Cell, Vol. 51, 1091–1104, December 24, 1987, Copyright © 1987 by Cell Press

The Sex-Determining Region of the Human Y Chromosome Encodes a Finger Protein

David C. Page,* Rebecca Mosher,* Elizabeth M. Simpson,* Elizabeth M. C. Fisher,* Graeme Mardon,*† Jonathan Pollack,* Barbara McGillivray,‡ Albert de la Chapelle,§ and Laura G. Brown*

* Whitehead Institute for Biomedical Research
Nine Cambridge Center
Cambridge, Massachusetts 02142
† Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139
‡ University of British Columbia Clinical Genetics Unit
4490 Oak Street
Vancouver, British Columbia V6H 3V5
Canada
§ Department of Medical Genetics
University of Helsinki
00290 Helsinki
Finland

Summary

The presence or absence of the Y chromosome determines whether a mammalian embryo develops as a male or female. In humans, genetic deletion analysis of "sex-reversed" individuals has identified a small portion of the Y chromosome necessary and sufficient to induce testicular differentiation of the bipotential gonad. We report the cloning of a 230-kilobase segment of the human Y chromosome that contains some or all of the testis-determining factor gene (*TDF*), the master sex-determining locus. The cloned region spans the deletion in a female who carries all but 160 kilobases of the Y. Certain DNA sequences within this region were highly conserved during evolution; homologs occur on the Y chromosomes of all mammals examined. In particular, homologous sequences are found within the sex-determining region of the mouse Y chromosome. The nucleotide sequence of this conserved DNA on the human Y chromosome suggests that it encodes a protein with multiple "finger" domains, as first described in frog transcription factor IIIA. The encoded protein probably binds to nucleic acids in a sequence-specific manner, and may regulate transcription. Very similar DNA sequences occur on the X chromosome of humans and other mammals. We discuss the possibility that the Y-encoded finger protein is the testis-determining factor, and propose models of sex determination accommodating the finding of a related locus on the X chromosome. The presence of similar sequences in birds suggests a possible role not only in the XX/XY sex determination system of mammals, but also in the ZZ/ZW system of birds.

Introduction

The mechanism by which the sex of an individual is determined has been a subject of scientific speculation since the time of Aristotle. Until 1900, it was generally thought that the sex of a human embryo is decided by environmental factors, such as maternal nutrition. Since the rediscovery of Mendel's findings, it has been shown that in many, though not all, vertebrate and invertebrate species the sex of an individual is determined by the individual's chromosomal constitution. Indeed, sexual dimorphism in insects was among the first traits shown to have a chromosomal basis (McClung, 1902; Stevens, 1905; Wilson, 1905; Morgan, 1910). In Drosophila, sex is determined by the number of X chromosomes: regardless of the presence or absence of the Y chromosome, diploid fly embryos with two X chromosomes (XX or XXY) develop as females. Diploid embryos with one X (XY or XO) develop as males (Bridges, 1916).

The existence of X and Y chromosomes in humans was demonstrated cytologically in 1923 (Painter, 1923). It was assumed, by analogy to Drosophila, that sex in mammals is decided by the number of X chromosomes. In 1959, however, studies of humans and mice with abnormal sex chromosomal constitutions (XXY and XO) revealed the critical role of the Y chromosome: regardless of the number of X chromosomes, mammalian embryos carrying a Y chromosome (XY or XXY) develop as males (Jacobs and Strong, 1959). Embryos lacking a Y chromosome (XX or XO) develop as females (Ford et al., 1959; Welshons and Russell, 1959).

The mammalian Y chromosome, by its presence or absence, constitutes a binary switch upon which hinge all sexually dimorphic characteristics. The execution of this binary decision, i.e., sex differentiation, is a complex physiologic process which must involve the products of many genes, some autosomal. The fates of the rudimentary internal accessory structures and external genitalia, for example, are decided by the hormonal output of the differentiated gonads, testes or ovaries (Jost, 1953). Even the differentiation of the bipotential gonad, where the effect of the Y chromosome is more direct, requires certain autosomal functions. For example, a mutation on chromosome 17 causes XY mouse embryos to develop ovaries (Washburn and Eicher, 1983). Many genes act in conjunction with or "downstream" of the Y chromosome in the pathways of mammalian sex differentiation.

There must exist on the Y chromosome one or more genes whose products, directly or indirectly, determine all aspects of sexual dimorphism. The biochemical nature of this putative "testis-determining factor" (TDF; referred to as Tdy in mouse) has been the subject of much speculation and inquiry. Virtually all biochemical differences between males and females have been considered possible clues to the identity of TDF. H-Y is a male-specific minor histocompatibility antigen originally defined by graft rejec-

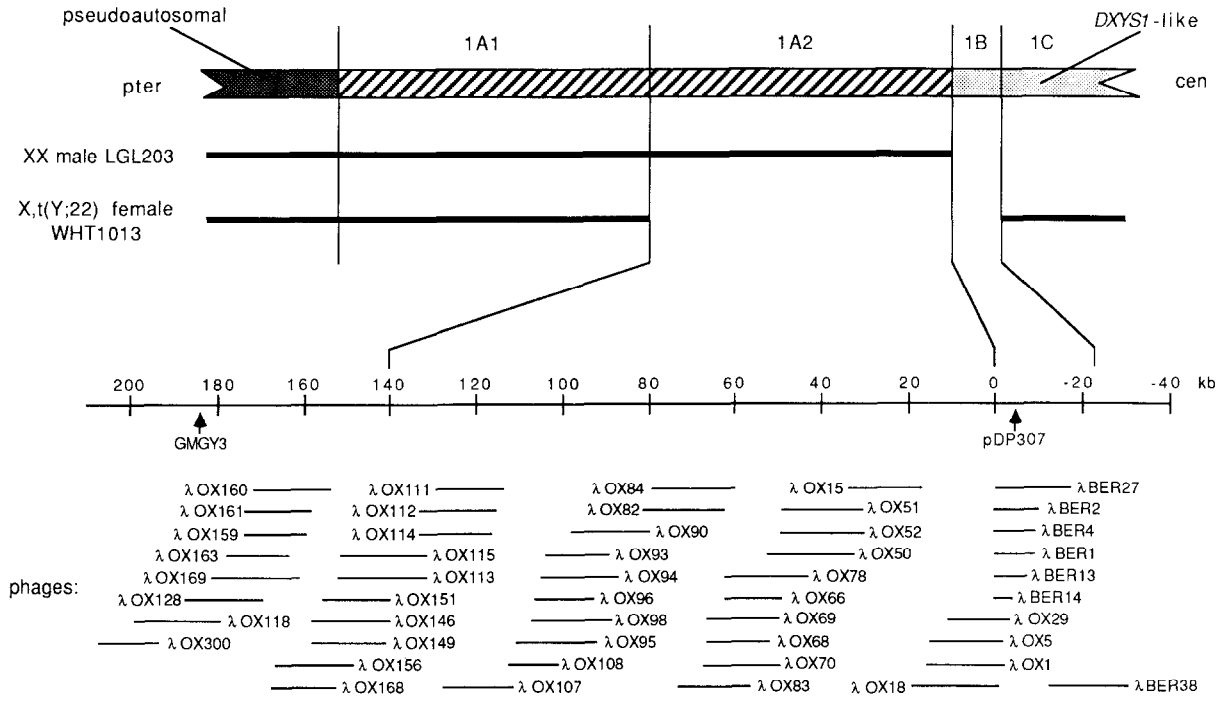

Figure 1. Chromosomal Walk of the Sex-Determining Region of the Human Y Chromosome

At top is a schematic representation of the distal short arm of the Y chromosome. The orientation with respect to the short-arm telomere (pter) and centromere (cen) is shown. The pseudoautosomal region is subject to recombination with the X chromosome during normal male meiosis. Intervals 1A1, 1A2, 1B, and 1C are defined by deletion analysis, and they show strictly sex-linked (as opposed to pseudoautosomal) inheritance. Sequences in intervals 1B and 1C are highly homologous to but do not recombine with sequences on the long arm of the X chromosome (see discussion of "*DXYS1*-like" sequences in text). Chromosomal deletions that demarcate the *TDF* locus are shown below the map of the chromosome. The black bars indicate the portions of the Y chromosome present in XX male LGL203 and X,t(Y;22) female WHT1013. At least part of the *TDF* gene must be located in interval 1A2, which is present in the XX male and absent in the X,t(Y;22) female. Shown below the deletions is a 230 kb block of DNA cloned from this region by chromosome walking. Distances from the chromosomal breakpoint in XX male LGL203 are shown in kilobases. Arrows indicate the starting points of the primary walk, clone pDP307, and of the secondary walk, clone GMGY3 (Affara et al., 1986). In all, 49 recombinant lambda phages with overlapping human inserts were isolated from genomic libraries. (The inserts of phages with the prefix "λBER" actually derive from the homologous, *DXYS1*-like region of the human X chromosome.) The single chromosomal breakpoint in LGL203 and the two chromosomal breakpoints in WHT1013 are located within the cloned region, as indicated. The breakpoint in XX male LGL203 corresponds roughly to the distal end of the *DXYS1*-like domain on Yp.

tion in mice (Eichwald and Silmser, 1955). It has been proposed and widely accepted that H-Y antigen is TDF (Wachtel et al., 1975). To be sure, H-Y antigen and TDF are both encoded or regulated by the Y chromosome. However, the *H-Y* and *TDF* genes map to different portions of the Y chromosome; they must be two different entities (McLaren et al., 1984; Simpson et al., 1987). Another model proposes that the *TDF* gene will be found among the "Bkm" DNA sequences, monotonous GATA and GACA repeats present in many copies on the sex chromosomes of some vertebrates (Epplen et al., 1983; Singh et al., 1984). However, the relative scarcity of Bkm sequences on the human Y chromosome casts doubt on this hypothesis (Kiel-Metzger et al., 1985).

We set out to characterize TDF with an approach that does not presuppose the nature of the gene or gene product. Despite our ignorance of the biochemical and cellular events regulated by TDF, we felt it possible to clone the *TDF* gene by precise determination of its chromosomal location. The natural occurrence of various structural abnormalities of the human Y chromosome made it possible to construct a deletion map by hybridization with Y-DNA probes, and to position *TDF* on this map (Vergnaud et al., 1986). Crucial to this deletion mapping of *TDF* was the presence of sex-determining Y chromatin in many "XX males" (Guellaen et al., 1984) and the deletion of that sex-determining portion of the Y chromosome in some "XY females" (Disteche et al., 1986), as assayed by Y-DNA hybridization. Y-DNA deletion analysis of these and other individuals suggests that a small portion of the short arm of the human Y chromosome ("interval 1") is necessary and sufficient to induce testicular differentiation of the bipotential gonad (Page, 1986).

