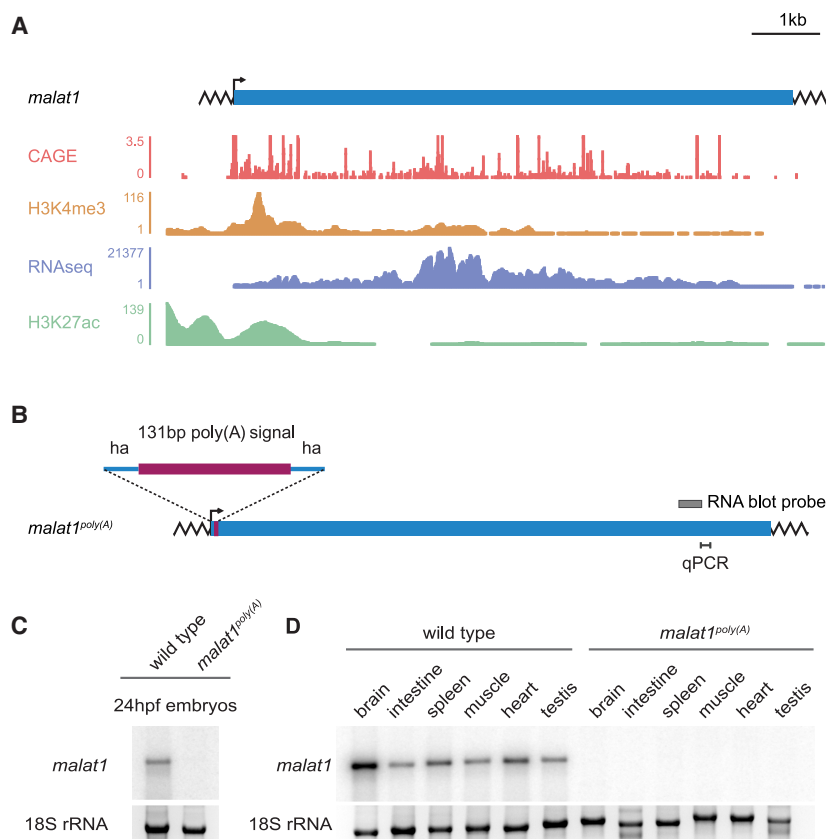


FIGURE 5. Effective inactivation of the lncRNA *malat1* in zebrafish by insertion of a premature polyadenylation signal. (A) The *malat1* locus in zebrafish. Shown are the corresponding CAGE (Nepal et al. 2013; Haberle et al. 2014), H3K4me3 ChIP-Seq (Ulitsky et al. 2011), RNA-seq, and H3K27ac ChIP-Seq (Pérez-Rico et al. 2017) tracks from WT zebrafish. (B) Generation of the *malat1^{poly(A)}* allele by targeted knock-in of the poly(A) signal. The hybridization site of the RNA blot probe is indicated as a gray box. ha, homology arms. (C) *malat1* expression in WT and homozygous *malat1^{poly(A)}* embryos detected by an RNA blot. 18S rRNA was used as a loading reference. hpf, hours postfertilization. (D) *malat1* expression across WT and homozygous *malat1^{poly(A)}* adult zebrafish tissues detected by RNA blot. 18S rRNA was used as a loading reference.

(131 bp) into the *malat1* locus. Briefly, one-cell stage embryos were injected with a single guide RNA (100 ng, Supplemental Table 1), Cas9 protein (50 ng/μL, a gift of the Concordet laboratory, Muséum d'Histoire Naturelle, Paris), a morpholino against *xrcc4* to suppress NHEJ (nonhomologous end joining) (3 ng/μL, Gene Tools LLC, Supplemental Table 1), and a 191 nt single-strand DNA oligo with 30 bp homology arms flanking both sides of the SV40 poly(A) sequence (200 ng, designed and manufactured by Ultramer IDT, Supplemental Table 1). Genomic DNA was extracted as described previously (Bitetti et al. 2018), and poly(A) insertion was detected by PCR using primers listed in Supplemental Table 2, DNA sequencing and mapping of genetic amplification product.

qRT-PCR

Total RNA was isolated from zebrafish embryos and adult tissues by TRIzol extraction (Invitrogen) followed by DNase treat-

ment (TURBO DNA-free Ambion). For individual replicates, RNA isolated from 30–100 embryos or tissues from one to six adult fish was used. cDNA was produced with SuperScript IV reverse transcriptase (Invitrogen) and amplified with PowerUp SYBR Green PCR Master Mix (ThermoFisher Scientific) using primers listed in Supplemental Table 3. For each biological replicate, qRT-PCRs were performed in technical triplicate. The *eef1α111* (*eukaryotic translation elongation factor 1 alpha 1, like 1*) was used as a reference gene (McCurley and Callard 2008).

