**DEVELOPMENTAL BIOLOGY**

**Testicular germ cell–specific IncRNA, Teshl, is required for complete expression of Y chromosome genes and a normal offspring sex ratio**

Seong Hyeon Hong¹, Gwidong Han¹, Seung Jae Lee¹, Julie Cocquet², Chunghee Cho¹*

Heat shock factor 2 (HSF2) regulates the transcription of the male-specific region of the mouse Y chromosome long arm (MSYq) multicopy genes only in testes, but the molecular mechanism underlying this tissue specificity remains largely unknown. Here, we report that the testicular germ cell–specific long noncoding RNA (IncRNA), **NR_038002**, displays a characteristic spatiotemporal expression pattern in the nuclei of round and elongating spermatids. **NR_038002**-knockout male mice produced sperm with abnormal head morphology and exhibited reduced fertility accompanied by a female-biased sex ratio in offspring. Molecular analyses revealed that **NR_038002** interacts with HSF2 and thereby activates expression of the MSYq genes. We designate **NR_038002** as a testicular germ cell–specific HSF2-interacting IncRNA (**Teshl**). Together, our study is the first to demonstrate that the testis specificity of HSF2 activity is regulated by the IncRNA **Teshl** and establishes a **Teshl**-HSF2-MSYq molecular axis for normal Y-bearing sperm qualities and consequent balanced offspring sex ratio.

**INTRODUCTION**

Long noncoding RNAs (lncRNAs) are a class of transcripts longer than 200 nucleotides (nt) with limited protein-coding potential (1). They have been implicated in diverse biological processes, including differentiation and development (2), and regulation of gene expression through cross-talk with DNA, RNA, and proteins (3). Compared with mRNAs, lncRNAs display less sequence conservation (4) and greater tissue-specific expression patterns (5). Although numerous lncRNAs have been reported to be specifically expressed in testes (6), only a few have been characterized (7–15).

Mammalian spermatogenesis is a tightly regulated series of developmental processes that occur in successive mitotic, meiotic, and postmeiotic phases (spermiogenesis) (16). During the pachytene stage of meiosis, autosomes are abundantly transcribed, whereas sex chromosomes are transcriptionally silent owing to meiotic sex chromatin inactivation (MSCI)—a surveillance mechanism for the largely nonhomologous regions of X and Y chromosomes (17–19). The repressed state of sex chromosomes partially persists throughout postmeiotic phases (20). Nevertheless, some X- and Y-linked genes, specifically those with indispensable functions in sperm differentiation, are transcribed during the postmeiotic phase. For example, loss of the male-specific region of the mouse Y chromosome long arm (MSYq), which contains the multicopy spermatid-expressed genes, **Syt1** and **Syt2** (spermiogenesis-specific transcript on the Y 1 and Y 2, respectively) and **Sly** (Syct3-like Y-linked), is associated with sperm head abnormalities and male infertility (21).

During spermatogenesis, heat shock factor 2 (HSF2), a transcription factor involved in corticogenesis and spermatogenesis (22), is expressed in a stage-specific manner in the nuclei of early pachytene spermatocytes and postmeiotically in round spermatids (23). **Hsf2**–knockout (KO) mice exhibit abnormal sperm chromatin compaction, sperm head abnormalities, and subfertility (24). HSF2 binds the promoters of a number of multicopy gene targets on sex chromosomes and regulates their transcription in the postmeiotic phase (24). However, how the ubiquitously expressed HSF2 regulates the expression of multicopy genes only in testes is largely unknown.

On the basis of these observations, we tested whether **NR_038002**, testicular germ cell–specific lncRNA, plays a critical role in regulating the expression of HSF2-dependent sex chromosome genes during postmeiotic phases. To explore the functional importance of this testicular germ cell–specific lncRNA, we used CRISPR-Cas9 genome editing to disrupt the genomic locus of **NR_038002**. Our investigation of transcriptome profiles and results of various biological analyses demonstrate that **NR_038002** interacts with HSF2 to activate Y chromosomal multicopy genes during the postmeiotic phase, a function that is critical for regulation of offspring sex ratio.

**RESULTS**

**NR_038002** is a testicular germ cell–specific IncRNA conserved between mice and humans

We previously reported that **NR_038002**, transcribed from the 1700027A15Rik gene on mouse chromosome 1, is a testicular germ cell–specific IncRNA predicted to have the highest expression levels among identified specific lncRNAs (6). For convenience, we use the term “**NR_038002**” for both the gene and transcript in this study. To further examine **NR_038002**, we performed Northern blot analyses on various mouse tissues (Fig. 1A). **NR_038002** exhibited a testis-specific expression pattern, with a distinct ~600-nt band. Rapid amplification of complementary DNA ends (RACE), used to identify potential isoforms and 5’ and 3’ ends of **NR_038002** (fig. S1A), showed the presence of a few transcripts with slight differences in the 5’ sequence. This is likely due to random variation because only a single TATA box was found. We found two isoforms for **NR_038002**, differing in the length of poly A tail (fig. S1B). We found that **NR_038002** transcription was terminated at position +416, followed...
by polyadenylation. In silico programs were used to investigate the coding potential of NR_038002. CPAT (25), but not CPC2 (26), predicted that translation of NR_038002 into a protein was highly unlikely (fig. S1C). Sucrose gradient–based polysome fractionation predicted that translation of NR_038002 into a protein was highly unlikely (fig. S1C). Sucrose gradient–based polysome fractionation predicted that translation of NR_038002 into a protein was highly unlikely (fig. S1C).