Results

Chromosomal Walk of the Sex-Determining Region of Yp

By extending this deletion analysis, the short arm of the Y chromosome (Yp) has been divided into 13 intervals, and their order has been determined (D. C. P. and colleagues, unpublished results). We constructed this map using 135 Y-DNA probes detecting, in total, 155 distinct Y-DNA loci. All but two of these 155 loci derive from the

euchromatic portion of the Y chromosome (the short arm, the centromere, and the proximal long arm; intervals 1 through 6; Vergnaud et al., 1986). As judged by its cytogenetic appearance, this euchromatic portion of the Y must be about 30 to 40 million bases in length.

If the deletion-mapped Y-DNA loci constitute a random sampling of the euchromatic portion of the Y, then it is likely that one locus is within a few hundred kilobases of *TDF*. Our deletion analysis unambiguously placed locus *DXYS42* (detected by probe pDP307) closest to *TDF* (D. C. P. and colleagues, unpublished results). Accordingly, we initiated a chromosomal walk from pDP307 in an effort to clone *TDF*. We identified recombinant phages with overlapping human inserts (Figure 1) by stepwise screening of human genomic libraries.

At first, we pursued the walk in both directions from pDP307, since the walk was not oriented with respect to either *TDF* or the centromere. We anticipated that crossing any one of three landmarks would provide an orientation: First, there were a few XX males in whom we had not detected any Y-specific DNA. If Y sequences obtained by walking were present in any such Y(-) XX male, i.e., if the Y-chromosomal breakpoint in such an XX male were crossed, it would suggest that we were moving toward *TDF* (and toward the Yp telomere; Petit et al., 1987; Page et al., 1987). Second, pDP307 derives from the "*DXYS1*-like" region, a large block of X-Y homology resulting from an Xq-to-Yp transposition during recent human evolution (Page et al., 1984; Geldwerth et al., 1985). It was apparent that the bulk of the *DXYS1*-like region is proximal to *TDF* on Yp (Vergnaud et al., 1986; D. C. P. and colleagues, unpublished results). Since the Y chromosome appears to be male-determining in all mammals, it seemed unlikely that *TDF* would be found among the *DXYS1*-like sequences, which have no homologs on the Y chromosomes of chimpanzees or gorillas. We suspected that *TDF* lies beyond the distal end of the *DXYS1*-like region, distal to pDP307. Third, pDP307 was the only Y-DNA probe detecting a deletion in WHT1013, a female with a Y;22 translocation (described below). If Y sequences obtained by walking from pDP307 were present in WHT1013, and if those sequences were *DXYS1*-like, we would deduce that we were moving away from *TDF* and toward the centromere.

All three landmarks were encountered within a 30-kilobase (kb) region surrounding pDP307, and they provided a consistent orientation with respect to *TDF* and the centromere. Autoradiograms documenting the crossing of these landmarks are shown in Figure 2. Walking "left" from pDP307 as shown in Figure 1, we crossed both the Y-chromosomal breakpoint in XX male LGL203 (case 6 in de la Chapelle et al., 1984) and the distal boundary of the *DXYS1*-like region. Prior to the walk, no Y DNA had been detected in the genome of XX male LGL203 (Vergnaud et al., 1986). This XX male lacks the Y-specific 3.4 kb PstI fragment detected by pDP307 (Figure 2, probe C). However, he carries the Y-specific 2.3-kb TaqI fragment detected by probe B, 12 kb to the left of pDP307 (Figure 2). Based on hybridizations with more than 50 probes, Y sequences to the left of probe B appear to be present in XX male LGL203, while those of the right of pDP307 (probe C) are absent. This XX male's Y-chromosomal breakpoint, which divides deletion intervals 1A2 and 1B, must lie between probes B and C. (Additional studies suggest that pDP307 is roughly 4 kb proximal to this breakpoint. This breakpoint will serve as the zero point from

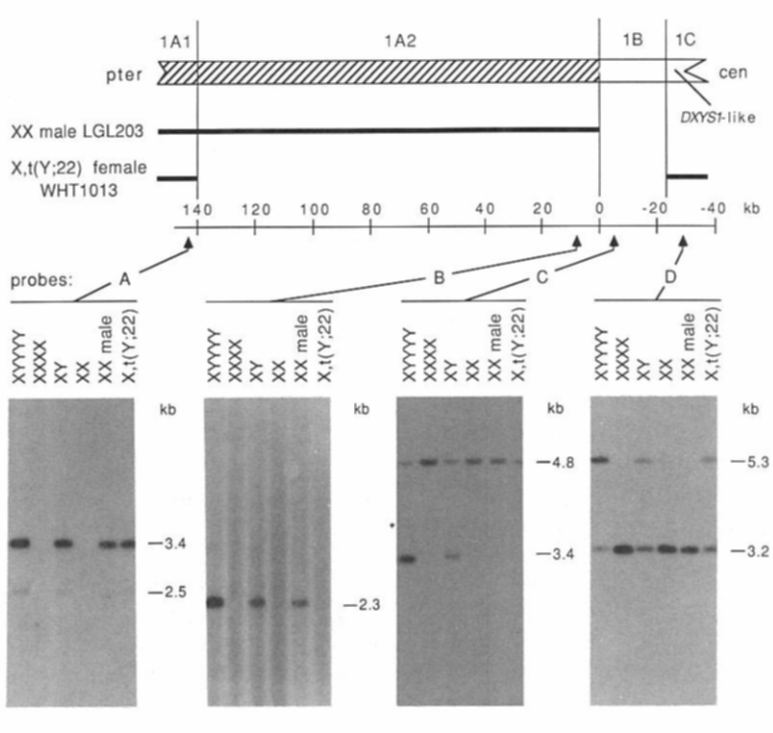

Figure 2. Deletion Analysis with DNA Probes from the Chromosomal Walk

At top is a representation of the distal short arm of the human Y chromosome and of the deletions that define the *TDF* locus. Distances from the Y-chromosomal breakpoint in XX male LGL203 are measured in kilobases. Probes (A) pDP1035, (B) 3.5 kb HindIII fragment purified from λOX18, (C) pDP307, and (D) 1.5 kb EcoRI fragment purified from λBER38 were hybridized to TaqI (A, B, D) or PstI (C) digests of DNAs from XYYYY male Oxen (Sirota et al., 1981), XXXX female GM1416, a normal male, a normal female, XX male LGL203, and X,t(Y;22) female WHT1013. The sizes in kilobases of the hybridizing restriction fragments are indicated. Probes A, B, and D: preannealing with human DNA, 47°C hybridization, 65°C wash; probe C: 47°C hybridization, 60°C wash.

which distances are measured; Figures 1 and 2.) Given that XX males carry terminal portions of the short arm of the Y chromosome (Petit et al., 1987; Page et al., 1987), these observations orient the walk: to the left is the Yp telomere, and to the right is the centromere (Figures 1 and 2).

The distal end of the *DXYS1*-like domain is near the Y-chromosomal breakpoint in XX male LGL203. As is typical of *DXYS1*-like sequences, pDP307, like probes proximal to it, detects not only a single-copy sequence on the Y chromosome but also a single-copy sequence on the X chromosome. The intensity of the 4.8 kb PstI fragment correlates well with the number of X chromosomes (Figure 2, probe C). In contrast, probe B, like most sequences distal to it, does not cross-hybridize to the X chromosome (Figure 2). Like the breakpoint in XX male LGL203, the distal end of the *DXYS1*-like region falls between probes B and C. (The *DXYS1*-like character of pDP307 and the sequences proximal to it will be more fully documented elsewhere). Precise definition of the *DXYS1*-like and LGL203 breakpoints will require comparative restriction mapping and nucleotide sequencing of the normal and translocated Y chromosomes.

Walking "right" from pDP307, we crossed a Y-chromosomal breakpoint in female WHT1013, who has a reciprocal translocation between Y and autosome 22. Her chromosomal constitution can be described as 46,X,t(Y;22)(p11.2;q11). At age 12, she was karyotyped because a gonadoblastoma had developed within her left gonad, which was "dysgenetic" (malformed and lacking germ cells). Subsequently, a gonadoblastoma was detected in her right gonad, which was also dysgenetic. She was otherwise a phenotypically normal female. (Gonadoblastomas are neoplasms that occur frequently in individuals with dysgenetic gonads and Y chromatin. It has been hypothesized that a gene in the centromeric region or long arm of the Y [deletion intervals 4B through 7] predisposes to gonadoblastoma formation in dysgenetic gonads; Page 1987.) Prior to the walk, pDP307 was the only Y probe detecting a deletion in this X,t(Y;22) female; she lacks the Y-specific 3.4 kb PstI fragment detected by pDP307 (Figure 2, probe C). However, she carries the Y-specific 5.3 kb TaqI fragment detected by probe D, at −28 kb, about 24 kb proximal to pDP307 (Figure 2). Probe D, like pDP307 and other sequences proximal to pDP307, is *DXYS1*-like; the intensity of the 3.2 kb TaqI fragment correlates well with the number of X chromosomes (Figure 2, probe D). One Y-chromosomal breakpoint in this X,t(Y;22) female falls within the *DXYS1*-like region, at about −23 kb (Figures 1 and 2; some studies not shown). This breakpoint defines the boundary between intervals 1B and 1C.

As a result of crossing these chromosomal landmarks, the walk was firmly oriented with respect to the centromere and short-arm telomere of the Y chromosome (Figures 1 and 2). We then extended the walk distally, toward the Yp telomere. About 140 kb distal to the breakpoint in XX male LGL203, a second Y-chromosomal breakpoint in X,t(Y;22) female WHT1013 was encountered, dividing intervals 1A1 and 1A2. While this female lacks all Y sequences examined between −23 and 140 kb, she carries all of the more distal sequences for which we tested. For example, she carries the Y-specific 3.4 kb and (barely visible) 2.5 kb TaqI fragments detected by probe A (Figure 2), at 143 kb.

Thus, X,t(Y;22) female WHT1013 appears to carry all of the Y chromosome except intervals 1A2 and 1B (Figures 1 and 2). We suspect that this deletion of about 160 kb occurred at the site of her reciprocal Y;22 translocation. (To the limited resolution of prometaphase chromosome banding, this translocation had been judged to be "balanced.") If the translocation sites and the deletion endpoints coincide, then interval 1A1 and the pseudoautosomal region of the Y chromosome should be grafted onto a centromere-containing portion of chromosome 22. The remainder of that chromosome 22 should be grafted onto intervals 1C through 7 of the Y chromosome.

