Developmental Biology

The Rapidly Evolving X-linked *miR-506* Family Finetunes Spermatogenesis to Enhance Sperm Competition

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint version 1
January 9, 2024 (this version)

Sent for peer review
August 9, 2023

Posted to preprint server
June 14, 2023

Zhuqing Wang, Yue Wang, Tong Zhou, Sheng Chen, Dayton Morris, Rubens Daniel Miserani Magalhães, Musheng Li, Shawn Wang, Hetan Wang, Yeming Xie, Hayden McSwiggin, Daniel Oliver, Shuiqiao Yuan, Huili Zheng, Jaaved Mohammed, Eric C. Lai, John R. McCarrey, Wei Yan

Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA • The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA • Department of Developmental Biology, Memorial Sloan-Kettering Institute, 1275 York Ave, Box 252, New York, NY 10065, USA • Department of Biology, University of Texas at San Antonio, San Antonio, TX, USA • Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA

Abstract

Despite rapid evolution across eutherian mammals, the X-linked *miR-506* family miRNAs are located in a region flanked by two highly conserved protein-coding genes (*Slitrk2* and *Fmr1*) on the X chromosome. Intriguingly, these miRNAs are predominantly expressed in the testis, suggesting a potential role in spermatogenesis and male fertility. Here, we report that the X-linked *miR-506* family miRNAs were derived from the MER91C DNA transposons and their sequences diverged *via* LINE1-mediated retrotransposition during evolution. Selective inactivation of individual miRNAs or clusters caused no discernable defects, but simultaneous ablation of five clusters containing nineteen members of the *miR-506* family led to reduced male fertility in mice. Despite normal sperm counts, motility and morphology, the KO sperm were less competitive than wild-type sperm when subjected to a polyandrous mating scheme. Transcriptomic and bioinformatic analyses revealed that these X-linked *miR-506* family miRNAs, in addition to targeting a set of conserved genes, have gained more targets that are critical for spermatogenesis and embryonic development during evolution. Our data suggest that the *miR-506* family miRNAs function to enhance sperm competitiveness and reproductive fitness of the male by finetuning gene expression during spermatogenesis.

Significance Statement

The X-linked *miR-506* family has rapidly evolved in mammals, but their physiological significance remains elusive. Given their abundant and preferential expression in the testis and sperm, these X-linked miRNAs likely play a functional role in spermatogenesis and/or early embryonic development. However, the deletion of either individual miRNA genes or all of the five miRNA clusters encoding 38 mature miRNAs did not cause major fertility defects in mice. When these mutant males were subjected to conditions resembling polyandrous mating, the mutant sperm were much less competitive than the wild-type sperm, rendering the mutant males “functionally infertile”. Our data suggest that the *miR-506* family of miRNAs regulates sperm competition and the reproductive fitness of the male.

eLife assessment

This study provides **important** findings on the evolution and function of the X-linked miR-506 miRNA cluster. The evidence supporting the conclusions is **convincing**, including the generation and characterization of an impressive number of the miRNA deletion mutants. This work will be of interest to RNA biologists, evolution biologists and reproductive biologists.

Introduction

Spermatogenesis is highly conserved among all vertebrates. Although it generally consists of three phases (mitotic, meiotic, and haploid), many characteristics appear to be species-specific, *e.g.*, the duration of each of the three phases, the seminiferous epithelial organization, and the shape and length of spermatozoa, likely reflecting the adaptive changes during evolution (1–3). Several cellular events are unique to the male germ cells, *e.g.*, meiosis, genetic recombination, and haploid male germ cell differentiation (also called spermiogenesis) (4). Unique cellular processes are often accompanied by a more complex yet unique transcriptome, which may explain why the testis expresses more genes than any other organs in the body, with the possible exception of the brain (5). Regulation of gene expression during spermatogenesis occurs at both transcriptional and post-transcriptional levels (6). As a post-transcriptional regulator, miRNAs are abundantly expressed in the testis and are required for spermatogenesis (7–11). miRNAs typically function at post-transcriptional levels by binding the complementary sequences in the untranslated regions (UTRs) of mRNAs – particularly in the 3'UTRs through the "seed sequence" (2^{nd}-7^{th} nucleotides)(12). Numerous miRNAs are subject to rapid evolution, probably in response to the accelerated rate of divergence of UTRs compared to the exonic sequences (13). Divergence of genomic sequences can be mediated by transposable elements (TEs), which are known as building blocks of the genome and mostly map to UTRs and intronic regions of protein-encoding genes (14). Each miRNA can bind numerous target mRNAs, and one mRNA can be targeted by multiple different miRNAs. This "one-to-multiple" relationship between miRNAs and mRNAs amplifies their potential to coordinate gene expression in the cell (12). Moreover, miRNA genes often exist in clusters, which are transcribed as a unit followed by nuclear and cytoplasmic cleavage events to generate individual miRNAs (12).

Multiple clusters of miRNA genes containing the same seed sequences are categorized as a miRNA family, and miRNAs within the same family likely evolved from a common ancestor sequence (15). Of great interest, many of the testis-enriched miRNA clusters map to the X chromosome (16). Sex-linked genes are generally subject to the male germline-specific phenomenon called meiotic sex chromosome inactivation (MSCI), which silences transcription during most, if not all, of meiosis (16). Indeed, prior to 2009, there were no confirmed reports of any sex-linked genes escaping the repressive effects of MSCI. Surprisingly, however, we found that many X-linked miRNA genes do escape MSCI, suggesting that they may contribute to particularly important functions during spermatogenesis (16). This notion has since been tested by the generation of knockouts (KOs) of individual miRNA genes or individual clusters of miRNA genes normally expressed during spermatogenesis. However, these KOs resulted in minimal, if any, phenotypic effects and did not appear to impede normal spermatogenesis or male fertility (15). This left the field facing multiple unanswered questions, including 1) why do so many X-linked miRNAs express uniquely or preferentially during spermatogenesis and escape MSCI, 2) what's their origin, and 3) how and why did they evolve rapidly?

To address these questions and better understand the functional role played by these X-linked miRNAs, we investigated the evolutionary history of this unique miRNA family and also generated KOs of individual, paired, triple, quadruple, or quintuple sets of miRNA clusters within this family and tested the effects on male fertility, initially by standard monandrous mating assays. Consistent with previous efforts to inactivate miRNA genes in *C. elegans* and mice (15, 17–19), KOs of either individual members or individual clusters of the *miR-506* family induced no discernable phenotypes and did not impact male fertility. This may reflect the level of functional redundancy inherent within the members and clusters of this miRNA family. It was only when four or more clusters of the *miR-506* family were ablated that relevant phenotype became detectable, which was manifested in the form of reduced litter size despite normal sperm counts, motility, and morphology. Interestingly, the most common male fertility testing for lab rodents is based on a monandrous mating scheme, *i.e.,* one fertility-proven female mated with one male. However, there are additional aspects of male fertility in many mammalian species, particularly those that are normally litter-bearing. In the wild, litter-bearing females often mate with multiple different males such that a single litter may include pups sired by more than one male (20, 21). This polyandrous mating introduces the potential for additional aspects of male reproductive fitness to accrue, one of which involves sperm competition (20, 21). Sperm competition can occur when sperm from more than one male are present in the female reproductive tract simultaneously, such that they then compete to fertilize each oocyte (22). Sperm competition has been now recognized as a significant evolutionary force directly impacting male reproductive success (22). Using experiments that mimic polyandrous mating, we found that the quinKO male mice indeed displayed compromised sperm competition. Hence, the X-linked *miR-506* family miRNAs appear to function to finetune spermatogenesis to enhance sperm competition and, consequently, male reproductive fitness.

Results

X-linked *miR-506* family miRNAs flanked by two highly conserved protein-coding genes *Slitrk2* and *Fmr1* rapidly evolved across species

X-linked genes are generally more divergent between species than autosomal ones, a phenomenon known as the “faster-X effect” (23). However, despite a high degree of conservation of two protein-coding genes, *Slitrk2* and *Fmr1*, on the X chromosome across species, the miRNA genes located between these two loci are divergent among clades across the eutherian mammals (15, 24). Through tracing the evolution of this genomic region, we found that *Slitrk2* and *Fmr1* were syntenically linked to autosomes in zebrafish and birds (**Fig. 1A**), but had migrated onto the X chromosome in most mammals, with divergence and multiplication of numerous miRNA genes that belong to the *miR-506* family in between (**Fig. 1A** and Table S1). By mapping these miRNAs of various species using the UCSC genome browser (25), we found that all members of the *miR-506* family are located in a region flanked by *Slitrk2* and *Fmr1* (**Fig. 1A**). Consistent with previous reports (15, 24), *Slitrk2* and *Fmr1* are usually on the positive strand, whereas the *miR-506* family miRNAs are in the reverse orientation (**Fig. 1A**). Based on the location of these miRNAs, we named the miRNAs proximal to *Slitrk2* (*miR-892c ~ miR-891a* in humans) and *Fmr1* (*miR-513c ~ miR-514a3* in humans) SmiRs (*Slitrk2*-proximal miRNAs) and FmiRs (*Fmr1*-proximal miRNAs), respectively.

To evaluate the sequence conservation of these miRNAs across species, we adopted the Multiz Alignment and Conservation pipeline, which utilizes PhastCons and PhyloP algorithms (25), to search miRNA datasets from 100 different species using the human genome as a reference (**Fig. 1B** and Fig. S1). The mean values of phyloP and phastCons of the *Fmr1* and *Slitrk2* coding sequences (CDS) were ~4.9 and ~0.97, respectively (**Fig. 1C** and **D**), indicating that these

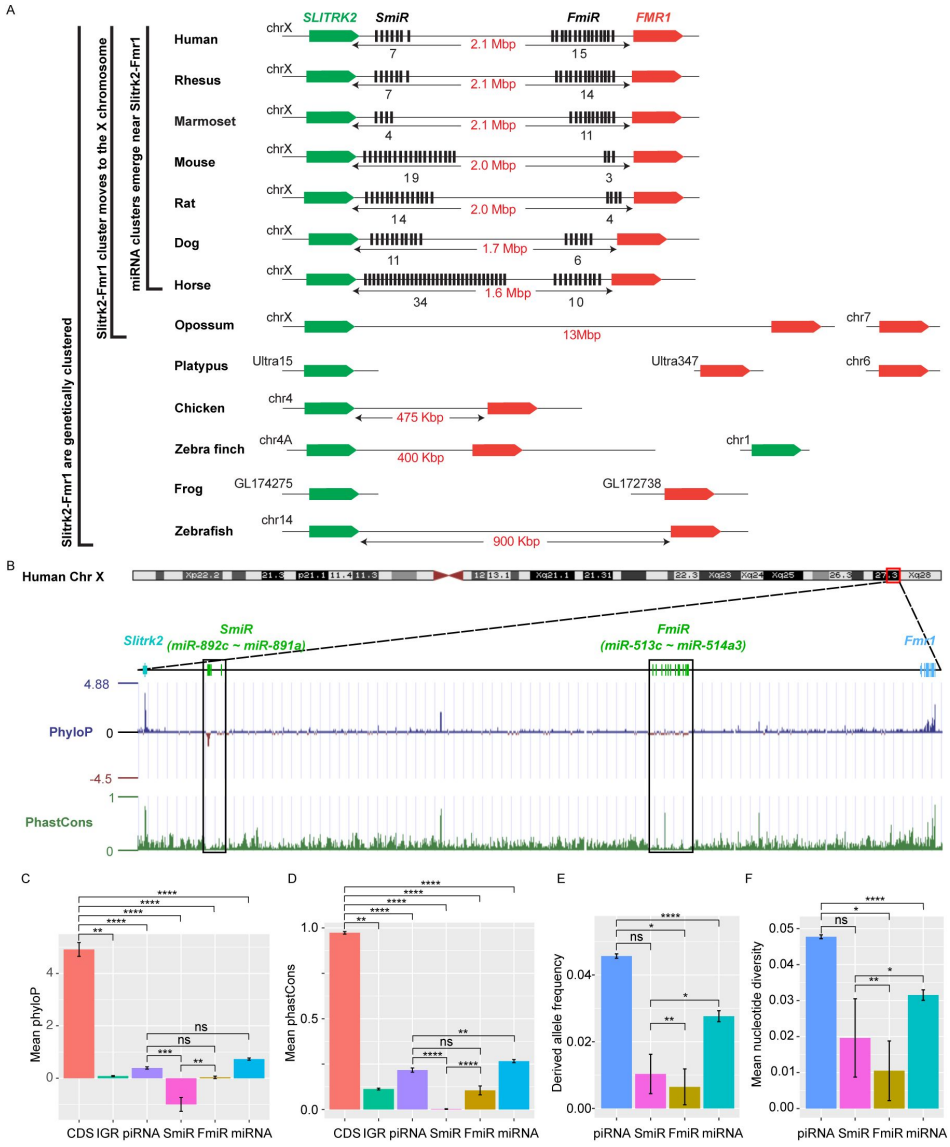

Figure 1.

