GENOMICS **29,** 353–363 (1995)

CpG Islands in Human *ZFX* and *ZFY* and Mouse *Zfx* Genes: Sequence Similarities and Methylation Differences

SHIUH-WEN LUOH,* KARIN JEGALIAN,* ANGELA LEE,† ELLSON Y. CHEN,‡ ANNE RIDLEY,*,[1] AND DAVID C. PAGE*,[2]

*Howard Hughes Research Laboratories at Whitehead Institute and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142; †Department of Medicine/Immunology, Stanford University, Stanford, California 94305-5487; and ‡Advanced Center for Genetic Technology, Applied Biosystems Division, Perkin–Elmer Corporation, Foster City, California 94404

Received March 22, 1995; accepted July 12, 1995

The human *ZFX*, human *ZFY*, and mouse *Zfx* genes have CpG islands near their 5′ ends. These islands are typical in that they span about 1.5 kb, contain transcription initiation sites, and encompass some 5′ untranslated exons and introns. However, comparative nucleotide sequencing of these human and mouse islands provided evidence of evolutionary conservation to a degree unprecedented among mammalian 5′ CpG islands. In one stretch of 165 nucleotides containing 19 CpGs, mouse *Zfx* and human *ZFX* are identical to each other and differ from human *ZFY* at only 9 nucleotides. In contrast, we found no evidence of homologous CpG islands in the mouse *Zfy* genes, whose transcription is more circumscribed than that of human *ZFX*, human *ZFY*, and mouse *Zfx*. Using the isoschizomers *Hpa*II and *Msp*I to examine a highly conserved segment of the *ZFX* CpG island, we detected methylation on inactive mouse X chromosomes but not on inactive human X chromosomes. These observations parallel the previous findings that mouse *Zfx* undergoes X inactivation while human *ZFX* escapes it.

INTRODUCTION

The existence of genes common to the mammalian X and Y chromosomes poses important evolutionary questions. In the past decade, investigators have discovered more than a dozen genes that are common to the X and Y chromosomes in humans or in mice. Several of these X–Y genes are located in the pseudoautosomal regions, where sequence identity between the two chromosomes is enforced by X–Y recombination during normal male meiosis. Other X–Y gene pairs are located in strictly sex-linked regions, where X–Y recombination does not occur (Affara *et al.*, 1994).

ZFX and *ZFY*, the first X–Y genes found in the strictly sex-linked portions of the mammalian X and Y chromosomes (Page *et al.*, 1987), encode proteins comprising an amino-terminal acidic domain, a putative nuclear localizing signal, and a carboxy-terminal domain of 13 zinc fingers (Ashworth *et al.*, 1989; Lau and Chan, 1989; Mardon and Page, 1989; Schneider-Gadicke *et al.*, 1989a). By analogy to other zinc-finger proteins, the ZFX and ZFY proteins probably bind DNA or RNA in a sequence-specific manner. They may function as transcription activators (Mardon *et al.*, 1990). The biological processes in which these proteins act are not known, although possible roles in gonadal sex determination, Turner syndrome, and spermatogenesis have been proposed and debated (Page *et al.*, 1987, 1990; Burgoyne, 1989; Koopman *et al.*, 1989, 1991; Mardon and Page, 1989; Palmer *et al.*, 1989; Simpson and Page, 1991; Zambrowicz *et al.*, 1994a).

We have previously suggested that *ZFX* and *ZFY* began diverging from a single common ancestral gene prior to the radiation of placental mammals, at least 60 to 80 million years ago. This proposal originated from the observation that most if not all placental mammals carry X-specific and Y-specific homologs of the human *ZFY* gene (Page *et al.*, 1987). Consistent with this hypothesis, the mammalian *ZFX* and *ZFY* genes were subsequently found to encode similar but distinct proteins (Ashworth *et al.*, 1989; Mardon and Page, 1989; Schneider-Gadicke *et al.*, 1989a; Mardon *et al.*, 1990; Palmer *et al.*, 1990).

The model also predicted that the intron/exon structures of *ZFX* and *ZFY* would have much in common. This prediction has not yet been definitively tested, since mouse *Zfx* is the only gene in this family whose structure has been comprehensively described (Luoh and Page, 1994). However, comparison with limited data available for mouse *Zfy*, human *ZFX*, and human *ZFY* (Schneider-Gadicke *et al.*, 1989a,b; Simpson and Page, 1991; Shimmin *et al.*, 1993) has revealed a num-

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. U00241–2.

[1] Present address: Ludwig Institute for Cancer Research, 91 Riding House Street, London WC1P 8BT, United Kingdom.

[2] To whom correspondence should be addressed. Telephone: (617) 258-5203. Fax: (617) 258-5578.

0888-7543/95 $12.00

ber of conserved splice sites and suggested that the intron/exon structures of mammalian *ZFX* and *ZFY* are very similar if not identical. The structures of the extant mammalian *ZFX* and *ZFY* genes are probably much like that of the ancestral gene from which they diverged (Luoh and Page, 1994).

Only small portions of the *ZFX* and *ZFY* genes, which span 50 to 70 kb (Schneider-Gadicke *et al.*, 1989a,b; Luoh and Page, 1994), have been sequenced in any mammal. Nonetheless, analysis of the limited genomic sequence available—and of more extensive cross-hybridization data (Page *et al.*, 1987; Schneider-Gadicke *et al.*, 1989b)—suggests that high similarity between mammalian *ZFX* and *ZFY* sequences is intermittent and restricted largely to coding exons. These findings are consistent with *ZFX* and *ZFY* facing shared evolutionary constraints operating at the level of the encoded proteins. As a rule, *ZFX/ZFY* introns, untranslated exons, and flanking sequences exhibit much less nucleotide similarity than their coding exons (Schneider-Gadicke *et al.*, 1989b; Shimmin *et al.*, 1993), presumably because shared evolutionary pressures are weaker or less extensive in those regions.

An exception to the generally low nucleotide similarity outside coding sequences has been noted within "CpG islands" located near the 5′ ends of the human *ZFX* and *ZFY* genes. CpG islands are unusual segments of the genome, typically about a kilobase in length and often containing transcription initiation sites, that have a high G + C content and in which the dinucleotide CpG is abundant (Bird, 1986; Larsen *et al.*, 1992). In the bulk of the human or mouse genome, CpGs are methylated, and such methylated CpGs are prone to mutate to TpG or CpA, resulting in CpG loss and underrepresentation. In CpG islands, by contrast, CpG dinucleotides are generally not methylated and are maintained at frequencies approximating those of the dinucleotide GpC. Clustering of recognition sites for the restriction endonucleases *Bss*HII, *Eag*I, and *Sac*II has revealed CpG islands near the 5′ ends of both the human *ZFY* and *ZFX* genes (Page *et al.*, 1987; Pritchard *et al.*, 1987; Schneider-Gadicke *et al.*, 1989b). More intensively studied was the CpG island of mouse *Zfx*, where nucleotide sequencing demonstrated that the island is 1.5 kb in length and has a G + C content of 74% (Luoh and Page, 1994) (corrected). The mouse *Zfx* CpG island was shown to possess promoter activity, to contain multiple transcription initiation sites, and to include the first two exons, both of which are untranslated (Luoh and Page, 1994). In Southern blotting experiments, the CpG islands of human *ZFY* and *ZFX* were found to cross-hybridize at high stringency, indicating that the nucleotide sequences of the two human genes are similar in this 5′ region (Schneider-Gadicke *et al.*, 1989b).

Given that *ZFY* and *ZFX* appear to have diverged from a single ancestral gene prior to the radiation of placental mammals, these results implied that the nucleotide sequences of the CpG islands either had been highly conserved on both the X and the Y chromosomes or had converged during human evolution (e.g., by gene conversion). Evidence for conservation rather than convergence was provided by the observation that the CpG island of human *ZFY* cross-hybridized at high stringency to genomic DNAs of a wide range of placental mammals. Indeed, the CpG island was one of four apparently conserved segments—the other three containing coding exons—whose hybridization to X- and Y-specific restriction fragments in diverse placental mammals established the rough outlines of the human *ZFY* transcription unit (Page *et al.*, 1987).

These results suggested that CpG islands of the *ZFX* and *ZFY* genes display unusual degrees of intraspecies and interspecies nucleotide similarity, and this possibility motivates the present study. Rather little is known about forces constraining or driving the evolution of mammalian CpG islands. Where CpG islands associated with homologous genes have been compared in two or more species, preservation of G + C-rich character and CpG content are often observed (Aissani and Bernardi, 1991). However, apart from recognized transcription factor binding sites (e.g., Zacksenhaus *et al.*, 1993), conservation of precise nucleotide sequence in the noncoding portion of CpG islands is generally unremarkable (A. Bird, Edinburgh, pers. comm., 11 Oct. 1994, confirmed by review of literature). Since the divergence of the mammalian *ZFY* and *ZFX* genes from a single ancestral gene began at least 60 million years ago, a detailed comparison of the associated CpG islands in various species should provide insight into the constraints within which these 5′ sequences evolved. In the present study, we compared the CpG islands of the human and mouse genes. We uncovered conservation whose extent and degree are, to our knowledge, unprecedented in the noncoding regions of mammalian structural genes.