Deletion intervals 1A1, 1A2, and 1B are defined by the Y-chromosomal breakpoints in LGL203 and WHT1013 (Figure 1). During the chromosomal walk of this region, we determined that an independently derived Y-chromosomal sequence, GMGY3 (Affara et al., 1986), maps in interval 1A1; it is present in both XX male LGL203 and X,t(Y;22) female WHT1013. Affara and colleagues had previously demonstrated that GMGY3 is present in most XX males. As GMGY3 was not contained within the portion of interval 1A1 we had cloned at that time, we initiated a second chromosomal walk from GMGY3. From this starting point, overlapping phages λOX118, 128, and 300 were obtained (Figure 1). This GMGY3 walk was soon joined to the walk initiated at pDP307. In all, serial screening of human genomic libraries yielded a set of 49 recombinant phages whose overlapping inserts encompass more than 230 kb of the Y chromosome (Figure 1).

In refining the deletion map of Yp we examined, by DNA hybridization, many individuals whose genomes contain portions of the Y chromosome. The two individuals we found to be most informative with respect to the localization of *TDF* are described here. Among the males found to carry part but not all of the Y chromosome, XX male LGL203 carries the smallest portion, interval 1A (1A1 plus 1A2; Figures 1 and 2). Indeed, interval 1A is present in each of the males found to have some but not all of the Y chromosome (D. C. P. and colleagues, unpublished results). We conclude that interval 1A of the Y chromosome is sufficient to induce testicular differentiation. Among the females found to carry part but not all of the Y chromosome, the smallest deletion occurs in X,t(Y;22) female WHT1013. She carries all of the Y chromosome except intervals 1A2 and 1B (Figures 1 and 2). Interval 1A (1A1 plus 1A2) is absent in each of the (nonmosaic) females found to carry part but not all of the Y chromosome (D. C. P. and colleagues, unpublished results). Taken as a group, these results argue that *TDF* is found in its entirety in interval 1A, the composite of intervals 1A1 and 1A2. The female phenotype of the X,t(Y;22) female implies that interval 1A2 contains an essential portion if not all of *TDF*.

(The fathers of XX male LGL203 and of X,t(Y;22) female WHT1013 appear to have intact, normal Y chromosomes as judged by hybridization with numerous Y-DNA probes, including probes for intervals 1A1, 1A2, 1B, and 1C [data not shown]. We conclude that the structural abnormalities

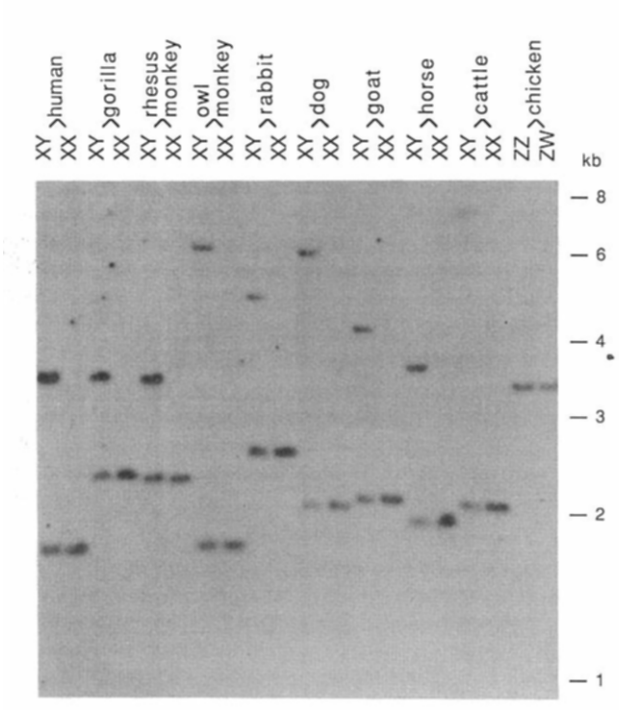

Figure 3. A Restriction Fragment from Interval 1A2 Detects Highly Conserved Sequences in Mammals and Birds

Probe pDP1007 was hybridized to EcoRI-digested DNAs from normal male and female humans, gorillas, rhesus monkeys, owl monkeys, rabbits, dogs, goats, horses, cattle, and chickens (42°C hybridization, 55°C wash). The sex chromosome constitutions of normal male and female mammals are, respectively, XY and XX. The sex chromosome constitutions of normal male and female birds are, respectively, ZZ and ZW. A scale in kilobases is provided at the right.

of the Y chromosome in LGL203 and WHT1013 are new mutations.)

Homologous Sequences on the Y and X Chromosomes of Other Mammals

Interval 1A2 is about 140 kb in size (Figure 1), or 0.2% of the Y chromosome. We supposed that only a small fraction of these 140 kb actually constitute TDF-encoding DNA sequences, and that these putative TDF-encoding sequences are situated among sequences with little or no function. Functional DNA sequences change slowly across evolutionary time as compared with nonfunctional DNA sequences. We reasoned that a search of interval 1A2 for DNA sequences showing a relatively high degree of evolutionary conservation could identify the *TDF* gene itself. Since testis determination appears to be a conserved function of the mammalian Y, we would search for 1A2 sequences having homologs on the Y chromosomes of a wide range of mammals. To this end, we prepared "Noah's ark blots." Onto these genomic DNA transfers "went one pair, male and female, of [many] beasts, clean and unclean, of birds and of [many things] that crawl on the ground, two by two" (Genesis 7:8–9), including humans, great apes (e.g., chimpanzee, gorilla), Old World monkeys (e.g., rhesus monkey), New World monkeys (e.g., owl monkey), rodents, rabbits, dogs, goats, horses, and cattle. Forty-eight restriction fragments collectively representing the entirety of deletion interval 1A2 were tested by hybridization to such blots. Under stringent conditions, many probes from interval 1A2 detect male-specific restriction fragments in humans and great apes, but do not cross-hybridize to DNAs from species more distantly related to humans. We conclude that most sequences in interval 1A2 have homologs on the Y chromosomes of chimpanzees and gorillas, but are not highly conserved among mammals. (Even this slight degree of Y-chromosomal conservation is in marked contrast to the *DXYS1*-like sequences in and proximal to interval 1B, which have no homologs on the Y chromosomes of apes.)

Against this backdrop of poorly conserved sequences, the high degree of evolutionary conservation displayed by certain sequences in interval 1A2 is striking. We have most closely examined the evolutionary conservation of a 1.3 kb HindIII fragment about 77 kb distal to the LGL203 breakpoint; plasmid pDP1007 contains this fragment. Under stringent hybridization conditions, pDP1007 detects a single- or low-copy sequence on the Y chromosome of all mammals examined. In each mammalian species shown in Figure 3, for example, pDP1007 detects one fragment that is present in the male but absent in the female. (The male-specific fragment is of the same length in human, gorilla, and rhesus monkey. In cattle, a male-specific fragment of between 7 and 8 kb is faintly detectable.) In each mammalian species shown in Figure 3, in addition to the Y-specific fragment, pDP1007 detects a fragment that is present in both males and females. In each mammalian species, this fragment hybridizes with about twice the intensity in the female as in the male. (Rehybridization with an α-tubulin probe confirmed that, in general, equal amounts of male and female DNA of each species were loaded. However, the female rhesus monkey lane is somewhat underloaded, such that the male-female-common fragment appears of equal intensity in male and female.) The simplest interpretation is that this locus is on the X chromosome in all mammals, such that females have two copies and males have one. Experiments described below confirm the X-chromosomal location of this fragment in humans. Thus, pDP1007 detects a highly conserved locus on the mammalian Y chromosome and a highly conserved locus on the mammalian X chromosome.

Probe pDP1007 also detects similar sequences in chickens (Figure 3). Chickens, like other birds, do not have the system of male heterogamety (XY male/XX female) seen in mammals. Instead, birds have a system of female heterogamety (ZZ male/ZW female). The fragment detected by pDP1007 is present in both male and female chickens. We have not observed a consistent difference between males and females in the intensity of hybridization of this fragment.

Homologous Sequences in the *Sxr* Region of the Mouse Y Chromosome

Probe pDP1007 detects a highly conserved locus on the mammalian Y chromosome, and it derives from the sex-determining region of the human Y. These findings are consistent with pDP1007 containing part of the *TDF* gene.

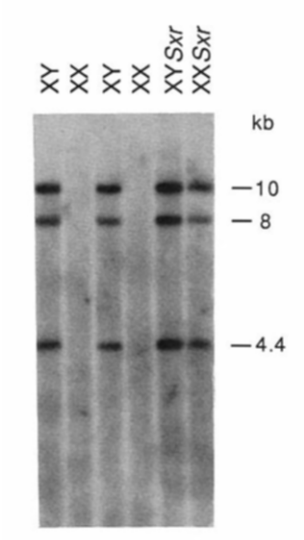

Figure 4. In Mice, Male-Specific Homologs of the Conserved Human Y Locus Map to the Sex-Determining (*Sxr*) Region of the Y Chromosome

Probe pDP1171 was hybridized to TaqI-digested DNAs from the following mice (from left to right): BALB/c normal male, BALB/c normal female, C57BL normal male, C57BL normal female, C57BL XY*Sxr* carrier male, C57BL XX*Sxr* male (47°C hybridization, 65°C wash). The sizes in kilobases of the hybridizing restriction fragments are indicated.

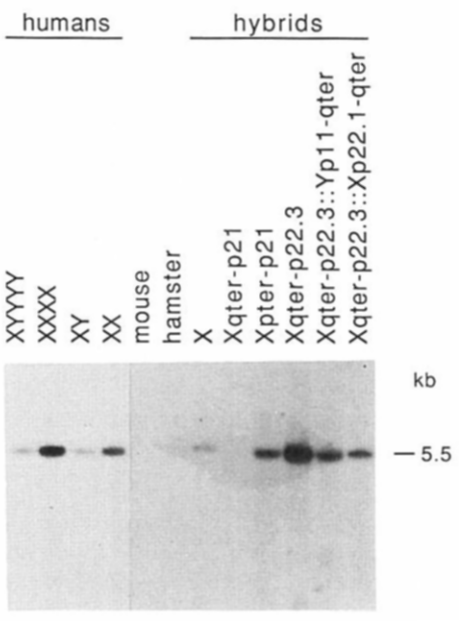

Figure 5. In Humans, the X-Chromosomal Homolog of the Evolutionarily Conserved Y-Chromosomal Sequence Maps to Xp21-p22.3

Probe pDP1039, which detects a 5.5-kb TaqI fragment on the human X chromosome, was hybridized to TaqI-digested DNAs from XYYYY male Oxen, XXXX female GM1416, a normal human male, a normal human female, mouse, hamster, t60-12 (somatic cell hybrid retaining human X chromosome; H. Willard, personal communication), A2-4 (somatic cell hybrid retaining a human X;21 translocation product carrying Xqter-p21; Worton et al., 1984), C1-T3 (somatic cell hybrid retaining a human X;21 translocation product carrying Xpter-p21; the reciprocal of A2-4), 85-13/8 (somatic cell hybrid retaining a terminally deleted human X, Xqter-p22.3; Curry et al., 1984), 97-1/5 (somatic cell hybrid retaining a human X-Y translocation product, Xqter-p22.3::Yp11-qter; Bernstein et al., 1978; Bernstein, 1985; Geller et al., 1986), and HY.22AZA1 (somatic cell hybrid retaining a dicentric X chromosome, Xqter-p22.3::Xp22.1-qter; Ferraro et al., 1980; M. Rocchi, personal communication); 47°C hybridization, 65°C wash.