RNA blots

Total RNA was isolated using TRIzol (Invitrogen), separated on 1% agarose gels containing 0.8% formaldehyde, and transferred to nylon membrane (Nytran SPC, GE Healthcare) by capillary action. Blots were hybridized with α-UTP ³²P-labeled RNA probes at 68°C in ULTRAhyb buffer (Ambion) as recommended by the manufacturer. RNA probe template was amplified from zebrafish brain cDNA by PCR using the primers listed in Supplemental Table 3 (the sequence of the T7 promoter is underlined) and in vitro transcribed (RNA Maxiscript, Ambion) in the presence of α-UTP ³²P. For each replicate, RNA isolated from 30–100 embryos or tissues from three to six adult fish was used. The gel blots and hybridizations in Figure 5C were performed in biological triplicates. The hybridizations in Figures 2F and 5D were performed once.

RNA ligase-mediated and oligo-capping rapid amplification of cDNA ends (5' RACE)

TSS usage was determined by rapid amplification of cDNA ends (RACE) according to manufacturer's instruction (GeneRacer kit, Life Technology). Gene specific primers listed in Supplemental Table 3 were used to amplify lncRNA 5' RACE products through PCR and nested PCR, subcloned into the PCR BLUNT II TOPO vector (Invitrogen), and transformed in the NEB TOP10 cells. A minimum of 12 colonies were sequenced, and the sequences were aligned to the corresponding lncRNA genomic locus.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We thank all members of the Shkumatava laboratory and Ines Drinnenberg for useful discussions. This work was supported by grants from the European Research Council (FLAME-337440), ATIP-Avenir, La Fondation Bettencourt Schueller, ANR-11-LABX-0044_DEEP, and ANR-10-IDEX-0001-02, as well as PSL and La Ligue Nationale Contre Le Cancer doctoral fellowships to P.L.

Author contributions: P.L. developed the protocol for the targeted knock-in in zebrafish and contributed to the design, generation, and analysis of the *Inc-sox4a*^{ATSS}, *Inc-sox4a*^{A3^{exon}}, and *malat1*^{poly(A)} alleles. H.E. contributed to the design, generation, and analysis of the *Inc-pou2af1*^{ATSS} and the maintenance and analyses of lncRNA alleles. L.D and F.C. contributed to lncRNA expression analyses and the maintenance of lncRNA alleles. S. M. contributed to the design, generation, and analyses of the *cyrano* alleles. A.B. contributed to the design, generation, and analyses of the *malat1*^{poly(A)} allele. A.G. contributed to the design and generation of the *cyrano* alleles. P.L. and A.S. wrote the final version of the manuscript. A.S. conceived and supervised the study.

Received November 2, 2018; accepted April 27, 2019.

