Positive Selection for Indel Substitutions in the Rodent Sperm Protein Catsper1

Ondrej Podlaha,[1] David M. Webb,*[1] Priscilla K. Tucker,*† and Jianzhi Zhang**

*Department of Ecology and Evolutionary Biology, University of Michigan; and
†Museum of Zoology, University of Michigan

Catsper1 is a voltage-gated calcium channel located in the plasma membrane of the sperm tail and is necessary for sperm motility and fertility in mice. We here examine the evolutionary pattern of Catsper1 from nine species of the rodent subfamily Murinae of family Muridae. We show that the rate of insertion/deletion (indel) substitutions in exon 1 of the gene is 4–15 times that in introns or neutral genomic regions, suggesting the presence of strong positive selection that promotes fixations of indel mutations in exon 1. The number of indel polymorphisms within species appears higher than expected from interspecific comparisons, although there are too little data to provide a statistically significant conclusion. These results, together with an earlier report in primates, indicate that positive selection promoting length variation in Catsper1 may be widespread in mammals. A structural model of Catsper1 suggested the importance of the exon 1–encoded region in regulating channel inactivation, which may affect sperm mobility and sperm competition. Our findings provide a necessary foundation for future experimental investigations of Catsper1's function in sperm physiology and role in sperm competition using rodent models.

Introduction

The occurrence of positive Darwinian selection at the molecular level has been reported in many genes, particularly among those involved in immunity or reproduction (Hughes and Nei 1988; Lee, Ota, and Vacquier 1995; Rooney and Zhang 1999; Wyckoff, Wang, and Wu 2000; Swanson and Vacquier 2002; Zhang and Rosenberg 2002; J. Zhang, Y. P. Zhang, and Rosenberg 2002; Wolfe and Li 2003; Wang and Zhang 2004; Zhang and Webb 2004). In all these genes, it was shown that positive selection promotes amino acid replacements. In theory, positive selection may also act upon other types of sequence changes. In a previous report, we showed that positive selection was responsible for elevated fixation rates of insertions and deletions (indels) in the primate *Catsper1* gene (Podlaha and Zhang 2003), which codes for a voltage-gated calcium ion channel essential for sperm mobility (Ren et al. 2001). Mice with the *Catsper1* gene deleted are sterile (Ren et al. 2001). Interestingly, the excess of indel substitutions was observed only in the first exon of *Catsper1*, which codes for the intracellular N-terminus of the channel. Based on the predicted structure of Catsper1 that includes six transmembrane domains and intracellular N- and C-termini (Ren et al. 2001), we hypothesized that the length of the N-terminus may affect the rate of channel inactivation (Podlaha and Zhang 2003), as in the so-called "ball-and-chain" model (Bezanilla and Armstrong 1977; Hoshi, Zagotta, and Aldrich 1990; Zagotta, Hoshi, and Aldrich 1990). Thus, the indels in Catsper1 may influence sperm mobility and consequently be an important factor in sperm competition (Podlaha and Zhang 2003). The mouse Catsper1 protein is found in the plasma membrane of the principle piece of the sperm tail (Ren et al. 2001). But this location does not tell whether the products of one allele or both alleles of the gene are present in an individual sperm, as it is possible that the progenitor diploid cells of sperm express both alleles whose products get incorporated into the sperm. A recent study, however, showed that *Catsper1* expression is confined to the postmeiotic stage of spermatogenesis (Schultz, Hamra, and Garbers 2003), strongly suggesting that only the product of one allele is present in each sperm. This provides the possibility of haploid selection on (Joseph and Kirkpatrick 2004) and meiotic drive of (Hartl and Clark 1997) *Catsper1* if the two alleles confer different fitnesses. These hypotheses about the function of and selection on *Catsper1* are difficult to test in primates because of the difficulties in housing animals and in obtaining biological materials. Rodents would be more convenient and suitable for addressing these questions. We here sequence the *Catsper1* gene in multiple species belonging to the rodent subfamily Murinae of family Muridae and multiple individuals of the species *Mus musculus* and *Mus macedonicus*. We show that the phenomenon of accelerated indel substitutions observed in primates also occurs in rodents. This opens the possibility for experimental examination of the functional and evolutionary significance of the indel substitutions in rodent Catsper1.

Materials and Methods

Polymerase Chain Reaction and Sequencing

Exon 1 (933 nt in *M. musculus*) of the *Catsper1* gene was amplified by polymerase chain reaction (PCR) from the genomic DNAs of *Mus pahari*, *Mus spretus*, *Mus spicilegus*, *M. macedonicus* (two individuals), *Mus cervicolor*, *Mus cookii*, *Mus caroli*, *Mus musculus musculus* (one individual from Loppi, Finland; one individual from Bratislava, Slovakia; one individual from Studenec, Czech Republic), *Mus musculus domesticus* (one individual from Centreville, Md.; one individual from Canton Vaud, Switzerland; one individual from Bern, Switzerland; one individual from Nurnberg, Germany; one individual from Erfoud, Morocco; one individual from Tubingen, Germany), *Mus musculus molossinus* (Fukuoka, Japan), and *Mus musculus castaneus* (Thailand). PCR products were purified and sequenced in both directions using an automated DNA sequencer. The nucleotide sequences from the above 19 individuals, mouse strain C57BL/6J (GenBank accession number NM_139301), and *Rattus norvegicus* (GenBank accession number XM_219698) were

[1] These authors contributed equally to this work.