Genomic location, sequence alignment and evolution conservation of the X-linked *miR-506* family.

A. Genomic location of the X-linked *miR-506* family miRNAs (black bars) and the two flanking coding genes, *Slitrk2* (green blocks) and *Fmr1* (red blocks). The number of miRNAs within each cluster is indicated underneath the miRNA clusters. B. Evolution conservation of X-linked *miR-506* family based on Multiz Alignment & Conservation using the human genome as a reference. Positive PhyloP scores indicate conservation and vice versa. PhastCons has a score between 0~1, and the higher the score, the more conserved the DNA region is. C-D. Comparison of mean phyloP (C) phastCons (D) scores among CDS of *Slitrk2* and *Fmr1*, intergenic region (IGR), pachytene piRNAs, SmiR, FmiR, and all miRNAs. **, ***, **** indicate adjusted p-value < 0.01, 0.001, and 0.0001, respectively. ns, not significant. Kruskal-Wallis test was used for statistical analyses. E-F. Comparison of derived allele (E) and mean nucleotide (F) frequencies among pachytene piRNAs, SmiR, FmiR, and all miRNAs. *, ** and **** indicate adjusted p-value < 0.05, 0.01 and 0.0001, respectively. ns, not significant. Kruskal-Wallis test was used for statistical analyses.

regions are highly conserved. In contrast, the mean values of phyloP and phastCons of SmiRs were ~1.0 and ~0.002, respectively, and those of FmiRs are ~0.03 and ~0.11, respectively (**Fig. 1 C** and **D**). The phyloP and phastCons values of the CDS, FmiRs, and SmiRs are significantly different from each other (adjusted p-value < 0.05, Kruskal-Wallis test) (**Fig. 1 C** and **D**), indicating that FmiRs and SmiRs are highly divergent, and SmiRs are more divergent than FmiRs.

We then assessed the genomic sequence similarity among various species using D-GENIES-based dot plot analyses (26). Although the *Slitrk2-Fmr1* genomic regions were highly variable among different species, the sequences within some clades shared a high degree of similarities, e.g., primates (rhesus monkeys, chimpanzees, and humans), *cetartiodactyla* (sheep and cows), *rodentia* (e.g., mice and rats) and *carnivora* (e.g., dogs and cats) (Fig. S2A). Although the *miR-506* family miRNAs were highly divergent, some orthologs displayed a higher degree of sequence conservation (Fig. S2 B and C), e.g., the SmiRs within primates were similar (Fig. S2B); *miR-891a* and *miR-891b* in rhesus monkeys were similar to *miR-891b* and *miR-891a* in humans and chimpanzees, respectively, and *miR-892b* in chimpanzees was homologous to *miR-892c* in humans and rhesus monkeys in terms of sequence and location (Figs. S2B and S3). The FmiRs, including *miR-201* (assigned as *Mir-506-P1* (paralogue 1) in MirGeneDB (27)), *miR-547* (*Mir-506-P2*), and *miR-509* (*Mir-506-P7*) in mice and rats, are orthologues of *miR-506*, *miR-507* and *miR-509* in humans, respectively (Fig. S2C, and Table S1). Of interest, although the *miR-506* family miRNAs are highly divergent, the seed sequences of some miRNAs, such as *Mir-506-P6* and *Mir-506-P7*, remain conserved (Fig. S2C), and these miRNAs represent the dominant mature miRNAs (28). It is noteworthy that the majority of the substitutions among the *miR-506* family are U➔C and A➔G (Fig. S2 B and C). Furthermore, we analyzed the conservation of the *miR-506* family in modern humans using data from the 1000 Genomes Project (1kGP) (**Fig. 1 E** and **F**) (29). We compared SmiRs, FmiRs, and all miRNAs with pachytene piRNAs, which are known to be highly divergent in modern humans but barely exert any biological functions (30). Of interest, the derived allele frequency (DAF) and mean nucleotide diversity (MND) of FmiRs and all miRNAs were significantly smaller than that of the pachytene piRNAs, whereas SmiRs were significantly smaller than all miRNAs (**Fig. 1 E** and **F**) (adjusted p-value < 0.05, Kruskal-Wallis test), suggesting that the *miR-506* family miRNAs are more conserved than pachytene piRNAs in modern humans. Taken together, these data indicate that the X-linked *miR-506* family, although rapidly evolving as a whole, contains both divergent and conserved miRNAs, suggesting both conserved and novel functions across species.

X-linked *miR-506* family miRNAs are derived from MER91C DNA transposons and expanded *via* LINE1 retrotransposition

To visualize the family history of these miRNAs, we built a phylogram for the *miR-506* family (Fig. S4). The phylogram suggests that these miRNAs shared a common ancestor and that the FmiRs emerged earlier than the SmiRs, which is also supported by the fact that some FmiRs exist in green sea turtles (Figs. S1 and S4). These data suggest that the *miR-506* family miRNAs arose much earlier than previously thought (25). The two subfamilies, FmiRs and SmiRs, despite their common ancestors, may have evolved at different paces and thus, might be functionally divergent.

Studies have shown that transposable elements (TEs) drive evolution through transpositions (31). CRISPR-Cas9/Cas12a genome editing can induce irreversible small indels at the cutting sites (also called “scars”)(32), which have been used for lineage tracing (33). Inspired by this strategy, we attempted to trace the evolution of the *miR-506* family miRNAs by searching the transposon database for the transpositional “scars” (partial TE sequences) after transposition. To search the potential TE sources of the *miR-506* family miRNAs, we downloaded all transposons in the human, horse, dog, and guinea pig genomes and aligned them to their corresponding *miR-506* family miRNAs using BLAST (Basic Local Alignment Search Tool) (34). The MER91C DNA transposon (~100-150 million years) (35) was the only transposon that aligned to FmiRs of the *miR-506* family (> 94% identical matches) in all four species (Table S2). The remaining miRNAs of

the *miR-506* family, mainly SmiRs, are flanked by LINEs (long interspersed elements) or SINEs (short interspersed element, including Alu) (Fig. S3) retrotransposons (Fig. S3 and Table S2). LINE1 is the most abundant and the only active TE in the human genome (36). Although SINEs cannot mobilize by themselves, they can hijack the LINE1 mechanism to achieve transpositions (36).

Given that the FmiRs (e.g., *miRs-506~509*) emerged much earlier than the SmiRs (fig. S1 and fig. S4) and that *miR-513* (belonging to FmiRs) and SmiRs (including *miR-891a* and *miR-891b*) share a common ancestor (Fig. S4), we reasoned that the X-linked *miR-506* family might be derived from the *MER91C* DNA transposon and further expanded through LINE1-mediated retrotransposition. To test this hypothesis, we first aligned the X-linked *miR-506* family miRNAs from several species to a human MER91C DNA transposon. Indeed, numerous FmiRs of almost all species analyzed aligned to the MER91C DNA transposon despite few mismatches (**Fig. 2A**), phylogenetic tree further confirmed that the MER91C is the sister group of the *miR-506* family miRNAs (**Fig. 2B** and Fig. S5), suggesting that the MER91C DNA transposon is the likely source of the older *miR-506* family miRNAs. Further supporting this notion, the MER91C DNA transposons could form hairpin structures, which is a prerequisite for miRNA biogenesis (**Fig. 2C** and Fig. S6A). Moreover, analyses of the testis small RNA datasets from humans, marmosets, dogs, and horses revealed the peaks corresponding to these miRNAs (**Fig. 2D** and Fig. S6A). Finally, by overexpressing several MER91C DNA regions randomly selected from humans, dogs, and horses in HEK293T cells, we found that these DNA regions were indeed capable of producing miRNAs (**Fig. 2 E** and **F**, and Fig. S6 B and C). Co-expression of MER91C DNA regions and AGO2 significantly increased the abundance of human MER91C miRNAs, as compared with overexpression of MER91C DNA regions alone (**Fig. 2** E and F), suggesting that these miRNAs could be loaded onto and protected by AGO2. Together, these data indicate that the MER91C DNA transposon serves as a source for the *miR-506* family miRNAs. As to the rest of the *miR-506* family miRNAs, we first compared the proportion of TEs in the miRNA-flanking regions (length of different classes of TEs/length of the genomic region) among humans, chimpanzees, horses, mice, and rats (**Fig. 2 G** and Fig. S6D). LINE1 retrotransposons were much more abundant in the *FmiR* genomic region when compared to either all the chromosomes or the *FmiR* proximal gene *Fmr1* (adjusted p-value < 0.05, one-way ANOVA). Similarly, LINE1s were more enriched in the *SmiR* region than in the *SmiR* proximal gene *Slitrk2* (adjusted p-value < 0.05, one-way ANOVA) (**Fig. 2 G**). Since the length of LINE1 retrotransposons (~6kb) is intrinsically longer than that of the SINEs (100-600bp), which may bias the results, we further calculated proportions of the nearest TE classes for the *miR-506* family and all miRNAs (**Fig. 2 H** and Fig. S6E and Table S2). In humans, chimpanzees, horses, mice, and rats, the *miR-506* family miRNAs displayed a significantly higher proportion of LINE1 when compared to all miRNAs (baseline) ($p < 0.05$, paired t-test) (**Fig. 2 H** and Table S2), suggesting that LINE1 may involve in *miR-506* family expansion. Taken together, these results suggest that the X-linked *miR-506* miRNAs were originally derived from the MER91C DNA transposon and further expanded through LINE1-mediated retrotransposition.

X-linked *miR-506* family miRNAs are predominantly expressed in spermatogenic cells and sperm

Several previous studies have shown that the X-linked *miR-506* family miRNAs are predominantly expressed in the testis of multiple species (15, 16, 24, 37–39). By analyzing the publicly available small RNA sequencing (sRNA-seq) datasets from multiple species, including humans, rhesus monkeys, mice, rats, rabbits, dogs, and cows (27, 40, 41), we further confirmed that *miR-506* family miRNAs were indeed highly abundant in the testis, but barely expressed in other organs (Fig. S7, and Table S3). To further determine whether these miRNAs are expressed in male germ cells in rodent testes, we conducted sRNA-seq using pachytene spermatocytes, round spermatids, and sperm purified from adult mice (Fig. S8A and Table S3). Consistent with previous data (15, 16, 24), these miRNAs were abundantly expressed in spermatogenic cells in murine testes (**Fig. 3A**). ~80% of these miRNAs were significantly

