

© Springer-Verlag New York Inc. 1992

A mouse Y Chromosome pseudogene is related to human ubiquitin activating enzyme *E1*

Priscilla K. Tucker, Kimberly S. Phillips, and Barbara Lundrigan

The Museum of Zoology and Department of Biology, University of Michigan, Ann Arbor, Michigan 48109, USA

Received July 19, 1991; accepted October 2, 1991

Abstract. A 2041 bp DNA fragment isolated from the Sxr (sex reversed) region of the mouse Y Chromosome (Chr) was sequenced and characterized. The sequence, pY8/b, contains four exons that are highly similar to 525 contiguous bases from the cDNA of human ubiquitin activating enzyme E1. Two of the exons contain stop codons, indicating that pY8/b is not part of a functional gene. Sequences related to pY8/b were amplified from the Y Chr of the inbred mouse strain, C57BL/6J. These sequences may be portions of the recently discovered functional equivalent of pY8/b. Despite a high degree of similarity with the human El gene, the functional equivalent of pY8/b is not the mouse El gene, because unlike El, the functional equivalent of pY8/b is expressed in a tissue-specific manner. These data are discussed with respect to theory on the evolution of the mammalian Y Chr, and in particular, to the prediction that functional genes on the Y Chr have a male-specific function.

Introduction

The Y Chr is unique among mammalian chromosomes because, with the exception of a small region that recombines with the X Chr in male meiosis, it is the only chromosome that does not recombine. Muller (1914) theorized that the Y Chr of *Drosophila* diverged from the X Chr as a result of two processes: crossover suppression between the ancestral pair of sex chromosomes, and the subsequent loss of gene function on the Y Chr. One prediction of Muller's hypothesis is that the non-recombining portion of the Y Chr should contain few functional loci. A second is that the Y Chr should have residual loci in the non-recombining region that are homologous to loci on the X Chr. These loci would reflect the ancestral homology between the X and Y Chrs.

Although Muller's hypothesis pertains to the evolution of the Drosophila Y Chr, it is also applicable to the evolution of the mammalian Y Chr. Recent advances in the molecular analysis of the Y Chr in mammals have allowed a direct test of the above predictions. For example, although numerous sequences have been isolated from the non-recombining region of the human and mouse Y Chrs, only three mouse and three human functional genes have been identified. These include the zinc finger genes, Zfy-1 and Zfy-2 in the mouse, and ZFY in human (Page et al. 1987; Mardon et al. 1989; Mardon and Page 1989; Mitchell et al. 1989; Nagamine et al. 1989); the male sex determining locus, Sry in mouse, and SRY in human (Sinclair et al. 1990; Gubbay et al. 1990; Koopman et al. 1991); and the human ribosomal protein gene, RPS4Y (Fisher et al. 1990). The mouse Zfy-1 and Zfy-2 loci appear to be homologous to a locus on the mouse X Chr, Zfx(Mitchell et al. 1989; Nagamine et al. 1989), and the human ZFY and RPS4Y genes appear to have homologs on the human X Chr, ZFX and RPS4X, respectively (Schneider-Gadicke et al. 1989; Palmer et al. 1990; Fisher et al. 1990). The ribosomal protein gene, Rps4, has also been mapped to the mouse X Chr (Hamuas et al. 1992). The presence of an X-specific homolog of $Zf_{V}(ZFY)$ in both mouse and human suggests ancestral homology between the mammalian X and Y Chrs.

To investigate further the evolutionary consequences of crossover suppression of the mammalian Y Chr, we sequenced and analyzed a DNA clone, pY8/b, reported to be a candidate Y-specific gene from the Sxr region of the mouse Y Chr (Bishop et al. 1987). The Sxr region on the Y Chr is a non-recombining region that contains both the testis determining locus, Tdy, and the H-Y transplantation antigen, Hya (reviewed in Eicher and Washburn 1986); it is located on

Offprint requests to: P.K. Tucker

the short arm of the Y Chr (McLaren et al. 1988; Roberts et al. 1988). We report that pY8/b contains part of a pseudogene related to the human ubiquitin activating enzyme E1. We provide evidence from amplification experiments that pY8/b-related sequences are located on the Y Chr in the inbred mouse strain C57BL/6J, and we discuss the relevance of these findings to the evolution of mammalian sex chromosomes.

Materials and methods

Source of pY8/b

pY8/b is a 2041 bp *Eco*RI fragment isolated from a flow sorted Y Chr library derived from a male of the inbred mouse strain designated 163H (Baron et al. 1986). This clone was selected as a candidate Y-specific gene because it hybridized with testis cDNA from male mice, but did not hybridize with liver cDNA from either male or female mice (Bishop et al. 1987). When used as a hybridization probe to *TaqI* restricted genomic DNA from male and female C57BL/6J, pY8/b detects four male-specific fragments. The same four fragments are also present when XX mice carrying the *Sxr* region of the male Y Chr are probed with pY8/b. Thus, pY8/b is from the *Sxr* region of the mouse Y Chr (Bishop et al. 1987).

Sequencing and analysis

pY8/b was sequenced in both directions using sequence-specific synthetic oligonucleotide primers and the Sanger dideoxy sequencing technique (Sanger et al. 1977). The nucleotide sequence was determined from double stranded products using Sequenase Version 2.0 polymerase under the conditions recommended by the manufacturer (United States Biochemical). The computer software program EuGene (release 3.2) was used to identify open reading frames of pY8/b and to search GenBank (release 67.0) and PIR (release 27.0) for similar nucleotide and peptide sequences, respectively. Regions of nucleotide and peptide similarity were identified (Lawrence et al. 1986; Lawrence and Goldman 1988) and percent similarities were calculated.

Amplification experiments

The pY8/b sequence was used to select three primer pairs for polymerase chain reaction (PCR) amplifications (Saiki et al. 1985) of related sequences in C57BL/6J inbred mice. The first primer pair spans the region between nucleotides 9 and 385; the second, the region between nucleotides 1640 and 1950; and the third, the region between nucleotides 1746 and 2030 (Fig. 1). Nine amplification experiments were conducted, one with each primer pair at each of three annealing temperatures (50°C, 48°C, and 38°C). Three samples were included in each experiment: a 1.5 μg sample of female C57BL/ 6J genomic DNA, a 1.5 µg sample of male C57BL/6J genomic DNA, and a blank to which no DNA was added. We used 5 µl of each 10 µM primer in each 100 µl reaction. The concentrations of Tag DNA polymerase, reaction buffer, and deoxynucleoside triphosphates per 100 µl reaction were as recommended by Perkin-Elmer Cetus. The amplification protocol was as follows: 25-35 cycles, with denaturation at 95°C for 1 min, annealing at 50°C, 48°C, or 38°C for 1 min, and elongation at 72°C for 1 min 15 s. PCR products were separated by size on a 3% Nusieve gel, stained with ethidium bromide, and visualized under UV light. To control for possible contamination of our template DNA, these experiments were repeated using two additional female and two additional male samples of C57BL/6J genomic DNA.

