

Critical roles of long noncoding RNAs in *Drosophila* spermatogenesis

Kejia Wen,^{1,6} Lijuan Yang,^{1,2,6} Tuanlin Xiong,^{1,6} Chao Di,¹ Danhui Ma,¹ Menghua Wu,¹ Zhaoyu Xue,¹ Xuedi Zhang,¹ Li Long,¹ Weimin Zhang,¹ Jiaying Zhang,¹ Xiaolin Bi,³ Junbiao Dai,^{1,4} Qiangfeng Zhang,^{1,4,5} Zhi John Lu,^{1,4} and Guanjun Gao^{1,4}

¹MOE Key Laboratory of Bioinformatics, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China; ²College of Animal Science, Tarim University, Xinjiang 843300, China; ³Institute of Cancer Stem Cell, Dalian Medical University, Dalian 116044, China; ⁴Center for Synthetic and Systems Biology, Tsinghua University, Beijing 100084, China; ⁵Beijing Advanced Innovation Center for Structural Biology, Tsinghua University, Beijing 100084, China

Long noncoding RNAs (lncRNAs), a recently discovered class of cellular RNAs, play important roles in the regulation of many cellular developmental processes. Although lncRNAs have been systematically identified in various systems, most of them have not been functionally characterized *in vivo* in animal models. In this study, we identified 128 testis-specific *Drosophila* lncRNAs and knocked out 105 of them using an optimized three-component CRISPR/Cas9 system. Among the lncRNA knockouts, 33 (31%) exhibited a partial or complete loss of male fertility, accompanied by visual developmental defects in late spermatogenesis. In addition, six knockouts were fully or partially rescued by transgenes in a *trans* configuration, indicating that those lncRNAs primarily work *in trans*. Furthermore, gene expression profiles for five lncRNA mutants revealed that testis-specific lncRNAs regulate global gene expression, orchestrating late male germ cell differentiation. Compared with coding genes, the testis-specific lncRNAs evolved much faster. Moreover, lncRNAs of greater functional importance exhibited higher sequence conservation, suggesting that they are under constant evolutionary selection. Collectively, our results reveal critical functions of rapidly evolving testis-specific lncRNAs in late *Drosophila* spermatogenesis.

[Supplemental material is available for this article.]

Thousands of long noncoding RNAs (lncRNAs; >200 nt) have been identified by genome-wide transcriptome analyses (Lee 2012; Iyer et al. 2015). In comparison to mRNA, lncRNAs are characterized by poorer conservation, lower expression levels, and more variable expression between tissues (Guttman et al. 2009; Ponting et al. 2009; Ulitsky and Bartel 2013). Deep studies have revealed the functionality of a dozen lncRNAs that play roles in biological processes such as dosage compensation, imprinting, apoptosis, immunity, cancer, and development (Pauli et al. 2011; Batista and Chang 2013; Carpenter et al. 2013). lncRNAs, such as *Xist* and *Air*, regulate transcription of neighboring genes, and thus function via a *cis*-acting mechanism (Lee and Bartolomei 2013; Mercer and Mattick 2013). lncRNAs have also been proposed to regulate gene expression *in trans*; examples include the TP53-induced lncRNAs *Dlx6os1* and *Hotair* (Feng et al. 2006; Huarte et al. 2010; Chu et al. 2011). However, despite the huge number of lncRNAs in the genome, the biological importance of most lncRNAs remains unknown.

In multiple species, the largest repertoire of lncRNAs is expressed in the testis, as revealed by genome-wide transcriptome analyses (Djebali et al. 2012; Nam and Bartel 2012; Brown et al. 2014; Morris and Mattick 2014). Moreover, lncRNAs participate with protein-coding genes in evolutionarily conserved coexpression networks during spermatogenesis (Necsulea et al. 2014), the process by which male germline stem cells (GSCs) divide and differentiate into mature sperm in sexual organisms. To date,

however, the functional significance of lncRNAs in spermatogenesis is unknown, with a few exceptions such as polymorphic derived intron-containing (*Pldi*) RNA (Heinen et al. 2009). To gain a comprehensive picture of lncRNA functionality in spermatogenesis, it is necessary to develop an efficient and large-scale gene knockout method for investigating the functions of lncRNAs in intact organisms.

The revolutionary CRISPR/Cas9-based genome editing system consists of two components—the Cas9 nuclease that cleaves DNA and the guide RNA that confers cleavage specificity (Garneau et al. 2010; Jinek et al. 2012; Wiedenheft et al. 2012). The RNA-guided DNA cleavage causes formation of double-strand breaks (DSBs), which leads to deletions or mutations via nonhomologous end joining (NHEJ) repair (Cong et al. 2013; Mali et al. 2013). However, frameshift mutations, nonsense mutations, or small deletions caused by NHEJ are unlikely to disrupt the functions of lncRNAs. Therefore, based on a previous study by Gratz et al. (2014), we developed a three-component Cas9 microinjection system consisting of the *Cas9* mRNA, a gene-specific gRNA, and a homologous recombination (HR) donor plasmid, simplifying the generation of knockout constructs; this system is applicable to deletion of almost any genomic locus. The efficiency of this system enabled us to successfully and rapidly delete 105 testis-specific *Drosophila* lncRNAs, 33 of which were revealed to play critical roles in the regulation of late spermatogenesis.

These authors contributed equally to this work.

Corresponding author: gaogu@mail.tsinghua.edu.cn

Article published online before print. Article, supplemental material, and publication date are at <http://www.genome.org/cgi/doi/10.1101/gr.199547.115>.

© 2016 Wen et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genome.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

We also investigated the origin and evolution of the functionality of these 105 lncRNAs.

Results

Systematic identification of testis-associated lncRNAs in *Drosophila*

To globally characterize the biological significance of lncRNAs in spermatogenesis, we developed a stepwise selection pipeline to identify testis-specific lncRNAs in *Drosophila*. To this end, we first analyzed published gene expression data and identified 120 lncRNAs specifically and/or highly expressed in fly testis (Brown et al. 2014). To comprehensively identify all lncRNAs expressed in fly, we used a previously developed computational algorithm (Lu et al. 2011; Gerstein et al. 2014; Hu et al. 2015) to predict 121 novel intergenic lncRNAs with no overlap with protein-coding genes; the predictions were based on RNA-seq data and RNA structure information (Fig. 1A; Supplemental Fig. S1A). From the resultant 241 lncRNAs, we ultimately identified the 128 testis-specific lncRNAs from a testis-specific expression screen and RNA in situ hybridization analysis. These lncRNAs were located on three different chromosomes (Fig. 1A). Meanwhile, our RT-PCR and lncRNA in situ hybridization results indicated that they were highly or specifically expressed in the *Drosophila* testis (Fig. 1B,C; Supplemental Fig. S1B,C). The majority of testis-specific lncRNAs were strictly expressed in the meiotic and post-meiotic region of the testis (Fig. 1C; Supplemental Fig. S2). Thus, we identified 128 testis-specific lncRNAs for further targeted mutational analysis.

The development of a three-component CRISPR/Cas9 system streamlines lncRNA gene knockout in *Drosophila*

To comprehensively analyze lncRNA functions in vivo, we need an efficient and robust experimental system to inactivate lncRNA genes by genomic deletion. Considering that the point mutations, small deletions, or inversions induced by a single gRNA are unlikely to disrupt lncRNA function unless the gRNA is targeted based on prior knowledge of the functional domains (Sauvageau et al. 2013; Yin et al. 2015), we tried to knock out an entire lncRNA gene by replacing it with an RFP marker. To this end, we developed an efficient three-component CRISPR system, including *Cas9* mRNA, gRNA, and an HR donor plasmid, to perform precisely targeted deletions via homologous recombination (Fig. 2A; Supplemental Fig. S3A).

To optimize the system for scalable genome-wide deletion analysis, we investigated three factors that might potentially affect HR efficiency, including the length of donor homology arms, the absence of NHEJ, and the distance between DSBs and homology arms. We first used the system to knock out the well-studied protein-coding gene *oskar* (Lehmann and Nüsslein-Volhard 1986). Donor plasmids containing homology arms with lengths of 1.5 or 3 kb, but not 0.3 kb, were sufficient to generate a 5-kb deletion in *oskar* (Fig. 2B; Supplemental Fig. S3B,C). Deletion in *oskar* caused defects in oocyte development and complete sterility in females, as previously reported (Lehmann and Nüsslein-Volhard 1986). To further test this strategy for large fragment deletions using 1.5-kb homology arms, we tried to knock out a 92-kb fragment in a known lncRNA gene, *iab-8*, in the genome. Surprisingly, one correct HR targeting event was obtained from

40 fertile crosses (Supplemental Fig. S3D). The *iab-8* knockout was pupa lethal.

In mammalian cells, inhibition of the NHEJ pathway can increase HR efficiency (Maruyama et al. 2015; Yu et al. 2015). To determine whether this is also the case in *Drosophila*, we compared the knockout efficiencies of five lncRNA genes in wild-type and *Lig4*^{-/-} mutant flies. The *Lig4* deletion did not dramatically increase HR frequencies in any of the tested knockouts, suggesting that inhibition of NHEJ has a negligible effect on HR efficiency in the fly (Fig. 2C). Furthermore, the distance between the double-strand break site and the homology arms had no dramatic effect on HR frequency (Fig. 2D). Knockout analyses of 18 lncRNAs indicated that HR frequencies were similar for HR donor arms adjacent to the breaks and for arms 700 bp away from the break sites.