We next examined the sequence conservation of NR_038002 based on PhastCons and PhyloP scores (27) and found that exonic regions of the NR_038002 gene are conserved among diverse mammalian species (fig. S1D). In particular, we identified a human locus, termed LINC01921, that transcribes an RNA with a nucleotide sequence very similar to that of NR_038002 (68% identity in exonic sequence) (Fig. 1C). Notably, such a level of conservation is observed in lncRNAs that are known to be well conserved between mice and humans, including Malat (79%), H19 (71%), Hotair (69%), and Xist (67%). We also found syntenic localization of NR_038002 and LINC01921 genes in mouse and human genomes, respectively, with both being located close to the genomic locus of the Tnp1 (transition protein 1) gene (28) (Fig. 1C). A structural analysis using RNAfold (29) predicted that both transcripts form stem-loops and have an overall similar structure (fig. S1E). Last, a tissue distribution analysis showed that LINC01921 is also specifically expressed in testes and surrounding tissues of the testis, indicating additional conservation at the expression level.

**NR_038002 is located in nuclei of round and elongating spermatids**

We previously reported that NR_038002 is expressed in a germ cell–specific manner, and its expression is limited to the postmeiotic phase (6). To determine exactly when and where NR_038002 is expressed in germ cell developmental stages, we performed in situ hybridization. Signals obtained using an antisense probe were distinctly detected in germ cells of seminiferous tubules, but not in testicular somatic cells, such as Sertoli and Leydig cells (Fig. 2A). In spermatogenic cells, NR_038002 was expressed at the I–III and VII–XII stages of spermatogenesis (7 to 15 steps of spermiogenesis), corresponding to late-stage round spermatids and entire-stage elongating spermatids (Fig. 2B). NR_038002 was located only in nuclei of the spermatids, indicating that NR_038002 is a developmentally regulated nuclear lncRNA expressed in round and elongating spermatids (Fig. 2C).

To investigate the transcriptional regulatory function of NR_038002 in mouse germ cells, we overexpressed NR_038002 transcripts in GC-2 cells and then performed a microarray analysis of the cells. However, this analysis showed no significant changes in the expression of genes related to spermatogenesis (table S1).

**Male mice lacking NR_038002 are subfertile and produce sperm with an abnormal head morphology**

To investigate the in vivo function of NR_038002, we generated NR_038002-mutant mice using CRISPR-Cas9 genome editing (Fig. 3A).
Fig. 2. Spatiotemporal expression patterns of NR_038002 in mouse testes. (A) In situ hybridization analysis of NR_038002 in mouse testes using sense and antisense probes. Scale bar, 20 μm. (B) Higher magnification of in situ hybridization in serial sections for determination of expression stages and localization of NR_038002 in seminiferous tubules. NR_038002 expression was detected in nuclei of late round spermatids (weak signal) and elongating spermatids (strong signal). Note that the NR_038002 nuclear signal is different from hematoxylin and eosin (H&E) stain in color and shape (blurred), as indicated with red arrow in VII to VIII. Scale bar, 20 μm. (C) Diagram of 12-stage spermatogenesis represented by Roman numerals and 16-step spermiogenesis indicated by Arabic numbers in mice (53) showing the stages in which NR_038002 (red) is expressed.
We ultimately obtained four founder mice, all of which were confirmed through cloning and sequencing of genomic DNAs (fig. S2, A and B). For phenotypic analyses, we selected a single mouse line in which the 894-bp (base pair) genomic region of NR_038002 was deleted and inbred these founder mice to obtain homozygous NR_038002-KO mice. The lack of NR_038002 transcripts was confirmed by Northern blotting and quantitative reverse transcription polymerase chain reaction (qRT-PCR) analyses (Fig. 3, B to D). NR_038002-KO mice were viable and developed normally. There was no significant difference in testis size or testis-to-body weight ratio between wild-type (WT) and KO mice (Fig. 3, E and F), and histological analyses revealed no disruptions in seminiferous tubules in NR_038002-KO mice (Fig. 3G). To further investigate the phenotype of NR_038002-KO mice, we evaluated mature sperm collected from the cauda epididymis and vas deferens of WT and KO mice. Although the number of sperm was not different between WT and KO mice (Fig. 3H), a microscopic analysis revealed defects in the head morphology of sperm from NR_038002-KO mice (Fig. 3, I and J). The most commonly observed abnormalities were small, flattened, and misshapen heads (Fig. 3I), implying the involvement of NR_038002 in sperm head formation. Quantitative analyses showed that sperm with abnormal heads accounted for ~46.7% of the total sperm population in NR_038002-KO mice compared with ~3.7% in WT mice. To determine whether the loss of...
**NR_038002** affects male fertility, we performed a fertility test, as detailed in fig. S2C. **NR_038002**-KO male mice were subfertile compared with WT mice (8.6 pups versus 5.9 pups) (Fig. 3K), and the degree of the sperm head abnormality was correlated with the decline in fertility in each individual KO mouse (Fig. 3L). These results demonstrate that an **NR_038002** deficiency leads to sperm head malformation and a consequent reduction in male fertility.

The abnormal head morphology of the sperm could be related to a defect in chromatin condensation, reflecting stepwise replacement of histones during spermiogenesis, first with transition proteins and lastly with protamines (30). To test this, we analyzed changes in the expression of transition proteins and protamines in **NR_038002**-KO sperm. We found that expression levels of transition proteins 1/2 (Tnp1/2) and protamine 1/2 (Prm1/2) were unchanged (fig. S3A). Consistently, PRM1 and PRM2 protein expression levels and processing were also normal in **NR_038002**-KO sperm (fig. S3B). We tested whether **NR_038002**-KO sperm are sensitive to chemical disruption, releasing genomic DNA fragments upon treatment with SDS under nonreducing conditions. We found that **NR_038002**-KO sperm did not differ from WT sperm with respect to this chromatin property (fig. S3C). Thus, altered sperm head formation in the absence of **NR_038002** is unlikely due to incomplete chromatin assembly.

