MOLECULAR AND CELLULAR BIOLOGY, Apr. 1994, p. 2485–2492
0270-7306/94/$04.00+0

Vol. 14, No. 4

Structure and Function of Ribosomal Protein S4 Genes on the Human and Mouse Sex Chromosomes

ANDREW R. ZINN,[1]† RAAJI K. ALAGAPPAN,[1] LAURA G. BROWN,[1] IRA WOOL,[2] AND DAVID C. PAGE[1]*

Howard Hughes Research Laboratories at Whitehead Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142,[1] and Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637[2]

Received 29 November 1993/Accepted 18 January 1994

The human sex-linked genes *RPS4X* and *RPS4Y* encode distinct isoforms of ribosomal protein S4. Insufficient expression of S4 may play a role in the development of Turner syndrome, the complex human phenotype associated with monosomy X. In mice, the S4 protein is encoded by an X-linked gene, *Rps4*, and is identical to human S4X; there is no mouse Y homolog. We report here the organization of the human *RPS4X* and *RPS4Y* and mouse *Rps4* genes. Each gene comprises seven exons; the positions of introns are conserved. The 5′ flanking sequences of human *RPS4X* and mouse *Rps4* are very similar, while *RPS4Y* diverges shortly upstream of the transcription start site. In chickens, S4 is encoded by a single gene that is not sex linked. The chicken protein differs from human S4X by four amino acid substitutions, all within a region encoded by a single exon. Three of the four substitutions are also present in human S4Y, suggesting that the chicken S4 gene may have arisen by recombination between S4X- and S4Y-like sequences. Using isoform-specific antisera, we determined that human S4X and S4Y are both present in translationally active ribosomes. S4Y is about 10 to 15% as abundant as S4X in ribosomes from normal male placental tissue and 46,XY cultured cells. In 49,XYYYY cells, S4Y is about half as abundant as S4X. In 49,XXXXY cells, S4Y is barely detectable. These results bear on the hypothesized role of S4 deficiency in Turner syndrome.

The eukaryotic ribosome is a complex of four RNA molecules and about 80 distinct proteins (43). As a rule, there are multiple processed pseudogenes but only one functional gene for each mammalian ribosomal protein (4, 6, 9, 25). Different functional ribosomal protein genes, though dispersed (11), are coordinately transcribed and translated. The dozen or so mammalian ribosomal protein genes that have been examined show several characteristic features that may be relevant for coordinate expression (1, 3–5, 7–9, 18, 19, 23, 30, 32, 40, 42). The genes are small, in most cases spanning only a few kilobases. Their promoters lack a canonical TATA sequence. Transcription initiates at a cytosine residue within a pyrimidine-rich sequence that is important for translational regulation (14, 22). The first exon usually encodes only a short 5′ untranslated sequence and one or a few amino acids. Nucleotide sequence analysis and biochemical studies suggest that specific transcription factors are shared by some but not all ribosomal protein genes (16, 23).

Human S4, a component of the 40S subunit, is unique among known mammalian ribosomal proteins in that two functional genes encode nonidentical isoforms (13). The X-linked gene, *RPS4X*, encodes a protein of 263 amino acids identical in sequence to rat S4 (44). Its Y-linked homolog, *RPS4Y*, encodes a protein that differs by 19 amino acid substitutions. Both *RPS4X* and *RPS4Y* are ubiquitously transcribed, and *RPS4X* is expressed from both the active and inactive X chromosomes. In humans, normal embryonic and fetal development may depend on the presence and function of two S4 genes per cell. Unlike humans, the mouse has only one functional S4 gene, *Rps4*. Mouse S4 is identical in sequence to human S4X and is also X linked (15), but in contrast to human *RPS4X*, mouse *Rps4* undergoes X inactivation (2, 46). To better understand the relationship of human *RPS4X*, *RPS4Y*, and mouse *Rps4*, we determined the genomic organization and transcription start sites of all three genes. For comparison, we also examined in less detail the S4 gene of chicken.

The function of human S4X was inferred from the absolute identity of its amino acid sequence to that of the previously characterized rat S4 protein. On the basis of sequence similarity, the S4Y protein was hypothesized to function interchangeably with S4X (13). We recently showed that human *RPS4Y* complemented a temperature-sensitive hamster S4 mutation, with human S4Y replacing hamster S4 in ribosomes (41). Using isoform-specific antisera, we now show that both S4X and S4Y are present in vivo in human ribosomes. We also examine the levels of S4X and S4Y in ribosomes from normal cells and from cells with supernumerary X or Y chromosomes. We find that the relative intracellular concentrations of the S4X and S4Y proteins reflect their gene dosage.

MATERIALS AND METHODS

Cloning, sequencing, and hybridization. Cloning of the human *RPS4X* and *RPS4Y* and mouse *Rps4* genomic loci was previously described (13, 15). For chicken S4, a λgt11 cDNA library made from White Leghorn day 8 embryonic eyes was screened by hybridization with a human *RPS4Y* cDNA (13). Three independent cDNAs were isolated, subcloned into Bluescript plasmid, and sequenced. Primers GATCGACGG CAAAGTTCG (forward) and GAGGGATTCCTTTGGT GCC (reverse) or GTTCCCTGTGTCAAACTTGATG (reverse) were chosen from the cDNA sequence to amplify by PCR a portion of the chicken S4 gene from genomic DNA. Reaction mixtures contained 1 μM each primer, 1 μg of

* Corresponding author. Mailing address: Whitehead Institute, Nine Cambridge Center, Cambridge, MA 02142. Phone: (617) 258-5203. Fax: (617) 258-5578. Electronic mail address: PAGE@WI.MIT.EDU.
† Present address: The University of Texas Southwestern Medical Center, Dallas, TX 75235.

template, 10 mM Tris (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl, 5 mM NH_4Cl, 100 μM deoxynucleoside triphosphates (dNTPs), and 2.5 U of *Taq* DNA polymerase per 100 μl. PCR conditions were 94°C for 1 min, 61°C for 1 min, and 72°C for 1 min for 30 cycles. The product was directly cycle sequenced by using ^{32}P-labeled primers and *Taq* DNA polymerase. Hybridization of chicken genomic DNAs (10 μg) with a random-primer labeled chicken cDNA probe was performed as previously described (13).

