Cell, Vol. 57, 1247–1258, June 30, 1989, Copyright © 1989 by Cell Press

ZFX has a Gene Structure Similar to *ZFY*, the Putative Human Sex Determinant, and Escapes X Inactivation

Ansbert Schneider-Gädicke,* Peggy Beer-Romero,* Laura G. Brown,* Robert Nussbaum,† and David C. Page*
* Whitehead Institute for Biomedical Research
Nine Cambridge Center
Cambridge, Massachusetts 02142 and
Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
† Department of Human Genetics
and Howard Hughes Medical Institute
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Summary

The *ZFX* gene on the human X chromosome is structurally similar to the *ZFY* gene, which may constitute the sex-determining signal on the human Y chromosome. *ZFY* and *ZFX* diverged from a common ancestral gene, as evidenced by similarities in their intron/exon organization and exon DNA sequences. The carboxy-terminal exons of *ZFY* and *ZFX* both encode 13 zinc fingers; 383 of 393 amino acid residues are identical, and there are no insertions or deletions. Thus, the ZFY and ZFX proteins may bind to the same nucleic acid sequences. *ZFY* and *ZFX* are transcribed in a wide variety of XY and (in the case of *ZFX*) XX cell lines. Transcription analysis of human–rodent hybrid cell lines containing "inactive" human X chromosomes suggests that *ZFX* escapes X inactivation. This result contradicts the "dosage/X-inactivation" model, which postulated that sex is determined by the total amount of functionally interchangeable ZFY and ZFX proteins.

Introduction

Analysis of humans with abnormal sex chromosomal constitutions has established the crucial role of the Y chromosome in sex determination. In the presence of the Y chromosome, regardless of the number of X chromosomes, the bipotential gonad develops as a testis and male differentiation ensues (e.g., Jacobs and Strong, 1959). In the absence of the Y chromosome, regardless of the number of X chromosomes, the gonad develops as an ovary and female differentiation ensues (e.g., Ford et al., 1959). These observations imply the existence on the human Y chromosome of one or more sex-determining genes (often referred to as the testis-determining factor, or *TDF*) whose presence or absence determines the fate of the bipotential gonad and, via gonadal hormones (Jost, 1953), all other sexually dimorphic characters.

By deletion analysis of human XX males, XY females, and other "sex-reversed" individuals, we found that sex is determined by the presence or absence of a very small portion of the Y chromosome (Page et al., 1987). The sex-determining function—at least an essential portion of that function—was mapped to a 140 kb portion of the short arm of the Y chromosome. Within this region (interval 1A2), which constitutes about 0.2% of the Y chromosome, we identified a gene encoding a protein with 13 Cys-Cys/His-His "zinc fingers" (Page et al., 1987), a nucleic acid binding motif first described in Xenopus transcription factor IIIA (Miller et al., 1985; Brown et al., 1985). This Y-encoded zinc finger protein, which we will refer to as *ZFY* (Page, 1988), may constitute the primary sex-determining signal. Curiously, on the short arm of the X chromosome, DNA sequences exist which cross-hybridize to the *ZFY* gene. Highly conserved homologs of the human *ZFY* gene are found not only on the Y but also on the X chromosomes of a wide range of placental mammals (Page et al., 1987).

These observations led to the formulation of several models of sex determination (Page et al., 1987), all supposing a central role for *ZFY*, but incorporating the finding of a related locus on the X chromosome. Simplicity made one of these models particularly attractive (Chandra, 1985; German, 1988; Ferguson-Smith, 1988). According to this "dosage/X-inactivation" model, *ZFY* and the related gene on the X chromosome produce functionally interchangeable proteins; XY cells would have two active copies of the gene while, because of X-inactivation, XX cells would have one active copy of the gene. Embryos with two active copies of the gene would develop as males, while embryos with one active copy would develop as females. Insofar as this model postulates gene dosage as the basis of sex determination, it would imply fundamental similarities among humans, Drosophila, and nematodes. (For reviews of sex determination in Drosophila and nematodes, see Baker and Belote, 1983; Cline, 1989; Hodgkin, 1987; and Meyer, 1988).

Here we report results that contradict the dosage/X-inactivation model. We have cloned the human X homolog of the *ZFY* gene. This X-chromosomal gene, which we will refer to as *ZFX*, apparently encodes a protein containing a zinc finger domain closely related to that found in the ZFY protein. The genomic organization of *ZFY* and *ZFX* is remarkably similar. Transcription analysis of human–rodent hybrid cell lines reveals that *ZFX* escapes X inactivation. These results allow us to reassess and restate models postulating roles for both *ZFY* and *ZFX* in gonadal sex determination.

Results

Cloning the *ZFX* Gene by Cross-Hybridization to *ZFY*

Scattered across 50 kb of Y interval 1A2 (Figure 1) are four restriction fragments that contain DNA sequences highly conserved on the Y chromosomes of all placental mammals examined (Page et al., 1987). One of these conserved segments, the human insert of plasmid pDP1007, appears to contain an exon encoding 13 zinc fingers. Pre-

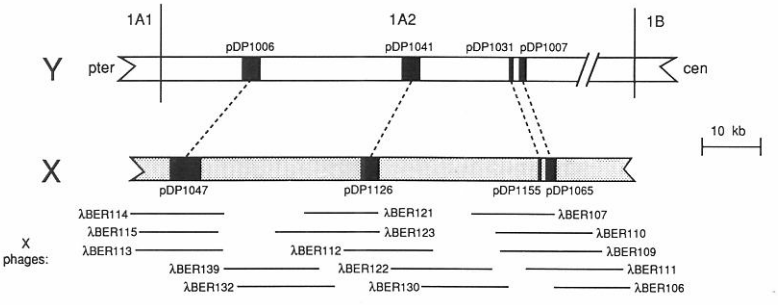

Figure 1. Cloning the Human *ZFX* Gene by Cross-Hybridization to *ZFY*

At top is a schematic representation of interval 1A2 and adjoining intervals on the short arm of the human Y chromosome (based on Page et al., 1987). Interval 1A2, which measures 140 kb, contains at least an essential portion of the sex-determining function. The orientation with respect to the short-arm telomere (pter) and centromere (cen) is shown. Within interval 1A2, four segments (in black) contain DNA sequences that were highly conserved during evolution, apparently contain exons of the *ZFY* gene, and cross-hybridize to sequences on the X chromosome. From left to right, these four segments represent the inserts of plasmids pDP1006, pDP1041, pDP1031, and pDP1007. Shown below is a nearly 90 kb block of DNA cloned from the human X chromosome by cross-hybridization to these four plasmids at high stringency (47°C, 50% formamide, 0.75 M NaCl). Eleven X-derived phages were identified in the initial screen: three with pDP1006, three with pDP1041, four with both pDP1031 and pDP1007, and one with pDP1007 alone. The four other X phages were isolated by chromosomal walking. Dotted lines connect strongly cross-hybridizing restriction fragments (in black) within these *ZFY* and *ZFX* loci. It is likely that each blackened segment contains but does not consist entirely of one or more exons with a high degree of X–Y sequence similarity. Apart from the blackened segments, there may exist additional, smaller regions of X–Y similarity not yet detected. As judged by nucleotide sequencing of pDP1007 (Page et al., 1987) and its X counterpart, pDP1065 (Figure 3), the direction of transcription is from left to right in both *ZFY* and *ZFX*. CpG islands are found in both pDP1006 (Page et al., 1987) and its X counterpart, pDP1047. It is possible, as is the case with many CpG islands (Bird, 1986), that transcription is initiated within these regions.

liminary transcription data suggested that at least two of the other three conserved segments are also exons of this *ZFY* gene (Page et al., 1987).

The vast majority of DNA sequences in interval 1A2 of the Y chromosome do not strongly cross-hybridize to the human X chromosome. However, the four conserved segments of 1A2 do hybridize to sequences highly conserved on the X chromosomes of placental mammals, leading to speculation that the X chromosome carries a gene closely related to *ZFY* (Page et al., 1987). To establish whether such a *ZFX* gene actually exists, we set out to clone the human X-chromosomal counterparts of the conserved segments in interval 1A2. To this end, the conserved segments were used as hybridization probes in screening a library of genomic DNA from a normal human female. Eleven different recombinant phages were identified in the initial screen, and four more phages were isolated by chromosomal walking (Figure 1). The human inserts of all the phages, 15 in total, form a single, overlapping cluster spanning almost 90 kb. That is, all 15 phages derive from a single locus, demonstrating that the counterparts of the various conserved portions of Y interval 1A2 occur in close proximity on the X chromosome.

