Evolutionary rate of a gene affected by chromosomal position

Jo Perry and Alan Ashworth

Genes evolve at different rates depending on the strength of selective pressure to maintain their function. Chromosomal position can also have an influence [1,2]. The pseudoautosomal region (PAR) of mammalian sex chromosomes is a small region of sequence identity that is the site of an obligatory pairing and recombination event between the X and Y chromosomes during male meiosis [3–6]. During female meiosis, X chromosomes can pair and recombine along their entire length. Recombination in the PAR is therefore approximately 10 times greater in male meiosis compared with female meiosis [4–6]. The gene *Fxy* (also known as *MID1* [7]) spans the pseudoautosomal boundary (PAB) in the laboratory mouse (*Mus musculus domesticus*, C57BL/6) such that the 5′ three exons of the gene are located on the X chromosome but the seven exons encoding the carboxy-terminal two-thirds of the protein are located within the PAR and are therefore present on both the X and Y chromosomes [8]. In humans [7,9], the rat, and the wild mouse species *Mus spretus*, the gene is entirely X-unique. Here, we report that the rate of sequence divergence of the 3′ end of the *Fxy* gene is much higher (estimated at 170-fold higher for synonymous sites) when pseudoautosomal (present on both the X and Y chromosomes) than when X-unique. Thus, chromosomal position can directly affect the rate of evolution of a gene. This finding also provides support for the suggestion that regions of the genome with a high recombination frequency, such as the PAR, may have an intrinsically elevated rate of sequence divergence.

Address: Section of Gene Function and Regulation, Chester Beatty Laboratories, Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK.

Correspondence: Alan Ashworth
E-mail: alana@icr.ac.uk

Received: 19 February 1999
Revised: 31 March 1999
Accepted: 21 July 1999

Published: 30 August 1999

Current Biology 1999, 9:987–989
http://biomednet.com/elecref/0960982200900987

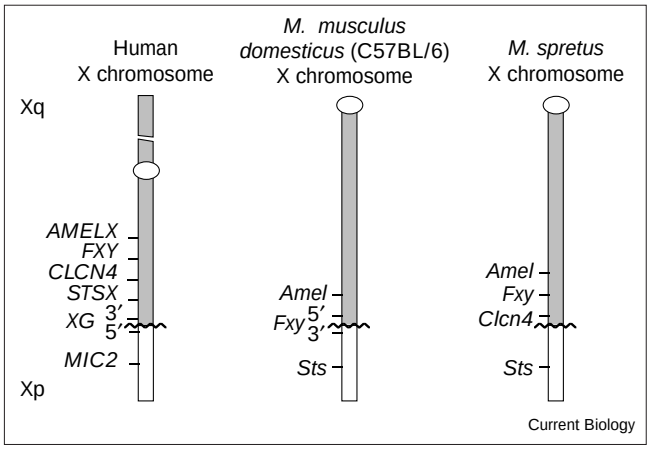
0960-9822/99/$ – see front matter

Results and discussion

In the laboratory mouse C57BL/6, the *Fxy* gene spans the PAB such that the 5′ third of the gene is located on the X chromosome but the 3′ two-thirds are located within the PAR [8] (Figure 1). The PAR is identical on the X and Y chromosomes [3–6] and the 3′ end of the *Fxy* gene, although apparently functionless on the Y, is maintained by recombination; deleterious mutations within the Y fragment would be rapidly transferred onto the X chromosome and selected against. In humans, the *FXY* gene (also known as *MID1*), mutations in which cause Opitz syndrome [7], is wholly X-unique [7,9]. To determine when the rearrangement occurred, which resulted in the placement of the PAB within the *Fxy* gene, we cloned and characterised the gene in the rat species *Rattus norvegicus*, which diverged from the ancestors of mice 10–35 million years ago [10–12], and in the mouse species *M. spretus*, which diverged from the ancestors of laboratory mice about 3 million years ago [13].

