

# Enhanced adaptive evolution of sperm-expressed genes on the mammalian X chromosome

DG Torgerson and RS Singh

Department of Biology, McMaster University, Hamilton, ON, Canada L8S 1J4

Genes on the mammalian X chromosome may be under unique evolutionary pressure due to their hemizygous expression in males. Since any recessive deleterious mutation would immediately be expressed in males and, therefore, efficiently removed from the population, selective constraint could be more pronounced in X-linked genes. Conversely, if a recessive mutation were beneficial, its immediate exposure to selection would be advantageous, and would facilitate adaptive evolution. We tested for positive selection in a total of 86 genes using a maximum likelihood approach, including 40 sperm-expressed and 46 non-sperm,

tissue-specific genes. We find evidence to suggest that X-linkage enhances the effects of positive selection in sperm-expressed genes in terms of the number of codons affected, and report a general trend for positively selected genes to reside on the X chromosome rather than on the autosomes. Our data suggest that hemizygous expression in males makes the X chromosome a preferred location for positively selected sperm genes that do not require postmeiotic transcription.

*Heredity* (2006) **96**, 39–44. doi:10.1038/sj.hdy.6800749; published online 21 September 2005

**Keywords:** positive selection; sperm genes; tissue-specific genes; X chromosome

## Introduction

In evolutionary biology, the X chromosome is of particular interest due to its unique role in sex determination and sexual dimorphism (Pask and Graves, 1999), segregation distortion (De La Casa-Esperon *et al*, 2002), hybrid sterility, and speciation (Zeng and Singh, 1993; Turelli and Begun, 1997). The evolutionary rate of the X chromosome may be notably different from the autosomes, with evidence for faster, slower, and similar nonsynonymous evolutionary rates (Wolfe and Sharp, 1993; McVean and Hurst, 1997; Betancourt *et al*, 2002; Torgerson and Singh, 2003; Counterman *et al*, 2004). It seems reasonable to assume that at least certain categories of genes on the mammalian X chromosome could evolve in a distinct manner because of its hemizygous expression in males.

Charlesworth *et al* (1987) proposed that if a new beneficial mutation arises that is at least partially recessive, haploid expression will facilitate its spread in a population by an immediate exposure to selection. On the other hand, haploid expression will also facilitate the immediate removal of recessive deleterious mutations, and so the evolutionary rates of X-linked genes may be more dependent on the nature of mutation and selection acting on individual genes. It is therefore likely that the evolutionary rate of the X chromosome, relative to the autosomes, is not simply faster or slower overall, but differs in a more complex manner.

There are conflicting reports on whether nonsynonymous mutation rates of X-linked genes differ from those

on the autosomes. Two studies in mammals found evidence of a lower nonsynonymous mutation rate ( $d_N$ ) in several X-linked genes compared to autosomal genes (Wolfe and Sharp, 1993; McVean and Hurst, 1997), whereas we previously found no significant difference in  $d_N$  between X-linked and autosomal genes in a study including 408 tissue-specific genes (Torgerson and Singh, 2003). A more recent genomic survey found that nonsynonymous substitutions are 30% more frequent on the X chromosome than on the autosomes in a comparison between human and chimpanzee (Lu and Wu, 2005). In *Drosophila*, one study failed to find any differences between X and autosomal nonsynonymous mutation rates (Betancourt *et al*, 2002), whereas a subsequent study reported evidence in support of faster X evolution (Counterman *et al*, 2004). Such studies on the evolutionary rates of X-linked genes may be confounded, as results may be largely dependent on the number and types of genes studied.

Genes that stand to gain an advantage from haploid expression should be those that experience a higher rate of beneficial mutations relative to deleterious, or genes that are under pressure to evolve rapidly. Genes under positive Darwinian selection are the most likely to have an advantage to being on the X chromosome, as selection can act immediately to retain adaptive recessive mutations in a population, and therefore counteract the effects of random genetic drift. There are several examples of genes involved in sex and reproduction that are subject to positive selection (Swanson *et al*, 2001, 2003; Torgerson *et al*, 2002), which may be driven by sexual selection, including but not limited to sexual coevolution, sexual conflict, and sperm competition (Singh and Kulathinal, 2000; Swanson and Vacquier, 2002). In mammals, positive selection has been reported in female reproductive proteins (Swanson *et al*, 2001), in genes responsible for sperm structure and function (Torgerson *et al*, 2002),

Correspondence: DG Torgerson. Current Address: Department of Molecular Biology and Genetics, Cornell University, 227 Biotechnology Bldg., Ithaca, NY 14850, USA. E-mail: dt224@cornell.edu

Received 25 August 2004; accepted 25 July 2005; published online 21 September 2005

and in genes involved in fertilization in general (Swanson *et al*, 2003).