Southern blot preparation/DNA labeling/hybridization conditions

PCR products were transferred from the Nusieve gel to nylon membranes in 0.4 M NaOH following the protocol of Reed and Mann (1985). Nylon membranes were prewashed at 65°C for 1 h in 0.1% SSC and 0.5% SDS, and prehybridized at 65°C for 4 h in 100 ml of $4 \times SSCP$, 10× Denhardt's solution, and 1% SDS. Membranes were hybridized overnight at 65°C in 20 ml of $4 \times SSCP$, 2× Denhardt's solution, 1% SDS, and approximately 1.0 × 10⁶ cpm/ml denatured probe labeled with $[\alpha$ -³³P]-dCTP to a specific activity of >1.0 × 10⁹ cpm/µg using random primer labeling (Feinberg and Vogelstein 1983, 1984). Following hybridization, the nylon membranes were rinsed in a solution of 0.5 × SSC, 0.1% SDS at 53°C (low stringency conditions) or 0.1 × SSC, 0.1% SDS at 55°C (high stringency conditions), and exposed to Kodak XAR-5 film for 4 h at room temperature.

Results

Sequence analysis

The pY8/b sequence is 2041 nucleotides in length (Fig. 1) and contains four separate domains that, when spliced together, are highly similar to 525 contiguous bases from the published cDNA of human ubiquitin activating enzyme E1 (Handley et al. 1991; Table 1, Fig. 2). Domain 1 encompasses 72 bases (positions 372-443), contains two stop codons, and is 86% identical at the nucleotide level and 83% identical at the amino acid level with a portion of E1 (positions 2064-2135). Domain 2 encompasses 198 bases (positions 680-877), contains three stop codons, and is 74% identical at the nucleotide level and 79% identical at the amino acid level with a portion of E1 (positions 2136-2330). Domain 2 contains an additional GAC codon at positions 791-793 that is absent in the E1 cDNA. Domain 3 encompasses 75 bases (positions 1438-1512), contains no stop codons, and is 84% identical at the nucleotide level and 81% identical at the amino acid level with a portion of E1 (positions 2331-2405). Domain 4 includes 183 bases (positions 1617-1799), contains no stop codons, and is 84% identical at the nucleotide level and 75% identical at the amino acid level with a portion of E1 (positions 2406–2588).

Amplification experiments

Amplification experiments using the first primer pair (which spans the region between nucleotides 9 and 385, Fig. 1) and annealing temperatures of 50° C, 48° C, and 38° C consistently resulted in the amplification of two male-specific fragments, 376 bp and 410 bp in length (Fig. 3a, top panel). No DNA products were amplified from female genomic DNA. Thus, both fragments are located on the Y Chr.

Amplification experiments using the second primer pair (which spans the region between nucleotides 1640 and 1950, Fig. 1) and annealing temperatures of 50°C and 48°C, resulted in two male-specific fragments, 310 bp and 350 bp in length. No products were amplified from female genomic DNA at these temperatures (not shown in figure). At a lower annealing temperature of

	GAAT TCCT <u>GTTGAGACCTGGGTACTG</u> TTCTGCCCAGTACTAAAAGTAACTGCCAATGGGTCATT	64
	GTTCTTTTAAGTGTGACTTCTTCCTTTCTGTCTCTCTTTCTT	128
	GT AT TGCCACGAA CAT TCCAGAT TCTCACA CATACATACT TGGCTGCAAAGTATGGCCTCTGGT	192
	CTGGAACACGTGGGTA TCCT TAT TGTA CAGGTTAACACAGT CT TCACCCAGAACAGT CT TACT	256
	AACTGGTTGTCTCACTAATGTGTACAAGAAGGGGTTTTTTGGGGGTTGTTGTTGTTGTTGTTGT	320
-	TGTTTTIGTTTGTTTGTTTGGCTATTGCTGACATGCATTTGTTTGCTCTGACTGTAGTGAGCTC	384
Domain 1	GGATGAGT TTGAAGAGCTCTTCAAGCAGT CAGCTGAAAATAT TAACCAATACCTCACGTAAGT	448
	AACAGACAAACCAACAGTGGACAGAAGGCAGTCAAACAAGGCTATCCTTGTCTGCT CTCCTCAG	512
	GCACATCCACGGTGGATGCTTTGGATAACTAGTGGTTGATTGA	576
	ATGGAAAATGGTCAGGATTGACATTTTCCCTGGAGAAAGAA	640
	TAGGCCTCTGAAGTAATCCAGGTCTTTGTATTGGATAGGGACCCCAAGTTCATGGAGCAGACAC	704
0	TGCAGCTAGCTGGCACCCAGCCTTTGTAAGTACTGGAGGCCATACACTGCAGCCTGGTCCTGCA	768
Domain 2	GAGGCCACAGACTTGGGCCGACGACTGTGTGACTTGGGCCTACCAGCACTGACACGCCGAGTAA	832
	TCTCACAACATCCAGCAGTTGTTGCACAACTTTCCTCCAGCTCAGGTATTACATACTTGGTGGA	896
	TT TAT TTGGCTGAGCACAT TTTCAGGAA AT TGGAT GCT TACT GT CCTCTGT CTGTACATGGCTT	960
	CTTCCAGGCAGTCTAGCTGCACAGGCACCTCAAGGCAAGGTGTAGGCCAGTACTTGGTGCAAGG	1024
	ATTTGTGTTTTCCTACTTCAGTCAGCCAGGCAAGGGGGCTTAGTTTCCTTCC	1088
	GGGCTGAGTCACAGCTGAGTTCTCTTTGAGGGCTGTAGTTGTAAGTGAGAAATTTTACAGTTAA	1152
	GCATGCAGTGATCCGTGGGTTTCAAGTCCCAGCTGTGTTTGACACTGGATTGTGACGTGCCTAA	1216
	T G T C T T C G T A G C C A T A T C T G T G A A G T A G G T G G G C T T G G G G T T T A G A A A G G A C T A C T T G C A	1280
	TTTCTTAGTGTTGCCTCATATAAATTGGTAAGTTGAACTTGCCGTGTGTTTGACTCATCGTGAA	1344
	TTTCCCAAGAATGGTGAACCTAGGCTTCTTGGCCTGAGGACTGCCTGAGAGTGCTTCTCCCTGG	1408
	GTCCTCTTCATGCGTTTGCATTCTCCTAGCTTACAAGCTCTGGATCACTTTTCTGGTCAGGACC	1472
Domain 3	AAAACGCTGTCCACATCTGCTCACCTTTGACATAAACAATGTAAGTGTCCTTGGGATCTCCCAG	1536
	GTGGGTATGTGGGAAGTGGGTCTCTCTACTCTTAACCTGTGGTTGTTAAGCTGACTTTCCTTTC	1600
	CATGGCTTTCTTCAGCCCCTGCATCTGGATTATGTGATGGCTGCCCAACCTGTTTGCTCAG	1664
	ATATACGGACTAGGAGGGTCCCAGGACTGTGCTGTGGTGGCCAAACTCCTGCAGTCTCTGCCGG	1728
Domain 4	TCCCCAAGTTTGCTCCCAAGTCTGGCATCAGGATCCATGTTTCTGAGCAGGAGCTGCAGAGTAC	1 79 2
	CAGTGCCACCACCATCAGTGAGAATGCTTCACCATTCGTTCAGCAGTCATCTGTTTGTGTTCTG	1856
	TCCTGCTCAGCCTTCCCTATGGTGTTTGCCACTGGCCAGGATTTTCACTATATTTCTGAGATTT	1920
	ACGTAGTCTGTGCTTAAGGGTTGATGGGTGGACAAACAAA	1984
	ATATGGACTTATGGAAATGGTTCATCC <u>CCTTCCTTGTGTGAGGACT</u> TTATAGAATTC	2041
Eta 1 Th	a nucleotide sequence of nV8/b from the mouse V Chr. Boxed regions represent the four domains ho	molo

Fig. 1. The nucleotide sequence of pY8/b from the mouse Y Chr. Boxed regions represent the four domains homologous to 525 contiguous bases of human ubiquitin activating enzyme E1. Sequences used as primers in the amplification experiments are underlined; in each case the second primer of each pair is the reverse complement of the actual sequence used.