Mapping 5′ ends of transcripts. Poly(A)$^+$ RNAs were prepared by using oligo(dT)-cellulose as previously described (35). Standard methods were used for primer extension and S1 analyses (34). For primer extension, antisense oligonucleotides AAACACACCGGTCAATTTAT (human *RPS4X*), AATAC ACCCGTTAGTTTGTC (human *RPS4Y*), and AAACACGC CAGTCAACTTAT (mouse *Rps4*) were labeled with [γ-^{32}P] ATP and T4 polynucleotide kinase. To generate S1 probes for *RPS4X* and *RPS4Y*, intronless minigenes were constructed by fusing exon 1 and 5′ flanking genomic sequences to downstream exons from cDNAs at an *Nco*I restriction site present in both genomic and cDNA clones. Single-stranded probes were generated by asymmetric PCR, using the same labeled oligonucleotides as for primer extensions and unlabeled oligonucleotide AGTCTCCAGCCCCAATTTCT (*RPS4X*) or GCTAA GAAATTCAGTTCCAG (*RPS4Y*) from 5′ flanking sequences. Reaction mixtures (50 μl) contained 2 μM unlabeled and 2 pM labeled primer, 1 ng of template, 10 mM Tris (pH 8.3), 1.5 mM $MgCl_2$, 100 mM KCl, 5 mM NH_4Cl, 100 μM dNTPs, and 1.25 U of *Taq* DNA polymerase. PCR conditions were 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min for 35 cycles. Single-stranded products were purified by nondenaturing polyacrylamide gel electrophoresis (PAGE).

Ribosomal protein preparation, electrophoresis, and immunoblotting. Human male placental 40S protein was purified as described previously (37), with the addition of deoxycholate to the homogenization buffer. Protein (100 μg) was analyzed by two-dimensional gel electrophoresis (37). The first-dimension gel was run at pH 8.6 in 4% acrylamide (2- by 70-mm tube gel) for 4 h at 90 V. Electrophoresis in the second dimension was for 4 h at 90 mA at pH 4.2 on a 15% acrylamide slab gel (10 by 30 by 2 mm). Proteins were visualized with Ponceau S and then transferred to nitrocellulose in 0.7% acetic acid at 70 mA for 16 h (39). Purification of ribosomal proteins from cultured cells, preparation of anti-S4X and anti-S4Y sera, expression of S4X and S4Y in *Escherichia coli*, and immunoblotting were done as previously described (41).

Nucleotide sequence data. Primary nucleotide sequence data in this report are available from GenBank under accession numbers L24368 (chicken *RPS4*), L24369 (human *RPS4X*), L24370 (human *RPS4Y*), and L24371 (mouse *Rps4*).

RESULTS

Mammalian S4 genes. Phage clones of the human *RPS4X* and *RPS4Y* (13) and mouse *Rps4* (15) genes were isolated previously. The genes spanned approximately 4 kb for human *RPS4X*, 25 kb for human *RPS4Y* (47), and 6 kb for mouse *Rps4*. Phage inserts were subcloned and partially sequenced to determine the positions of splice sites (Fig. 1). Each gene comprised seven exons. All splice donor and acceptor sites conformed to the GT/AG rule. The positions but not the lengths of introns were completely conserved. Limited sequencing of introns did not reveal any obvious nucleotide similarity among the three genes past the immediate vicinity of splice sites (not shown). The first exon of each gene consisted of just the 5′ untranslated sequences and the initiator methionine codon. For human *RPS4X* and *RPS4Y*, the entirety of each cDNA sequence was determined in genomic phage clones. Exon sequences from genomic and cDNA clones were identical except for a previously noted silent nucleotide substitution in *RPS4X* (13), which is a polymorphism (unpublished data). The mouse *Rps4* phage insert ended at an *Mbo*I restriction site within the last exon. The remainder of the mouse gene was sequenced from a PCR product obtained by amplification of genomic DNA with one primer from the last intron and the other from the 3′ end of the cDNA sequence (not shown).

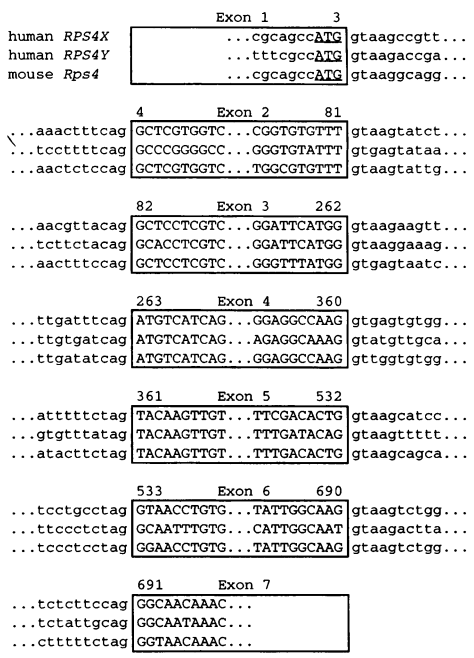

FIG. 1. Splice sites of human *RPS4X* and *RPS4Y* and mouse *Rps4* genes. Exons are boxed. Translated sequences are capitalized. Numbers refer to nucleotide positions within cDNA sequences, beginning with the initiator methionine codon (underlined).

We next mapped the 5′ ends of mammalian *RPS4* transcripts by primer extension of poly(A)$^+$ RNA from male or female human lymphoblastoid cells or male mouse liver tissue (Fig. 2). Primers corresponding to nucleotides 81 to 62 (human *RPS4X* and mouse *Rps4*) or nucleotides 80 to 61 (*RPS4Y*) of exon 2 were chosen. Primer extension yielded products of 105 and 104 bp for *RPS4X*, 103 and 102 bp for *RPS4Y*, and 104 bp for *Rps4*. Extension with the *RPS4Y* primer using RNA from 46,XX cells yielded no product, confirming the primer's specificity. In repeated experiments, the smaller products were not always observed (not shown), so we considered the 105-bp *RPS4X* and 103-bp *RPS4Y* products to reflect the major transcription start sites.