Key words: evolutionary rate, indel, positive selection, sperm, rodents.

E-mail: jianzhi@umich.edu.

Mol. Biol. Evol. 22(9):1845–1852. 2005
doi:10.1093/molbev/msi178
Advance Access publication June 1, 2005

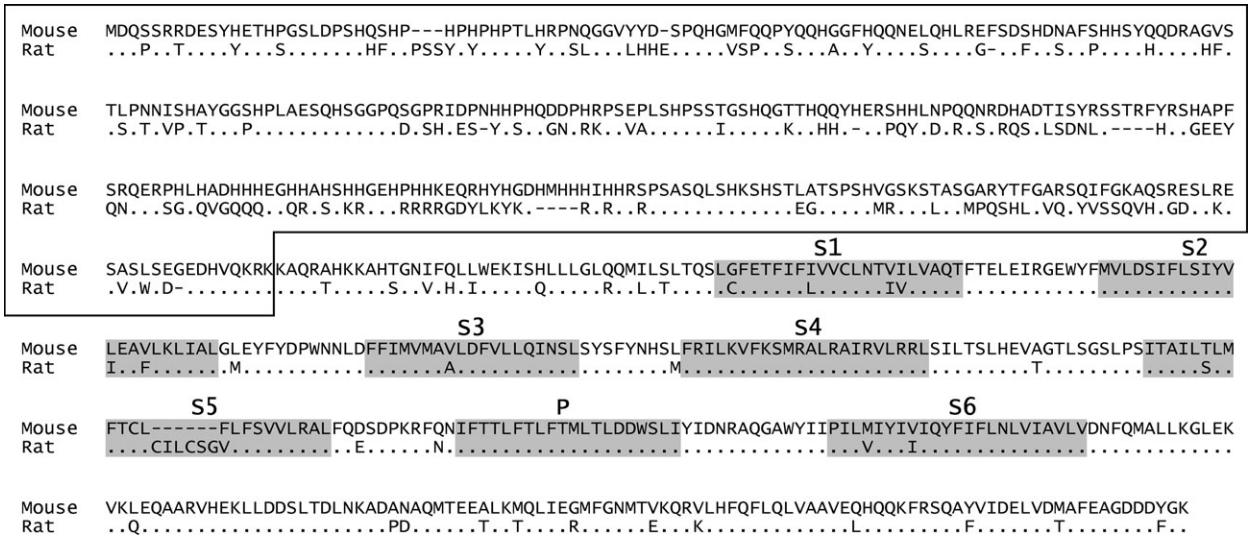

FIG. 1.—Alignment of the full-length Catsper1 sequences of the mouse (*Mus musculus*) and rat (*Rattus norvegicus*). Dots represent amino acids identical to the first sequence, and dashes represent alignment gaps. The region encoded by exon 1 is boxed. The six transmembrane domains (S1–S6) and the pore-forming region (P) are shaded.

aligned based on the protein sequence alignment made by ClustalX (Thompson et al. 1997). The complete exon 1 sequences except for the first 18 and last 17 codons were obtained and analyzed.

Estimating Rates of Indel Substitutions

We assumed that the phylogenic relationships of the rodent species used in this study are the same as those presented in figure 2*B* of Tucker, Sandstedt, and Lundrigan (2005), which was the maximum likelihood tree from a combined analysis of eight nuclear and mitochondrial genes. The number of indel substitutions that occurred in the rodent *Catsper1* sequences was counted using the parsimony principle based on the species tree. In cases where multiple equally parsimonious solutions for a single alignment gap were encountered, all solutions were weighted equally. Because there are no unequivocal estimates of absolute divergence times among the *Mus* taxa included in our study, we calculated the rate of indel substitutions per unit of d_S distance (number of synonymous substitutions per synonymous site), rather than per unit of time. This approach allows us to make reliable comparisons of indel rates across all lineages and avoid the imprecision inherent in divergence time estimates. Branch lengths for a given tree topology were calculated using the BNBS program (Zhang, Rosenberg, and Nei 1998). The genomic average rate of indel substitutions in mouse and rat (Rat Genome Sequencing Consortium 2004) was used as an approximation of the neutral rate of indel substitutions.

Testing for Positive Selection

Positive selection promoting single-nucleotide and indel substitutions in exon 1 of *Catsper1* was tested. The maximum likelihood method implemented in PAML (Yang 1997) was used to test positive selection for single-nucleotide substitutions. We tested whether indel mutations have been fixed at a higher than neutral rate in *Catsper1* exon 1 by comparing it with the genomic indel rate following Podlaha and Zhang (2003). We also modified the Mcdonald and Kreitman (1991) test to examine whether the ratio of the number of indels to the number of synonymous differences is the same for intraspecific and interspecific data. A significant difference between these ratios suggests variable selective pressures on indel mutations at the intraspecific and interspecific levels.