To locate the *Fxy* gene in *M. spretus*, we analysed the segregation of variants of the gene in an interspecific backcross (European Collaborative Interspecific Backcross, EUCIB) [14]. Variants derived from both the 5′ and 3′ ends of the gene segregated in an X-unique fashion. Therefore, in *M. spretus*, the *Fxy* gene is completely X-unique, confirming a previous suggestion [15], and is located between the *Amelogenin* (*Amel*) [16] and *Chloride channel* (*Clcn4*) genes [17,18] (see Supplementary material; Figure 1). The *Fxy* gene in rat was also found to be entirely X-unique as shown by Southern blotting (data not

Figure 1



Organisation of the distal part of the X chromosomes of *M. musculus* domesticus and *M. spretus* and the equivalent region in humans. The shaded areas are X-unique and the unshaded areas are pseudoautosomal and common to the X and Y chromosomes. In *M. musculus*, the *Clcn4* gene is located on chromosome 7 [17,18]. The *Fxy* gene spans the PAB in C57BL/6 mice such that the 5′ end of the gene is X-unique but the 3′ end is present on the X and Y chromosomes. In humans and *M. spretus*, all of the *Fxy* gene is X-unique (see Supplementary material). Only relevant genes are indicated and the diagram is not to scale.

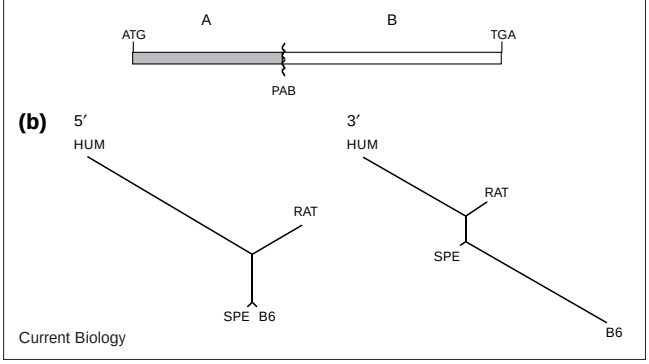
shown). Thus, the rearrangement that led to *Fxy* spanning the PAB in C57BL/6 mice must have occurred within the last 3 million years [13].

We cloned cDNAs for *Fxy* from the rat and *M. spretus* and analysed the sequence divergence of the portions of the gene equivalent to those that are X-unique or located within the PAR in C57BL/6 mice. The part of the gene that is X-unique in both C57BL/6 and *M. spretus* was equally divergent relative to humans and to the rat, as one would expect given the relatively recent divergence time of these two mouse species [13]. Nevertheless, when the 3′ two-thirds of the coding portion of the gene in the four species was compared, it was clear that the C57BL/6 *Fxy* gene was much more divergent from the human and rat sequences than was the *M. spretus* gene (Figure 2). We calculated the number of synonymous and non-synonymous substitutions between these four species in the regions of the *Fxy* gene that correspond to the portion of the gene that is either X-unique or pseudoautosomal in C57BL/6 mice (Figure 2a). As the rodent lineage appears to have a higher intrinsic rate of nucleotide substitution [12,19], we used the rat nucleotide sequence as an outgroup for statistical analysis. In addition, we analysed a partial sequence of the *Fxy* gene from the mouse species *Mus caroli* (see Supplementary material). Calculating the branch lengths for the unrooted tree connecting rat, *M. spretus* and C57BL/6 allowed us to compare the C57BL/6 and *M. spretus* rates directly (see Supplementary material for details). This comparison indicated that there has been a huge acceleration (estimated as 170-fold) in the rate of synonymous substitution in the 3′ part of the C57BL/6 *Fxy* gene since the divergence of *M. musculus* from *M. spretus*, which coincided with the introduction of this part of the gene into the PAR. This increase in the synonymous substitution rate was associated with a modest increase in the non-synonymous rate (estimated as 3.3-fold).

Figure 2

(a)

Region	Comparison	Aligned length (bp)	Differences (bp)	K_s	K_a
A	HUM/SPE	756	49	0.240 (0.050)	0.007 (0.006)
	HUM/B6	756	49	0.240 (0.050)	0.007 (0.006)
	HUM/RAT	756	48	0.235 (0.046)	0.005 (0.006)
	RAT/SPE	756	20	0.098 (0.026)	0.002 (0.002)
	RAT/B6	756	22	0.091 (0.025)	0.004 (0.003)
	B6/SPE	756	4	0.019 (0.019)	0.004 (0.003)
B	HUM/SPE	1245	115	0.439 (0.061)	0.002 (0.002)
	HUM/B6	1245	178	0.672 (0.089)	0.034 (0.009)
	HUM/RAT	1245	119	0.428 (0.063)	0.003 (0.003)
	RAT/SPE	1245	54	0.120 (0.023)	0.003 (0.003)
	RAT/B6	1245	167	0.493 (0.062)	0.035 (0.009)
	B6/SPE	1245	134	0.377 (0.051)	0.032 (0.008)