Our longest *RPS4X* cDNA, pDP1284 (13), extends 12 bp further 5′ than the start site indicated by primer extension. Because of this discrepancy, we confirmed the 5′ ends of *RPS4X* and *RPS4Y* mRNAs by S1 nuclease protection. Asymmetric PCR was used to generate single-stranded DNA probes predicted to protect products of the same length as the primer extension products. As shown in Fig. 2, the S1 protection and primer extension products were within a few base pairs of each other in size. S1 analysis also suggested the presence of longer *RPS4X* transcripts of lower abundance. The cDNA pDP1284

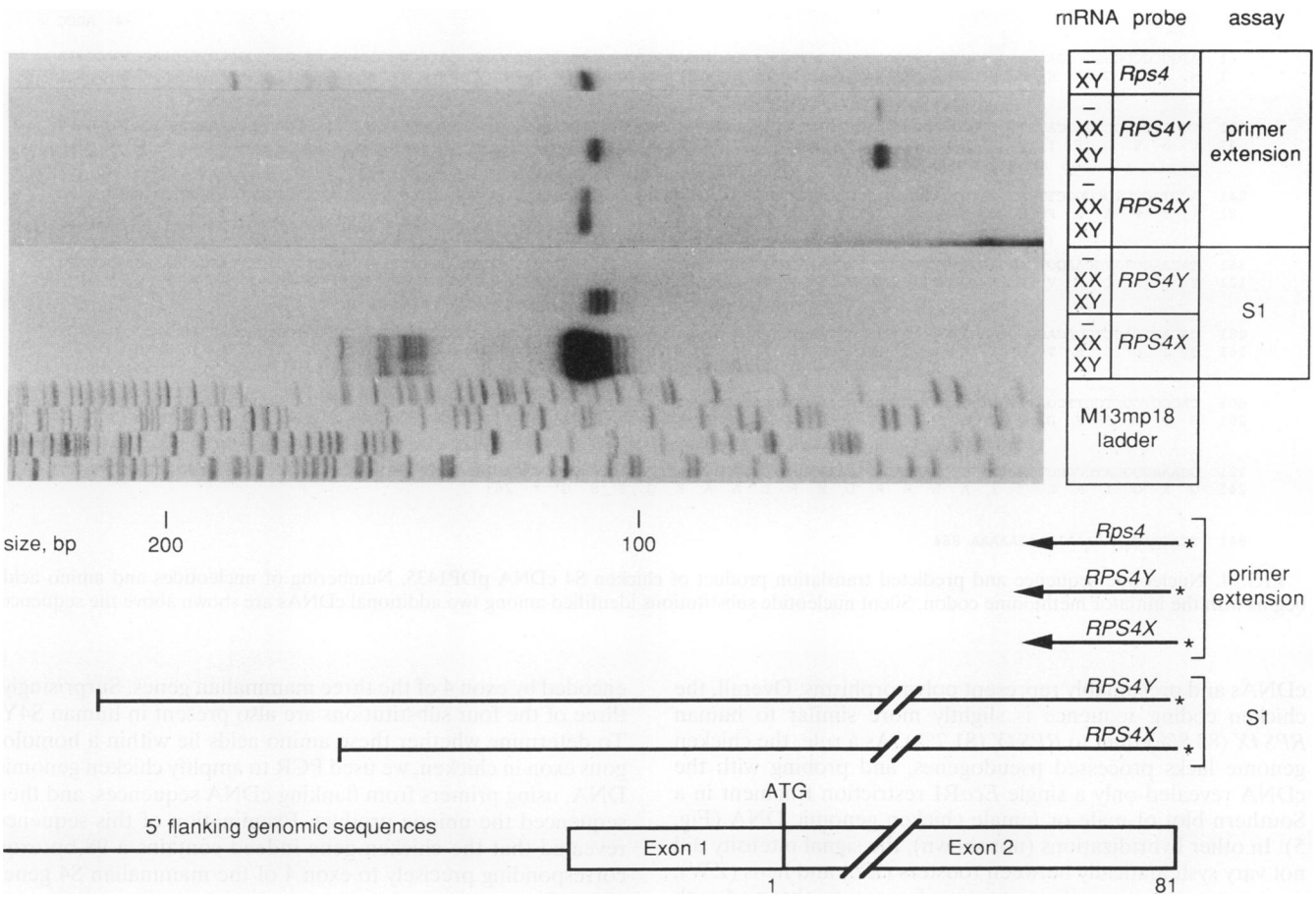

FIG. 2. 5′ ends of mammalian S4 transcripts. The results of primer extension and S1 nuclease protection experiments are shown. An M13 sequencing ladder is a size standard. The positions and lengths of the primers and unprotected S1 probes are shown schematically; numbering is as in Fig. 1. Asterisks indicate the ^{32}P-labeled sites.

was presumably derived from such a transcript. The S1 nuclease protection experiments with poly(A)$^+$ RNA from 46,XY cells indicated that the *RPS4X* mRNA was approximately 10-fold more abundant than the *RPS4Y* mRNA. This difference was not seen with primer extension, but the much weaker signal implies that the 20-mer oligonucleotides did not anneal quantitatively. Northern (RNA) analysis also suggested that in tissues the *RPS4X* transcript is more abundant than that of *RPS4Y* (13).

The 5′ flanking sequences of mouse *Rps4*, human *RPS4X*, and human *RPS4Y* are aligned in Fig. 3. The 5′ untranslated sequences of human *RPS4Y* and mouse *Rps4* are 23 nucleotides in length; the human *RPS4X* mRNA has one additional nucleotide. All three genes adhere to the rule for mammalian ribosomal proteins that transcription initiates at a cytosine residue in a pyrimidine-rich tract. The human *RPS4X* and mouse *Rps4* sequences show extensive similarity for at least 100 bp upstream of the start site, while *RPS4Y* diverges past nucleotide −18. None of the genes has a canonical TATA sequence at the −30 position. The sequence (A/G)CCG GAA(A/G) is present near the transcription start site in all three genes. The same sequence is found in the mouse L32 gene at position −78 (16), where it binds an ETS domain-containing transcription factor (45). This sequence is also present in the opposite orientation at position −71 in the mouse L7 gene (23) and at −51 in the mouse L30 gene (16). A closely related sequence is repeated twice in the promoter region of the human and mouse small nuclear ribonucleoprotein E genes (10). A consensus Sp1 recognition site (20) is present in the human *RPS4X* and mouse *Rps4* genes at position −74 but absent in human *RPS4Y*.