---

**Loss of **NR_038002** causes up-regulation of X chromosomal gene expression**

To comprehensively investigate the function of **NR_038002** during male germ cell development, we performed RNA sequencing (RNA-seq) analyses of WT and KO testes. Deletion of **NR_038002** was accompanied by significant changes in the expression (fold change >1.5 and \( P < 0.05 \)) of a total of 82 genes in testes (Fig. 4A and table S2), 65 of which were up-regulated and 17 of which were down-regulated. Notably, 69 of 82 differentially expressed genes (DEGs) were located on X or Y chromosome. All 54 X-linked genes were up-regulated in **NR_038002**-KO testes, whereas expression of the 15 Y-linked genes was reduced in KO testes (Fig. 4B). Because some sex-linked genes were previously known as multicopy genes, we analyzed the DEGs whether they are multicopy genes. All of the 15 Y-linked DEGs belonged to multicopy gene families, such as **Rbm31y**, **Sly**, and **Ssty**, whereas 28 of 54 (51.9%) X-linked genes were members of multicopy gene families, including **Btbd35** (BTB domain containing 35), **Gm6812**, **H2al1** (H2A histone family member L1), **Mageb** (melanoma antigen, family b), **Samt1** (spermatogenesis-associated multipass transmembrane protein 1), **Sly/Sixl1** (Scyp3-like, X-linked/Slx-like 1), and **Tgfl2lx** (TGFβ-induced factor homeobox 2–like, X-linked) (Fig. 4C and table S2). Gene Ontology analyses of these DEGs showed that the corresponding encoded proteins are
localized to the nucleus, nuclear chromatin, and nucleosomes and are functionally involved in DNA binding, chromatin silencing, spermatid development, and gamete generation (Fig. 4D and table S3). Using qRT-PCR, we validated changes in the expression of some X-linked DEGs with fold change > 2 (Fig. 4E). The 54 X-linked DEGs included H2af1 variants and H2af3, both of which are known to be highly expressed in elongating spermatids (31). A recent study demonstrated that H2af2, which is closely related to H2af1, participates in chromatin organization during spermiogenesis (32). These data suggest that NR_038002 plays a role in transcriptional repression of X chromosomal genes in mouse testes.

**Loss of NR_038002 down-regulates Y chromosomal multicopy gene expression in the MSYq and in turn activates X chromosomal gene expression**

In the context of repression of X chromosomal gene expression, it should be noted that the protein encoded by the Y chromosome–located multicopy gene, Sly, has been found to regulate the expression of X and Y chromosome–encoded genes in mice (33). In addition to Sly, MSYq contains several multicopy genes, including Srsy (serine-rich secreted, Y-linked) as well as Ssty1 and Ssty2. These genes exist as hundreds of copies of an ampliconic unit (Fig. 5A) (34) and, with the exception of Srsy, are expressed predominantly in mouse testes (35). Similarly, multicopy genes encoding SLX and SLXL1 are located on the mouse X chromosome and regulate X- and Y-linked genes, but their products antagonize the regulatory influence of SLY (36). SLY acts as a transcription corepressor, interacting with TBL1XR1, a member of the SMRT/N-CoR repressor complex (37), and represses the expression of a number of X and Y chromosomal genes (33), including histone variants, similar to our findings in NR_038002 KO mice. SLY–knockdown (KD) mice are subfertile and produce sperm with abnormal head morphology (33).

Our RNA-seq data demonstrated the down-regulation of Sly and Ssty in NR_038002–KO testes. Using qRT-PCR to obtain a more accurate measurement of expression levels, we reinvestigated the expression of multicopy sex chromosomal genes in NR_038002–KO testes. Both Sly mRNA and protein were slightly, but significantly, down-regulated in NR_038002–KO testes (Fig. 5, B to D). Ssty1 and Ssty2 transcript levels were also markedly reduced in NR_038002–KO testes (Fig. 5B). Unlike the case for Sly, Ssty1, and Ssty2, expression of Slx/Slxl1 was normal in the absence of NR_038002 (Fig. 5, B to D). To analyze the relationship of NR_038002 with Sly, we examined the expression of NR_038002 in Sly–KD testes. NR_038002 transcript levels were not changed in Sly–KD testes (Fig. 5E), suggesting that the regulation of X chromosomal genes by SLY occurs downstream of NR_038002 activity. Thus, it is likely that the up-regulation of X chromosomal genes in NR_038002–KO testes is directly attributable to Sly down-regulation. A profile analysis of DEGs (fold change > 1.5 and \( P < 0.05 \)) revealed that 12 sex-linked genes were misregulated in both NR_038002–KO and Sly–KD testes (Fig. 5F and table S4). Not all X and Y chromosomal genes were similarly regulated in NR_038002–KO and Sly–KD testes, likely owing to the modest reduction in the amount of Sly mRNA and protein in NR_038002–KO mice (Fig. 5, B to D). Moreover, expression levels of Ssty1 and Ssty2, which are up-regulated in the absence of Sly (33), were decreased in NR_038002–KO testes (Fig. 5B). Thus, it is possible that NR_038002 regulates the expression of Ssty1 and Ssty2 as well as Sly, but the regulation of Ssty1 and Ssty2 by NR_038002 is independent of Sly. STY1 and STY2 appear to be functionally important, with properties similar to the H3K4me3 reader function of spindlin (38). Moreover, SLY and SLX/SLXL1 were recently reported to compete for interaction with SSTY1 at the promoter of thousands of spermatid-specific multicopy genes, exerting an opposite regulatory effect on their expression (38). Our results, together with previous Sly studies, suggest that NR_038002 activates the expression of the Y chromosomal multicopy genes, Sly and Ssty1/2, which in turn repress X chromosomal gene expression.