REFERENCES

- Amândio AR, Necsculea A, Joye E, Mascrez B, Duboule D. 2016. *Hotair* is dispensible for mouse development. *PLoS Genet* **12**: e1006232. doi:10.1371/journal.pgen.1006232
- Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R, Olson EN. 2016. Transcription of the non-coding RNA upperhand controls *Hand2* expression and heart development. *Nature* **539**: 433–436. doi:10.1038/nature20128
- Ballarino M, Cipriano A, Tita R, Santini T, Desideri F, Morlando M, Colantoni A, Carrieri C, Nicoletti C, Musarò A, et al. 2018. Deficiency in the nuclear long noncoding RNA *Charme* causes myogenic defects and heart remodeling in mice. *EMBO J* **37**: e99697. doi:10.15252/emj.201899697
- Bitetti A, Mallory AC, Golini E, Carrieri C, Carreño Gutiérrez H, Perlas E, Pérez-Rico YA, Tocchini-Valentini GP, Enright AJ, Norton WHJ, et al. 2018. MicroRNA degradation by a conserved target RNA regulates animal behavior. *Nat Struct Mol Biol* **25**: 244–251. doi:10.1038/s41594-018-0032-x
- Bond AM, Vangompel MJ, Sametsky EA, Clark MF, Savage JC, Disterhoft JF, Kohtz JD. 2009. Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat Neurosci* **12**: 1020–1027. doi:10.1038/nn.2371
- Eissmann M, Gutschner T, Hämmerle M, Günther S, Caudron-Herger M, Gross M, Schirmacher P, Rippe K, Braun T, Zörnig M, et al. 2012. Loss of the abundant nuclear non-coding RNA *MALAT1* is compatible with life and development. *RNA Biol* **9**: 1076–1087. doi:10.4161/ma.21089
- Fitzpatrick GV, Soloway PD, Higgins MJ. 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of *KvDMR1*. *Nat Genet* **32**: 426–431. doi:10.1038/ng988
- Goff LA, Rinn JL. 2015. Linking RNA biology to lncRNAs. *Genome Res* **25**: 1456–1465. doi:10.1101/gr.191122.115
- Goudarzi M, Berg K, Pieper LM, Schier AF. 2019. Individual long non-coding RNAs have no overt functions in zebrafish embryogenesis, viability and fertility. *Elife* **8**: e40815. doi:10.7554/eLife.40815
- Grote P, Wittler L, Hendrix D, Koch F, Währisch S, Beisaw A, Macura K, Bläss G, Kellis M, Werber M, et al. 2013. The tissue-specific lncRNA *Fendrr* is an essential regulator of heart and body wall development in the mouse. *Dev Cell* **24**: 206–214. doi:10.1016/j.devcel.2012.12.012
- Haberle V, Li N, Hadzhiev Y, Plessy C, Previti C, Nepal C, Gehrig J, Dong X, Akalin A, Suzuki AM, et al. 2014. Two independent transcription initiation codes overlap on vertebrate core promoters. *Nature* **507**: 381–385. doi:10.1038/nature12974
- Han P, Li W, Lin CH, Yang J, Shang C, Nuernberg ST, Jin KK, Xu W, Lin CY, Lin CJ, et al. 2014. A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* **514**: 102–106. doi:10.1038/nature13596
- Han X, Luo S, Peng G, Lu JY, Cui G, Liu L, Yan P, Yin Y, Liu W, Wang R, et al. 2018. Mouse knockout models reveal largely dispensable but context-dependent functions of lncRNAs during development. *J Mol Cell Biol* **10**: 175–178. doi:10.1093/jmcb/mjy003
- Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP, Ulitsky I. 2015. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep* **11**: 1110–1122. doi:10.1016/j.celrep.2015.04.023
- Hosono Y, Niknafs YS, Prensner JR, Iyer MK, Dhanasekaran SM, Mehra R, Pitchiaya S, Tien J, Escara-Wilke J, Poliakov A, et al. 2017. Oncogenic role of THOR, a conserved cancer/testis long non-coding RNA. *Cell* **171**: 1559–1572 e1520. doi:10.1016/j.cell.2017.11.040
- Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, Yeh JR. 2013. Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One* **8**: e68708. doi:10.1371/journal.pone.0068708
- Ip JY, Sone M, Nashiki C, Pan Q, Kitaichi K, Yanaka K, Abe T, Takao K, Miyakawa T, Blencowe BJ, et al. 2016. Gomafu lncRNA knockout mice exhibit mild hyperactivity with enhanced responsiveness to the psychostimulant methamphetamine. *Sci Rep* **6**: 27204. doi:10.1038/srep27204
- Isoda T, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP, et al. 2017. Non-coding transcription instructs chromatin folding and compartmentalization to dictate enhancer-promoter communication and T cell fate. *Cell* **171**: 103–119 e118. doi:10.1016/j.cell.2017.09.001
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**: 253–310. doi:10.1002/aja.1002030302
- Kleaveland B, Shi CY, Stefano J, Bartel DP. 2018. A network of non-coding regulatory RNAs acts in the mammalian brain. *Cell* **174**: 350–362 e317. doi:10.1016/j.cell.2018.05.022
- Kok FO, Shin M, Ni CW, Gupta A, Grosse AS, van Impel A, Kirchmaier BC, Peterson-Maduro J, Kourkoulis G, Male I, et al. 2015. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell* **32**: 97–108. doi:10.1016/j.devcel.2014.11.018
- Kotzin JJ, Spencer SP, McCright SJ, Kumar DBU, Collet MA, Mowel WK, Elliott EN, Uyar A, Makiya MA, Dunagin MC, et al. 2016. The long non-coding RNA *Morrbid* regulates *Bim* and short-lived myeloid cell lifespan. *Nature* **537**: 239–243. doi:10.1038/nature19346
- Lai KM, Gong G, Atanasio A, Rojas J, Quispe J, Posca J, White D, Huang M, Fedorova D, Grant C, et al. 2015. Diverse phenotypes and specific transcription patterns in twenty mouse lines with ablated lincRNAs. *PLoS One* **10**: e0125522. doi:10.1371/journal.pone.0125522
- Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM. 1995. Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* **375**: 34–39. doi:10.1038/375034a0

- Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, Carlson JC, Lin M, Fang F, Gupta RA, et al. 2013. Targeted disruption of *Hotair* leads to homeotic transformation and gene derepression. *Cell Rep* **5**: 3–12. doi:10.1016/j.celrep.2013.09.003
- Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R. 1997. *Xist*-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev* **11**: 156–166. doi:10.1101/gad.11.2.156
- McCurley AT, Callard GV. 2008. Characterization of housekeeping genes in zebrafish: male-female differences and effects of tissue type, developmental stage and chemical treatment. *BMC Mol Biol* **9**: 102. doi:10.1186/1471-2199-9-102
- Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, Prasanth KV. 2012. *Malat1* is not an essential component of nuclear speckles in mice. *RNA* **18**: 1487–1499. doi:10.1261/rna.033217.112
- Nakagawa S, Shimada M, Yanaka K, Mito M, Arai T, Takahashi E, Fujita Y, Fujimori T, Standaert L, Marine JC, et al. 2014. The lncRNA *Neat1* is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. *Development* **141**: 4618–4627. doi:10.1242/dev.110544
- Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grützner F, Kaessmann H. 2014. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature* **505**: 635–640. doi:10.1038/nature12943
- Nepal C, Hadzhiev Y, Previti C, Haberle V, Li N, Takahashi H, Suzuki AM, Sheng Y, Abdelhamid RF, Anand S, et al. 2013. Dynamic regulation of the transcription initiation landscape at single nucleotide resolution during vertebrate embryogenesis. *Genome Res* **23**: 1938–1950. doi:10.1101/gr.153692.112
- Pérez-Rico YA, Boeva V, Mallory AC, Bitetti A, Majello S, Barillot E, Shkumatava A. 2017. Comparative analyses of super-enhancers reveal conserved elements in vertebrate genomes. *Genome Res* **27**: 259–268. doi:10.1101/gr.203679.115
- Ripoche MA, Kress C, Poirier F, Dandolo L. 1997. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* **11**: 1596–1604. doi:10.1101/gad.11.12.1596
- Sado T, Wang Z, Sasaki H, Li E. 2001. Regulation of imprinted X-chromosome inactivation in mice by *Tsix*. *Development* **128**: 1275–1286.
- Sarangdhar MA, Chaubey D, Srikakulam N, Pillai B. 2018. Parentally inherited long non-coding RNA *Cyrano* is involved in zebrafish neurodevelopment. *Nucleic Acids Res* **46**: 9726–9735. doi:10.1093/nar/gky628
- Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, et al. 2013. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife* **2**: e01749. doi:10.7554/eLife.01749
- Sleutels F, Zwart R, Barlow DP. 2002. The non-coding air RNA is required for silencing autosomal imprinted genes. *Nature* **415**: 810–813. doi:10.1038/415810a
- Tuck AC, Natarajan KN, Rice GM, Borawski J, Mohn F, Rankova A, Flemr M, Wenger A, Nutiu R, Teichmann S, et al. 2018. Distinctive features of lincRNA gene expression suggest widespread RNA-independent functions. *Life Sci Alliance* **1**: e201800124. doi:10.26508/lsa.201800124
- Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. 2011. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**: 1537–1550. doi:10.1016/j.cell.2011.11.055
- Wu M, Zhang S, Chen X, Xu H, Li X. 2018. Expression and function of lncRNA *MALAT-1* in the embryonic development of zebrafish. *Gene* **680**: 65–71. doi:10.1016/j.gene.2018.09.037
- Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, et al. 2012. The lncRNA *Malat1* is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep* **2**: 111–123. doi:10.1016/j.celrep.2012.06.003