If positive selection acts frequently on sex and reproduction-related genes, and if hemizygous expression facilitates positive selection, then X-linked genes expressed in sperm cells should show a larger effect of positive selection than those on the autosomes. Furthermore, if hemizygous expression facilitates positive selection, thus giving an advantage to X-linkage, then we would expect to find a higher proportion of positively selected genes on the X chromosome. In this study, we present results from testing the hypothesis that positive selection acts more strongly on sperm genes on the X chromosome, and that positively selected genes are preferentially found on the X chromosome.

## Materials and methods

Sperm genes were defined as those expressed in sperm cells at some time throughout spermatogenesis from the spermatogonia up to the mature sperm; this information was retrieved from the primary literature. Tissue-specific genes were defined as genes limited in their expression to a single normal human tissue, and were retrieved from UniGene at NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) by searching for genes represented in only 1–3 EST databases. Tissue-specific genes were then filtered to remove any male- or female-specific genes (ie showing expression in the testes, prostate, ovary, uterus, etc.), and to remove any genes showing expression in more than one normal tissue. Chromosomal locations were also retrieved from UniGene.

From our list of sperm and non-sperm tissue-specific genes, homologous human, mouse, and rat sequences were retrieved from HomoloGene at NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)), by using a reciprocal best hits criterion to determine orthology. Genes that did not match this criterion, or that did not have orthologs in all three species were discarded. DNA sequences were aligned using RevTrans version 1.0 (Wernersson and Pedersen, 2003), which aligns translated amino-acid sequences, and then untranslates amino-acid sequences back into the original nucleotide sequences while maintaining the protein alignment. By using this method of alignment, homologous codons can be more accurately aligned for subsequent analyses, and any frame-shifts are easily avoided. Alignments were then visually inspected to check for errors.

Tests of positive selection for each gene were performed using a maximum likelihood approach as implemented in the program codeml in version 3.14b of PAML (Yang, 1997). This program utilizes a model of evolutionary change of codons, and calculates the likelihoods of specified models. Twice the difference in likelihoods of two nested models can be compared to a  $\chi^2$  distribution to test for significance, with the number of degrees of freedom equal to the difference in the number of estimated free parameters. We utilized a conservative test for positive selection by comparing two models that assume a  $\beta$  distribution on the parameter  $\omega$  ( $d_N/d_S$ ): Model 7 (M7), where  $\omega$  is constrained to be less than one, and Model 8 (M8) that contains an additional class of sites for which  $\omega$  is allowed to exceed one. Positively

selected codons were identified using empirical Bayes Factors (Yang *et al*, 2005).

We used a Fisher Exact test to compare (1) the proportions of positively selected genes in sperm and non-sperm-expressed genes, (2) the proportions of affected codons in positively selected X-linked and autosomal genes, and (3) the proportions of positively selected genes on the X chromosome compared to autosomes for both sperm and non-sperm tissue-specific genes.

## Results

Our final data set consisted of 40 sperm-expressed and 46 tissue-specific genes making 86 genes in total. Using a maximum likelihood framework, we found evidence for positive selection in 20 of these genes at a 99% confidence level (Table 1). A complete list of all 86 genes along with their likelihood ratio tests and parameter estimates are available as online supplementary material. It is important to note that no pair of X-linked sperm genes included in this study is known to belong to the same gene family and, consequently, the sample of X-linked genes has not been biased by the selection of multiple members of the same family. The majority of genes identified as positively selected were sperm-expressed ( $n=16$ ); however, four other genes with tissue-specific expression – in the pancreas (GPR119), retina (GPR112), kidney (NYX), and eye (TEXTB) – were also identified to be under positive selection. The proportion of positively selected sperm genes is significantly higher than that found in the tissue-specific, or non-sperm genes (40 *vs* 9.5%; Fisher Exact test,  $P=0.0006$ ), consistent with the frequent findings of positive selection acting on sex and reproduction-related genes (Swanson *et al*, 2001, 2003; Torgerson *et al*, 2002). However, considering only X-linked genes, the proportion of positively selected sperm- and tissue-specific genes is not significantly different ( $P=0.12$ ; Table 2).