Sequence divergence of the *Fxy* gene in mice, rats and humans. (a) Evolutionary analysis of the human (HUM), rat, C57BL/6 (B6) and *M. spretus* (SPE) *Fxy* genes. The coding region of *Fxy* has been divided into two regions, A and B, corresponding, respectively, to parts of the gene that are X-unique or located in the PAR in C57BL/6 mice. *K* values for synonymous substitutions (K_S) and non-synonymous substitutions (K_a), expressed as substitutions per site, were calculated as described in the Supplementary material. Standard deviations are shown in brackets. Figures are rounded to three decimal places. (b) Phylogenetic tree analysis of the sequence divergence of the *Fxy* gene. Branch lengths are indicative of sequence divergence. The unrooted tree was calculated using the Maximum Likelihood analysis method version 3.5c which uses the program dnam14 on the PIE interface to the Phylip phylogenetic analysis programs at the Human Genome Mapping Programme (HGMP) Resource Centre, Hinxton, UK.

Our observation of the accelerated divergence of the part of *Fxy* located in the PAR may be generally true of PAR genes. Human PAR genes, where analysed, are autosomal in rodents and appear in general to be particularly divergent (for example, see [20]). This effect is difficult to quantify, however, as it is not known when the relevant chromosomal rearrangements took place. Furthermore, these genes may be intrinsically rapidly evolving. As the rearrangement that led to *Fxy* becoming partially pseudoautosomal was relatively recent, we can be much more confident about the rapid sequence divergence.

Interestingly, the sequence divergence in the 3′ region of the gene was accompanied by an increase in G+C content (47% overall for *M. spretus* and 57% for C57BL/6). This was particularly marked in the third, 'wobble' position of codons: 50% G+C for *M. spretus* and 73% for C57BL/6. There was only a small (3.3%) increase in the first and second codon positions and the 5′ part of the gene was similar in G+C content in all codon positions between the two species. High G+C content has been associated with regions of genomes that have a high frequency of recombination [21,22]. The only other mouse PAR gene isolated, *Sts*, is also highly G+C rich [23] and this may be a property of genes in this part of the genome.

What might be responsible for the accelerated evolution that we have observed? One possibility is that the PAR-located portion of *Fxy* behaves like an autosomal gene; these have been shown to evolve significantly faster than X-linked genes [24–26]. Suggested reasons for the lower rate of mutation of X-linked genes include the exposure of deleterious recessive mutations on the hemizygous X chromosome in males [27] and the greater number of divisions of male than female germ cells [24,28]. Nevertheless, the effect that we observed was much larger than

the approximately twofold effect seen in rodents for either of these possibilities [26,29]. An alternative possibility is that the rapid divergence that we observed is due to the much higher recombination frequency seen in the PAR relative to the rest of the X chromosome. Studies in yeast have suggested that recombinational events *per se* are mutagenic because of errors in double-stranded DNA repair [30]. Thus, it seems possible that the accelerated evolutionary rate that we observed is due to errors in double-strand break repair associated with recombination. Other regions of vertebrate genomes with high recombination rates may also be subject to the same effect and this might result in different rates of sequence divergence of genes in separate evolutionary lineages dependent on their chromosomal context.

Supplementary material

Supplementary material including methods used for the isolation of the rat and *M. spretus Fxy* cDNAs and mapping of the gene in *M. spretus* and additional methodological details are available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

This work was supported by the Wellcome Trust. We thank the HGMP, UK for provision of DNA and computing resources.