Results

We determined from the mouse (*M. musculus*) genome sequence that *Catsper1* is located in chromosome 19 and is composed of 12 exons totaling 2,061 nt. A comparison of the mouse *Catsper1* and its human ortholog revealed a large number of alignment gaps in the first exon, whereas the rest of the sequences remained relatively conserved. Because our previous study in primates showed that exon 1, which codes for the majority of the N-terminal intracellular region of the ion channel, has been under positive selection for indel substitutions, we here focus on the same exon in rodents. In fact, the alignment of the mouse and rat Catsper1 sequences showed a high frequency of both amino acid replacements and indels in the region encoded by exon 1 (fig. 1). We obtained DNA sequences of *Catsper1* exon 1 from 19 individuals of 9 *Mus* species. No heterozygous individuals were found. In addition, we downloaded from GenBank one *Catsper1* sequence from *M. musculus* and one from *R. norvegicus*. Thirteen different nucleotide sequences and 12 different protein sequences were found among the 21 *Catsper1* exon 1 sequences obtained (fig. 2). These raw sequences varied in length from 808–856 nt, with the open reading frame being intact in each of them. Sequence alignment by ClustalX was straightforward and robust, without the need for manual adjustments. The alignment of the sequences, spanning 910 nt, contained numerous gaps ranging in number from five in *M. pahari* to 10 in one of the *M. musculus*

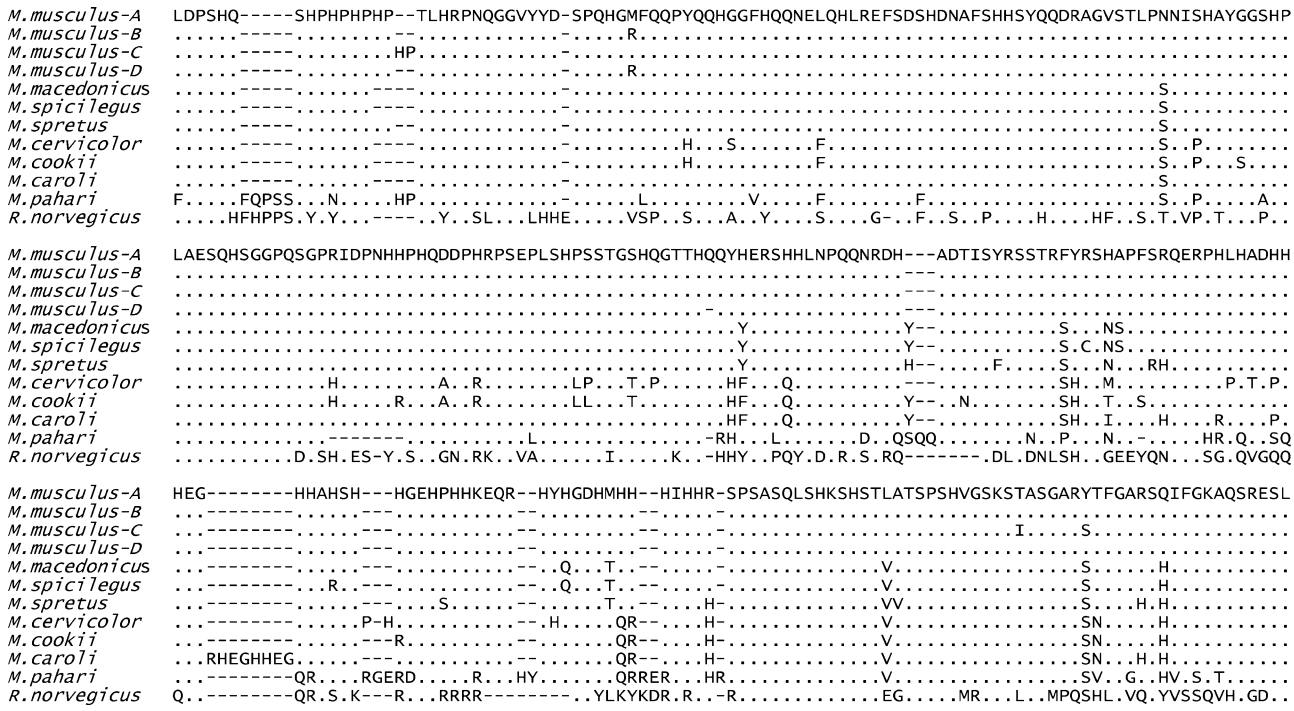

FIG. 2.—Alignment of translated rodent *Catsper1* exon 1 sequences obtained in this study. Only nonredundant protein sequences are shown. Dots represent amino acids identical to the first sequence, and dashes represent alignment gaps. *Mus musculus-A* sequence was downloaded from GenBank (mouse strain C57BL/6J, GenBank accession number NM_139301) and represents a haplotype shared by all six *Mus musculus domesticus* mice and one *Mus musculus castaneus* individual. *Mus musculus-B* represents a haplotype carried by two *Mus musculus musculus* individuals, one from Czech Republic and the other from Slovakia. *Mus musculus-C* represents a haplotype of one *M. m. musculus* from Finland. *Mus musculus-D* represents a unique haplotype found in *Mus musculus molossinus* from Japan. Sequences from two *Mus macedonicus* individuals were identical and are represented by a single *M. macedonicus* haplotype.

