Parallel Evolution of Sperm Hyper-Activation Ca^{2+} Channels

Jacob C. Cooper* and Nitin Phadnis

Department of Biology, University of Utah

*Corresponding author: E-mail: jacob.cooper@utah.edu.

Accepted: July 9, 2017

Data deposition: This project has been deposited at GenBank under the accession # KY990034-KY990049

Abstract

Sperm hyper-activation is a dramatic change in sperm behavior where mature sperm burst into a final sprint in the race to the egg. The mechanism of sperm hyper-activation in many metazoans, including humans, consists of a jolt of Ca^{2+} into the sperm flagellum via CatSper ion channels. Surprisingly, all nine *CatSper* genes have been independently lost in several animal lineages. In *Drosophila*, sperm hyper-activation is performed through the cooption of the polycystic kidney disease 2 (*pkd2*) Ca^{2+} channel. The parallels between *CatSpers* in primates and *pkd2* in *Drosophila* provide a unique opportunity to examine the molecular evolution of the sperm hyper-activation machinery in two independent, nonhomologous calcium channels separated by $>$ 500 million years of divergence. Here, we use a comprehensive phylogenomic approach to investigate the selective pressures on these sperm hyper-activation channels. First, we find that the entire *CatSper* complex evolves rapidly under recurrent positive selection in primates. Second, we find that *pkd2* has parallel patterns of adaptive evolution in *Drosophila*. Third, we show that this adaptive evolution of *pkd2* is driven by its role in sperm hyper-activation. These patterns of selection suggest that the evolution of the sperm hyper-activation machinery is driven by sexual conflict with antagonistic ligands that modulate channel activity. Together, our results add sperm hyper-activation channels to the class of fast evolving reproductive proteins and provide insights into the mechanisms used by the sexes to manipulate sperm behavior.

Key words: sexual conflict, sperm hyper-activation, CatSper, reproductive proteins, parallel evolution, evolutionary arms race.

Introduction

Sexual conflict shapes sperm development and sperm dynamics (Swanson and Vacquier 2002; Clark et al. 2006; Turner et al. 2008; Wilburn and Swanson 2016). Both male–female interactions and intermale competition drive rapid changes in male reproductive proteins, whose constant innovation has been likened to a molecular arms race. These rapid changes in reproductive proteins have the potential to establish barriers to fertilization between populations and lead to the evolution of new species (Parker and Partridge 1998; Gavrilets 2000; Howard et al. 2009; Moyle et al. 2014). The best-known examples of this phenomenon include the rapid evolution of reproductive proteins in abalone, mammals, and *Drosophila* (Lee et al. 1995; Swanson and Vacquier 1997; Kresge et al. 2001; Swanson et al. 2003; Gomendio et al. 2006; Clark et al. 2007; Hamm et al. 2007; Findlay et al. 2014; Vicens et al. 2015). Molecular evolutionary studies on how sexual conflicts shape reproductive proteins have focused on several aspects of sperm biology such as direct sperm-egg interactions, seminal fluid proteins and sperm behavior (Swanson and Vacquier 2002; Panhuis et al. 2006; Fisher et al. 2016). Here, we uncover the patterns of molecular evolution of the sperm hyper-activation machinery in animals, which remains a fundamental but largely unexplored aspect of sperm biology.

When spermatogenesis is complete, the resulting mature sperm are motile but quiescent. After copulation, however, sperm cease normal swimming and burst into a sprint. This dramatic postmating acceleration of sperm is known as sperm hyper-activation (Suarez et al. 1991). Sperm hyper-activation was first observed in mammals through studies on the golden hamster (Yanagimachi 1969, 1970), and has since been since described in many other taxa (Cosson et al. 2008; Suarez and Ho 2003). When sperm hyper-activate, they go through a cellular change that alters the motion of the sperm flagellum from a slow, low amplitude, symmetric beat to a whip-like, high amplitude, asymmetric beat (Ooi et al. 2014). This transition is not subtle; hyper-activated sperm swimming at top

speed are propelled with a force several times that of normal swimming (Ishijima 2011). This hyper-activated acceleration of sperm is necessary for successful fertilization, and is an integral part of sperm capacitation.

The proximate molecular mechanism of sperm hyper-activation consists of a jolt of Ca^{2+} ions to the sperm flagellum, which triggers a complex intracellular chain of events to drive accelerated swimming (Ho et al. 2002; Carlson et al. 2009). The use of Ca^{2+} influx as a sperm hyper-activation trigger is an ancient and widely conserved mechanism across metazoans (Cai et al. 2014). In most metazoans, the Ca^{2+} influx arrives via the cation channels of sperm (CatSper) complex, which forms ion channels on the sperm flagellum (Cai et al. 2014). The CatSper complex consists of four core proteins that form the Ca^{2+} pore and five auxiliary proteins (Chung et al. 2017). Sperm that do not hyper-activate fail to fertilize eggs, and males with mutations that disable sperm hyper-activation are often sterile despite producing morphologically normal sperm (Ren et al. 2001).

All CatSper proteins—core and auxiliary—are necessary for sperm hyper-activation and male fertility (Loux et al. 2013). Despite the conserved role of the CatSper complex in sperm hyper-activation across metazoan taxa, the sequences of CatSper proteins provide hints of being evolutionarily labile. Previous analyses of *CatSper* evolution have been focused on the first exon of *CATSPER1*, which shows signs of adaptive evolution in the form of an accelerated accumulation of indels and nonsynonymous changes in mammals (Cai and Clapham 2008; Podlaha and Zhang 2003; Podlaha et al. 2005). More surprisingly, the entire suite of *CatSper* genes has been lost in several animal taxa, including those leading to arthropods, nematodes, mollusks, jawless fishes, bony fishes, birds, frogs, etc. (Cai and Clapham 2008). Some taxa that have lost the *CatSper* complex, such as frogs, no longer hyper-activate their sperm during fertilization (Dziminski et al. 2009). In contrast, many other taxa, such as flies and birds, are known to hyper-activate sperm despite lacking a functional *CatSper* complex (Köttgen et al. 2011; Yang and Lu 2011; O'Brien et al. 2011; Nguyen et al. 2014; Zhou et al. 2015). These patterns suggest that some taxa that are missing the CatSper complex may have compensated for the loss through the cooption of other mechanisms to perform sperm hyper-activation. The evolutionary forces that drive the repeated turnover of the sperm hyper-activation machinery remain unaddressed.

The mechanism of CatSper-independent sperm hyper-activation remains unknown in many taxa, but is best understood in *Drosophila melanogaster*. None of the *CatSper* genes are present in the *D. melanogaster* genome. Yet, *Drosophila* males hyper-activate their sperm postcopulation as a necessary step for successful fertilization. CatSper-independent sperm hyper-activation in *Drosophila* is performed by the protein polycystic kidney disease 2 (*pkd2*) (Yang and Lu 2011; Köttgen et al. 2011). Similar to the CatSper proteins in mammals, *Drosophila pkd2* is a Ca^{2+} ion channel protein on the fly sperm flagellum. *Drosophila pkd2* null sperm are morphologically normal, but do not hyper-activate after transfer to the female reproductive tract. As a result, *Drosophila pkd2*-deficient sperm fail to reach the storage organs and are not retained in the female (Gao et al. 2003; Watnick et al. 2003). *Drosophila*, therefore, appear to have compensated for the loss of CatSper ion channels by using the pkd2 channel to trigger sperm hyper-activation.