References

1. Wolfe KH, Sharp PM, Li WH: Mutation rates differ among regions of the mammalian genome. *Nature* 1989, 337:283-285.
2. Charlesworth D, Charlesworth B: Sequence variation: looking for effects of genetic linkage. *Curr Biol* 1998, 8:R658-R661.
3. Burgoyne PS: Genetic homology and crossing over in the X and Y chromosomes of mammals. *Hum Genet* 1982, 61:85-90.
4. Ellis N, Goodfellow PN: The mammalian pseudoautosomal region. *Trends Genet* 1989, 5:406-410.
5. Rappold GA: The pseudoautosomal regions of the human sex chromosomes. *Hum Genet* 1993, 92:315-324.
6. Blaschke RJ, Rappold GA: Man to mouse — lessons learned from the distal end of the human X chromosome. *Genome Res* 1997, 7:1114-1117.
7. Quaderi NA, Schweiger S, Gaudenz K, Franco B, Rugarli EI, Berger W, *et al.*: Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. *Nat Genet* 1997, 17:285-291.
8. Palmer S, Perry J, Kipling D, Ashworth A: A gene spans the pseudoautosomal boundary in mice. *Proc Natl Acad Sci USA* 1997, 94:12030-12035.
9. Perry J, Feather S, Smith A, Palmer S, Ashworth A: The human ***FXY*** gene is located within Xp22.3: implications for evolution of the mammalian X chromosome. *Hum Mol Genet* 1998, 7:299-305.
10. Wilson AC, Ochman H, Prager EM: Molecular time scale for evolution. *Trends Genet* 1987, 3:241-247.
11. Catzeflis FM, Aguilar J-P, Jaeger J-J: Muroid rodents: phylogeny and evolution. *Trends Ecol Evol* 1992, 7:122-126.
12. Li W-H: *Molecular Evolution*. Sunderland, Massachusetts: Sinauer; 1997.
13. Ferris SD, Sage RD, Prager EM, Ritte U, Wilson AC: Mitochondrial DNA evolution in mice. *Genetics* 1983, 105:681-721.
14. The European Backcross Collaborative Group: Towards high resolution maps of the mouse and human genomes — a facility for ordering markers to 0.1 cM resolution. *Hum Mol Genet* 1994, 3:621-627.
15. Dal Zotto L, Quaderi NA, Elliott R, Lingerfelter PA, Carrel L, Valsecchi V, *et al.*: The mouse ***Mid1*** gene: implications for the pathogenesis of Opitz syndrome and the evolution of the mammalian pseudoautosomal region. *Hum Mol Genet* 1998, 7:489-499.
16. Snead ML, Lau EC, Zeichner-David M, Fincham AG, Woo SL, Slavkin HC: DNA sequence for cloned cDNA for murine amelogenin reveal the amino acid sequence for enamel-specific protein. *Biochem Biophys Res Commun* 1985, 129:812-818.
17. Palmer S, Perry J, Ashworth A: A contravention of Ohno's law in mice. *Nat Genet* 1995, 10:472-476.
18. Rugarli EI, Adler DA, Borsani G, Tsuchiya K, Franco B, Hauge X, *et al.*: Different chromosomal localization of the ***Clcn4*** gene in ***Mus spretus*** and C57BL/6J mice. *Nat Genet* 1995, 10:466-471.
19. Wu CI, Li WH: Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc Natl Acad Sci USA* 1985, 82:1741-1745.
20. Ellison JW, Li X, Francke U, Shapiro LJ: Rapid evolution of human pseudoautosomal genes and their mouse homologs. *Mamm Genome* 1996, 7:25-30.
21. Eyre-Walker A: Recombination and mammalian genome evolution. *Proc R Soc Lond B Biol Sci* 1993, 252:237-243.
22. Charlesworth B: Genetic recombination: patterns in the genome. *Curr Biol* 1994, 4:182-184.
23. Salido EC, Li XM, Yen PH, Martin N, Mohandas TK, Shapiro LJ: Cloning and expression of the mouse pseudoautosomal steroid sulphatase gene (***Sts***). *Nat Genet* 1996, 13:83-86.
24. Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T: Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harb Symp Quant Biol* 1987, 52:863-867.
25. Wolfe KH, Sharp PM: Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. *J Mol Evol* 1993, 37:441-456.
26. McVean GT, Hurst LD: Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. *Nature* 1997, 386:388-392.
27. Charlesworth B, Coyne JA, Barton NH: The relative rates of evolution of sex chromosomes and autosomes. *Am Nat* 1987, 130:113-146.
28. Shimmin LC, Chang BH, Li WH: Male-driven evolution of DNA sequences. *Nature* 1993, 362:745-747.
29. Chang BH, Shimmin LC, Shyue SK, Hewett-Emmett D, Li WH: Weak male-driven molecular evolution in rodents. *Proc Natl Acad Sci USA* 1994, 91:827-831.
30. Strathern JN, Shafer BK, McGill CB: DNA synthesis errors associated with double-strand-break repair. *Genetics* 1995, 140:965-972.