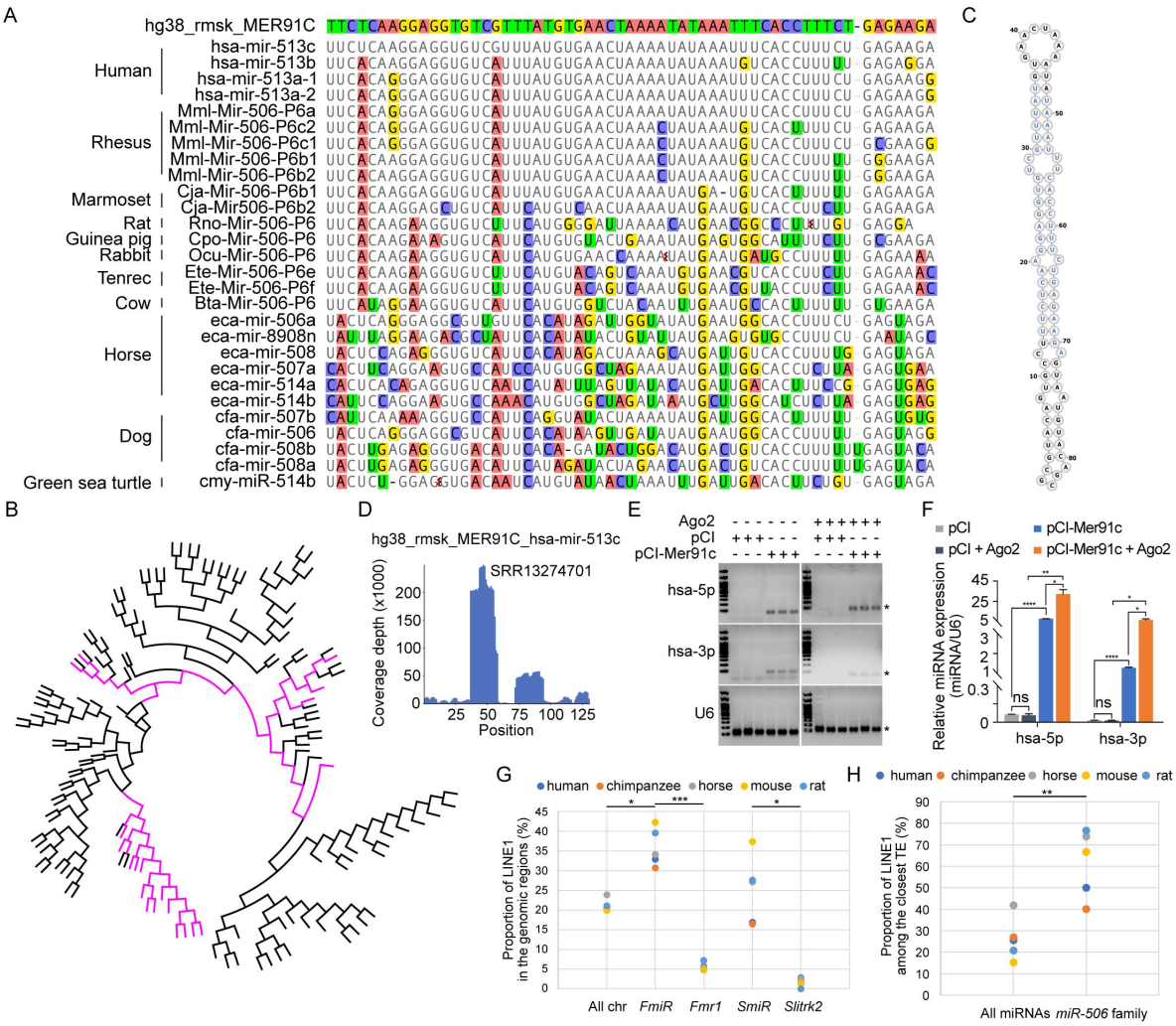

Figure 2.

Evolutionary history of the X-linked *miR-506* family.

A. Sequences alignment of FmiRs from various species using human MER91C DNA transposon as the reference. The first line is the human MER91C DNA transposon, and below are the miRNAs of various species. Mismatched nucleotides are highlighted with various colors. B. A phylogenetic tree of the MER91C DNA transposons and the X-linked *miR-506* family miRNAs. The MER91C DNA transposons are labeled in purple. C. RNA structure of the MER91C DNA transposon-derived miRNA (hsa-miR-513c). D. sRNA-seq reads (lower panel) of the MER91C DNA transposon-derived miRNA, hsa-miR-513c. E. Representative gel images showing expression levels of the MER91C DNA transposon-derived miRNA (hsa-miR-513a1) in HEK293T cells. The asterisk (*) indicates the expected miRNA size. U6 was used as the loading control. F. qPCR analyses of expression levels of MER91C DNA transposon-derived miRNA (hsa-miR-513a1) in HEK293T cells. *, **, **** indicate adjusted p-value < 0.05, 0.01, 0.0001, respectively. One-way ANOVA was used for statistical analyses. G. The proportions of LINE1 in the genomic regions of all chromosomes, *Fmr1*, *FmiR*, *Slitrk2*, and *SmiR* in humans, chimpanzees, horses, mice, and rats. * and *** indicate adjusted p-value < 0.05 and p < 0.001, respectively. One-way ANOVA was used for statistical analyses. H. The proportion of LINE1 between the closest TEs to all miRNAs and to *miR-506* family miRNAs. ** indicates p < 0.01. Paired t-test was used for statistical analyses.

upregulated (FDR < 0.05) when pachytene spermatocytes developed into round spermatids, and ~83.3% were significantly upregulated (FDR < 0.05) when developed into cauda sperm (**Fig. 3A**). By analyzing the publicly available sRNA-seq datasets from humans (42), marmosets (37) and horses (43), we determined the expression patterns of the *miR-506* family miRNAs in the testes of these species (**Fig. 3 B, C** and **D**, Table S3). The significantly increasing abundance of the SmiRs from immature to mature testes in horses (43) supports the elevated expression in haploid male germ cells (round, elongating/elongated spermatids, and sperm) compared to meiotic male germ cells (spermatocytes) (**Fig. 3D**). More interestingly, the SmiRs and FmiRs appear to be differentially expressed in the testes of various species, e.g., relative levels of the SmiRs were greater than those of the FmiRs in mice (**Fig. 3 A** and Table S3). Both the SmiRs and FmiRs were highly expressed in horses (**Fig. 3D** and Table S3). Levels of the FmiRs in marmoset (37) and human testes (42) were much greater than in those of the SmiRs (**Fig. 3 B** and **C** and Table S3). Overall, the *miR-506* family miRNAs are abundant in the testis and predominantly expressed in haploid male germ cells, *i.e.*, spermatids and spermatozoa.

Ablation of X-linked *miR-506* family miRNAs compromises male fertility due to reduced sperm competitiveness

To define the physiological role of the *miR-506* family, we sequentially deleted these miRNA genes using CRISPR-Cas9-based genome editing (**Fig. 4A**) (15). We first generated the KO mice lacking either the *miR-883* single cluster (*miR-883* sKO) or the *miR-465* single cluster (*miR-465* sKO) (**Fig. 4A**), as these two clusters are the most abundantly expressed in the mouse testes (**Fig. 3 A** and Fig. S7). No discernable defects were observed, and these KO males developed normally and were fertile (**Fig. 4 B** and **C**). On the *miR-883* sKO background, we further deleted the *miR-741* cluster, which we termed double KO (dKO) (**Fig. 4 A**), but no discernable abnormalities were observed in the dKO males either (Fig. S8 B-E). On the dKO background, we next deleted either the *miR-465*, termed triple KO (tKO), or the *miR-471* and *miR-470* clusters, termed quadruple KO (quadKO) (**Fig. 4 A**). Lastly, we ablated the *miR-465* cluster on the quadKO background, named quintuple KO (quinKO) (**Fig. 4 A**). To reduce potential off-target effects due to multiple rounds of CRISPR-Cas9 targeting, we also generated a KO mouse line with only 4 guide RNAs (gRNAs) flanking the SmiR region, named X-linked SmiR KO (XS) (**Fig. 4 A**). The XS mice were genetically equivalent to the quinKO mice, and phenotypically identical to quinKOs (**Fig. 4 B** and **C**), suggesting the phenotype observed was not due to the accumulating off-target effects. To further exclude the potential off-target effect, all KO mouse strains were backcrossed with WT C57BL/6J mice for at least 5 generations before data collection. In addition, T7 endonuclease I (T7EI) assays showed no discernable off-target effects in the quinKO mice (Fig. S8F).

While the litter size was still comparable between the tKO and WT control mice, the quadKO, quinKO, and XS males produced significantly smaller litters (~5 *vs.* ~8 pups/litter) (adjusted p-value < 0.05, One-Way ANOVA) (**Fig. 4 B**). Of interest, no significant changes were detected in litter interval, testis weight or histology in any of the four types of KOs, as compared to WT controls (**Fig. 4 C, D** and **G**). Computer-assisted sperm analyses (CASA) revealed no significant differences in sperm counts and motility parameters among the four types of KOs (**Fig. 4 E** and **F** and Fig. S8H). Overall, there appears to be an inverse correlation ($R^2 = 0.9139$, $p < 0.05$, F-test) between the number of miRNAs inactivated and the litter size (Fig. S8G). Interestingly, several human studies have correlated the dysregulated *miR-506* family miRNAs with impaired male fertility due to maturation arrest and oligo-asthenozoospermia (Table S4)(44–48). These data suggest that the *miR-506* family may play an important role in spermatogenesis and male fertility.

Most of the protein-coding genes that are exclusively or preferentially expressed in the testis with an essential role in spermatogenesis are highly conserved across species (49). Despite their male germ cell-predominant expression, the *miR-506* family miRNAs appear to have evolved rapidly to diverge their sequences, suggesting that these miRNAs might control certain “non-conserved”

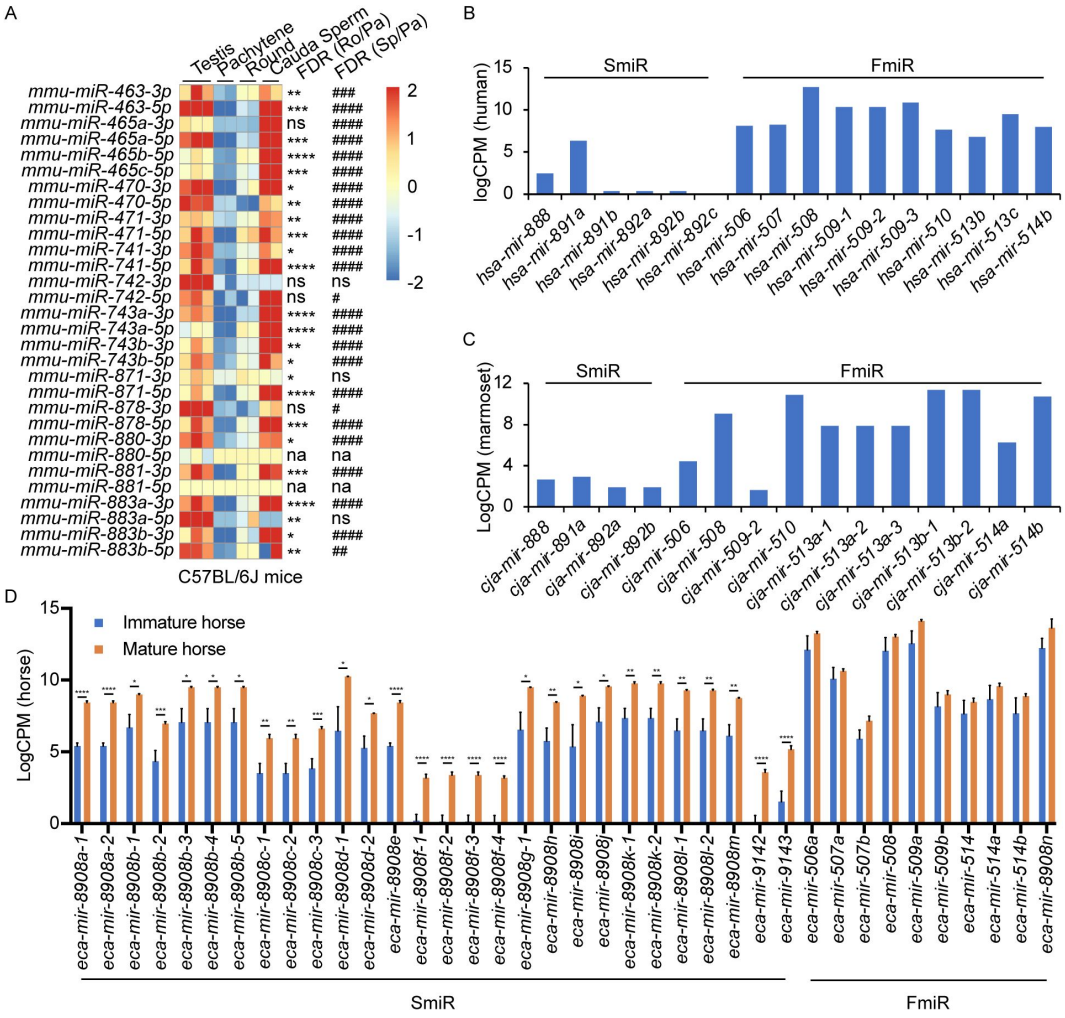

Figure 3.