Chicken S4 gene. For comparison, we examined the S4 gene of a nonmammalian vertebrate, the chicken. The sequence of the insert of cDNA pDP1435 is shown in Fig. 4. Two silent nucleotide substitutions were identified among two additional

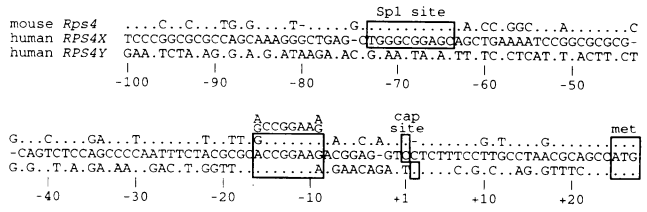

FIG. 3. Alignment of sequences of mammalian S4 promoter regions. Dots indicate nucleotide sequence identity to human *RPS4X*; dashes are insertion or deletions. Numbering is according to the human *RPS4X* transcript and is relative to the transcription start site (+1). Sp1 sites, an (A/G)CCGGAA(A/G) motif, major transcription start sites, and initiator methionine codons are boxed.

```
                                                                    -4  AGCC
  1  ATGGCCCGCGGCCCGAAGAAGCACCTGAAGCGCGTGGCGGCGCCGAAGCACTGGATGCTGGACAAGCTGACGGGCGTCTTCGCGCCCCGTCCCTCCACGGGCCCTCACAAGCTGAGGGAA
  1  M  A  R  G  P  K  K  H  L  K  R  V  A  A  P  K  H  W  M  L  D  K  L  T  G  V  F  A  P  R  P  S  T  G  P  H  K  L  R  E

121  TGCCTCCCGCTGATCATCTTCCTGCGGAACAGGCTGAAGTACGCCCTGACCGGAGATGAGGTGAAGAAGATTTGCATGCAGAGGTTCATCAAGATCGACGGCAAAGTTCGCACCGACATC
 41  C  L  P  L  I  I  F  L  R  N  R  L  K  Y  A  L  T  G  D  E  V  K  K  I  C  M  Q  R  F  I  K  I  D  G  K  V  R  T  D  I

241  ACCTACCCTGCGGGCTTCATGGATGTCATCAGCATTGAGAAGACAGGGGAACACTTCCGCTTGGTGTACGACACCAAGGGCCGGTTTGCTGTTCACCGCATCACAGCTGAGGAGGCCAAG
 81  T  Y  P  A  G  F  M  D  V  I  S  I  E  K  T  G  E  H  F  R  L  V  Y  D  T  K  G  R  F  A  V  H  R  I  T  A  E  E  A  K

361  TACAAGCTGTGCAAGGTGAGGAAGATCTTTGTGGGCACCAAAGGAATCCCTCACCTGGTCACCCACGATGCCCGCACCATCCGCTATCCGGATCCCCTCATCAAGGTGAACGATACGATC
121  Y  K  L  C  K  V  R  K  I  F  V  G  T  K  G  I  P  H  L  V  T  H  D  A  R  T  I  R  Y  P  D  P  L  I  K  V  N  D  T  I

                    C                                                                                   T
481  CAGATTGACCTGGAGACGGGCAAGATCACAGACTTCATCAAGTTTGACACAGGGAACCTGTGCATGGTGACTGGCGGTGCTAACTTGGGCCGGATCGGGGTGATCACCAACCGGGAGAGA
161  Q  I  D  L  E  T  G  K  I  T  D  F  I  K  F  D  T  G  N  L  C  M  V  T  G  G  A  N  L  G  R  I  G  V  I  T  N  R  E  R

601  CACCCTGGCTCCTTCGATGTGGTTCATGTGAAGGATGCCAATGGCAACAGCTTTGCCACCAGGCTCTCCAACATCTTCGTTATTGGCAAGGGCAACAAGCCATGGATCTCCCTGCCTCGT
201  H  P  G  S  F  D  V  V  H  V  K  D  A  N  G  N  S  F  A  T  R  L  S  N  I  F  V  I  G  K  G  N  K  P  W  I  S  L  P  R

721  GGAAAGGGCATCCGCCTGACCATCGCCGAAGAGAGAGATAAGAGACTGGCAGCCAAACAGAGCAGCGGGTAAACTGCAGCTTGGGCCGGCTGTGGTCCTCAAAGTATGTCCAATTAAATT
241  G  K  G  I  R  L  T  I  A  E  E  R  D  K  R  L  A  A  K  Q  S  S  G  *  263

841  TTATTACCAAAAAAAAAAAAAAAA 864
```

FIG. 4. Nucleotide sequence and predicted translation product of chicken S4 cDNA pDP1435. Numbering of nucleotides and amino acids begins with the initiator methionine codon. Silent nucleotide substitutions identified among two additional cDNAs are shown above the sequence.

cDNAs and presumably represent polymorphisms. Overall, the chicken coding sequence is slightly more similar to human *RPS4X* (83.8%) than to *RPS4Y* (81.7%). As a rule, the chicken genome lacks processed pseudogenes, and probing with the cDNA revealed only a single *Eco*RI restriction fragment in a Southern blot of male or female chicken genomic DNA (Fig. 5). In other hybridizations (not shown), the signal intensity did not vary systematically between roosters (ZZ) and hens (ZW), nor were any sex-specific restriction fragments observed with use of several additional enzymes. The chicken S4 gene is therefore unlikely to be linked to the W or Z sex chromosome.

The sequence of the chicken S4 protein differs from that of human S4X (and rodent S4) by four amino acid substitutions (Fig. 6A). All four substitutions fall within the sequence encoded by exon 4 of the three mammalian genes. Surprisingly, three of the four substitutions are also present in human S4Y. To determine whether these amino acids lie within a homologous exon in chicken, we used PCR to amplify chicken genomic DNA, using primers from flanking cDNA sequences, and then sequenced the unique product. Examination of this sequence revealed that the chicken gene indeed contains a 98-bp exon corresponding precisely to exon 4 of the mammalian S4 genes (Fig. 6B).

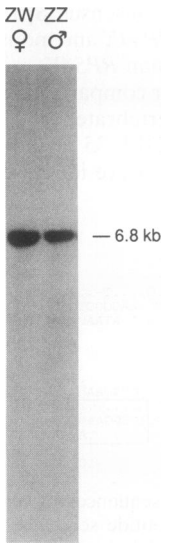

FIG. 5. Southern blot of chicken genomic DNA digested with *Eco*RI and hybridized with the chicken S4 cDNA insert. A single fragment is detected in both hen (ZW) and rooster (ZZ) DNA.

Expression of human S4 isoforms. S4 was originally identified by two-dimensional gel electrophoresis of proteins from rat liver 40S ribosomal subunits (37). The electrophoretic patterns of ribosomal proteins from humans and rats are virtually identical. To test whether S4X and S4Y are both present in ribosomes of human males, proteins from highly purified 40S ribosomal subunits from human male placenta were fractionated by two-dimensional gel electrophoresis, visualized with Ponceau S, transferred to nitrocellulose, and probed with isoform-specific antisera (41). As expected, immunoblotting with the anti-S4X serum confirmed that the spot previously assigned as S4 indeed contains the S4X protein (Fig. 7, arrow). Several lower-molecular-weight spots that did not correspond to major ribosomal proteins were also detected; these presumably represent S4X proteolytic products (see also Fig. 8).