Sxr (sex-reversed) mice provide a test of this hypothesis. XX *Sxr* mice do not have a Y chromosome, yet they develop as males because of the presence of a small, sex-determining portion of the Y, attached to the paternally derived X chromosome (Singh and Jones, 1982; Evans et al., 1982). In these respects, XX *Sxr* mice are similar to human XX males. However, while human XX males generally occur sporadically, each as the result of a new mutation, *Sxr* is passed through XY *Sxr* carrier males. These XY *Sxr* males have an abnormal Y chromosome carrying a duplication of the sex-determining region, one copy of which is frequently transferred to the X during meiosis.

If the mouse Y-chromosomal homolog of pDP1007 is sex-determining, then that homolog should be present in one copy in normal XY mice, in two copies in XY *Sxr* mice, and in one copy in XX *Sxr* mice. In this study, we used as a hybridization probe plasmid pDP1171, which carries a sequence cloned from the mouse Y by DNA sequence similarity to pDP1007. (The hybridization of pDP1007 to male-specific fragments in mouse DNA, though detectable at high stringency, is not as intense as its hybridization to male-specific fragments in other mammals.) When hybridized to TaqI digests, pDP1171 detects fragments of 4.4, 8, and 10 kb, all of which are present in normal males but absent in normal females (Figure 4). (Clone pDP1171 contains little conserved sequence and, at high stringency, there is no detectable hybridization to the mouse X chromosome.) The hybridization of these three Y-specific fragments is about twice as intense in the XY *Sxr* male as in the normal XY and XX *Sxr* males. (Rehybridization with an α-tubulin probe confirmed that equal amounts of DNA were loaded in the various lanes; not shown.) We conclude that a homolog of pDP1007 maps to the sex-determining region of the mouse Y chromosome.

(As discussed in Experimental Procedures, we suspect that the *Sxr* region of the mouse Y chromosome carries two distinct sequences homologous to pDP1007. We have no evidence of there being more than one pDP1007-homologous locus on the Y chromosome of any other mammal.)

A Homologous Locus on the X Chromosome

Probe pDP1007 detects highly conserved sequences not only on the Y chromosome but also, it seems, on the X chromosome (Figure 3). We set out to confirm the existence of and to map more precisely this putative X homolog in the human. First, the putative human X locus was cloned. We screened a genomic library from a human XX female by hybridization with pDP1007. Five different recombinant phages were identified, and their human inserts were found to overlap. This confirms that pDP1007 detects a single locus in the human female genome. We purified from one of these phages a human 1.3 kb HindIII

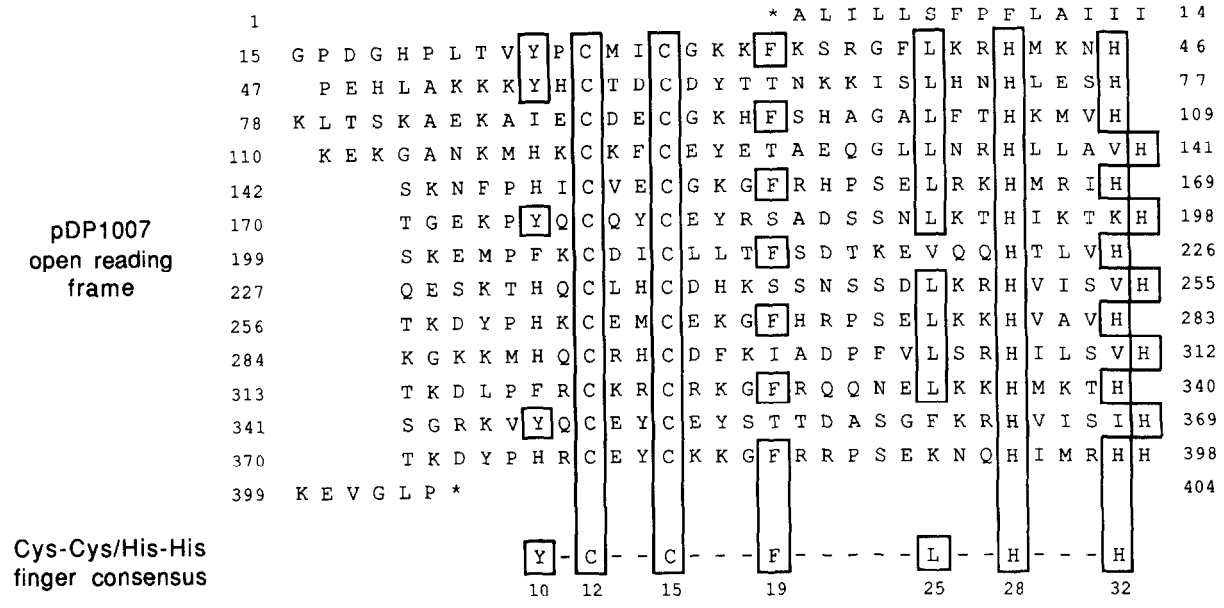

Figure 6. Nucleotide Sequence of pDP1007 Predicts a Protein with Multiple Cys-Cys/His-His Fingers

At the top is the sequence of 404 amino acids corresponding to the long open reading frame in the human insert of pDP1007. The sequence is aligned to show 13 repeats of 28 to 32 residues each. It is likely that the actual exon is shorter than the open reading frame and, therefore, that not all of the residues shown are found in the protein. The numbers at the very bottom indicate position in the repeat unit. Asterisks denote stop codons that bound the open reading frame. Below the open reading frame sequence is the amino acid sequence characteristic of repetitive Cys-Cys/His-His finger domains, as found in human transcription factor Sp1 (Kadonaga et al., 1987) and in the proteins encoded by Drosophila *serendipity* (Vincent et al., 1985), Drosophila *Krüppel* (Rosenberg et al., 1986), Drosophila *hunchback* (Tautz et al., 1987), and yeast *ADR1* (Hartshorne et al., 1986). Such domains are also found in the products of Drosophila *Kr h* (Schuh et al., 1986), mouse *mkr1* and *mkr2* (Chowdhury et al., 1987), and frog *Xfin* (Altaba et al., 1987), all cloned by DNA sequence similarity to Drosophila *Krüppel*. The best-conserved residues are boxed. Dashes mark positions which are more variable. Frog transcription factor IIIA (Miller et al., 1985; Brown et al., 1985) has four residues between the paired cysteines in eight of its nine repeat units, but is otherwise characterized by the consensus given here. As in the 13 repeats shown here for pDP1007, the other repetitive Cys-Cys/His-His finger domains have three, four, or five residues between the paired histidines.

fragment that is not highly conserved among mammals (and that does not cross-hybridize to pDP1007 or to any other Y DNA sequence); this fragment constitutes the insert of plasmid pDP1039. Probe pDP1039 detects a single-copy locus in humans. When hybridized to human genomic DNAs, the intensity of the 5.5 kb TaqI fragment detected by pDP1039 correlates well with the number of X chromosomes (XXXX > XX > XY, XYYYY; Figure 5), as expected for an X-chromosomal locus. We also hybridized probe pDP1039 to human–rodent somatic cell hybrids retaining intact or partially deleted X chromosomes (Figure 5). No cross-hybridization to rodent DNA was observed. We obtained positive results with a hybrid in which the X is the only human chromosome detectable cytogenetically. (No hybridization was observed with hybrids retaining the human Y but not the human X chromosome; data not shown.) Positive results were also obtained with hybrids retaining Xpter-p21 or Xqter-p22.3. We conclude that the pDP1007-related sequence present in the human female maps to region p21-p22.3 of the X chromosome. Thus, pDP1007 detects related sequences on Xp and Yp. Since these Xp and Yp loci are characterized by X-specific and Y-specific restriction fragments, they must not be subject to X-Y recombination during male meiosis. Apart from the pseudoautosomal region, where X-Y recombination is a frequent event (Simmler et al., 1985; Cooke et al., 1985), this is the first example of similar DNA sequences on Xp and Yp.

A Protein with Multiple "Finger" Domains

We have determined the nucleotide sequence of pDP1007's insert, which derives from the human Y chromosome and was highly conserved during mammalian evolution. In one of the six possible reading frames, nearly the entire length of this 1.3 kb segment is free of stop codons. The existence of a 1.2 kb open reading frame suggests that most of the human insert of pDP1007 is an exon of a gene encoding a protein.

We compared the amino acid sequence predicted for this pDP1007 open reading frame (Figure 6) with previously determined protein sequences. (Of course, the actual exon is probably shorter than the open reading frame in pDP1007. Some of the inferred amino acid residues, especially those near the ends of the open reading frame, may not be found in the protein.) We found a repetitive structure in the putative pDP1007 protein which is remarkably similar to the cysteine- and histidine-rich "finger" domains that are tandemly repeated in proteins such as frog transcription factor IIIA (TFIIIA). The sequence of 404 amino acids corresponding to the pDP1007 open reading frame is shown in Figure 6, arranged so as to align 13 repeats, each 28 to 32 residues in length. The amino acid

sequence of these 13 repeats from pDP1007 is compared to the consensus sequence for the multiple finger domains in frog TFIIIA (Miller et al., 1985; Brown et al., 1985), human transcription factor Sp1 (Kadonaga et al., 1987), and the products of the Drosophila *serendipity* (Vincent et al., 1985), Drosophila *Krüppel* (Rosenberg et al., 1986), Drosophila *hunchback* (Tautz et al., 1987), and yeast *ADR1* (Hartshorne et al., 1986) genes.

In each of these multiple-finger proteins, as in the protein predicted for pDP1007, the most conserved feature of the repeat is a pair of cysteines (positions 12 and 15) separated by 12 residues from a pair of histidines (positions 28 and 32); two residues (four in TFIIIA) separate the paired cysteines, and three to five residues separate the paired histidines (Figure 6). These paired cysteines and paired histidines are thought to be pulled into a tetrahedral coordination complex with a zinc cation, leaving the intervening residues (16 through 27) to form a DNA-binding loop (Miller et al., 1985). These domains, including those found in pDP1007, typically have large hydrophobic groups at residues 10 (Tyr), 19 (Phe), and 25 (Leu) (Figure 6). As in the other repetitive Cys-Cys/His-His fingers, there are many potentially DNA-binding residues (Lys, Arg, His, Asn, Gln, Thr; Ohlendorf and Matthews, 1983) throughout, including the putative DNA-binding loop (residues 16 through 27).