**Lack of NR_038002 results in a female-biased offspring sex ratio**

Sly and Slx/Slxl1 are transmission distorters, involved in an intragenomic conflict that favors their own transmission (38). Because the expression levels of SLY and Slx/Slxl1 were unbalanced in NR_038002–KO mice owing to down-regulation of SLY, we rigorously examined offspring for a distortion in their sex ratio. A count of the number of offspring obtained from mating with female mice showed that the proportions of male and female pups for KO males were 0.39 and 0.61, respectively (\( n = 13 \)), compared with 0.5 and 0.5 for WT males (\( n = 18 \)) (Fig. 5G). We then calculated the sex ratio distortion based on the total number of offspring. This analysis showed that KO male mice sired more female offspring (60.4%; 198 of 328) compared with WT male mice (50.2% female; 139 of 277), a difference that approached statistical significance in a Fisher’s exact test (\( P < 0.05 \)) (Fig. 5H). A female-biased sex ratio was also observed in Sly–KO mice, 2/3MSYq KO mice, and Y\(^{Hsf2}\)-deleted mice, findings at odds with the male-biased ratio observed in Slxl1–KO mice (33, 39, 40). Our results suggest that deregulation of multicopy genes on the X or Y chromosome in NR_038002–KO male mice causes an X-versus-Y sperm fitness difference, thereby resulting in distortion of the sex ratio toward female offspring. In this regard, it should be noted that the sex ratio distortion produced by the absence of NR_038002 appears to reflect a decrease in the number of male offspring (4.0 and 2.9 male offspring from WT and KO males, respectively), rather than a change in female progeny number (3.9 and 4.3 female offspring from WT and KO males, respectively) (Fig. 5I and table S5). A decrease in male progeny number is congruent with the degree of reduced fertility in male NR_038002–KO mice (Fig. 3K), implying impaired fertility of Y-bearing sperm.

**NR_038002 interacts with HSF2 to activate Y chromosomal genes during spermiogenesis**

Previous studies have shown that transcription of the sex chromosomal multicopy genes, Slx, Sly, Ssty1, and Ssty2, is regulated by HSF1 and HSF2 (24, 41). In Hsf2–KO testes, both Ssty2 and Sly mRNA levels are reduced, but Slx mRNA levels are increased (24). To investigate how NR_038002 regulates the expression level of genes on sex chromosomes, we first examined whether NR_038002 regulates mRNA and protein expression levels of Hsf1 and Hsf2 in mouse testes, but found no significant change in expression of either at the mRNA or protein level in NR_038002–KO testes (Fig. 6, A to C). Next, we performed a biotin-labeled RNA pull-down assay to test whether NR_038002 interacts with HSF1 and HSF2 in mouse testes (Fig. 6D). Notably, we found that biotinylated sense of NR_038002 specifically bound to HSF2, but not HSF1, in testes (Fig. 6E). To identify the HSF2-binding region of NR_038002, we constructed three truncated fragments of NR_038002 and tested them in RNA pull-down assay (Fig. 6F). These analyses revealed that the 5’ fragment (1 to 100 nt) clearly interacted with HSF2, whereas the other
fragments (101 to 230 nt and 231 to 416 nt) did not (Fig. 6F). To validate the interaction between NR_038002 and HSF2, we performed RNA immunoprecipitation (RIP) in mouse testes using an anti-HSF2 antibody. We found that NR_038002 was highly enriched in anti-HSF2 antibody immunoprecipitates compared with control immunoglobulin G (IgG) immunoprecipitates (Fig. 6, G and H). A previous chromatin immunoprecipitation (ChIP)-chip analysis showed that promoters located upstream of Sly, Ssty1, and Ssty2 are occupied by HSF2 only in testes, and not in the brain, muscle, or kidney, suggesting that the target promoter sequences alone are not sufficient to specify HSF2 binding (24). A comparison of HSF2 binding to the promoter sequences of Sly, Ssty1, and Ssty2 in NR_038002 WT and KO tests showed greater occupancy of these promoters by HSF2 in WT testes than in KO testes (Fig. 6I), indicating the importance of NR_038002 for HSF2 binding to Sly, Ssty1, and Ssty2 promoters. Further investigation is needed to determine whether NR_038002 and HSF2 together bind to the promoters or NR_038002 only stabilizes the binding of HSF2 to the promoters. Collectively, our results suggest that the interaction of NR_038002 with HSF2 is critical for the expression of Y chromosomal multicopy genes, whose postmeiotic expression, in turn, is crucial for normal sperm maturation and fertility (Fig. 6I). On the basis of these findings, we newly designate NR_038002 as testicular germ cell–specific HSF2-interacting lncRNA (Teshl).
DISCUSSION

In the present study, we report the novel finding that the conserved testicular germ cell–specific IncRNA NR_038002, hereafter Teshl, activates the Y chromosomal multicopy genes, Sly and Ssty, through interaction with the transcription factor HSF2 in mouse testes, and showed that this process is required for production of normal Y-bearing sperm. Deletion of Teshl in mouse testes caused impaired sperm head formation and resulted in sex ratio distortion toward females in offspring (Fig. 7). Disruption of Hsf2 exhibited similar phenotypes, such as altered expression of multicopy genes, sperm head abnormalities, and subfertility (24, 42).

The deletion of Teshl led to deregulation of distinct sets of sex chromosomal multicopy genes, as evidenced by the significant up-regulation of X-linked genes and down-regulation of Y-linked genes. Sex chromosomes are transcriptionally silenced during meiosis by MSCI and then partially reactivated at genes (17) with important roles during sperm maturation. Although sex chromosomal multicopy genes have been proposed to reactivate sex chromosomes
during this postmeiotic period, the detailed molecular mechanism is largely unknown. Our study provides the first evidence for operation of a Teshl-HSF2-MSYq axis in this process, showing that Teshl-dependent HSF2 activity promotes expression of Y chromosomal multicopy genes in MSYq during the postmeiotic phase. Both HSF1 and HSF2 are transcription factors that regulate the transcription of multicopy genes located in the MSYq and share ~15% of targets (24, 41, 43). Our data showed that Teshl interacts with HSF2 but not HSF1 (Fig. 6E), indicating that Teshl-dependent HSF2 transcriptional regulation is independent of HSF1. HSF1 activation by heat shock is mediated by a ribonucleoprotein complex containing the translation elongation factor eEF1A and the lncRNA, HSR1 (44), which may serve as a thermosensor with its constitutive expression pattern. However, HSF2 is transiently activated by Teshl to regulate postmeiotic target genes during spermiogenesis. Teshl may act as a developmental cue with its spatiotemporal expression pattern.