When positively selected genes are examined at the level of the codon, we find a significantly higher proportion of codons to be affected within sperm genes on the X chromosome relative to the autosomes ( $P<0.0001$ ; Table 3). Although our sample size is limited in the number of tissue-specific genes, we also find that within positively selected X-linked genes, the proportion of affected codons is significantly higher in sperm genes compared to non-sperm, tissue-specific genes that do not have male limited expression ( $P<0.0001$ ; Table 4).

Out of 40 sperm-expressed genes, we find the proportion of positively selected genes on the X chromosome to be higher when compared to the autosomes (62.5 *vs* 34.4%); however, this trend is not significantly different ( $P=0.11$ ; Table 2). Combining these sperm genes with tissue-specific genes we again find a higher proportion of X-linked genes to be under positive selection (42.1 *vs* 17.9%;  $P=0.025$ ); however, this relationship is now significant. Although we only detected positive selection in four out of 46 non-sperm, tissue-specific genes, three of these genes are found on the X chromosome whereas one is on the autosomes.

**Table 1** Genes predicted to be under positive selection using a maximum likelihood framework in PAML (Yang, 1997)

Gene	Tissue	Map	$L_c^a$	$2\Delta l$ M7 vs M8	P-value	Parameter estimates <sup>a</sup> under M8: $p_1(\omega > 1)$ , $\omega$
<i>Cyl1</i>	Sperm	X	277	67.1	<0.0001	$p_1 = 0.15$ , $\omega = 37.8$
<i>Nxf</i>	Sperm	X	552	23.7	<0.0001	$p_1 = 0.05$ , $\omega = 12.6$
<i>Taf2q</i>	Sperm	X	349	33.7	<0.0001	$p_1 = 0.09$ , $\omega = 28.2$
<i>Tex11</i>	Sperm	X	507	49.5	<0.0001	$p_1 = 0.14$ , $\omega = 10.1$
<i>Usp26</i>	Sperm	X	817	32.3	<0.0001	$p_1 = 0.03$ , $\omega = \infty$
<i>Acrv1</i>	Sperm	11	239	13.2	0.0014	$p_1 = 0.03$ , $\omega = 232.9$
<i>Adam18</i>	Sperm	8	580	25.5	<0.0001	$p_1 = 0.04$ , $\omega = \infty$
<i>Adam2</i>	Sperm	8	728	22.6	<0.0001	$p_1 = 0.13$ , $\omega = 2.8$
<i>Izumo</i>	Sperm	19	336	14.8	0.0006	$p_1 = 0.14$ , $\omega = 4.1$
<i>Ldhc</i>	Sperm	11	332	11.3	0.0036	$p_1 = 0.02$ , $\omega = 369.5$
<i>Odf1</i>	Sperm	8	89	15.2	0.0005	$p_1 = 0.04$ , $\omega = 57.6$
<i>Odf4</i>	Sperm	17	216	14.1	0.0008	$p_1 = 0.17$ , $\omega = 5.3$
<i>Prm1</i>	Sperm	16	49	15.1	0.0005	$p_1 = 0.03$ , $\omega = 844.0$
<i>Spam1</i>	Sperm	7	510	15.7	0.0004	$p_1 = 0.03$ , $\omega = 73.6$
<i>Tex14</i>	Sperm	17	889	32.7	<0.0001	$p_1 = 0.03$ , $\omega = \infty$
<i>Tex15</i>	Sperm	8	2628	31.8	<0.0001	$p_1 = 0.08$ , $\omega = 3.2$
<i>Gpr112</i>	Retina	X	2280	66.5	<0.0001	$p_1 = 0.02$ , $\omega = 160.1$
<i>Gpr119</i>	Pancreas	X	335	14.8	0.0006	$p_1 = 0.04$ , $\omega = \infty$
<i>Nyx</i>	Kidney	X	476	15.0	0.0006	$p_1 = 0.03$ , $\omega = 6.3$
<i>Tectb</i>	Eye	10	308	12.9	0.0016	$p_1 = 0.01$ , $\omega = 126.3$

The number of taxa was limited to three for every gene, and included the orthologous human, mouse, and rat sequences as determined by reciprocal best hits.