TFIIIA, Sp1, and the *serendipity*, *Krüppel*, *hunchback*, and *ADR1* products contain tandem arrays of between 2 and 9 such repeats, each about 30 residues in length. There are 13 fingers in an uninterrupted array in pDP1007, constituting nearly the entirety of its human insert (Figure 6). Of course, more fingers may be encoded in portions of the gene yet to be sequenced. (Frog *Xfin*, a gene of unknown function isolated by DNA sequence similarity to Drosophila *Krüppel*, encodes a protein with 37 fingers. These 37 fingers are divided into six arrays of three to eight fingers each [Altaba et al., 1987].)

Among the array of 13 fingers in pDP1007, there is a second-order repeat unit consisting of a pair of fingers. For example, there is a phenylalanine at position 19 in every other finger, and all but the first two fingers alternate in having three versus four residues between the paired histidines (Figure 6). This two-finger repeat has not been observed in the other proteins with multiple Cys-Cys/His-His fingers.

Outlines of a Transcription Unit on the Y Chromosome

Interval 1A2 of the Y chromosome contains some or all of a finger-protein gene, of which pDP1007 is essentially one exon. Several observations provide additional insights into the structure of that transcription unit. In addition to pDP1007, three other DNA segments from interval 1A2 were found to contain sequences highly conserved during mammalian evolution. The positions of the conserved segments, four in total, are indicated in Figure 7; pDP1007, 77 kb distal to the 1A2-1B boundary, is the most centromere-proximal. The other conserved sequences are located approximately 80, 97, and 124 kb distal to the 1A2-1B boundary.

At least three of the four conserved segments appear to fall within a single transcription unit. Partial cDNA clones were obtained by screening a newborn bovine testis cDNA library (Cate et al., 1986) with pDP1007. On gel transfers of bovine male and female genomic DNAs, these bovine testis cDNAs hybridize exclusively to the restriction fragments previously detected by pDP1007 (at 77 kb; Figure 3) and by the conserved sequences at 80 and 97 kb (not shown). Therefore, these cDNAs derive from the bovine Y- or X-chromosomal homolog of pDP1007. When hybridized to gel transfers of restriction digested DNA clones of human Y interval 1A2, the bovine cDNAs hybridize strongly to the conserved segments at 77, 80, and 97 kb (unpublished results). We suspect that these three conserved segments, and possibly the conserved segment at 124 kb, represent exons of a single gene. The DNA sequence of the "finger" domains at 77 kb (pDP1007) suggests that the direction of transcription is toward the centromere. The number, size, and detailed structure of transcripts have yet to be determined.

"CpG islands" are unusual segments of the genome, on the order of a kilobase in length, with a high G + C content and in which the dinucleotide CpG is abundant and unmethylated. In many but not all mammalian genes, a CpG island overlaps the transcription start site (Bird, 1986). The presence of sites for certain restriction enzymes, including BssHII (GCGCGC), EagI (CGGCCG), and SacII (CCGCGG), is highly characteristic of CpG islands (Lindsay and Bird, 1987). Within interval 1A2, we found five BssHII sites, four EagI sites, and two SacII sites by restriction digestion of cloned DNA. As shown in Figure 7, these 11 sites fall in two clusters, one at about 53 kb and the other at about 125 kb. It is not certain whether either of these two CpG islands corresponds to the 5′ end of a gene. The distal CpG island, at 125 kb, overlaps the most distal of the four conserved segments in interval 1A2. It is tempting to speculate that the distal CpG island overlaps the site of initiation of a transcript which extends, 5′ to 3′, from perhaps the conserved sequence at 124 kb to at least the conserved sequence (pDP1007) at 77 kb. CpG islands occur near the 3′ ends of several tissue-specific genes, some of which also have a second island at the 5′ end (Gardiner-Garden and Frommer, 1987). This suggests the possibility that the proximal CpG island, at 53 kb, is part of the same transcription unit (including pDP1007). Alternatively, the CpG island at 53 kb may be associated with a different transcript, or it may have no functional significance. We have not yet confirmed that the CpG dinucleotides in these islands are unmethylated in human genomic DNA.

(By means of pulsed-field gel electrophoresis, Pritchard et al. [1987] inferred the existence of a CpG island on distal Yp, 110 to 210 kb proximal to the pseudoautosomal region and 0 to 80 kb proximal to GMGY3. This inferred CpG island, which Pritchard and colleagues speculated might be *TDF*, likely corresponds to the more distal of the two CpG islands we identified by cloning interval 1A2; our more distal CpG island is about 60 kb proximal to GMGY3 [compare Figures 1 and 7]. This CpG island is about 15 kb proximal to the 1A1-1A2 boundary [Figure 7]. The size of interval 1A1, which by definition extends to the boundary

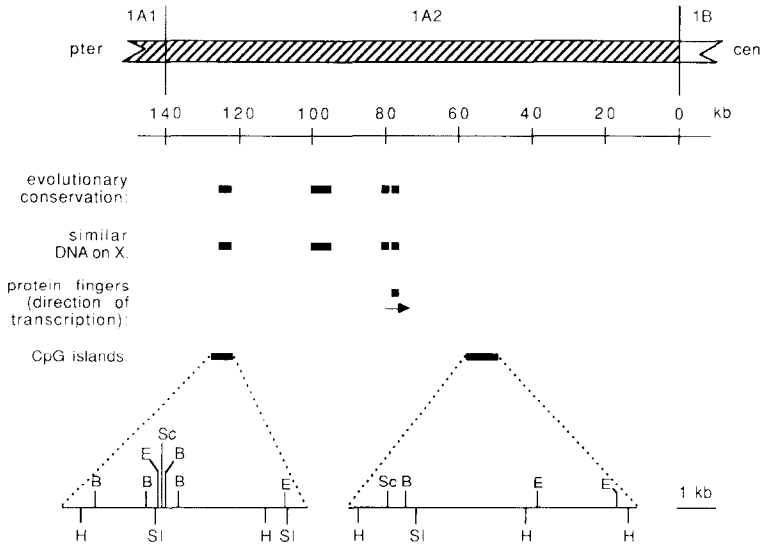

Figure 7. Detail of Interval 1A2 to Highlight Regions Showing Evolutionary Conservation, Cross-Hybridization to the X Chromosome, Encoded Protein Fingers (with direction of transcription inferred from nucleotide sequencing of pDP1007), and CpG Islands (with restriction maps drawn to a larger scale)

This summary of positive findings should be qualified in two ways. First, we do not mean to imply that the entirety of each highlighted region is, for example, highly conserved in evolution. Second, we do not mean to imply that other portions of interval 1A2 will not in the future be found, for example, to be highly conserved in evolution or to encode finger domains. B, Bss HII; E, Eag I; H, Hind III; Sc, Sac II; Sl, Sal I.

of the pseudoautosomal region, can then be estimated at 95 to 195 kb.)

A Related Gene on the X Chromosome

As already described, DNA sequences very similar to pDP1007 (at 77 kb) are found on the human X chromosome. Each of the three other conserved segments, at 80, 97, and 124 kb, also detect related sequences on the human X chromosome (Figure 7). This is in sharp contrast to the majority of sequences within interval 1A2, which do not detect similar sequences on the X chromosome. Thus, among DNA sequences within interval 1A2, there is a close correlation between evolutionary conservation and the presence of similar sequences on the X chromosome. It seems unlikely that either the Y locus or the X locus is a pseudogene, for the DNA sequences of both were highly conserved during mammalian evolution, as judged by hybridization with pDP1007 (Figure 3). Just as the conserved segments of interval 1A2 of the Y chromosome may constitute exons of a single gene, we speculate that their counterparts on the X chromosome constitute exons of a closely related gene.

Testicular Differentiation in the Absence of Interval 1A2

Human hermaphrodites are individuals whose gonads contain both testicular and ovarian tissue. In contrast to the vast majority of XX males, all XX hermaphrodites and a few XX males we have studied lack the entirety of interval 1A2 of the Y chromosome, including the Y-specific sequences detected by pDP1007 (not shown). Indeed, we have not detected DNA sequences from any part of the Y chromosome in these XX hermaphrodites and rare Y(-) XX males. We conclude that testicular differentiation can occur in the absence of the sex-determining portion of the Y chromosome. In such cases, testicular differentiation may be the result of mutations in autosomal or X-linked genes whose products function in conjunction with or downstream of TDF.

Discussion

A Gene in the Sex-Determining Region of the Y Chromosome

Our studies suggest that the sex of an individual is determined by the presence or absence of a very small portion of the human Y chromosome, perhaps a single gene (*TDF*, the testis-determining factor). XX male LGL203 carries only intervals 1A1 and 1A2, which together comprise between 235 and 335 kb of the distal short arm of the Y chromosome (Figure 1). (Judging by its cytogenetic appearance, the entire human Y chromosome is about 70,000 kb in length.) Though LGL203 carries less than 0.5% of the Y chromosome, he is a male. This finding, together with deletion analysis of other males carrying part but not all of the Y chromosome, demonstrates that the entirety of *TDF* is found in intervals 1A1 and 1A2. Conversely, X,t(Y;22) female WHT1013 carries 99.8% of the Y chromosome; she lacks only the 160 kb that comprise intervals 1A2 and 1B—yet she is a female. This result, together with deletion analysis of other females carrying part but not all of the Y chromosome, argues that an essential portion of *TDF* resides in interval 1A2, which measures 140 kb.

By chromosome walking, we have cloned a 230 kb portion of the human Y chromosome which includes all of interval 1A2 (Figure 1). Within interval 1A2, we have identified DNA sequences that appear to encode a protein, which we suspect is the testis-determining factor. Most of our studies of this putative gene have focused on the clone pDP1007, which carries a 1.3 kb portion of interval 1A2. Several findings argue that pDP1007 is essentially an exon of a gene. First, as judged by DNA–DNA hybridization, its sequence was highly conserved during evolution (Figure 3). Second, determination of its nucleotide sequence reveals an open reading frame stretching across most of its length. Third, as judged by the presence of homologous clones in a bovine testis cDNA library, it is transcribed. Fourth, these bovine cDNAs hybridize not only to

pDP1007 but also to other conserved sequences within interval 1A2, which probably constitute other exons of the same gene.