To better characterize the DNA sequence similarity between *ZFY* and the related X locus, the four conserved segments from Y interval 1A2 were hybridized to restriction digests of phage DNAs spanning the X locus. Each of the four segments from 1A2 cross-hybridized to a single, small segment of the X locus. This one-to-one correspondence is summarized in Figure 1, where it can be seen that the X–Y counterparts occur in the same order and with roughly the same spacing on the two chromosomes. There is a modest expansion of the locus on the X as compared with that on the Y: while the four conserved segments on the Y span about 50 kb, their counterparts on the X span about 70 kb. The most distal of the four conserved segments on the Y, the human insert of plasmid pDP1006, contains a "CpG island" (Page et al., 1987), with a high GC content and an abundance of CpG dinucleotides (Bird, 1986). As judged by a concentration of sites recognized by the restriction enzymes BssHII, EagI, and SacII (Lindsay and Bird, 1987), the X counterpart of pDP1006 also contains a CpG island. Apart from the four X–Y-homologous segments indicated in Figure 1, there is little cross-hybridization between Y interval 1A2 and the related X locus. Even when genomic DNA clones from the X and Y loci are directly compared, e.g., by two-dimensional DNA blot (Southern cross) hybridization (Potter and Dressler, 1986), we find little evidence of additional X–Y sequence similarity (data not shown). Of course, there may exist additional, small regions of X–Y sequence similarity not detected by these techniques. Because of the remarkably extensive, albeit discontinuous, sequence similarity between *ZFY* and the related locus on the X chromosome—and further evidence of homology, as described below—we will refer to the X locus as *ZFX*.

Transcription of the *ZFY* and *ZFX* Genes

By Northern analysis, we examined human cells cultured from a variety of tissues, both male and female, for transcripts from *ZFY* and *ZFX*. The male cells examined included primary culture fibroblasts as well as B-lymphoblastoid, T cell leukemia, and neuroblastoma cell lines. The female cells tested included primary culture fibroblasts as well as B-lymphoblastoid, retinoblastoma, cervical carcinoma, and osteosarcoma cell lines. On Southern blots, the human genomic insert of plasmid pDP1007 (the zinc finger domain of *ZFY*) cross-hybridizes to *ZFX*, even under stringent conditions. Therefore, on Northern blots, pDP1007 might be expected to hybridize to transcripts from either *ZFY* or *ZFX*. In fact, pDP1007 hybridized to transcripts of 6.3 and 8 kb in all cell lines tested, both male and female (Figure 2A, upper panel). In addition, a 5.5 kb transcript was detected in all cell lines containing the Y chromosome. These observations suggested that the 6.3 and 8 kb transcripts derive from *ZFX*; the 5.5 kb transcript either derives from *ZFY* or from *ZFX* and is regulated by the Y chromosome.

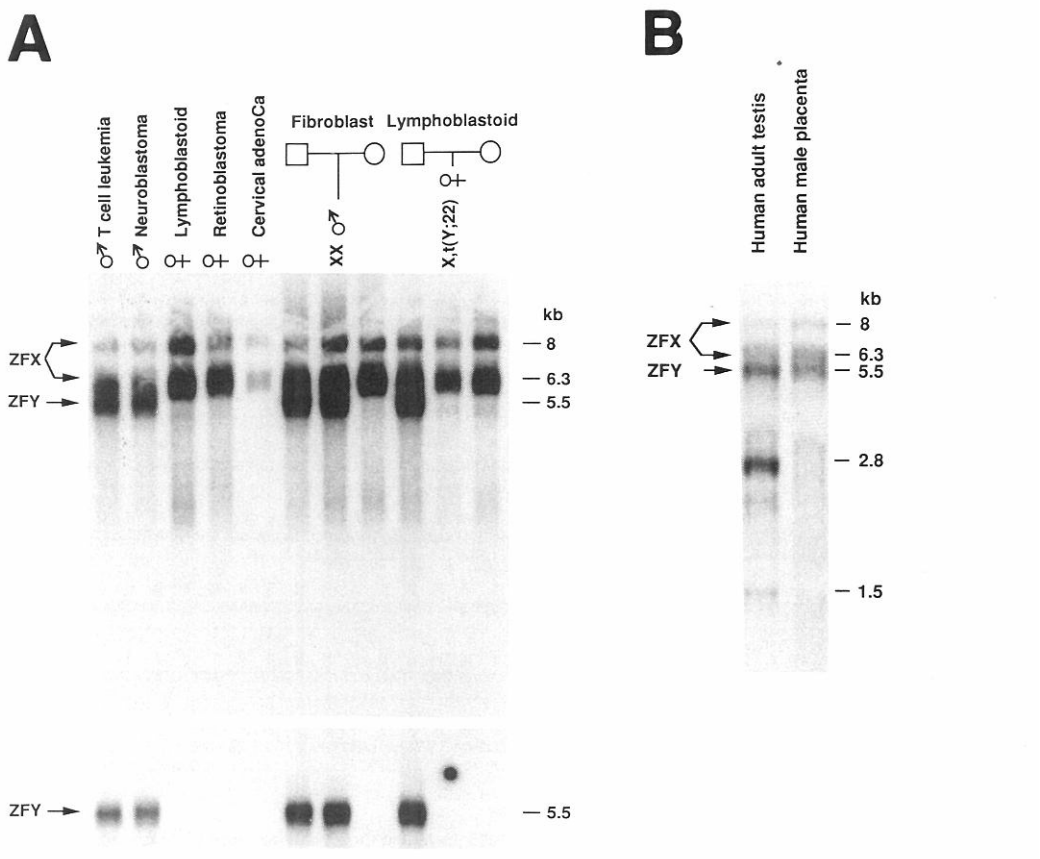

Figure 2. The *ZFX* and *ZFY* Genes are Widely Transcribed

(A) A Northern blot of poly(A)$^+$ RNAs prepared from human cell lines was hybridized sequentially with *ZFY* genomic probes containing coding sequences (pDP1007, above) and 3′ untranslated sequences (pDP1100, below). While pDP1007 cross-hybridizes to *ZFX*, pDP1100 does not. The 5.5 kb *ZFY* transcript, detected by both probes, is present in all cell lines containing an intact Y chromosome as well as in the XX male, who carries only intervals 1A1 and 1A2 of the Y chromosome. This *ZFY* transcript is absent in the X,t(Y;22) female, who carries all but intervals 1A2 and 1B of the Y chromosome. The 6.3 and 8 kb *ZFX* transcripts, detected by pDP1007 but not by pDP1100, are present in all cell lines. Longer exposures reveal additional transcripts of much lower abundance. Lanes contain RNAs from the following cell lines (left to right): Jurkat, NGP, WHT1660, Weri, HeLa, LGL201 (father of LGL03), LGL203 (XX male; case 6 in de la Chapelle et al., 1984; Page et al., 1987), LGL202 (mother of LGL203), WHT1011 (father of WHT1013), WHT1013 (X,t (Y;22) female; Page et al., 1987), and WHT1012 (mother of WHT1013). With the exception of the HeLa lane, which was underloaded, all lanes contain about 5 μg of poly(A)$^+$ RNA. Similar results were obtained with a female osteosarcoma cell line (data not shown). The presence or absence of the Y chromosome was assessed by hybridization of the Y-specific repetitive sequence *DYZ4* (probe pDP105; (D. C. P., unpublished data) to Southern blots of cell line genomic DNAs.

(B) A Northern blot of cellular RNAs (20 μg/lane) prepared from human adult testis and male placenta was hybridized with pDP1007. In addition to the 5.5 kb *ZFY* and 6.3 and 8 kb *ZFX* transcripts observed in human cell lines (A), other smaller transcripts, including an abundant 2.8 kb transcript in testis, are seen in the tissues.

To assign these 5.5, 6.3, and 8 kb transcripts of *ZFY* or *ZFX* with greater certainty, we sought X- or Y-specific probes that would detect one or more of the transcripts. Plasmid pDP1100, for example, contains sequences located immediately 3′ of pDP1007 on the Y chromosome. Unlike pDP1007, pDP1100 hybridizes only to the Y chromosome on genomic Southern blots (not shown). On Northern blots, pDP1100 hybridizes only to the 5.5 kb transcript (Figure 2A, lower panel), demonstrating that this transcript derives from *ZFY*. Conversely, plasmid pDP1112 contains sequences located just 3′ of the pDP1007-homologous sequences in *ZFX*, and, on genomic Southern blots, it hybridizes only to the X chromosome. On Northern blots, pDP1112 hybridizes to the 6.3 and 8 kb (but not the 5.5 kb) transcripts (data not shown), confirming their *ZFX* origin. (Results described below suggest that pDP1100 and pDP1112 contain 3′ untranslated sequences from, respectively, *ZFY* and *ZFX*).