We also set out to determine whether mouse *Zfx* and human *ZFX*, which differ markedly with respect to X inactivation, might also differ with respect to methylation of their CpG islands. Like most genes on the mouse X chromosome, *Zfx* undergoes X inactivation; it is transcribed on active but not inactive X chromosomes (Ashworth *et al.*, 1991; Zinn *et al.*, 1991). In contrast, the human *ZFX* gene escapes inactivation; it is transcribed on both active and inactive X chromosomes (Schneider-Gadicke *et al.*, 1989b). In mice and humans, most X-linked CpG islands are unmethylated on active X chromosomes but heavily methylated on inactive X chromosomes (Wolf *et al.*, 1984; Pfeifer *et al.*, 1990; Tribioli *et al.*, 1992; Singer-Sam and Riggs, 1993). Indeed, on inactive mouse X chromosomes, the CpG island of *Zfx* appears to be methylated (Erickson *et al.*, 1993). Thus, for many X-linked CpG islands, methylation correlates with the inactivation status of the "host" chromosome. Given that human *ZFX* escapes X inactivation while mouse *Zfx* does not, we set out to compare methylation of the associated CpG islands on inactive and active X chromosomes in the two species. We hoped to learn whether *Zfx/ZFX* CpG island methylation more closely

reflects the inactivation status of the gene or that of the host chromosome.

MATERIALS AND METHODS

Sequencing of human ZFX and ZFY genomic DNA clones. Portions of the human genomic inserts of two plasmids were sequenced. Plasmid pDP1047 (Schneider-Gadicke *et al.*, 1989b) contains a 5.7-kb *Hin*dIII fragment from the human X chromosome (phage λBER113) subcloned into pBluescript (Stratagene). Plasmid pDP1024 contains a 5.2-kb *Hin*dIII fragment from the human Y chromosome (phage λOX107; Page *et al.*, 1987) subcloned into pBluescript.

These human genomic DNA clones were sequenced by dideoxynucleotide chain termination (Sanger *et al.*, 1977; Chen *et al.*, 1991) using synthetic oligonucleotide primers and modified T7 polymerase (Sequenase II, United States Biochemical). To sequence a 3.5-kb portion of *ZFX* plasmid pDP1047, plasmid subclones were constructed by restriction digestion or by using *Exo*III and S1 nucleases (Henikoff, 1984). Sequencing templates were (1) single-stranded DNAs rescued from pBluescript KS(−) constructs using helper phage VCS-M13 or (2) supercoiled, double-stranded DNAs prepared by alkaline lysis. To sequence the entirety of *ZFY* plasmid pDP1024, its 5.2-kb *Hin*dIII insert was digested at a unique *Sal*I site to yield 3.2-kb and 2.0-kb fragments. After subcloning into plasmid pUC119, these fragments were sequenced using synthetic "walking" primers spaced every 400 to 500 bp.

High G + C content made the sequencing difficult. Ambiguities caused by compression and other artifacts were resolved using nucleotide analogs (e.g., 7-deaza-dGTP and dITP; United States Biochemical; Mizusawa *et al.*, 1986), modified reaction conditions (Sanger *et al.*, 1977; Chen *et al.*, 1991), or 6% polyacrylamide/8 *M* urea/20% formamide gels.

Cloning of human ZFY cDNA. We previously constructed a cDNA library (Fisher *et al.*, 1990) using poly(A)$^+$ RNA from OXEN, a human lymphoblastoid cell line derived from a 49,XYYYY male (Sirota *et al.*, 1981). This library was screened by hybridization with the insert of plasmid pDP1207, which contains a 0.3-kb *Pst*I–*Sal*I fragment from the 5′ portion of the human *ZFY* genomic locus (nucleotides 6 through 314 as numbered in Fig. 2). (The insert of pDP1207 was subcloned from phage λOX107; Page *et al.*, 1987.) One cDNA clone was identified in this screen, and its 1.4-kb insert was transferred into the *Not*I site of Bluescript SK(+), generating plasmid pDP1297. Partial sequencing of pDP1297 revealed that it was collinear with but extended further 5′ than previously described *ZFY* cDNAs (Lau and Chan, 1989; Palmer *et al.*, 1990). The cDNA insert of pDP1297 appeared to be truncated at its 3′ end, probably because of internal priming during reverse transcription.

Southern blot analysis of mouse genomic DNAs. Genomic DNAs prepared from livers of male and female FVB/N mice were digested with restriction endonucleases, subjected to electrophoresis in 0.8% agarose, and transferred (Southern, 1975) to nylon membrane. A 395-bp *Bss*HII genomic fragment (nucleotides −303 to 92 in Fig. 2) from human *ZFY* was labeled with ^{32}P by random-primer synthesis (Feinberg and Vogelstein, 1984) and hybridized overnight to the genomic DNA transfer at 47°C in 50% formamide, 5× SSC (1× SSC = 0.15 *M* NaCl, 15 m*M* Na citrate, pH 7.4), 1× Denhardt's (0.02% Ficoll 400, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin), 20 m*M* Sodium phosphate, pH 6.6, 50 μg/ml denatured salmon sperm DNA, 1% SDS (sodium dodecyl sulfate). Following hybridization, the transfer membrane was washed three times for 15 min each at 65°C in 0.1× SSC, 0.1% SDS and exposed at −80°C with X-ray film backed with an intensifying screen for 4 days.

Methylation analysis of human and mouse genomic DNAs. One hundred nanograms mouse or human genomic DNA was incubated with 10 units *Hin*dIII or *Hpa*II or *Msp*I for 4 h at 37°C in buffers recommended by the manufacturer (New England Biolabs). The digested genomic DNAs were then used as template in PCR with primers (CTACCCTTCCGCATTTTCCT and GAGCTCGGAGCTGACAAAAA) chosen from sequences conserved between mouse *Zfx* and human *ZFX* and spanning, in both species, a 105-bp region containing two CCGG sites. PCR using 100 ng template DNA was carried out in 20 μl of 12.5 m*M* Tris, pH 8.2, 50 m*M* KCl, 12.5 m*M* NaCl, 5 m*M* NH_4Cl, 2.5 m*M* $MgCl_2$, and 1 m*M* each of the two primers. After heating to 100°C for 5 min, the four deoxyribonucleotides (to a final concentration of 0.125 m*M* each) and 2 units *Taq* polymerase were added. Thirty cycles of 1 min at 94°C, 1 min at 62°C, and 1 min at 72°C were followed by extension for 2 min at 72°C.

PCR products were subjected to electrophoresis in 4% NuSieve agarose (FMC Corporation), 90 m*M* Tris–borate, 2 m*M* EDTA, 0.5 μg/ml ethidium bromide, visualized with UV light, and transferred to nylon membrane in preparation for Southern hybridization. The hybridization probe was an oligonucleotide, GGTGACGTGACGTGCTGACG, chosen from sequence within the PCR product that was conserved completely between mouse and human. The oligonucleotide was labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase and allowed to hybridize with the filter overnight at 42°C in 6× SSC, 5× Denhardt's, 0.05% $Na_4P_2O_7$, 100 μg/ml tRNA, and 0.5% SDS. The filter was then washed three times for 20 min each at 42°C in 6× SSC, 0.1% SDS and exposed with X-ray film for 2 days.

RESULTS AND DISCUSSION

Sequencing of CpG Islands at 5′ Ends of Human ZFX and ZFY

We had previously characterized the CpG island of mouse *Zfx* (Luoh and Page, 1994). To compare the 5′ CpG islands of human *ZFX*, human *ZFY*, and mouse *Zfx* in detail, we sequenced a 3.5-kb portion of the human *ZFX* genomic locus (Fig. 1) and a 5.2-kb portion of the human *ZFY* genomic locus (Fig. 2). In the sequenced region of *ZFX*, the frequency of CpG dinucleotides ranges from 0/150 to 31/150 nucleotides (just as in mouse *Zfx;* Fig. 3) and serves to demarcate the CpG island. Applying an arbitrary threshold of 10 CpG dinucleotides per 150 nucleotides, the human *ZFX* CpG island measures 1.5 kb in length, almost identical in length to the mouse *Zfx* CpG island. Within this 1.5-kb CpG island of human *ZFX*, 76% of the nucleotides are either G or C, and CpG and GpC dinucleotides are comparably abundant. In *ZFY*, the CpG island spans 1.3 kb, and the incidence of CpGs, which ranges from 0/150 to 28/150 nucleotides, is slightly lower than that in its human and mouse X homologs (Fig. 3). Within this 1.3-kb CpG island, G+C content is 68%, and CpG and GpC dinucleotides occur at similar frequencies. [The segment sequenced was larger for *ZFY* than for human *ZFX* and mouse *Zfx;* the additional *ZFY* material sequenced contains a second, smaller region (centered at +1900; Fig. 3) that clears the arbitrary threshold of 10 CpGs per 150 nucleotides].