<sup>a</sup> $L_c$  = total number of homologous codons after indels removed,  $p_1$  = proportion of codons belonging to a site class of  $\omega > 1$ ,  $\omega = d_N/d_S$ .

**Table 2** The proportions of X-linked vs autosomal sperm and non-sperm genes under positive selection

Tissue	X-linked	Autosomal	Fisher exact
Sperm	62.5% (5/8)	34.4% (11/32)	$P = 0.11$
Non-sperm	27.3% (3/11)	2.9% (1/35)	$P = 0.035$
Fisher exact	$P = 0.12$	$P = 0.00075$	

**Table 3** The proportions of positively selected codons in sperm genes that are under positive selection on the X chromosome vs autosomes

Tissue	Location	Number of codons under positive selection	Number of codons not under positive selection	Proportion	Fisher exact test
Sperm	X-linked	221	2502	8.8%	$P < 0.0001$
	Autosome	327	6596	5.0%	

**Table 4** The proportions of positively selected codons in X-linked sperm and non-sperm (tissue-specific) genes under positive selection

Tissue	Number of codons under positive selection	Number of codons not under positive selection	Proportion	Fisher exact test
Sperm	221	2502	8.8%	$P < 0.0001$
Non-sperm	84	3091	2.7%	

## Discussion

### X-linkage enhances positive selection in sperm genes

If there were an evolutionary advantage for positively selected genes to be X-linked, we would expect to find evidence that positive selection is somehow enhanced in X-linked genes. We tested two predictions of this theory. First, if positive selection were acting more strongly on X-linked genes, we might expect to find a difference in the proportion of affected codons within positively selected X-linked and autosomal sperm genes. Consistent with this theory, we found a significantly higher proportion of affected codons in positively selected sperm genes on the X chromosome (compared to those on the autosomes), suggesting that positive selection is enhanced in X-linked genes.

A second prediction of this theory is that genes with male-limited expression should be more strongly affected by positive selection than genes that are expressed in both male and female subjects. Genes on the X chromosome are hemizygotously expressed in male but not in female subjects, so recessive beneficial mutations in genes expressed in both genders are only exposed to selection in female subjects once they reach sufficient frequency in the population to appear in homozygotes. Our data are also consistent with this prediction, as the proportion of affected codons is significantly higher in positively selected X-linked sperm genes than in positively selected X-linked genes without male limited expression. Our data, therefore, supports the theory that X-linkage, or hemizygotous expression can enhance the effects of positive selection in terms of the number of codons affected.

### Evidence for a selective advantage of X-linkage in premeiotic sperm genes

Although we found a higher proportion of positively selected sperm genes on the X chromosome, this relationship was not significantly different from the proportion of those found on the autosomes. Although our data suggest we reject the hypothesis that X-linked sperm genes are more likely to be under positive selection, other factors may affect the location of positively selected sperm genes. Genes expressed later in spermatogenesis generally evolve faster, and are more likely to be under positive selection than those expressed earlier in spermatogenesis, possibly due to stronger sexual selection acting on genes influencing sperm structure and function (Good and Nachman, 2005). However, due to the earlier inactivation of the X chromosome compared to autosomes in the mammalian germline (Lifschytz and Lindsley, 1972), it is impossible for postmeiotic transcribed sperm genes to be located on the X chromosome (Wu and Xu, 2003). Therefore, even if hemizygous expression confers an advantage for positively selected sperm genes, those expressed during late spermatogenesis are restricted to the autosomes. Owing to their higher abundance, these genes may have influenced our test of the hypothesis that X-linked sperm genes are more likely to be under positive selection.