Apart from the zinc finger domain itself (plasmid pDP-1007), do the other conserved sequences within Y interval 1A2 and the corresponding sequences on the X (Figure 1) constitute exons of the *ZFY* and *ZFX* genes? To address this question, we hybridized Northern blots of RNAs from cultured cells with the other conserved DNA fragments from the Y chromosome. Probes pDP1031 and pDP1041 hybridized strongly (data not shown) with the same 5.5 kb *ZFY* and 6.3 and 8 kb *ZFX* transcripts detected using pDP-1007 (Figure 2A). We conclude that these three conserved segments (pDP1007, pDP1031, and pDP1041) and the homologous portions of the X chromosome (Figure 1) do contain exons of similar *ZFY* and *ZFX* transcription units. We have not detected hybridization of pDP1006, the most dis-

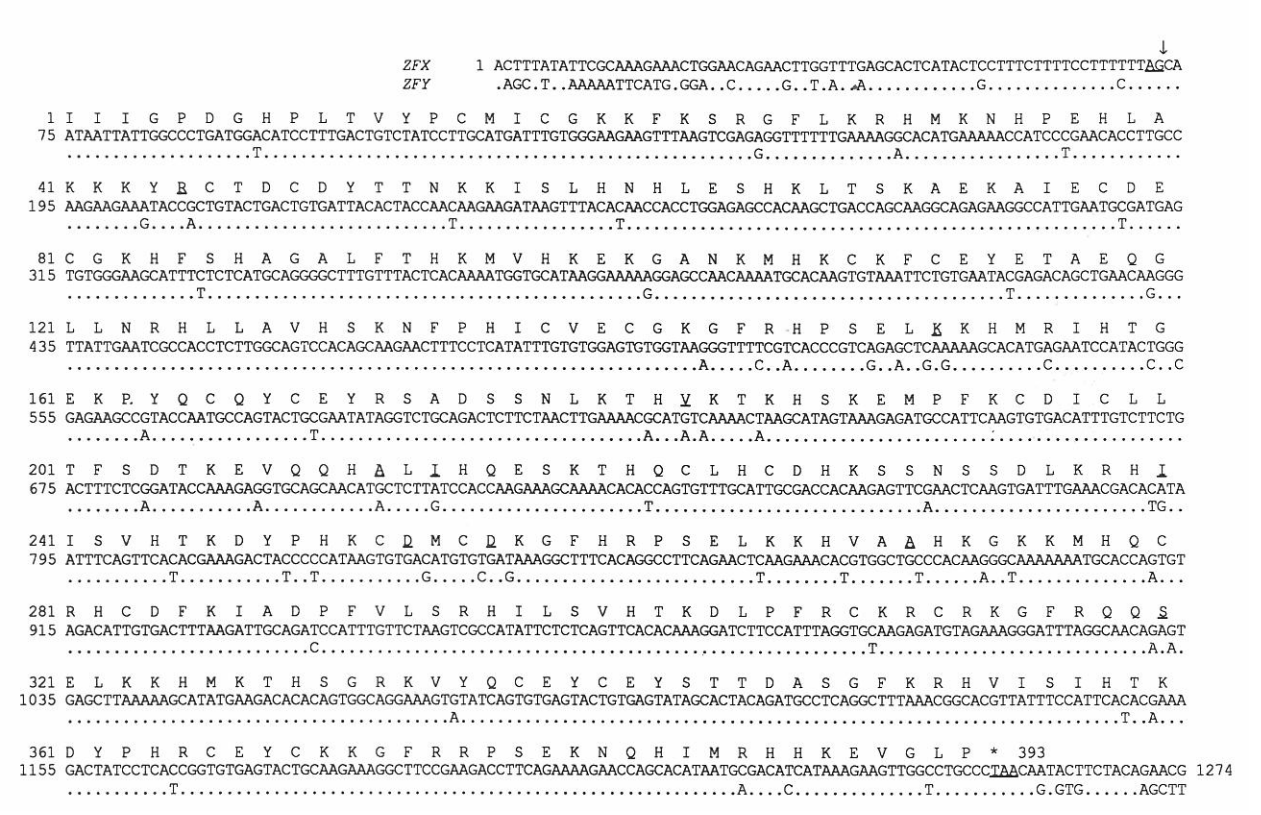

Figure 3. Nucleotide Sequences of the Zinc Finger Exons of the *ZFX* and *ZFY* Genes

The nucleotide sequence of a 1274 bp portion of the genomic insert of plasmid pDP1065, including the zinc finger exon of *ZFX*, is shown. Below this *ZFX* sequence is the nucleotide sequence of the corresponding portion of *ZFY* (i.e., the genomic insert of plasmid pDP1007; Page et al., 1987); dots indicate identical nucleotides. The underlined TAA stop codon at nucleotides 1254–1256, also present in cDNAs from the two genes (A. S. G. and co-workers, unpublished data), closes a long open reading frame present in both *ZFX* and *ZFY* . Analysis of cDNAs also reveals that, in both *ZFX* and *ZFY*, a splice (indicated by arrow) is accepted immediately 3′ to the underlined AG at nucleotides 71–72. Above the *ZFX* nucleotide sequence is the 393 amino acid sequence predicted for this portion of the ZFX protein. Ten residues at which the zinc finger domains of the ZFX and ZFY proteins differ are underlined in the ZFX amino acid sequence. (In addition to the 1274 bp shown here, the insert of pDP1065 contains several hundred bp of additional 5′ and 3′ sequences that show no similarity to pDP1007.)

tal of the conserved, X-homologous segments in interval 1A2, to these *ZFY* and *ZFX* transcripts on Northern blots. Nonetheless, some *ZFY* and *ZFX* cDNA clones (A. S. G. and co-workers, unpublished data) hybridize to pDP1006 and its X counterpart, indicating that pDP1006 and its X homolog do form part of the *ZFY* and *ZFX* transcription units.

We also carried out Northern analysis of cell lines from two sex-reversed individuals instrumental in defining the sex-determining region of the human Y chromosome: XX male LGL203, who carries only intervals 1A1 and 1A2 of the Y chromosome, and X,t(Y;22) female WHT1013, who carries all but intervals 1A2 and 1B of the Y chromosome (Page et al., 1987). The XX male possesses all three transcripts, including the 5.5 kb *ZFY* transcript (Figure 2A). Conversely, the X,t(Y;22) female carries the 6.3 and 8 kb *ZFX* transcripts but lacks the 5.5 kb transcript. These results demonstrate that, as expected, the 5.5 kb *ZFY* transcript is dependent on the presence of interval 1A2 of the human Y chromosome, and that the *ZFY* gene, at least as judged by the presence of the 5.5 kb transcript, is entirely contained within intervals 1A1 and 1A2.

Our Northern analysis of cells in culture (Figure 2A) suggested that *ZFY* and *ZFX* are transcribed in a wide range of human tissues. To ascertain whether the genes are transcribed not only in cultured cells but also in vivo, we extended the Northern analysis of RNAs prepared from human male placenta and adult testis. Both tissues show the same 5.5 kb *ZFY* and 6.3 and 8 kb *ZFX* transcripts found in cell lines (Figure 2B). In addition, several smaller transcripts not observed in the cultured cells are present in the tissues; a 2.8 kb transcript is particularly abundant in adult testis. As we have not detected hybridization of gene-specific probes to these shorter transcripts, it remains to be determined whether they derive from *ZFY* or from *ZFX*. In summary, *ZFY* and *ZFX* are both transcribed in primary culture fibroblasts, in transformed cells, and in tissues. The distribution of the various transcripts as a function of tissue and developmental stage remains to be described in detail.

Similar Zinc Finger Domains in the ZFX and ZFY Proteins

The human insert of plasmid pDP1007, the most centromere-proximal of the four conserved segments identified in Y interval 1A2 (Figure 1), contains a long open reading

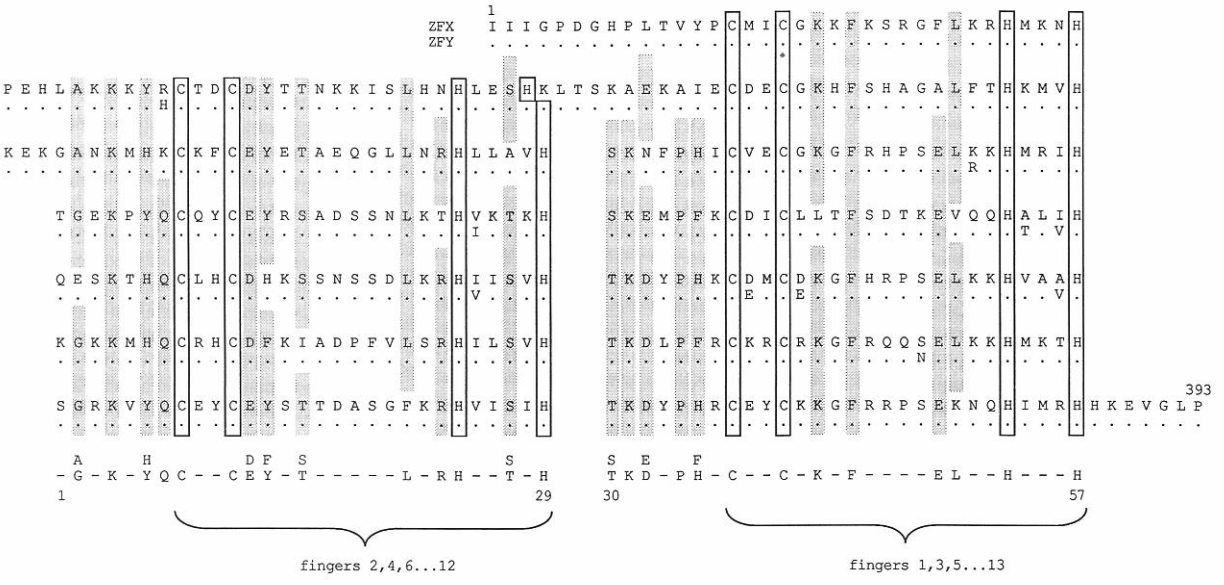

Figure 4. The Zinc Finger Domains of the Putative ZFX and ZFY Proteins Are Very Similar; Both Contain a Two-Finger Repeat