The anti-S4Y serum detected the same major S4 spot as well as minor lower-molecular-weight products. S4Y is predicted to be slightly more basic than S4X (net charge difference of about +1 at pH 8.6); the slight difference in electrophoretic mobility of the two isoforms (observed more clearly in proteolytic products) was consistent with this charge difference. The 40S ribosomal proteins were purified from polysomes, implying that both S4X and S4Y are present in translationally active ribosomes.

We used sodium dodecyl sulfate-PAGE and immunoblotting to estimate the relative abundance of S4X and S4Y in human male placenta and to examine the effect of gene dosage on the levels of the two isoforms (Fig. 8). Duplicate gels containing 40S ribosomal proteins from male placenta or 80S ribosomal proteins from cultured lymphoblastoid cell lines

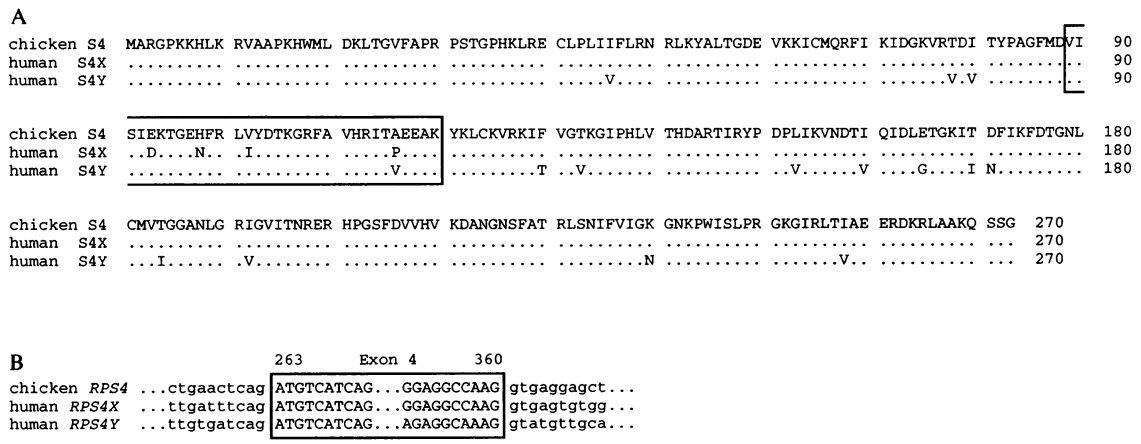

FIG. 6. (A) Alignment of chicken and human S4 amino acid sequences. Dots indicate identity to chicken S4. Amino acids encoded by one exon (exon 4 in the human genes) are boxed. (B) Splice sites surrounding exon 4 of human *RPS4X* and *RPS4Y* and the homologous exon of chicken. Exon sequences are capitalized; intron sequences are in lowercase; numbering is as in Fig. 1.

were immunoblotted with the anti-S4X or anti-S4Y serum. As standards for quantitation, equal amounts (as judged by Coomassie staining of serial dilutions) of S4X and S4Y expressed in *E. coli* (41) were also loaded. The amount of S4X or S4Y in each sample relative to the amount in the corresponding *E. coli* extract was estimated by densitometry. The results indicate that there was about 15% as much S4Y as S4X in samples from either placenta or 46,XY lymphoblastoid cells. S4Y was barely detectable in 49,XXXXY cells. In 49,XYYYY cells, S4Y was about 45% as abundant as S4X. Reprobing of filters with antiserum to S14, another small subunit protein, confirmed that the loads of ribosomal proteins in the different lymphoblastoid samples were approximately equal.

DISCUSSION

S4 appears to be unique among mammalian ribosomal proteins in that two genes encode isoforms, at least in humans. Nonetheless, human *RPS4X* and *RPS4Y* and mouse *Rps4* are for the most part typical of ribosomal protein genes. Their promoters lack a TATA box; transcription initiates at a cytosine residue in a polypyrimidine sequence; there is a short 5′ untranslated sequence; and an intron is present near the 5′ end of the gene. One difference is that human *RPS4Y* is unusually large compared with other known ribosomal protein genes.

Several other X-Y genes have been discovered in the strictly sex-linked portions of mammalian sex chromosomes. In each case, the X- and Y-linked genes encode closely related but nonidentical proteins, and, where studied, the genes have similar structures. The *ZFX* and *ZFY* genes encode zinc finger transcriptional activators and are present on the sex chromosomes of all eutherian mammals examined (29). The human *AMGX* and *AMGY* genes encode tooth extracellular matrix proteins (33). Mouse *Ube1x* and *Ube1y* encode isoforms of ubiquitin-activating enzyme (21, 24).

Human X-Y genes like *RPS4X* and *RPS4Y* may have diverged from allelic copies present on the ancestral autosome from which the mammalian X and Y chromosomes are derived. The fact that no gene has yet been found that is sex linked in both mammals and birds is taken as evidence that the avian and mammalian sex chromosomes evolved independently from different ancestral autosomes. Consistent with this notion, our data indicate that in chickens, unlike mammals, S4 is encoded by a single gene that is most likely autosomal. The chicken S4 gene encodes a protein identical to human S4X and to mouse S4 except for four amino acid residues. All four residues are encoded by a single exon in both chicken and

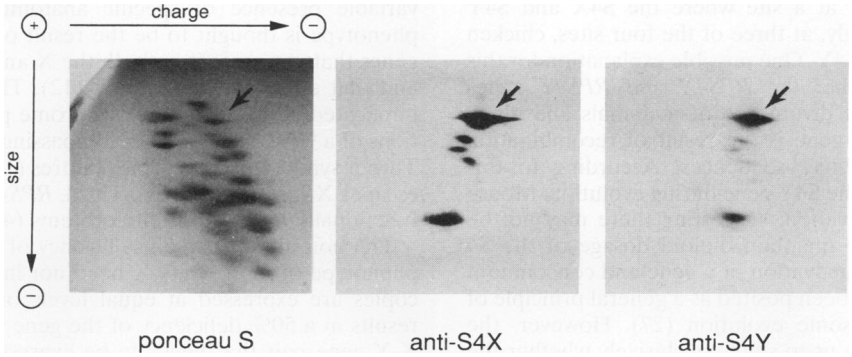

FIG. 7. Both S4X and S4Y are present in human male ribosomes. Male placental 40S subunit proteins were resolved by two-dimensional gel electrophoresis (left). The spot corresponding to S4 is indicated (arrow). The same spot was detected with anti-S4X (middle) or anti-S4Y (right) serum. Note also the slight difference in horizontal migration of the S4X and S4Y proteolytic products.