Last, we developed a new high-throughput cloning strategy for constructing plasmids carrying homology arms in 96-well plates (Supplemental Fig. S3E). Using this system, we successfully generated deletion strains corresponding to 105 of the 128 testis-specific lncRNAs identified as described above (Supplemental Table S1). The deletions were confirmed independently by PCR and sequencing (Fig. 2E; Supplemental Fig. S4). The deleted genomic regions ranged from 200 to 11 kb (Fig. 2F; Supplemental Table S1), and the HR efficiency was as high as 47% (Fig. 2G; Supplemental Table S1). As expected, all 105 homozygous lncRNA deletion strains were viable, because these RNAs are primarily expressed in late germ cells of the *Drosophila* testis.

To determine whether our CRISPR/Cas9 knockout system induced off-target mutagenesis by introducing DSBs at unintended genomic sequences, we performed PCR amplification and sequencing analysis of all potential off-target cleavage sites in 22 lncRNA knockout mutants (see online tool, <http://crispr.mit.edu>). We did not detect mutations at any of these sites (Supplemental Fig. S5). Taken together, these results indicate that our three-component CRISPR-based knockout system is suitable for rapid large-scale functional investigation of lncRNA genes in vivo.

Some testis-specific lncRNAs are required for male fertility and late spermatogenesis

Given the testis-specific expression of the lncRNAs we identified, we examined the fertility of knockout (KO) mutant males using a standard fertility test (Sitnik et al. 2014). After virgin females mated with wild-type control or lncRNA knockout males, the numbers of progeny from each female fly were counted daily over a 15-d period. In 32 of 105 lncRNA knockouts (30%), we observed a substantial reduction in male fertility (Fig. 3A; Supplemental Fig. S6). One lncRNA KO (*CR44455/6*^{-/-}) strain had a more severe phenotype and was completely male-sterile (Fig. 3B,C). In contrast, lncRNA KO mutant females did not exhibit any obvious reduction in fertility (Fig. 3B,C). These results suggested that 30% of the testis-specific lncRNAs we identified have important functions during *Drosophila* spermatogenesis.

To further characterize KO phenotypes, we microscopically examined testis morphology and mature sperm in seminal vesicles of lncRNA KO males. Ten of the 33 deletion mutants with diminished male fertility, including *CR44455/6*^{-/-}, *lncRNA:TS23*^{-/-}, and *CR43282*^{-/-}, exhibited severe morphological defects in testis. The abnormalities included an accumulation of cotton-like white flocculus in the distal part of the testis, which corresponds to the later stages of sperm development (Fig. 4A; Supplemental Fig. S7A). The flocculus contained a large number of tightly packed small cells (Supplemental Fig. S7B). In addition, the sperm in seminal vesicles

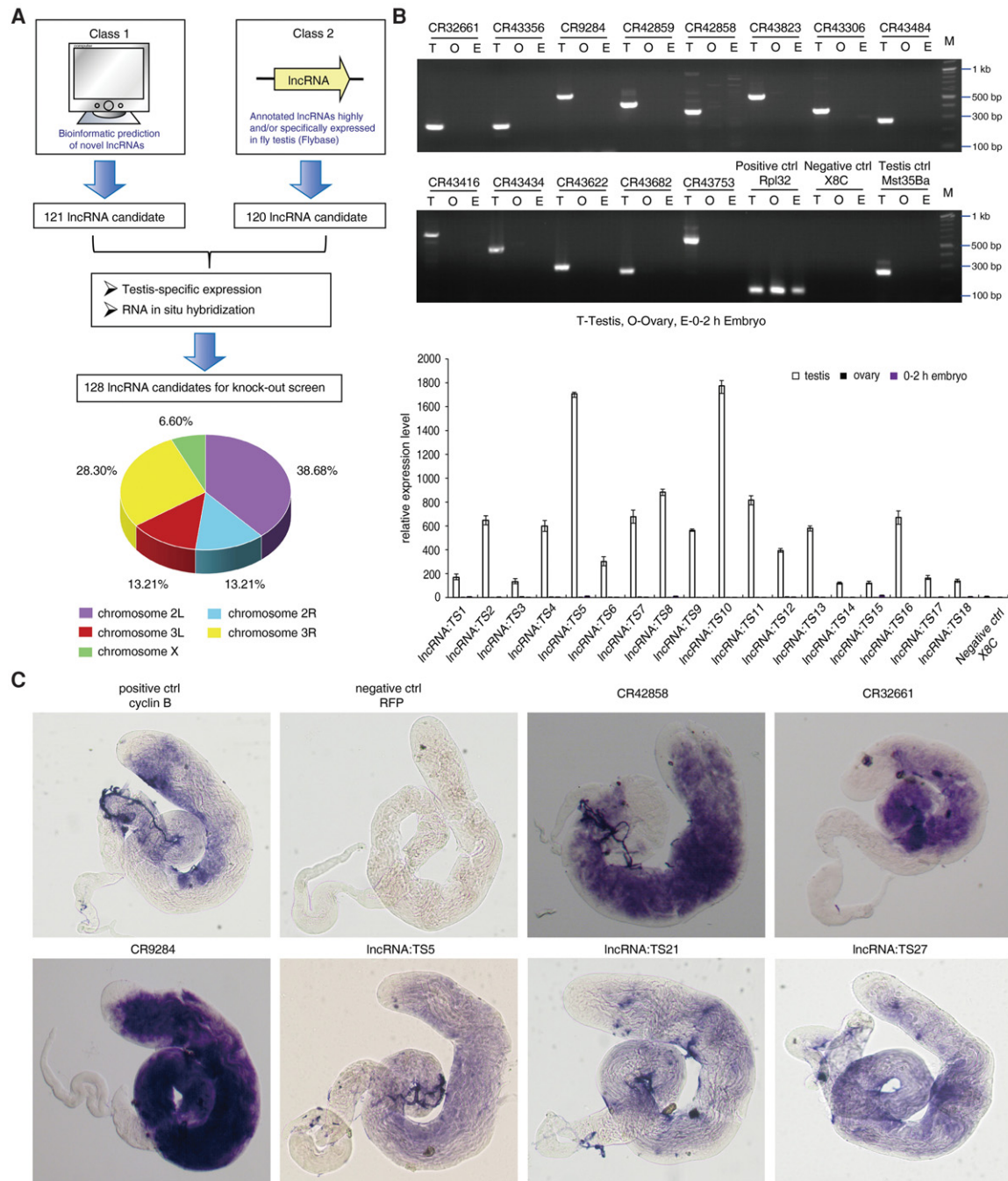


Figure 1. Systematic identification and validation of *Drosophila* lncRNAs involved in spermatogenesis. (A) Flowchart of identification and selection of testis-specific lncRNAs for the knockout study. Novel lncRNA prediction using bioinformatics and analysis of annotated lncRNAs in FlyBase were combined to build the lncRNA starting pool. Then, through a testis-specific expression screen and RNA in situ hybridization, 128 testis-specific lncRNA candidates were selected for targeted knockout. These lncRNAs were located on three different chromosomes, including the left and right arms of Chromosome 2, the left and right arms of Chromosome 3, and Chromosome X. (B) Testis-specific expression screen of predicted lncRNAs and annotated lncRNAs by quantitative RT-PCR and semiquantitative RT-PCR, respectively. *Rpl32* was used as an internal control. Values represent means \pm SEM for three biological replicates. *Mst35Ba* was used as a testis-specific control. *X8C* (a Chromosome X-linked intergenic region that has been determined to be silent for transcription) was used as a negative control to rule out contamination of RNA by genomic DNA. (C) Expression of selected lncRNAs in *Drosophila* testis, analyzed by whole-mount in situ hybridization. *Cyclin B* RNA was used as a positive control, and *RFP* RNA in nontransgenic testis (w1118) was used as a negative control.

of *CR44455/6*^{-/-} males were smaller than those in wild-type flies (Fig. 4A). Other lncRNA mutants with severe white flocculus, such as *CR42858*^{-/-} and *lncRNA:TS23*^{-/-}, contained significantly

smaller numbers of mobile mature spermatozoa in their seminal vesicles (Fig. 4A; Supplemental Fig. S7A). These observations explain the reduced male fertility of some lncRNA deletion mutants.