The predicted 393 amino acid sequence of the carboxy-terminal, zinc finger domain of ZFX is given. Below ZFX is the sequence of the homologous portion of ZFY (Page et al., 1987); identical residues are indicated by dots. (For comparison with mouse Zfy-2, see Mardon and Page, 1989). The sequences are arranged so as to highlight a 57 residue repeat, the consensus for which is shown at the bottom. Each repeat contains an "even-numbered" finger (2, 4, 6, . . . 12; on the left) and an "odd-numbered" finger (1, 3, 5, . . . 13; on the right). Boxes are drawn around the invariant cysteins and histidines characteristic of zinc finger domains, as in Xenopus transcription factor IIIA (Miller et al., 1985; Brown et al., 1985). The shaded residues are highly conserved among the repeats.

frame encoding 13 zinc fingers (Page et al., 1987). *ZFY* plasmid pDP1007 strongly hybridizes to a 1.75 kb EcoRI fragment in *ZFX* (see Figure 3 in Page et al., 1987). We found that the nucleotide sequence of this *ZFX* genomic fragment is very similar to that of pDP1007 (Figure 3). Both sequences contain a long open reading frame. Nucleotide sequence analysis of *ZFX* and *ZFY* cDNAs (A. S. G. and co-workers, unpublished data) revealed that, in both genes, a splice is accepted immediately 3′ to the AG at nucleotides 71–72. Furthermore, the in-frame TAA stop codon at nucleotides 1254–1256 is present in cDNAs from both genes. We conclude that these genomic fragments contain exons encoding carboxy-terminal domains of the ZFX and ZFY proteins. Within the putative coding region (nucleotides 73–1253), the *ZFX* and *ZFY* nucleotide sequences are 95% identical, with no deletions or insertions. This nucleotide similarity wanes in the 5′ intron and 3′ to the stop codon.

Based on the predicted translation of these nucleotide sequences, we conclude that the ZFX and ZFY proteins have very similar carboxy-terminal domains (Figure 4). In this 393 amino acid domain, the ZFX and ZFY proteins differ by only ten substitutions, nine of which are conservative, and there are no insertions, deletions, or rearrangements. High stringency cross-hybridization of the more 5′ exons of *ZFY* and *ZFX* (Figure 1) suggests that the amino-terminal portions of the proteins may also be similar.

As in *ZFY*, the carboxy-terminal exon of *ZFX* appears to encode a tandem array of 13 zinc fingers, each with two cysteines and two histidines (Figure 4). The zinc finger domain of the ZFY protein is characterized by a secondary repeat, comprising a pair of fingers (Page et al., 1987). Apart from the mouse homolog of ZFY (Mardon and Page, 1989), we know of no other protein with multiple Cys-Cys/His-His fingers in which such a two-finger repeat is present. Nonetheless, the same structure is readily apparent in ZFX (Figure 4). The zinc finger domain in ZFX, like that of ZFY, consists of six and one-half tandem repeats of a 57 or 63 residue unit. As shown in Figure 4, this repeat unit consists of two fingers, each 21 or 22 residues long, and two "linkers," each seven or more residues long.

Transcription of *ZFX* Is Not Subject to X Inactivation

An X-linked gene can be said to be dosage compensated if its level of expression is equal in cells with one X chromosome and cells with two X chromosomes. In mammalian somatic cells, dosage compensation is achieved by transcriptional inactivation of genes on all but one X chromosome (Lyon, 1988). Transcription of *ZFX* in B-lymphoblastoid cells and fibroblasts enables us to assay the gene for X inactivation directly by Northern blot analysis. If *ZFX* were subject to X inactivation, its level of transcription should be independent of the number of X chromosomes. However, Northern blot analysis of human cell lines with one or more X chromosomes suggests that this is not the case. We find that the level of *ZFX* transcription increases with the number of X chromosomes (Figures 5A and 5B; Table 1), suggesting that *ZFX* escapes X inactivation. As a control, the blot was rehybridized with *HPRT*, an X-linked gene previously demonstrated to be dosage compensated (Graves and Gartler, 1986). Though the level of

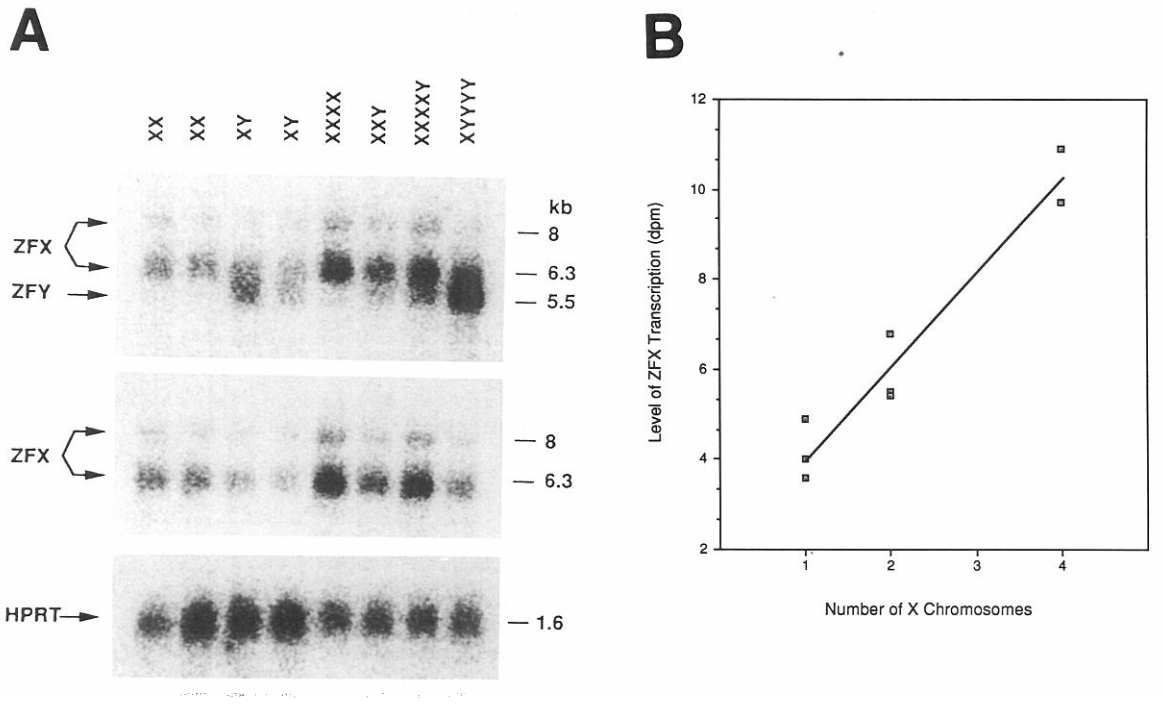

Figure 5. The Level of *ZFX* Transcription Increases with the Number of X Chromosomes

(A) Northern blot analysis of poly(A)$^+$ RNAs from human cell lines with differing numbers of X and Y chromosomes. Shown here are Polaroid photographs of β-emission images obtained using a Betascope Blot Analyzer. In these images, which otherwise closely resemble the images produced by conventional X-ray autoradiography, the β-emission/darkness ratio is linear. The lanes contain RNAs from the following B-lymphoblastoid cell lines (left to right): two normal females, two normal males, GM1416 (XXXX female), WHT1002 (XXY Klinefelter male), WHT0706 (XXXXY male), and OXEN (XYYYY male; Sirota et al., 1981). The blot was hybridized sequentially with the following probes (from top to bottom):

(1) pDP1007, which hybridizes to transcripts from both *ZFX* and *ZFY*.

(2) A 1.3 kb EcoRI–HaeIII fragment purified from pDP1112, a 3′ untranslated *ZFX* probe that does not cross-hybridize to *ZFY* transcripts.

(3) pHPT31 (Brennand et al., 1983), an *HPRT* cDNA.

(B) Level of *ZFX* transcription (in dpm; data from Table 1) plotted versus number of X chromosomes per cell. As explained in Table 1, these data derive from the Northern blot shown in (A).

HPRT transcription is somewhat variable from cell line to cell line, we found, as expected, that the level of *HPRT* transcription does not correlate with the number of X chromosomes (Figure 5A; Table 1). Finally, we note that the relative abundance of *ZFX* and *ZFY* transcripts mirrors the ratio of X to Y chromosomes (Figure 5A, Table 1). These results (corroborated on several other Northern blots; not shown) suggest that transcription of the human *ZFX* gene might not be dosage compensated.