individuals (fig. 2). Interestingly, within the *M. musculus* species, there are two indel polymorphisms. One of them involves an indel of one codon (found in *M. m. molossinus* from Japan), whereas the other involves two codons (found in *M. m. musculus* from Finland). A gene tree of the 21 sequences was reconstructed using the neighbor-joining method (Saitou and Nei 1987), which shows a branching pattern that is largely consistent with the assumed species phylogeny (fig. 3), indicating that the sequences are indeed orthologous to each other. Furthermore, all allelic sequences within species cluster in species-specific clades. Using the parsimony principle, we inferred the number of indels in *Catsper1* exon 1 and mapped them onto the species tree (fig. 3). For example, the branch that leads to *M. musculus* has 2.17 indels because this branch is representative of 12 individuals from the *M. musculus* species, and while all of them share two indels, two individuals show one unique indel each in the alignment. Therefore, averaged across all 12 individuals, there are 0.17 individual-specific indels in addition to the two indels shared by all. Note that parsimony makes our inference of the number of indels conservative. A total of 22.47 indel substitutions were found throughout the tree (fig. 3). We then calculated the average number of indel events for different parts of the tree, dividing the taxa into groups M1 (*M. musculus*, *M. macedonicus*, *M. spicilegus*, and *M. spretus*), M2 (*M. cervicolor*, *M. cookii*, and *M. caroli*), M3 (*M. pahari*), and R (*R. norvegicus*) (table 1 and fig. 3). The average number of indel substitutions from the common ancestor of a group to all terminal nodes in the group is computed by averaging the number of indels between sister lineages in the group in a hierarchical manner. For example, the number for M1 is ((((0 + 0)/2 + 0) + 1)/2 + 0 + 2.17)/2 = 1.34. The average

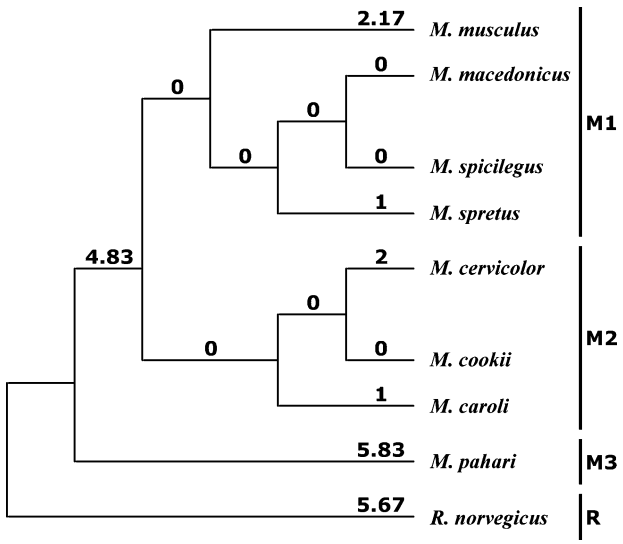

FIG. 3.—Indel substitutions in the evolution of rodent *Catsper1* exon 1. The phylogeny of the nine rodent species included in our study was assumed to follow Tucker, Sandstedt, and Lundrigan (2005). Numbers above branches represent parsimony-inferred numbers of indel substitution events. Capital letters on the right depict groupings of rodent species.

Table 1
Test of Positive Selection for Indel Substitutions in Rodent *Catsper1* Exon 1

Comparison Between Groups[a]	d_S[b]	Expected Number of Indels in *Catsper1* Exon 1[c]		Observed Number of Indels in *Catsper1* Exon 1	Probability[d]	
		Genomic	*Catsper1*		Genomic	*Catsper1*
M1-M2	0.0383	0.33	0.41	2.34	4.4×10^{-2}	6.3×10^{-2}
M1-M3	0.0865	0.75	0.92	12.00	5.1×10^{-10}	4.3×10^{-9}
M1-R	0.2648	2.29	2.81	11.84	2.8×10^{-5}	1.7×10^{-4}
M2-M3	0.0728	0.63	0.77	11.66	8.6×10^{-11}	7.3×10^{-10}
M2-R	0.2511	2.17	2.67	11.50	1.7×10^{-5}	1.1×10^{-4}
M3-R	0.2609	2.25	2.77	11.50	2.5×10^{-5}	1.5×10^{-4}

[a] See figure 3 for definitions of the groups.

[b] Average d_S distance between groups.

[c] Expected numbers of indels were calculated by multiplying the neutral substitution rate of 3*n* indels per d_S, average d_S between groups, and the length of the sequence concerned. To make the statistical test more conservative, we used the sequence length of 759 nt, after removing all alignment gaps. Two neutral rates, one estimated from the mouse-rat genomic comparison and the other from introns of *Catsper1*, were used.

[d] The probability of the observation given the expectation was computed under the assumption that the number of indels follows a Poisson distribution.

number of indels between species of two groups was subsequently computed by adding the number of indels between the common ancestors of the two groups and the numbers from the common ancestors to terminal nodes (table 1). For example, the number of indels between M1 and R is 4.83 + 5.67 + 1.34 = 11.84 (table 1 and fig. 3).

To test whether the rate of indel substitutions in *Catsper1* exon 1 is significantly greater than the neutral expectation, it is necessary to first estimate the neutral rate of indel substitutions. For this we used the recently published mouse-rat genomic comparison (Cooper et al. 2004). It was estimated that indel substitutions in bulk genomic DNA have a frequency of 12.46 indels/kb between mouse and rat. This rate of indel substitution should be very close to the neutral rate as only 1%–1.5% of the mammalian genome are coding sequences (Human Genome Sequencing Consortium 2001). Because indels of sizes other than multiples of 3 nt (3*n*) will almost certainly be highly deleterious in coding regions, it is necessary to compute the neutral substitution rate for 3*n* indels. The proportion of 3*n* indels is less than one-third of all indels because the frequency of indels declines rapidly with size in a nonlinear fashion. Studies of the indel size distribution in mammalian genomes show that approximately 17%–19% of all indels are 3*n* indels (Britten 2002; Silva and Kondrashov 2002; Makova, Yang, and Chiaromonte 2004). In particular, the proportion of 3*n* indels for mouse and rat was 17.3% (Makova, Yang, and Chiaromonte 2004). Applying this value to the neutral rate of indel substitutions, we estimated that the neutral substitution rate of 3*n* indels is 2.156 indels/kb between mouse and rat.