Expression profiles of X-linked *miR-506* family in mammalian testes and male germ cells.

A. Heatmaps showing the *miR-506* family expression in the testis, pachytene spermatocytes, round spermatids and sperm in mice. Biological triplicates of the testis samples (n=3) and duplicates of pachytene spermatocytes, round spermatids and sperm samples isolated from 2-4 mice were used for sRNA-seq. *, **, ***, and **** indicate FDR < 0.05, 0.01, 0.001, and 0.0001, respectively, when comparing round spermatids to pachytene spermatocytes. #, ##, ###, and #### indicate FDR < 0.05, 0.01, 0.001, and 0.0001, respectively, when comparing cauda sperm to pachytene spermatocytes. ns and na indicate not significantly and not not applicable, respectively.B-C. LogCPM bar graphs showing the *miR-506* family expression in the testis of humans (B) and marmosets (C). D. LogCPM bar graph showing the *miR-506* family expression in sexually immature and mature horse testes. n=3. *, **, ***, and **** indicate FDR < 0.05, 0.01, 0.001, and 0.0001, respectively.

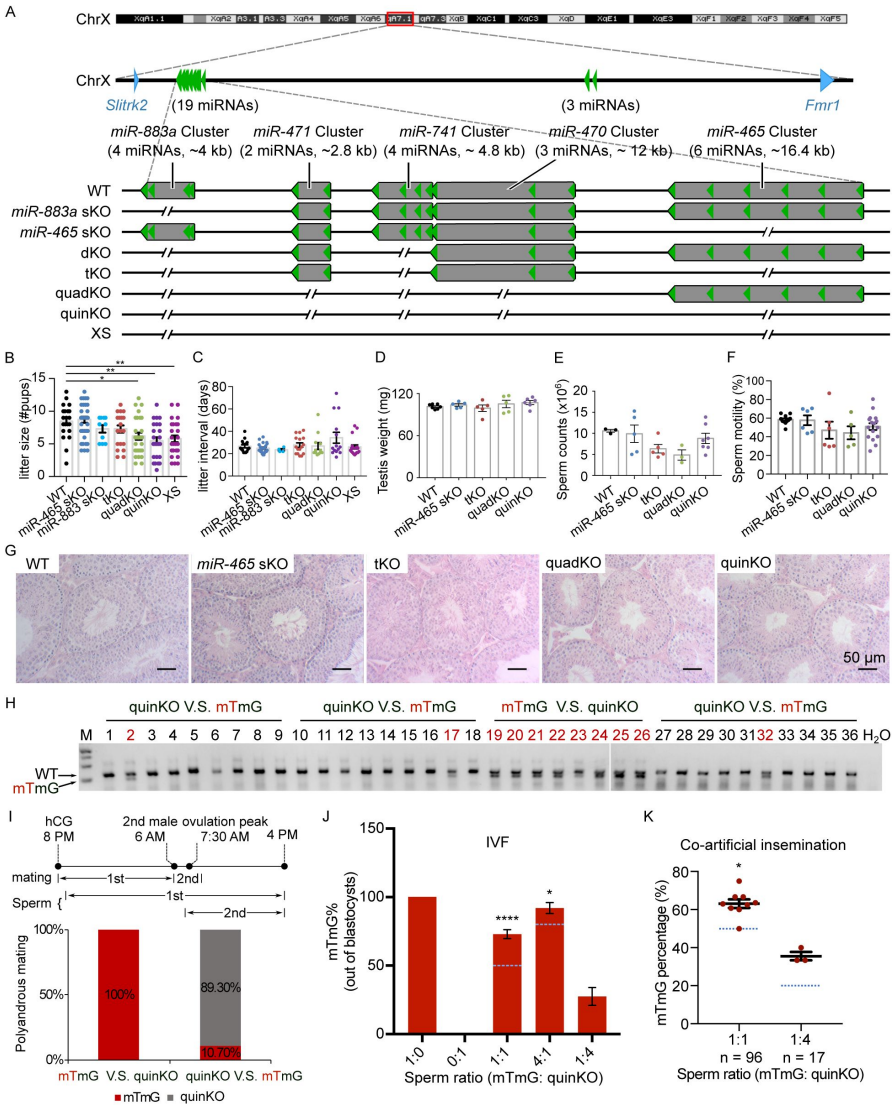

Figure 4.

Ablation of X-linked *miR-506* family miRNAs compromised sperm competitiveness and reproductive fitness in male mice.

A. Schematics showing the strategy used to generate six lines of KO mice lacking individual or combined miRNA clusters within the *miR-506* family using CRISPR-Cas9. B-C. Litter size (B) and litter interval (C) of six *miR-506* KO lines, at least 10 litters from 3 different breeding pairs for each KO line were counted. Dunnett's multiple comparisons test as the posthoc test following one-way ANOVA was used for the statistical analysis. ns, not significant. * and **indicate adjusted p-value <0.05 and 0.01, respectively. D-F. Analyses of testis weight (D), sperm counts (E) and sperm motility (F) in four *miR-506* KO lines. n ≥ 3 and Dunnett's multiple comparisons test as the posthoc test following one-way ANOVA was used for the statistical analysis. G. Testicular histology of WT and four *miR-506* KO lines showing largely normal spermatogenesis. Scale bars = 50 μm. H. Representative genotyping results of the sequential mating experiments. I. Sequential mating of WT female mice with mTmG and quinKO males. Upper panel, an overview of the polyandrous mating scheme. "mTmG V.S. quinKO": mTmG male mice mated first; "quinKO V.S. mTmG": quinKO male mice mated first. J. Percentage of mTmG blastocysts obtained from IVF using WT MII oocytes and mixed sperm from mTmG (control) and quinKO males at different ratios. Data were based on three independent IVF experiments. The expected ratio was indicated as the blue line. Chi-squared test was used for statistical analyses. * and **** indicate p<0.05 and 0.0001, respectively. K. Percentage of mTmG embryos obtained from co-artificial insemination using different ratios of mTmG and quinKO sperm. Data were based on 9 and 3 independent AI experiments for the 1:1 and 1:4 sperm ratio (mTmG: quinKO), respectively. The expected ratio was indicated as the blue line. Chi-squared test was used for statistical analyses. *, p<0.05.

aspects of spermatogenesis, leading to enhanced sperm competitiveness for male reproductive success. Supporting this hypothesis, previous reports have documented that females of most species throughout the animal kingdoms mate with multiple males before pregnancy, suggesting that sperm competition may serve as a selection mechanism to bias the birth of offspring sired by the males with more competitive sperm (20, 21). Studies have also shown female rodents in the wild mate with multiple males and produce litters of mixed paternity, and that pups born to the females following such polyandrous mating display greater survival rates than those produced from females following monandrous mating (50). Given that CASA detected no difference in swimming patterns between quinKO and WT sperm (Fig. S8H), we next carried out sperm competition experiments that mimic polyandrous mating in the wild. Since the *miR-506* family miRNAs are X-linked, the Y sperm from the quinKO mice are genetically indistinguishable from those of WT controls. We, therefore, adopted the mTmG male mice (51) for sperm competition experiments because the embryos or offspring fathered by the mTmG males can be easily identified based on the constitutively expressed membrane-tagged tomato red (mT) fluorescence and/or PCR genotyping.

We first conducted sequential mating with two mating events ~6-8 hours apart. Interestingly, all of the pups born were fathered by mTmG males (n=8) when the WT females were mated first with mTmG males and subsequently with the quinKO males. In contrast, when the WT females were mated first with quinKO males and subsequently with mTmG males, ~89% of the pups born were fathered by quinKO males, and the remaining ~11% of pups were from mTmG males (n=28) (**Fig. 4 H-I**). It is noteworthy that in the sequential mating experiments, the two coituses occurred ~6-8 hours apart due to practical reasons, whereas in the wild, polyandrous mating may take place much faster. To better mimic polyandrous mating *in vitro*, we mixed the WT and quinKO sperm in different ratios and used the mixed sperm to perform *in vitro* fertilization (IVF). MII oocytes fertilized by mTmG, quinKO, or a mixture of two types of sperm at three ratios (mTmG: quinKO = 1:1, 4:1 and 1:4) all displayed comparable rates at which fertilized oocytes developed into blastocysts (Fig. S8I). Interestingly, when a 1:1 ratio (mTmG sperm: quinKO sperm) was used, ~73% of the resulting blastocysts were derived from mTmG sperm, whereas the remaining ~27% were from quinKO sperm (n=179) ($p < 0.0001$, Chi-squared test) (**Fig. 4 J**). When a 4:1 sperm ratio (mTmG: quinKO) was used, ~92% of the blastocysts were from mTmG sperm and only 8% were from quinKO sperm (n=170) ($p < 0.05$, Chi-squared test) (**Fig. 4 J**). In contrast, when a 1:4 sperm ratio (mTmG: quinKO) was used, blastocysts derived from mTmG and quinKO sperm represented ~28% and ~72% of the total, respectively (n=135) (**Fig. 4 J**). We also performed co-artificial insemination (AI) using mTmG and quinKO sperm. When a 1:1 sperm ratio (mTmG: quinKO) was used, ~62.5% of the embryos (n=96) were derived from mTmG sperm ($p< 0.05$, Chi-squared test) (**Fig. 4 K**). When a 1:4 ratio was used, ~35.3% of the embryos (n=17) were from the mTmG mice (**Fig. 4K** and Fig. S8J). Together, these results indicate that the quinKO sperm are less competitive than the control mTmG sperm both *in vivo* and *in vitro*. Previous studies suggest that sperm aggregation and midpiece size might be involved in sperm competitiveness (52, 53), but no changes in these two parameters were observed in the quinKO sperm (Fig. S8 K and L). Although the blastocyst rate (out of 2-cell embryos) of quinKO was comparable to that of the mTmG mice, the 2-cell rates (out of zygotes) were significantly reduced ($p< 0.05$, paired t-test) in the quinKO mice (~39%, n=67) when compared to the mTmG mice (~88%, n=58) (Fig. S8M), implying that the quinKO sperm are indeed less efficient in fertilizing eggs and/or supporting early embryonic development, especially the first cleavage of the zygotes.

X-linked *miR-506* family miRNAs mostly target the genes involved in spermatogenesis and embryonic development and compensate for each other

To identify the target genes of these X-linked *miR-506* family miRNAs, we performed RNA-seq analyses using testis samples from the five types of KOs (*miR-465* sKO, dKO, tKO, quadKO, and quinKO) (Fig. S9A and Table S5). Comparisons between the KO and WT testes revealed thousands

of differentially expressed genes (DEGs) (fold change ≥ 2, FDR< 0.05, Fig. S9A and Table S5). The DEGs identified were then compared with the predicted *miR-506* target genes using four different databases, including TargetScan (54), microrna.org (55), miRWalk (56) and mirDB (57), to predict the differentially expressed targets (DETs) of the *miR-506* family miRNAs (Fig. S9A and Table S5). We obtained 2,692, 2,028, 1,973, 3,405, and 1,106 DETs from *miR-465* sKO, dKO, tKO, quadKO and quinKO testes, respectively. GO terms of DETs from each KO testis revealed that the DETs were mostly involved in embryonic development, response to stimulus, centrosome cycle, epithelium morphogenesis, organelle organization, cell projection, RNA metabolic process, and DNA repair (Fig. S9B). The 431 DETs identified to be shared across all five KO testes were also enriched in similar pathways (**Fig. 5** An and B). Several genes, including *Crisp1*, *Egr1*, and *Trpv4*, were selected for validation using qPCR, Western blots and luciferase-based reporter assays. Consistent with the RNA-seq data, qPCR showed that *Crisp1*, *Egr1*, and *Trpv4* were significantly downregulated in the quinKO testes (Fig. S9C). Western blots also confirmed that *CRISP1* is downregulated in the quinKO testis when compared to the WT testis (Fig. S9D). Luciferase assays further confirmed that *Egr1* and *Crisp1* are targets of the *miR-506* family members (Fig. S9 E and F). *Egr1* 3'UTR luciferase activity was upregulated by *miR-465c*, while downregulated by *miR-743b* (fig. S9E). *miR-465a*, *miR-465c*, *miR-470*, *miR-741*, and *miR-743a* upregulated *Crisp1* 3'UTR luciferase activity, while *miR-743b* exerted the opposite effect (Fig. S9F). Of interest, KO of *Crisp1* in mice or inhibition of *CRISP1* in human sperm appears to phenocopy the quinKO mice (58, 59), e.g., sperm motility in the *Crisp1* KO mice is comparable to that in WT mice, but their ability to penetrate the eggs was reduced in the *Crisp1* KO mice (58); a similar effect was also observed in human sperm treated with anti-hCRISP1 antibody (59).