To assess directly whether *ZFX* is subject to X inactivation, we tried to determine whether *ZFX* is transcribed in human–rodent hybrid cell lines containing "inactive" human X chromosomes (Figure 6). In control human XX cell lines, hybridization with pDP1007 reveals *ZFX* transcripts of 6.3 and 8 kb. In control hamster and mouse cell lines, cross-hybridization to transcripts of 6 and 7.5 kb (from rodent *Zfx*; G. Mardon, S. -W. Luoh, and D. C. Page, unpublished data) is observed. In a human–hamster hybrid cell line retaining two human X chromosomes, one "active" and one "inactive," both human *ZFX* and rodent *Zfx* are transcribed. More importantly, in each of three hybrid cell lines retaining only an "inactive" human X chromosome, both human *ZFX* and rodent *Zfx* are transcribed. It should be noted that each of these three hybrids contains a different and independent "inactive" human X chromosome (see legend to Figure 6). These results suggest that the

Table 1. Quantitation of *ZFX*, *ZFY*, and *HPRT* Transcripts in Northern Blot of Figure 5A

	XX	XX	XY	XY	XXXX	XXY	XXXXY	XYYYY
ZFX (dpm)	5.5	5.4	4.0	3.6	10.9	6.8	9.7	4.9
ZFY (dpm)	–	–	2.3	1.3	–	1.3	2.2	6.4
HPRT (dpm)	13.8	26.5	23.1	26.6	15.9	15.5	15.7	15.1

β emissions from the *ZFX*, *ZFY*, and *HPRT* bands in each of the eight lanes were directly quantitated using the Betascope Blot Analyzer. (*ZFX* data derive from hybridization with the 1.3 kb EcoRI–HaeIII fragment from pDP1112). β emissions were counted for 15 hr (*ZFX* and *ZFY*) or 3 hr (*HPRT*). The numbers shown (in disintegrations per min; dpm) are corrected for filter background. The *ZFX* data are also plotted in Figure 5B.

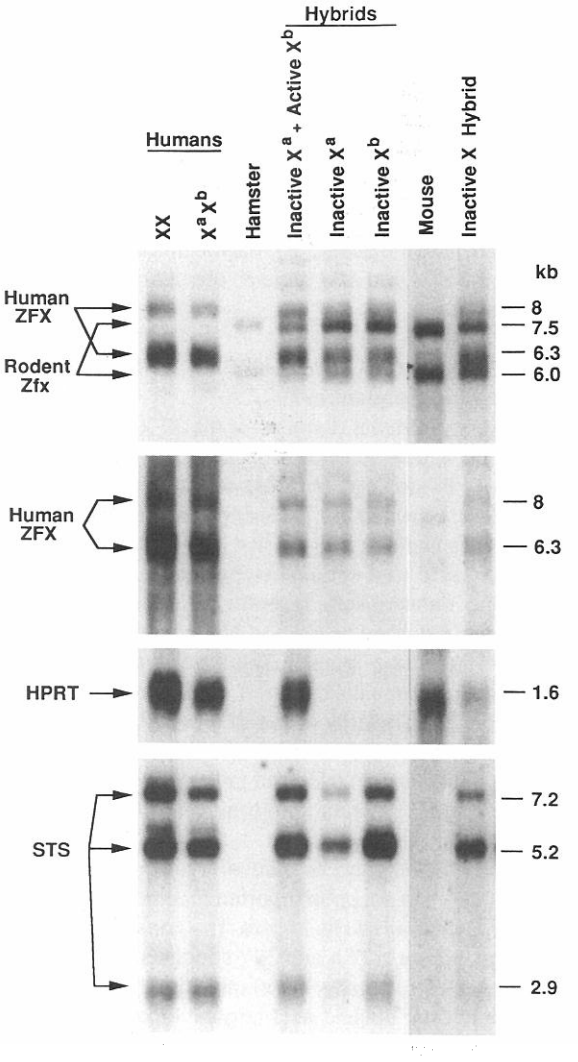

Figure 6. *ZFX* Is Transcribed from "Inactive" X Chromosomes

Northern blot analysis of poly(A)$^+$ RNAs from human–rodent hybrid cell lines retaining "inactive" human X chromosomes. The lanes contain RNAs from the following cell lines (left to right): WHT1628 (human female B-lymphoblastoid control), VTL142, PHL4, PHL21, PHL5, PHL18, A9, and 37-26R-D. PHL21, PHL5, and PHL18 are hybrids derived from fusion of VTL142 B-lymphoblastoid cells, from a human XX female, with PHL4 fibroblasts, from Chinese hamster (Ledbetter et al., 1986). The *Hprt* gene is deleted in PHL4. The two X chromosomes in VTL142 can be distinguished by restriction fragment length polymorphisms and are here referred to as X^a and X^b. Hybrid PHL21 retains both human X chromosomes: an "inactive" X^a and an "active" X^b. PHL5, derived from PHL21 by treatment with 6-thioguanine, has lost the "active" X^b but retains the "inactive" X^a. PHL18 derives from a hybrid in which X^a was "active" and X^b "inactive"; PHL18 has lost the "active" X^a but retains the "inactive" X^b. (Retained human X chromosomes were judged to be "active" or "inactive" on the basis of assays for the X-linked enzymes G6PD, PGK, and HPRT.) Hybrid 37-26R-D (Mohandas et al., 1981) derives from fusion of XX human female fibroblasts with mouse line A9. Hybrid 37-26R-D retains an "inactive" human X chromosome. The blot was hybridized sequentially with the following probes (from top to bottom):

(1) pDP1007, which hybridizes to human *ZFX* and its rodent homolog. The 6.3 and 8 kb transcripts, derived from human *ZFX*, are present in three hybrid cell lines that lack "active" human X chromosomes but retain "inactive" human X chromosomes.

(2) pDP1112, a 3′ untranslated human *ZFX* probe which, under the stringent conditions used, displays no significant hybridization to rodent transcripts. Hybridization to the 6.3 and 8 kb transcripts present in the hybrids verifies that these transcripts derive from human *ZFX*.

(3) pHPT31 (Brennand et al., 1983), a human *HPRT* cDNA. The absence of *HPRT* transcripts in the two human–hamster hybrids retaining "inactive" human X chromosomes (and the low level of transcription, presumably from the mouse gene, in the human–mouse hybrid retaining an "inactive" X chromosome) confirms the "inactive" state of the human X chromosomes.

(4) pSTS76 (Yen et al., 1987), a human *STS* cDNA that does not hybridize to rodent transcripts. Like *ZFX*, *STS* is transcribed from the "inactive" X chromosomes.

human *ZFX* gene is transcribed regardless of whether it is borne by an "active" or "inactive" X chromosome. This conclusion is supported by the following controls (Figure 6):

—To demonstrate unequivocally that the 6.3 and 8 kb transcripts present in the hybrids are of human and not rodent origin, the Northern blot was rehybridized with a *ZFX* 3′ untranslated DNA fragment under conditions that produce little cross-hybridization to rodent *Zfx*.

—To verify the "active" or "inactive" status of the human X chromosomes in the hybrids, we looked for transcription of human *HPRT*. As expected, human *HPRT* transcripts are not detected in either of the two human–hamster hybrids retaining "inactive" human X chromosomes. (These hybrids derive from a fusion with hamster cells deleted for the homologous *Hprt* gene.) The weak signal seen in the human–mouse hybrid probably derives from the mouse gene, which is transcribed in the mouse cell line used to generate the hybrid.

—In contrast to *HPRT*, the human steroid sulfatase (*STS*) gene escapes X inactivation (Shapiro et al., 1979; Mohandas et al., 1980; Migeon et al., 1982). As expected, human *STS* transcripts are detected not only in human cells and hybrids retaining "active" human X chromosomes, but also in hybrids retaining only "inactive" human X chromosomes. No cross-hybridization to rodent *Sts* is detected.

ZFY and *ZFX* Transcription in XY Females, XX Males, and XX Hermaphrodites

The sex-determining region of the Y chromosome, including *ZFY*, is deleted in some XY females, but it is grossly intact, as judged by Southern blotting, in many others (Page et al., 1987). Such "unexplained" sex reversal might nonetheless be due to mutations in *ZFY*—or perhaps even in *ZFX*. In search of such mutations, we examined, by Northern analysis, transcription of the two genes in fibroblasts or B-lymphoblastoid cells cultured from nine such XY females. In all nine we detected (data not shown) the same 5.5 kb *ZFY* and 6.3 and 8 kb *ZFX* transcripts seen in normal XY male controls (Figures 2 and 5A).

Conversely, while the vast majority of XX males carry the sex-determining region of the Y chromosome, including *ZFY*, a few XX males and all XX hermaphrodites tested do not (Page et al., 1987; Ramsay et al., 1988). Again, searching for causes of such "unexplained" sex reversal, we looked for altered *ZFX* transcription in fibroblasts or B-lymphoblastoid cells from three XX hermaphrodites and two XX males in whom the *ZFY* gene is absent. In all five we detected (data not shown) the same 6.3 and 8 kb

ZFX transcripts seen in normal XX female controls (Figures 2 and 5).