When we examine the literature on the transcription of positively selected sperm genes, we found, as expected, that all five X-linked genes show premeiotic transcription (Wang *et al.*, 2001), which is prior to germline X-inactivation. If there were no advantage for positively selected sperm genes to be X-linked, we would expect to identify some positively selected premeiotic transcribed sperm genes on the autosomes (along with all postmeiotic transcribed genes). However, the majority of positively selected autosomal sperm genes have at least some postmeiotic (or post germline X-inactivation) transcription. Out of 11 positively selected autosomal sperm genes, nine show postmeiotic transcription, and the expression of two is currently unknown (see online supplementary material for references). Genes that require postmeiotic transcription are restricted to an autosomal location, as the X chromosome has become transcriptionally inactivated at this point in spermatogenesis. Overall, it appears as though the majority (and potentially all) of positively selected sperm genes that could be on the X chromosome actually are. Our data therefore suggest that X-linkage could pose as a selective advantage for positively selected sperm genes that are expressed prior to germline X-inactivation.

### Evidence for a selective advantage of X-linkage in non-male-biased genes

Owing to the tendency for tissue-specific genes to evolve faster than more broadly expressed genes (Coulthart and Singh, 1988; Duret and Mouchiroud, 2000; Zhang and Li, 2004), we considered the possibility that positive selection is occurring more frequently in this class of genes. We selected a group of tissue-specific genes and excluded those having male or female limited expression, therefore removing genes limited to expression in the male germline where early X-inactivation is a factor. Although the number of positively selected tissue-specific genes is only four, almost a third of the X-linked

tissue-specific genes show signatures of positive selection. More convincing data comes from a recent genome-wide survey that has also found there to be an overall bias for positively selected genes to be X-linked, even after controlling for male-expressed genes (Nielsen *et al.*, 2005). This pattern suggests that hemizygous expression may confer a selective advantage, even in genes that are not limited to male-specific expression.

An unexpected finding was that the proportions of positively selected X-linked sperm genes are not significantly different from the proportion of tissue-specific genes, despite our previous findings that X-linked sperm proteins have both a higher divergence and lower selective constraints (measured by  $K_a/K_s$ ; Torgerson and Singh, 2003). Our initial data are consistent with our current findings though, because if hemizygous expression is resulting in a higher incidence of positive selection on X-linked genes, it may also enhance negative selection by increasing the rate at which recessive deleterious mutations are removed. The overall effect would be to lower the average divergence and  $K_a/K_s$  values for X-linked genes that evolve more slowly (non-sperm) than those that evolve more rapidly (sperm) despite the inclusion of a similar proportion of positively selected genes. Our findings therefore highlight the difficulties of examining the evolutionary rate of the X chromosome using larger-scale data sets with no *a priori* hypothesis about the effects of X-linkage on different classes of genes.

### Overall bias for positively selected genes to be on the X chromosome

When all genes are combined without regard to expression patterns, we find a higher proportion of X-linked genes to be under positive selection compared to autosomal genes, suggesting there may be a general bias for positively selected genes to be on the X chromosome. These findings are again consistent with a genome-wide survey using ratios of  $K_a/K_s$  between human and chimp to infer positive selection (Nielsen *et al.*, 2005). Genes on the X chromosome have a relatively high turnover rate compared to the autosomes, with many genes moving on and off through retroposition events (Emerson *et al.*, 2004). Given this higher turnover rate, it possible that genes under selection to evolve rapidly would obtain an advantage by remaining on the X chromosome, whereas those that require conservation may benefit from moving to an autosomal location where recessive deleterious mutations can hide.

Previous studies indicate that the mammalian X chromosome has an over-representation of premeiotic sperm-expressed genes (Wang *et al.*, 2001; Khil *et al.*, 2004), prostate expressed genes (Lercher *et al.*, 2003), genes involved in cognitive ability (Zechner *et al.*, 2001), brain expressed (Qiu *et al.*, 2002), skeletal muscle expressed (Bortoluzzi *et al.*, 1998), and possibly even tissue-specific expressed genes in general (Lercher *et al.*, 2003). Our data suggest that the X chromosome harbors a higher proportion of positively selected genes, which may offer a more general explanation for the unusually biased gene content of the mammalian X chromosome. Positive selection may be important in the evolution of a wide range of proteins, and is unlikely to be restricted to genes involved in male reproduction.