Sperm hyper-activation is a powerful and tightly controlled behavioral switch for the final sprint in the race to the egg (Montoto et al. 2011). The sperm hyper-activation machinery may, therefore, be vulnerable intermale sperm competition, female choice based selection, or both. An evolutionary arms race over the modulation of sperm hyper-activation can manifest as the rapid evolution of sperm hyper-activation channels. The parallels between *CatSpers* in primates and *pkd2* in *Drosophila* provide a unique opportunity to examine the how selection has shaped sperm hyper-activation machinery in two independent, nonhomologous calcium channels. Here, we use a comprehensive phylogenomic approach with *CatSper* in primates and *pkd2* in *Drosophila* to investigate the selective pressures on the Ca^{2+} channels required for sperm hyper-activation. First, we find that all core and auxiliary proteins of the *CatSper* complex evolve rapidly under recurrent positive selection in primates. Second, we find that *pkd2* has similar patterns of positive selection in *Drosophila*, including increased amino-acid substitution and an accumulation of indels. Third, we show that the selective pressures of *Drosophila pkd2* and primate *PKD2* are radically different; primate *PKD2*, which has no role in sperm hyper-activation, is highly conserved rather than rapidly evolving. This provides a unique example where an otherwise slow-evolving "housekeeping" gene is dragged into an evolutionary conflict and experiences adaptive evolution. Together, our study provides the first comprehensive analysis of the molecular evolutionary patterns of the sperm hyper-activation Ca^{2+} channels in primates and flies.

Materials and Methods

Collection of Primate *and Drosophila* Gene Sequences

We acquired sequences from up to 17 species of primates and flies each for our phylogenetic analyses. Previous methods for finding homologs for phylogenetic analyses have relied on direct sequencing or manual curating of annotated genomes. Because a reliance on gene annotation models is prone to errors due to mis-annotation, and also limits the number of species that may be analyzed (Markova-Raina and Petrov 2011), we used a different approach to scan sequenced genomes and identify high quality sequences. This method is similar to the one used recently to study positive selection in the synaptonemal complex in *Drosophila* (Hemmer and Blumenstiel 2016), and utilizes the broad range of available sequenced genomes to maximize the power of our analyses.

Table 1
Species Used in This Study

A		B	
Primates	**Group**	**Drosophila**	**Subgroup**
Pan troglodyte	Hominid	*D. melanogaster*	melanogaster
Pan paniscus	Hominid	*D. simulans*	melanogaster
Homo sapiens	Hominid	*D. sechellia*	melanogaster
Gorilla gorilla	Hominid	*D. mauritiana*	melanogaster
Pongo abeli	Hominid	*D. yakuba*	melanogaster
Nomascus leucogenys	Hominid	*D. tessieri*	melanogaster
Colobus angolensis	Old World Monkeys	*D. santomea*	melanogaster
Chorocebus sabaeus	Old World Monkeys	*D. erecta*	melanogaster
Cerocebus atys	Old World Monkeys	*D. orena*	melanogaster
Mandrillis leucophaeus	Old World Monkeys	*D. eugracillis*	eugracillis
Macaca fascicularis	Old World Monkeys	*D. biarmipes*	suzukii
Macaca mulatta	Old World Monkeys	*D. takahashii*	takahashii
Macaca nemestrina	Old World Monkeys	*D. ficusphila*	ficusphila
Saimiri boliviensis	New World Monkeys	*D. rhopaloa*	montium
Aotus nancymae	New World Monkeys	*D. elegans*	elegans
Callithirx jacchus	New World Monkeys	*D. ananassae*	ananassae
		D. bipectinata	ananassae

First, we gathered the annotated *H. sapiens* CDS sequence for each *CatSper* gene to use as bait for the searches. For each *CatSper* gene, we gathered homologous sequences from up to 17 primate species that provide a well distributed phylogenetic sampling (table 1*A*). Next, we identified the genomic region containing each gene using tBlastn (NCBI) (Madden 2003). Using Exonerate (Slater and Birney 2005) on the narrowed genomic region, we identified the homologous coding regions for each gene, and predicted intron–exon splice sites. We used this two-step method because while exonerate is highly accurate at identifying homologs, it is slow at searching through full genomes. Focusing on smaller contigs that contained our gene of interest accelerated our analyses.

For analyses with *Drosophila* species, we restricted our survey of *Drosophila pkd2* sequences to the melanogaster group of species because including more divergent sequences saturated dS and raised the rate of false positives. For species with annotated genomes, we acquired homologous sequences from Flybase, whereas for species with unannotated genomes, we used our two-step method homology search (table 1*B*). We also PCR amplified, subcloned and directly Sanger sequenced *Drosophila pkd2* from *D. tessieri*, *D. santomea*, *D. mauritiana*, and *D. orena*. These sequences have been deposited in Genbank under accession numbers KY990034-KY990049.

Tests of Recurrent Positive Selection

We developed a pipeline to obtain high quality sequences from available genomic resources, and to modify our analysis to ignore low quality sequences (supplementary fig. 6, Supplementary Material online). Codon based methods of phylogenetic analysis require the accurate alignment of homologous gene sequences. We, therefore, only included sequences from a species when the homolog accounted for at least 90% of the total length of the reference coding region. This method ensured that we did not lose power in our tests because of incomplete gene sequences. To detect patterns of recurrent positive selection in these genes, we prepared and analyzed sequences with the following pipeline. First, we translated all coding sequences to amino acid sequences and aligned the amino acid sequences using CLUSTAL Omega (Sievers et al. 2011). Next, we back-translated the amino acid sequences to their corresponding CDS using Pal2Nal (Suyama et al. 2006), using settings to remove all gaps and stop codons. We then manually reviewed each alignment to confirm that there were no gaps or stop codons. We constructed the phylogenies for our alignment based on modEncode data (Perelman et al. 2011; Chen et al. 2014).

We used the same alignments for two separate analyses: Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang 2007) and Baysean Unrestricted Test For Episodic Diversification (BUSTED) (Murrell et al. 2015). To test for recurrent selection using PAML, we compared NSsites models M7 and M8, using the branch model 0. We calculated a *P* value using a log-ratio test between the log-likelihood scores for each model (Yang 2007). PAML is highly sensitive to mis-alignments; even slight mis-alignments can easily create false positive signals (Markova-Raina and Petrov 2011). When we observed a P value < 0.05, we repeated our analyses by starting at the beginning of the pipeline, but this time using the

T-coffee (Notredame et al. 2000) and Muscle (Edgar 2004) aligners, ensuring that our result was not an artifact of alignment error. We only considered true rejections of the null where a P value < 0.05 was observed with all three aligners, and reported the least significant P value. We ran M8a using the same alignment that generated the least significant P value. Like PAML, BUSTED compares models of selection for homologous sequences over a phylogenetic distribution. Unlike PAML, BUSTED takes a Bayesian approach to build these models. This framework makes BUSTED an independent test from PAML to analyze the molecular evolution of a gene. We ran BUSTED using the Data Monkey server (http://www.datamonkey.org/). When the program did not correctly compute the base tree, we reoriented nodes to correctly reconstruct the phylogeny.

To compare d*N*/d*S* between the *CatSper* genes (fig. 1*A*), we removed the percent alignment threshold for each species so that we could analyze d*N*/d*S* over a common phylogeny of 16 species. To calculate d*N*/d*S* along branch lengths, we used the branch model 1 with NSsites model 0 (Yang 2007).

We also used the McDonald–Kreitman (MK) test for positive selection on the narrower timescale of *D. melanogaster* and *D. simulans* divergence. For this test, we subcloned and resequenced *Drosophila pkd2* from eight lines of *D. melanogaster* and *D. simulans* using the method described above. We ran MK tests for *Drosophila pkd2* using DnaSP (Librado and Rozas 2009). To identify protein domains, we submitted amino acid sequences to the SMART server (Letunic et al. 2015) using all available databases.