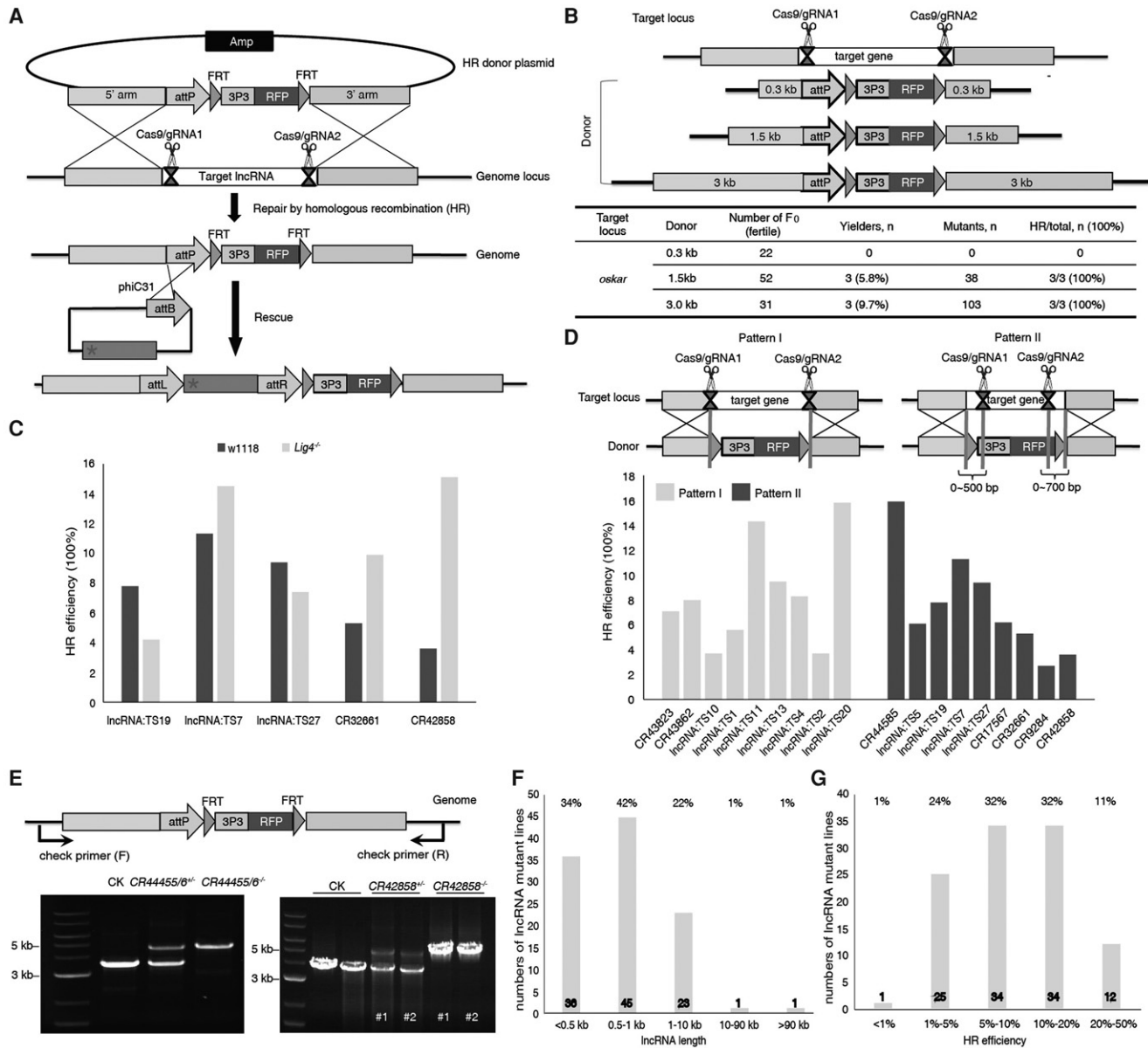


Figure 2. CRISPR/Cas9-mediated high-throughput mutagenesis of 105 lncRNAs. (A) Schematic representation of the strategy for generating lncRNA knockout and rescue lines using CRISPR. (B) Determination of the optimal HR donor length for efficient gene replacement. The plasmid containing a reporter cassette (attP-FRT-3P3-RFP) flanked by 0.3, 1.5, and 3 kb HR sequences was provided as the donor DNA. The *oskar* gene was used to test the system. (C) DNA Ligase4 deficiency had no dramatic effect on HR efficiency at different target loci. (D) Within a range of 0–700 bp, the distance between the double-strand break site and the homology arm had no dramatic effect on HR frequency. (E) Genotyping of w1118, *lncRNA*^{+/+}, and *lncRNA*^{-/-} flies to confirm lncRNA knockouts. Insertion of the reporter cassette results in a size increase in the PCR products, which is visualized only in heterozygous and homozygous flies. (F) Length distribution of the targeted lncRNAs. (G) The distribution of HR efficiency when generating the 105 lncRNA mutants using CRISPR. F₀ injected flies were crossed to w1118, and their progeny were screened for RFP-positive eyes. Crosses producing one or more transgenic progeny were considered as founder lines. The homologous recombination efficiency was calculated as the ratio between the number of founder lines and the number of F₀ crosses.

Testis-specific lncRNAs are required for nuclear condensation and morphogenesis

To further characterize any possible defect in late germ cell development in lncRNA KO testes, we examined spermatid morphology of the 105 lncRNA KO mutants by testis squash and DAPI staining. During spermatid development, the initially round nuclei synchronously elongate and condense to form long, straight, needle-shaped structures. Twenty-two of the lncRNA KO mutants

lost synchronization and exhibited defects in spermatid morphogenesis (Fig. 4B,C; Supplemental Fig. S8). The developmental defects could be classified into three major types. First, ~10% of spermatids in *lncRNA:TS1*^{-/-} and ~20% of sperm heads in *CR43484*^{-/-} testes adopted a tadpole shape, in which the sperm nucleus was concentrated at one end of the cell (Fig. 4B; Supplemental Fig. S8). This phenotype is similar to those of protamine mutations (Rathke et al. 2010). Second, *CR45542*^{-/-} and *CR44420*^{-/-} mutant testes contained some round uncondensed

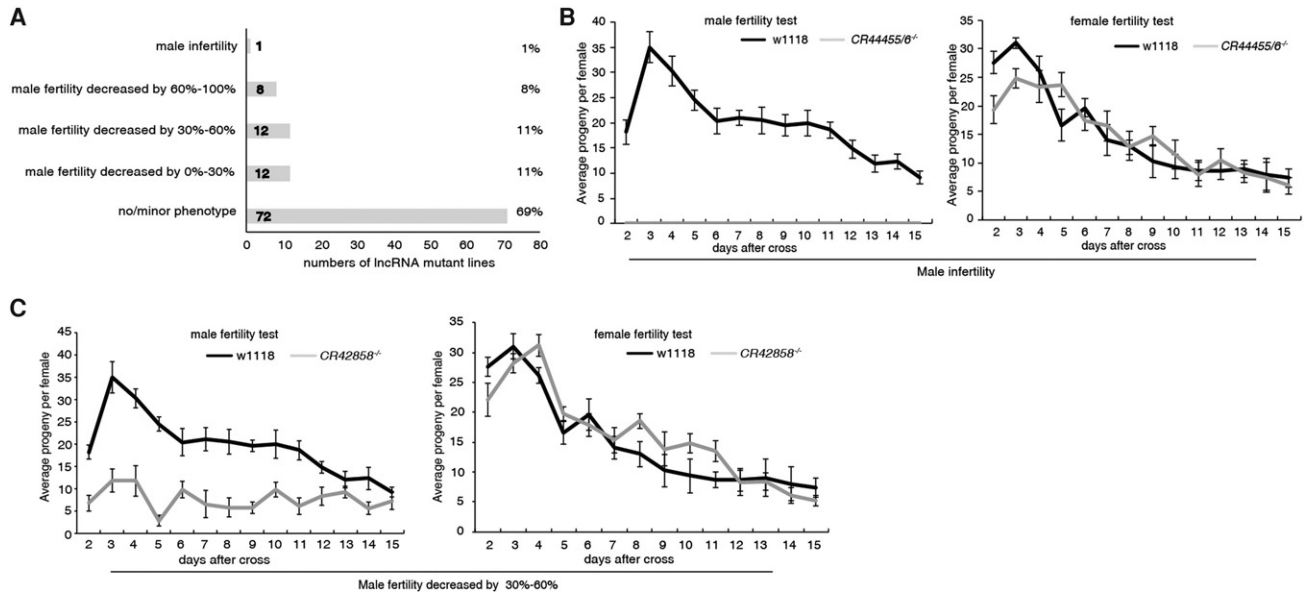


Figure 3. IncRNA mutants develop male-specific fertility defects. (A) Fertility profiles of 105 IncRNA mutants. (B) Deletion of IncRNA *CR44455/6* results in male sterility, whereas *CR44455/6*^{-/-} females were fully fertile. (C) A qualitative fertility assay was performed for both male and female mutant flies of *CR42858*^{-/-}. Deletion of *CR42858* substantially reduced fertility in males, but not in females. Values represent means ± SEM for 15 crosses each.

nuclei (~30% in *CR45542*^{-/-} and ~20% in *CR44420*^{-/-}) at the mature stage (Fig. 4B,C), indicating that these two lncRNAs are required for nuclear condensation. Third, 17 other lncRNA mutants, including *CR42858*^{-/-} and *CR43416*^{-/-}, had scattered or curled sperm (~25% in *CR42858*^{-/-} and ~30% in *CR43416*^{-/-}) (Fig. 4B,C). This phenotype is reminiscent of mutants in the testis-specific proteasome subunit *Prosa6T* (Zhong and Belote 2007) or *yuri* (Texada et al. 2008).

Next, we used HIS2AV-RFP and Protamine B-GFP to mark early spermatids and mature sperm, respectively (Rathke et al. 2007, 2010). The round uncondensed sperm of *CR44420*^{-/-}, tadpole-shape sperm of *lncRNA:TS1*^{-/-}, and bent sperm of *CR43416*^{-/-} were positive for Protamine B-GFP but not HIS2AV-RFP, confirming that they corresponded to the late elongate and/or mature stage (Fig. 4C). This result is consistent with a previous study of *Prosa6T* (Zhong and Belote 2007). Surprisingly, the mature sperm in the seminal vesicles of these lncRNA mutants did not have abnormal morphologies, suggesting that mutant sperm with abnormally condensed chromatin might not enter the seminal vesicle.

In contrast to the spermatid nuclear phenotype described above, the spermatids in *CR44455/6*^{-/-} mutant testis were smaller than those of the wild type from the meiotic stage onward (Fig. 4D); specifically, mature *CR44455/6*^{-/-} sperm were about half the size of wild-type sperm. The small nuclei in the mutant could be caused by overcondensation. Because *CR44455/6*^{-/-} mutant males were completely sterile, these small mature sperm were defective in fertilization. Taken together, these data indicate that some of the testis-specific lncRNAs play important roles in controlling proper nuclear condensation during late spermatogenesis.