38°C, the same two male-specific fragments were amplified, and four female-specific fragments were detected; these were approximately 1500 bp, 1100 bp, 315 bp and 285 bp in length (Fig. 3b, top panel).

Amplification experiments using the third pair of primers (which spans the region between nucleotides 1746 and 2030) and annealing temperatures of 50°C, 48°C, or 38°C resulted in two male-specific fragments,

Table 1. Percent similarity at the nucleotide and amino acid levelbetween portions of pY8/b from mouse and ubiquitin activating en-zyme E1 from human.

Nucleotide H	Positions	Percent Ident	ity
Mouse pY8/b	Human E1	Nucleotide	Amino Acid
372-443	2064-2135	86	83
680-877	2136-2330	74	79
1438-1512	2331-2405	84	81
1617-1799	2406-2588	84	75
	Nucleotide I Mouse pY8/b 372-443 680-877 1438-1512 1617-1799	Nucleotide Positions Mouse Human E1 372-443 2064-2135 680-877 2136-2330 1438-1512 2331-2405 1617-1799 2406-2588	Nucleotide Positions Percent Ident Mouse PY8/b Human E1 Nucleotide 372-443 2064-2135 86 680-877 2136-2330 74 1438-1512 2331-2405 84 1617-1799 2406-2588 84

284 bp and 320 bp in length, and two fragments shared between male and female, approximately 1300 bp and 1000 bp in length (Fig. 3c, top panel).

The PCR products from the three amplification experiments shown in Fig. 3 were hybridized with radioactively labeled pY8/b using the Southern blot technique (Fig. 3a-c, bottom panel). All male-specific PCR products hybridized at both low and high stringencies, confirming our assumption that these Y-specific PCR products are related to pY8/b. In contrast, the femalespecific fragments, and the fragments shared between male and female C57BL/6J did not hybridize with pY8/b at either low or high stringencies, suggesting that they are not related to pY8/b.

Discussion

The high similarity between four domains in pY8/b and 525 contiguous bases of the cDNA from human ubiquitin activating enzyme EI (Handley et al. 1991) suggests that pY8/b contains exons of a gene related to El. However, the presence of stop codons in the first and second pY8/b domains indicates that these exons are part of a nonfunctional gene. pY8/b was originally isolated from the Y Chr because it hybridizes with a cDNA synthesized from mouse testis mRNA (Bishop et al. 1987). This suggests that pY8/b has a functional equivalent in the mouse genome.

Our amplification data indicate that there are at least two pY8/b related sequences on the Y Chr of C57BL/6J. This finding is consistent with the recent discovery of a functional gene related to pY8/b on the Y Chr in the mouse (Mitchell et al. 1991). It is doubtful, however, that the functional equivalent of pY8/b on the Y Chr is the mouse El gene. Ubiquitin activating enzyme E1 is a regulatory molecule that catalyzes the activation of ubiquitin. Ubiquitin has been isolated from a variety of eukaryotic organisms, and from a wide variety of tissues (reviewed in Wilkinson 1988). One would expect that E1 would also be found in a wide variety of mammalian tissues, and indeed, Northern blot analysis has shown that E1 is present in several human tissues as a single transcript approximately 3.5 kb in length (Handley et al. 1991). In contrast, the functional equivalent of pY8/b on the Y Chr appears to be expressed in a tissue-specific manner, that is, when pY8/b was used as a hybridization probe to poly(A)⁺ RNA from adult male mouse liver and testis, it identified a testis-specific transcript approximately 3.5 kb in length (Bishop et al. 1987). This tissue specificity suggests that the Y-specific functional equivalent of pY8/b is not the mouse E1 gene. Rather, it may be a member of a family of E1-like regulatory genes.

Our amplification and Southern blot experiments indicate that there are no pY8/b-related sequences in the female C75BL/6J genome. However, Mitchell and co-workers (1991) identified a ubiquitously expressed gene on the X Chr which they hypothesize to be the mouse EI gene. If pY8/b is related to the purported mouse EI gene on the X Chr, one would expect to amplify a pY8/b-related sequence in both females and males. Our failure to detect this gene using PCR can be explained in the following way. Most of the primers used in our experiments were constructed from introns, which are evolutionarily less conserved than exons. It is probable, therefore, that our primers were designed from a region of the pseudogene that is not conserved in the functional EI gene on the X Chr.

The presence of a pair of related functional genes in the nonpseudoautosomal portion of the X and Y Chrs is an intriguing find because, like the apparent homology between mouse Zfx and Zfy-1/Zfy-2, and between human ZFX and ZFY, and RPS4X and RPS4Y, it suggests ancestral homology between the mammalian X and Y Chrs. Because recombination is absent between such loci, they are effectively non-allelic genes, and divergence at both the nucleotide and amino acid level resulting in functional differences is expected. Studies of the zinc finger genes in mouse have determined that Zfx and Zfy-1/Zfy-2 are 70% identical at the amino acid level and differ markedly in timing and specificity of expression (Mardon et al. 1990; Nagamine et al. 1990). In contrast, studies of the zinc finger genes in human have determined that ZFX and ZFY are on average 92% identical at the amino acid level and both genes are ubiquitously expressed (Schneider-Gadicke et al. 1989; Lau and Chan 1989; Palmer et al. 1990). The ribosomal protein genes, RPS4X and RPS4Y, located on the human X and Y Chrs are 93% identical at the amino acid level (Fisher et al. 1990). The El-related genes on the X and Y Chrs in mouse are 85% identical at the amino acid level and differ in specificity of expression (Mitchell et al. 1991).

These differences between residual homologs on the X and Y Chrs may be best understood using an evolutionary perspective. For example, Charlesworth (1991) hypothesized a two-step process for the suppression of crossing-over between the X and Y Chrs which has implications for the function of Y-linked loci. First, suppression of crossing-over evolved between genes controlling male and female sex function; and second, genetic exchange was restricted between other sex chromosome-linked genes for which allelic variants were advantageous in males but disadvantageous in females.