The temporal pattern of Teshl expression is very similar to that of its nearest gene, Tnp1 (elongating spermatids), suggesting the possibility that both genes are transcribed by a coregulatory mechanism. If so, the histone-to-proteins exchange process necessary for chromatin condensation and the transient expression of sex chromosomal multicopy genes involved in spermiogenesis may be temporally coordinated.

The current model for equal transmission of X- and Y-bearing sperm to offspring posits that a balanced copy number of Slx/Slxl1 and Sly/Sty genes is necessary for normal X- and Y-bearing sperm differentiation (38). For example, Slx/Slxl1-KO male mice exhibit a male-biased sex ratio in offspring, whereas males carrying Slx/Slxl1 duplications display a distortion toward female offspring (40), indicating the importance of Slx/Slxl1 copy number. Male mice with a deletion of the MSYq exhibit severely impaired sperm differentiation and a female-skewed offspring sex ratio (39). Sly-KD male mice also show a female-biased sex ratio in offspring (33).
distortion toward females in MSYq-deleted mice is likely attributable to the absence of Y chromosomal multicopy genes rather than the partial loss of the Y chromosome. Likewise, our Teshl-KO mice exhibited a female-biased distortion in sex ratio among offspring owing to substantial inactivation of the Y chromosomal multicopy genes, Sly and Ssyt. Misregulation of most of the MSYq multicopy genes in Teshl KO mice suggests chromosome-wide–level regulation of gene expression by Teshl-HSF2. Cytoplasmic bridges connecting spermatids are known to facilitate the sharing of haploid gene products (see also Fig. 7) (45). The question is, how does down-regulation of Y chromosomal multicopy genes cause a distortion in the sex ratio in offspring? A recent analysis of sex chromosomal transcripts at the single-cell level after meiosis showed that thousands of transcripts, including Ssx, Sly, and Ssyt, are not equally shared despite the presence of cytoplasmic bridges (46). This suggests that spermatids do not display absolute phenotypic equivalence with regard to gene products of sex chromosomes. The sex ratio distortion observed in Teshl-KO mice is not attributable to an imbalance in the proportion of X- or Y-bearing sperm because the total number of sperm was not changed in Teshl-KO mice (Fig. 3H). Furthermore, the sex ratio changes in offspring of male Teshl-KO mice compared with WT male mice was attributable to a decrease in the number of male progeny, not an increase in the number of female progeny (table S5). X- and Y-bearing sperm from male mice with a deletion of the MSYq have altered morphology and motility (47). Similarly, the distortion in the sex ratio toward females in offspring of Teshl-KO mice may be related to the abnormal head morphology and/or motility defects of Y-bearing sperm. Therefore, our study strongly suggests that, by activating Y chromosomal multicopy genes whose products contribute to normal spermatogenesis, Teshl plays a crucial role in conferring the normal fertility and integrity of Y-bearing sperm (Fig. 7).

One of the most common genetic causes of spermatogenesis failure in humans is deletion of the MSYq, which leads to copy number variations of Y chromosomal multicopy genes and, ultimately, male infertility (48). In this study, we identified LINC01921 as a human ortholog of Teshl (Fig. 1D). It was recently reported that LINC01921 is significantly decreased or differentially expressed in patients with nonobstructive azoospermia (49). Although further investigation is required to determine whether a decrease in the expression levels of LINC01921 is directly involved in nonobstructive azoospermia, this human IncRNA Teshl ortholog could be considered a suitable biomarker for diagnosis of human infertility.

In conclusion, we identified a new testicular germ cell–specific IncRNA and defined its biological and molecular nature, demonstrating the importance of the Teshl-Hsfl2-MSYq axis in postmeiotic gene expression mechanisms on the Y chromosome. Our study further contributes to our understanding of sex ratio variations and suggests the potential for the development of future fertility therapeutics.

**MATERIALS AND METHODS**

**Animals**
All animal experiments were performed in accordance with Korean Food and Drug Administration guidelines. Protocols were reviewed and approved by the Institutional Animal Care and Use Committee of Gwangju Institute of Science and Technology (GIST) (permit number: GIST-2020-034). Each tissue sample was snap frozen in liquid nitrogen for later extraction.

**Northern blotting**
Total RNA was isolated from each mouse tissue using TRIzol reagent (Ambion). RNA was heated at 65°C for 5 min and resolved on a denaturing agarose gel. After equilibrating the gel for 15 min in distilled water and then twice for 10 min each in 10× saline sodium citrate (SSC) buffer, total RNA was transferred to Hybond-XL membranes (GE Healthcare). After ultraviolet cross-linking, the blot was prehybridized for 1 hour at 65°C in Rapid-hyb Buffer (Amersham), followed by hybridization for 2 hours at 65°C in the presence of a cDNA probe. The probe, derived from products amplified using gene-specific primers (table S6), was purified using ProbeQuant G-50 Micro Columns (GE Healthcare) and labeled with [α-32P]dCTP (Perkin Elmer) using a Prime-It II Random Primer Labeling Kit (Agilent) according to the manufacturer’s instructions. Blots were washed two to four times in 2× SSC and 0.05% SDS at 37°C for 10 min and twice in 0.1× SSC and 0.1% SDS at 65°C for 10 min, after which they were exposed to FUJI Medical X-ray Film (Fujifilm) at −70°C.