Thus, by Northern analysis of cultured cells, we did not detect any alteration of *ZFY* or *ZFX* transcription in "unexplained" human XY females, XX males, or XX hermaphrodites. These results are compatible with sex reversal in these cases being due to missense or nonsense mutations in *ZFY* or *ZFX*; mutations affecting the expression of *ZFY* or *ZFX* during development; or mutations in other genes functioning in conjunction with or downstream of *ZFY*.

Discussion

ZFX and *ZFY* Derive from a Common Ancestral Gene and Encode Proteins with Similar Zinc Finger Domains

The human *ZFX* and *ZFY* genes have a similar organization (Figure 1). Each gene contains four evolutionarily conserved DNA segments, the most 5′ of which is a CpG island. The four conserved segments are similarly arranged in the two genes and show a high degree of X–Y nucleotide sequence similarity as evidenced by high stringency cross-hybridization and, in the case of the most 3′ segment, nucleotide sequence analysis (Figure 3). All four segments contain exons of the *ZFX* and *ZFY* transcription units. Between these evolutionarily conserved, X–Y-homologous portions of *ZFX* and *ZFY* are sequences that lack these properties and which constitute, at least in large part, introns. The regions encoding the carboxy-terminal domains of ZFX and ZFY are bounded by clearly homologous splice acceptor sites and stop codons, are of precisely the same length, and have 95% identical nucleotide sequences (Figure 3). In all respects, the remarkably similar structure and sequence of the *ZFX* and *ZFY* genes strongly support the conjecture (Page et al., 1987) that they diverged from a common ancestral gene prior to the radiation of placental mammals at least 60 to 80 million years ago. This ancestral gene was probably located on both members of a pair of autosomes—an autosomal pair from which evolved the X and Y chromosomes of placental mammals (Page et al., 1987).

If the *ZFX* and *ZFY* genes began to diverge so long ago, why are their nucleotide sequences so similar? *ZFX/ZFY* similarity is not maintained by conventional crossing over between the X and Y chromosomes, as evidenced by the fact that nucleotide similarity is limited to the exons; the introns of *ZFX* and *ZFY* do not cross-hybridize. Analysis of the zinc finger domains suggests that functional constraints on the proteins are the preeminent force in maintaining this similarity: of 60 nucleotide substitutions that distinguish the zinc finger domains of *ZFX* and *ZFY*, 50 are silent. Of the ten resulting amino acid substitutions, nine are conservative. These analyses argue that the products of the *ZFX* and *ZFY* genes face a common evolutionary selective pressure that indirectly serves to maintain DNA sequence similarity. Gene conversion may play a supplementary role in the maintenance of *ZFX/ZFY* sequence similarity. Gene conversion could account, for example, for the extension of X–Y sequence similarity about 50 nucleotides into the intron 5′ of the zinc finger exon (Figure 3).

In *ZFX* and *ZFY*, not only the carboxy-terminal, zinc finger exons but also the more 5′ exons show a high degree of X–Y cross-hybridization and, therefore, sequence similarity. This suggests that the more amino-terminal portions of the proteins might also be similar. However, we cannot exclude the possibility that the amino-terminal portions of the ZFX and ZFY proteins are structurally or functionally distinct.

Do the ZFX and ZFY Proteins Bind to the Same Nucleic Acid Sequences?

The carboxy-terminal domains of the ZFX and ZFY proteins each appear to contain 13 Cys-Cys/His-His fingers (Figure 4). By analogy of proteins such as TFIIIA (Miller et al., 1985; Brown et al., 1985) and Sp1 (Kadonaga et al., 1987), these findings suggest that both ZFX and ZFY bind to nucleic acid in a sequence-specific fashion, perhaps regulating transcription. In both ZFX and ZFY, the 13 fingers are arranged as six and one-half tandem repeats of a two-finger unit (Figure 4). This repetitive structure may have arisen by multiplication of a primordial two-finger unit. If so, the similarity of the ZFX and ZFY zinc finger domains—and of the exons that encode them—argues that the multiplication occurred prior to the divergence of the *ZFX* and *ZFY* genes from their common ancestor.

The two-finger repeat, first described in human *ZFY* (Page et al., 1987) and its mouse homolog (Mardon and Page, 1989), has not been reported for other proteins with multiple Cys-Cys/His-His fingers. The conservation of the two-finger repeat in ZFX and ZFY suggests that the repeat itself may be structurally important in the binding of these proteins to their nucleic acid targets. Indeed, the occurrence of the two-finger repeat agrees well with a proposed model of the interaction of multiple finger proteins with DNA (model II, Fairall et al., 1986). According to this model, the protein lies on one face of the DNA double helix, with every other finger making structurally equivalent contacts. However, a two-finger repeating structure is not incompatible with other proposed models (model I, Fairall et al., 1986; Berg, 1988; Gibson et al., 1988). Further studies will be required to clarify the structure of the putative protein–DNA complex.

The remarkable similarity of amino acid sequence in the zinc finger domains of ZFX and ZFY (Figure 4) suggests that the two proteins may bind to the same nucleic acid sequences. ZFX and *ZFY* may regulate transcription of the same gene(s). However, we cannot exclude the possibility that the few amino acid substitutions by which the zinc finger domains of ZFX and ZFY differ do affect binding specificity or affinity. In any case, it seems likely that interactions with similar or identical nucleic acid binding sites have dramatically constrained the divergence of these ZFX and ZFY protein domains during evolution.

ZFX Escapes X Inactivation

In a variety of organisms, the level of transcription of

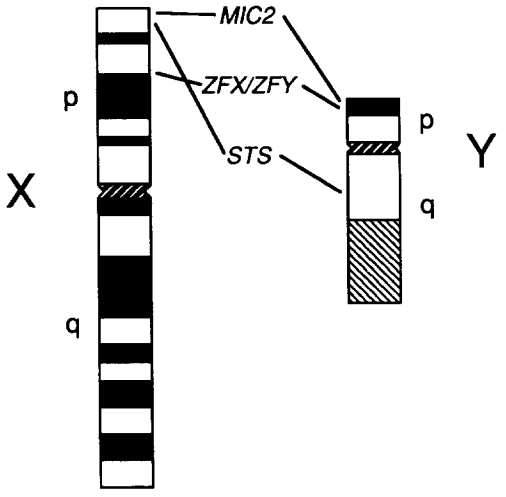

Figure 7. Three Cloned Human Genes Known to Escape X Inactivation Have Homologs on the Y Chromosome

MIC2 maps to the pseudoautosomal region (Goodfellow et al., 1986) of Xp22.3 and Yp (Buckle et al., 1985) and escapes X inactivation (Goodfellow et al., 1984). The *STS* gene is located in Xp22.3 (Yen et al., 1987) and escapes X inactivation (Shapiro et al., 1979; Mohandas et al., 1980; Migeon et al., 1982); a pseudogene is found on Yq (Yen et al., 1988). (Mouse *Sts*, presumably the homolog of human *STS*, is functional on both X and Y [Keitges et al., 1985] and escapes X inactivation [Keitges and Gartler, 1986].) *ZFX* and *ZFY* map to Xp21.3/22.1 and Yp, respectively (Page et al., 1987; D. C. Page, A. de la Chapelle, U. Francke, and P. Green, unpublished data). Though not cloned, the human *XG* gene also escapes X inactivation (Race and Sanger, 1975; Fialkow, 1978), and there may be a related gene on the Y chromosome (Goodfellow and Tippett, 1981).

X-linked genes is roughly equal in males (with one X chromosome) and females (with two X chromosomes). Such dosage compensation can be achieved by various mechanisms, including turning up expression of genes on the single X chromosome in males, as in Drosophila (reviewed by Jaffe and Laird, 1986). In mammals, dosage compensation is accomplished by transcriptional inactivation of all but one X chromosome (Lyon, 1988). In the case of mammalian X-linked genes whose expression is limited to certain cell types, indirect assays (e.g., methylation, DNAase I sensitivity, etc.) are often the only means of examining inactivation. Because of the widespread transcription of *ZFX*—particularly in B-lymphoblastoid cells and fibroblasts—we were able to assay its transcription directly by Northern blot analysis. Evidence of two sorts suggests that the *ZFX* gene on the human X chromosome is not dosage compensated. First, in human B-lymphoblastoid cell lines containing one, two, or four X chromosomes, the level of *ZFX* transcription increases with the number of X chromosomes (Figures 5A and 5B; Table 1). Second, human *ZFX* is transcribed in human–rodent hybrids retaining "inactive" human X chromosomes (Figure 6). Thus, *ZFX* is transcribed from "inactive" as well as "active" X chromosomes.