One prediction of this hypothesis is that Y-linked loci, other than the sex determining locus, will function differently than their homologs on the X Chr. Y-linked genes would originally have been alleles of homologous X-Y loci that were subject to different

cor coc arr can arr can and cro the one one one of an arr arr are can are cro and arr are are gin an value pre lys gin pro aia giu ann value an gin tyr leu the lys gin pro ease giu ann value and gin his cys arr and giu arr gin are are gin arr and and are			* *	* * *						Ŭ	glu				-	BUT				ile						
<pre>ccr cos car car car car car and car car and an giu ann car car car and met gin arg arg giu pine giu giy leu pine lys gin pro ala giu ann car car car car and car cas can car car can car car can car can car car can car car car rer arg arg giu arg thr leu ala giy thr gin pro leu giu ann car ca can car car car can can car can car can car can car can car tra car can car car can rer arg and can car car can car can car can car tra car can car car can rer arg can can car car car car can car can car tra car car can car car can car can can can car car car car car car car can car tra car car car can car car can car can can can car car car car car car car car car car</pre>	CTG TAG TGA	TAG TGA	TGA	-	E C J	990	CAT (CAG	LTT.	GAA	GAG	DE E	~ DEI	AAG	CAG	LCA	E SCI	GAA	AAT	ATT	AAC	CAA	TAC	CTC	ACG	
<pre>at a try any gut priver of the try gut prover gut and variant with any gut priver the try gut and the try and</pre>	CTG CAG TGG	CAG TGG	TGG		CT.	2000	GAT	C 9 C	EL.	GAA C	290	D E E	LIC .	AAG	CAG	N CO	CCA -	GAA	AAT	GTC	AAC	CAG	TAC	CTC	ACA + hr	
TTC ANG GAG ANG ATG GAG GAT GAT GAT TA T	leu gin trp	gln trp	trp		ala	arg	d a t	glu]	bhe	- n16	д1у	Ten	phe	- в. т.Х.в	d T D	bro Dro	ara	nīb	aen	Val	aen	uтб	суг	nər		
TTC ATG CAG CAG ACA TG CAG CTA CCTAG CAT TTG TTA CTA CTA CAG CGC ATA CAT TGC TTT GTG GAG CAA ACA TGT CGT CAT CAG CAT TAG CAT CAG CTA CAG CTA CAG TTA TTT GTG GAG CAA ACA TGT CAG CAT CAG CAT TAG CAT CAG CTA CAG CTA CAG CTA CTG CAG CAA ACA TGT CGG CAT CAG CAT TAG CAG CTAG CAG CAG CAG CAG CTG CAG CAA CAG TGG CCT TAG CAC TGG CAG CAG CAG CAG CAG CTG CAG CAA CAG TGG CAT TGG CAC TGG CAT TAG CAG CAG CAG CAG Leu gin arg pro gin thr trp ala asp cys val thr trp ala cys his his trp his thr gin AAC ATC CAG CAG TGG CTT CAA CAT CCT CAG CT CG CAC CAG CAG CAG CTG CAG CAG TGG TGG CAA CAT CTT CCT CAG CT CG CAT CAG CAT TGG CAG CAG AAC ATC CAG CAG TTG TGC CAA CAT TTC CT CAG CT CAG CAT CAG CAT CAG CAT AAC ATC CAG CAG TTG TG CAC AAC TTT CCT CAG CT CAG CAT CAG CAT CAG AAC ATC CAG CAG TTG CAC CAC AAC TTT CCT CAG CT CAG CAT CAG AAT AAC ATC CAG CAG TTG CAG CAA AA CGC TGT CAG CAT CAG CAT CAG AAT AAC ATC CAG CAT TTC TGG TCA CAG CAT CCT CAG CAT CAG CAT CAG AAT AAC ATC CAG CAT TTC TGG TCA CAG CT CCG CAG CAG CAG CAG CAG CAG CAG CAG CAG				1		met		aln			aln								* *					ile	his	cys
TTT GTC GAG CGA ACT CGG CTG GCG ACT CAG CCC TTG GAG GTG CTG GAG GCT GTG CAG CGG GAG CTG CAG CGG CAG CTG CAG CAG CTG CAG CGG CGG CGG CGG CGG CGG CGG CGG CG	GAC CCC AAG	CCC AAG	AAG		TTC	ATG	GAG	CAG	ACA	CTG	CAG	CTA	ECT.	000	ACC	CAG	CCT	TTG	TAA	GTA	CTG	GAG	000	ATA	CAC	1GC
phe val glu arg thr leu arg leu ala gly thr gln pro leu glu val leu glu ala val glu pro cre cad acc cad cad cad cad cad cad cad cad	GAC CCC AAG	CCC AAG	AAG		TTT	GTG	GAG	CGA	ACA	CLC	990	CIG	GCA	200	ACT	CAG	200	TTG	GAG	GTG	D C T C	GAG	EC CC	CTC CTC	CAG	8 S S
<pre>erg cha wee cent and the trip alls asp tree erg act tree erg and the erg and the</pre>	авр рго 1ув	pro lys	lув		phe	val	glu	arg	thr	leu	arg	leu	ala	gly	thr	gln	pro	leu	glu	val	leu	glu	ala	val	gln	pro
CFC CAG AGG CCA CAG AGC TGC GAC GAC GAC GAC TGG GAC TGG CAG GAC GAC GAC GAC GAC GAC GAC GAC G														asp						tyr	gln		* * *		ala	glu
<pre>cruc cua cua acc rue der euc rue crea acc rue dec rue crea cua cua cua alla Auc arc cua cua rue rue alla apper cys val thr trp alla cyp his his trp his thr glu Auc arc cua cua rue rue alla apper cys val thr trp alla cyp his his trp his thr glu Auc arc cua cua cua rue cua cua rue cua cua cua cua cua cua and arc cua cua cua cua rue cua cua cua cua cua cua and arc cua cua cua cua cua cua cua cua cua cu</pre>	AGC CTG GTC	CTG GTC	GTC		cro	CAG	AGG	CCA	CAG	ACT	TGG	220	GAC	GAC	TGT	GTC	ACT	TGG	000	TAC	CAG	CAC	TGA	CAC	с С С	GAG
And are program are the transmission of the tr	AGC CTG GTG	CTG GTG	GTG		CIG.	CAG	CGA	CCA	CAG	ACC	U U U U U U U	ະ ບິ	GAC		LGC L	010	ACC	L L C		TGC	CAC A	CAC		CAC his	ACC	
ANC ATC CAG CAG TTG CAC ANC TTT CCT CCG GGC CAG ANC ATC CGG CAG CTG GAC ANT TT CCT CCT CGG GGC CAG ann ile arg gin leu leu his asn phe pro pro asp gin aer leu TCT GGA TCG CTT TGG TCG GGA CCA ANA CGC TGT CAG CTT GAG ATT ANA ANT TCG GGA CGG TTT TGG TCG GGA CCA ANA CGC TGT CGA CTT GAG GTC AAC ANT TCG GGA GGG CGG TTT GGG TCG GGG CCC ANA CGC TGT GGA GGA TTG GAT GTC AAC ANT TCG GGA GGG CGG TTT GGG CCC ANA CGC TGT CGA CGC TTT GAG GTC AAC ANT TCG GGA GGG CGG TTT GGG CCC ANA CGC TGT CGA CGG TCG CAG GAG TC TCG GGA TTAT GTG ATT GGG CCC ANA CGC TGT CGG CTC ACC TTT GAT GTC AAC ANT TCG GGA TTAT GTG ATT GGG CCC ANA CGC TGT CGG CTC CGG CTC CGG GGG TCC CAG GAG TTAT GTG ATT GTG ATT GCC CAG CTG ATA TAC GGA CTT GAT GTC AAC ANT TCG GGC TAT GTG ATT GCC ACC CTG TTT GGT CGG CTG GGG TCC CAG GAG TTAT GTG ATT GTG CTG GCC ANA CCC TTT GGT CGG CTG GGG TCC CAG GAG TTT GTG GCC ANA CTC TGG CGC AAC CTG TTT GCT CGG CTG GGG TCC CAG GAG TTT GTG GCC AAA CTC CTG CCC AAC CTG TTT GCT CGG CTG GGG TCC CAG GAG TTT GTG GCC AAA CTC CTG CGG GTC CCC AAG TTT GCT CCC AAG TTT GGT CCC AAG ATT GCT CCC AAG TTT GCT CCC CAG TTT GCT CCC CAG TTC AGG ATC AGG ATC CAT TCG GGC AAA CTC CTG CGG GTC CCC CAG TTT GCT CCC AAG TTT GCT CCC AAG TTT GCT CCC CAG TTT GCT CCC AAG TTT GCT CCC CAG TTT GCT CCC CAG TTT GCT CCC AAG TTT GCT CCC CAG TTT GCT CCC CAG TTT GCT CCC CAG TTT GCT CCC AAG TTT GCC CCC AAG TTT GCC CCC AAG TTT GCC CCC AAG TTT GCT CCC AAG TTT GCT CCC AAG TTT GCC CCC AAG TTT GCC CAG TTT GCC CCC AAG TTT GCC CAG TTT GCC CAG TTT GCC CAG TTT GCC CAG TTT GCCC	ser leu val	leu val	VAL		leu	dTn	arg	pro	gın	thr	τrp	818	asp		сув	V B L	LUE	d 1 1	e Te	c V B	8 T U	9 T U	d 1 1	8717		11T 6