**Rapid amplification of cDNA ends**
Deoxyribonuclease I (DNase I)–treated total RNA from a mouse testis was used for first-strand cDNA synthesis. The SMARTer RACE cDNA Amplification Kit (Clontech) was used to perform 5′- and 3′-RACE according to the manufacturer’s instructions. The gene-specific primers used for RACE are listed in table S6. RACE products were cloned into the linearized pRACE vector using the In-Fusion HD cloning kit and then sequenced.

**Sucrose gradient–based polysome fractionation**
Beckman ultracentrifuge tubes were loaded with 2.2 ml of a 10 to 50% sucrose (Sigma-Aldrich) gradient in salt buffer [2 mM tris (pH 7.5), 10 mM NaCl, and 0.5 mM MgCl2] and incubated horizontally at 4°C overnight. Two pieces of mouse testes were homogenized in 1 ml of polysome extraction buffer [20 mM tris (pH 7.5), 100 mM KCl, 5 mM MgCl2, 0.5% (v/v) NP-40, protease inhibitor cocktail, RNasin (Promega), and cycloheximide (100 μg/ml)]. After a 10-min centrifugation, clarified lysates were layered onto ultracentrifuge tubes containing 10 to 50% sucrose gradients and centrifuged at 39,000 rpm for 90 min in an SW41 rotor (Beckman Coulter) at 4°C. Gradients were collected in 12 × 1–ml fractions using a Pasteur pipet, and absorbance was measured at 254 nm. RNA was isolated by extraction using TRIzol reagent. RNA quality and amount were determined using an ND-1000 spectrophotometer (Thermo Fisher Scientific).

**Prediction of RNA secondary structure**
Secondary structure analyses of IncRNAs were performed using RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) with default parameters.

**Conventional RT-PCR**
DNase I–treated tissue RNA was reverse transcribed using Omniscript RT kit (Qiagen) according to the manufacturer’s instructions. RT-PCR was carried out using primers listed in table S6, and results were normalized to mouse Gapdh mRNA. Analyses of a human cDNA panel used 5 μl of cDNA purchased from Clontech (Human MTC panels I and II cDNA panel). A human GAPDH primer pair in the kit was used as a control.
Quantitative RT-PCR
Testicular samples from Sly-KD mice were prepared as previously described (33). Total testis RNA was reverse transcribed using a QuantiNova RT kit (Qiagen) according to the manufacturer’s instructions. qRT-PCR was carried out using TOPreal qPCR 2X premix (Enzymonics) with primers listed in table S6. Amplifications were performed in at least triplicate for each sample. Relative gene expression levels were evaluated using the 2−ΔΔCt method (50) and were normalized to the level of Gapdh mRNA. Data are presented as means ± SD. Statistical analyses were performed using two-tailed Student’s t tests, and P ≤ 0.05 was taken as indicating a significant difference.

In situ hybridization
Sense and antisense NR_038002 probes were amplified from mouse testis cDNA using the primer pair indicated in table S6. After cloning into the pGEM T Easy Vector (Promega), the plasmid was linearized by digesting with Sac II or Pst I, generating sense and antisense probes, respectively. Digoxigenin (DIG)–labeled RNA probes were synthesized using a DIG RNA Labeling Kit (Roche) with SP6 RNA polymerase as a sense probe and T7 RNA polymerase as an antisense probe. Adult mouse testes were fixed in Bouin’s solution (Sigma-Aldrich) overnight at room temperature (RT), then dehydrated in an ethanol series, embedded in paraffin, cut into 5-μm-thick sections, and loaded onto slide glasses. Sections were deparaffinized, rehydrated, and stained with hematoxylin and eosin (H&E; Sigma-Aldrich). For in situ hybridization, sections were deparaffinized and air dried for 10 min, and then fixed with 4% paraformaldehyde (PFA) for 10 min and treated with proteinase K (1 μg/ml) (GeneAll) at 37°C for 5 min. After washing in 1× phosphate-buffered saline (PBS), sections were postfixed by incubating with 4% PFA for 5 min. Thereafter, sections were acetylated by incubating in acetylation buffer [1.35% (v/v) triethanolamine (Sigma-Aldrich), 1.75% (v/v) concentrated hydrochloric acid (Sigma-Aldrich), and 0.25% (v/v) acetic anhydride (Sigma-Aldrich)] for 10 min, and then washed with 1× PBS. The sections were prehybridized with hybridization buffer [deionized formamide (500 μl/ml) (Sigma-Aldrich), Denhardt’s solution (20 μl/ml) (Roche), baker’s yeast RNA (0.25 μg/ml) (Sigma-Aldrich), salmon sperm DNA (0.5 μg/ml) (Roche), and 1× SSC] for 1 hour at RT. After denaturing DIG-labeled probes (1 μg) at 80°C for 5 min, sections were hybridized with DIG-labeled probes in hybridization buffer by incubating overnight at 68°C. After hybridization, the sections were washed once in 5× SSC for 10 min at 65°C, twice in 0.2× SSC for 1 hour at 65°C, and once in 0.2× SSC for 5 min at RT. After washing in B1 buffer [0.1 M tris (pH 7.4) and 1.67 M NaCl] for 5 min at RT, sections were blocked by incubating with 10% (v/v) heat-inactivated normal goat serum (HINGS) in B1 buffer for 1 hour, and then incubated overnight at 4°C in B1 buffer containing 1% HINGS and anti-DIG antibody (Roche), diluted 1:5000. Sections were then washed three times in B1 buffer and equilibrated in B2 buffer [0.1 M tris (pH 9.5) and 0.1 M NaCl]. Last, sections were incubated with NBT/BCIP (Promega), diluted 1:50, in B2 buffer for several hours until the signal was detectable and then observed under an Aperio ScanScope (Leica).