Supplementary material

Evolutionary rate of a gene affected by chromosomal position

Jo Perry and Alan Ashworth

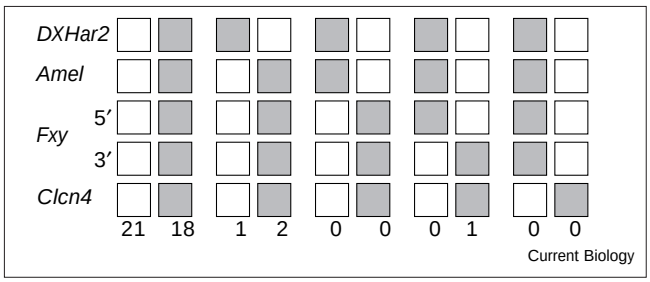
Current Biology 30 August 1999, 9:987–989

Supplementary results and discussion

Location of the *Fxy* gene in *M. spretus*

To locate the *Fxy* gene in *M. spretus*, we analysed the segregation of variants of the gene in an interspecific backcross (EUCIB). Variants derived from both the 5′ and the 3′ end of the gene segregated in an X-unique fashion. Therefore, in *M. spretus*, the *Fxy* gene is completely X-unique, confirming a previous suggestion [S1], and is located between the *Amelogenin* (*Amel*) [S2] and *Chloride channel* (*Clcn4*) gene [S3,S4] (Figure S1). A single recombinant was found between the 5′ and the 3′ ends of the *Fxy* gene, which located the gene between *Amel* and *Clcn4* (see Supplementary materials and methods below for details of the mapping assays).

Figure S1



Genetic mapping of *Fxy* in *M. spretus*. Haplotype analysis of *Fxy* in the EUCIB interspecific backcross. Parental and recombinant X chromosomes are shown for the total of 43 mice analysed. Shaded boxes indicate the presence of the C57BL/6 allele and unshaded boxes the *M. spretus* variant. The markers tested are listed on the left.

Calculation of branch lengths of the 5′ and 3′ portions of the *Fxy* gene in *R. norvegicus*, *M. spretus* and *M. musculus* (C57BL/6)

The program DIVERGE was used to calculate between-species distances, and the branch lengths then calculated by solving pairwise simultaneous equations (See Supplementary materials and methods). We did this using rat as an outgroup. The result is summarised in Table S1.

Comparing *M. musculus* and *M. spretus* directly (possible because of the equal divergence time of these two species), gave a 170-fold increase in the K_s value in region B. K_s was very similar for *M. musculus* and *M. spretus* in region A.

The acceleration in K_a was difficult to quantify, as the solution to the simultaneous equations for the K_a value of the 3′ portion (region B) of *M. spretus* is a negative value. Nevertheless, when we estimated the K_a value by comparing the ratio of region B to region A of the rat gene to *M. musculus*, we obtained a ratio of 0.0032/0.0009 for rat and 0.032/0.003 for *M. musculus*, which is a 3.3-fold increase, a much lower value than when this ratio is compared with humans (40-fold).

Table S1

Branch lengths between rat and mouse *Fxy* genes.

		A	B
K_s	*R. norvegicus*	0.0855	0.1178
	M. musculus	0.0059	0.3751
	M. spretus	0.0128	0.0022
K_a	*R. norvegicus*	0.0009	0.0032
	M. musculus	0.003	0.032
	M. spretus	0.0012	–0.001

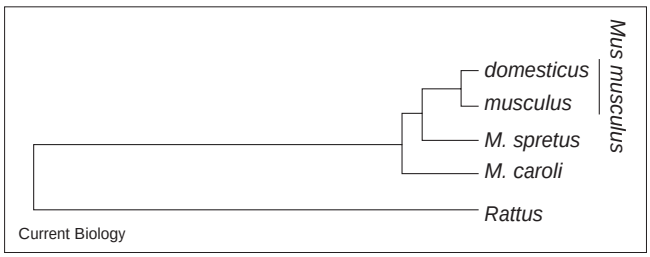
Cloning and evolutionary analysis of the *Fxy* gene from *M. caroli*

To provide a closer outgroup for the analysis of the *Fxy* gene in *M. musculus* (C57BL/6) and *M. spretus*, we analysed part of the gene in the mouse species *M. caroli*. This species separated from the branch leading to *M. musculus* and *M. spretus* approximately 5 million years (Myr) ago [S5] (Figure S2). PCR primers derived from the *M. spretus* cDNA (5′-ATGCCATTGGCCTTGCATAC-3′ and 5′-GGCCTCAACCTTTGTAATTCC-3′) were used to amplify part of exon 10 of the *Fxy* gene from *M. caroli*. An alignment of this sequence with exon 10 from the C57BL/6 and *M. spretus Fxy* gene is shown in Figure S3. It is clear that this part of the gene is very similar in *M. spretus* and *M. caroli*. This suggests that most of the divergence seen in the C57BL/6 gene has occurred since the separation of *M. musculus* and *M. spretus*.