Insertion/Deletion Polymorphism Analysis

To quantitatively assess the extent of indel differences between *Drosophila pkd2* in the *D. melanogaster* group, we developed a null prediction for indel variation per amino-acid site. To do this, we first measured the differences in length for *Drosophila pkd2* for the full gene, and determined the length variance per base pair. We measured the difference in length of each region of the gene, using the N-terminus, first trans-membrane domain, PKD channel, and C-terminus as alignment anchors. We then calculated the standard deviation of each of these regions, and divided by the length of the region to calculate a deviation per amino-acid value. For each region, we present the ratio of one deviation per amino-acid of the gene region divided by the full length of the gene.

Computational Modeling of the *Drosophila pkd2* Three Dimensional Structure

To model the structure of *Drosophila pkd2* we accessed i-Tasser via the web server at http://zhanglab.ccmb.med.umich.edu/I-TASSER/(Roy et al. 2010). We provided the program with the amino acid sequence of *Drosophila pkd2* from positions 233-810 to correspond with the published high resolution cryo-EM structure for human PKD2 (Shen et al. 2016). We provided the human PKD2 structure as a scaffold for i-Tasser, and used the option to align the two sequences before structural prediction. Due to the limits of i-Tasser, we modeled a single monomer of *Drosophila pkd2*. To arrange these monomers in a tetrameric complex, we aligned the *Drosophila pkd2* monomer to the four positions of the human PKD2 monomers in the solved structure using PyMOL. We highlighted all nonsynonymous positions between *D. melanogaster* and *D. simulans* in the predicted structure.

Results

Accelerated Amino Acid Evolution of the Entire CatSper Complex in Primates

Despite the critical role of the *CatSper* complex in sperm hyperactivation across a wide variety of metazoa, little is known about its molecular evolution. At the core of the CatSper complex lies a Ca^{2+} pore composed of a hetero-tetramer of CATSPER1-4. Each core CatSper protein contains a six-pass transmembrane domain with polycystic kidney disease (PKD) domains (Jin et al. 2005; Lobley et al. 2003; Qi et al. 2007, 20; Quill et al. 2001; Ren et al. 2001) (fig. 1*B*). In contrast, four of the five auxiliary proteins, CATSPERβ, CATSPERΔ, CATSPERγ, and CATSPERϵ, have a large extracellular region with one or two transmembrane domains (Liu et al. 2007; Wang et al. 2009; Chung et al. 2011, 2017). The fifth auxiliary protein, CATSPERζ, is a small intracellular scaffold that helps assemble the complex (Chung et al. 2017). It is found only on the sperm flagellum, with a substantial portion exposed to the external environment of the sperm, and its only known function is in fertilization. The CatSper channel is, in theory, well positioned to be involved in sperm competition. We were, therefore, interested in a comprehensive understanding of the molecular evolutionary patterns of the *CatSper* complex and in uncovering the evolutionary forces that drive the changes in these genes.

To examine the evolutionary forces that shape the CatSper complex in primates, we gathered and aligned homologous sequences for each *CatSper* gene from 16 primate species (table 1). This sampling of primate species represents ~40 million years of divergence. To test if any of any of the CatSper proteins show signatures of rapid accumulation of amino acid changes, we used Phylogenetic Analysis by Maximum Likelihood (PAML) to calculate the synonymous to nonsynonymous substitution rate ratios (d*N*/d*S*) (Yang 2007). For each of the nine genes that constitute the *CatSper* complex, we asked if any branch in the primate phylogeny had a d*N*/d*S* >1. We first analyzed the four core proteins CATSPER1-4 that form the Ca^{2+} pore. We detected elevated rates of d*N*/d*S* in each of these four core proteins (fig. 1*A*, supplementary fig. 1, Supplementary Material online). In addition to the core proteins, all five auxiliary transmembrane proteins CATSPERβ, CATSPERΔ, CATSPERγ, CATSPERϵ and CATSPERζ also show elevated rates of d*N*/

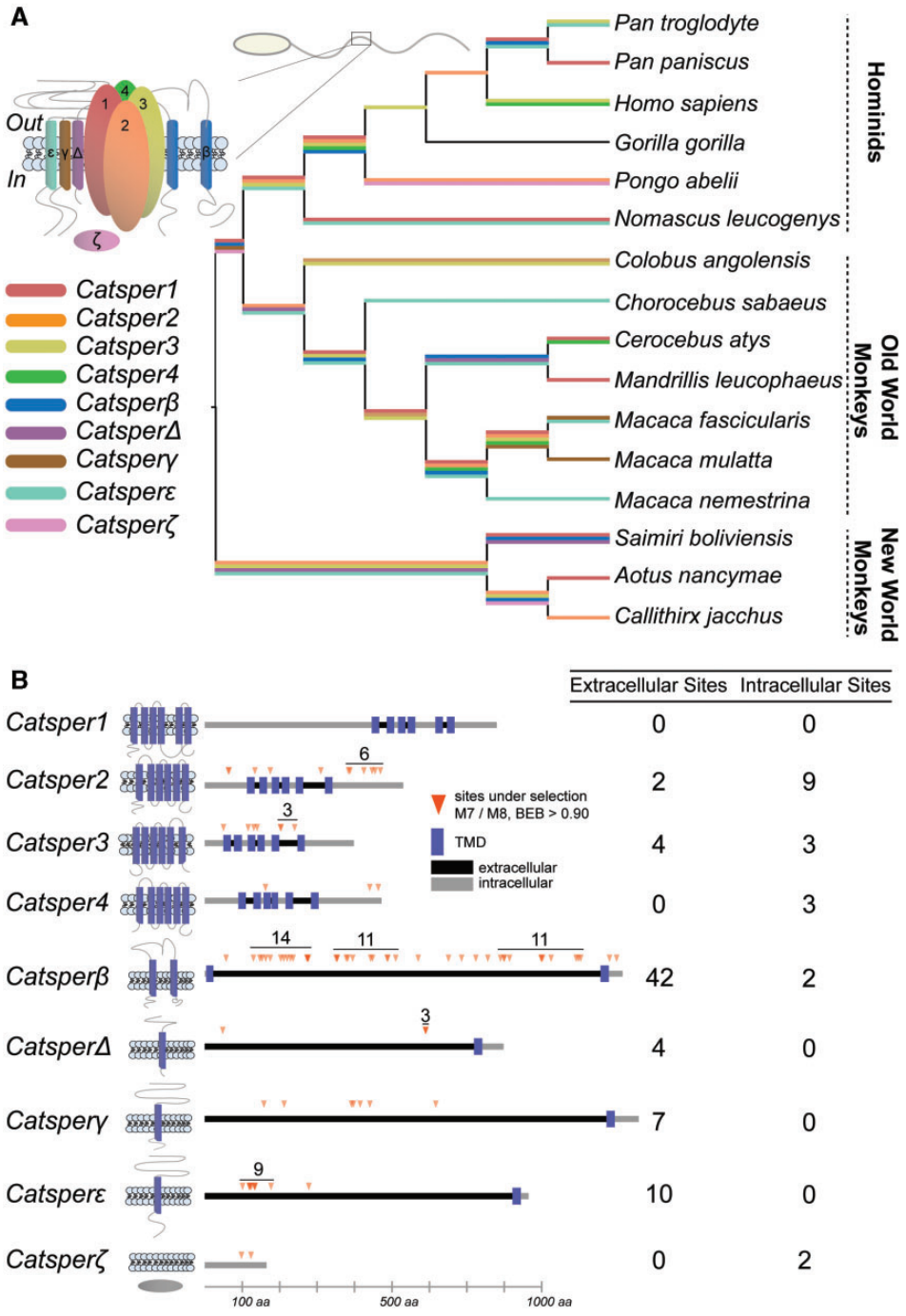


Fig. 1.—The entire *CatSper* complex evolves adaptively in primates. (*A*) The *CatSper* complex evolves under positive selection in almost every lineage in the primate phylogeny. For each *CatSper* gene, branches where d*N*/d*S* >1 are highlighted with a color corresponding to each gene. In cases where multiple genes have a d*N*/d*S* >1 for a single branch, the colors are stacked. (*B*) We detect specific sites under selection in almost every *CatSper* gene. Extracellular domains are marked as dark segments and intracellular domains as light segments. Orange arrows indicate the sites under selection by the Bayes-Empirical-Bayes test in PAML, with a posterior-probability >0.90 (Yang et al. 2005). Selected sites grouped close together are labeled with a bar specifying the number of sites under selection. The number of extracellular and intracellular sites under selection for each gene are tabulated.