In mouse *Zfx*, the CpG island comprises the 5′-most portion of the transcription unit (Luoh and Page, 1994). Similarly, human *ZFX* transcripts appear to have high G+C content near their 5′ termini, and the 5′ portions of four human *ZFX* cDNA clones whose sequences have been reported (Schneider-Gadicke *et al.*, 1989a; Palmer *et al.*, 1990) are identical in sequence to the human *ZFX* CpG island (Fig. 1). The 5′ ends of these four cDNAs all lie within 142 nucleotides near the CpG island's midpoint. All four *ZFX* cDNAs employ the same first splice donor, at nucleotide 49 (Fig. 1), although they exhibit diverse patterns of splicing downstream

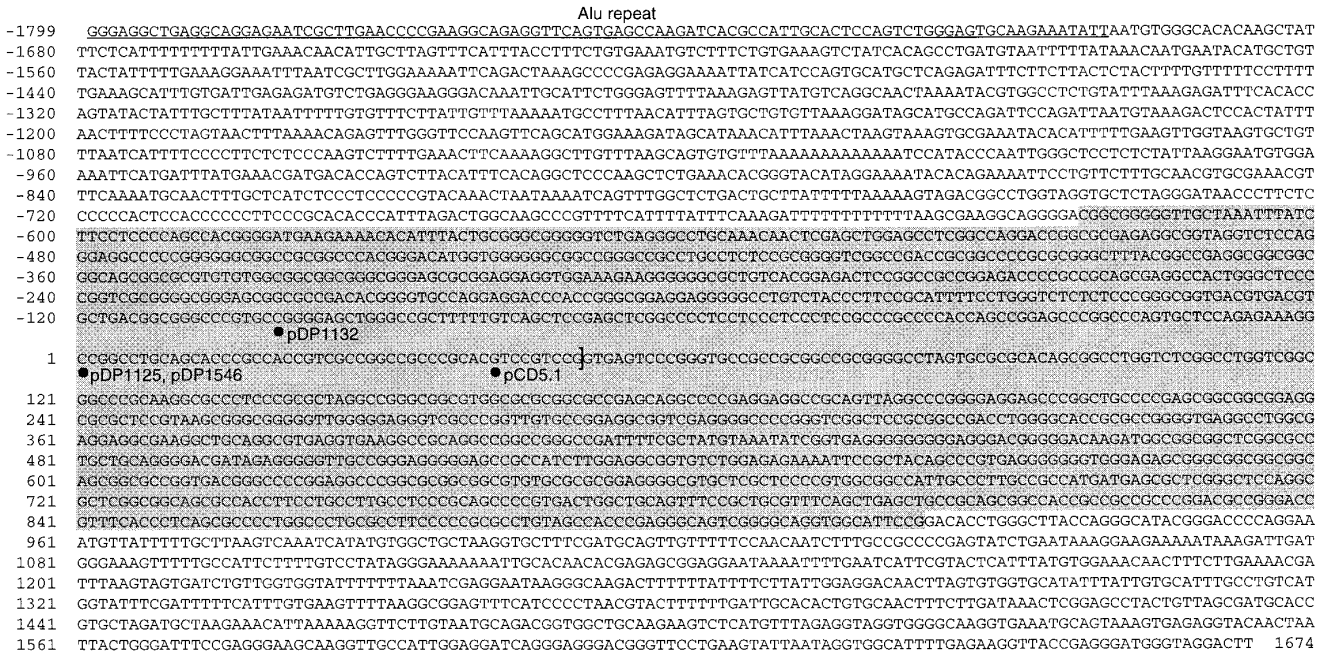


FIG. 1. Nucleotide sequence of the 5′ portion of the human *ZFX* gene. This 3.5-kb portion of the human X chromosome (GenBank Accession No. U00241; plasmid pDP1047) includes (1) an *Alu* repeat (underlined), (2) a CpG island (shaded), (3) the 5′ ends of four cDNA clones (dots beneath sequence), and (4) the 3′ boundary of exon 1 (right bracket). Nucleotides are numbered according to 5′ nucleotide in cDNAs pDP1125 and pDP1546. The numbering scheme brings human *ZFX* and mouse *Zfx* into register within a region of sequence identity; human *ZFX* nucleotide 1 (as numbered here) is homologous to mouse *Zfx* nucleotide 1 [in Fig. 5A of Luoh and Page (1994)]. Origin of cDNA clones: pDP1132, pDP1125, and pDP1546 correspond to cDNAs 1, 3, and 2, respectively, of Schneider-Gadicke *et al.* (1989a); pCD5.1 was described by Palmer *et al.* (1990).

(Schneider-Gadicke *et al.*, 1989a). It appears that in human *ZFX*, as in mouse *Zfx*, the CpG island comprises the 5′-most portion of the transcription unit.

For human *ZFY*, no evidence that the 5′ CpG island is part of the transcription unit had been reported. A single human *ZFY* cDNA clone extending 5′ of the initiator codon had been described (Lau and Chan, 1989), but that cDNA's 5′ portion did not have a remarkably high G+C content. We identified one additional cDNA clone that extended 5′ of the initiator codon. Partial sequencing revealed that this clone, pDP1297, was collinear with previously described human *ZFY* cDNAs (Lau and Chan, 1989; Palmer *et al.*, 1990) but extended 44 nucleotides further 5′. The most 5′ portion of cDNA pDP1297 has a high G+C content and is identical in sequence to a portion of the *ZFY* CpG island (Fig. 2). Sequence alignment with pDP1297 reveals that the 5′ terminus of the previously reported *ZFY* cDNA (Lau and Chan, 1989) actually falls within the CpG island (Fig. 2). Both cDNAs employ the same first splice donor, at nucleotide 49 (Fig. 2). These findings collectively demonstrate that the 5′ CpG island is part of the human *ZFY* transcription unit, as in human *ZFX* and mouse *Zfx*.

Conservation of Nucleotide Sequence in CpG Island

Previous studies of evolutionary conservation of the mammalian *ZFX/ZFY* CpG islands had relied exclusively upon Southern blot hybridization ("ark blots"; Page *et al.*, 1987). Sequencing of the human *ZFX* and *ZFY* CpG islands (Figs. 1 and 2), together with the prior sequencing of the mouse *Zfx* CpG island (Luoh and Page, 1994), provided an opportunity to explore this conservation with greater precision.

Pairwise comparisons between the sequences of the three genes revealed that their 5′ CpG islands are remarkably similar (Fig. 4). For example, the mouse *Zfx* CpG island is highly similar in sequence to the human *ZFX* CpG island along nearly all of its 1.5-kb length (Fig. 4A). This similarity is not a trivial corollary of the two sequences having high G+C contents, as dot-plot analysis reveals little similarity off the diagonal. Neither is this the result of the mouse and human X chromosomes having had little time to diverge; the mouse–human similarity is restricted to the CpG island and does not extend to flanking sequences.

It has been suggested that gene conversion during primate evolution accounts for much of the present-day similarity between the coding sequences of human *ZFX* and *ZFY* (Hayashida *et al.*, 1992). Indeed, in most of the coding exons (exons 5 through 9 and exon 11, the last of these encoding the zinc-finger domain), the DNA sequence of human *ZFX* is more closely related to that of human *ZFY* than to that of mouse *Zfx* (Table 1). If gene conversion were a major force in the evolution of the CpG islands, then one might again expect the greatest similarity to be exhibited by the two human genes. Alternatively, if sequence similarities among the *ZFX/ZFY* CpG islands were due primarily to simple conservation, i.e., a low rate of fixation of mutations,