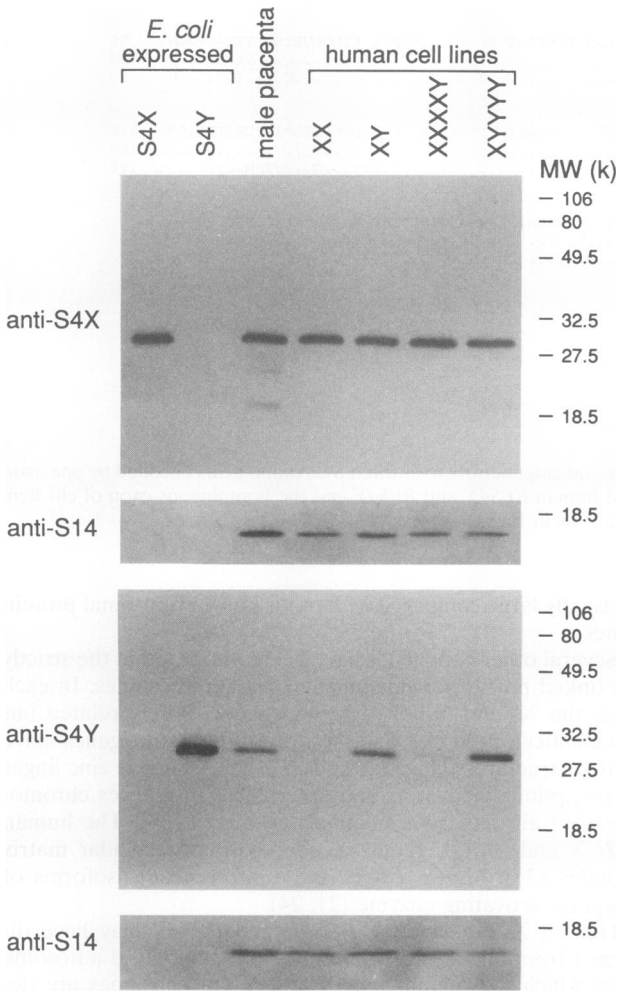

FIG. 8. Relative amounts of S4X and S4Y isoforms in ribosomes. Duplicate Western blots (immunoblots) containing 140 ng of *E. coli*-expressed S4X or S4Y, 5 μg of placental 40S ribosomal protein, or 10 μg of 80S ribosomal protein from human lymphoblastoid cell lines from a normal 46,XX female (WHT1660), a normal 46,XY male (WHT1659), 49,XXXXY male GM1202 (NIGMS Human Mutant Cell Repository), or a 49,XYYYY male (38) were probed with anti-S4X or anti-S4Y serum. Blots were stripped and reprobed with anti-S14 serum.

mammals, and each lies at a site where the S4X and S4Y proteins differ. Surprisingly, at three of the four sites, chicken S4 is identical to human S4Y. One possible explanation for this observation would be that the *RPS4X* and *RPS4Y* genes diverged even before the divergence of mammals and birds, and the present chicken gene is the result of recombination between S4X- and S4Y-like sequences. According to this scenario, the mouse lost the S4Y gene during evolution. Mouse *Rps4* undergoes X inactivation, suggesting there may not be any selective pressure to maintain diploid dosage of the S4 gene in this species. X inactivation of a gene and concomitant loss of its Y homolog has been posited as a general principle of mammalian sex chromosome evolution (27). However, the present data do not allow us to state conclusively whether the common ancestor of human and mouse had one or two S4 genes.

As evidenced by S1 nuclease protection and Northern blotting, the *RPS4Y* transcript is only about 10% as abundant as the *RPS4X* transcript. Lower abundance of the *RPS4Y* mRNA could be due to differences in the rate of transcription, mRNA stability, or both. The absence of an Sp1 consensus site in the *RPS4Y* promoter region may be significant in this regard. The greater length of the *RPS4Y* transcription unit could also play a role in its lower level of expression. Interestingly, the steady-state mRNA level of another Y-linked gene, *AMGY*, is also about 10% of that of its X-linked counterpart (33).

The relative abundance of the S4X and S4Y proteins reflects their steady-state mRNA levels. Assuming that (i) individual *RPS4X* and *RPS4Y* genes are transcribed with a fixed ratio and (ii) the S4X and S4Y mRNAs are translated and the proteins are incorporated into ribosomes (and thereby protected from degradation) with similar efficiencies, we would expect the ratio of S4X to S4Y to be about 10:1 in 46,XY cells, 40:1 in 49,XXXXY cells, and 10:4 in 49,XYYYY cells. The levels that we observed are in reasonable agreement with these predictions. The difference in the levels of S4X and S4Y proteins may be due simply to the difference in the levels of *RPS4X* and *RPS4Y* transcripts, in turn most likely due to a difference in their transcription rates.

The S4Y protein is present in translationally active ribosomes from normal human male placental tissue and cultured cells. S4Y is thus a true isoform of mammalian ribosomal protein S4. These results confirm the conjecture that some ribosomes are structurally distinct in human males and females (13). The consequences, if any, of this sexual dimorphism are unknown. The fact that S4Y is present in only a minority of ribosomes may explain the failure to detect S4 isoforms in the past. Although S4Y is normally a minor species, its abundance is increased in 49,XYYYY cells without any obvious cellular phenotype. Thus, S4Y appears to be functionally interchangeable with S4X. These data agree with our recent finding that in transfection studies, *RPS4Y* complemented a temperature-sensitive hamster *Rps4* mutation (41). At the organismal level, several individuals with the 49,XYYYY karyotype have been reported (26, 31, 36, 38), one of whom was the source of the cell line used in this study. The 49,XYYYY syndrome includes mental retardation and dysmorphic features such as facial anomalies, brachydactyly, and radioulnar synostosis. While S4X and S4Y appear to be interchangeable in vitro, we cannot rule out the possibility that overexpression of *RPS4Y* contributed to the phenotype of these individuals.