Cell culture and transfection
Mouse GC-2 cells were obtained from the American Type Culture Collection (ATCC) and cultured with medium and serum as recommended by the ATCC. For transfection, full-length NR_038002 was amplified from mouse testis cDNA using primers incorporating Kpn I and Age I restriction sites (table S6), and cloned into the pcDNA3.1A vector. Transfection was performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer’s instructions.

Microarray
RNA purity and integrity were evaluated using an ND-1000 spectrophotometer (Thermo Fisher Scientific) and Agilent 2100 Bioanalyzer (Agilent). The Affymetrix whole transcript expression array process was executed according to the manufacturer’s protocol (Affymetrix). cDNA was synthesized using the GeneChip WT (Whole Transcript) Amplification kit as described by the manufacturer. The sense cDNA was then fragmented and biotin labeled with Tdt (terminal deoxynucleotidyl transferase) using a GeneChip WT Terminal Labeling Kit. Approximately 5.5 μg of labeled DNA target was hybridized to the Affymetrix GeneChip Mouse 2.0 ST Array at 45°C for 16 hours. Hybridized arrays were washed and stained on a GeneChip Fluidics Station 450 and scanned on a GCS3000 Scanner (Affymetrix). Signal values were computed using Affymetrix GeneChip Command Console software. Raw data were extracted automatically with the Affymetrix data extraction protocol using the provided Affymetrix GeneChip Command Console Software. After importing CEL files, the data were summarized and normalized using the robust multiarray method implemented in Affymetrix Expression Console Software. Comparative analyses of test and control samples were carried out using fold change >2 and a P < 0.05 as criteria. Statistical analyses and visualization of DEGs were implemented in R software and Galaxy (https://usegalaxy.org/).

Generation of NR_038002-KO mice
Cas9 protein tagged with a nuclear localization signal and guide RNAs (gRNAs) targeting exon 1 and exon 2 of the NR_038002 gene were purchased from Macrogen. The activity of gRNAs (exon 1_gRNA: CGT GTC ACA GAC TCT TAT AAA GG; exon 2_gRNA: GAG GTT TTA TTT AGA GAC CCT GG) was validated by assessing deletion of NR_038002 using in vitro cleavage reactions. Briefly, amplified NR_038002, used as a template, was incubated for 90 min at 37°C with Cas9 protein (20 nM) and single gRNA (sgRNA) (40 nM) in 1× NEB 3.1 buffer. Reactions were stopped with 6× stop solution containing 30% glycerol, 1.2% SDS, and 100 mM EDTA. Cleavage activity was confirmed by electrophoresis of reaction mixtures. NR_038002-KO mice, generated by Macrogen, were interbred and maintained at Macrogen facilities under pathogen-free conditions. For breeding, C57BL/6 female mice were treated with pregnant mare serum gonadotropin and human chorionic gonadotropin. After 48 hours, treated female mice were mated with C57BL/6 male mice. The next day, female mice were checked for vaginal plugs, after which mice were euthanized and fertilized embryos were harvested. One-cell embryos were then microinjected with a mixture of sgRNA and Cas9 protein and incubated at 37°C for 1 to 2 hours. Fourteen to 16 injected one-cell-stage embryo were transplanted into oviducts of pseudopregnant recipient mice (ICR). After founder offspring were born, they were genotyped by PCR analysis of tail-clip samples using the primers listed in table S6, followed by sequencing.

Phenotypic analyses of NR_038002-KO mice
Testis integrity was analyzed by fixed testes from WT and NR_038002-KO male mice by immersion in Bouin’s solution overnight. The tissue was then dehydrated, embedded in paraffin, and sectioned (5 μm thick). Testis sections from male WT and NR_038002-KO
mice were H&E stained and scanned using an Aperio ScanScope CS2 System (Leica Biosystems). For sperm counts, sperm from the cauda epididymis and vas deferens of male WT and NR_038002-KO mice were collected and counted in a hemocytometer under a light microscope. For analysis of sperm head morphology, mature sperm were isolated from the cauda epididymis. Sperm from male WT (n = 8) and NR_038002-KO (n = 8) mice were spread onto microscope slides and allowed to air dry. About 200 sperm on each slide were analyzed under a light microscope. For fertility tests, adult C57BL/6 WT females (8 weeks old) were mated with sexually mature male WT (n = 10) and NR_038002 KO (n = 8) mice for 14 days and checked daily for vaginal plugs. The fertility rate was calculated from the number of fetuses produced by pregnant females, counted 12 to 16 days after mating.

**Analysis of DNA fragmentation and sperm nuclear basic proteins**

Genomic DNA (gDNA) was obtained by incubating isolated sperm in lysis buffer containing 2% SDS and protease K (0.5 mg/ml) for 2 hours at 55°C. gDNAs in remaining sperm nuclei were released by incubating nuclei in lysis buffer containing 20 mM dithiothreitol for 6 hours. Released gDNA was subsequently purified on a DNeasy column (Qiagen). Chromatin was isolated from sperm from the cauda epididymis and vas deferens by treatment with cetyltrimethylammonium bromide, and basic proteins were extracted with NaCl, urea, and guanidine hydrochloride. Proteins were precipitated with trichloroacetic acid, separated on a 15% acid-urea polyacrylamide gel, and stained with Coomassie blue.