d*S*. We found that every branch in the primate lineage, with the exception of gorillas, shows a signature of accelerated evolution for at least one of the *CatSper* genes. These results show that all nine proteins that form the CatSper complex may have evolved under pervasive and strong positive selection across 40 million years of primate evolution.

Patterns of Positively Selected Sites in CatSper Channel Proteins

The positions of the adaptively evolving amino acid sites within a protein can provide insights into the functional properties that are under selection. To identify the adaptively evolving amino acid sites in each of the *CatSper* genes, we used the NSsites models of PAML. For each *CatSper* gene, we tested whether its evolution over the primate phylogeny is consistent with neutral evolution (model M7 and M8a) or with recurrent positive selection (model M8). PAML also uses a Bayesian framework to identify the specific sites in a gene that evolve adaptively under recurrent positive selection. We found that all nine *CatSper* genes show significant evidence of recurrent positive selection with both of these tests (supplementary table 1, Supplementary Material online). In CATSPER1, the site models of PAML did not detect any specific sites under selection, indicating that the selective signature may be broadly dispersed across the gene. *CATSPER2*, *CATSPER3*, and *CATSPER4* each have several sites that evolve under positive selection, but these are not clustered in any particular functional domain (fig. 1*B*; supplementary table 2, Supplementary Material online). The patterns of selection in the auxiliary genes, however, are remarkably different. *CATSPERΔ*, *CATSPERγ*, and *CATSPERϵ* all have several sites under selection, and *CATSPERβ* has a dramatic excess of sites under selection compared to any other *CatSper* gene (supplementary table 2, Supplementary Material online). The extracellular domain CATSPERβ is full of adaptively changing amino acid sites. Because *CATSPERβ* is such a clear outlier with the greatest number of sites under positive selection, we were concerned about spurious false positives generated from alignment errors. A manual inspection of the *CATSPERβ* alignment makes it clear that these are not false positives—we find practically no mis-alignment between the 16 homologous protein sequences (supplementary fig. 2, Supplementary Material online). Little is known about the precise molecular properties of CATSPERβ, other than that it requires CATSPER1 for stable localization to the tail of sperm (Liu et al. 2007). These results show that among all components of the CatSper complex, CATSPERβ is the most frequent target of adaptive evolution.

If the evolution of the sperm hyper-activation machinery is driven by factors in the external environment of sperm, this would manifest as an enrichment of adaptively evolving sites in the extra-cellular regions of the CatSper proteins. We observed different patterns of extra and intracellular changes between the core and auxiliary proteins. Consistent with the patterns observed with CATSPERβ, all of the sites under selection in the auxiliary proteins CATSPERΔ, CATSPERγ, and CATSPERϵ are in extracellular domains. In contrast, only one third of the sites under selection in the core proteins are extracellular (supplementary table 2, Supplementary Material online). These results suggest that in the auxiliary CatSper proteins, adaptive evolution is driven by extracellular interactions. The CatSper complex, therefore, appears locked in an evolutionary conflict that drives its rapid evolution, with *CATSPERβ* being most directly placed at the interface of this conflict.

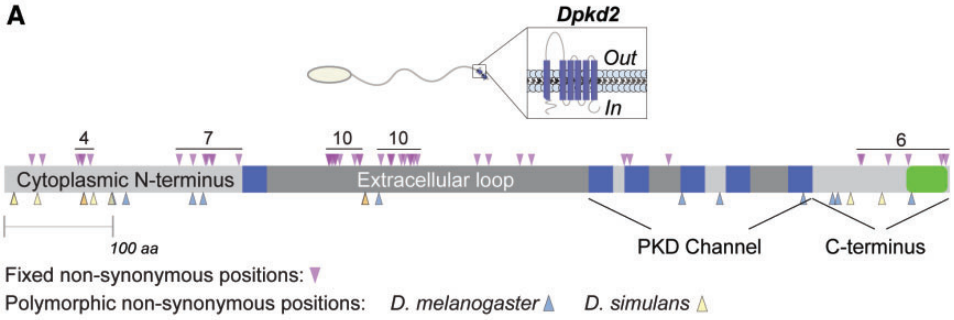
The Sperm Hyper-Activation Channel pkd2 *Evolves Adaptively in* Drosophila

Although all *CatSper* genes are missing in *Drosophila*, sperm hyper-activation remains an essential step in successful fertilization (Köttgen et al. 2011). For successful fertilization, *Drosophila* sperm have to swim from the uterus to the sperm storage organs in the female (Neubaum and Wolfner 1999). Several lines of evidence have shown strong selection on different aspects of *Drosophila* reproductive biology, including the length of sperm and ducts (Lüpold et al. 2016), and the peptides that control sperm storage and release (Findlay et al. 2014). In *D. melanogaster*, the gene *pkd2* is required to hyper-activate sperm so that they can navigate the ducts that lead to the sperm storage organs (Gao et al. 2003; Köttgen et al. 2011; Watnick et al. 2003). *Drosophila pkd2* is highly enriched in the testis, and the primary mutant phenotype is male sterility, along with a subtle effect on smooth muscle contraction (Gao et al. 2004). Like *CatSper*, *Drosophila pkd2* is a Ca^{2+} channel with a PKD domain. If the *CatSper*-independent sperm hyper-activation machinery in *Drosophila* is also involved in a molecular arms race, we predict to find similar signatures of positive selection in *Drosophila pkd2*.

First, we investigated whether *Drosophila pkd2* evolves under positive selection between *D. melanogaster* and its sister species *D. simulans*. Indeed, *Drosophila pkd2* is notable for having the most significant *P* value of all genes in the *Drosophila* genome in previous McDonald–Kreitman comparisons between *D. melanogaster* and *D. simulans* (Begun et al. 2007). To conduct more detailed analyses, we performed direct Sanger sequencing of *Drosophila pkd2* from eight strains of each species (supplementary table 3, Supplementary Material online). While *Drosophila pkd2* is highly polymorphic within populations, there is also a significant excess of fixed nonsynonymous differences between species (McDonald and Kreitman 1991) (fig. 2*A* and *B*). The fixed nonsynonymous changes form discrete clusters. Both the polymorphic and fixed nonsynonymous differences are located mostly outside of the PKD domain, indicating the channel pore function may be well conserved while channel activity modulating sites in *Drosophila pkd2* change rapidly.