ANC ATC CNG CNG TTG TTG CTC ANC TTT CCT CCA GCT CNG ANC ATC CGC CNG CTG TTG CAC ANC TTT CCT CCT GAC CAG ann ile arg gin leu leu his asn phe pro pro arg gin tTCT GGA TCA TTTTTTTTTTTTTTTTTTTTTTTTTTTT	*** his	hia	hia				gln									ala										
ANC ATC CGG CAG CTG CTG CAG ANG TTC CGT CGA CAG ann lie arg gin leu leu his ann phe pro pro asp gin TCT GGA TCH TTC TGG TCA GGA CCA ANA CGC TGT CCA CAT TTG GAT GTC ANG ANT TCA GGA GCG CGG TTC GGG CCA ANA CGC TGT CCA CAT GAT GTC ANG ANT FCA GGA GCG CGG TTC GGG CCA ANA CGC TGT CCA CAT GAT GTC ANG ANT FTC GGA TTA GTG TTC GGG TCA GGA CCA ANA CGC TGT CCA CAC TTT GAT GTC ANG ANT FCA GGA GCG CGG TTC GGG CCC ANA CGC TGT CCA CAT GAT GTC ANG ANT FTC GGA TTA GTG ATG GCT GGG CCC ANA CGC TCA CGA CTA GGG GGT CCAG GAG FTC GAT TAT GTG ATG GCT GCT GCC ANA CGC TAG GGA GGT TCC CGG GAG FTC GAC TAT GTG ATG GCT GCT GCC ANG TTT GGT GAG GGT TCC CGG GAG FTC GAC TAT GTG ATG GCT GCC AAC CTG TTT GGT GAG GGT TCC CGG GAG FTC GAC TAT GTG ATG GCT GCC GAG TTT GGT GGG GCT CAG GGG TCC CGG GAG FTC GCC ANA CTC CTG GCT GCC GAG TTT GGT CGG ATA GGG GGC TCC CGG GAG FTC GCC ANA CTC CTG GGT GCC GAG TTT GCT CGG GTT GGG GTC CGG GAG FTC GCC ANA CTC CTG GGG GTC CCC GAA TTT GGT CCC CAG GTT GGG GTC GGG GAG FTC GCC ANA CTC CTG GGG GTC CCC GAA TTT GGT CGC CTG GGG GTC AGG GTC CAG FTC GCC ANA CTC CTG GGG GTC CCC GAA TTT GGT CCC CAG GTT GGC GTC AGG GTC CGC FTC GGC ANA CTC CTG GGG GTC CCC GAA TTT GGT CCC CAG GTT GGC GTC AGG GTC CAG FTC GGC ANA CTC CTG GGG GTC CCC GAA TTT GGT CCC CAG GTT GGC GTC AGG GTC CAG FTC GCC AAA CTC CTG GGG GTC CCC GAA TTT GGT CCC CAG TTT GGC GTC AGG GTC CAG FTC GCC AAA CTC CTG GGG GTC CCC GAA TTTC GCC CAG GTC CCC CAG TTT GGC GTC AGG GTC CAG FTC GGC ACA CTC CTG GGG GTC CCC GAA TTTC GGC GTC AGG GTC CAG FTC GCC CCC AAA CCC CCC CAG TTT GCC CCC CAG TTT GGC GTC AGG GTC CAG FTC GCC CCC CCC CAG TTT GCC CCC CAG TTT GGC GTC AGG GTC CCC FTC GGC CCC CCC CCC CAG TTT GCC CCC CAG TTT GGC GTC AGG TTC CCC CAG CCC CAG CCC CCC CAG TTT GGC GTC CAG CCC CAG CCC CCC CAG CCC CCC CAG TTT CCCC CAG TTT CCCC CAG TTT CCCC CAG CCC CCC CAG TTT CCCC CAG CCCC CCC CCC CCC CCC CCC CCC C	TAA TCT CAC	TCT CAC	CAC		AAC	ATC	CAG	CAG	11G	TIG	CAC	AAC	TTT	H C C C H	CCA	ECE CCH	CAG									
<pre>asn ile arg gln leu leu his asn phe pro pro asp gln TCT GGA TCA CTT TTC TGG TCA GGA CCA AAA CGC TGT CTG CTC ACC TTT GAT GAT AAC AAT TCA GGA GCG CCG TTC TGG TCT GGC CCA AAA CGC TGT CCA CCA CTT GAT GAT GAT AAC AAT Ser gly ala pro phe trp ser gly pro lys arg cys pro his pro leu thr phe asp val asn asn tcr GAT TTAT GT GGT GCT GCC AAA CGC TGT CCA CGC TTT GAT GAT GAT AAC AAT ser gly ala pro phe trp ser gly pro lys arg cys pro his pro leu thr phe asp val asn asn tcr GAT TTAT GT GGT GCT GCT CCA CGT CTA GGA GGG TCT GAG GAC TTAT GT GT GT GCT GCT GCT CAG TTT GGT CAG GAT GAG GGT CAG GAC TTAT GT GT GT GCT GCT GCT CAG TTT GGT CAG GAT GAG GGT CAG GAC TTAT GTG TAT GTT GT GCT GCT CAG ATA TAC GGA GGA GGG TCC CAG GAC Ieu asp tyr val met ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp tr val ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ise can val ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ise can thr met ala asn asn ala</pre>	TAC TCG AAC	TCG AAC	AAC		AAC	ATC	g	CAG	CIG	CIG	CAC	AAC	TTC	ដ្រ	CCI	GAC	CAG									
TCT GGA TCA CTT TTC TGG TCA GGA CCA ANA CGC TGT CCA CTT TGAC TTT GAC TTT GAC TTA ANC ANT TCA GGA GCG CTC TGG TCT GGG CCC ANA CGC TGT CCA CCT TTGA CTT GAT GAT ANC ANT SET GJY ala pro phe trp ser gJy pro Jys arg cys pro his pro leu thr phe asp val asn asn the asp tyr val met ala ala ala an leu phe ala gJu thr tyr gJy leu thr gJy ser gJu asp tal ala thr phe asp tro ccc And CCC CAG CCC ANG CCC CAG CTG GGA GGG TCT CAG GAC CTG CGA GGA GGG TCT CAG GAC TAT GTG ANT TAC CTG GAT TAT CTG TC TAT GTT GCT GCT GCA TAT GCT CAG GAT TAT GTT GTT GTT GTT GCT GCA TAT GTT GTT GTT GTT GTT GTT GTT GTT GT	tyr ser asn	ser asn	a 8 n		8 8n	ile	arg	gln	leu	leu	his	asn	phe	pro	pro	asp	gìn									
TCT GGA TCA CTT TC TGG TCA GGA CCA TAA CGC TGT CTG CTC TTT GAC ATA AAC AAT TCA GGA GCG CCG TTC TGG TCA GGA CCA TT GGA CTT GAT AAC AAT EET g1y ala pro phe trp ser g1y pro 1ys arg cys pro his pro leu thr phe asp val asn asn er g1y and tr tyr g1y er g1y asn true asp tyr val met ala ala ala asn leu phe ala g1n thr tyr g1y leu thr g1y ser g1y arg ta thr phe leu g1y er g1y er g1y er g1y er g1y er g1n asp true thr phe leu g1n er val g1n val pro g1u phe thr pro 1ys er g1y val lys ite his true dat cros cros cros cros and cros cros and cros cros and thr tyr g1y leu thr g1y ser g1n asp true thr phe leu g1n er val g1n val pro g1u phe thr pro 1ys er g1y val lys ite his thr g1a er ala an				1																						
TCT GGA TCA CTT TTC TGG TCA GGA CCA AA GGC TGT CGA CTT GG TTT GAC AT AA AA AT TCA GG GCG TTC TGG TCT GGG CCG AA GG TGC TGA CTT GAT AT AA AT TCA GG GC CC AA GG CC TAT GGA GG TCG GT AC AC AT AA AA AT TCA GG GC CC TAT GG GC TTT GAT TT GAT AT TA TA AA AT AT TA TA TA TA TA TA T	7						36 r	leu											leu					11e		