We have shown that sperm genes have a relatively high incidence of positive selection in general, which may be the case for other categories of genes that are over-represented on the X chromosome. For example, there are several genes expressed in the brain that may be involved in cognitive ability or speech that show evidence of positive selection in human evolution (Zhang, 2003; Evans *et al*, 2004). Moreover, gene duplications may often become tissue specific in their expression patterns through the subfunctionalization of a multifunctional ancestor (Force *et al*, 1999), and may have retained a new adaptive function to avoid being lost (Ohno, 1970). Positive selection has, in fact, been shown to occur frequently following gene duplication (eg's, see Moore and Purugganan, 2003; Rodriguez-Trelles *et al*, 2003), so if the X chromosome enhances the effects of positive selection it may be promoting the retention of duplicated tissue-specific genes. Therefore, a bias in the number of tissue-specific genes, as well as brain expressed, and premeiotic sperm-expressed genes on the X chromosome could be a byproduct of a bias in the number of positively selected genes on the X chromosome.

However, the process of adaptive evolution involves a number of variables other than hemizygous expression, and therefore not all positively selected genes should be expected to reside on the X or even the Y chromosome. A combination of many other factors could affect a gene's location, for example, genes that have similar function or coordinated expression tend to cluster together (Hurst *et al*, 2004), genes with similar GC content tend to form chromosomal isochors (Matassi *et al*, 1999), and members of gene families are often found on the same chromosome (Friedman and Hughes, 2004). Adaptive dominant mutations would not have a selective advantage or disadvantage to being hemizygously expressed, and for genes under diversifying selection or with a heterozygous advantage, an autosomal location may be preferred, which may be the case for genes involved in the immune response such as the major histocompatibility complex (Penn *et al*, 2002).

## Conclusion

Hemizygous expression of the mammalian X chromosome in males can theoretically enhance the effects of positive selection, making the X chromosome a preferred location for positively selected genes. We found evidence to suggest that in sperm-expressed genes, X-linkage can enhance the effects of positive selection in terms of the number of codons affected, and that positively selected genes may obtain an advantage from being on the X chromosome unless postmeiotic transcription is required. Selection acting on hemizygously expressed genes is therefore an important factor when discussing the unique nature of the evolution of X-linked genes, and may offer a more general explanation for the unusual gene content of the mammalian X chromosome.

## Acknowledgements

We thank Dr JP Xu, the Associate Editor, and two anonymous reviewers for their many helpful suggestions. This project was funded through the Natural Sciences and Engineering Research Council of Canada by

a grant to RSS, and by an Ontario Graduate Scholarship to DGT.

## References

- Betancourt AJ, Presgraves DC, Swanson WJ (2002). A test for faster X evolution in *Drosophila*. *Mol Biol Evol* **19**: 1816–1819.
- Bortoluzzi S, Rampoldi L, Simionati B, Zimbello R, Barbon A, d'Alessi F *et al* (1998). A comprehensive, high-resolution genomic transcript map of human skeletal muscle. *Genome Res* **8**: 817–825.
- Charlesworth B, Coyne JA, Barton N (1987). The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* **140**: 126–148.
- Coulthart MB, Singh RS (1988). High level of divergence of male-reproductive-tract proteins, between *Drosophila melanogaster* and its sibling species, *D. simulans*. *Mol Biol Evol* **5**: 182–191.
- Counterman BA, Ortíz-Barrientos D, Noor MAF (2004). Using comparative genomic data to test for fast-X evolution. *Evolution* **58**: 656–660.
- De La Casa-Esperon E, Loredó-Osti JC, Pardo-Manuel de Villena F, Briscoe TL, Malette JM, Vaughan JE *et al* (2002). X chromosome effect on maternal recombination and meiotic drive in the mouse. *Genetics* **161**: 1651–1659.
- Duret L, Mouchiroud D (2000). Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol Biol Evol* **17**: 68–74.
- Emerson JJ, Kaessmann H, Betrán E, Long M (2004). Extensive gene traffic on the mammalian X chromosome. *Science* **303**: 537–540.
- Evans PD, Anderson JR, Vallender EJ, Choi SS, Lahn BT (2004). Reconstructing the evolutionary history of microcephalin, a gene controlling human brain size. *Hum Mol Genet* **13**: 1139–1145.
- Force A, Lynch M, Pickett FB, Amores A, Yan Y-L, Postlethwait J (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531–1545.
- Friedman R, Hughes AL (2004). Two patterns of genome organization in mammals: the chromosomal distribution of duplicate genes in human and mouse. *Mol Biol Evol* **21**: 1008–1013.
- Good JM, Nachman MW (2005). Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol Biol Evol* **22**: 1044–1052.
- Hurst LD, Pál C, Lercher MJ (2004). The evolutionary dynamics of eukaryotic gene order. *Nat Rev Genet* **5**: 299–310.
- Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD (2004). The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* **36**: 642–646.
- Lercher MJ, Urrutia AO, Hurst LD (2003). Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol Biol Evol* **20**: 1113–1116.
- Lifshytsz É, Lindsley DL (1972). The role of X-chromosome inactivation during spermatogenesis. *Proc Natl Acad Sci USA* **69**: 182–186.
- Lu J, Wu C-I (2005). Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. *Proc Natl Acad Sci USA* **102**: 4063–4067.
- Matassi G, Sharp PM, Gautier C (1999). Chromosomal location effects on gene sequence evolution in mammals. *Curr Biol* **9**: 786–791.
- McVean GT, Hurst LD (1997). Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. *Nature* **386**: 388–392.
- Moore RC, Purugganan MD (2003). The early stages of duplicate gene evolution. *Proc Natl Acad Sci USA* **100**: 15682–15687.