Some lncRNAs regulate spermatid individualization

In spermatogenesis, after completion of meiosis, 64 cells in each spermatid cyst begin differentiating into individual sperm. Once fully elongated, spermatids undergo the process of individualization, which separates individual sperm tails and removes excess cy-

toplasm (Ma et al. 2010). During individualization, the actin-based investment cones (ICs) form and translocate mature spermatid nuclei down axonemes. To obtain insight into the nature of the spermatogenesis defects of lncRNA mutants, we further analyzed ICs by labeling testes for actin bundles, myosin VI, and DNA. In wild-type testes, ICs assemble above spermatid nuclei and coordinately move along spermatid bundles as a complex. However, the testes of 19 lncRNA mutants contained poorly aligned or lagging ICs (Fig. 4E; Supplemental Fig. S9). In the testes of the *CR43282*^{-/-}, *CR42859*^{-/-}, and *CR44371*^{-/-} mutants, the ICs were severely disorganized, and the actin cone structures were scattered (Fig. 4E; Supplemental Fig. S9). In the *CR43484*^{-/-} and *CR44420*^{-/-} mutants, the nuclei failed to remain tightly clustered and were displaced distally along the cyst, resulting in lagging ICs (Fig. 4E; Supplemental Fig. S9). These observations indicate that some of the testis-specific lncRNAs are required for developmental synchronization of the 64-cell cyst and the shaping and differentiating of spermatids during late spermatogenesis.

Testis-specific lncRNAs function *in trans* to regulate late spermatogenesis

To rule out the possibility that the phenotypes of lncRNA mutants were caused by off-target events, we first performed *in cis* rescue experiments on *CR42858*, *CR43484*, *CR44585*, *lncRNA:TS2*, *CR43416*, and *CR43862* deletion mutants by inserting the wild-type lncRNAs under the control of their endogenous promoters via PhiC31-mediated attB/attP exchange (Fig. 5A; Supplemental Fig. S10). The late spermatogenesis defects of these six lncRNA deletion mutants were rescued by *in cis* lncRNA restoration, indicating that the late spermatogenesis phenotypes were indeed caused by deletion of the lncRNAs (Fig. 5A). To determine whether these lncRNAs function *in trans*, we transgenically rescued the *CR42858*, *CR43484*, *CR44585*, *lncRNA:TS2*, *CR43416*, and *CR43862* deletion mutants by expressing the corresponding

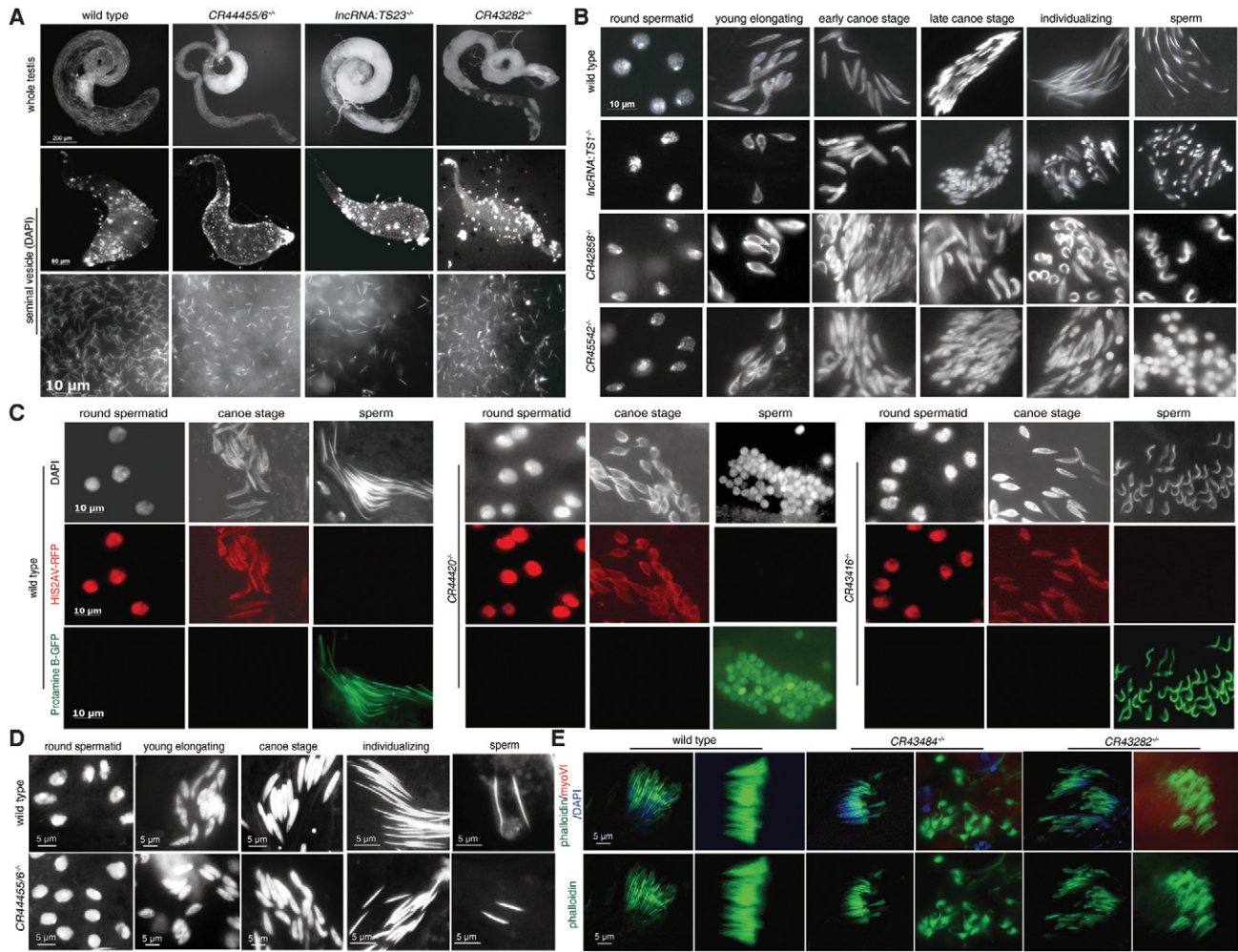


Figure 4. Deletion of lncRNAs lead to various defects in spermatogenesis. (A) lncRNA knockout mutants cause malformation and obstruction of the testis: (top rows) whole testis; (middle rows) seminal vesicle stained with DAPI; (bottom rows) sperm in seminal vesicle. Scale bars are noted. Testes of *CR44455/6^{-/-}*, *lncRNA:TS23^{-/-}*, and *CR43282^{-/-}* contained an accumulation of abnormal white flocculus, although these lncRNA mutants had a normal spherical testis shape. Seminal vesicles from *CR44455/6^{-/-}* males contained smaller sperm relative to wild type. *lncRNA:TS23^{-/-}* contained significantly reduced numbers of mature sperm in seminal vesicles, whereas the numbers in *CR43282^{-/-}* were comparable to those in the wild type. (B) lncRNA knockouts affect male germ cell development. Testes squash preparations were stained with DAPI to visualize DNA of the wild-type and lncRNA mutants. In the wild type, the initially round spermatids nuclei elongated and condensed to form long, straight, and needle-shaped mature sperm. In *lncRNA:TS1^{-/-}*, some mature-stage sperm adopted a tadpole shape in which the nucleus was concentrated at one end of the spermatid head. Deletion of *CR42858* led to scattered or curled sperm in which some nuclei did not fully condense. *CR45542^{-/-}* exhibited some round uncondensed sperm at a very mature stage. (C) Chromatin condensation defects of lncRNA mutants appear in late spermatogenesis. HIS2AV-RFP and Protamine B-GFP were used to distinguish early spermatid nuclei and mature sperm, respectively. The round uncondensed nuclei of *CR44420^{-/-}* and the bent sperm of *CR43416^{-/-}* were labeled by Protamine B-GFP but not HIS2AV-RFP, indicating that these abnormal germline cell phenotypes appeared in late elongate to mature sperm stage. (D) Spermatogenesis in wild type and *CR44455/6^{-/-}* (male infertility). Spermatids in *CR44455/6^{-/-}* were smaller than those in the wild type from the meiotic stages onward. The nuclei of mature sperm in *CR44455/6^{-/-}* were half the size of those in wild-type sperm. (E) lncRNA mutants exhibit individualization defects. Phalloidin was used to stain investment cones (ICs) in wild-type and lncRNA mutants. Wild-type testis contained ordered and associated ICs. In *CR43484^{-/-}* and *CR43282^{-/-}*, ICs were severely disorganized or lagged, and individual actin cone structures were scattered. Scale bars are noted.

lncRNAs on a different chromosome via PhiC31-mediated attB/attP exchange. The spermatogenesis defects of all six lncRNA mutants were also rescued by expression of the lncRNAs *in trans*, ruling out the possibility that the phenotypes of the deletions are caused by the disruption of the regulatory DNA elements (Fig. 5A). These results indicate that these testis-specific lncRNAs primarily function *in trans* to regulate late spermatogenesis.