Many of the fixed differences in *Drosophila pkd2* are clustered in the first extracellular loop. A recently solved cryo-electron microscopy (cryo-EM) structure of human PKD2 found that this extracellular loop is critical for gating the homo-tetrameric PKD2 channel (Shen et al. 2016). The orientation of the channel exposes this loop to the external

A



B

Region	Length (nt)	MK Table		FE *p*-value	*α*-value
Full sequence	2685	Nf=47 Sf=52	Np=22 Sp=126	$p < 0.0001$	$\alpha = 0.807$
N-terminus	678	Nf=13 Sf=9	Np=11 Sp=23	$p = 0.059$	$\alpha = 0.669$
Extracellular loop	873	Nf=25 Sf=26	Np=3 Sp=48	$p < 0.0001$	$\alpha = 0.935$
PKD channel	711	Nf=3 Sf=10	Np=3 Sp=38	$p = 0.141$	$\alpha = 0.737$
C-terminus	354	Nf=6 Sf=4	Np=5 Sp=15	$p = 0.108$	$\alpha = 0.778$

C

Species	MK Table		FE *p*-value	*α*-value
D. melanogaster	Nf=15 Sf=15	Np=1 Sp=23	$p = 0.002$	$\alpha = 0.957$
D. simulans	Nf=15 Sf=11	Np=3 Sp=25	$p = 0.0075$	$\alpha = 0.868$

D

Test	2ΔlnL	*p*-value
M7/M8	15.02	0.00027
M8a/M8	14.37	0.00038
BUSTED	-	0.0020

Fig. 2.—*pkd2* evolves adaptively in *Drosophila*. (*A*) Fixed nonsynonymous fixed differences between *D. melanogaster* and *D. simulans* are marked with purple arrows above the domain structure. Clusters of fixed nonsynonymous changes are labeled with a bar specifying the number of sites. Polymorphic nonsynonymous changes within each species are marked below the domain structure. The gene span of *Drosophila pkd2* is annotated as a grey bar with blue rectangles marking the transmembrane domains and a green box marking the coiled-coil domain. Extracellular and intracellular domains are in different shades of grey. (*B*) McDonald–Kreitman tests show that *Drosophila pkd2* evolves under positive selection between *D. melanogaster and D. simulans*, and this signal is generated by changes in the extracellular domain. The MK table details nonsynonymous fixed (Nf), synonymous fixed (Sf), nonsynonymous polymorphic (Np), and synonymous polymorphic (Sp) sites. We report the Fisher's exact (FE) *P* value, and the *alpha*-value for each segment of the gene. (*C*) A polarized McDonald–Kreitman test for the extracellular domain of *Drosophila pkd2* demonstrates that this region evolves under positive selection along both lineages. Fixed changes were polarized to the *D. melanogaster* or *D. simulans* lineages using *D. yakuba* as an outgroup species. (*D*) NSsites model tests using PAML show that *Drosophila pkd2* evolves adaptively across many species of *Drosophila*.

environment of the sperm. To see if *Drosophila pkd2* has evolved adaptively in the extracellular loop, we separately analyzed *Drosophila pkd2* in four regions. Analyses of these separate regions show that it is this extracellular loop that drives the signature of positive selection in *Drosophila pkd2*. To further identify the lineages that experienced adaptive evolution, we polarized our MK test using *D. yakuba* as an outgroup species. Our results show that *Drosophila pkd2* underwent adaptive evolution along each of the lineages that lead to *D. melanogaster* and *D. simulans* (fig. 2*C*).

To better understand the special positioning of the fixed nonsynonymous sites, we modeled a three-dimensional molecular structure of *Drosophila pkd2* with the i-Tasser software (Roy et al. 2010) (fig. 3*A*). We used the cryo-EM structure of human PKD2 (Shen et al. 2016) as a template, which includes the region from the first transmembrane domain to the end of the Ca^{2+} channel domain. A plot of the nonsynonymous changes between *D. melanogaster* and *D. simulans* on this predicted structure shows that these sites are not buried in the pore, but are instead directly accessible to the environment of the sperm (fig. 3*B*).

Because our results provide strong evidence that *Drosophila pkd2* evolves under positive selection between *D. melanogaster* and *D. simulans*, we next investigated whether a similar signature of recurrent positive selection is seen across a broader range of *Drosophila* species. We curated sequences

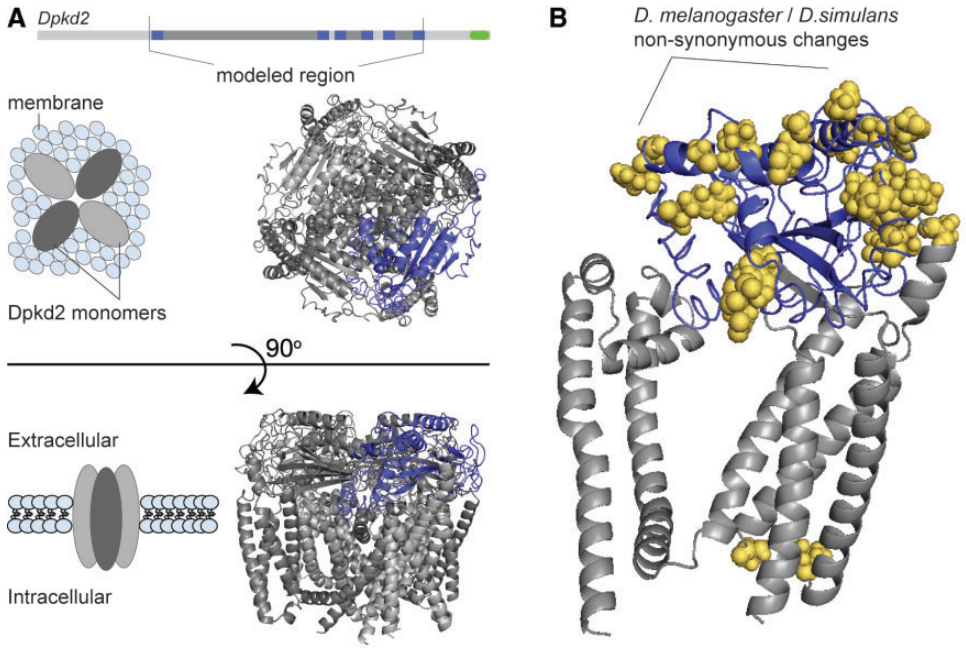


Fig. 3.—A predicted structural model of *Drosophila pkd2* shows that nonsynonymous changes between *D. melanogaster* and *D. simulans* reside on the extracellular faces. (*A*) The homo-tetramer of our predicted *Drosophila* pkd2 structure. The 2-D gene diagram is shaded for the region that could be successfully modeled. The monomers are alternated in shades of grey for contrast. The extracellular loop of one monomer is colored blue. The diagram to the right describes the orientation of the channel. (*B*) Sites that diverge between *D. melanogaster* and *D. simulans* are on the extracellular region of the channel. The base of the monomer is grey, and the extracellular region is blue. The nonsynonymous changes between the two species shown in yellow.

for *Drosophila pkd2* from 17 species of *Drosophila* by identifying homologs from reference genomes and by directly Sanger sequencing *Drosophila pkd2* from additional species. We used the branch and NSsites models of PAML, and BUSTED to test for signatures of selection in *Drosophila pkd2* (fig. 2*D*, supplementary fig. 3, Supplementary Material online) (Murrell et al. 2015). Our results from NSsites and BUSTED analyses provide evidence for episodic positive selection on *pkd2* across *Drosophila* species.