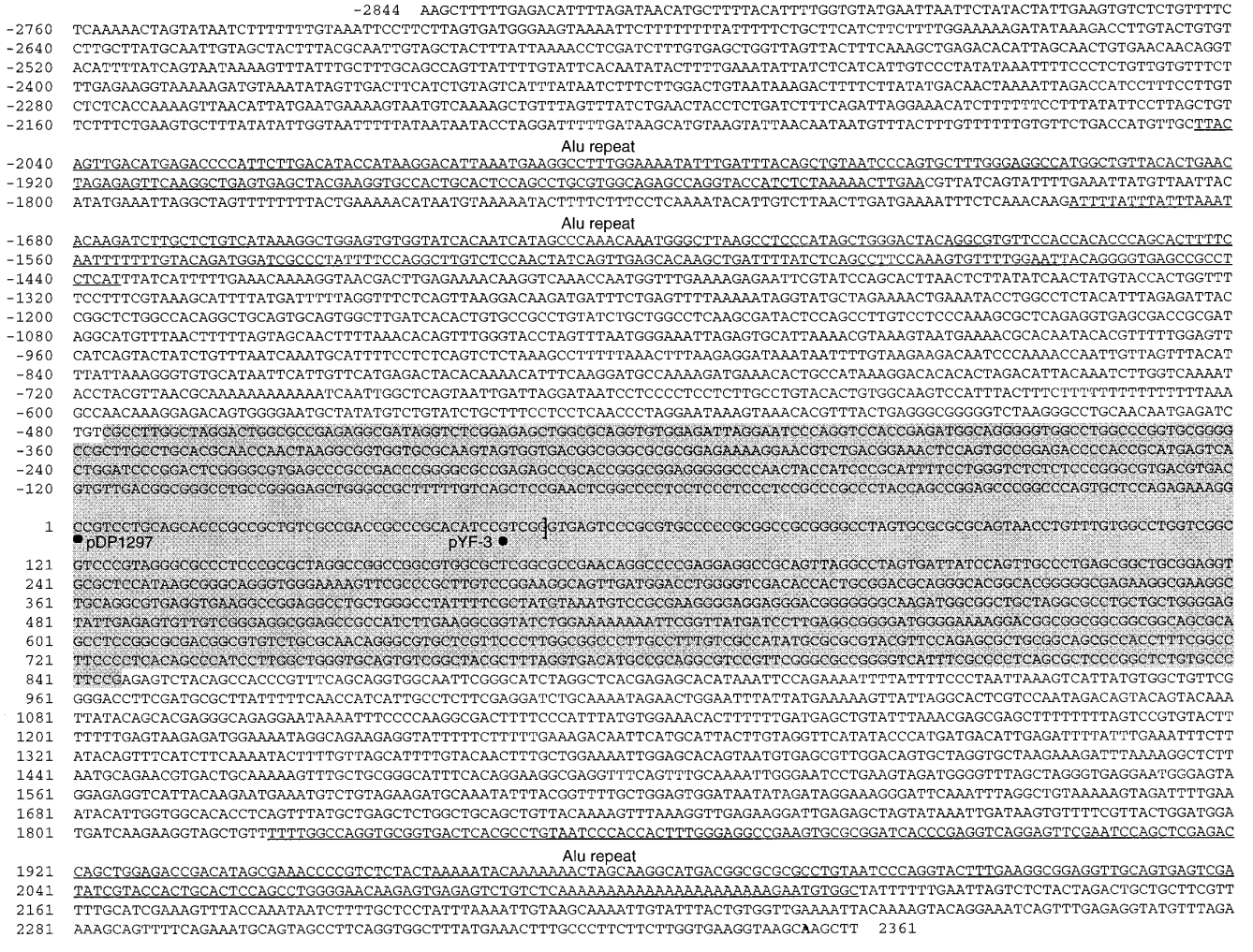


FIG. 2. Nucleotide sequence of the 5′ portion of the human *ZFY* gene. This 5.2-kb portion of the human Y chromosome (GenBank Accession No. U00242; plasmid pDP1024) includes (1) three *Alu* repeats (underlined), (2) a CpG island (shaded), (3) the 5′ ends of two cDNA clones (dots beneath sequence), and (3) the 3′ boundary of exon 1 (right bracket). Nucleotides numbered according to the 5′ nucleotide in cDNA pDP1297. The numbering scheme brings human *ZFY* into register with human *ZFX* and mouse *Zfx;* human *ZFY* nucleotide 1 (as numbered here) is homologous to human *ZFX* nucleotide 1 (Fig. 1) and to mouse *Zfx* nucleotide 1 [Fig. 5A of Luoh and Page (1994)]. Origin of cDNA clones: pDP1297, this paper; pYF-3 described by Lau and Chan (1989).

one might expect the greatest similarity to be exhibited by the two X-linked genes, since the split between murine and human lineages appears to have occurred after divergence of *ZFX* from *ZFY* began. The latter expectation is borne out by our analysis: The greatest similarity is exhibited by the mouse *Zfx* and human *ZFX* CpG islands (Fig. 4A). In the case of mouse *Zfx* and human *ZFY* (Fig. 4B), and also in the case of human *ZFX* and human *ZFY* (Fig. 4C), sequence similarity is somewhat less extensive and less uniform, although nonetheless striking in its extent, nearly 1 kb in both instances. Again, in each pairwise comparison, nucleotide similarity is limited to the CpG island and does not extend to flanking sequences. Thus, dot-plot analysis (1) reveals remarkable mouse–human and human X–human Y similarities, (2) demonstrates that the similarities are restricted to the CpG island itself, and (3) provides evidence that similarity has not been created anew by gene conversion but instead has been maintained by restricted divergence of homologs from their common ancestor. (Of course, the present data do not allow us to rule out small gene conversion events. Comparative study of a Y-linked mouse homolog could strengthen our conclusion, but such study is not possible in the case of the CpG island, as discussed below.)

Closer examination of the dot plots suggested that the most striking sequence similarities are found in or about the first 5′-untranslated exon (Fig. 4). This was confirmed by direct comparison of nucleotide sequences, which revealed near identity among mouse *Zfx*, human *ZFX*, and human *ZFY* about the transcription initiation sites, in the first 5′-untranslated exon, and extending downstream into the first intron (Fig. 5). Again, the two X-linked genes show the highest similarity. Mouse *Zfx* and human *ZFX* are absolutely identical to each other and are 95% identical to human *ZFY* in a 165-bp portion of the first 5′ untranslated exon (nucleotides −136 through +29, as numbered in Fig.

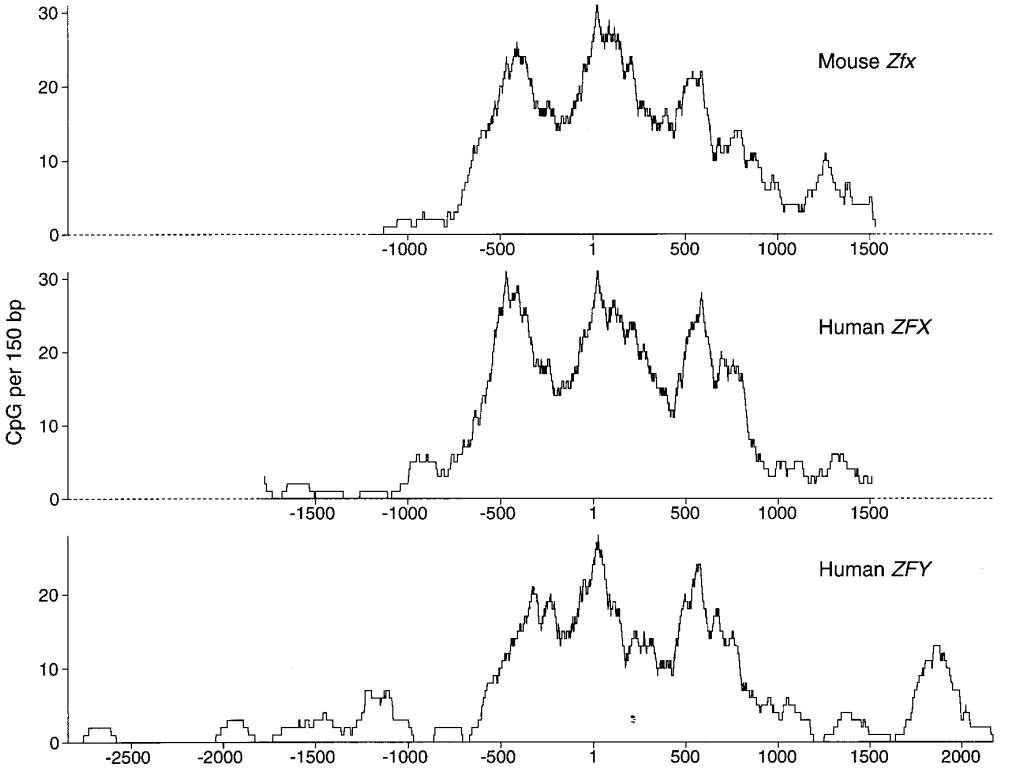


FIG. 3. CpG dinucleotide frequency in 5′ portions of mouse *Zfx*, human *ZFX*, and human *ZFY* genes. Graphs depict number of CpGs per 150 nucleotides immediately following indicated nucleotide. Sequences analyzed: mouse *Zfx*, 2.9-kb segment in Fig. 5A of Luoh and Page (1994); human *ZFX*, 3.5-kb segment in Fig. 1 of this paper; human *ZFY*, 5.2-kb segment in Fig. 2 of this paper. Sequences aligned at nucleotide 1, as described in Figs. 1 and 2. Outside sequenced regions of mouse *Zfx* and human *ZFX*, the *x* axis is shown as dashed line.