Next, we investigated the function of the lncRNA *CR42858* in more detail. *CR42858* was efficiently transcribed from transgenes either at its endogenous locus (*in cis*) or a different chromosome

(*in trans*) (Fig. 5B). The abnormal testis morphology and poorly aligned IC phenotypes of *CR42858^{-/-}* were rescued by transgenes *in cis* and *in trans* (Fig. 5C). To further separate RNA from DNA sequence-dependent effects, we individually introduced *CR42858* DNA sequences without the promoter and *CR42858* promoter-driven eGFP sequences into *CR42858* mutant flies *in situ* (Supplemental Fig. S11). The phenotypic defects in *CR42858^{-/-}* were not rescued by *CR42858* DNA without the promoter or by *CR42858* promoter-driven eGFP (Supplemental Fig. S11). All these results rule out the possibility that the defective phenotypes of

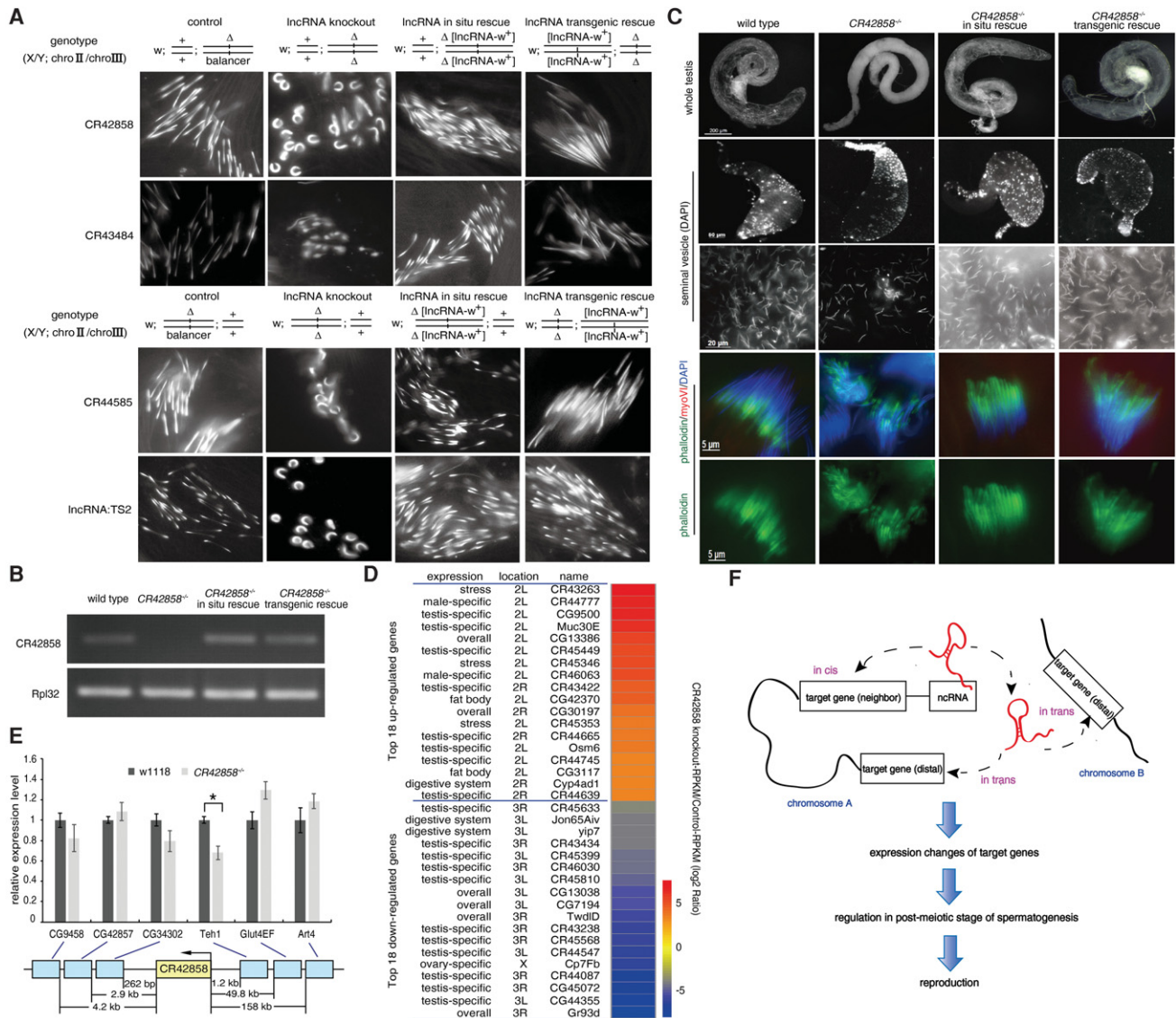


Figure 5. Testis-specific lncRNAs function *in trans* to regulate late spermatogenesis. (A) Male germ cell development defects of four lncRNA knockout mutants were rescued by restoration of the corresponding lncRNAs *in cis* or *in trans*. For *in cis* rescue, a construct bearing an lncRNA gene fragment under the control of the endogenous promoter was inserted back into the genomic locus from which the original lncRNA had been deleted. For rescue *in trans*, the lncRNA was placed on another chromosome (i.e., different from the original chromosomal locus of the lncRNA) through Phic31-mediated attB/attP exchange. The abnormal morphological phenotype of scattered and curled sperm nuclei in late spermatogenesis of *CR42858*^{-/-}, *CR43484*^{-/-}, *CR44585*^{-/-}, and *lncRNA:TS2*^{-/-} were rescued by restoration of these lncRNAs either *in cis* or *in trans*. (B) lncRNA *CR42858* was transcribed in flies rescued with *CR42858* *in cis* and *in trans*, but not in *CR42858*^{-/-}. RNA was isolated from the testes of wild-type, *CR42858*^{-/-}, and flies rescued with *CR42858* *in cis* or *in trans*, and subjected to semiquantitative RT-PCR. *Rpl32* was used as a control. (C) *In cis* or *in trans* rescue of *CR42858* recovered the abnormal morphology of the testis, lower density of sperm, and lagging and poorly aligned IC phenotypes of *CR42858*^{-/-}. (D) RNA-seq analysis of wild-type and *CR42858*^{-/-} flies. The heat map shows fold changes in expression of each of the top 18 differentially expressed genes (DEGs) in *CR42858*^{-/-} relative to those in the wild type, and the DEGs were corrected for the biological variability by combining biological replicates (NOISeqBIO method in NOISeq, version 2.14.1) (Tarazona et al. 2015). (E) Expression changes of six neighboring genes of *CR42858*^{-/-}. RNA was isolated from testes of *CR42858*^{-/-} and wild type and subjected to quantitative RT-PCR. *Rpl32* was used as an internal control. Results are represented as means ± SEM of three biological replicates. Asterisks indicate a significant difference between the samples (*P* < 0.05; *t*-test). (F) Schematic illustration of *in vivo* functions of lncRNAs in fly spermatogenesis. lncRNAs can regulate the expression of both local and distal genes in testis, and function in post-meiotic stages of spermatogenesis.

CR42858^{-/-} are caused by the deletion of regulatory DNA elements. RNA sequencing (RNA-seq) revealed that hundreds of genes were up- or down-regulated (|log₂ Ratio| ≥ 1), respectively, in *CR42858*^{-/-} testis in comparison with wild-type testis. These *CR42858*-regulated genes include protein-coding genes as well as lncRNA genes (Fig. 5D). Genes exhibiting the most dramatic ex-

pression changes in the *CR42858* mutant testis were distant from the *CR42858* locus, whereas genes located close to *CR42858*, including *Teh1*, exhibited moderate but significant down-regulation (Fig. 5E). These results further support the idea that these testis-specific lncRNAs primarily function *in trans* in late spermatogenesis (Fig. 5F).

To gain insight into the transcriptional regulation of other lncRNAs, we performed massively parallel RNA-seq of testes from wild-type and four lncRNA KO strains (Supplemental Fig. S12). Consistent with the RNA-seq results of *CR42858*^{-/-}, the expression of dozens of genes changed significantly in *CR44585*^{-/-}, *lncRNA:TS1*^{-/-}, *lncRNA:TS2*^{-/-}, and *CR45542*^{-/-} (Supplemental Fig. S13). The lncRNA-regulated genes are located distantly from the lncRNAs, and many of them are associated with reproductive or metabolic processes, suggesting that lncRNAs control the expression of genes important for germ cell development *in trans*. Further studies will be needed to understand at the molecular level how testis-specific lncRNAs control the expression of neighboring and distantly located genes in the *Drosophila* testis.

lncRNAs with defective knockout phenotypes evolve much more slowly than those without phenotypic effects

We next investigated the origin and evolution of all testis-specific lncRNAs through sequence conservation and phylogenetic analysis. By comparative genomic analysis of 12 close *Drosophila* relatives, we first dated the evolutionary origin of all surveyed lncRNAs, with or without phenotypic effects, via their phylogenetic distribution. Although different internal nodes of the phylogenetic tree exhibited some fluctuations in the proportions of functional lncRNAs (18%–57%) (Fig. 6A), we observed a constant fraction of functional ones (~35%) within each evolutionary age group (Supplemental Table S2), including old (>40 million years ago [Mya]) and young RNAs. A previous study found that ~30% of new protein-coding genes quickly become essential (Chen et al. 2010). In addition, we investigated the sequence conservation of all surveyed lncRNAs and compared the conservation of the RNAs with that of protein-coding genes. Consistent with other studies (Necsulea et al. 2014), old lncRNAs were more conserved than younger ones (Supplemental Fig. S14). Overall, in comparison to protein-coding genes (median score, 0.986), lncRNAs evolved much more rapidly (Mann-Whitney *U* test, *P*-value <10⁻²⁰). Furthermore, as expected, testis-biased coding DNA sequences (CDSs) (median score, 0.965) were under a slightly more relaxed selection constraint than CDSs overall (Mann-Whitney *U* test, *P*-value <10⁻¹⁶) (Haerty et al. 2007). Meanwhile, the group

of lncRNAs whose knockouts had defective phenotypes (median score, 0.536) evolved more slowly than lncRNAs whose knockouts did not confer defects (median score, 0.353) (Mann-Whitney *U* test, *P*-value <10⁻²⁰) (Fig. 6B). These results suggest that lncRNAs evolve more slowly once they acquire functional importance.