TCA GGA GCG CTC TGG TCT GGG CCC AAA CGC TGT CAA CCG CTC AC ATT GAT GTC AAC AAT ser gly ala pro phe trp ser gly pro lys arg cys pro his pro leu thr phe asp val asn asn asn CTG GAT TAT GTG ATG GCT GCT GCT AAC AAT TAC GGA CTA GGA GGG TCC CAG GAC CTG GAT TAT GTG ATG GCT GCT GCT AAC TAT GTA TAC GGA CTA GGA GGG TCC CAG GAC leu asp tyr val met ala ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp trg GCG GCC AAA TAC GCG CCC AAG TTT GCT CAG GAG TCT CAG GAC TAC GCC TAC GCT GCT GCC CAG ATA TAC GGA CTA GGA GGG TCC CAG GAC leu asp tyr val met ala ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp trg GCG GCC AAA TTC CTG GCT GCG GAG TTT GCT CAG GAC TTT GCT CAG GAC TAT GTC CTG CAG GTC CCC AAG TTT GCT CCC AAG TTT GCT CAG GAC AAA TTC CTG CAG GTC CCC AAG TTT GCT CCC AAG TTT GCT CCC AAG TTT GCT CAG GAC ATC AGG ATC CAG TAT ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ile his thr ser ala asn asn ala	CTT ACA AGO	ACA AGO	AGO		HOT.	GGA	TCA	CIT	TTC	TGG	TCA	CGA CGA	CCA	AAA	ы С	TGT	CCA	CAT	CIG	CIC	ACC	TTT	GAC	ATA :	AAC	TAA
<pre>ser gly ala pro phe trp ser gly pro lys arg cys pro his pro leu thr phe asp val asn asn crg GAT TAT GTG ATG GCT GCT GCT AGC TTT GCT CAG ATA TAC GGA CTA GGA GCG TTC CAG GAC CTG GAC TAT GTG ATG GCT GCT GTT GCT CAG ATA TAC GGA CTA GGA GCG TTC CAG GAC Leu asp tyr val met ala ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp GTG GCC AAA TTC CTG CGG GTC CAG GTC CAG GTG ACA GGA TTC CAG GAC GTG GCC AAA TTC CTG CGG GTC CAG TTT GCT CCC AAG TTT GCT CGG CTG AGG TTC CGG TT ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ile his thr set cdf GCC CGG AAT CCC CAG GTC CCC CAG TTT GCT CCC AAG TTT GCT CCC CAG GTC CCC CAG TTT CCCC CAG TTT CCCCC CAG TTT CCCCC CAG TTT CCCCC CAG TTT CCCCC CAG TTT CCCCCCCCCC</pre>	CTC ACA AGC	ACA AGC	AGO		TCA	GGA	с С С	500	TIC	166	TCT	39	2000	AAA	ပ္ပ	TGT	CCA	CAC	8 S	CTC CTC	ACC	LLL	GAT	CTC .	AAC	TAA
CTG GAT TAT GTG ATG GCT GCT GCT ATT GCT CAG ATA TAC GGA GTG GTG GAG GTG TAT GTG ATG TAT GTG ATG TAT GTG ATG TAT GTG ATG TAT GTG GAG GTG TAT GTG ATG TAT GTG ATG TAT GTG ATG TAT GTG TAG GGA TCT CAG GAC I u asp tyr val met ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp tyr gtg GCC AAA CTC CTG CAG ATG TAT GTG TAT GTG TAT GTG TAG GGG TAG GAG TAT GTG TAG GGG TAG ATG TAG GAG TAT GTG TAG GAG TAT GTG TAG GAG TAT GTG TAG GAG TAT TAG TAG	leu thr ser	thr ser	801		ser	glγ	ala	pro	phe	trp	86ľ	glγ	pro	lув	arg	сув	pro	his	pro	leu	thr	phe	asp	val	asn	asr
CTGGATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	4			2														ile				gly				
CTG GAC TAT GTG ATG GCT GCT GCT AT GCC CAG ACC TAC GGG CTG ACA GGC TCT CAG GAC ILU asp tyr val met ala ala ala asn leu phe ala gln thr tyr gly leu thr gly ser gln asp upe ala leu pro leu pro leu pro leu pro da ctr crG CAG ATC CAG ATC CAG ATC TC CG AAC TCT CTG CG CTG ATG ATG ATG ATG ATG ATG ATG ATG ATG A	CCC CTG CN	CTG CA	C	E	CTG	GAT	TAT	GTG	ATG	GCT	GCT	000	AAC	0 LO	TTT	GCT	CAG	ATA	TAC	GGA	CTA	GGA	89	TCC	CAG	GAC
leu asp tyr val met ala ala ala ala ala ala ala gun thr tyr gly leu thr gly ser gln asp lys leu GTG GCC AAA CTC CTG CG TCT CTG CCG AGG TTT GCT CCC AAG TTT GCT AGG ATC AGG ATC CAT val ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ile his thr ser CTG AGT AGT AGT CCC AAG TTT AGT CCC AAG TTT AGT AGT AGT AGT AGT AGT AGT AGT	CCC CTG CA	CTG CA:	CA	64	CIG	GAC	TAT	GTG	ATG	E C E	5 GCH	200	AAC	CTC CTC	TTT	50	CAG	ACC	TAC	999	0 TO	ACA	g	E TC	CAG	GAG
Ive leuIveleuprolysalaileargGTGGCCAAACTCCTGCGGGTCCAGTCTCTGAGGATCGTGGCCAAACTCCTGCGGGTCCGGGTCAGGATCAGGATCGTGGCCAAACTCCTGCGGGTCGCGGTCAGGATCAGGATCValalathrpheItrproJuphethrproJyvalJysilehisCAGGAGCTGCAGGTCGAGTCTGCCAAGTCTGCCAAGATCCATValalathrphothrproJysserglyvalJysilehisCAGGAGCTGCAGAGTGCCGATGCCGCCAAGGCCCATGCCGAGTCTGCCAAGGCCCATGCCCATGCCCATGCCCATGCCCATGCCGCCAAGGCCCATGCCGCCAAGGCCGCCGCCAAGGCCGCCAAGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC<	pro leu his	leu hi	hid	80	leu	asp	tуr	val	met	ala	ala	ala	8 80	leu	phe	ala	gln	thr	tyr	gly	leu	thr	ζ19	T BEI	ulg :	186 1
GTG GCC AAA CTC CTG CAG TCT CTG CCG GTC CCC AAG TCT GGC ATC AGG ATC CAT GTG GCC ACA TTC CTG CAG TCT GTG CAG GTC AGG TCT GGC ATC AGG ATC CAT val ala thr phe leu gln ser val gln val pro glu phe thr pro lys ser gly val lys ile his thr ser CAG GAG CTG CAG AGT ACC AGT TCAG CCC AAG TCT GGC GTC AAG ATC CAT thr ser CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGT ACC AAT GCC GIN glu leu gln ser ala asn ala	CVB VD	8 >	2				lys	leu				leu	рго			1ys		ala					ile	arg	-	
GTG GCC ACA TTC CTG CAG TCT GTG CAG GTC CCC GAA TTC ACC CCC AAG TCT GGC GTC AAG ATC CAT val ala thr phe leu gin mer val gin val pro glu phe thr pro lym mer gly val lym ile him thr mer CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGT GCC GIN glu leu gin mer ala asn ala	TGT GCT GTC	GCT GTC	610	75	010	555	AAA	5 E	CTG	CAG	TCT	0 TO	g	GTC	22	AAG	TTT	CCI	22	AAG	TCT	50	ATC	AGG	ATC	CA1
<pre>val ala thr phe leu gin ser val gin val pro glu phe thr pro lys ser gly val lys ile his thr ser thr ser CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGC GCC AAT GCC gin glu leu gin ser ala asn ala</pre>	CGA GCT GCT	GCT GCT	E C C		DID	000	ACA	TTC	CTC CTC	CAG	TCT	010	CAG	GTC	200	GAA	TTC	ACC	20	AAG	E P F	80	GTO :	SAAG	ATC	8
thr ser CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGC GCC AAT GCC gln glu leu gln ser ala asn ala	arg ala ala	ala ala	ala		val	ala	thr	phe	leu	gln	86L	Val	gln	val	pro	glu	phe	thr	pro	lув	398 1	g1y	Lav .	1 JYS	-	i hii
CAG GAG CTG CAG AGT ACC AGT GCC CAG GAG CTG CAG AGC GCC AAT GCC gin glu leu gin ser ala asn ala	glu	glu	glu							thr	ser															
CAG GAG CTG CAG AGC GCC AAT GCC gin giu leu gin ser ala asn ala	GTT TCT GAG	TCT GAG	GAG		CAG	GAG	CIG	CAG	AGT	ACC	AGT	200														
gin giu leu gin ser aia asn aia	GTT TCT GAC	TCT GAC	GAC		CAG	GAG	р Б	CAG	AGC	000	AAT	200														
	val ser asp	ser asp	A 8 p	-	gln	glu	leu	gln	Ber	ala	88 0	ala														
	1 sequence	lence		from	humai	a is gi	ven b	elow .	the E	l segu	lence.	Diffe	rence	s in t	he am	ino a(cid of	the n	nouse	are g	jven :	above	the 1	pY8/b	sequi	ence.