The divergence times for most taxa included in our study are either unknown or ambiguous. Therefore, rather than considering substitution rates of indels per site per year, we calculated the number of indel substitutions per site per unit of synonymous distance (d_S). Using d_S instead of absolute time is a more reliable way for comparing indel substitution rates among different lineages because it allows us to avoid errors in divergence time estimation. Applying the genomic average $d_S = 0.19$ between mouse and rat (Rat Genome Sequencing Consortium 2004), the neutral substitution rate of 3*n* indels becomes $2.156 \times 10^{-3}/0.19 = 0.01135$ per site per unit of d_S.

With the above neutral substitution rate of 3*n* indels, we computed the expected number of 3*n* indels in exon 1 of *Catsper1* if all 3*n* indels are neutral. For example, we expect to observe $0.01135 \times 759 \times 0.2648 = 2.29$ indels between groups M1 and R (table 1). Here, 0.01135 is the neutral substitution rate of 3*n* indels per site per unit of d_S, 759 is the length of exon 1 (alignment gaps excluded for conservative statistical tests), and 0.2648 is the average d_S in *Catsper1* exon 1 between groups M1 and R (fig. 3). The observed number of indels between M1 and R is 11.84, which is significantly greater than the neutral expectation ($P < 0.001$, Poisson test; table 1). Similarly, we computed the expected numbers of indels for other between-group comparisons (table 1). The observed numbers of indels in exon 1 of rodent *Catsper1* are 5–18 times the neutral expectations (table 1). Except for one comparison that yields marginal significance, all other comparisons show highly significant differences between the observed and expected numbers of indels (table 1).

The presence of increased number of indels in exon 1 of *Catsper1*, when compared to the genomic average, could be explained by positive selection promoting the fixation of indel mutations and/or by a higher indel mutation rate at the *Catsper1* locus. To discriminate between these two possibilities, we assessed the indel mutation rate in the *Catsper1* locus by examining the intron sequences. Because the majority of sites in introns are more or less neutral, the indel substitution rate in *Catsper1* introns reflects the local indel mutation rate at the locus. We found a total of 106 indels in the 6,739-nt-long alignment of the 11 introns of *Catsper1* between mouse and rat. Applying the genomic neutral indel substitution rate to the *Catsper1* introns, we would expect to see 86.2 indels between mouse and rat, which is significantly different from the observed number ($P = 0.02$, Poisson test). The indel mutation rate at *Catsper1* is 23% greater than the genomic average. However, even when this elevated mutation rate is considered, we still find the observed number of indels in *Catsper1* exon 1 to be 4–15 times the neutral expectation at the locus. Except for

the M1-M2 comparison, all other group comparisons are statistically significant (table 1). These results strongly suggest that 3*n* indels have been favored by positive selection in rodent *Catsper1* exon 1. We further tested whether indels of a particular size are preferentially fixed in exon 1 of *Catsper1* by comparing the *Catsper1* data to a genomic indel size distribution (Makova, Yang, and Chiaromonte 2004). No significant difference was detected ($\chi^2 = 0.97$, df = 2, $P > 0.1$), although this could be due to a relatively small sample size.

Comparison of intraspecific polymorphism data and interspecific divergence data is often used to test for natural selection. In particular, one can compare the ratio of the number of nonsynonymous differences to that of synonymous differences within species and between species (Mcdonald and Kreitman 1991). Here we modify this test by replacing nonsynonymous differences with indel differences, where each indel is counted as a single event regardless of the indel size. From the intraspecific data of *M. musculus* and *M. macedonicus*, we found two indel polymorphisms and one synonymous polymorphism. The number of fixed differences between the two species is two and four for indels and synonymous changes, respectively. The interspecific indel/synonymous ratio of 2:4 is not significantly different from the intraspecific ratio of 2:1 (Fisher's exact test, $P > 0.2$). Therefore, this modified Mcdonald and Kreitman test does not reject the hypothesis that the selective intensity on indels is identical between the interspecific and intraspecific data. The above test has a low power here because both intraspecific and interspecific differences are small in our data. When the entire tree of figure 3 is considered, the indel/synonymous ratio of 22:61 is found. However, this ratio is again not significantly different from the intraspecific indel/synonymous ratio of 2:1 found in *M. musculus* and *M. macedonicus* (Fisher's exact test, $P > 0.1$). Nevertheless, the ratio appears higher for the intraspecific data than for the interspecifc data at face value.