The inverse correlation between the number of miRNAs inactivated and the severity of the phenotype strongly hints that these miRNAs compensate for each other (**Fig. 4B** and Fig. S8G). To test this hypothesis, we performed sRNA-seq on four KO (*miR-465* sKO, tKO, quadKO and quinKO) testes. The sRNA-seq data showed that these miRNAs were no longer expressed in the corresponding KOs, confirming the successful deletion of these miRNAs in these KOs (**Fig. 5C** and Table S6). Interestingly, in *miR-465* sKO testes, *miR-201*, *miR-547*, *miR-470*, *miR-471*, *miR-742*, *miR-871*, *miR-881*, *miR-883a*, and *miR-883b* were all significantly upregulated (FDR < 0.05). Similarly, *miR-201*, *miR-547*, *miR-470*, *miR-471*, *miR-871*, and *miR-883b* were all significantly upregulated in the tKO testes (FDR < 0.05); *miR-201* and *miR-547* were all significantly upregulated in the quadKO and the quinKO testes (FDR < 0.05) (**Fig. 5 C** and Table S6). These results support the notion that genetic compensation exists among the X-linked *miR-506* family miRNAs.

Rapid evolution of the *miR-506* family is not driven by the increased complexity of 3'UTRs of the conserved targets but rather adaptive to targeting more genes

To minimize false positives, the DETs in mice were selected using the following criteria: (1) dysregulated by fold change ≥ 2 and FDR < 0.05. (2) Falling within the predicted targets. (3) Intersected with at least two different KO mouse samples. Using the 3,043 DETs identified in mice as a reference, we searched the predicted targets of the *miR-506* family miRNAs in rats and humans to determine if these target genes were shared across species (**Fig. 5A**). While 2,098 (~69%) target genes were shared among all three species, 2,510 (~82%) were common to both humans and mice, and 2,202 (~72%) were shared between mice and rats (**Fig. 6 A** and Table S7). To test the accuracy of the predicted targets, we selected several genes in humans and performed luciferase assays using human *miR-506* family miRNAs and their corresponding target genes (Fig. S10 A and B). Among these targets, *Crisp1* and *Fmr1* were shared among humans, mice, and rats, and confirmed to be targeted by the *miR-506* family miRNAs in mice (fig. S9C, S9D, S9F, and ref 11 and 15)(11, 15). Luciferase assays also confirmed that human *CRISP1* (*hCRISP1*) and *FMR1*(*hFMR1*) were targets of the *miR-506* family, and *miR-510* and *miR-513b* both could activate

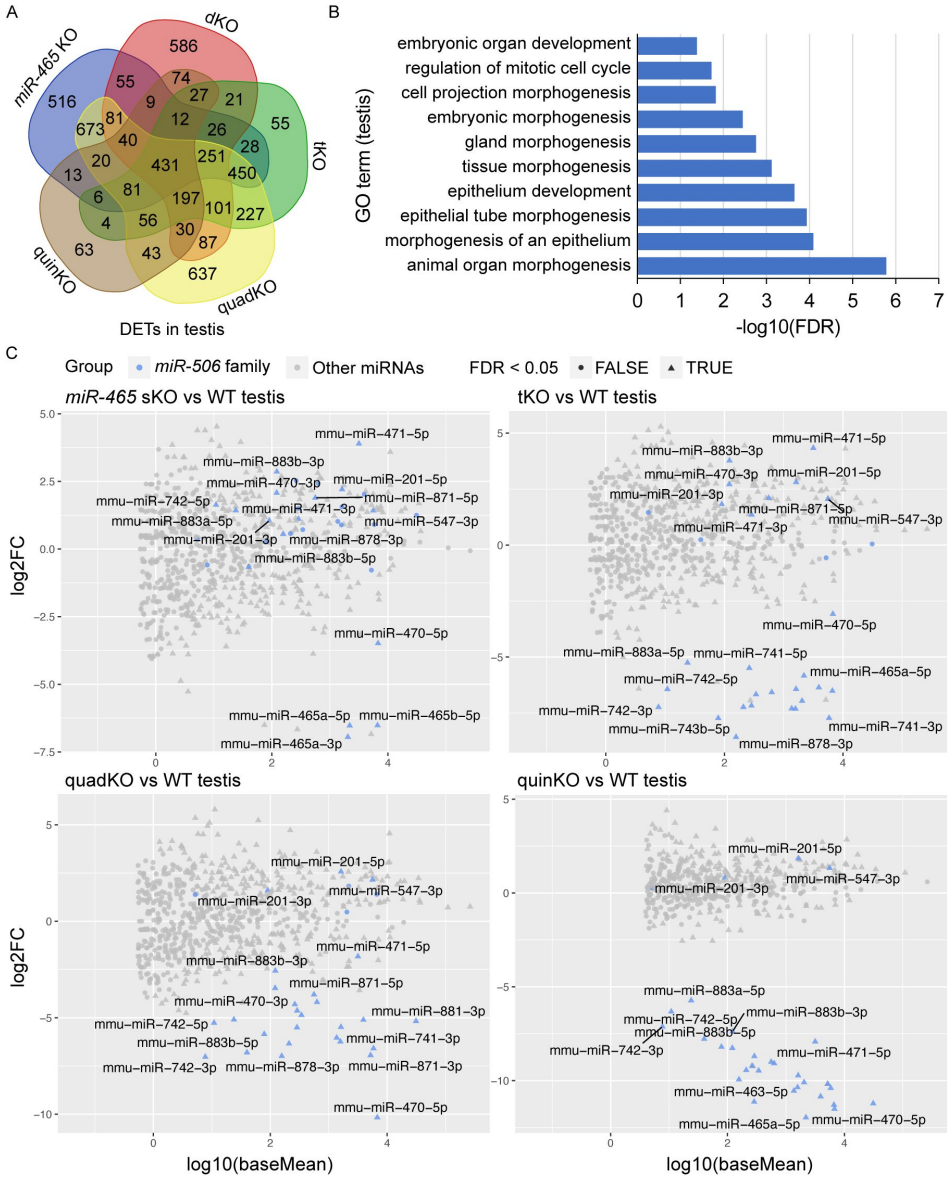

Fig. 5.

Target genes and genetic compensation of the X-linked *miR-506* family miRNAs.

A. Intersections of the DETs among different KO testes. B. GO term enrichment analyses of the 431 DETs shared among the four different miR-506 KO testes. C. MA plots showing the expression levels of the miR-506 family miRNAs in WT, sKO, tKO, quadKO, and quinKO testes. Three biological replicates (n=3) were used for sRNA-seq analyses.

hCRISP1 3'UTR luciferase activity (Fig. S10A), whereas *miR-509-1*, *miR-509-2*, *miR-509-3*, *miR-513b*, *miR-514a*, and *miR-514b* could enhance *hFMR1* 3'UTR luciferase activity (Fig. S10B). These results confirmed the accuracy of our predicted targets in humans.

We considered two likely explanations for the paradox where the majority of their target genes were shared across species despite the rapid evolution of the *miR-506* family miRNAs: 1) The 3'UTR sequences in extant target genes became increasingly divergent during evolution such that the *miR-506* family miRNAs had to adapt to maintain their ability to bind these 3'UTRs; or 2) that the *miR-506* family miRNAs evolved rapidly in a manner that allowed them to target mRNAs encoded by additional genes involved in spermatogenesis. To distinguish the two possibilities, we first compared the extent of similarities among the 3'UTR sequences of the 2,510 shared target genes between humans and mice (**Fig. 6 A** and Table S7). We adopted the PhyloP scores to measure the evolutionary conservation at individual nucleotide sites in the 3'UTRs of the shared target genes. The overall conservation appeared to be greater in the regions targeted by *miR-506* family miRNAs than in the non-target regions in the 3'UTRs of the shared target genes in both mice and humans (**Fig. 6 B** and **C**) with few exceptions (**Fig. 6 D** and **E**) ($p < 0.05$, t-test). These data suggest that the regions targeted by the X-linked *miR-506* family miRNAs are under relatively stronger purifying, rather than adaptive, selection. We then tested the second hypothesis that the rapid evolution of the *miR-506* family resulted in more extant mRNAs being targeted by these miRNAs. We first compared the average target numbers of each *miR-506* family miRNA between humans and mice using the 2,510 shared targets between predicted targets in humans and the dysregulated targets in mice (**Fig. 6 A** and **F**). Among these shared targets, the human *miR-506* family members could target ~1,268 targets for each miRNA, whereas the miRNAs in mice could only target ~1,068 ($p < 0.05$, t-test) (**Fig. 6 F**), indicating that the *miR-506* family miRNAs had gained more targets in humans when compared to those in mice. Furthermore, we analyzed the number of all potential targets of the *miR-506* family miRNAs predicted by the aforementioned four algorithms among humans, mice and rats. The total number of targets for all the X-linked *miR-506* family miRNAs among different species did not show significant enrichment in humans (Fig. S10C), suggesting the sheer number of target genes does not increase in humans. We then compared the number of target genes per miRNAs. When comparing the number of target genes per miRNA for all the miRNAs (baseline) between humans and mice, we found that human miRNAs have more target genes than murine miRNAs ($p<0.05$, t-test) (Fig. S10D), implicating a higher biological complexity in humans. This became even more obvious for the X-linked *miR-506* family ($p<0.05$, t-test) (fig. S10D). In humans, the X-linked *miR-506* family targets significantly more genes than the average of all miRNAs combined ($p<0.05$, t-test) (Fig. S10D). In contrast, in mice, we did not observe a significant difference in target mRNA number between X-linked miRNAs and all of the mouse miRNAs (mouse baseline) (Fig. S10D). These results imply that human's X-linked *miR-506* family has gained additional target genes relative to mice during recent mammalian evolution. We also investigated the number of *miR-506* family miRNA targeting sites within the individual target genes in both humans and mice, but no significant differences were found between humans and mice (**Fig. 6 G**). Taken together, these results suggest the human X-linked *miR-506* family has undergone additional selective pressure causing them to gain additional regulatory functions after the evolutionary divergence between primates and rodents, and that this has been manifested as X-linked miRNAs targeting additional, novel mRNAs expressed during spermatogenesis, rather than by these miRNAs simply targeting additional sites within mRNAs that already targeted (**Fig. 6** H).