Indel variation is known to drive the rapid diversification of the first exon of *CATSPER1* (Podlaha and Zhang 2003). Because there are several fixed indels between *D. melanogaster* and *D. simulans*, we wanted to know if the variation in the length of the extracellular region is greater than that in the other domains of *Drosophila pkd2*. We compared the variance in length of each segment of *Drosophila pkd2* to the variance in length of the whole gene. We find that most of the change in the length of *Drosophila pkd2* comes from the extracellular domain (supplementary fig. 4, Supplementary Material online). Therefore, the region of *Drosophila pkd2* that contains the most indel variability is the same region that contains the majority of the divergent sites between *D. melanogaster* and *D. simulans*. Together, these results show that, similar to *CatSper* genes, both amino acid substitutions and indel differences in regions exposed to the extracellular environment have played an important role in the rapid evolution of *Drosophila pkd2*.

Primate PKD2, Which Does Not Have a Role in Sperm Hyper-Activation, Does Not Evolve under Positive Selection

In primates, *PKD2* is primarily known for its role in autosomal dominant polycystic kidney disease (Cai et al. 1999). Loss of function in primate *PKD2* causes male sterility, but this sterility is due to abnormal formation of cysts within the testes rather than a sperm hyper-activation defect as in *Drosophila* (Nie and Arend 2014). Because *PKD2* is required for both somatic and gamete development, it is unclear if its evolution would also be driven by sexual selection. Therefore, we analyzed primate PKD2 to see if we could find patterns of selection similar to *Drosophila pkd2* and *CatSpers*.

We first confirmed that primate PKD2 and *Drosophila pkd2* are true orthologs by constructing a maximum-likelihood phylogeny using sequences from the primate *PKD2* family, primate *PKD1* family, *CatSpers*, and *Drosophila pkd2* (supplementary fig. 5*A*, Supplementary Material online). Next, we examined the molecular evolution of primate *PKD2* by collecting sequences from 11 primate species and using the same PAML based approach as with the *CatSper* complex and *Drosophila pkd2*. Consistent with our predictions, neither the branch analysis nor the NSsites models suggest any pattern of positive selection (supplementary table 5 and fig. 5*B* and *C*, Supplementary Material online). Like CatSper proteins, primate PKD2 is part of a complex that contains several other PKD proteins (Tsiokas et al. 1997). We collected sequences and analyzed PKD1, PKD1L1, PKD1L3, PKD2, PKD2L1, and

PKD2L2 by the same tests (we did not obtain enough homologous sequences for *PKD1L2* to complete our analysis). We found no evidence for positive selection in primate PKD genes, with the exception of *PKD1L3* (supplementary table 5, Supplementary Material online). Interestingly, while the rest of the PKD genes are expressed primarily in the heart and kidney, transcript profiling indicates that *PKD1L3* is most strongly expressed in the placenta, suggesting that it may be subject to a different set of selective pressures than the other PKD genes (Li et al. 2003). Nevertheless, our data show that *PKD2* is well conserved in primates, and suggests that the adaptive evolution of *pkd2* in *Drosophila* is driven by its role in sperm hyper-activation.

Discussion

Hyper-activation of sperm, triggered by the opening of flagellar Ca^{2+} channels, is a critical behavioral switch in the race to fertilization. Our analyses show that these Ca^{2+} channels are not conserved, but rather are shaped by recurrent bouts of positive selection. These findings have several implications for understanding the evolutionary and mechanistic aspects of sperm behavior.

First, while much attention has focused on the evolution of the first exon of *CatSper1*, this region reflects only a small part of a much larger pattern. Our analyses show that the entire CatSper complex, including all core and auxiliary proteins, show robust signatures of positive selection in primate lineages. By viewing the molecular evolutionary patterns of the *CatSper* complex as a whole, we find that *CATSPERβ* is the most prevalent target of selection. Little is know about the molecular function of CATSPER*β*, and our results suggest that it plays a key role in the regulation of channel activity. Because we find the majority of positively selected sites in the extracellular region of CATSPER*β* and the other auxiliary proteins, it is likely that interactions with proteins in the seminal fluid or the female reproductive tract modulate CatSper activity.

Second, we find that the nonorthologous sperm hyper-activation Ca^{2+} channels in primates and flies, taxa that are separated by >500 million years, experience remarkably similar selective pressures. We find that the positively selected sites in *Drosophila pkd2* are not buried within the pore, but instead interface with the external environment. Similar to the *CatSper* complex, our results predict that the forces that drive changes in the *Drosophila pkd2* complex involve interacting proteins that modulate the activity of the sperm hyper-activation machinery. These parallel patterns suggest that the sperm hyper-activation machinery may be engaged in the same evolutionary conflicts in a broad diversity of taxa. Interestingly, these are not the first polycystic kidney disease domain containing genes that have been shown to evolve rapidly in the context of reproductive functions. PKDREJ is a sperm receptor for egg-coat proteins which controls the timing of the acrosome reaction (Sutton et al. 2008), and it has evolved rapidly under positive selection across primates and murine species (Hamm et al. 2007; Vicens et al. 2015). PKDREJ may act as an ion channel via its PKD domain, though its molecular function is not well characterized. The later stages of sperm behavior appear to involve two rapidly evolving PKD domain meditated switches: first *CatSper/pkd2* for sperm hyper-activation, and then PKDREJ for the acrosome reaction.

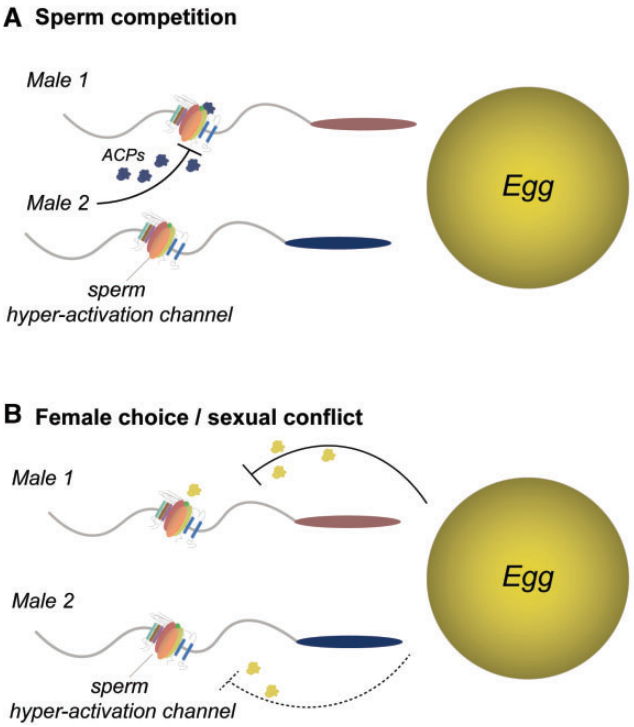


Fig. 4.—Both intermale sperm competition and female choice can drive the rapid evolution of the sperm hyper-activation channels. (*A*) Males that secrete seminal peptides that block the hyper-activation channels of sperm from competing males gain a selective advantage. (*B*) Females may secrete peptides that inhibit sperm hyper-activation channels to control fertilization rates. Males that can prevent the inhibition of sperm hyper-activation gain a selective advantage.

Both sperm competition and female choice have the potential to engage the sperm hyper-activation machinery in sexual conflict. Males may deploy seminal fluid peptides to inactivate the sperm hyper-activation channels of competing sperm, thus providing a massive advantage for their own sperm (fig. 4*A*) (Deborah M. Neubaum and Wolfner 1999; Qazi and Wolfner 2003). Alternatively, females may secrete peptides to inhibit sperm hyper-activation, thus providing a mechanism to modulate fertilization rates and avoid polyspermy (fig. 4*B*) (Aagaard et al. 2013). Under either scenario, the evolutionary arms race between secreted reproductive proteins and sperm hyper-activation channels drives the patterns of recurrent positive selection that we observe in *CatSpers* and *Drosophila pkd2*.