Moreover, this gene is probably *TDF*. First, deletion mapping argues that some or all of *TDF* is in interval 1A2, where the pDP1007 gene is also located; at present, there is no convincing evidence of a second gene in 1A2. Second, the presence of pDP1007-homologous sequences on the Y chromosome has been conserved throughout mammalian evolution (Figure 3). Third, in the mouse, homologs of pDP1007 not only map to the Y chromosome, they map specifically to the sex-determining (*Sxr*) region of the mouse Y (Figure 4). Despite concerted efforts in several laboratories to identify single-copy DNA sequences in the *Sxr* region of the mouse Y, the sequences we describe are among the first single- or low-copy number DNA sequences to be localized there.

A Protein with Multiple Finger Domains

The gene that includes pDP1007 encodes a protein with a tandem array of cysteine- and histidine-rich finger domains (Figure 7) similar to those found in frog transcription factor IIIA (TFIIIA; Miller et al., 1985; Brown et al., 1985), human transcription factor Sp1 (Kadonaga et al., 1987), and the proteins encoded by the Drosophila *serendipity* (Vincent et al., 1985), Drosophila *Krüppel* (Rosenberg et al., 1986), Drosophila *hunchback* (Tautz et al., 1987), and yeast *ADR1* (Hartshorne et al., 1986) genes. It is postulated that, in each such finger domain, a pair of cysteines and a pair of histidines are arranged about a central zinc ion in a tetrahedral coordination complex (Miller et al., 1985; Brown et al., 1985); the existence of this zinc complex has been demonstrated in the case of TFIIIA (Diakun et al., 1986).

It is postulated that such Cys-Cys/His-His finger domains (Figure 7) bind to nucleic acids in a sequence-specific manner (Miller et al., 1985). TFIIIA, which is comprised largely of finger domains, binds to an approximately 50-base region of the 5S RNA gene (Engelke et al., 1980; Sakonju et al., 1981) and to 5S RNA itself (Pelham and Brown, 1980; Honda and Roeder, 1980). Transcription factor Sp1 binds to a characteristic 10-base sequence in the promoter region of many mammalian genes, again via Sp1's finger domains (Kadonaga et al., 1987). If one relaxes the consensus sequence criteria that define putative cysteine- and histidine-rich, metal-binding finger domains, such domains appear to be characteristic of a host of nucleic acid binding proteins (Berg, 1986). By analogy to TFIIIA and Sp1, the presence of multiple Cys-Cys/His-His finger domains in the pDP1007 protein strongly implies that it binds to DNA or RNA in a sequence-specific manner.

TDF and Gonadal Sex Determination

It has often been speculated that TDF functions as a component of the cell surface or following its release from cells. However, the presence of a putative nucleic acid binding domain in the pDP1007 protein argues against such possibilities. That the pDP1007 protein likely binds to DNA or RNA suggests that it directly affects only those cells in which it is expressed; the pDP1007 protein probably functions in a cell-autonomous fashion. Nonetheless, the pDP1007 protein may have other domains that will be important to consider in more fully comprehending its function. Our findings do not resolve the question of whether the entirety of the process of sex determination is cell-autonomous. Indeed, the results of mouse XX↔XY aggregation chimera studies seem to exclude the possibility of cell-autonomous sex determination in all cell types except Sertoli cells and germ cells (Burgoyne et al., 1988). While our findings suggest that the first step in mammalian sex determination is cell-autonomous, subsequent steps need not be.

Among the other proteins containing tandemly repeated Cys-Cys/His-His finger domains like those found in the pDP1007 protein (Figure 7), the best characterized are TFIIIA and Sp1. Both proteins have been purified and shown, in vitro, to regulate transcription of the genes to which they bind (Engelke et al., 1980; Briggs et al., 1986). By analogy, we postulate that the pDP1007 protein binds to the regulatory sequences of one or more genes and positively or negatively regulates their transcription. If the pDP1007 protein is TDF, as we strongly suspect, then this transcriptional regulation constitutes the first step in mammalian sex determination. If the pDP1007 protein proves to be a transcriptional regulatory factor, it will be of interest to characterize the gene(s) whose expression it controls. The nucleic acid binding properties predicted for this protein may provide a biochemical means to identify genes that function downstream of *TDF* in the pathway of gonadal differentiation. Of course, the possibility that the pDP1007 protein functions as an RNA-binding protein should not be dismissed.

During the earliest stages of mammalian organogenesis, the developing gonads of XX and XY embryos are histologically indistinguishable; they are said to be "indifferent." By the sixth or seventh week of human development (day 12 or 13 in mice), the gonads are recognizable as testes or ovaries. *TDF* must be expressed by the time of this first histological appearance of sexual dimorphism. Otherwise, little is known as to either the temporal or spatial distribution of *TDF* expression. Given the cloning of what is likely the *TDF* gene, such issues should become more approachable: In what cells is *TDF* expressed, and at what stage(s) of development? Is *TDF* required at more than one stage of development to not only establish but also maintain the sexually determined state? Does *TDF* have functions apart from gonadal sex determination?

Indeed, it remains to be demonstrated that *TDF* is a single gene. It is possible that two or more transcription units, all contained within intervals 1A1 and 1A2, are required to induce testicular differentiation of the bipotential gonad. The haploid state of the Y chromosome has precluded, even in mice, genetic analysis that might allow one to determine the number of genes required. However, given that the gene we have cloned is postulated to act in a dominant fashion, transgenic animals should provide a biological assay for the sufficiency of this gene in inducing testicular differentiation. If the gene we have cloned is *TDF*, one would expect the transgene to have a dominant effect

on sexual phenotype, so that either XX or XY embryos carrying the transgene would develop as males.

Functional Relationship of Homologous Genes on Y and X Chromosomes

The human DNA insert of pDP1007, derived from the Y chromosome, strongly cross-hybridizes to a single locus in the human female genome on the X chromosome (Figure 5). The high degree of evolutionary conservation of both the X and Y loci (Figure 3) argues against either locus being a pseudogene. Rather, our studies suggest the possibility that the X and Y loci may encode similar proteins (Figure 7).

For purposes of discussion, let us assume that the Y-chromosomal gene we have identified is *TDF*, and that the related locus on the X chromosome encodes a structurally similar protein. Then several models are possible:
(1) The X-encoded protein does not function in gonadal sex determination, while the Y-encoded protein is sex-determining simply by its presence or absence.
(2) The X and Y loci act antagonistically in sex determination. For example, the X and Y genes could encode transcription factors which bind to the same regulatory sequences, one factor being a negative regulator and the other a positive regulator.
(3) The X and Y loci act in concert to bring about testis determination but are not functionally interchangeable. For example, the X and Y genes could encode subunits of a multimeric structure, perhaps producing a testis-determining heterodimer in XY embryos as opposed to a homodimer or monomers in XX embryos.

Models 1, 2, and 3 all fit well with the prevailing notion of a dominantly acting sex-determining factor unique to the Y chromosome. A fourth model does not fit with this prevailing notion, but its simplicity is attractive (see also German, 1988):
(4) The X and Y loci are functionally interchangeable, both are testis determining, and the X locus is subject to X-chromosome inactivation. According to this model, sex is determined by the total number of expressed X and Y loci: a single dose is female determining, while a double (or greater) dose is male determining. Assuming that the X locus is subject to X inactivation, there should be one active X locus per cell in all individuals, regardless of the number of X chromosomes. (When more than one X chromosome is present in a cell, all but one X chromosome are inactivated [Barr and Carr, 1962].) The number of functional copies—and therefore sex—are determined by the presence or absence of the Y chromosome, which carries one active copy. According to this scheme, sex determination is a function of X inactivation (see also Chandra, 1985).

Model 4 makes two testable predictions. First, the addition of an X-derived transgene to the genome of an XX embryo should result in testis differentiation, so long as that transgene is not subject to X inactivation. Second, increased expression of the X-chromosomal locus, perhaps due to a failure of X-inactivation (generalized or specific to this locus) could result in testicular differentiation in the absence of the Y-chromosomal locus, *TDF*. This might explain the presence of testicular tissue in XX hermaphrodites and the rare Y(-) XX males, who lack interval 1A2 of the Y chromosome and, therefore, *TDF*.

As judged by Y-DNA analysis, some XY females lack the sex-determining region of the Y chromosome (Disteche et al., 1986; Müller et al., 1986; Magenis et al., 1987; D.C.P. and colleagues, unpublished results). However, most XY females do not appear to have such large deletions; the short arm of the Y chromosome, including interval 1A2, is grossly intact in many XY females we have examined (unpublished results). These unexplained XY females may have point mutations in *TDF* or in genes that function in conjunction with or downstream of *TDF*. Models 3 and 4 predict that mutation of the X-chromosomal locus we have identified (at Xp21-p22.3) could cause XY embryos to develop as females. Indeed, transmission and cytogenetic findings suggest that female development of human XY embryos can be the result of X-chromosomal defects (German et al., 1978) and, in particular, of mutations involving region Xp21-pter (Bernstein et al., 1980). Mutation of the X-chromosomal locus we have identified may also account for the dimorphic X chromosomes of wood lemmings, in which XX, X*X, and X*Y embryos develop as females, while XY embryos develop as males (Fredga et al., 1976).

Evolution of Sex Determination in Vertebrates

It is likely that both the X- and Y-chromosomal genes detected by probe pDP1007 evolved from a single, common ancestral gene. We postulate that this ancestral gene was located on both homologs of a pair of autosomes, and that from this pair of autosomes evolved the mammalian X and Y chromosomes. These conjectures stem from the hypothesis, first forwarded by Muller to explain the paucity of genes on the Drosophila Y chromosome (1914), that morphologically and genetically distinct X and Y chromosomes evolved from an identical pair, a pair of autosomes. According to this hypothesis, the nascent Y chromosome tends to lose genes not involved in sex determination, resulting in divergence of the X and Y chromosomes. On the nascent Y—at least in the region where recombination with the X chromosome is suppressed—only the gene or genes that have acquired a sex-determining function should escape this fate; the sex-determining gene or genes may retain homology to X-chromosomal genes. As we have detected X-specific and Y-specific genes in a range of mammals (Figure 3), it appears that the divergence of the X and Y genes from a common ancestor occurred prior to the radiation of placental mammals. (We have yet to examine marsupials and monotremes.)