The sequences of rodent *Catsper1* exon 1 show many amino acid replacements in addition to indels. Over 50% of pairwise comparisons have more nonsynonymous substitutions per nonsynonymous site (d_N) than synonymous substitutions per synonymous site (d_S) (fig. 4), with an average d_N/d_S ratio of all comparisons being 1.18. The standard Mcdonald and Kreitman test, comparing nonsynonymous/synonymous ratios within species (3:1 for *M. musculus* and *M. macedonicus*) and between species (229:61 for the entire tree), showed no significant differences ($P > 0.5$), suggesting that the selective pressure on nonsynonymous changes is indistinguishable between the intraspecific and interspecific levels. The average d_N/d_S ratio of approximately 1 in the pairwise comparisons may be a result of several factors. One possibility is that several sites in exon 1 of *Catsper1* are under positive selection while others are under purifying selection. Alternatively, the region may be experiencing a relaxation of functional constraint. It is unlikely that exon 1 of *Catsper1* would evolve completely neutrally because deleting the gene in mice leads to infertility (Ren et al. 2001). Also, despite many indels present in the first exon of *Catsper1*, the open reading frame remains undisrupted. It is therefore more plausible that several sites are evolving under positive selection for amino acid replacements, while other sites are under functional constraints, rendering the overall d_N/d_S ratio approximately 1. To test this hypothesis, we performed a likelihood analysis using PAML (Yang 1997). Specifically, we compared models M7 and M8 (Yang et al. 2000). M7 is a null model, which assumes that the d_N/d_S ratio across codons follows a beta distribution in the range of 0–1. M8 is identical to M7 except for an additional class of codons with a free d_N/d_S ratio. Our likelihood ratio test showed that M8 fits our data significantly better than M7 ($\chi^2 = 46.77$, df = 2, $P < 0.001$), with an additional class of sites having $d_N/d_S = 5.78$. It has been suggested that a better way to test for positive selection is to compare the alternative hypothesis of M8 with the null hypothesis of M8a (Swanson, Nielsen, and Yang 2003). M8a is a special case of M8, where instead of estimating one additional class of sites with any d_N/d_S ratio, M8a fixes this d_N/d_S ratio to 1. The use of M8a seems to reduce false positives (Suzuki and Nei 2001; Zhang 2004) caused by codon sites under weak or no selection because these get absorbed in the neutral class of sites with $d_N/d_S = 1$ (Wong et al. 2004). In our analysis, M8 fit the data significantly better than M8a ($\chi^2 = 45.03$, df = 1, $P < 0.001$), suggesting that positive selection for nonsynonymous substitutions may indeed be acting on the first exon of rodent *Catsper1*, a result consistent with that found in primates.

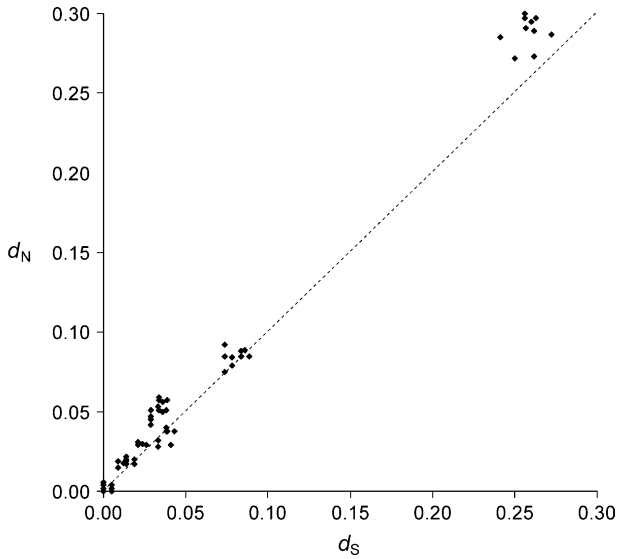

FIG. 4.—Number of synonymous (d_S) and nonsynonymous (d_N) substitutions per site between rodent *Catsper1* exon 1 sequences. The 13 nonredundant nucleotide sequences were used in this pairwise comparison. Dotted line represents $d_N/d_S = 1$.

Discussion

Our analysis of exon 1 of *Catsper1* from 22 individuals of 9 rodent species revealed strong evidence for positive selection that promotes fixation of indel mutations. These substitutions resulted in different lengths in the N-terminus of Catsper1 among different species. Within species, the level of indel polymorphism also appears higher than expected from interspecific comparisons, although our data are too limited to provide a statistically

significant result. It is interesting to note that some genes that have been shown to evolve rapidly between species also show exceptionally high levels of intraspecific polymorphism, possibly due to balancing selection (e.g., Drosophila accessory gland protein Acp26Aa; Tsaur, Ting, and Wu 2001).

It is known that insertion and deletion mutations in DNA sequences are often due to slippage during DNA replication (Hancook 1999). This type of mutation is particularly common in regions containing simple nucleotide repeats. With the exception of a short C_5A repeat in the 5′ end, *Catsper1* exon 1 does not have apparent sequence repeats. However, it should be noted that *Catsper1* exon 1 has a strong amino acid compositional bias, with histidine and serine having a total frequency of 32%. With the exception of the first two alignment gaps that are within C_5A repeats, the majority of the indels observed are not located near repeats, suggesting that the unusually high abundance of indels in *Catsper1* exon 1 is not caused by repeat expansion/contraction.

What selective agents are behind the extraordinarily rapid pace of indel substitutions in the sperm ion channel encoded by *Catsper1*? Genes belonging to the *Catsper* family encode putative voltage-gated Ca^{2+} (Ca_V) ion channels. A recent study showed that Catsper1 is necessary for depolarization-evoked Ca^{2+} entry and for hyperactive sperm movement (Carlson et al. 2003). The onset of hyperactive movement occurs near the site of fertilization and is crucial for penetrating the egg's zona pellucida. Although direct evidence showing that Catsper1 forms a voltage-gated ion channel is still lacking, the requirement of Catsper1 for depolarization-evoked entry of Ca^{2+} strongly supports such a hypothesis (Carlson et al. 2003). Several other members of the Catsper family have been identified, and all share a common expression in sperm (Quill et al. 2001; Lobley et al. 2003). It was thus conjectured that Catsper1, 2, 3, and 4 form a functional heterotetramer in sperm; however, empirical evidence for this hypothesis is lacking (Lobley et al. 2003). The putative ion channel formed by Catsper1 resembles K_V channels more than Ca_V channels in structure, as Catsper1 and K_V channels are formed by four peptides rather than a single long peptide with four repeats. This structural similarity to K_V ion channels prompted us to hypothesize a regulatory mechanism for Catsper1 similar to the ball-and-chain model (Bezanilla and Armstrong 1977), in which strong positive selection could occur for indel substitutions in the N-terminus of Catsper1 (Podlaha and Zhang 2003). Specifically, it was found that the action of the K_V channel Shaker in *Drosophila* follows the ball-and-chain model, in which the length of the N-terminus directly affects the rate of the channel's inactivation (Hoshi, Zagotta, and Aldrich 1990). If a similar regulatory mechanism indeed operates in Catsper1, different N-terminus lengths could have a direct functional effect on the channel's performance and consequently on sperm motility (Podlaha and Zhang 2003).