Discussion

Although a large number of lncRNAs have been identified in various tissues, especially in testis, their biological functions remain largely unexplored. In this study, we identified 128 testis-specific lncRNAs in *Drosophila*. Using a three-component CRISPR/Cas9-based HR system, we deleted 105 lncRNAs and found that males exhibited reduced fertility and late spermatogenesis developmental defects in 33 of the deletion mutants. Our evolutionary analysis revealed that the functional lncRNAs tend to be under stronger selection constraints. Despite the fact that several studies report testis-enriched or testis-specific transcription of lncRNAs (Nam and Bartel 2012; Necsulea et al. 2014), this is one of the first studies to demonstrate the general relevance of lncRNAs to testis function.

The three-component CRISPR/Cas9-based system facilitates generation of lncRNA knockout animal models

Recently, genome-scale loss-of-function studies have indicated that lncRNAs are key regulators of cellular processes and development (Batista and Chang 2013; Mercer and Mattick 2013; Flynn and Chang 2014). Those studies were generally conducted in cell-based *in vitro* systems; however, phenotypes at the cellular and organismal levels are frequently discrepant. For example, loss-of-function studies of *Malat1* or *Dlx6os1* in mouse revealed subtle or undetectable phenotypes (Bond et al. 2009; Eißmann et al. 2012), despite the fact that these lncRNAs appear to be important at the cellular level. Therefore, the gold standard in the field is the targeted *in vivo* silencing or deletion of specific lncRNAs (Mattick 2013).

Compared to other strategies for generating gene deletions, our optimized CRISPR system offers several advantages. First, direct injection of the three components (*Cas9* mRNA, gRNA, and HR plasmid) into *Drosophila* embryos greatly simplifies targeted

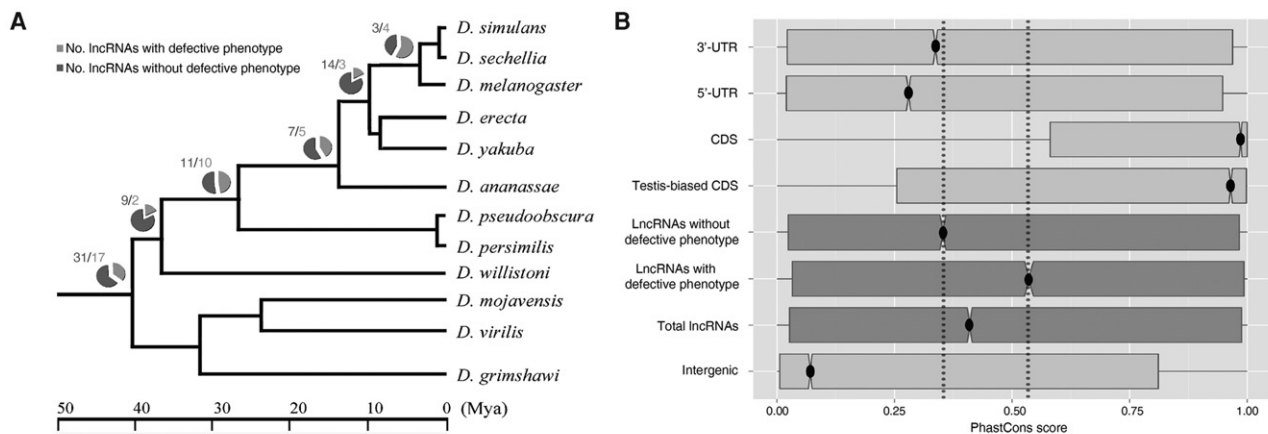


Figure 6. Evolutionary age and sequence conservation of testis-specific lncRNAs. (A) A simplified phylogenetic tree to illustrate the distribution of functional and dysfunctional lncRNAs. (B) Sequence conservation (15-way PhastCons score) for lncRNAs, with or without knockout phenotypes, and different characteristic genomic regions. PhastCons scores represent probabilities of negative selection (range between 0 and 1) at single-nucleotide resolution. The smaller the divergence of a DNA segment across species, the more likely it is that the segment belongs to a conserved element maintained by negative selection.

gene replacement. Second, our knockout method can efficiently delete genes of up to 92 kb, and such deletions lead to a complete loss of gene function (Supplemental Fig. S3D). Third, our knockout method has a low rate of off-target effects, as demonstrated by direct sequencing of potential off-targeting loci in 22 lncRNA mutants (Supplemental Fig. S5) and rescue of six lncRNA KO mutants. Consequently, of 128 lncRNAs, we were able to successfully delete 105 (82%) with an average HR frequency of 10%. Taken together, these observations indicate that our CRISPR/Cas9 system is suitable for large-scale gene deletion screens with low off-target effects in *Drosophila* and should also be applicable to other organisms.

The phenotype of testis-specific lncRNAs are mainly manifested in late *Drosophila* spermatogenesis

Recent studies show that lncRNAs are often predominantly transcribed in testis in both vertebrates (Necsulea et al. 2014) and *Drosophila* (Young et al. 2012), suggesting that these RNAs play similar functional roles in spermatogenesis across a broad range of animal taxa. However, it remains unclear whether lncRNAs are truly involved in male reproduction *in vivo*. Three mouse lncRNAs are involved in spermatogenesis *in vitro* (Zhang et al. 2010; Ni et al. 2011; Arun et al. 2012), but the aforementioned *Pldi* RNA represents the only case in which functional significance in spermatogenesis has been demonstrated in lncRNA knockout models (Heinen et al. 2009). Thus, we attempted to survey the functional roles of lncRNAs in spermatogenesis using our optimized CRISPR system.

We found that 33 lncRNA knockouts exhibited developmental defects in late spermatogenesis, resulting in low or no male fertility. The mutant phenotypes in late germ cell development indicate that these testis-specific lncRNAs play critical roles in the regulation of nuclear condensation and sperm individualization. During late spermatogenesis in both mammals and *Drosophila*, spermatids need to remodel and condense chromatin by replacing histones with protamines and also require the removal of excess cytoplasm for individualization (Rathke et al. 2010). *CR44455/6*, *CR45542*, and *CR44420* mutant testes exhibited scattered nuclei and round uncondensed nuclei, whereas *lncRNA:TS1* and *CR43484* mutant testes exhibited a crumpled nucleus phenotype similar to those of protamine mutants (Rathke et al. 2010). In regard to sperm individualization, the testes of 19 lncRNA KO mutants, including *CR42858*^{-/-} and *CR43282*^{-/-}, exhibited defects in coordinated actin cone movement, resulting in poorly aligned or lagging ICs. Similar phenotypes have been reported for the mutants in the genes encoding the testis-specific proteasome subunit *Prosa6T*, myosin VI, myosin V, and dynein (Hicks et al. 1999; Li et al. 2004; Mermall et al. 2005; Zhong and Belote 2007). It remains to be determined whether these lncRNAs are directly functional in late spermatogenesis or instead play a role in the early spermatogenesis that is only manifest in the late stage.

Like protein-coding genes, lncRNAs also exhibit redundancy of function. For example, the male-specific lncRNAs *roX1* and *roX2* paint the X Chromosome of male fly, thereby contributing to equalization of X Chromosome-linked gene expression. Flies lacking *roX1* or *roX2* separately have no phenotype, whereas simultaneous removal of both *roX1* and *roX2* causes a striking male-specific reduction in viability, indicating that these lncRNAs are functionally redundant (Meller and Rattner 2002). Similarly, some lncRNAs without phenotypes in this study may have redundant counterparts elsewhere in the genome. Meanwhile, the

lncRNAs without discernable phenotypes in this study should be further investigated using more sensitive assays, such as sperm exhaustion techniques and sperm competence tests (Yeh et al. 2012).

Testis-specific lncRNAs affect late spermatogenesis primarily by regulating gene expression *in trans*

The next obvious question is how lncRNAs affect late spermatogenesis. Our RNA-seq results revealed that lncRNA *CR42858* controls the expression of hundreds of genes, most of which are highly expressed in the testis or are testis-specific. These differentially transcribed genes consisted of both coding genes and lncRNAs, suggesting that the general role of testis-specific lncRNAs in late spermatogenesis may involve transcriptional regulation, as proposed for other functionally characterized lncRNAs, e.g., *Paupar* and *Pantr1* (Vance et al. 2014; Goff et al. 2015).

Some lncRNAs regulate the transcription of neighboring genes in a *cis*-acting manner (Lai et al. 2013; Melo et al. 2013), whereas others regulate gene expression *in trans*, e.g., the TP53-induced lncRNAs *Dlx6os1* and *Hotair* (Feng et al. 2006; Huarte et al. 2010; Chu et al. 2011). We showed that *CR42858* could regulate the expression of neighboring genes as well as many more distant genes. This alteration of transcription of both neighboring and distal genes upon deletion of an lncRNA is consistent with a recent study of mouse lncRNAs (Goff et al. 2015). To determine whether these testis-specific lncRNAs primarily function *in cis* or *in trans*, we performed rescue experiments on six lncRNA deletion mutants by inserting the rescue transgenes either in the endogenous locations (*in cis*) or in other genomic locations (*in trans*). Strikingly, the transgenes could rescue the spermatogenesis defects in the six lncRNA mutants both *in cis* and *in trans*, suggesting that these testis-specific lncRNAs affect late spermatogenesis by regulating the expression of target genes *in trans*.