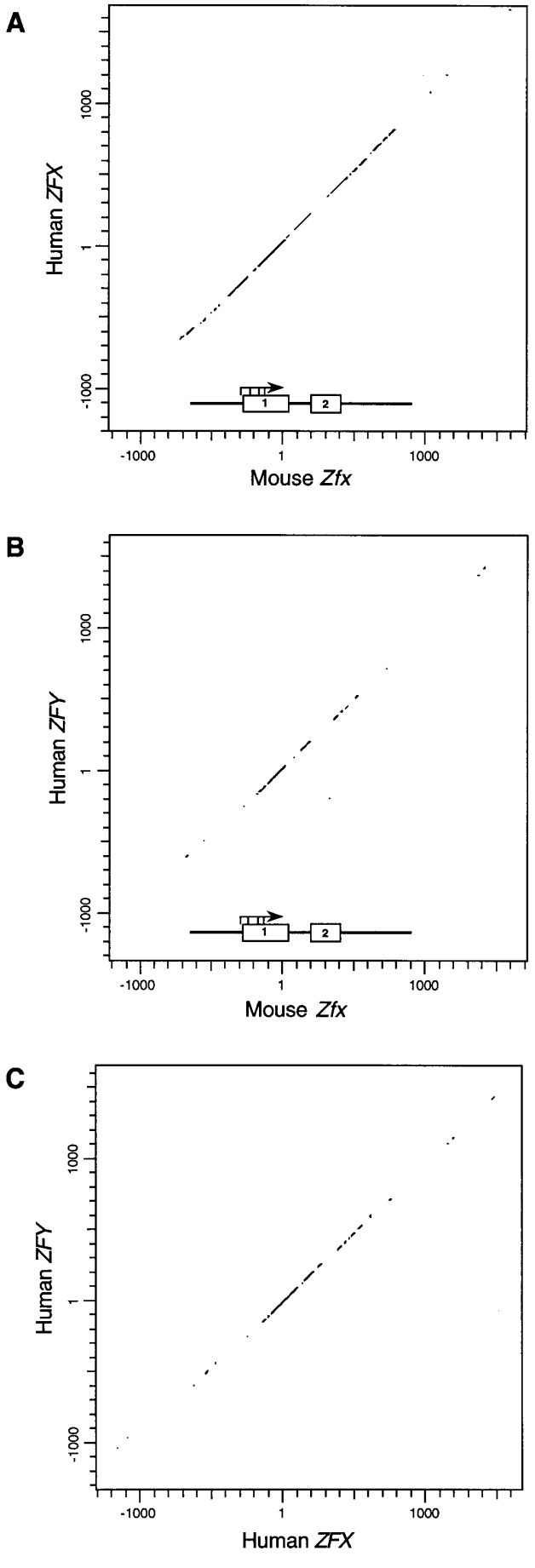
5), a region that contains 19 CpGs and a mouse *Zfx* transcription initiation site. This degree of nucleotide sequence conservation rivals or exceeds that seen in the most conserved portions of the *ZFX/ZFY* coding regions, i.e., exons 10 and 11, the latter encoding the zinc-finger domain (Table 1).

The alignment of mouse *Zfx*, human *ZFX*, and human *ZFY* sequences shown in Fig. 5 also helped us recognize two striking parallels in the structure of the three genes' transcripts as captured in cDNA clones. First, all three genes employ the same (or, more strictly, a homologous) first splice donor, at nucleotide 49. Second, for each of the three genes, at least one cDNA clone has been obtained whose 5′ terminus is at the same (homologous) nucleotide [position 1 as numbered in Figs. 1, 2, and 5; also see Fig. 5A of Luoh and Page (1994)]. We would not be surprised if the transcription start sites employed by human *ZFX* and *ZFY*, which have not yet been identified, were homologous to the sites used by mouse *Zfx*, most of which fall within conserved portions of the CpG island (Fig. 5 and Luoh and Page, 1994).

Why does the 5′ CpG island of the mammalian *ZFX* and *ZFY* genes display such extensive conservation of nucleotide sequence? The importance of this question is underscored by the general lack of such conservation in the few other 5′ CpG islands whose sequence has been determined in two or more mammalian species (e.g., Zacksenhaus *et al.*, 1993; A. Bird, Edinburgh, pers. comm., 11 Oct. 1994). In other conserved portions of *ZFX* and *ZFY*, nucleotide sequence is maintained across evolutionary time by functional constraints on the encoded proteins (Schneider-Gadicke *et al.*, 1989b; Mardon *et al.*, 1990). This explanation seems not to apply to the CpG island, since we find no evidence that it encodes a protein (i.e., no conserved long open reading frame). Such conservation of nucleotide sequence is most uncommon in the noncoding regions of protein-encoding mammalian genes, and we would appreciate readers bringing other such examples to our attention. Might the *ZFX/ZFY* CpG island have duties that are broader or more demanding than those of most mammalian CpG islands? If so, the conservation of precise nucleotide sequence suggests to us that this CpG island functions as a polynucleotide, either DNA or RNA, perhaps in transcriptional or translational regulation of *ZFX*, *ZFY*, or other genes. Perhaps this polynucleotide forms a functionally important secondary structure, which might account for the identical (or homologous) 5′ ends found in several mouse *Zfx*, human *ZFX*, and human *ZFY* cDNA clones (at position +1; Fig. 5).

No Closely Related CpG Island on Mouse Y Chromosome

The Y chromosomes of most placental mammals, apart from rodents, appear to carry a single homolog of the human *ZFY* gene (Page *et al.*, 1987). In contrast,



the mouse Y chromosome carries two homologous genes, *Zfy1* and *Zfy2* (Page *et al.*, 1987; Mardon and Page, 1989; Mitchell *et al.*, 1989; Nagamine *et al.*, 1989; Simpson and Page, 1991), the result of an intrachromosomal duplication that occurred during rodent evolution (Mardon *et al.*, 1989).

Sequences closely related to the CpG island of human *ZFY* were previously detected on the Y chromosomes of most placental mammals tested, but not mice (Page *et al.*, 1987). The negative results obtained with mice in these earlier Southern blotting experiments could not be interpreted with confidence, since the CpG island was not well defined then, and the hybridization probes employed were not optimal, extending well beyond what we now know to be the bounds of the CpG island. We repeated these experiments using a smaller, 395-bp hybridization probe derived from the most conserved portion of the CpG island. Under conditions where this CpG island probe hybridized strongly to mouse *Zfx* (and to X- and Y-specific restriction fragments in humans; not shown), the probe detected no other locus in the mouse genome (Fig. 6). We conclude that there are no closely related CpG islands on the mouse Y chromosome. This is in accord with recent data, which suggest that the mouse *Zfy1* and *Zfy2* genes employ TATA box rather than CpG island promoters (Zambrowicz *et al.*, 1994b).

The absence of highly conserved 5′ CpG islands adds to the list of characteristics known to distinguish the mouse *Zfy1* and *Zfy2* genes from their most thoroughly studied homologs: mouse *Zfx*, human *ZFX*, and human *ZFY*. Although the mouse Zfy proteins exhibit the usual two-domain structure—an amino-terminal acidic half and a carboxy-terminal string of 13 zinc fingers—their amino acid sequences differ substantially from those of mouse Zfx, human ZFX, and human ZFY. In the zinc-finger domain, for example, the mouse Zfx protein is 99.5% identical to human ZFX, 97% identical to human ZFY, but only 80% identical to mouse Zfy1 or Zfy2 (Page *et al.*, 1987; Ashworth *et al.*, 1989; Mardon and Page, 1989; Schneider-Gadicke *et al.*, 1989b; Mardon *et al.*, 1990).

Perhaps of more direct relevance to the issue of 5′ CpG islands are dramatic differences in patterns of expression. While the human *ZFX* and *ZFY* and mouse *Zfx* genes seem to be ubiquitously expressed (Schnei-

FIG. 4. Pairwise comparisons of DNA sequences from 5′ portions of mouse *Zfx*, human *ZFY*, and human *ZFX* genes. (**A**) Dot-matrix comparison of a 3.0-kb portion of human *ZFX* (nucleotides −1300 through +1674 as shown in Fig. 1) with a 2.9-kb portion of mouse *Zfx* (entire sequence in Fig. 5A of Luoh and Page, 1994). Analysis (Maizel and Lenk, 1981) employed a "window" of 19 nucleotides and "stringency" of 17. Above the *x* axis is a drawing of the 5′ portion of mouse *Zfx* indicating the locations of a CpG island (black line), exons 1 and 2 (open boxes), and four transcription initiation sites defined by S1 nuclease analysis (Luoh and Page, 1994). (**B**) Comparison of a 3.0-kb portion of human *ZFY* (nucleotides −1344 through +1656 as shown in Fig. 2) with mouse *Zfx*. (**C**) Comparison of human *ZFY* with human *ZFX*.