In contrast, there is genetic evidence to suggest that underexpression of S4 plays a role in Turner syndrome, the phenotype associated with monosomy X in humans. The cardinal features of the Turner syndrome phenotype include frequent intrauterine lethality, short stature, gonadal failure, and the variable presence of specific anatomic abnormalities. The phenotype is thought to be the result of haploinsufficiency of genes that are present on both the X and the Y chromosomes and that escape X inactivation (12). Three lines of evidence implicated S4 in the Turner syndrome phenotype. First, deletions of a 90-kb Y interval encompassing *RPS4Y* correlate with Turner syndrome phenotypic features (13, 28). Second, *RPS4X* escapes X inactivation (13). Third, *RPS4X* and *RPS4Y* encode functionally interchangeable proteins (41).

The notion that haploinsufficiency of X-Y genes causes the phenotype of monosomy X need not imply that the X and Y copies are expressed at equal levels or that haploid dosage results in a 50% deficiency of the gene products. Consider an X-Y gene pair that needs to be expressed at some minimum level for normal development. Each X gene copy must produce at least 50% of that level to ensure that a 46,XX individual would have 100%. If a single X gene copy produces more than

50% of the minimum level, then the Y gene can produce correspondingly less, and the deficiency associated with monosomy X would be less than 50%. This appears to be the case for the *RPS4X* and *RPS4Y* genes. In normal males, the single copy of *RPS4X* supplies 85 to 90% of the S4 mRNA and protein, while *RPS4Y* supplies only about 10 to 15%. We predict that an embryo with either a deletion of *RPS4Y* or an XO karyotype would have at most a 10 to 15% deficiency of S4 mRNA. The biochemical consequence of this deficiency would depend on whether the S4 mRNA becomes rate limiting for ribosome assembly. In most tissues, ribosomal protein genes are transcribed in excess and translationally regulated. Maximal translation of ribosomal protein mRNAs has been observed only in certain circumstances, such as in regenerating hepatocytes or during early development (22). *RPS4X* and *RPS4Y* both have a pyrimidine-rich motif at the 5′ end of the mRNAs that has been shown to mediate translational regulation of other ribosomal protein genes. Increased translation of *RPS4X* mRNA may help to compensate for the absence of *RPS4Y*, so that the rate of S4 synthesis and ribosome assembly may be minimally affected. We have failed to detect any marked alteration in polysome profiles or ribosomal subunit ratios in cultured cells with deletions of *RPS4Y* (unpublished observations), but this could be due to the assay's lack of sensitivity.

Lethality associated with monosomy X occurs most often early in gestation (17). Two copies of the S4 gene may be needed at early development stages when the rate of ribosome synthesis is greatest. A better understanding of the phenotypic consequences of ribosomal protein haploinsufficiency in mammals will require development of an experimental system such as targeted disruption of a ribosomal protein gene in the mouse.

ACKNOWLEDGMENTS

We thank Lin Zuo and Ellson Chen for communicating unpublished nucleotide sequence for human *RPS4X* and *RPS4Y*, David Dietcher for the chicken cDNA library, Steven Bloom for chicken blood samples, Donald Roufa for anti-S14 serum, and Mark Behlke, Naoya Kenmochi, and Renee Reijo for helpful comments on the manuscript.

This work was supported by the Howard Hughes Medical Institute, the Searle Scholars Program/Chicago Community Trust (D.C.P.), and Damon Runyon-Walter Winchell Cancer Research Fund Fellowship DRG-052 (A.R.Z.).

REFERENCES

1. **Antoine, M., and M. Fried.** 1992. The organization of the intron-containing human S6 ribosomal protein (rpS6) gene and determination of its location at chromosome 9p21. Hum. Mol. Genet. **1**:565–570.
2. **Ashworth, A., S. Rastan, R. Lovell-Badge, and G. Kay.** 1991. X-chromosome inactivation may explain the difference in viability of XO humans and mice. Nature (London) **351**:406–408.
3. **Chan, Y. L., and I. G. Wool.** 1991. The structure of a gene containing introns and encoding rat ribosomal protein P2. Nucleic Acids Res. **19**:4895–4900.
4. **Chen, I. T., and D. J. Roufa.** 1988. The transcriptionally active human ribosomal protein S17 gene. Gene **70**:107–116.
5. **Colombo, P., J. Yon, and M. Fried.** 1991. The organization and expression of the human L7a ribosomal protein gene. Biochim. Biophys. Acta **1129**:93–95.
6. **Davies, B., S. Feo, E. Heard, and M. Fried.** 1989. A strategy to detect and isolate an intron-containing gene in the presence of multiple processed pseudogenes. Proc. Natl. Acad. Sci. USA **86**:6691–6695.
7. **Davies, B., and M. Fried.** 1993. The structure of the human intron-containing S8 ribosomal protein gene and determination of its chromosomal location at 1p32-p34.1. Genomics **15**:68–75.
8. **De Falco, S., G. Russo, A. Angiolillo, and C. Pietropaolo.** 1993. Human L7a ribosomal protein sequence, structural organization, and expression of a functional gene. Gene **126**:227–235.
9. **Dudov, K. P., and R. P. Perry.** 1984. The gene family encoding the mouse ribosomal protein L32 contains a uniquely expressed intron-containing gene and an unmutated processed gene. Cell **37**:457–468.
10. **Fautsch, M. P., M. A. Thompson, E. L. Holicky, P. J. Schultz, J. B. Hallett, and E. D. Wieben.** 1992. Conservation of coding and transcriptional control sequences within the snRNP E protein gene. Genomics **14**:883–890.
11. **Feo, S., B. Davies, and M. Fried.** 1992. The mapping of seven intron-containing ribosomal protein genes shows they are unlinked in the human genome. Genomics **13**:201–207.
12. **Ferguson-Smith, M. A.** 1965. Karyotype-phenotype correlations in gonadal dysgenesis and their bearing on the pathogenesis of malformations. J. Med. Genet. **2**:142–155.
13. **Fisher, E. M., P. Beer-Romero, L. G. Brown, A. Ridley, J. A. McNeil, J. B. Lawrence, H. F. Willard, F. R. Bieber, and D. C. Page.** 1990. Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. Cell **63**:1205–1218.
14. **Hammond, M. L., W. Merrick, and L. H. Bowman.** 1991. Sequences mediating the translation of mouse S16 ribosomal protein mRNA during myoblast differentiation and in vitro and possible control points for the in vitro translation. Genes Dev. **5**:1723–1736.
15. **Hamvas, R. M., A. Zinn, J. T. Keer, E. M. Fisher, R. P. Beer, S. D. Brown, and D. C. Page.** 1992. *Rps4* maps near the inactivation center on the mouse X chromosome. Genomics **12**:363–367.
16. **Hariharan, N., D. E. Kelley, and R. P. Perry.** 1989. Equipotent mouse ribosomal protein promoters have a similar architecture that includes internal sequence elements. Genes Dev. **3**:1789–1800.
17. **Hassold, T. J.** 1986. Chromosome abnormalities in human reproductive wastage. Trends Genet. **2**:105–110.
18. **Huxley, C., and M. Fried.** 1990. The mouse rpL7a gene is typical of other ribosomal protein genes in its 5′ region but differs in being located in a tight cluster of CpG-rich islands. Nucleic Acids Res. **18**:5353–5357.
19. **Huxley, C., T. Williams, and M. Fried.** 1988. One of the tightly clustered genes of the mouse surfeit locus is a highly expressed member of a multigene family whose other members are predominantly processed pseudogenes. Mol. Cell. Biol. **8**:3898–3905.
20. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. **11**:20–23.
21. **Kay, G. F., A. Ashworth, G. D. Penny, M. Dunlop, S. Swift, N. Brockdorff, and S. Rastan.** 1991. A candidate spermatogenesis gene on the mouse Y chromosome is homologous to ubiquitin-activating enzyme E1. Nature (London) **354**:486–489.
22. **Levy, S., D. Avni, N. Hariharan, R. P. Perry, and O. Meyuhas.** 1991. Oligopyrimidine tract at the 5′ end of mammalian ribosomal protein mRNAs is required for their translational control. Proc. Natl. Acad. Sci. USA **88**:3319–3323.
23. **Meyuhas, O., and A. Klein.** 1990. The mouse ribosomal protein L7 gene. Its primary structure and functional analysis of the promoter region. J. Biol. Chem. **265**:11465–11473.
24. **Mitchell, M. J., D. R. Woods, P. K. Tucker, J. S. Opp, and C. E. Bishop.** 1991. Homology of a candidate spermatogenic gene from the mouse Y chromosome to the ubiquitin-activating enzyme E1. Nature (London) **354**:483–486.
25. **Monk, R., O. Meyuhas, and R. Perry.** 1981. Mammals have multiple genes for individual ribosomal proteins. Cell **24**:301–306.
26. **Noel, B., M. Bénézech, M. T. Bouzon, and D. Coudrat.** 1988. A 7-year-old boy with the 49,XYYYY syndrome. Ann. Genet. (Paris) **31**:111–116.
27. **Ohno, S.** 1967. Sex chromosomes and sex-linked genes. Springer-Verlag, New York.
28. **Page, D. C., E. M. Fisher, B. McGillivray, and L. G. Brown.** 1990. Additional deletion in sex-determining region of human Y chromosome resolves paradox of X,t(Y;22) female. Nature (London) **346**:279–281.
29. **Page, D. C., R. Mosher, E. M. Simpson, E. M. Fisher, G. Mardon, J. Pollack, B. McGillivray, A. de la Chapelle, and L. G. Brown.** 1987. The sex-determining region of the human Y chromosome