Of interest in this context are the results we obtained by hybridizing pDP1007 to genomic DNAs of male and female chickens (Figure 3). While mammals are characterized by male heterogamety (XX females, XY males), birds are characterized by female heterogamety (ZZ males, ZW females). Unfortunately, it is not known whether sex in birds is determined by the presence of the W chromosome or by the number of Z chromosomes. (Diploid birds with ZO and ZZW sex chromosomes have not been observed.) We have not, as yet, detected any consistent

qualitative or quantitative difference between the patterns of hybridization in male and female chickens. Thus we cannot determine whether the pDP1007 homologs are on the Z and W chromosomes or on an autosome. One intriguing possibility is that pDP1007 homologs are located on the avian Z and W chromosomes, but are expressed only on Z. There appears to be no dosage compensation of Z-linked genes in birds (Baverstock et al., 1982). Rather, the W chromosome in females appears to be heterochromatic and late replicating (for review, see Bloom, 1974). Therefore, it is possible that inactivation of the gene on the W chromosome is sex-determining in birds, such that two active copies lead to male development and one active copy leads to female development. (This model is quite analogous to model 4 for mammalian sex determination as presented earlier. It should also be recalled that in Drosophila and nematodes sex is determined by a dosage effect, females [or hermaphrodites] having two X chromosomes and males having one X chromosome; Bridges, 1916; Madl and Herman, 1979.) In any case, the finding of an avian homolog of the putative mammalian *TDF* gene holds out the prospect of understanding the genetic basis of sex determination not only in mammals, but perhaps in birds and other vertebrates as well.

Experimental Procedures

Origin of Probes pDP307, pDP1007, pDP1035, and pDP1039

The human insert of plasmid pDP307 (locus *DXYS42*) derives from a library of genomic HindIII restriction fragments prepared from flow-sorted human Y chromosomes. This library, provided by Marvin Van Dilla, was constructed in λ vector Charon 21A at the Biomedical Sciences Division, Lawrence Livermore National Laboratory, Livermore, CA, under the auspices of the National Laboratory Gene Library Project, sponsored by the U.S. Department of Energy. Plasmid pDP307 was constructed by inserting a 0.9 kb genomic HindIII fragment into the HindIII site of pUC13 (Vieira and Messing, 1982). These and subsequent recombinant DNA manipulations were carried out as described by Maniatis et al. (1982).

The human insert of plasmid pDP1007 is a 1.3 kb genomic HindIII fragment purified from phage λOX82 (Figure 1); this fragment, which derives from the human Y chromosome, is inserted into the HindIII site of pUC13. The human insert of plasmid pDP1035 is a 1.3 kb genomic EcoRI fragment purified from phage λOX113 (Figure 1); this fragment, which derives from the human Y chromosome, is inserted into the EcoRI site of Bluescript (Stratagene). The human insert of plasmid pDP1039 is a 1.2 kb genomic HindIII fragment purified from λBER110 (not shown); this fragment, which derives from the human X chromosome, is inserted into the HindIII site of Bluescript.

Human Chromosomal Walk

Recombinant phage with the prefix "λBER" were isolated from a genomic library of 46,XX female DNA constructed in λ vector EMBL3A (Frischauf et al., 1983); this library was a gift from Stuart Orkin. Recombinant phages with the prefix "λOX" were isolated from a genomic library of 49,XYYYY male (Sirota et al., 1981) DNA constructed in λ2010, a derivative of λ2001 (Karn et al., 1984); λ2010 was a gift from Jay Short. The overlaps among the clones (Figure 1) were determined by restriction mapping and hybridization.

DNA Extraction and Gel-Transfer Hybridization

Vertebrate genomic DNAs were prepared from blood, tissue, or cultured cell lines (Kunkel et al., 1977), digested with restriction endonucleases, electrophoresed on 0.7% agarose gels, and transferred (Southern, 1975) to nylon membrane. Whole plasmids or restriction fragments purified from the mammalian inserts of recombinant plasmids or phages were labeled with ^{32}P by nick translation (Rigby et al., 1977) or random-primer synthesis (Feinberg and Vogelstein, 1984) and hybridized overnight to genomic DNA transfers at 42°C or 47°C in 50% formamide, 5× SSC (1× SSC = 0.15 M NaCl, 15 mM Na citrate [pH 7.4]), 1× Denhardt's (0.02% Ficoll 400, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin), 20 mM $NaPO_4$ (pH 6.6), 50 μg/ml denatured salmon sperm DNA, 1% SDS, and 10% dextran sulfate. Prior to hybridization, some probes were prehybridized with an excess of sonicated human genomic DNA (Litt and White, 1985). Following hybridization, transfer membranes were washed three times for 15 min each at either 55°C, 60°C, or 65°C in 0.1× SSC, 0.1% SDS and exposed at −80°C for 1 to 4 days with X-ray film backed by an intensifying screen.

Cloning of Mouse Y Homolog

When hybridized to EcoRI digests of mouse genomic DNA, pDP1007 detects male-specific fragments of 5 and 11 kb. Genomic DNA from a normal BALB/c male was digested to completion with EcoRI and fractionated by agarose gel electrophoresis. The 4 to 6 kb fraction was ligated to EcoRI-digested λgt10 DNA (Huynh et al., 1985), and a clone containing the 5 kb EcoRI fragment isolated; pDP1171 is a subclone of this 5 kb EcoRI fragment in plasmid pBR322. When hybridized to EcoRI digests of mouse DNA, pDP1171 hybridizes to both male-specific fragments, 5 and 11 kb, detected by pDP1007. Indeed, it detects a Y-chromosomal TaqI restriction fragment length polymorphism (RFLP) also detected by pDP1007. The fact that pDP1171, which contains a 5 kb EcoRI fragment specific to the mouse Y, hybridizes on EcoRI digests to Y-specific fragments of 5 and 11 kb suggests that there are two pDP1007-homologous loci on the mouse Y. Both loci are in the *Sxr* region of the Y chromosome.

Nucleotide Sequencing and Protein Sequence Comparisons

Both strands of the 1.3 kb human insert of plasmid pDP1007 were sequenced by dideoxy chain termination (Sanger et al., 1977) using double-stranded plasmid DNA as template (Chen and Seeburg, 1985) and synthesized oligonucleotides as primers. Using the FastP algorithm (Lipman and Pearson, 1985), the amino acid sequence predicted from the nucleotide sequence of pDP1007 was compared with protein sequences stored in the National Biomedical Research Foundation Protein Identification Resource (NBRF/PIR) database (version 4.2/1.0).

Acknowledgments

We thank Kris Bieker and Dorothy Pettay for technical assistance; Marvin Van Dilla for a flow-sorted Y-chromosome library; Stuart Orkin for a human female genomic library; Nabeel Affara and Malcolm Ferguson-Smith for GMGY3; Jay Short for λ2010; Rich Cate for a bovine testis cDNA library; Tim Bestor for help with DNA sequencing; and Stephen Bloom, Don Doering, Robert Lofstedt, Robin Lovell-Badge, Nancy Ma, Paul Matsudaira, Anne McLaren, T. K. Mohandas, Donald Patterson, Hee-Sup Shin, Huntington Willard, Mariano Rocchi, and the MIT Department of Comparative Medicine for cell lines, blood and tissue samples, and genomic DNAs. We are grateful to Eric Lander and Don Brown for helpful discussions and to Bob Tjian, James German, and Paul Burgoyne for communication of results prior to publication. E. M. S. is the recipient of a Medical Research Council of Canada fellowship. E. M. C. F. is the recipient of a Public Health Service International Research Fellowship (TWO3951) and a Wellcome Trust travel grant. This study was supported by grants from the National Institutes of Health, the Dupont Center of Molecular Genetics at the Whitehead Institute, the Sigrid Jusélius Foundation, the Folkhälsan Institute of Genetics, and the Academy of Finland.

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 November 3, 1987; revised November 17, 1987.

References

Affara, N. A., Ferguson-Smith, M. A., Tolmie, J., Kwok, K., Mitchell, M., Jamieson, D., Cooke, A., and Florentin, L. (1986). Variable transfer of

Y specific sequences in XX males. Nucl. Acids Res. *14*, 5375–5387.

Altaba, A. R., Perry-O'Keefe, H., and Melton, D. A. (1987). *Xfin*: an embryonic gene encoding a multifingered protein in *Xenopus*. EMBO J. *6*, 3065–3070.

Barr, M. L., and Carr, D. H. (1962). Correlations between sex chromatin and sex chromosomes. Acta Cytol. *6*, 34–45.

Baverstock, P. R., Adams, M., Polkinghorne, R. W., and Gelder, M. (1982). A sex-linked enzyme in birds—Z-chromosome conservation but no dosage compensation. Nature *296*, 763–766.

Berg, J. M. (1986). Potential metal-binding domains in nucleic acid binding proteins. Science *232*, 485–487.

Bernstein, R. (1985). X;Y chromosome translocations and their manifestations. In The Y Chromosome, Part B: Clinical Aspects of Y Chromosome Abnormalities, A. A. Sandberg, ed. (New York: Alan R. Liss), pp. 171–206.

Bernstein, R., Wagner, J., Isdale, J., Nurse, G. T., Lane, A. B., and Jenkins, T. (1978). X-Y translocation in a retarded phenotypic male. J. Med. Genet. *15*, 466–474.

Bernstein, R., Koo, G. C. and Wachtel, S. S. (1980). Abnormality of the X chromosome in human 46,XY female siblings with dysgenetic ovaries. Science *207*, 768–769.

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

Bloom, S. E. (1974). Current knowledge about the avian W chromosome. BioScience *24*, 340–344.

Bridges, C. B. (1916). Non-disjunction as proof of the chromosome theory of heredity. Genetics *1*, 1–51, 107–163.

Briggs, M. R., Kadonaga, J. T., Bell, S. P., and Tjian, R. (1986). Purification and biochemical characterization of the promoter-specific transcription factor Sp1. Science *234*, 47–52.

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.

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 chimeric mouse testes. Development, in press.

Cate, R. L., Mattaliano, R. J., Hession, C., Tizard, R., Farber, N. M., Cheung, A., Ninfa, E. G., Frey, A. Z., Gash, D. J., Chow, E. P., Fisher, R. A., Bertonis, J. M., Torres, G., Wallner, B. P., Ramachandran, K. L., Ragin, R. C., Manganaro, T. F., MacLaughlin, D. T., and Donahoe, P. K. (1986). Isolation of the bovine and human genes for Müllerian inhibiting substance and expression of the human gene in animal cells. Cell *45*, 685–698.

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

Chen, E. Y., and Seeburg, P. H. (1985). Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA *4*, 165–170.

Chowdhury, K., Deutsch, H., and Gruss, P. (1987). A multigene family encoding several "finger" structures is present and differentially active in mammalian genomes. Cell *48*, 771–778.

Cooke, H. J., Brown, W. R. A., and Rappold, G. A. (1985). Hypervariable telomeric sequences from the human sex chromosomes are pseudoautosomal. Nature *317*, 687–692.

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.

Curry, C. J. R., Magenis, R. E., Brown, M., Lanman, J. T., Jr., Tsai, J., O'Lague, P., Goodfellow, P., Mohandas, T., Bergner, E. A., and Shapiro, L. J. (1984). Inherited chondrodysplasia punctata due to a deletion of the terminal short arm of an X chromosome. N. Eng. J. Med. *311*, 1010–1015.

Diakun, G. P., Fairall, L., and Klug, A. (1986). EXAFS study of the zinc-binding sites in the protein transcription factor IIIA. Nature *324*, 698–699.