TABLE 1

Identity among Human *ZFX*, Human *ZFY*, and Mouse *Zfx* in Coding Regions

	Human *ZFX* vs mouse *Zfx*		Human *ZFY* vs mouse *Zfx*		Human *ZFY* vs human *ZFX*	
Domain	DNA	Protein	DNA	Protein	DNA	Protein
Acidic (exons 5–9)[a]	955/1072 (89%)	325/359 (91%)	922/1075 (86%)	303/359 (84%)	981/1075 (92%)	319/364 (88%)
Acidic + nuclear localization (exon 10)	141/141 (100%)	47/47 (100%)	138/141 (98%)	46/47 (98%)	138/141 (98%)	46/47 (98%)
Zinc finger (exon 11)	1104/1184 (93%)	391/393 (99%)	1089/1184 (92%)	383/393 (97%)	1124/1184 (95%)	383/393 (97%)

Note. Nucleotide and amino acid identities counted after optimal alignment of each pair of sequences; insertions and deletions were not considered. Sources of nucleotide and predicted amino acid sequence: human *ZFX,* Schneider-Gadicke *et al.,* 1989a; mouse *Zfx,* Mardon *et al.,* 1990; human *ZFY,* Page *et al.,* 1987, and Lau and Chan, 1989.

[a] Exons numbered as in mouse *Zfx* (Luoh and Page, 1994).

der-Gadicke *et al.*, 1989b; Mardon *et al.*, 1990; Palmer *et al.*, 1990), mouse *Zfy1* and *Zfy2* appear to be transcribed only in the testes, at least in the adult (Ashworth *et al.*, 1989; Mardon and Page, 1989; Nagamine *et al.*, 1989, 1990). (Evidence of other sites of *Zfy* expression in mouse embryos has been obtained by RT-PCR, immunohistochemistry, and *lacZ* transgene studies; Koopman *et al.*, 1989; Nagamine *et al.*, 1990; Su and Lau, 1992; Zwingman *et al.*, 1993; Zambrowicz *et al.*, 1994a.) Given that mammalian genes with CpG islands tend to be more widely expressed than those without CpG islands (Larsen *et al.*, 1992), perhaps it is not surprising that the loss of the CpG island during mouse *Zfy* evolution would be associated with more circumscribed expression. Three other genes known to be associated with a CpG island in human but not in mouse show restricted expression in both species (Antequera and Bird, 1993).

Interspecies Differences in Methylation of CpG Island Parallel Differences in X Inactivation

Most X-linked CpG islands are unmethylated on active X chromosomes but heavily methylated on inactive X chromosomes (Wolf *et al.*, 1984; Pfeifer *et al.*, 1990; Tribioli *et al.*, 1992; Singer-Sam and Riggs, 1993). We examined methylation of the mouse *Zfx* and human *ZFX* CpG islands. Using the experimental strategy outlined in Fig. 7A, we assayed the methylation status of two consecutive *Hpa*II/*Msp*I (CCGG) recognition sites in the most highly conserved portion of the CpG island. *Hpa*II cleaves only when the central CpG dinucleotide is unmethylated, while *Msp*I cleaves the site regardless of methylation status. Mouse and human genomic DNAs were digested with *Hpa*II or *Msp*I and used as template in PCR reactions with primers flanking the two CCGG sites. In principle, only methylated DNA

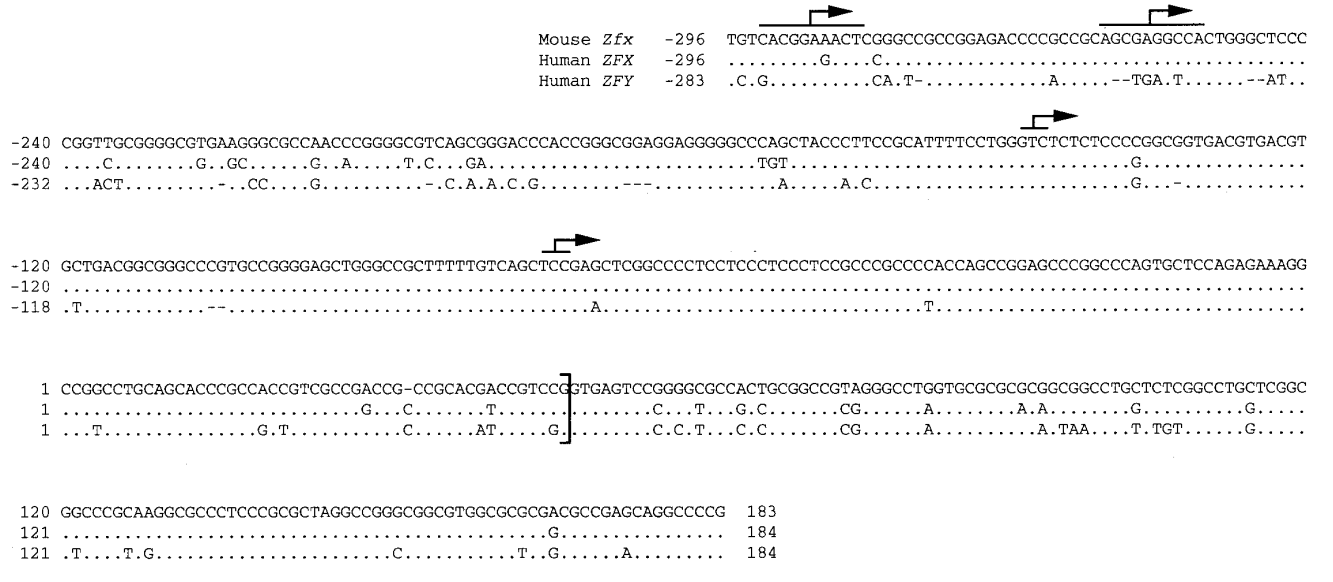


FIG. 5. Alignment of mouse *Zfx,* human *ZFX,* and human *ZFY* genomic DNA sequences in a region of high similarity. Dots represent identity to *Zfx.* Dashes indicate gaps in one sequence compared with another. Four transcription initiation sites in mouse *Zfx* (Luoh and Page, 1994) are indicated by arrows. Right bracket indicates 3′ boundary of exon 1 in all three genes. Human *ZFX* and *ZFY* nucleotides numbered as in Figs. 1 and 2. Mouse *Zfx* nucleotides numbered as in Fig. 5A of Luoh and Page (1994).

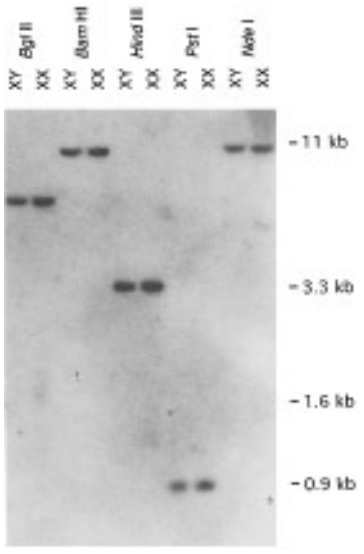


FIG. 6. A single mouse homolog of human *ZFY* CpG island detected by Southern blot analysis. A 395-bp *Bss*HII fragment from a human *ZFY* insert of pDP1024 was ^{32}P-labeled using random oligonucleotide primers and hybridized to Southern transfer of male and female mouse genomic DNAs digested with five different restriction endonucleases. Sequencing and restriction mapping of mouse *Zfx* genomic DNA clones (Luoh and Page, 1994) strongly suggest that 0.9-kb *Pst*I and 3.3-kb *Hin*dIII fragments observed derive from *Zfx*.

incubated with *Hpa*II should support subsequent PCR amplification.

Using DNAs from mouse XY males or human XYp- females, little or no amplification was seen after *Hpa*II (or *Msp*I) digestion (Figs. 7B and 7C). Since such cells carry a single, active X chromosome, this indicated that the two sites tested are unmethylated on active X chromosomes in both mouse and human, as one would have predicted. Female DNAs, which derive from cells with one active and one inactive X chromosome, yielded quite different results depending upon the species. As one might have expected (especially given the results of Erickson *et al.*, 1993), amplification was seen after *Hpa*II (but not *Msp*I) digestion of mouse female DNA, indicating the presence of methylated sites, presumably on the inactive X chromosome. The result of greatest interest was the lack of amplification after *Hpa*II digestion of human female DNA, indicating the absence of methylated sites. Thus, for human *ZFX*, the CpG island—or at least the two sites tested—are unmethylated on both active and inactive X chromosomes. We conclude that methylation of the *Zfx*/*ZFX* CpG island reflects the expression status of the gene rather than the inactivation status of the host chromosome.