Separating RNA-dependent lncRNA functions from DNA sequence-dependent effects

Three independent experiments were performed to discriminate between RNA and DNA sequence-dependent effects. First, *in situ* rescue experiments demonstrated that the defective phenotypes of *CR42858*^{-/-} were due to loss of RNA-dependent lncRNA functions rather than to loss of DNA regulatory elements on the *CR42858* promoter or *CR42858* DNA sequences (Supplemental Fig. S11). Second, the results of RNAi against nine lncRNAs with clear knockout phenotypes revealed that the *CR44585*, *CR43416*, and *CR44456* knockdown phenotypes were similar to the phenotypes of the corresponding knockouts (Supplemental Fig. S15), indicating that the phenotypes of some lncRNA KOs were indeed due to the removal of the lncRNA transcripts rather than the absence of the endogenous DNA. Third, transgenic rescue experiments on six lncRNAs showed that the spermatogenesis defects in the knockouts could be rescued by expression of the lncRNAs *in trans* (Fig. 5A; Supplemental Fig. S10). Collectively, this evidence argues that the phenotypes of these lncRNA KO mutants are more likely to be due to the loss of the lncRNA transcripts themselves than to changes in chromosomal DNA sequences. However, we cannot entirely rule out the possibility that DNA regulatory elements play a role in all lncRNA mutants. Further investigation will be required to more rigorously distinguish RNA from DNA sequence-dependent effects.

Functionality may be only accumulated within a constant proportion of lncRNAs

Although dozens of lncRNAs have been implicated in various biological processes (Pauli et al. 2011; Batista and Chang 2013; Carpenter et al. 2013), the functions of the vast majority of other putative lncRNAs are largely unexplored, and it remains unclear how many lncRNAs are functional (Moran et al. 2012; Doolittle 2013). One conservative but reliable benchmark for the functionality of biological macromolecules is their conservation over the course of evolution (Graur et al. 2013). Indeed, our results showed that in testis, lncRNAs with a defective KO phenotype were more conserved than those without such a phenotype (Fig. 6B). The conservation levels for protein-coding genes and intergenic regions were consistent with expectations defined in other studies (Necsulea et al. 2014).

Remarkably, among testis-specific lncRNAs, the proportion of functional lncRNAs was similar along the sampled evolutionary ages (Supplemental Table S2), suggesting that, as in the formation of essential genes (Chen et al. 2010), a constant proportion of young lncRNAs quickly acquire important functions. Furthermore, although our analysis was limited to testis-biased lncRNAs, we hypothesize that across the entire pool of lncRNAs, functional sequences may be likely to accumulate in a stationary tempo and maintained at a constant proportion, around 30%. More comprehensive studies combining functional and evolutionary analysis will provide further insights into this issue.

In summary, we developed an efficient CRISPR/Cas9-based gene deletion system to systematically delete 105 testis-lncRNAs in *Drosophila*, of which 31% exhibited strong phenotypes, especially in late spermatogenesis, and an equivalent proportion quickly becoming functional independent of their age. Thus, our study provides important insights into the functions and evolution of tissue-specific lncRNAs, and the mutant lncRNAs generated by this study will be a valuable resource for future studies of spermatogenesis and the functions of lncRNA.

Methods

Curation of known and novel lncRNAs in fly

We first collected the annotated ncRNAs from FlyBase r5.45 and then adopted a machine learning method to predict novel lncRNAs in *Drosophila* (Lu et al. 2011; Gerstein et al. 2014; Hu et al. 2015). In this method, we used multiple features (e.g., sequence, structure, and expression data) to train a random forest model. The model used the lncRNAs annotated in FlyBase as the training set, and then made predictions throughout the whole genome (Hu et al. 2015). Subsequently, the annotated (known) and predicted novel lncRNAs were filtered and classified based on their genomic locations. They were subtyped into antisense, intronic, ambiguous, and intergenic ncRNAs (Di et al. 2014). To remove the ambiguity, we only retained intergenic lncRNAs (i.e., lincRNAs) for further studies. To select testis-specific lncRNAs, we used expression profiles derived from ENCODE RNA-seq data (for details, see Hu et al. 2015, supplemental table). These testis-specific lncRNAs were verified by qRT-PCR and whole-mount RNA in situ hybridization as described in the Supplemental Methods.

Generation of lncRNA knockout flies

In vitro transcription of *Cas9* mRNA was performed using the Sp6 mMACHINE Kit (Ambion), according to Yu et al. (2013). In vitro transcription of the designed gRNAs was per-

formed using the RiboMAX Large Scale RNA Production Systems-T7 Kit (Promega). Purified *Cas9* mRNA, gRNA, and donor plasmid were mixed at final concentrations of 1 $\mu\text{g}/\mu\text{L}$, 50 $\text{ng}/\mu\text{L}$, and 0.8 $\mu\text{g}/\mu\text{L}$, respectively, followed by injection into w1118 embryos (Supplemental Table S3). The details of donor plasmid construction and gRNA design, Cas9/gRNA-mediated lncRNA deletion screen, *in cis* and *in trans* rescue of lncRNA knockout flies, and off-target analysis are listed in the Supplemental Methods.

Qualitative fertility assays

Fertility tests for males were always performed in batches of 15. For each lncRNA knockout mutant, one lncRNA homozygous mutant virgin male was placed in a vial with one wild-type virgin female at 27.5°C. For the next 15 d, the flies from each mating were transferred to new vials every 24 h. Upon eclosion, all progeny from each vial were counted (Sitnik et al. 2014). The average number of flies per parental pair and standard errors were calculated for each combination of genotypes. The details of testis imaging, phalloidin staining, and immunohistochemistry are listed in the Supplemental Methods.

RNA-seq analysis

RNA libraries were prepared for sequencing using standard Illumina protocols. Library products were sequenced on an Illumina HiSeq 2000 at the BGI (<http://www.genomics.cn/index>). The differentially expressed genes between two samples with biological replicates were identified using NOISeq, version 2.14.1 (Tarazona et al. 2015). See the Supplemental Methods for more details.

Evolutionary age and sequence conservation analysis of lncRNAs

For lncRNAs with successful KO mutants, we estimated their evolutionary age based on their phylogenetic distribution on a reference tree: ((((((droMel, (droSim, droSec)), (droYak, droEre)), droAna), (droPer, droPse)), droWil), ((droVir, droMoj), droGri)) (Stark et al. 2007). All lncRNAs emerging more than 40 million years ago were regarded as old lncRNAs (Chen et al. 2010), and the young lncRNAs were further divided into two or three age groups to calculate the relative proportion of functional lncRNAs and study the emergence of lncRNA functionalization. Sequence conservation was assessed by PhastCons score. See the Supplemental Methods for additional details.

Data access

The RNA-seq data sets generated in this study have been submitted to the NCBI Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra/>) under accession numbers SRX1542553 for *lncRNA:TS1* knockout, SRX1542555 for *lncRNA:TS2* knockout, SRX1512980 for *CR44585* knockout, SRX1542556 for *CR45542* knockout, SRX1542557 for *CR42858* knockout, and SRX1542554 for wild type.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (31171278, 31271542). We thank Dr. Ting Xie, Dr. Xiaohua Shen, and Dr. Yong Zhang for critical suggestions on manuscript preparation.

Author contributions: K.W. and G.G. conceived the project, designed the experiments, and wrote the manuscript; L.Y., K.W., M.W., Z.X., and X.Z. generated lncRNA mutants using the CRISPR system; L.L., W.Z., J.Z., X.B., and J.D. constructed the plasmids;

T.X. and Q.Z. performed evolutionary analyses; C.D. and Z.J.L. predicted novel lncRNAs in *Drosophila*; K.W., L.Y., D.M., and L.L. performed phenotype classifications and analyzed the data.