ZFX is not the only human X-chromosomal gene to escape X inactivation. While the large majority of X-chromosomal genes appear to be subject to X inactivation, a few genes mapping to the extreme distal short arm of the X chromosome (band Xp22.3) are known to escape X inactivation (Figure 7). These observations helped foster the notion of a single small, discrete region of the human X chromosome that escapes X inactivation. However, *ZFX* is located more proximally on the short arm of the human X chromosome, in the region of bands Xp21.3 and p22.1 (Page et al., 1987; D. C. Page, A. de la Chapelle, U. Francke, and P. Green, unpublished data). Thus, escape from X-inactivation is not limited to genes in Xp22.3. We conclude that either a single large, contiguous portion of the short arm of the human X chromosome escapes X inactivation, or, more likely, multiple, noncontiguous portions of the X chromosome escape X inactivation. Among human X-chromosomal genes of which we are aware, there is a striking correlation between escape from X inactivation and the existence of homologs on the Y chromosome. That is, each of the three cloned X-chromosomal genes known to escape X inactivation has a homolog—either a gene or a pseudogene—on the Y (Figure 7). None of the many genes subject to X inactivation is known to have a Y homolog.

Sex Determination and the Functional Relationship of *ZFY* and *ZFX*

Experiments with XX–XY chimeric mice indicate that the sex determining function of the Y chromosome is required only in certain cells of the embryonic gonad (Burgoyne et al., 1988). If *ZFY* and *ZFX* function only in gonadal sex determination, then one might expect to find them expressed only in the embryonic gonad. However, we find that *ZFY* and *ZFX* are transcribed in cultured cells originating from a variety of human tissues (Figure 2A). Analysis of adult testis and male placenta (Figure 2B) verifies that transcription of *ZFY* and *ZFX* is not limited to cultured cells. The wide transcription of *ZFX* and *ZFY* does not argue against a role in gonadal sex determination; several genes involved in sex determination in Drosophila or C. elegans are expressed in tissues in which no sexual dimorphism is apparent (for reviews, see Cline, 1989, and Meyer, 1988). The widespread transcription of human *ZFX* and *ZFY* may indicate that the genes are not tightly regulated despite having quite restricted functions, or have functions in addition to gonadal sex determination.

The situation is somewhat different in mice. Transcription of the mouse *Zfy* genes has been detected only in adult testis, where there is an abundant 3 kb transcript (Mardon and Page, 1989), suggesting a possible function in male reproduction. In this connection, it should be pointed out that there is an abundant 2.8 kb message in human adult testis (Figure 2B), the X or Y origin of which remains to be determined.

What is the functional relationship of *ZFY* and *ZFX*? We previously discussed four models (Page et al., 1987), assuming, for purposes of discussion, that *ZFY* constitutes the primary sex-determining signal. In brief, the first three models postulate that the ZFY and ZFX proteins are functionally distinct and that either *ZFX* does not function in sex determination; or *ZFX* and *ZFY* act antagonistically; or

ZFX and *ZFY* act in concert. No evidence excludes any of these models, but some qualifications are in order. If we suppose that the first model is correct, then we must account for the likelihood that the ZFX and ZFY proteins bind to and regulate the same gene(s). If the second model is correct, and ZFX and ZFY proteins compete, then we must suppose that the competition is not finely balanced: while the relative levels of *ZFX* and *ZFY* transcription are markedly different in, for example, XXXXY and XYYYY (Figure 5A), the bipotential gonad undergoes testicular differentiation in either case.

According to model 4, *ZFX* and *ZFY* are differentially regulated but encode functionally interchangeable proteins. Specifically, it was proposed (see also Chandra, 1985, and German, 1988) that sex is determined by the dosage of active *ZFY* and *ZFX* genes, this dosage depending critically upon X inactivation of *ZFX*. In embryos with two active copies (e.g., *ZFX* + *ZFY* in XY), testicular differentiation would occur, while in embryos with one active copy (e.g., one active *ZFX* in XX), ovarian differentiation would occur. This specific model, however, is contradicted by the finding that *ZFX* escapes X inactivation (assuming the *ZFX* escapes X inactivation not only in the cultured adult cells we examined, but also in somatic cells of the embryo). Of course, X inactivation is only one possible mode of gene regulation. Functionally interchangeable ZFY and ZFX proteins might be differentially expressed during embryonic development, at least in certain tissues or at certain stages, as a result of other regulatory mechanisms, either transcriptional or posttranscriptional. Sex determination might hinge upon such differential regulation of *ZFX* and *ZFY*. Of course, it remains to be determined whether not only the carboxy-terminal, zinc finger domains but also the amino-terminal domains of *ZFY* and *ZFX* are similar.

In conclusion, all four models of sex determination remain viable, although they are refined by the data presented. To be sure, the postulated sex-determining role of *ZFY* remains to be demonstrated. Such issues may be resolved by studies of the expression of *ZFX* and *ZFY* in embryos, by the search for *ZFX* or *ZFY* mutations in sex-reversed individuals, or by analysis of transgenic mice.

Experimental Procedures

Human Chromosomal Walk

Recombinant phage with the prefix λBER have inserts derived from the human X chromosome and were isolated from a partial digest library of female genomic DNA constructed in λ vector EMBL3A (Frischauf et al., 1983). This library was a gift from Stuart Orkin. Overlaps among the λBER clones (Figure 1) were determined by restriction mapping and hybridization. Phage with the prefix λOX have inserts derived from the human Y chromosome and were isolated in an earlier study (Page et al., 1987).

DNA Hybridization Probes

Listed below are plasmids whose inserts were used as hybridization probes in screening genomic libraries or in Northern or Southern blot analysis. All inserts are genomic sequences derived from the human X or Y chromosomes. Stated below are the recombinant phage from which the inserts were subcloned, the sizes and restriction site termini of the inserts, and the plasmid cloning vectors. All restriction site termini indicated are present in the human genome except the SalI sites in pDP1041 and pDP1031, which derive from the phage cloning vector.

	Origin		Insert		
Plasmid	X or Y	Phage	Size (kb)	Enzyme(s)	Vector
pDP1006	Y	λOX107	3.0	HindIII–SalI	pUC13
pDP1041	Y	λOX90	3.0	HindIII–SalI	Bluescript
pDP1031	Y	λOX84	0.6	EcoRI–SalI	Bluescript
pDP1007	Y	λOX82	1.3	HindIII	pUC13
pDP1100	Y	λOX82	0.8	HindIII–XbaI	Bluescript
pDP1065	X	λBER110	1.7	EcoRI	Bluescript
pDP1112	X	λBER106	1.4	EcoRI–PstI	Bluescript

In the human genome, the inserts of plasmids pDP1100 and pDP1112 lie immediately 3′ of the inserts of, respectively, pDP1007 and pDP1065.

Northern Blot Analysis

RNAs were prepared from 2 × 10^8 cultured cells of each type. In each case, the cell pellet was homogenized with a Polytron in 40 ml of solution containing 200 μg/ml proteinase K, 0.5% SDS, 0.1M NaCl, 20 mM Tris–HCl (pH 7.4), 1mM EDTA. The homogenate was incubated for 1 hr at 37°C. After adjusting the NaCl concentration to 0.4 M, we added 0.5 ml of oligo(dT)-cellulose (Collaborative Research) equilibrated in "high salt buffer" (0.4 M NaCl, 0.1% SDS, 10 mM Tris–HCl (pH 7.4), 1mM EDTA). After shaking for 1 hr at room temperature, the oligo(dT)-cellulose was washed four times with 50 ml of high salt buffer and poured into a column. Poly(A)$^+$ RNA was eluted with 4 × 0.5 ml TES (10 mM Tris [pH 7.4], 1 mM EDTA, 0.1% SDS).

RNAs were transferred to GeneScreen Plus membrane (New England Nuclear) after electrophoretic separation in 1% agarose–formaldehyde gels. DNA inserts of plasmids were purified, radiolabeled with ^{32}P by random-primer synthesis (Feinberg and Vogelstein, 1983), and hybridized to Northern blots at concentrations of 3 × 10^6 cpm/ml. Hybridizations were carried out at 47°C in 50% formamide, 5× SSC (1× SSC = 0.15 M NaCl, 15 mM sodium citrate [pH 7.4]), 2× Denhardt's (1× Denhardt's = 0.02% Ficoll 400, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin), and 0.01% yeast tRNA. After washing four times for 20 min each at 73°C in 2× SSC, 0.1% SDS, the filters were exposed at −80°C with X-ray films and intensifying screens. In some cases, β emissions from Northern blots were imaged and quantitated directly using a Betascope 603 Blot Analyzer (Betagen; Sullivan et al., 1987). Between hybridizations with different probes, Northern blots were stripped by incubation for 20 min at 80°C in TES.

Nucleotide Sequence Analysis

Single-stranded DNA templates were generated from Bluescript-based plasmid pDP1065 using the R408 helper phage as recommended by Stratagene. Both strands were sequenced by dideoxynucleotide chain termination (Sanger et al., 1977) using oligonucleotide primers.