- Nielsen R, Clark AG, Bustamante C, Glanowski S, Hubisz MJ, Tanenbaum DM *et al* (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol* **3**: e170.
- Ohno S (1970). *Evolution by Gene Duplication*. Springer Verlag: New York.
- Pask A, Graves JA (1999). Sex chromosomes and sex-determining genes: insights from marsupials and monotremes. *Cell Mol Life Sci* **55**: 864–875.
- Penn DJ, Damjanovich K, Potts WK (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proc Natl Acad Sci USA* **99**: 11260–11264.
- Qiu P, Benbow L, Liu S, Greene JR, Wang L (2002). Analysis of a human brain transcriptome map. *BMC Genom* **3**: 10.
- Rodriguez-Trelles F, Tarrío R, Ayala FJ (2003). Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proc Natl Acad Sci USA* **100**: 13413–13417.
- Singh RS, Kulathinal RJ (2000). Sex gene pool evolution and speciation: a new paradigm. *Genes Genet Syst* **75**: 119–130.
- Swanson WJ, Nielsen R, Yang Q (2003). Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* **20**: 18–20.
- Swanson WJ, Vacquier VD (2002). The rapid evolution of reproductive proteins. *Nat Rev Genet* **3**: 137–144.
- Swanson WJ, Yang Z, Wolfner MF, Aquadro CF (2001). Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc Natl Acad Sci USA* **98**: 2509–2514.
- Torgerson DG, Kulathinal RJ, Singh RS (2002). Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. *Mol Biol Evol* **19**: 1973–1980.
- Torgerson DG, Singh RS (2003). Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Mol Biol Evol* **20**: 1705–1709.
- Turelli M, Begun DJ (1997). Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* **147**: 1799–1815.
- Wang PJ, McCarrey JR, Yang F, Page DC (2001). An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* **27**: 422–426.
- Wernersson R, Pedersen AG (2003). RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res* **31**: 3537–3539.
- Wolfe KH, Sharp PM (1993). Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. *J Mol Evol* **37**: 441–456.
- Wu C-I, Xu EY (2003). Sexual antagonism and X inactivation – the SAXI hypothesis. *Trends Genet* **19**: 243–247.
- Yang Z (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* **13**: 555–556.
- Yang Z, Wong WS, Nielsen R (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Mol Biol Evol* **22**: 1107–1118.
- Zechner U, Wilda M, Kehrer-Sawatzki H, Vogel W, Fundele R, Hameister H (2001). A high density of X-linked genes for general cognitive ability: a run-away process shaping human evolution? *Trends Genet* **17**: 697–701.
- Zeng LW, Singh RS (1993). The genetic basis of Haldane's rule and the nature of asymmetric hybrid male sterility among *Drosophila simulans*, *Drosophila mauritiana*, and *Drosophila sechellia*. *Genetics* **134**: 251–260.
- Zhang J (2003). Evolution of the human ASPM gene, a major determinant of brain size. *Genetics* **165**: 2063–2070.
- Zhang L, Li W-H (2004). Mammalian housekeeping genes evolve more slowly than tissue-specific genes. *Mol Biol Evol* **21**: 236–239.

Supplementary Information accompanies the paper on the Heredity website (<http://www.nature.com/hdy>)