Disteche, C. M., Casanova, M., Saal, H., Friedman, C., Sybert, V., Graham, J., Thuline, H., Page, D. C., and Fellous, M. (1986). Small deletions of the short arm of the Y chromosome in 46, XY females. Proc. Natl. Acad. Sci. USA *83*, 7841–7844.

Eichwald, E. J., and Silmser, C. R. (1955). Untitled communication. Transpl. Bull. *2*, 148–149.

Engelke, D. R., Ng, S.-Y., Shastry, B. S., and Roeder, R. G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell *19*, 717–728.

Epplen, J. T., Cellini, A., Romero, S., and Ohno, S. (1983). An attempt to approach the molecular mechanisms of primary sex determination: W- and Y-chromosomal conserved simple repetitive DNA sequences and their differential expression in mRNA. J. Exp. Zoology *228*, 305–312.

Evans, E. P., Burtenshaw, M. D., and Cattanach, B. M. (1982). Meiotic crossing-over between the X and Y chromosomes of male mice carrying the sex-reversing (*Sxr*) factor. Nature *300*, 443–445.

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

Ferraro, M., de Capoa, A., Mostacci, C., Pelliccia, F., Zulli, P., Baldini, M. A., and di Nisio, Q. (1980). Cytogenetic and clinical studies in gonadal dysgenesis with 46,X,Xt(qter-p221::p223-qter) karyotype: review and phenotype/karyotype correlations. J. Med. Genet. *17*, 457–463.

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 *I*, 711.

Fredga, K., Gropp, A., Winking, H., and Frank, F. (1976). Fertile XX- and XY-type females in the wood lemming *Myopus schisticolor*. Nature *261*, 225–227.

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

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. J. Mol. Biol. *196*, 261–282.

Geldwerth, D., Bishop, C., Guellaën, G., Koenig, M., Vergnaud, G., Mandel, J.-L., and Weissenbach, J. (1985). Extensive DNA sequence homologies between the human Y and the long arm of the X chromosome. EMBO J. *4*, 1739–1743.

Geller, R. L., Shapiro, L. J., and Mohandas, T. K. (1986). Fine mapping of the distal short arm of the human X chromosome using X/Y translocations. Am. J. Hum. Genet. *38*, 884–890.

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., in press.

German, J., Simpson, J. L., Chaganti, R. S. K., Summitt, R. L., Reid, L. B., and Merkatz, I. R. (1978). Genetically determined sex-reversal in 46,XY humans. Science *202*, 53–56.

Guellaen, G., Casanova, M., Bishop, C., Geldwerth, D., Andre, G., Fellous, M., and Weissenbach, J. (1984). Human XX males with Y single-copy DNA fragments. Nature *307*, 172–173.

Hartshorne, T. A., Blumberg, H., and Young, E. T. (1986). Sequence homology of the yeast regulatory protein ADR1 with *Xenopus* transcription factor TFIIIA. Nature *320*, 283–287.

Honda, B. M., and Roeder, R. G. (1980). Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell *22*, 119–126.

Huynh, T. V., Young, R. A., and Davis, R. W. (1985). Construction and screening cDNA libraries in lambda gt10 and lambda gt11. In DNA Cloning Techniques: A Practical Approach, D. Glover, ed. (Oxford: IRL Press), pp. 49–78.

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

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., Masiar, 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.

Karn, J., Matthes, H. W. D., Gait, M. J., and Brenner, S. (1984). A new selective phage cloning vector, λ2001, with sites for XbaI, BamHI, HindIII, EcoRI, SstI and XhoI. Gene *32*, 217–224.

Kiel-Metzger, K., Warren, G., Wilson, G. N., and Erickson, R. P. (1985). Evidence that the human Y chromosome does not contain custered DNA sequences (Bkm) associated with heterogametic sex determination in other vertebrates. New Eng. J. Med. *313*, 242–245.

Kunkel, L. M., Smith, K. D., and Boyer, S. H. (1976). Human Y-chromosome-specific reiterated DNA. Science *191*, 1189–1190.

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

Lipman, D. J., and Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. Science *227*, 1435–1441.

Litt, M., and White, R. L. (1985). A highly polymorphic locus in human DNA revealed by cosmid-derived probes. Proc. Natl. Acad. Sci. USA *82*, 6206–6210.

Madl, J. E., and Herman, R. K. (1979). Polyploids and sex determination in *Caenorhabditis elegans*. Genetics *93*, 393–402.

Magenis, R. E., Casanova, M., Fellous, M., Olson, S., and Sheehy, R. (1987). Further cytologic evidence for Xp-Yp translocation in XX males using in situ hybridization with Y-derived probe. Hum. Genet. *75*, 228–233.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Marcus, M., Tantravahi, R., Dev, V. G., Miller, D. A., and Miller, O. J. (1976). Human–mouse cell hybrid with human multiple Y chromosomes. Nature *262*, 63–65.

McLaren, A., Simpson, E., Tomonari, K., Chandler, P., and Hogg, H. (1984). Male sexual differentiation in mice lacking H-Y antigen. Nature *312*, 552–555.

McClung, C. E. (1902). The accessory chromosome—sex determinant? Biol. Bull. *3*, 43–84.

Miller, J., McLachlan, A. D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. *41*, 1609–1614.

Morgan, T. H. (1910). Sex-limited inheritance in *Drosophila*. Science *32*, 120–122.

Muller, H. J. (1914). A gene for the fourth chromosome of *Drosophila*. J. Exp. Zoology *17*, 325–336.

Müller, U., Donlon, T., Schmid, M., Fitch, N., Richer, C.-L., Lalande, M., and Latt, S. A. (1986). Deletion mapping of the testis determining locus with DNA probes in 46,XX males and in 46,XY and 46,X,dic(Y) females. Nucl. Acids Res. *14*, 6489–6505.

Ohlendorf, D. H., and Matthews, B. W. (1983). Structural studies of protein–nucleic acid interactions. Ann. Rev. Biophys. Bioeng. *12*, 259–284.

Page, D. C. (1986). Sex reversal: deletion mapping the male-determining function of the human Y chromosome. Cold Spring Harbor Symp. Quant. Biol. *51*, 229–235.

Page, D. C. (1987). Hypothesis: a Y-chromosomal gene causes gonadoblastoma in dysgenetic gonads. Development, in press.

Page, D. C., Harper, M. E., Love, J., and Botstein, D. (1984). Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. Nature *311*, 119–123.

Page, D. C., Brown, L. G., and de la Chapelle, A. (1987). Exchange of terminal portions of X- and Y-chromosomal short arms in human XX males. Nature *328*, 437–440.

Painter, T. (1923). Studies in mammalian spermatogenesis. II. The spermatogenesis of man. J. Exp. Zoology *37*, 291–335.

Pelham, H. R. B., and Brown, D. D. (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. Proc. Natl. Acad. Sci. USA *77*, 4170–4174.

Petit, C., de la Chapelle, A., Levilliers, J., Castillo, S., Noel, B., and Weissenbach, J. (1987). An abnormal X-Y interchange accounts for most but not all cases of human XX maleness. Cell *49*, 595–602.

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

Rigby, P. W. J., Dieckmann, M., Rhodes, C., and Berg, P. (1977). Labelling DNA to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. *113*, 237–251.

Rosenberg, U. B., Schröder, C., Preiss, A., Kienlin, A., Côté, S., Riede, E., and Jäckle, H. (1986). Structural homology of the product of the *Drosophila Krüppel* gene with *Xenopus* transcription factor IIIA. Nature *319*, 336–339.

Sakonju, S., Brown, D. D., Engelke, D. R., Ng, S.-Y., Shastry, B. S., and Roeder, R. G. (1981). The binding of a transcription factor to deletion mutants of a 5S ribosomal RNA gene. Cell *23*, 665–669.

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

Schuh, R., Aicher, W., Gaul, U., Côté, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R., and Jäckle, H. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by *Krüppel*, a Drosophila segmentation gene. Cell *47*, 1025–1032.

Simmler, M.-C., Rouyer, F., Vergnaud, G., Nystrom-Lahti, M., Ngo, K. Y., de la Chapelle, A., and Weissenbach, J. (1985). Pseudoautosomal DNA sequences in the pairing region of the human sex chromosomes. Nature *317*, 692–697.

Simpson, E., Chandler, P., Goulmy, E., Disteche, C. M., Ferguson-Smith, M. A., and Page, D. C. (1987). Separation of the genetic loci for the H-Y antigen and for testis determination on human Y chromosome. Nature *326*, 876–878.

Singh, L., and Jones, K. W. (1982). Sex reversal in the mouse (*Mus musculus*) is caused by a recurrent nonreciprocal crossover involving the X and an aberrant Y chromosome. Cell *28*, 205–216.

Singh, L., Phillips, C., and Jones, K. W. (1984). The conserved nucleotide sequences of Bkm, which define *Sxr* in the mouse, are transcribed. Cell *36*, 111–120.

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

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

Stevens, N. M. (1905). Studies in spermatogenesis with especial reference to the "accessory chromosome." Carnegie Institute Report 36 (Washington, D.C.).

Tautz, D., Lehmann, R., Schnürch, H., Schuh, R., Seifert, E., Kienlin, A., Jones, K., and Jäckle, H. (1987). Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. Nature *327*, 383–389.

Vergnaud, G., Page, D. C., Simmler, M-C., Brown, L., Rouyer, F., Noel, B., Botstein, D., de la Chapelle, A., and Weissenbach, J. (1986). A deletion map of the human Y chromosome based on DNA hybridization. Am. J. Hum. Genet. *38*, 109–124.

Vieira, J., and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene *19*, 259–268.

Vincent, A., Colat, H. V., and Rosbash, M. (1985). Sequence and structure of the *serendipity* locus of *Drosophila melanogaster*. J. Mol. Biol. *186*, 149–166.

Wachtel, S. S., Ohno, S., Koo, G. C., and Boyse, E. A. (1975). Possible role for H-Y antigen in the primary determination of sex. Nature *257*, 235–236.

Washburn, L. L., and Eicher, E. M. (1983). Sex reversal in XY mice caused by dominant mutation on chromosome 17. Nature *303*, 338–340.

Welshons, W. J., and Russell, L. B. (1959). The Y chromosome as the bearer of male determining factors in the mouse. Proc. Natl. Acad. Sci. USA *45*, 560–566.

Wilson, E. B. (1905). The chromosomes in relation to the determination of sex in insects. Science *22*, 500–502.

Worton, R. G., Duff, C., Sylvester, J. E., Schmickel, R. D., and Willard, H. F. (1984). Duchenne muscular dystrophy involving translocation of the *dmd* gene next to ribosomal RNA genes. Science *224*, 1447–1449.