Discussion

Successful reproduction is pivotal for the perpetuation of species, and sperm are constantly facing selective pressures (60). To enhance their chance of fertilizing the eggs, sperm need to adapt accordingly, and miRNAs-mediated regulation of gene expression in spermatogenesis provides a rapidly adaptable mechanism for this purpose. Although miRNAs were initially believed to be

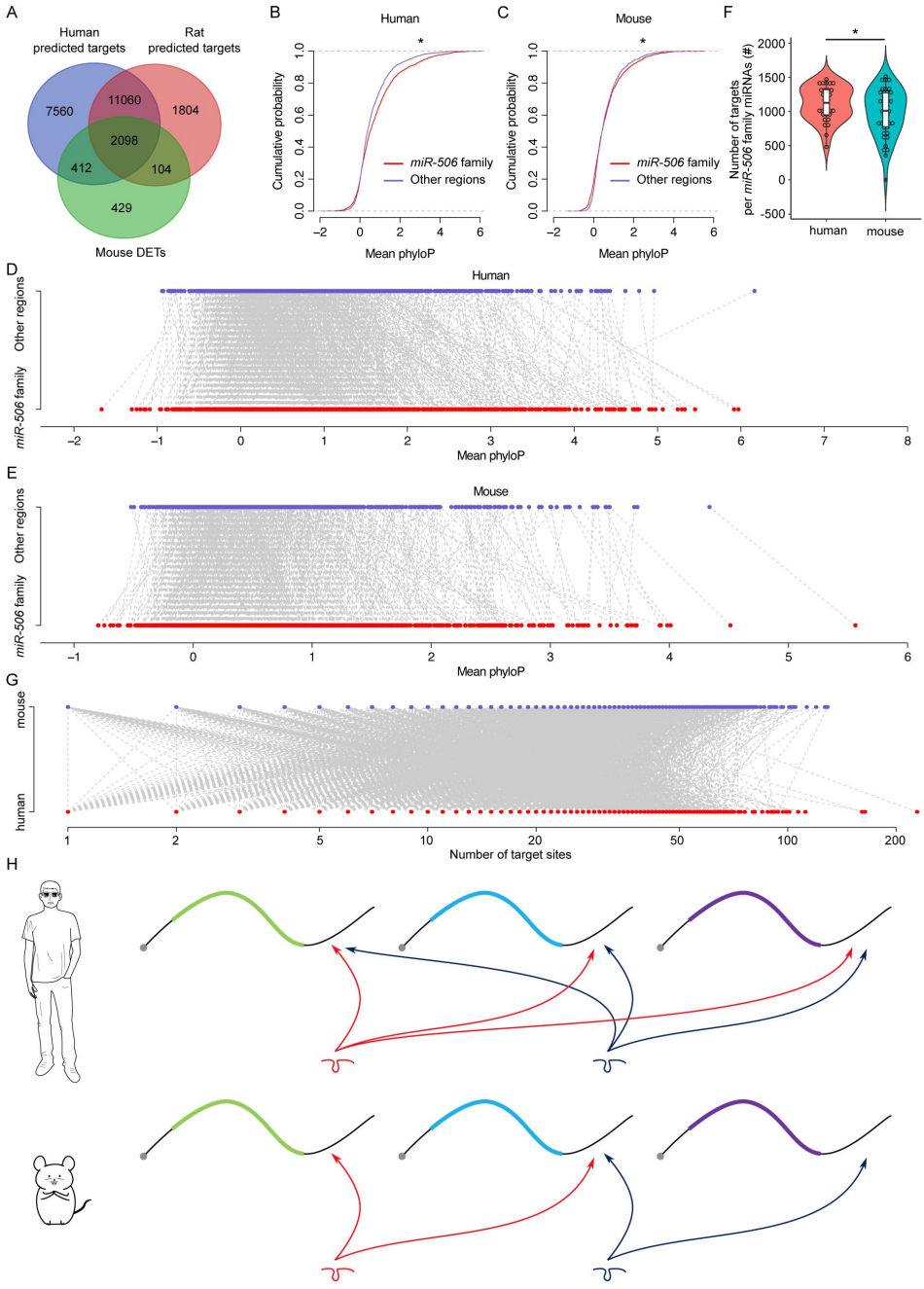

Figure 6.

Rapid evolution of the X-linked *miR-506* family miRNAs correlates with increased complexity of genetic networks that regulate spermatogenesis across mammalian species.

A. Overlap between the dysregulated targets in mice and the predicted targets in humans and rats. **B**. Comparison of the cumulative distribution between the *miR-506* family targeting sites and the other regions in humans. *: p< 0.05; t-test was used for statistical analyses. **C.** Comparison of the cumulative distribution between the *miR-506* family targeting sites and the other regions in mice. *: p< 0.05; t-test was used for statistical analyses. **D.** Paired comparison of the PhyloP score between the *miR-506* family targeting sites and the other regions in humans. **E.** Paired comparison of the PhyloP score between the *miR-506* family targeting sites and the other regions in mice. **F.** Comparison of the number of the targets per miRNA for the X-linked *miR-506* family in mice and humans. *: p< 0.05; t-test was used for statistical analyses. **G.** The number of target sites within individual target mRNAs in both humans and mice. **H.** Schematics show that human *miR-506* family miRNAs gained more targets relative to those of mice during evolution.

evolutionarily conserved, the number of non-conserved miRNAs has been steadily increasing (61). Among the non-conserved miRNAs, many are derived from TEs, suggesting that TEs serve as a major source of miRNA sequences (61). The delayed recognition of TE-derived miRNAs, in part, results from the fact that repetitive sequences were usually excluded during the computational annotation of miRNAs. In theory, TEs can serve as a good donor of miRNA sequences for the following reasons: 1) TEs are ubiquitous and abundant in the genome and are known to contribute to the regulatory elements of the coding genes, e.g., UTRs (62). As one of the regulatory factors that target mainly the 3'UTRs, TE-derived miRNAs can regulate a larger number of mRNAs with multiple miRNA-targeting sites. 2) TEs are among the most rapidly evolving sequences in the genome (63) and thus, can continuously produce species/lineage-specific miRNA genes to diversify their regulatory effects. Consistent with these notions, the present study provides evidence supporting that the *miR-506* family miRNAs originated from the MER91C DNA transposons and subsequently underwent sequence divergence *via* LINE1-mediated retrotransposition during evolution. Of more interest, the *miR-506* family miRNAs, despite their rapid evolution, are all expressed in spermatogenic cells in the testis, supporting a lineage-specific functional diversification of TE-derived miRNAs. In fact, the *miR-506* family miRNAs were among the first reported TE-derived miRNAs because of their location in a small region of the X chromosome and their confined, abundant expression in the testis (16).

It has been demonstrated that under some circumstances, genes that evolve under sexual conflicts should preferentially move to the X chromosome, especially when they are male-beneficial, female-deleterious, and act recessively (64, 65). The X chromosome is enriched with genes associated with male reproductions (15, 16, 66). The rapid evolution of the X-linked *miR-506* family strongly suggests that these miRNA genes were under selection to expand and diversify their regulatory effects on spermatogenesis. Indeed, our data strongly suggest that targeting new mRNAs was likely the driving force for the rapid evolution of the *miR-506* family of miRNAs. However, the expansion and sequence divergence of the X-linked *miR-506* family may simply reflect natural drifting without functional significance, similar to some of the pachytene piRNA clusters (30). We argue that the neutral drifting theory may not be true to the *miR-506* family for the following reasons: 1) Despite highly divergent overall sequences of the *miR-506* family, some miRNAs share the same seed regions across multiple species, suggesting that these regions may undergo strong selections. 2) The *miR-506* family miRNAs, especially the FmiRs, are highly conserved in modern humans, implying a strong selection of these miRNAs. 3) Knockout the *miR-506* family (either quinKO or XS) results in male subfertility, reflecting a biological function. 4) The quinKO sperm are less competitive than the WT sperm both in *vivo* and *in vitro*. 5) Several human studies have linked the dysregulation of the *miR-506* family with male infertility/subfertility (44–48). Similarly, one study in *Drosophila* also showed that the rapidly evolving testis-restricted miRNAs underwent adaptive evolution rather than neutral drifting (67).

Since TEs are abundant in UTRs, TE-derived miRNAs can target a much greater number of mRNAs than those derived from distinct non-repetitive genomic loci (61). Indeed, thousands of the dysregulated genes detected in the *miR-465* sKO, dKO, tKO, quadKO and quinKO testes are involved in multiple pathways of spermatogenesis. By analyzing the sequences divergence, we noticed that the most common sequence substitutions among all of the *miR-506* miRNA sequences were U-to-C and A-to-G, which were likely mediated by ADARs (adenosine deaminases acting on RNA) that can change A to I (which is functionally equivalent to G) (68). Interestingly, ~90% of the A-to-I editing appears to have occurred in Alu elements (belonging to the SINE family), and some of the edits occurred in miRNAs (68). Since G and U can form the so-called G-U wobble base pair (69), those U-to-C or A-to-G substitutions can, in theory, target similar sequences and exert regulatory functions (70), suggesting that the evolving miRNA sequences could target not only the original sequences but also new sites with similar sequences. Consistent with this notion, the predicted target genes of the *miR-506* family in mice can also be found in rats and humans, suggesting the target genes are shared across species despite the quick divergence of the miRNA sequences across species. This is also supported by our data showing that the binding sites for the

miR-506 family of miRNAs are more conserved than the surrounding, non-targeting regions in the 3'UTRs of the predicted target mRNAs. Furthermore, seed sequences among some *miR-506* family miRNAs remain the same despite the high divergence of these miRNAs, and these conserved seed sequences appear to be present in the dominant mature miRNAs. Thus, the seed region of these miRNAs appears to have undergone strong selection. Supporting this notion, previous studies have shown correlations between miRNA expression and the evolution of miRNAs and target sites (71, 72). Of interest, miRNAs with the same seed sequences may exert divergent functions. For example, the mature *miR-465a*, *miR-465b*, and *miR-465c* only have a few mismatches outside of the seed region, but only *miR-465c* exerts functional activation of *Egr1*. A similar effect has also been reported in the *miR-465* cluster on the *Alkbh1* 3'UTR activity (18). Similarly, despite the same seed sequences in the *miR-465* or *miR-743* cluster, *miR-465a* and *miR-465c* have differential activating effects on the 3'UTR of *Crisp1*, and *miR-743a* and *miR-743b* exert opposite effects on the *Crisp1* 3'UTR, further confirming their functional divergence. Therefore, the sequences outside of the miRNA seed region may play an important role in their functions, which have also been observed in *C. elegans* and human HEK293 cells (18, 73, 74).

To unequivocally demonstrate the physiological role of miRNAs, it would be ideal to delete not only the miRNAs but also their binding sites in their target transcripts *in vivo*. A few previous studies have established the miRNA:target relationship by deleting the miRNA-binding sites in target transcripts in *C. elegans*, *Drosophila*, and cell lines (75–77), but similar studies have not been reported in mice or humans. The strategy may work in mRNAs with 3'UTRs containing only one or two miRNA-binding sites, but for more complex 3'UTRs of mRNAs in mice and humans that often contain multiple binding sites for the same or different miRNAs, deletion of one miRNA binding site may not cause any discernable effects as the loss of function can easily be compensated by other miRNAs.Nevertheless, by deleting all five clusters of the *miR-506* family one by one, we were able to overcome the compensatory effects among the family members/clusters and successfully revealed the physiological role of this miRNA family.

It is well-known that mice in the wild are promiscuous, and one female often mates with multiple males sequentially, giving rise to polyandrous litters derived from sperm from more than one sire (20, 21). Polyandrous mating establishes a situation where sperm from multiple males co-exist in the female reproductive tract, with the most competitive ones fertilizing eggs and producing offspring (22, 78). Therefore, a male that may be fertile in the monandrous mating scheme, may rarely sire offspring in a polyandrous mating scenario, rendering this male functionally "infertile". Therefore, sperm competitiveness reflects the general reproductive fitness of the male (78). Although the quinKO males tend to produce smaller litters under the monandrous mating scheme, their sperm counts, sperm motility and morphology are indistinguishable from those of WT sperm. This is not surprising given that miRNAs of the *miR-506* family most likely function to control certain non-essential aspects of spermatogenesis. Sperm can be subject to competition at multiple steps during fertilization, including their migration through the female reproductive tract (cervix, uterine cavity and oviduct), binding the cumulus-oocyte complexes, penetration of zona pellucida, etc. Therefore, IVF may not be ideal for evaluating sperm competition in the real world as it bypasses several key sites where sperm competition likely takes place. Artificial insemination (AI) may represent a better way to assess sperm competition than IVF, but it is probably less desirable than polyandrous mating for the following reasons: First, in the wild, sperm from two males rarely, if not never, enter the female reproductive tract simultaneously. We had tried to place two males into the cage with one female, but the two males ended up fighting, and the submissive one never mated. Second, sperm are delivered directly into the uterus or oviduct during AI (79, 80), thus bypassing the potential sites for sperm competition (e.g., cervix and uterine cavity). Although our breeding scheme also involves sperm competition, by shortening the time between the two mating events in a laboratory setting, the sequential mating method reported here may be further improved to better mimic the natural polyandrous mating in the future. Moreover, future analyses of the quinKO sperm may help identify biochemical or molecular biomarkers for sperm competitiveness.