Asterisks indicate stop codons. A dashed line in the E1 sequence indicates a gap.



Fig. 3. Results of the amplification experiments using three pairs of primers from pY8/b. PCR products were separated by size on a 3% Nuseive gel and stained with ethidium bromide. Female (\mathcal{Q}) and male (\mathcal{J}) genomic DNA was used in each experiment as well as a blank to control for contamination. The size standard (S) is *HaeIII* digested phiX174 RF DNA. Top panel: (a) Amplification of the malespecific target size fragment, 376 bp in length, and an additional male-specific fragment, 410 bp in length, from C57BL/6J using an annealing temperature of 48°C and primers spanning nucleotides 9–385. Identical results are obtained at annealing temperatures of 38°C and 50°C. (b) Amplification of the target size 310 bp fragment and an additional 350 bp fragment in male C57BL/6J, and four fragments, approximately 1500 bp, 1100 bp, 285 bp and 315 bp in length in female C57BL/6J using an annealing temperature of 38°C and primers spanning nucleotides 1640–1950. The presence of female-specific

fragments can be explained by hypothesizing that, in the absence of a good match with female template DNA, the primers annealed to other, less well matched genomic sequences unrelated to pY8/b. This is further supported by the fact that the female-specific fragments did not amplify at annealing temperatures of 48° C and 50° C (not shown in figure). (c) Amplification of the target size 284 bp male-specific fragment, an additional 320 bp male-specific fragment, and two fragments shared between male and female, 1300 bp and 1000 bp in length using an annealing temperature of 48° C and primers spanning nucleotides 1746–2030. Identical results are obtained at 38° C and 50° C. **Bottom panel**: Corresponding Southern blot hybridized with pY8/b under high stringency conditions. No femalespecific fragments or fragments shared between males and females hybridized with pY8/b under low or high stringency conditions. selection in males and females. With the evolution of crossover suppression between the ancestral X and Y Chrs, alleles fixed on the Y Chr would become independently evolving loci that code for male-specific characters. If Zfy, RPS4Y, and the EI-related locus on the Y evolved from the ancestral mammalian X and Y, then they should function in a male-specific manner. A direct test of this prediction will come from two kinds of studies: phylogenetic studies to determine whether these loci were members of the ancestral mammalian X and Y Chrs, and functional studies of these genes at the protein level.

The zinc finger genes on the sex chromosomes are the first for which there are data of this kind. Although these genes are present on the sex chromosomes in a variety of eutherian orders (Page et al. 1987), they are not present on the sex chromosomes in metatherians (Sinclair et al. 1988), suggesting that they were not part of the ancestral mammalian X and Y Chrs (assuming a single origin for the mammalian XY pair). In addition, the zinc finger genes on the X and Y Chrs in humans are highly similar and ubiquitously expressed, indicating that the zinc finger gene on the Y may not necessarily function in a male-specific manner (Schneider-Gadicke et al. 1989; Palmer et al. 1990). Together, these data do not violate the prediction that Y-linked loci will function differently than their residual homologs on the X Chr, because ZFY may not have been part of the ancestral mammalian Y Chr. However, the testis-specific expression of Zfy-1 and Zfy-2 in adult mice (Nagamine et al. 1990) suggests that these genes function in a male-specific manner in some species. If the male-specific function of Z_{fy-1} and Z_{fy-2} in the mouse evolved subsequent to the translocation of the zinc finger genes onto the mammalian X and Y Chrs, as the phylogenetic data indicate, then the malespecific function of Zfy-1/Zfy-2 in the mouse cannot be explained by Charlesworth's hypothesis. Rather, a different explanation is needed for the male-specific expression of Zfy-1/Zfy-2 in the mouse. Additional research on the expression of the zinc finger genes across a variety of mammalian taxa should clarify whether differential function of these genes, as seen in the mouse, is indeed an evolutionarily derived trait. More generally, these kinds of studies, investigating the history and function of residually homologous loci from the X and Y Chrs, should significantly further our understanding of sex chromosome evolution.