The physiological function and molecular evolutionary pattern of *Catsper1* suggest that it is a candidate for direct involvement in sperm competition. Further understanding of the evolutionary forces acting on this gene requires that questions about Catsper1 expression and fitness effects be addressed. Because *Catsper1* is expressed in sperm, could it possibly be under haploid selection? For a gene to be haploid selected, it is necessary to show that each sperm contains the product of only one allele of the gene and that the two alleles confer different fitnesses (Joseph and Kirkpatrick 2004). Rare examples of haploid selection have been documented in sperm adhesion molecule 1 *Spam1* (Zheng, Deng, and Martin-DeLeon 2001), segregation distortion responder *Rsp* (Kusano et al. 2003), and the *t*-locus responder *Tcr* (Herrmann et al. 1999). Although no direct demonstration of haploid selection in *Catsper1* exists, two lines of evidence lead us to believe that this is the case. First, *Catsper1* (as well as *Catsper2*, *Catsper3*, *Catsper4*) is solely expressed in testis and no other reproductive tissue (Genomics Institute of the Novaris Research Foundation SymAtlas; http://symatlas.gnf.org/SymAtlas/). Second, the timing of *Catsper1* expression is confined to the postmeiotic stage of spermatogenesis (Schultz, Hamra, and Garbers 2003), strongly suggesting that only one allele is expressed in each sperm. Haploid selection can result in meiotic drive, which refers to the phenomenon of nonrandom segregation in heterozygotes (Hartl and Clark 1997). Such haploid selection and meiotic drive could generate an "arms-race" between alleles and result in high levels of intraspecific polymorphism and interspecific divergence. Haploid selection and meiotic drive can be tested by genotyping the offspring of heterozygous males and examining the competitive advantage of a particular length variant of Catsper1 over another variant. Such experiments are feasible in mice, where it is faster and easier to obtain individuals of a particular genotype and acquire data on reproductive success than in primates. *Catsper1* is a promising candidate gene for sperm competition as sperm motility has been shown to be one of its decisive factors (Birkhead et al. 1999). The present study, as well as our previous analysis in primates, reveals strong positive selection on the length of Catsper1. Furthermore, it provides a necessary foundation for future experimental investigations of Catsper1's function in sperm physiology and role in sperm competition using rodent models.

Supplementary Material

GenBank accession numbers of the DNA sequences reported in this paper are DQ021482–DQ021500.

Acknowledgments

We thank A. Brege for technical assistance, M. Potter (National Cancer Institute Contract N01-CB-71085) and E. Eicher (The Jackson Laboratory) for providing frozen tissue/DNA from various stocks and strains of wild mice, and R. D. Sage for providing frozen tissue/DNA from wild mice collected in the field. This work was supported in part by the National Institutes of Health grant GM67030 to J.Z.

Literature Cited

Bezanilla, F., and C. M. Armstrong. 1977. Inactivation of the sodium channel. I. Sodium current experiments. J. Gen. Physiol. **70**:549–566.

Birkhead, T. R., J. G. Martinez, T. Burke, and D. P. Froman. 1999. Sperm mobility determines the outcome of sperm competition in the domestic fowl. Proc. R. Soc. Lond. B Biol. Sci. **266**: 1759–1764.

Britten, R. J. 2002. Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. Proc. Natl. Acad. Sci. USA **99**:13633–13635.

Carlson, A. E., R. E. Westenbroek, T. Quill, D. J. Ren, D. E. Clapham, B. Hille, D. L. Garbers, and D. F. Babcock. 2003. CatSper1 required for evoked Ca2+ entry and control of flagellar function in sperm. Proc. Natl. Acad. Sci. USA **100**:14864–14868.

Cooper, G. M., M. Brudno, E. A. Stone, I. Dubchak, S. Batzoglou, and A. Sidow. 2004. Characterization of evolutionary rates and constraints in three mammalian genomes. Genome Res. **14**: 539–548.

Hancook, J. M. 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. Pp. 1–9 *in* D. B. Goldstein and C. Schlotterer, eds. Microsatellites, evolution, and applications. Oxford University Press, New York.

Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics. 3rd edition. Sinauer, Sunderland, Mass.

Herrmann, B. G., B. Koschorz, K. Wertz, K. J. McLaughlin, and A. Kispert. 1999. A protein kinase encoded by the t complex responder gene causes non-Mendelian inheritance. Nature **402**:141–146.

Hoshi, T., W. N. Zagotta, and R. W. Aldrich. 1990. Biophysical and molecular mechanisms of Shaker potassium channel inactivation. Science **250**:533–538.