To determine the rate of synonymous and non-synonymous substitutions, we used the DIVERGE program (HGMP, Hinxton, UK; see Supplementary materials and methods), which gives the results, expressed as substitutions per site, shown in Table S2. Solving simultaneous equations for the branch lengths of the unrooted tree for these values gives the values shown in Table S3.

Comparing C57BL/6 and *M. spretus* directly and using *M. caroli* as an outgroup gave a 29-fold increase in the rate

Figure S2



Simplified diagram of rodent evolution. Divergence times are controversial but are approximately: *M. musculus domesticus* and *M. musculus musculus*, < 1 Myr; *M. musculus* and *M. spretus*, 3 Myr; *M. musculus*/*M. spretus* and *M. caroli*, 5 Myr; *Mus species* and *Rattus*, 10–35 Myr [S5,S16–S19].

of synonymous substitution of the C57BL/6 sequence when compared with *M. spretus*.

Although these figures are not directly comparable with the figures derived from the entire gene presented in the main paper, they do strongly support our contention that the divergence of the 3′ part of the *Fxy* gene has occurred coincident with the movement of this portion of the gene into the PAR.

Supplementary materials and methods

Cloning of cDNAs

The *M. spretus Fxy* DNA sequence was obtained by a combination of PCR on *M. spretus* brain cDNA with primers derived from the human nucleotide sequence [S6] and analysis of genomic DNA from the 3′ end of the gene by anchored PCR. Oligonucleotides derived from the human *FXY* nucleotide sequence (5′-TGGCCTTCCTGGAATTGTTG-3′ and 5′-ATCATAGTCCAGCAGGATGC-3′) [S6] amplify nucleotides 1–2074 of the *M. spretus* cDNA using PCR conditions of 94°C, 30 sec; 52°C, 30 sec; 72°C, 90 sec for 30 cycles. The remaining 3′ portion of the gene was obtained by PCR on a *M. spretus* genomic library constructed using a Clontech Universal Genome Walker kit. Nested oligonucleotides derived from exon 10 of the *M. spretus* (5′-ATGCCATTGGCCTTG-CATAC-3′) and the human nucleotide sequence (5′-TGAACTCCATC-CACCTCTAC-3′) [S6] and adapter-specific oligonucleotides (Clontech) amplify a 1.2 kb band corresponding to nucleotides 2102–3348 of the *M. spretus* cDNA from the genomic library.

The rat *Fxy* gene was cloned by PCR from rat liver cDNA. Oligonucleotides derived from the human *FXY* nucleotide sequence (5′-TGGC-CTTCCTGGAATTGTTG-3′ and 5′-GGCCTCAACCTTTGTAATTCC-3′) [S6], amplify a product of 2.2 kb which includes the 5′ UTR and termination codon of the gene. Southern analysis showed that the gene is X-unique in this species as is the case in *M. spretus* and humans (data not shown). Automated DNA sequencing was performed by cycle sequencing with Applied Biosystems Big Dye-containing terminators and analysed on an Applied Biosystems 377 sequencer.

Table S2

Synonymous and non-synonymous substitution rates.

	K_s	K_a
M. spretus/*M. caroli*	0.0276	0
C57BL/6/*M. caroli*	0.6616	0.0848
C57BL/6/*M. spretus*	0.6786	0.0848

Genetic mapping

To map the 5′ of the *M. spretus Fxy* gene, oligonucleotides derived from intronic sequence 5′ to exon 2 of the *M. spretus Fxy* gene (5′-CAAAACTCTCCGGGGTCATC-3′ and 5′-GCCTCTTCACCCTAAT-GAAC-3′) were used to amplify a 1.1 kb product from C57BL/6 and *M. spretus* genomic DNA with the PCR conditions of 94°C, 30 sec; 52°C, 30 sec; 72°C, 60 sec for 38 cycles. Digestion of this PCR product with *Hinf*I yields two bands of 480 bp and 620 bp from C57BL/6 genomic DNA and three bands of 480, 330 and 290 bp from *M. spretus* genomic DNA. To map the 3′ end of the *M. spretus Fxy* gene, oligonucleotides derived from exon 9 of human *FXY* gene [S6] (5′-ACTG-GATCCCAAATCTGCTC-3′ and 5′-GTGCTTCCACTTATGACCAC-3′, which amplify nucleotides 1547–1739 of the published human sequence) were used to amplify a 192 bp product from *M. spretus* genomic DNA using the PCR conditions of 94°C, 30 sec; 60°C, 30 sec; 72°C, 30 sec for 28 cycles. These oligonucleotides do not amplify a corresponding product in C57BL/6 genomic DNA.