References

- Arun G, Akhade VS, Donakonda S, Rao MR. 2012. mrhl RNA, a long non-coding RNA, negatively regulates Wnt signaling through its protein partner Ddx5/p68 in mouse spermatogonial cells. *Mol Cell Biol* **32**: 3140–3152.
- Batista PJ, Chang HY. 2013. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**: 1298–1307.
- 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.
- Brown JB, Boley N, Eisman R, May GE, Stoiber MH, Duff MO, Booth BW, Wen J, Park S, Suzuki AM. 2014. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**: 393–399.
- Carpenter S, Aiello D, Atianand MK, Ricci EP, Gandhi P, Hall LL, Byron M, Monks B, Henry-Bezy M, Lawrence JB. 2013. A long noncoding RNA mediates both activation and repression of immune response genes. *Science* **341**: 789–792.
- Chen S, Zhang YE, Long M. 2010. New genes in *Drosophila* quickly become essential. *Science* **330**: 1682–1685.
- Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. 2011. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* **44**: 667–678.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823.
- Di C, Yuan J, Wu Y, Li J, Lin H, Hu L, Zhang T, Qi Y, Gerstein MB, Guo Y. 2014. Characterization of stress-responsive lncRNAs in *Arabidopsis thaliana* by integrating expression, epigenetic and structural features. *Plant J* **80**: 848–861.
- Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F. 2012. Landscape of transcription in human cells. *Nature* **489**: 101–108.
- Doolittle WF. 2013. Is junk DNA bunk? A critique of ENCODE. *Proc Natl Acad Sci* **110**: 5294–5300.
- Eißmann M, Gutschner T, Hämmerle M, Günther S, Caudron-Herger M, Groß M, Schirmacher P, Rippe K, Braun T, Diederichs S. 2012. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol* **9**: 1076–1087.
- Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD. 2006. The *Evf-2* noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* **20**: 1470–1484.
- Flynn RA, Chang HY. 2014. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell* **14**: 752–761.
- Garneau JE, Dupuis MÈ, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadán AH, Moineau S. 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67–71.
- Gerstein MB, Rozowsky J, Yan KK, Wang D, Cheng C, Brown JB, Davis CA, Hillier L, Sisu C, Li JJ. 2014. Comparative analysis of the transcriptome across distant species. *Nature* **512**: 445–448.
- Goff LA, Groff AF, Sauvageau M, Traves-Gibson Z, Sanchez-Gomez DB, Morse M, Martin RD, Elcavage LE, Liapis SC, Gonzalez-Celeiro M, et al. 2015. Spatiotemporal expression and transcriptional perturbations by long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci* **112**: 6855–6862.
- Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, O'Connor-Giles KM. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* **196**: 961–971.
- Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA, Elhaik E. 2013. On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol Evol* **5**: 578–590.
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**: 223–227.
- Haerty W, Jagadeeshan S, Kulathinal RJ, Wong A, Ravi Ram K, Sirot LK, Levesque L, Artieri CG, Wolfner MF, Civetta A, et al. 2007. Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*. *Genetics* **177**: 1321–1335.
- Heinen TJ, Staubach F, Haming D, Tautz D. 2009. Emergence of a new gene from an intergenic region. *Curr Biol* **19**: 1527–1531.
- Hicks JL, Deng WM, Rogat AD, Miller KG, Bownes M. 1999. Class VI unconventional myosin is required for spermatogenesis in *Drosophila*. *Mol Biol Cell* **10**: 4341–4353.
- Hu L, Di C, Kai M, Yang YC, Li Y, Qiu Y, Hu X, Yip KY, Zhang MQ, Lu ZJ. 2015. A common set of distinct features that characterize noncoding RNAs across multiple species. *Nucleic Acids Res* **43**: 104–114.
- Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M. 2010. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**: 409–419.
- Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, et al. 2015. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* **47**: 199–208.
- Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**: 816–821.
- Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, Shiekhattar R. 2013. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* **494**: 497–501.
- Lee JT. 2012. Epigenetic regulation by long noncoding RNAs. *Science* **338**: 1435–1439.
- Lee JT, Bartolomei MS. 2013. X-inactivation, imprinting, and long noncoding RNAs in health and disease. *Cell* **152**: 1308–1323.
- Lehmann R, Nüsslein-Volhard C. 1986. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*. *Cell* **47**: 141–152.
- Li MG, Serr M, Newman EA, Hays TS. 2004. The *Drosophila* tctex-1 light chain is dispensable for essential cytoplasmic dynein functions but is required during spermatid differentiation. *Mol Biol Cell* **15**: 3005–3014.
- Lu ZJ, Yip KY, Wang G, Shou C, Hillier LW, Khurana E, Agarwal A, Auerbach R, Rozowsky J, Cheng C. 2011. Prediction and characterization of non-coding RNAs in *C. elegans* by integrating conservation, secondary structure, and high-throughput sequencing and array data. *Genome Res* **21**: 276–285.
- Ma Z, Liu Z, Huang X. 2010. OSBP- and FAN-mediated sterol requirement for spermatogenesis in *Drosophila*. *Development* **137**: 3775–3784.
- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* **339**: 823–826.
- Maruyama T, Dougan SK, Truttman MC, Bilate AM, Ingram JR, Ploegh HL. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**: 538–542.
- Mattick JS. 2013. Probing the phenomics of noncoding RNA. *eLife* **2**: e01968.
- Meller VH, Rattner BP. 2002. The *roX* genes encode redundant *male-specific lethal* transcripts required for targeting of the MSL complex. *EMBO J* **21**: 1084–1091.
- Melo CA, Drost J, Wijchers PJ, van de Werken H, de Wit E, Vrieling JAO, Elkon R, Melo SA, Léveillé N, Kalluri R. 2013. eRNAs are required for p53-dependent enhancer activity and gene transcription. *Mol Cell* **49**: 524–535.
- Mercer TR, Mattick JS. 2013. Structure and function of long noncoding RNAs in epigenetic regulation. *Nat Struct Mol Biol* **20**: 300–307.
- Mermall V, Bonafe N, Jones L, Sellers JR, Cooley L, Mooseker MS. 2005. *Drosophila* myosin V is required for larval development and spermatid individualization. *Dev Biol* **286**: 238–255.
- Moran VA, Perera RJ, Khalil AM. 2012. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic Acids Res* **40**: 6391–6400.
- Morris KV, Mattick JS. 2014. The rise of regulatory RNA. *Nat Rev Genet* **15**: 423–437.
- Nam JW, Bartel DP. 2012. Long noncoding RNAs in *C. elegans*. *Genome Res* **22**: 2529–2540.
- 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.
- Ni MJ, Hu ZH, Liu Q, Liu MF, Lu MH, Zhang JS, Zhang L, Zhang YL. 2011. Identification and characterization of a novel non-coding RNA involved in sperm maturation. *PLoS One* **6**: e26053.
- Pauli A, Rinn JL, Schier AF. 2011. Non-coding RNAs as regulators of embryogenesis. *Nat Rev Genet* **12**: 136–149.
- Ponting CP, Oliver PL, Reik W. 2009. Evolution and functions of long non-coding RNAs. *Cell* **136**: 629–641.
- Rathke C, Baarends WM, Jayaramaiah-Raja S, Bartkuhn M, Renkawitz R, Renkawitz-Pohl R. 2007. Transition from a nucleosome-based to a protamine-based chromatin configuration during spermiogenesis in *Drosophila*. *J Cell Sci* **120**: 1689–1700.
- Rathke C, Barckmann B, Burkhard S, Jayaramaiah-Raja S, Rote J, Renkawitz-Pohl R. 2010. Distinct functions of Mst77F and protamines

- in nuclear shaping and chromatin condensation during *Drosophila* spermiogenesis. *Eur J Cell Biol* **89**: 326–338.
- Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M. 2013. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *eLife* **2**: e01749.
- Sitnik JL, Francis C, Hens K, Huybrechts R, Wolfner MF, Callaerts P. 2014. Neprilysins: an evolutionarily conserved family of metalloproteases that play important roles in reproduction in *Drosophila*. *Genetics* **196**: 781–797.
- Stark A, Lin MF, Kheradpour P, Pedersen JS, Parts L, Carlson JW, Crosby MA, Rasmussen MD, Roy S, Deoras AN, et al. 2007. Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**: 219–232.
- Tarazona S, Furió-Tarí P, Turrà D, Pietro AD, Nueda MJ, Ferrer A, Conesa A. 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res* **43**: e140.
- Texada MJ, Simonette RA, Johnson CB, Deery WJ, Beckingham KM. 2008. *Yuri gagarin* is required for actin, tubulin and basal body functions in *Drosophila* spermatogenesis. *J Cell Sci* **121**: 1926–1936.
- Ulitsky I, Bartel DP. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**: 26–46.
- Vance KW, Sansom SN, Lee S, Chalei V, Kong L, Cooper SE, Oliver PL, Ponting CP. 2014. The long non-coding RNA *Paupar* regulates the expression of both local and distal genes. *EMBO J* **33**: 296–311.
- Wiedenheft B, Sternberg SH, Doudna JA. 2012. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**: 331–338.
- Yeh SD, Do T, Chan C, Cordova A, Carranza F, Yamamoto EA, Abbassi M, Gandasetiawan KA, Librado P, Damia E. 2012. Functional evidence that a recently evolved *Drosophila* sperm-specific gene boosts sperm competition. *Proc Natl Acad Sci* **109**: 2043–2048.
- Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, Zhao Y, Shen B, Huang X, Zhu H. 2015. Opposing roles for the lincRNA *Haunt* and its genomic locus in regulating *HOXA* gene activation during embryonic stem cell differentiation. *Cell Stem Cell* **16**: 504–516.
- Young RS, Marques AC, Tibbit C, Haerty W, Bassett AR, Liu JL, Ponting CP. 2012. Identification and properties of 1,119 candidate lincRNA loci in the *Drosophila melanogaster* genome. *Genome Biol Evol* **4**: 427–442.
- Yu Z, Ren M, Wang Z, Zhang B, Rong YS, Jiao R, Gao G. 2013. Highly efficient genome modifications mediated by CRISPR/Cas9 in *Drosophila*. *Genetics* **195**: 289–291.
- Yu C, Liu Y, Ma T, Liu K, Xu S, Zhang Y, Liu H, La Russa M, Xie M, Ding S. 2015. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**: 142–147.
- Zhang L, Lu H, Xin D, Cheng H, Zhou R. 2010. A novel ncRNA gene from mouse chromosome 5 trans-splices with *Dmrt1* on chromosome 19. *Biochem Biophys Res Commun* **400**: 696–700.
- Zhong L, Belote JM. 2007. The testis-specific proteasome subunit Prosa6T of *D. melanogaster* is required for individualization and nuclear maturation during spermatogenesis. *Development* **134**: 3517–3525.

Received September 14, 2015; accepted in revised form July 12, 2016.