Nothing is known about the molecular mechanism by which *ZFX* and a minority of other X-linked human genes escape inactivation (Race and Sanger, 1975; Shapiro *et al.*, 1979; Migeon *et al.*, 1982; Goodfellow *et al.*, 1984; Schneider-Gadicke *et al.*, 1989b; Brown and Willard, 1990; Fisher *et al.*, 1990; Ellison *et al.*, 1992; Schiebel *et al.*, 1993; Slim *et al.*, 1993; Agulnik *et al.*, 1994). Nor is it understood why certain homologous genes on the human and mouse X chromosomes, e.g., *ZFX* and *Zfx*, *RPS4X* and *Rps4*, *UBE1* and *Ube1x*, differ with respect to inactivation (Adler *et al.*, 1991; Ashworth *et al.*, 1991; Kay *et al.*, 1991; Zinn *et al.*, 1991). On both issues, speculation has centered on the genes' promoters: might the promoters or other regulatory se-

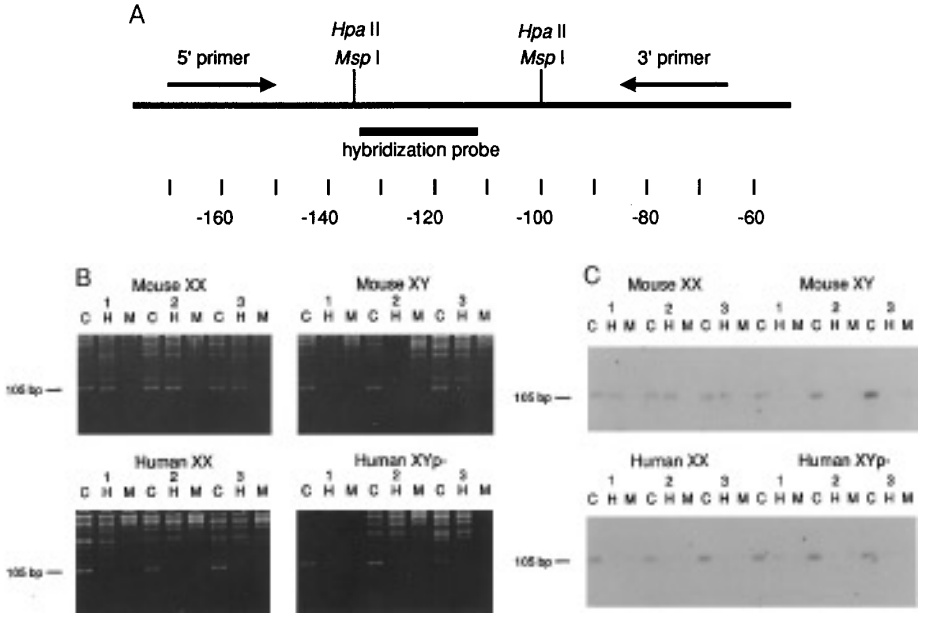


FIG. 7. Male–female and mouse–human differences in methylation of *Zfx*/*ZFX* CpG islands. Mouse and human genomic DNAs digested with *Hpa*II (H), *Msp*I (M), or *Hin*dIII (C, "control") were used as alternate templates in PCR assay (**A**). Locations of *Hpa*II/*Msp*I recognition (CCGG) sites and of PCR primers and hybridization probe (all perfectly complementary to both mouse and human) are indicated; there are no *Hin*dIII sites between the primers. Nucleotides are numbered as in Fig. 5. (**B**) PCR products from three individuals of each species and sex chromosome constitution visualized by ethidium bromide/UV staining after agarose gel electrophoresis. Human XYp- individuals lack the *ZFY* gene but retain *ZFX* (Blagowidow *et al.*, 1989; Cantrell *et al.*, 1989). (**C**) Southern blot autoradiogram of the same gel hybridized with oligonucleotide internal to PCR primers.

quences of genes that escape X inactivation differ in some fundamental way from those that do not (e.g., Lyon, 1993)? The results presented here do not resolve these difficult questions but may help in structuring experimental approaches to them. Let us consider two observations. First, the 5′ portion of the mouse *Zfx* CpG island displays promoter activity (Luoh and Page, 1994), and it seems likely that this would also be true of the human *ZFX* CpG island. Second, the nucleotide sequences of the mouse and human CpG islands are remarkably similar (Figs. 4 and 5). If the mouse–human dichotomy with respect to inactivation is due to differences in regulatory sequences, then those differences must either lie outside the CpG island or involve rather subtle changes in sequence. Experiments by which to validate or refute these possibilities should now be considered.

ACKNOWLEDGMENTS

We are grateful to Rebecca Mosher, whose studies of *ZFY/ZFX* CpG island conservation provided motivation for the present work. We thank Laura Brown for valuable contributions; Robert Dredge for graphical analysis of CpG content; Adrian Bird for thoughtful advice; and Paul Bain, Bruce Lahn, Renee Reijo, and Elizabeth Simpson for comments on the manuscript. This work was supported by the National Institutes of Health. A.R. was supported by a postdoctoral fellowship from the European Molecular Biology Organization.

REFERENCES

Adler, D. A., Bressler, S. L., Chapman, V. M., Page, D. C., and Disteche, C. M. (1991). Inactivation of the *Zfx* gene on the mouse X chromosome. *Proc. Natl. Acad. Sci. USA* **88:** 4592–4595.

Affara, N. A., Lau, Y. F., Briggs, H., Davey, P., Jones, M. H., Khwaja, O., Mitchell, M., and Sargent, C. (1994). Report of the first international workshop on Y chromosome mapping. *Cytogenet. Cell Genet.* **67:** 359–402.

Agulnik, A. I., Mitchell, M. J., Mattei, M.-G., Borsani, G., Avner, P. A., Lerner, J. L., and Bishop, C. E. (1994). A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. *Hum. Mol. Genet.* **3:** 879–884.

Aissani, B., and Bernardi, G. (1991). CpG islands: Features and distribution in the genomes of vertebrates. *Gene* **106:** 173–183.

Antequera, F., and Bird, A. (1993). Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA* **90:** 11995–11999.

Ashworth, A., Swift, S., and Affara, N. (1989). Sequence of cDNA for murine *Zfy-1,* a candidate for *Tdy. Nucleic Acids Res.* **17:** 2864.

Ashworth, A., Rastan, S., Lovell-Badge, R., and Kay, G. (1991). X-chromosome inactivation may explain the difference in viability of XO humans and mice. *Nature* **351:** 406–408.

Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209–213.

Blagowidow, N., Page, D. C., Huff, D., and Mennuti, M. T. (1989). Ullrich–Turner syndrome in an XY female fetus with deletion of the sex-deermining portion of the Y chromosome. *Am. J. Med. Genet.* **34:** 159–162.

Brown, C. J., and Willard, H. F. (1990). Localization of a gene that escapes inactivation to the X chromosome proximal short arm: Implications for X inactivation. *Am. J. Hum. Genet.* **46:** 273–279.

Burgoyne, P. S. (1989). Mammalian sex determination: Thumbs down for zinc finger? *Nature* **342:** 860–862.

Cantrell, M. A., Bicknell, J. N., Pagon, R. A., Page, D. C., Walker, D. C., Saal, H. M., Zinn, A. B., and Disteche, C. M. (1989). Molecular analysis of 46,XY females and regional assignment of a new Y-chromosome-specific probe. *Hum. Genet.* **83:** 88–92.

Chen, E. Y., Kuang, W.-J., and Lee, A. L. (1991). Overview of manual and automated DNA sequencing by the dideoxy chain termination method. *Methods* **3:** 3–19.

Ellison, J. W., Ramos, C., Yen, P. H., and Shapiro, L. J. (1992). Structure and expression of the human pseudoautosomal gene XE7. *Hum. Mol. Genet.* **1:** 691–696.

Erickson, R. P., Zwingman, T., and Ao, A. (1993). Gene expression, X-inactivation, and methylation during spermatogenesis: The case of *Zfa, Zfx,* and *Zfy* in mice. *Mol. Reprod. Dev.* **35:** 114–120.

Feinberg, A. P., and Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. *Anal. Biochem.* **137:** 266–267.

Fisher, E. M. C, Beer-Romero, P., Brown, L. G., Ridley, A., McNeil, J. A., Lawrence, J. B., Willard, H. F., Bieber, F. R., and Page, D. C. (1990). Homologous ribosomal protein genes on the human X and Y chromosomes: Escape from X inactivation and possible implications for Turner syndrome. *Cell* **63:** 1205–1218.

Goodfellow, P., Pym, B., Mohandas, T., and Shapiro, L. J. (1984). The cell surface antigen locus, *MIC2X,* escapes X-inactivation. *Am. J. Hum. Genet.* **36:** 777–782.

Hayashida, H., Kuma, K., and Miyata, T. (1992). Interchromosomal gene conversion as a possible mechanism for explaining divergence patterns of *ZFY*-related genes. *J. Mol. Evol.* **35:** 181–183.

Henikoff, S. (1984). Unindirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28:** 351–359.

Kay, G. F., Ashworth, A., Penny, G. D., Dunlop, M., Swift, S., Brockdorff, N., and Rastan, S. (1991). A candidate spermatogenesis gene on the mouse Y chromosome is homologous to ubiquitin-activating enzyme E1. *Nature* **354:** 486–489.