Acknowledgments

We thank T. K. Mohandas for a human–rodent hybrid cell line; Albert de la Chapelle and Robert Weinberg for human cell lines; Stuart Orkin for a human female genomic library; George Mutter for a Northern blot; Pauline Yen for the STS probe; Daniel Sullivan for help with β-emission imaging of Northern blots; Patrick Baeuerle, Jeremy Berg, Aaron Klug, and Barbara Meyer for stimulating discussions; and Elizabeth Fisher, Eric Lander, Shiuh-Wen Luoh, Graeme Mardon, Elizabeth Simpson, and Steve Swendeman for comments on the manuscript. This study was supported by grants from the National Institutes of Health, the Whitaker Health Sciences Fund, and the Lucille P. Markey Charitable Trust. A. S. G. was supported by a fellowship from the European Molecular Biology Organization.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 13, 1989; revised May 18, 1989.

References

Baker, B. S., and Belote, J. M. (1983). Sex determination and dosage compensation in *Drosophila melanogaster*. Annu. Rev. Genet. *17*, 345–393.

Berg, J. (1988). Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc. Natl. Acad. Sci. USA *85*, 99–102.

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

Brennard, J., Konecki, D. S., and Caskey, C. T. (1983). Expression of human and Chinese hamster hypoxanthine–guanine phosphoribosyltransferase cDNA recombinants in cultured Lesch–Nyhan and Chinese hamster fibroblasts. J. Biol. Chem. *258*, 9593–9596.

Brown, R. S., Sander, C., and Argos, P. (1985). The primary structure of transcription factor TFIIIA has 12 consecutive repeats. FEBS Lett. *186*, 271–274.

Buckle, V., Mondello, C., Darling, S., Craig, I. W., and Goodfellow, P. N. (1985). Homologous expressed genes in the human sex chromosome pairing region. Nature *317*, 739–741.

Burgoyne, P. S., Buehr, M., Koopman, P., Rossant, J., and McLaren, A. (1988). Cell-autonomous action of the testis-determining gene: Sertoli cells are exclusively XY in XX–XY chimaeric mouse testes. Development *10*, 443–450.

Chandra, H. S. (1985). Is human X chromosome inactivation a sex-determining device? Proc. Natl. Acad. Sci. USA *82*, 6947–6949.

Cline, T. W. (1989). Exploring the role of the gene, *Sex-lethal*, (*Sxl*), in the genetic programming of *Drosophila* sexual dimorphism. In Evolutionary Mechanisms in Sex Determination, S. S. Wachtel, ed. (Boca Raton, FL: CRC Press), pp. 23–36.

de la Chapelle, A., Savikurki, H., Herva, R., Tippett, P. A., Knutar, F., Grohn, P., Siponen, H., Huovinen, K., and Korhonen, T. (1984). Aetiological studies in males with the karyotype 46,XX. In Aspects of Human Genetics with Special Reference to X-linked Disorders, C. San Roman and A. McDermott, eds. (Basel: Karger), pp. 125–142.

Fairall, L., Rhodes, D., and Klug, A. (1986). Mapping the sites of protection on a 5S RNA gene by the *Xenopus* transcription factor IIIA: a model for the interaction. J. Mol. Biol. *192*, 577–591.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction fragments to high specific activity. Anal. Biochem. *13*, 6–13.

Ferguson-Smith, M. A. (1988). Genes on the X and Y chromosome controlling sex: genetic sex is a matter of quantity. Brit. Med. J. *297*, 635–636.

Fialkow, P. J. (1978). X-chromosome inactivation and the *Xg* locus. Am. J. Hum. Genet. *22*, 460–463.

Ford, C. E., Miller, O. J., Polani, P. E., de Almeida, J. C., and Briggs, J. H. (1959). A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). Lancet *1*, 711.

Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. (1983). Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. *170*, 827–842.

German, J. (1988). Gonadal dimorphism explained as a dosage effect of a locus on the sex chromosomes, the gonad-differentiation locus (GDL). Am. J. Hum. Genet. *42*, 414–421.

Gibson, T. J., Postma, J. P. M., Brown, R. S., and Argos, P. (1988). A model for the tertiary structure of the 28 residue DNA-binding motif ("zinc finger") common to many eukaryotic transcriptional regulatory proteins. Protein Eng. *2*, 209–218.

Goodfellow, P. N., and Tippett, P. (1981). A human quantitative polymorphism related to Xg blood groups. Nature *289*, 404–405.

Goodfellow, P. N., 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.

Goodfellow, P. J., Darling, S. M., Thomas, N. J., and Goodfellow, P. N. (1986). A pseudoautosomal gene in man. Science *234*, 740–743.

Graves, J. A. and Gartler, S. M. (1986). Mammalian X chromosome inactivation: testing the hypothesis of transcriptional control. Som. Cell Mol. Genet. *12*, 275–280.

Hodgkin, J. (1987). Sex determination and dosage compensation in *Caenorhabditis elegans*. Annu. Rev. Genet. *21*, 133–154.

Jacobs, P. A., and Strong, J. A. (1959). A case of human intersexuality having a possible XXY sex-determining mechanism. Nature *183*, 302–303.

Jaffe, E., and Laird, C. (1986). Dosage compensation in *Drosophila*. Trends Genet. *2*, 310–315.

Jost, A. (1953). Problems of fetal endocrinology: the gonadal and hypophyseal hormones. Rec. Progr. Horm. Res. *8*, 379–418.

Kadonaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987). Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell *51*, 1079–1090.

Keitges, E., and Gartler, S. M. (1986). Dosage of the *sts* gene in the mouse. Am. J. Hum. Genet. *39*, 470–476.

Keitges, E., Rivest, M., Siniscalco, M., and Gartler, S. M. (1985). X-linkage of steroid sulfatase in the mouse is evidence for a functional Y-linked allele. Nature *315*, 226–227.

Ledbetter, D. H., Airhart, S. D., and Nussbaum, R. L. (1986). Somatic cell hybrid studies of fragile (X) expression in a carrier female and transmitting male. Am J. Med. Genet. *23*, 429–444.

Lindsay, S., and Bird, A. P. (1987). Use of restriction enzymes to detect potential gene sequences in mammalian DNA. Nature *327*, 336–338.

Lyon, M. F. (1988). X-chromosome inactivation and the location and expression of X-linked genes. Am. J. Hum. Genet. *42*, 8–16.

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.

Meyer, B. J. (1988). Primary events in C. *elegans* sex determination and dosage compensation. Trends Genet. *4*, 337–342.

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

Miller, J., McLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. *4*, 1609–1614.

Mohandas, T., Sparkes, R. S., Hellkuhl, B., Grzeschik, K. H., and Shapiro, L. J. (1980). Expression of an X linked gene from an inactive human X-chromosome in mouse-human hybrid cells: further evidence for the non-inactivation of the steroid sulfatase locus in man. Proc. Natl. Acad. Sci. USA *77*, 6759–6773.

Mohandas, T., Sparkes, R. S., and Shapiro, L. J. (1981). Reactivation of an inactive human X chromosome: evidence for inactivation by DNA methylation. Science *211*, 393–296.

Mohandas, T., Geller, R. L., Yen, P. H., Rosendorff, J., Bernstein, R., Yoshida, A., and Shapiro, L. J. (1987). Cytogenetic and molecular studies on a recombinant human X chromosome: implications for the spreading of X chromosome inactivation. Proc. Natl. Acad. Sci. USA *84*, 4954–4958.

Page, D. C. (1988). Is *ZFY* the sex-determining gene on the human Y chromosome? Phil. Trans. R. Soc. (Lond.) B *322*, 155–157.

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.

Potter, H., and Dressler, D. (1986). A "Southern Cross" method for the analysis of genome organization and the localization of transcription units. Gene *48*, 229–239.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683–689.

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

Ramsay, M., Bernstein, R., Zwane, E., Page, D. C., and Jenkins, T. (1988). XX true hermaphroditism in Southern African blacks: an enigma of primary sexual differentiation. Am J. Hum. Genet. *43*, 4–13.

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

Shapiro, L. J., Mohandas, T., Weiss, R., and Romero, A. (1979). Non-

inactivation of an X chromosome locus in man. Science *204*, 1224–1226.

Sullivan, D. E., Auron, P. E., Quigley, G. J., Watkins, P., Stanchfield, J. E., and Bolon, C. (1987). The nucleic acid blot analyzer. I: High speed imaging and quantitation of ^{32}P-labeled blots. Biotechniques *5*, 672–678.

Yen, P. H., Allen, E., Marsh, B., Mohandas, T., Wang, N., Taggart, R. T., and Shapiro, L. J. (1987). Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X–Y interchange. Cell *49*, 443–454.

Yen, P. H., Marsh, B., Allen, E., Tsai, S. P., Ellison, J., Connolly, L. Neiswanger, K., and Shapiro, L. J. (1988). The human X-linked steroid sulfatase gene and a Y-encoded pseudogene: evidence for an inversion of the Y chromosome during primate evolution. Cell *55*, 1123–1135.