Hughes, A. L., and M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335**:167–170.

Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. Nature **409**: 860–921.

Joseph, S. B., and M. Kirkpatrick. 2004. Haploid selection in animals. Trends Ecol. Evol. **19**:592–597.

Kusano, A., C. Staber, H. Y. Chan, and B. Ganetzky. 2003. Closing the (Ran)GAP on segregation distortion in Drosophila. Bioessays **25**:108–115.

Lee, Y. H., T. Ota, and V. D. Vacquier. 1995. Positive selection is a general phenomenon in the evolution of abalone sperm lysin. Mol. Biol. Evol. **12**:231–238.

Lobley, A., V. Pierron, L. Reynolds, L. Allen, and D. Michalovich. 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.

Makova, K. D., S. Yang, and F. Chiaromonte. 2004. Insertions and deletions are male biased too: a whole-genome analysis in rodents. Genome Res. **14**:567–573.

Mcdonald, J. H., and M. Kreitman. 1991. Adaptive protein evolution at the Adh locus in Drosophila. Nature **351**: 652–654.

Podlaha, O., and J. Zhang. 2003. Positive selection on protein-length in the evolution of a primate sperm ion channel. Proc. Natl. Acad. Sci. USA **100**:12241–12246.

Quill, T. A., D. J. Ren, D. E. Clapham, and D. L. Garbers. 2001. A voltage-gated ion channel expressed specifically in spermatozoa. Proc. Natl. Acad. Sci. USA **98**:12527–12531.

Rat Genome Sequencing Consortium. 2004. Genome sequence of the brown Norway rat yields insights into mammalian evolution. Nature **428**:493–521.

Ren, D. J., B. Navarro, G. Perez, A. C. Jackson, S. F. Hsu, Q. Shi, J. L. Tilly, and D. E. Clapham. 2001. A sperm ion channel required for sperm motility and male fertility. Nature **413**:603–609.

Rooney, A. P., and J. Zhang. 1999. Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection? Mol. Biol. Evol. **16**:706–710.

Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

Schultz, N., F. K. Hamra, and D. L. Garbers. 2003. A multitude of genes expressed solely in meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive targets. Proc. Natl. Acad. Sci. USA **100**:12201–12206.

Silva, J. C., and A. S. Kondrashov. 2002. Patterns in spontaneous mutation revealed by human-baboon sequence comparison. Trends Genet. **18**:544–547.

Suzuki, Y., and M. Nei. 2001. Reliabilities of parsimony-based and likelihood-based methods for detecting positive selection at single amino acid sites. Mol. Biol. Evol. **18**:2179–2185.

Swanson, W. J., R. Nielsen, and Q. Yang. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol. Biol. Evol. **20**:18–20.

Swanson, W. J., and V. D. Vacquier. 2002. The rapid evolution of reproductive proteins. Nat. Rev. Genet. **3**:137–144.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.

Tsaur, S. C., C. T. Ting, and C. I. Wu. 2001. Sex in Drosophila mauritiana: a very high level of amino acid polymorphism in a male reproductive protein gene, Acp26Aa. Mol. Biol. Evol. **18**:22–26.

Tucker, P. K., S. A. Sandstedt, and B. L. Lundrigan. 2005. Phylogenetic relationships in the subgenus Mus (genus Mus, family Muridae, subfamily Murinae): examining gene trees and species trees. Biol. J. Linn. Soc. Lond. **84**:653–662.

Wang, X., and J. Zhang. 2004. Rapid evolution of mammalian X-linked testis-expressed homeobox genes. Genetics **167**: 879–888.

Wolfe, K. H., and W. H. Li. 2003. Molecular evolution meets the genomics revolution. Nat. Genet. **33**(Suppl.):255–265.

Wong, W. S., Z. Yang, N. Goldman, and R. Nielsen. 2004. Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. Genetics **168**:1041–1051.

Wyckoff, G. J., W. Wang, and C. I. Wu. 2000. Rapid evolution of male reproductive genes in the descent of man. Nature **403**:304–309.

Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13**:555–556.

Yang, Z., R. Nielsen, N. Goldman, and A. M. Pedersen. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155**:431–449.

Zagotta, W. N., T. Hoshi, and R. W. Aldrich. 1990. Restoration of inactivation in mutants of Shaker potassium channels by a peptide derived from ShB. Science **250**:568–571.

Zhang, J. 2004. Frequent false detection of positive selection by the likelihood method with branch-site models. Mol. Biol. Evol. **21**:1332–1339.

Zhang, J., and H. F. Rosenberg. 2002. Diversifying selection of the tumor-growth promoter angiogenin in primate evolution. Mol. Biol. Evol. **19**:438–445.

Zhang, J., H. F. Rosenberg, and M. Nei. 1998. Positive Darwinian selection after gene duplication in primate

ribonuclease genes. Proc. Natl. Acad. Sci. USA **95**: 3708–3713.

Zhang, J., and D. M. Webb. 2004. Rapid evolution of primate antiviral enzyme APOBEC3G. Hum. Mol. Genet. **13**: 1785–1791.

Zhang, J., Y. P. Zhang, and H. F. Rosenberg. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat. Genet. **30**:411–415.

Zheng, Y., X. Deng, and P. A. Martin-DeLeon. 2001. Lack of sharing of Spam1 (Ph-20) among mouse spermatids and transmission ratio distortion. Biol. Reprod. **64**:1730–1738.

Pekka Pamilo, Associate Editor

Accepted May 23, 2005