In summary, our data suggest that the *miR-506* family miRNAs are derived initially from the MER91C DNA transposon and their rapid evolution results from LINE1-meditated transposition to diversify their regulatory effects on both conserved and non-conserved genes involved in spermatogenesis. These miRNAs share many of their targets and can compensate for each other's absence, and they work jointly through regulating their target genes in spermatogenesis to ensure sperm competitiveness and reproductive fitness of the male.

Materials and Methods

KO mice were generated using CRISPR-Cas9-based genome editing, as described (15). F1 KO mice were backcrossed for at least five generations before data collection. Proportions of TEs were quantified based on the genomes retrieved from the UCSC genome browser. PhyloP and phastCons were applied to compare the conservation among miRNAs, and phyloP was used to evaluate the conservation of miRNA target genes. RNA-seq and sRNA-seq were performed to identify differential expressed mRNAs and miRNAs, respectively. qPCR, WB, and luciferase assays were carried out to validate miRNA targets. Detailed methods can be found in supplementary methods.

Acknowledgements

We would like to thank Dr. Kevin J. Peterson, Dartmouth College, Hanover, NH, for his help with phylogenetic analyses of the *miR-506* miRNA genes. This work was supported by grants from the NIH (HD071736, HD085506, HD098593, HD099924, and P30GM110767 to WY), NIH/National Center for Advancing Translational Science (NCATS) UCLA CTSI (UL1TR001881-01 to ZW and WY), and the Templeton Foundation (PID: 61174 to WY). Work in the Lai lab was supported by the NIH (R01-GM083300 and R01-HD108914) and Memorial Sloan-Kettering Institute Grant P30-CA008748.

References

1. Oakberg E. F. (1957) **Duration of spermatogenesis in the mouse** *Nature* **180**:1137–1138

2. Muciaccia B., et al. (2013) **Novel stage classification of human spermatogenesis based on acrosome development** *Biol Reprod* **89**

3. Gu N. H., Zhao W. L., Wang G. S., Sun F. (2019) **Comparative analysis of mammalian sperm ultrastructure reveals relationships between sperm morphology, mitochondrial functions and motility** *Reprod Biol Endocrinol* **17**

4. Hermo L., Pelletier R. M., Cyr D. G., Smith C. E. (2010) **Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 1: background to spermatogenesis, spermatogonia, and spermatocytes** *Microsc Res Tech* **73**:241–278

5. Khaitovich P., Enard W., Lachmann M., Paabo S. (2006) **Evolution of primate gene expression** *Nat Rev Genet* **7**:693–702

6. Idler R. K., Yan W. (2012) **Control of messenger RNA fate by RNA-binding proteins: an emphasis on mammalian spermatogenesis** *J Androl* **33**:309–337

7. Ro S., Park C., Sanders K. M., McCarrey J. R., Yan W. (2007) **Cloning and expression profiling of testis-expressed microRNAs** *Dev Biol* **311**:592–602

8. Wu Q., et al. (2012) **The RNase III enzyme DROSHA is essential for microRNA production and spermatogenesis** *J Biol Chem* **287**:25173–25190

9. Papaioannou M. D., et al. (2009) **Sertoli cell Dicer is essential for spermatogenesis in mice** *Dev Biol* **326**:250–259

10. Zhang Y., et al. (2017) **MicroRNAs control mRNA fate by compartmentalization based on 3' UTR length in male germ cells** *Genome Biol* **18**

11. Guo M., et al. (2022) **Uncoupling transcription and translation through miRNA-dependent poly(A) length control in haploid male germ cells** *Development* **149**

12. Bartel D. P. (2009) **MicroRNAs: target recognition and regulatory functions** *Cell* **136**:215–233

13. Mayr C. (2016) **Evolution and Biological Roles of Alternative 3'UTRs** *Trends Cell Biol* **26**:227–237

14. Thompson P. J., Macfarlan T. S., Lorincz M. C. (2016) **Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire** *Mol Cell* **62**:766–776

15. Wang Z., et al. (2020) **X-linked miR-506 family miRNAs promote FMRP expression in mouse spermatogonia** *EMBO Rep* **21**

16. Song R., et al. (2009) **Many X-linked microRNAs escape meiotic sex chromosome inactivation** *Nat Genet* **41**:488–493

17. Bao J., et al. (2012) **MicroRNA-449 and microRNA-34b/c function redundantly in murine testes by targeting E2F transcription factor-retinoblastoma protein (E2F-pRb) pathway** *J Biol Chem* **287**:21686–21698

18. Wang Z., et al. (2022) **Ablation of the miR-465 Cluster Causes a Skewed Sex Ratio in Mice** *Front Endocrinol (Lausanne* **13**

19. Miska E. A., et al. (2007) **Most Caenorhabditis elegans microRNAs are individually not essential for development or viability** *PLoS Genet* **3**

20. Dean M. D., Ardlie K. G., Nachman M. W. (2006) **The frequency of multiple paternity suggests that sperm competition is common in house mice (Mus domesticus)** *Mol Ecol* **15**:4141–4151

21. Firman R. C., Simmons L. W. (2008) **Polyandry facilitates postcopulatory inbreeding avoidance in house mice** *Evolution* **62**:603–611

22. Parker G. A. (1970) **Sperm competition and its evolutionary consequences in the insects** *Biological reviews* **45**:525–567

23. Meisel R. P., Connallon T. (2013) **The faster-X effect: integrating theory and data** *Trends Genet* **29**:537–544

24. Zhang F., et al. (2019) **Evolution of an X-Linked miRNA Family Predominantly Expressed in Mammalian Male Germ Cells** *Mol Biol Evol* **36**:663–678

25. Casper J., et al. (2018) **The UCSC Genome Browser database: 2018 update** *Nucleic Acids Res* **46**:D762–D769

26. Cabanettes F., Klopp C. (2018) **D-GENIES: dot plot large genomes in an interactive, efficient and simple way** *PeerJ* **6**

27. Fromm B., et al. (2020) **MirGeneDB 2.0: the metazoan microRNA complement** *Nucleic Acids Res* **48**

28. Kozomara A., Birgaoanu M., Griffiths-Jones S. (2019) **miRBase: from microRNA sequences to function** *Nucleic Acids Res* **47**:D155–D162

29. Byrska-Bishop M., et al. (2022) **High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios** *Cell* **185**:3426–3440

30. Ozata D. M., et al. (2020) **Evolutionarily conserved pachytene piRNA loci are highly divergent among modern humans** *Nat Ecol Evol* **4**:156–168

31. Fedoroff N. V. (2012) **Presidential address. Transposable elements, epigenetics, and genome evolution** *Science* **338**:758–767

32. Wang Z., et al. (2020) **Efficient genome editing by CRISPR-Mb3Cas12a in mice** *J Cell Sci* **133**

33. McKenna A., et al. (2016) **Whole-organism lineage tracing by combinatorial and cumulative genome editing** *Science* **353**

34. Altschul S. F., Gish W., Miller W., Myers E. W., Lipman D. J. (1990) **Basic local alignment search tool** *J Mol Biol* **215**:403–410

35. Giordano J., et al. (2007) **Evolutionary history of mammalian transposons determined by genome-wide defragmentation** *PLoS Comput Biol* **3**

36. Wang Z., et al. (2019) **Insertion of a chimeric retrotransposon sequence in mouse Axin1 locus causes metastable kinky tail phenotype** *Mob DNA* **10**

37. Hirano T., et al. (2014) **Small RNA profiling and characterization of piRNA clusters in the adult testes of the common marmoset, a model primate** *RNA* **20**:1223–1237

38. Zhang R., Peng Y., Wang W., Su B. (2007) **Rapid evolution of an X-linked microRNA cluster in primates** *Genome Res* **17**:612–617

39. Koenig E. M., et al. (2016) **The beagle dog MicroRNA tissue atlas: identifying translatable biomarkers of organ toxicity** *BMC Genomics* **17**

40. Keller A., et al. (2022) **miRNATissueAtlas2: an update to the human miRNA tissue atlas** *Nucleic Acids Res* **50**:D211–D221

41. Bushel P. R., et al. (2020) **RATEmiRs: the rat atlas of tissue-specific and enriched miRNAs for discerning baseline expression exclusivity of candidate biomarkers** *RNA Biol* **17**:630–636

42. Gainetdinov I. V., et al. (2018) **Assessment of piRNA biogenesis and function in testicular germ cell tumors and their precursor germ cell neoplasia in situ** *BMC Cancer* **18**

43. Li B., et al. (2019) **Identification of piRNAs and piRNA clusters in the testes of the Mongolian horse** *Sci Rep* **9**

44. Abu-Halima M., et al. (2013) **Altered microRNA expression profiles of human spermatozoa in patients with different spermatogenic impairments** *Fertil Steril* **99**:1249–1255

45. Heidary Z., Zaki-Dizaji M., Saliminejad K., Khorram Khorshid H. R. (2019) **MicroRNA profiling in spermatozoa of men with unexplained asthenozoospermia** *Andrologia* **51**

46. Tian H., et al. (2018) **Semen-specific miRNAs: Suitable for the distinction of infertile semen in the body fluid identification?** *Forensic Sci Int Genet* **33**:161–167

47. Qing X., et al. (2017) **Dysregulation of an X-linked primate-specific epididymal microRNA cluster in unexplained asthenozoospermia** *Oncotarget* **8**:56839–56849

48. Wang C., et al. (2011) **Altered profile of seminal plasma microRNAs in the molecular diagnosis of male infertility** *Clin Chem* **57**:1722–1731

49. Xia B., et al. (2020) **Widespread Transcriptional Scanning in the Testis Modulates Gene Evolution Rates** *Cell* **180**:248–262

50. Firman R. C. (2011) **Polyandrous females benefit by producing sons that achieve high reproductive success in a competitive environment** *Proc Biol Sci* **278**:2823–2831

51. Muzumdar M. D., Tasic B., Miyamichi K., Li L., Luo L. (2007) **A global double-fluorescent Cre reporter mouse** *Genesis* **45**:593–605

52. Fisher H. S., Hoekstra H. E. (2010) **Competition drives cooperation among closely related sperm of deer mice** *Nature* **463**:801–803

53. Fisher H. S., Jacobs-Palmer E., Lassance J. M., Hoekstra H. E. (2016) **The genetic basis and fitness consequences of sperm midpiece size in deer mice** *Nat Commun* **7**

54. Agarwal V., Bell G. W., Nam J. W., Bartel D. P. (2015) **Predicting effective microRNA target sites in mammalian mRNAs** *Elife* **4**

55. Betel D., Koppal A., Agius P., Sander C., Leslie C. (2010) **Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites** *Genome Biol* **11**

56. Dweep H., Gretz N. (2015) **miRWalk2.0: a comprehensive atlas of microRNA-target interactions** *Nat Methods* **12**

57. Chen Y., Wang X. (2020) **miRDB: an online database for prediction of functional microRNA targets** *Nucleic Acids Res* **48**:D127–D131

58. Da Ros V. G., et al. (2008) **Impaired sperm fertilizing ability in mice lacking Cysteine-RIch Secretory Protein 1 (CRISP1)** *Dev Biol* **320**:12–18

59. Maldera J. A., et al. (2014) **Human fertilization: epididymal hCRISP1 mediates sperm-zona pellucida binding through its interaction with ZP3** *Mol Hum Reprod* **20**:341–349

60. Morrow E. H. (2004) **How the sperm lost its tail: the evolution of aflagellate sperm** *Biol Rev Camb Philos Soc* **79**:795–814

61. Piriyapongsa J., Marino-Ramirez L., Jordan I. K. (2007) **Origin and evolution of human microRNAs from transposable elements** *Genetics* **176**:1323–1337

62. Jordan I. K., Rogozin I. B., Glazko G. V., Koonin E. V. (2003) **Origin of a substantial fraction of human regulatory sequences from transposable elements** *Trends Genet* **19**:68–72