Koopman, P., Gubbay, J., Collignon, J., and Lovell-Badge, R. (1989). *Zfy* gene expression patterns are not compatible with a primary role in mouse sex determination. *Nature* **342:** 940–942.

Koopman, P., Ashworth, A., and Lovell-Badge, R. (1991). The *ZFY* gene family in humans and mice. *Trends Genet.* **7:** 132–136.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13:** 1095–1107.

Lau, Y. F., and Chan, K. M. (1989). The putative testis-determining factor and related genes are expressed as discrete-sized transcripts in adult gonadal and somatic tissues. *Am. J. Hum. Genet.* **45:** 942–952.

Luoh, S. W., and Page, D. C. (1994). The structure of the *Zfx* gene on the mouse X chromosome. *Genomics* **19:** 310–319.

Lyon, M. F. (1993). Epigenetic inheritance in mammals. *Trends Genet.* **9:** 123–128.

Maizel, J. V., Jr., and Lenk, R. P. (1981). Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc. Natl. Acad. Sci. USA* **78:** 7665–7669.

Mardon, G., Mosher, R., Disteche, C. M., Nishioka, Y., McLaren, A., and Page, D. C. (1989). Duplication, deletion, and polymorphism in the sex-determining region of the mouse Y chromosome. *Science* **243:** 78–80.

Mardon, G., and Page, D. C. (1989). The sex-determining region of the mouse Y chromosome encodes a protein with a highly acidic domain and 13 zinc fingers. *Cell* **56:** 765–770.

Mardon, G., Luoh, S. W., Simpson, E. M., Gill, G., Brown, L. G., and Page, D. C. (1990). Mouse Zfx protein is similar to Zfy-2: Each contains an acidic activating domain and 13 zinc fingers. *Mol. Cell. Biol.* **10:** 681–688.

Migeon, B. R., Shapiro, L. J., Norum, R. A., Mohandas, T., Axelman, J., and Dabora, R. L. (1982). Differential expression of steroid sulphatase locus on active and inactive human X chromosome. *Nature* **299:** 838–840.

Mitchell, M., Simon, D., Affara, N., Ferguson-Smith, M., Avner, P., and Bishop, C. (1989). Localization of murine X and autosomal

sequences homologous to the human Y located testis-determining region. *Genetics* **121:** 803–809.

Mizusawa, S., Nishimura, S., and Seela, F. (1986). Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. *Nucleic Acids Res.* **14:** 1319–1324.

Nagamine, C. M., Chan, K. M., Kozak, C. A., and Lau, Y. F. (1989). Chromosome mapping and expression of a putative testis-determining gene in mouse. *Science* **243:** 80–83.

Nagamine, C. M., Chan, K., Hake, L. E., and Lau, Y. F. (1990). The two candidate testis-determining Y genes (*Zfy-1* and *Zfy-2*) are differentially expressed in fetal and adult mouse tissues. *Genes Dev.* **4:** 63–74.

Page, D. C., Mosher, R., Simpson, E. M., Fisher, E. M. C, Mardon, G., Pollack, J., McGillivray, B., de la Chapelle, A., and Brown, L. G. (1987). The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* **51:** 1091–1104.

Page, D. C., Fisher, E. M., McGillivray, B., and Brown, L. G. (1990). Additional deletion in sex-determining region of human Y chromosome resolves paradox of X,t(Y;22) female. *Nature* **346:** 279–281.

Palmer, M. S., Sinclair, A. H., Berta, P., Ellis, N. A., Goodfellow, P. N., Abbas, N. E., and Fellous, M. (1989). Genetic evidence that *ZFY* is not the testis-determining factor. *Nature* **342:** 937–939.

Palmer, M. S., Berta, P., Sinclair, A. H., Pym, B., and Goodfellow, P. N. (1990). Comparison of human *ZFY* and *ZFX* transcripts. *Proc. Natl. Acad. Sci. USA* **87:** 1681–1685.

Pfeifer, G. P., Tanguay, R. L., Steigerwald, S. D., and Riggs, A. D. (1990). In vivo footprint and methylation analysis by PCR-aided genomic sequencing: Comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human *PGK-1*. *Genes Dev.* **4:** 1277–1287.

Pritchard, C. A., Goodfellow, P. J., and Goodfellow, P. N. (1987). Mapping the limits of the human pseudoautosomal region and a candidate sequence for the male-determining gene. *Nature* **328:** 273–275.

Race, R. R., and Sanger, R. (1975). "Blood Groups in Man," Blackwell Scientific, London.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

Schiebel, K., Weiss, B., Wohrle, D., and Rappold, G. (1993). A human pseudoautosomal gene, ADP/ATP translocase, escapes X-inactivation whereas a homologue on Xq is subject to X-inactivation. *Nature Genet.* **3:** 82–87.

Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Mardon, G., Luoh, S. W., and Page, D. C. (1989a). Putative transcription activator with alternative isoforms encoded by human *ZFX* gene. *Nature* **342:** 708–711.

Schneider-Gadicke, A., Beer-Romero, P., Brown, L. G., Nussbaum, R., and Page, D. C. (1989b). *ZFX* has a gene structure similar to *ZFY*, the putative human sex determinant, and escapes X inactivation. *Cell* **57:** 1247–1258.

Shapiro, L. J., Mohandas, T., Weiss, R., and Romeo, G. (1979). Non-inactivation of an X-chromosome locus in man. *Science* **204:** 1224–1226.

Shimmin, L. C., Chang, B. H., and Li, W. H. (1993). Male-driven evolution of DNA sequences. *Nature* **362:** 745–747.

Simpson, E. M., and Page, D. C. (1991). An interstitial deletion in mouse Y chromosomal DNA created a transcribed *Zfy* fusion gene. *Genomics* **11:** 601–608.

Singer-Sam, J., and Riggs, A. D. (1993). X chromosome inactivation and DNA methylation. *In* "DNA Methylation: Molecular Biology and Biological Significance" (J. P. Jost and H. P. Saluz, Eds.), Vol. **64,** pp. 358–384, Birkhäuser Verlag, Basel.

Sirota, L., Zlotogora, Y., Shabtai, F., Halbrecht, I., and Elian, E. (1981). 49,XYYYY: A case report. *Clin. Genet.* **19:** 87–93.

Slim, R., Levilliers, J., Ludecke, H. J., Claussen, U., Nguyen, V. C., Gough, N. M., Horsthemke, B., and Petit, C. (1993). A human pseudoautosomal gene encodes the ANT3 ADP/ATP translocase and escapes X-inactivation. *Genomics* **16:** 26–33.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98:** 503–517.

Su, H., and Lau, Y. F. (1992). Demonstration of a stage-specific expression of the ZFY protein in fetal mouse testis using anti-peptide antibodies. *Mol. Reprod. Dev.* **33:** 252–258.

Tribioli, C., Tamanini, F., Patrosso, C., Milanesi, L., Villa, A., Pergolizzi, R., Maestrini, E., Rivella, S., Bione, S., Mancini, M., *et al.* (1992). Methylation and sequence analysis around *Eag*I sites: Identification of 28 new CpG islands in Xq24–Xq28. *Nucleic Acids Res.* **20:** 727–733.

Wolf, S. F., Jolly, D. J., Lunnen, K. D., Friedman, T., and Migeon, B. R. (1984). Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: Implications for X-chromosome inactivation. *Proc. Natl. Acad. Sci. USA* **81:** 2806–2810.

Zacksenhaus, E., Gill, R. M., Phillips, R. A., and Gallie, B. L. (1993). Molecular cloning and characterization of the mouse *RB1* promoter. *Oncogene* **8:** 2343–2351.

Zambrowicz, B. P., Zimmerman, J. W., Harendza, C. J., Findley, S. D., Simpson, E. M., Page, D. C., Brinster, R. L., and Palmiter, R. D. (1994a). Expression of a *Zfy-1/lacZ* transgene in the somatic cells of the embryonic gonad and germ cells of the adult testis. *Development* **120:** 1549–1559.

Zambrowicz, B. P., Findley, S. D., Simpson, E. M., Page, D. C., and Palmiter, R. D. (1994b). Characterization of the murine *Zfy1* and *Zfy2* promoters. *Genomics* **24:** 406–408.

Zinn, A. R., Bressler, S. L., Beer-Romero, P., Adler, D. A., Chapman, V. M., Page, D. C., and Disteche, C. M. (1991). Inactivation of the *Rps4* gene on the mouse X chromosome. *Genomics* **11:** 1097–1101.

Zwingman, T., Erickson, R. P., Boyer, T., and Ao, A. (1993). Transcription of the sex-determining region genes *Sry* and *Zfy* in the mouse preimplantation embryo. *Proc. Natl. Acad. Sci. USA* **90:** 814–817.