encodes a finger protein. Cell **51**:1091–1104.
30. **Pata, I., S. Hoth, J. Kruppa, and A. Metspalu.** 1992. The human ribosomal protein S6 gene: isolation, primary structure and location in chromosome 9. Gene **121**:387–392.
31. **Plauchu, H., C. Charrin, and J. C. Kossman.** 1984. The 49,XYYYY syndrome: apropos of a case detected at birth and followed for 2 1/2 years. J. Genet. Hum. **32**:299–306.
32. **Rhoads, D. D., A. Dixit, and D. J. Roufa.** 1986. Primary structure of human ribosomal protein S14 and the gene that encodes it. Mol. Cell. Biol. **6**:2774–2783.
33. **Salido, E. C., P. H. Yen, K. Koprivnikar, L. C. Yu, and L. J. Shapiro.** 1992. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. Am. J. Hum. Genet. **50**:303–316.
34. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
35. **Schneider-Gädicke, A., P. Beer-Romero, L. G. Brown, R. Nussbaum, and D. C. Page.** 1989. *ZFX* has a gene structure similar to ZFY, the putative human sex determinant, and escapes X inactivation. Cell **57**:1247–1258.
36. **Shanske, A., J. Rinsky, R. Kazi, and D. Warburton.** 1991. An adult with 49,XYYYY karyotype, abstr. 1510, p. 273. Abstr. 41st Annu. Meet. Am. Soc. Hum. Genet. 1991.
37. **Sherton, C. C., and I. G. Wool.** 1972. Determination of the number of proteins in liver ribosomes and ribosomal subunits by two-dimensional polyacrylamide gel electrophoresis. J. Biol. Chem. **247**:4460–4467.
38. **Sirota, L., Y. Zlotogora, F. Shabtai, I. Halbrecht, and E. Elian.** 1981. 49,XYYYY. A case report. Clin. Genet. **19**:87–93.
39. **Towbin, J., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.
40. **Wagner, M., and R. P. Perry.** 1985. Characterization of the multigene family encoding the mouse S16 ribosomal protein: strategy for distinguishing an expressed gene from its processed pseudogene counterparts by an analysis of total genomic DNA. Mol. Cell. Biol. **5**:3560–3576.
41. **Watanabe, M., A. R. Zinn, D. C. Page, and T. Nishimoto.** 1993. Functional equivalence of human X- and Y-encoded isoforms of ribosomal protein S4 consistent with a role in Turner syndrome. Nat. Genet. **4**:268–271.
42. **Wiedemann, L. M., and R. P. Perry.** 1984. Characterization of the expressed gene and several processed pseudogenes for the mouse ribosomal protein L30 gene family. Mol. Cell. Biol. **4**:2518–2528.
43. **Wool, I. G.** 1979. The structure and function of eukaryotic ribosomes. Annu. Rev. Biochem. **48**:719–754.
44. **Wool, I. G., Y. L. Chan, V. Paz, and J. Olvera.** 1990. The primary structure of rat ribosomal proteins: the amino acid sequences of L27a and L28 and corrections in the sequences of S4 and S12. Biochim. Biophys. Acta **1050**:69–73.
45. **Yoganathan, T., N. K. Bhat, and B. H. Sells.** 1992. A positive regulator of the ribosomal protein gene, β factor, belongs to the ETS oncoprotein family. Biochem. J. **287**:349–353.
46. **Zinn, A. R., S. L. Bressler, P. Beer-Romero, D. A. Adler, V. M. Chapman, D. C. Page, and C. M. Disteche.** 1991. Inactivation of the *Rps4* gene on the mouse X chromosome. Genomics **11**:1097–1101.
47. **Zuo, L., E. Chen, L. G. Brown, and D. C. Page.** Unpublished data.