To map the *Amel* gene in *M. spretus*, oligonucleotides derived from intronic sequence 5′ to exon 4 and from exon 6 of the *Amel* genomic DNA sequence [S2] (5′-CTGACACATAGAAAGCATCC-3′ and 5′-TTGGTCTTGTCTGTCGCTGG-3′, respectively) were used to amplify an 890 bp product from both C57BL/6 and *M. spretus* genomic DNA with the PCR conditions of 94°C, 30 sec; 52°C, 30 sec; 72°C, 60 sec for 38 cycles. Digestion of the *M. spretus* PCR product with *Bst*EII results in two bands of 205 and 685 bp with no corresponding restriction site in the C57BL/6 PCR product. The *Clcn4* gene in *M. spretus* was mapped using an assay which has been previously described [S3].

These assays were used to type mice from the EUCIB. Animals from the portion of the backcross to C57BL/6 mice were scored and the data analysed using the Mbx program [S7].

Evolutionary analysis

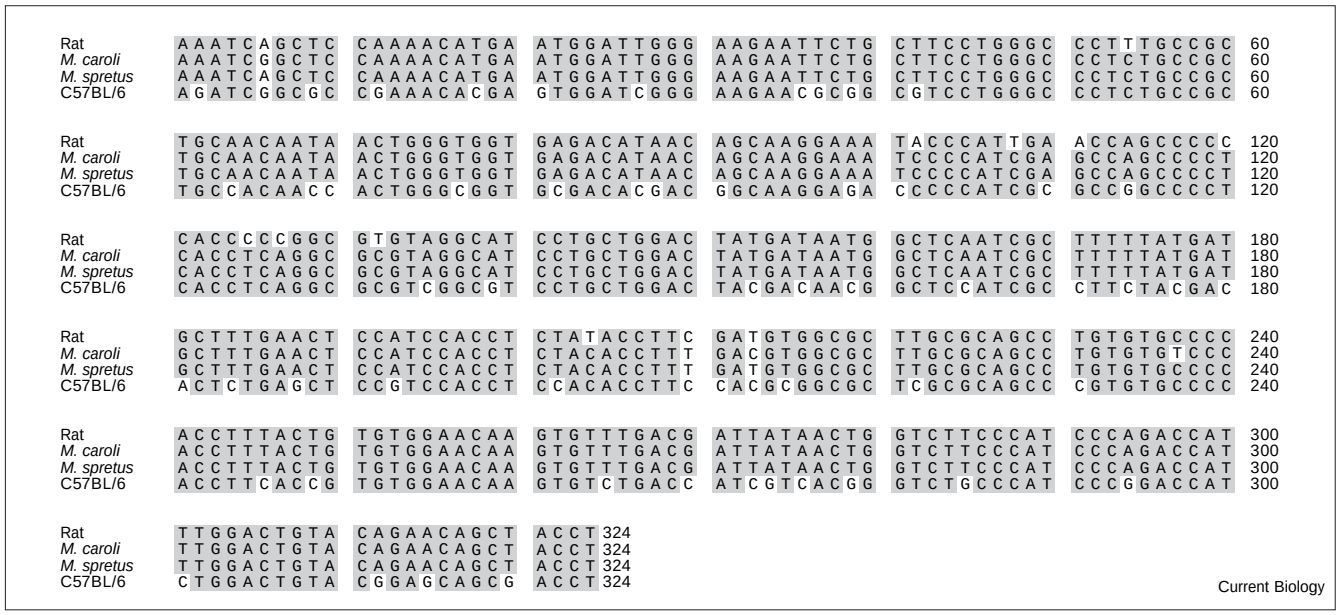
Sequence comparisons were performed using BLASTN [S8,S9] on the NCBI website [S10]. Sequences were aligned using the PILEUP program of the GCG suite of programs at the HGMP Resource Centre, Hinxton, UK. The evolutionary distance for synonymous (silent) substitutions (K_s) or non-synonymous substitutions (K_a) between two sequences was calculated with the GCG program DIVERGE which is based on the method described by Li *et al.* [S11] as modified by Li [S12] and Pamilo and Bianchi [S13] and applies Kimura's two parameter method [S14] to correct for multiple hits and to account for the difference in substitution rates for translesions and transversions. The branch lengths for the unrooted tree connecting rat, *M. spretus* and C57BL/6 were calculated by solving the appropriate simultaneous equations for these three species which allowed us to compare the C57BL/6 and *M. spretus* values directly using the rat sequence as an outgroup (See Table S1). Phylogenetic analysis was performed using