**RNA sequencing**

Libraries were prepared from total RNA using the NEBNext Ultra II Directional RNA-Seq Kit (NEB). mRNA used for cDNA synthesis and shearing was isolated using a Poly(A) RNA Selection Kit (LEXOGEN) following the manufacturer’s instructions. Indexing was performed using Illumina indexes 1 to 12. After performing a PCR-based enrichment step, libraries were evaluated for mean fragment size using the Agilent 2100 Bioanalyzer (DNA High Sensitivity Kit), and the library was quantified using a StepOne Real-Time PCR System (Life Technologies). High-throughput paired-end sequencing (100-bp reads) was performed using HiSeq 2500 (Illumina). Alignment files were obtained by mapping mRNA-Seq reads using the HISAT2 software tool. Alignment files, in turn, were used as input for the Cufflinks program for assembling transcripts, estimating their abundances, and detecting DEGs and isoforms. Expression levels of gene regions were determined using the FPKM (fragments per kilobase of exon per million fragments) method, and the resulting FPKM data were normalized by trimmed mean of M values (TMM) normalization using the EdgeR module in R software. Raw and processed RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) database at the National Center for Biotechnology Information (NCBI) and can be accessed using the GEO accession number, GSE162652. To predict the gene families of the DEGs, we first obtained the sequence information of each gene and subsequently investigated a copy number with an identity greater than 95%, using BLASTN program with reference selected from NCBI Reference Sequence Database (Refseq). Isoforms, pseudogenes, and noncoding genes were excluded. If more than one gene was detected through BLASTN search, we referred to such a gene as a multicopy gene (gene family).

**Western blotting**

Testis lysates were prepared by homogenizing testes in radioimmunoprecipitation assay (RIPA) lysis buffer (Pierce) containing a protease inhibitor cocktail, and protein concentration was determined using a bicinchoninic acid (BCA) assay (Pierce). Equivalent amounts of protein in each sample were resolved by SDS–polyacrylamide gel electrophoresis (PAGE) and transferred to polyvinylidene difluoride membranes (Millipore). Membranes were blocked in 5% nonfat milk and incubated at 4°C overnight with primary antibodies against the following proteins, as indicated in the text: SLY (diluted 1:2000) (51), SLX/SXL/1 (diluted 1:1000) (52), α-tubulin (diluted 1:10,000; Millipore, #05-829), HSF1 (diluted 1:100; Santa Cruz Biotechnology, #sc-17756), and HSF2 (diluted 1:200; R&D Systems, #AF5227). Membranes were subsequently incubated with horseradish peroxidase–conjugated anti-rabbit IgG (diluted 1:10,000; Promega, #W401B), anti-mouse IgG (diluted 1:1000; Santa Cruz Biotechnology, #sc-516102), or anti-goat IgG (diluted 1:10,000; Santa Cruz Biotechnology, #sc-2354) secondary antibody, as appropriate. Scans of full-size immunoblots are presented in Source Data.

**RNA pull-down assay**

RNA pull-down assays were performed using a Pierce Magnetic RNA-Protein Pull-Down Kit (Thermo Fisher Scientific) according to the manufacturer’s instructions. Briefly, RNA was transcribed in vitro using a HiScribe T7 High Yield RNA Synthesis Kit (NEB), followed by DNase I (NEB) digestion and LiCl (Sigma-Aldrich) purification. RNA was biotinylated using a Pierce RNA 3'-End Desthiobiotinylnation Kit (Thermo Fisher Scientific) according to the manufacturer’s instructions and purified using a NucAway spin column (Invitrogen). Testis extracts were prepared using the Pierce T-PER Tissue protein extraction reagent (Thermo Fisher Scientific). Fifty picomol of biotinylated RNA was captured with streptavidin magnetic beads and then incubated with 1 mg of testis lysates at 4°C for 3 hours. The beads were then washed three times with wash buffer and boiled in 2× Laemmli loading buffer. The retrieved proteins were resolved by SDS-PAGE and Western blotting.

**RNA immunoprecipitation**

RIP was performed using an EZ-Magna RIP Kit (Millipore) according to the manufacturer’s instructions. Briefly, RNAs were immunoprecipitated using 10 μg of goat anti-HSF2 antibody (R&D Systems, #AF5227); normal goat IgG (R&D Systems, #AB-108-C) was used as a control. After proteinase K digestion, immunoprecipitated RNA was extracted, purified, and used as a template for preparation of cDNA using a QuantiNova Reverse Transcription Kit (Qiagen). Enrichment of immunoprecipitated RNA was assessed by qRT-PCR using Gapdh mRNA as a negative control. The primer pairs used are shown in table S6.

**Chromatin immunoprecipitation**

ChIP assays were performed using a Magna ChIP A/G Kit (Millipore) according to the manufacturer’s instructions. In brief, adult testes from WT and NR_038002-deficient mice were collected and cross-linked with 1% formaldehyde, followed by quenching with glycine. Chromatin was sheared on ice to an average length of 200 to 1000 bp with a Vibra-Cell VCX-600 ultrasonic processor (20% amplitude, 10-s pulse on/30-s off, 20 cycles), and centrifuged at 10,000g at 4°C for 10 min. Immunoprecipitation with 5 μg of normal goat IgG (R&D Systems, #AB-108-C) and goat anti-HSF2 antibody
(R&D Systems, #AF5227), washing, and elution were carried out according to the protocol. Regarding ribonuclease (RNase) treatment, no RNase was treated before or during ChIP. Primers for the ChIP assay were designed to amplify the promoter regions of the target genes in table S6. Fold enrichment was calculated by PCR products with anti-His2 antibody to those with the normal goat IgG.

**Statistical analysis**
All experiments were performed at least in triplicate for each sample. Data are presented as the means ± SD. After analysis with F-test, statistical analysis was performed using Student’s t test, and P values were indicated in the figure legends.

**SUPPLEMENTARY MATERIALS**
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/24/eaab5177/DC1

**REFERENCES AND NOTES**


Testicular germ cell–specific lncRNA, Teshl, is required for complete expression of Y chromosome genes and a normal offspring sex ratio
Seong Hyeon Hong, Gwidong Han, Seung Jae Lee, Julie Cocquet, Chunghee Cho

Sci. Adv., 7 (24), eabg5177. • DOI: 10.1126/sciadv.abg5177