Third, the molecular mechanisms of sperm competition remain unclear despite ample evidence for genetic variation

in sperm competitive abilities both within and between species (Price et al. 1999; Matute and Coyne 2010; Sweigart 2010; Castillo and Moyle 2014). The patterns of recurrent positive selection on *CatSper* and *Drosophila pkd2* complexes make them strong candidates for a role in sperm competition. While testing the role of *CatSpers* in sperm competition in primates has obvious experimental limitations, testing the role of *pkd2* in sperm competition in *Drosophila* is imminently feasible. Our results set the stage for experiments involving transgenic allelic swaps between *D. melanogaster pkd2* and divergent *Drosophila pkd2* alleles from other species in an otherwise *D. melanogaster* genetic background. If the sperm of individuals bearing these interspecies allele swaps compete poorly with other sperm, this may reveal an important role for *Drosophila pkd2* in the molecular mechanisms of sperm competition.

Together, our study introduces the sperm hyper-activation genes as a new class of male reproductive proteins that evolve rapidly. We find parallel patterns of adaptive evolution in non-orthologous proteins that serve as the sperm hyper-activation calcium channels in primates and flies. Identifying the factors that modulate the sperm hyper-activation machinery promises to provide insights into the molecular mechanisms used by the sexes to manipulate sperm behavior to their own advantage.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by the National Institutes of Health (Developmental Biology Training Grant 5T32 HD0741 [to J.C.C.]; R01 GM115914 [to N.P.], a Mario Capecchi endowed assistant professorship [N.P.], and the Pew Biomedical Scholars Program. We thank Nora Brown, Chris Leonard, Mia Levine, and two anonymous reviewers for helpful comments with this manuscript.

Literature Cited

Aagaard JE, Springer SA, Soelberg SD, Swanson WJ, 2013. Duplicate abalone egg coat proteins bind sperm lysin similarly, but evolve oppositely, consistent with molecular mimicry at fertilization. PLoS Genet. 9(2):e1003287.

Begun DJ, 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. 5(11):e310.

Cai X, Clapham DE, 2008. Evolutionary genomics reveals lineage-specific gene loss and rapid evolution of a sperm-specific ion channel complex: CatSpers and CatSperβ. PLoS One 3(10):e3569.

Cai X, Wang X, Clapham DE, 2014. Early evolution of the eukaryotic Ca2+ signaling machinery: conservation of the catsper channel complex. Mol Biol Evol. 31:2735–2740.

Cai Y, 1999. Identification and characterization of polycystin-2, the PKD2 gene product. J Biol Chem. 274(40):28557–28565.

Carlson AE, 2009. Pharmacological targeting of native catsper channels reveals a required role in maintenance of sperm hyperactivation. PLoS One 4(8):e6844.

Castillo DM, Moyle LC, 2014. Intraspecific sperm competition genes enforce post-mating species barriers in Drosophila. Proc Biol Sci. 281(1797):2014–2050.

Chen Z-X, et al. 2014. Comparative validation of the *D. melanogaster* modENCODE transcriptome annotation. Genome Res. 24(7):1209–1223.

Chung J-J, et al. 2017. CatSperζ regulates the structural continuity of sperm Ca2+ signaling domains and is required for normal fertility. eLife 6:e23082.

Chung J-J, Navarro B, Krapivinsky G, Krapivinsky L, Clapham DE, 2011. A novel gene required for male fertility and functional CATSPER channel formation in spermatozoa. Nat Commun. 2:153.

Clark NL, Findlay, GD, Yi X, MacCoss MJ, Swanson WJ, 2007. Duplication and selection on abalone sperm lysin in an allopatric population. Mol Biol Evol. 24(9):2081–2090

Clark NL, Aagaard JE, Swanson WJ, 2006. Evolution of reproductive proteins from animals and plants. Reproduction 131(1):11–22.

Cosson J, et al. 2008. Marine fish spermatozoa: racing ephemeral swimmers. Reproduction 136(3):277–294.

Dziminski MA, Roberts JD, Beveridge M, Simmons LW, 2009. Sperm competitiveness in frogs: slow and steady wins the race. Proc Biol Sci. 276:3955–3961.

Edgar RC, 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113

Findlay GD, et al. 2014. Evolutionary rate covariation identifies new members of a protein network required for *Drosophila melanogaster* female post-mating responses. PLoS Genet. 10(1):e1004108.

Fisher HS, Jacobs-Palmer E, Lassance J-M, Hoekstra HE, 2016. The genetic basis and fitness consequences of sperm midpiece size in deer mice. Nat Commun. 7:13652.

Gao Z, Joseph E, Ruden DM, Lu X, 2004. Drosophila Pkd2 is haploid-insufficient for mediating optimal smooth muscle contractility. J Biol Chem. 279(14):14225–14231.

Gao Z, Ruden DM, Lu X, 2003. PKD2 cation channel is required for directional sperm movement and male fertility. Curr Biol. 13(24):2175–2178.

Gavrilets S, 2000. Rapid evolution of reproductive barriers driven by sexual conflict. Nature 403(6772): 886–889.

Gomendio M, Martin-Coello J, Crespo C, Magana C, Roldan ERS, 2006. Sperm competition enhances functional capacity of mammalian spermatozoa. Proc Natl Acad Sci U S A. 103(41):15113–15117.

Hamm D, Mautz BS, Wolfner MF, Aquadro CF, Swanson WJ, 2007. Evidence of amino acid diversity: enhancing selection within humans and among primates at the candidate sperm-receptor gene PKDREJ. Am J Hum Genet. 81(1):44–52.

Hemmer LW, Blumenstiel JP, 2016. Holding it together: rapid evolution and positive selection in the synaptonemal complex of Drosophila. BMC Evol Biol. 16:91.

Ho H-C, Granish KA, Suarez SS, 2002. Hyperactivated motility of bull sperm is triggered at the axoneme by Ca2+ and not cAMP. Dev Biol. 250(1):208–217.

Howard DJ, Palumbi SR, Birge LM, Manier MK. 2009. 9 - Sperm and speciation. In: Birkhead TR, Hosken DJ, and Pitnick S, editors. Sperm Biology. London: Academic Press. p. 367–403 [cited 2017 Jan 24]. Available from: www.sciencedirect.com/science/article/pii/B9780123725684000094.

Ishijima S, 2011. Dynamics of flagellar force generated by a hyperactivated spermatozoon. Reproduction 142(3):409–415.

Jin J-L, et al. 2005. Catsper3 and Catsper4 encode two cation channel-like proteins exclusively expressed in the testis. Biol Reprod. 73(6):1235–1242.