Table S3

Divergence of exon 10 of the *Fxy* gene in mice.

K_s	*M. caroli*	0.0053
	M. spretus	0.0223
	C57BL/6	0.6563
K_a	*M. caroli*	0
	M. spretus	0
	C57BL/6	0.0848

Figure S3



Alignment of sequences derived from exon 10 of the *Fxy* gene from various *Mus* species and the rat. This region corresponds to nucleotides 1781–2174 of the published *M. musculus* (C57BL/6) *Fxy* sequence. Sequences were aligned using the PILEUP program of the GCG suite of programs at the HGMP Resource Centre, Hinxton, UK.

the Maximum Likelihood method version 3.5c which uses the program dnam14 on the PIE interface to the Phylip phylogenetic analysis programs at the HGMP [S15].

Supplementary references

S1. Dal Zotto L, Quaderi NA, Elliott R, Lingerfelter PA, Carrel L, Valsecchi V, *et al.*: The mouse ***Mid1*** gene: implications for the pathogenesis of Opitz syndrome and the evolution of the mammalian pseudoautosomal region. *Hum Mol Genet* 1998, 7:489-499.

S2. Snead ML, Lau EC, Zeichner-David M, Fincham AG, Woo SL, Slavkin HC: DNA sequence for cloned cDNA for murine amelogenin reveal the amino acid sequence for enamel-specific protein. *Biochem Biophys Res Commun* 1985, 129:812-818.

S3. Palmer S, Perry J, Ashworth A: A contravention of Ohno's law in mice. *Nat Genet* 1995, 10:472-476.

S4. Rugarli EI, Adler DA, Borsani G, Tsuchiya K, Franco B, Hauge X, *et al.*: Different chromosomal localization of the ***Clcn4*** gene in ***Mus spretus*** and C57BL/6J mice. *Nat Genet* 1995, 10:466-471.

S5. Ferris SD, Sage RD, Prager EM, Ritte U, Wilson AC: Mitochondrial DNA evolution in mice. *Genetics* 1983, 105:681-721.

S6. Perry J, Feather S, Smith A, Palmer S, Ashworth A: The human ***FXY*** gene is located within Xp22.3: implications for evolution of the mammalian X chromosome. *Hum Mol Genet* 1998, 7:299-305.

S7. The European Backcross Collaborative Group: Towards high resolution maps of the mouse and human genomes – a facility for ordering markers to 0.1 cM resolution. *Hum Mol Genet* 1994, 3:621-627.

S8. Karlin S, Altschul SF: Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc Natl Acad Sci USA* 1990, 87:2264-2268.

S9. Karlin S, Altschul SF: Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc Natl Acad Sci USA* 1993, 90:5873-5877.

S10. National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov/

S11. Li WH, Wu CI, Luo CC: A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol* 1985, 2:150-174.

S12. Li WH: Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol* 1993, 36:96-99.

S13. Pamilo P, Bianchi NO: Evolution of the ***Zfx*** and ***Zfy*** genes: rates and interdependence between the genes. *Mol Biol Evol* 1993, 10:271-281.

S14. Kimura M: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980, 16:111-120.

S15. Human Genome Mapping Programme: http://menu.hgmp.mrc.ac.uk/menu-bin/PIE/pie.pl

S16. Li W-H: *Molecular Evolution*. Sunderland, Massachusetts: Sinauer; 1997.

S17 Makalowski W, Boguski MS: Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc Natl Acad Sci USA* 1998, 95:9407-9412.

S18. Wilson AC, Ochman H, Prager EM: Molecular time scale for evolution. *Trends Genet* 1987, 3:241-247.

S19. Catzeflis FM, Aguilar J-P, Jaeger J-J: Muroid rodents: phylogeny and evolution. *Trends Ecol Evol* 1992, 7:122-126.