63. Gonzalez J., Petrov D. A. (2012) **Evolution of genome content: population dynamics of transposable elements in flies and humans** *Methods Mol Biol* **855**:361–383

64. Rice W. R. (1984) **Sex Chromosomes and the Evolution of Sexual Dimorphism** *Evolution* **38**:735–742

65. Gibson J. R., Chippindale A. K., Rice W. R. (2002) **The X chromosome is a hot spot for sexually antagonistic fitness variation** *Proc Biol Sci* **269**:499–505

66. Wang P. J., McCarrey J. R., Yang F., Page D. C. (2001) **An abundance of X-linked genes expressed in spermatogonia** *Nat Genet* **27**:422–426

67. Mohammed J., et al. (2014) **Adaptive evolution of testis-specific, recently evolved, clustered miRNAs in Drosophila** *RNA* **20**:1195–1209

68. Nishikura K. (2016) **A-to-I editing of coding and non-coding RNAs by ADARs** *Nat Rev Mol Cell Biol* **17**:83–96

69. Varani G., McClain W. H. (2000) **The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems** *EMBO Rep* **1**:18–23

70. Doench J. G., Sharp P. A. (2004) **Specificity of microRNA target selection in translational repression** *Genes Dev* **18**:504–511

71. Simkin A., Geissler R., McIntyre A. B. R., Grimson A. (2020) **Evolutionary dynamics of microRNA target sites across vertebrate evolution** *PLoS Genet* **16**

72. Meunier J., et al. (2013) **Birth and expression evolution of mammalian microRNA genes** *Genome Res* **23**:34–45

73. Broughton J. P., Lovci M. T., Huang J. L., Yeo G. W., Pasquinelli A. E. (2016) **Pairing beyond the Seed Supports MicroRNA Targeting Specificity** *Mol Cell* **64**:320–333

74. Helwak A., Kudla G., Dudnakova T., Tollervey D. (2013) **Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding** *Cell* **153**:654–665

75. Pinzon N., et al. (2017) **microRNA target prediction programs predict many false positives** *Genome Res* **27**:234–245

76. Ecsedi M., Rausch M., Grosshans H. (2015) **The let-7 microRNA directs vulval development through a single target** *Dev Cell* **32**:335–344

77. Garaulet D. L., Zhang B., Wei L., Li E., Lai E. C. (2020) **miRNAs and Neural Alternative Polyadenylation Specify the Virgin Behavioral State** *Dev Cell* **54**:410–423

78. Gomendio M., Martin-Coello J., Crespo C., Magana C., Roldan E. R. (2006) **Sperm competition enhances functional capacity of mammalian spermatozoa** *Proc Natl Acad Sci U S A* **103**:15113–15117

79. Nagy A., Gertsenstein M., Vintersten K., Behringer R. (2003) **Manipulating the mouse embryo: a laboratory manual**

80. Stone B. J., Steele K. H., Fath-Goodin A. (2015) **A rapid and effective nonsurgical artificial insemination protocol using the NSET device for sperm transfer in mice without anesthesia** *Transgenic Res* **24**:775–781

Article and author information

Zhuqing Wang

Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA
ORCID iD: 0000-0002-3988-0733

Yue Wang

Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Tong Zhou

Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Sheng Chen
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Dayton Morris
The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Rubens Daniel Miserani Magalhães
The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Musheng Li
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Shawn Wang
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Hetan Wang
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Yeming Xie
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Hayden McSwiggin
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Daniel Oliver
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA

Shuiqiao Yuan
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA
ORCID iD: 0000-0003-1460-7682

Huili Zheng
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA

Jaaved Mohammed
Department of Developmental Biology, Memorial Sloan-Kettering Institute, 1275 York Ave, Box 252, New York, NY 10065, USA

Eric C. Lai
Department of Developmental Biology, Memorial Sloan-Kettering Institute, 1275 York Ave, Box 252, New York, NY 10065, USA
ORCID iD: 0000-0002-8432-5851

John R. McCarrey
Department of Biology, University of Texas at San Antonio, San Antonio, TX, USA

Wei Yan
Department of Physiology and Cell Biology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA 90502, USA, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA
For correspondence: wei.yan@lundquist.org
ORCID iD: 0000-0001-9569-9026

Copyright

Editors

Reviewing Editor
Diane Harper
University of Michigan, Ann Arbor, United States of America

Senior Editor
Diane Harper
University of Michigan, Ann Arbor, United States of America

Reviewer #1 (Public Review):

Wang et al investigated the evolution, expression, and function of the X-linked miR-506 miRNA family. They showed that the miR-506 family underwent rapid evolution. They provided evidence that miR-506 appeared to have originated from the MER91C DNA transposons. Human MER91C transposon produced mature miRNAs when expressed in cultured cells. A series of mouse mutants lacking individual clusters, a combination of clusters, and the entire X-linked cluster (all 22 miRNAs) were generated and characterized. The mutant mice lacking four or more miRNA clusters showed reduced reproductive fitness (litter size reduction). They further showed that the sperm from these mutants were less competitive in polyandrous mating tests. RNA-seq revealed the impact of deletion of miR-506 on the testicular transcriptome. Bioinformatic analysis analyzed the relationship among miR-506 binding, transcriptomic changes, and target sequence conservation. The miR-506-deficient mice did not have apparent effect on sperm production, motility, and morphology. Lack of severe phenotypes is typical for miRNA mutants in other species as well. However,

the miR-506-deficient males did exhibit reduced litter size, such an effect would have been quite significant in an evolutionary time scale. The number of mouse mutants and sequencing analysis represent a tour de force. This study is a comprehensive investigation of the X-linked miR-506 miRNA family. It provides important insights into the evolution and function of the miR-506 family.

The conclusions of this preprint are mostly supported by the data except being noted below. Some descriptions need to be revised for accuracy.

L219-L285: The conclusion that X-linked miR-506 family miRNAs are expanded via LINE1 retrotransposition is not supported by the data. LINE1s and SINEs are very abundant, accounting for nearly 30% of the genome. In addition, the LINE1 content of the mammalian X chromosome is twice that of the autosomes. One can easily find flanking LINE1/SINE repeat. Therefore, the analyses in Fig. 2G, Fig. 2H and Fig. S3 are not informative. In order to claim LINE1-mediated retrotransposition, it is necessary to show the hallmarks of LINE1 retrotransposition, which are only possible for new insertions. The X chromosome is known to be enriched for testis-specific multi-copy genes that are expressed in round spermatids (PMID: 18454149). The conclusion on the LINE1-mediated expansion of miR-506 family on the X chromosome is not supported by the data and does not add additional insights. I think that the LINE1 related figure panels and description (L219-L285) need to be deleted. In discussion (L557-558), "...and subsequently underwent sequence divergence via LINE1-mediated retrotransposition during evolution" should also be deleted. This section (L219-L285) needs to deal only with the origin of miR-506 from MER91C DNA transposons, which is both convincing and informative.

Fig. 3A: can you speculate/discuss why the miR-506 expression in sperm is higher than in round spermatids?

- https://doi.org/10.7554/eLife.90203.1.sa2

Reviewer #2 (Public Review):

In this paper, Wang and collaborators characterize the rapid evolution of the X-linked miR-506 cluster in mammals and characterize the functional reference of depleting a few or most of the miRNAs in the cluster. The authors show that the cluster originated from the MER91C DNA transposon and provide some evidence that it might have expanded through the retrotransposition of adjacent LINE1s. Although the animals depleted of most miRNAs in the cluster show normal sperm parameters, the authors observed a small but significant reduction in litter size. The authors then speculate that the depletion of most miRNAs in the cluster could impair sperm competitiveness in polyandrous mating. Using a successive mating protocol, they show that, indeed, sperm lacking most X-linked miR-506 family members is outcompeted by wild-type sperm. The authors then analyze the evolution of the miR-506 cluster and its predicted targets. They conclude that the main difference between mice and humans is the expansion of the number of target sites per transcript in humans.

The conclusions of the paper are, in most cases, supported by the data; however, a more precise and in-depth analysis would have helped build a more convincing argument in most cases.

1. In the abstracts and throughout the manuscript, the authors claim that "... these X-linked miRNA-506 family miRNA [...] have gained more targets [...] " while comparing the human miRNA-506 family to the mouse. An alternative possibility is that the mouse has lost some targets. A proper analysis would entail determining the number of targets in the mouse and human common ancestor.

2. The authors claim that the miRNA cluster expanded through L1 retrotransposition. However, the possibility of an early expansion of the cluster before the divergence of the species while the MER91C DNA transposon was active was not evaluated. Although L1 likely contributed to the diversity within mammals, the generalization may not apply to all species. For example, SINEs are closer on average than L1s to the miRNAs in the SmiR subcluster in humans and dogs, and the horse SmiR subcluster seems to have expanded by a TE-independent mechanism.

3. Some results are difficult to reconcile and would have benefited from further discussion. The miR-465 sKO has over two thousand differentially expressed transcripts and no apparent phenotype. Also, the authors show a sharp downregulation of CRISP1 at the RNA and protein level in the mouse. However, most miRNAs of the cluster increase the expression of Crisp1 on a reporter assay. The only one with a negative impact has a very mild effect. miRNAs are typically associated with target repression; however, most of the miRNAs analyzed in this study activate transcript expression.

4. More information is required to interpret the results of the differential RNA targeting by the murine and human miRNA-506 family. The materials and methods section needs to explain how the authors select their putative targets. In the text, they mention the use of four different prediction programs. Are they considering all sites predicted by any method, all sites predicted simultaneously by all methods, or something in between? Also, what are they considering as a "shared target" between mice and humans? Is it a mRNA that any miR-506 family member is targeting? Is it a mRNA targeted by the same miRNA in both species? Does the targeting need to occur in the same position determined by aligning the different 3'UTRs?

5. The authors highlight the particular evolution of the cluster derived from a transposable element. Given the tendency of transposable elements to be expressed in germ cells, the family might have originated to repress the expression of the elements while still active but then remained to control the expression of the genes where the element had been inserted. The authors did not evaluate the expression of transcripts containing the transposable element or discuss this possibility. The authors proposed an expansion of the target sites in humans. However, whether this expansion was associated with the expansion of the TE in humans was not discussed either. Clarifying whether the transposable element was still active after the divergence of the mouse and human lineages would have been informative to address this outstanding issue.

Post-transcriptional regulation is exceptionally complex in male haploid cells, and the functional relevance of many regulatory pathways remains unclear. This manuscript, together with recent findings on the role of piRNA clusters, starts to clarify the nature of the selective pressure that shapes the evolution of small RNA pathways in the male germ line.

- https://doi.org/10.7554/eLife.90203.1.sa1

Reviewer #3 (Public Review):

Summary:

In this manuscript, the authors conducted a comprehensive study of the X-linked miR-506 family miRNAs in mice on its origin, evolution, expression, and function. They demonstrate that the X-linked miR-506 family, predominantly expressed in the testis, may be derived from MER91C DNA transposons and further expanded by retrotransposition. By genetic deletion of different combinations of 5 major clusters of this miRNA family in mice, they found these miRNAs are not required for spermatogenesis. However, by further examination, the mutant

mice show mild fertility problem and inferior sperm competitiveness. The authors conclude that the X-linked miR-506 miRNAs finetune spermatogenesis to enhance sperm competition.

Strengths:

This is a comprehensive study with extensive computational and genetic dissection of the X-linked miR-506 family providing a holistic view of its evolution and function in mice. The finding that this family miRNAs could enhance sperm competition is interesting and could explain their roles in finetuning germ cell gene expression to regulate reproductive fitness.

Weaknesses:

The authors specifically addressed the function of 5 clusters of X-link miR-506 family containing 19 miRNAs. There is another small cluster containing 3 miRNAs close to the Fmr1 locus. Would this small cluster act in concert with the 5 clusters to regulate spermatogenesis? In addition, any autosomal miR-506 like miRNAs may compensate for the loss of X-linked miR-506 family. These possibilities should be discussed.

Direct molecular link to sperm competitiveness defect remains unclear but is difficult to address.

- https://doi.org/10.7554/eLife.90203.1.sa0