Köttgen M, et al. 2011. Drosophila sperm swim backwards in the female reproductive tract and are activated via TRPP2 ion channels. PLoS One 6(5):e20031.
Kresge N, Vacquier VD, Stout CD, 2001. Abalone lysin: the dissolving and evolving sperm protein. BioEssays 23(1):95–103.
Lee Y-H, Ota T, Vacquier VD, 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. Mol Biol Evol. 12(2):231–238.
Letunic I, Doerks T, Bork P, 2015. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 43:D257–D260.
Li A, Tian X, Sung S-W, Somlo S, 2003. Identification of two novel polycystic kidney disease-1-like genes in human and mouse genomes. Genomics 81(6): 596–608.
Librado P, Rozas J, 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11):1451–1452.
Liu J, Xia J, Cho K-H, Clapham DE, Ren D, 2007. CatSperβ, a novel transmembrane protein in the CatSper channel complex. J Biol Chem. 282(26):18945–18952.
Lobley A, Pierron V, Reynolds L, Allen L, Michalovich D, 2003. Identification of human and mouse CatSper3 and CatSper4 genes: characterisation of a common interaction domain and evidence for expression in testis. Reprod Biol Endocrinol. 1:53.
Loux SC, et al. 2013. CatSper and the relationship of hyperactivated motility to intracellular calcium and pH kinetics in equine sperm. Biol Reprod. 89(5):123.
Lüpold S, et al. 2016. How sexual selection can drive the evolution of costly sperm ornamentation. Nature 533(7604):535–538.
Madden T. 2003. The BLAST Sequence Analysis Tool. National Center for Biotechnology Information (US) [cited 2017 Mar 23]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK21097/.
Markova-Raina P, Petrov D, 2011. High sensitivity to aligner and high rate of false positives in the estimates of positive selection in the 12 Drosophila genomes. Genome Res. 21(6):863–874.
Matute DR, Coyne JA, 2010. Intrinsic reproductive isolation between two sister species of Drosophila. Evolution 64:903–920.
McDonald JH, Kreitman M, 1991. Adaptive protein evolution at the Adh locus in Drosophila. Nature 351(6328):652–654.
Montoto LG, et al. 2011. Sperm competition, sperm numbers and sperm quality in muroid rodents. PLoS One 6(3):e18173.
Moyle LC, Jewell CP, Kostyun JL, 2014. Fertile approaches to dissecting mechanisms of premating and postmating prezygotic reproductive isolation. Curr Opin Plant Biol. 18:16–23.
Murrell B, et al. 2015. Gene-wide identification of episodic selection. Mol Biol Evol. 32(5):1365–1371.
Neubaum DM, Wolfner MF, 1999. Mated *Drosophila melanogaster* females require a seminal fluid protein, Acp36DE, to store sperm efficiently. Genetics 153(2):845–857.
Neubaum DM, Wolfner MF, 1999. Wise, winsome, or weird? Mechanisms of sperm storage in female animals. Curr Top Dev Biol. 41:67–97.
Nguyen TMD, et al. 2014. Central role of 5′-AMP-activated protein kinase in chicken sperm functions. Biol Reprod. 91(5):121.
Nie X, Arend LJ, 2014. Novel roles of Pkd2 in male reproductive system development. Differentiation 87(3–4):161–171.
Notredame C, Higgins DG, Heringa J, 2000. T-coffee: a novel method for fast and accurate multiple sequence alignment1. J Mol Biol. 302(1):205–217.
O'Brien ED, Krapf D, Cabada MO, Visconti PE, Arranz SE, 2011. Transmembrane adenylyl cyclase regulates amphibian sperm motility through protein kinase A activation. Dev Biol. 350:80–88.
Ooi EH, Smith DJ, Gadelha H, Gaffney EA, Kirkman-Brown J, 2014. The mechanics of hyperactivation in adhered human sperm. R Soc Open Sci. 1(2):140–230
Panhuis TM, Clark NL, Swanson WJ, 2006. Rapid evolution of reproductive proteins in abalone and Drosophila. Philos Trans R Soc B Biol Sci. 361(1466):261–268.
Parker GA, Partridge L, 1998. Sexual conflict and speciation. Philos Trans R Soc Lond B Biol Sci. 353(1366):261–274.
Perelman P, et al. 2011. A molecular phylogeny of living primates. PLoS Genet. 7(3):e1001342.
Podlaha O, Webb DM, Tucker PK, Zhang J, 2005. Positive selection for indel substitutions in the rodent sperm protein Catsper1. Mol Biol Evol. 22(9):1845–1852.
Podlaha O, Zhang J, 2003. Positive selection on protein-length in the evolution of a primate sperm ion channel. Proc Natl Acad Sci U S A. 100(21):12241–12246.
Price CSC, Dyer KA, Coyne JA, 1999. Sperm competition between Drosophila males involves both displacement and incapacitation. Nature 400:449–452.
Qazi MCB, Wolfner MF, 2003. An early role for the *Drosophila melanogaster* male seminal protein Acp36DE in female sperm storage. J Exp Biol. 206(Pt 19):3521–3528.
Qi H, et al. 2007. All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. Proc Natl Acad Sci U S A. 104(4):1219–1223.
Quill TA, Ren D, Clapham DE, Garbers DL, 2001. A voltage-gated ion channel expressed specifically in spermatozoa. Proc Natl Acad Sci U S A. 98(22):12527–12531.
Ren D, et al. 2001. A sperm ion channel required for sperm motility and male fertility. Nature 413(6856):603–609.
Roy A, Kucukural A, Zhang Y, 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 5(4):725–738.
Shen PS, et al. 2016. The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. Cell 167(3):763–773.e11.
Sievers F, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7:539
Slater GSC, Birney E, 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.
Suarez S, Ho H-C, 2003. Hyperactivated motility in sperm. Reprod Domest Anim. 38(2):119–124.
Suarez SS, Katz DF, Owen DH, Andrew JB, Powell RL, 1991. Evidence for the function of hyperactivated motility in sperm. Biol Reprod. 44(2):375–381.
Sutton KA, Jungnickel MK, Florman HM, 2008. A polycystin-1 controls postcopulatory reproductive selection in mice. Proc Natl Acad Sci. 105(25):8661–8666.
Suyama M, Torrents D, Bork P, 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34(Web Server):W609–W612.
Swanson WJ, Nielsen R, Yang Q, 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol Biol Evol. 20(1):18–20
Swanson WJ, Vacquier VD, 1997. The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. Proc Natl Acad Sci U S A. 94(13):6724–6729.
Swanson WJ, Vacquier VD, 2002. The rapid evolution of reproductive proteins. Nat Rev Genet. 3(2):137–144.
Sweigart AL, 2010. The genetics of postmating, prezygotic reproductive isolation between *Drosophila virilis* and *D. americana*. Genetics 184(2):401–410.
Tsiokas L, Kim E, Arnould T, Sukhatme VP, Walz G, 1997. Homo- and heterodimeric interactions between the gene products of PKD1 and PKD2. Proc Natl Acad Sci U S A. 94(13):6965–6970.
Turner LM, Chuong EB, Hoekstra HE, 2008. Comparative analysis of testis protein evolution in rodents. Genetics 179(4):2075–2089.
Vicens A, Montoto LG, Couso-Ferrer F, Sutton KA, Roldan ERS, 2015. Sexual selection and the adaptive evolution of PKDREJ protein in primates and rodents. Mol Hum Reprod. 21(2):146–156.

Wang H, Liu J, Cho K-H, Ren D, 2009. A novel, single, transmembrane protein CATSPERG is associated with CATSPER1 channel protein. Biol Reprod. 81(3):539–544.
Watnick TJ, Jin Y, Matunis E, Kernan MJ, Montell C, 2003. A flagellar polycystin-2 homolog required for male fertility in Drosophila. Curr Biol. 13(24):2179–2184.
Wilburn DB, Swanson WJ, 2016. From molecules to mating: rapid evolution and biochemical studies of reproductive proteins. J Proteomics 135:12–25.
Yanagimachi R, 1970. In vitro capacitation of golden hamster spermatozoa by homologous and heterologous blood sera. Biol Reprod. 3(2):147–153.
Yanagimachi R, 1969. In vitro capacitation of hamster spermatozoa by follicular fluid. J Reprod Fertil. 18(2):275–286.
Yang Y, Lu X, 2011. Drosophila sperm motility in the reproductive tract. Biol Reprod. 84(5):1005–1015.
Yang Z, 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24(8):1586–1591.
Yang Z, Wong WSW, Nielsen R, 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22(4):1107–1118.
Zhou J, et al. 2015. The semen pH affects sperm motility and capacitation. PLoS One 10(7):e0132974.

Associate editor: Esther Betran