Acknowledgments. We thank C.E. Bishop for providing us with the pY8/b clone. C.R. Friedman assisted in sequencing experiments. L. Vawter, M. Mitchell, and C.E. Bishop provided comments on the manuscript. This research was supported by funds from a University of Michigan Rackham faculty grant, the Office of the Vice-President for Research, and a grant from the National Science Foundation (BSR-9009806) to P.K. Tucker.

References

mosomes purified by flow cytometry. Som Cell Mol Gen 12: 289-295, 1986.

- Bishop, C.E., Roberts, C., Michot, J.-L., Nagamine, C., Winking, H., Guénet, J.-L., and Weith, A.: The use of specific DNA probes to analyse the Sxr mutation in the mouse. *Development* 101(Suppl): 167-175, 1987.
- Charlesworth, B.: The evolution of sex chromosomes. *Science 251:* 1030-1033, 1991.
- Eicher, E.M. and Washburn, L.L.: Genetic control of primary sex determination in mice. Ann Rev Genet 20: 327-360, 1986.
- Feinberg, A.P. and Vogelstein, B.: A technique for radiolabeling DNA restriction endonuclease fragments to a high specific activity. *Anal Biochem 132*: 6–13, 1983.
- Feinberg, A.P. and Vogelstein, B.: A technique for radiolabeling DNA restriction endonuclease fragments to a high specific activity. *Anal Biochem 137:* 266–267, 1984.
- Fisher, E.M.C., Beer-Romero, P., Brown, L.G., Ridley, A., Mc-Neil, J.A., Lawrence, J.B., Willard, H.F., Bieber, F.R., and Page, D.C.: 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, 1990.
- Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R.: A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature 346*: 245-250, 1990.
- Hamuas, R.M., Brown, S.D.M., Kerr, J.T., Fisher, E.M., Romero, P., Vinn, A., and Page, D.: The mapping of the locus *Rps4* to the X inactivation region in the mouse. *Cytogenet Cell Genet*, in press, 1992.
- Handley, P.M., Mueckler, M., Siegel, N.R., Ciechanover, A., and Schwartz, A.L.: Molecular cloning, sequence, and tissue distribution of the human ubiquitin-activating enzyme E1. Proc Natl Acad Sci USA 88: 258–262, 1991.
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P.N., and Lovell-Badge, R.: Male development of chromosomally female mice transgenic for Sry. Nature 351: 117-121, 1991.
- Lau, Y.-F.C. and Chan, K.: 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, 1989.
- Lawrence, C.B., Goldman, D.A., and Hood, R.A.: Optimized homology searches of the gene and protein sequence banks. *Bull Math Biol 48*: 569-583, 1986.
- Lawrence, C.B. and Goldman, D.A.: Definition and identification of homology domains. CABIOS 4: 25–33, 1988.
- Mardon, G. and Page, D.C.: 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, 1989.
- Mardon, G., Mosher, R., Disteche, C.M., Nishioka, Y., McLaren, A., and Page, D.C.: Duplication, deletion and polymorphism in the sex-determining region of the mouse Y chromosome. *Science* 243: 78-80, 1989.
- Mardon, G., Luoh, S.-W., Simpson, E.M., Gill, G., Brown, L.G., and Page, D.C.: Mouse Zfx protein is similar to Zfy-2: Each contains an acidic activating domain and 13 zinc fingers. Mol Cell Biol 10: 681-688, 1990.
- McLaren, A., Simpson, E., Epplen, J.T., Studer, R., Koopman, P., Evans, E.P., and Burgoyne, P.S.: Location of the genes controlling H-Y antigen expression and testis determination on the mouse Y chromosome. *Proc Natl Acad Sci USA 85:* 6442-6445, 1988.
- Mitchell, M.M., Simon, D., Affara, N., Ferguson-Smith, M., Avner, P., and Bishop, C.E.: Localization of murine X and autosomal sequences homologous to the human Y located testisdetermining region. *Genetics* 121: 803-809, 1989.
- Mitchell, M.M., Woods, D.R., Tucker, P.K., Opp, J.S., and Bishop, C.E.: A candidate spermatogenic gene from the mouse Y chromosome may be a ubiquitin activating enzyme. *Nature 354:* 483-486, 1991.
- Muller, H.J.: A gene for the fourth chromosome of Drosophila. J Exp Zool 17: 325-336, 1914.
- Baron, B., Metezeau, P., Hatat, D., Roberts, C., Goldberg, M.E., and Bishop, C.E.: Cloning of DNA libraries from mouse Y chro-
- Nagamine, C.M., Kaimin, C., Kozak, C.A., and Lau, Y.-F.C.:

Chromosome mapping and expression of a putative testisdetermining gene in the mouse. *Science 342:* 80-83, 1989.

- Nagamine, C.M., Chan, K., Hake, L.E., and Lau, Y.-F.C.: The two candidate testis-determining Y genes (*Zfy-1* and *Zfy-2*) are differentially expressed in fetal and adult mouse tissues. *Genes Develop* 4: 63–74, 1990.
- Palmer, M.S., Berta, P., Sinclair, A.H., Pym, B., and Goodfellow, P.N.: Comparison of human ZFY and ZFX transcripts. Proc Natl Acad Sci USA 87: 1681-1685, 1990.
- 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.: The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* 51: 1091–1104, 1987.
- Reed, K.C. and Mann, D.A.: Rapid transfer of DNA from agarose gels to nylon membranes. *Nucl Acids Res* 13: 7207-7221, 1985.
- Roberts, C., Weith, A., Passage, E., Michot, J.L., Mattei, M.G., and Bishop, C.E.: Molecular and cytogenetic evidence for the location of *Tdy* and *Hya* on the mouse Y chromosome short arm. *Proc Natl Acad Sci USA* 85: 6446–6449, 1988.
- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, A., and Arnheim, N.: Enzymatic amplification of β -globin

genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230:* 1350-1354, 1985.

- Sanger, F., Nicklen, S., and Coulson, A.R.: DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463– 5467, 1977.
- Schneider-Gadicke, A., Beer-Romero, P., Brown, L.G., Nussbaum, R., and Page, D.C.: ZFX has a gene structure similar to ZFY, the putative human sex determinant, and escapes X inactivation. Cell 75: 1247-1258, 1989.
- Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.-M., Lovell-Badge, R., and Goodfellow, P.N.: A gene from the human sexdetermining region encodes a protein with homology to a conserved DNA-binding motif. *Nature 346*: 240-244, 1990.
- Sinclair, A.H., Foster, J.W., Spencer, J.A., Page, D.C., Palmer, M., Goodfellow, P.N., and Graves, J.A.: Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. *Nature 336*: 780–783, 1988.
- Wilkinson, K.D.: Purification and structural properties of ubiquitin. In M. Rechsteiner (ed.); Ubiquitin, pp. 5–38, Plenum Press, New York, 1988.