Isolation of a sequence which maps close to the human sex determining gene

C.A.Pritchard, P.J.Goodfellow and P.N.Goodfellow

Laboratory of Human Molecular Genetics, Imperial Cancer Research Fund, Lincoln's Inn Fields,
London WC2A 3PX, UK

Received June 1, 1987; Revised and Accepted July 9, 1987

ABSTRACT

A sequence mapping close to the human sex determining gene (TDF) has been isolated from a lambda library constructed with DNA derived from a chromosome transfectant hybrid cell line. This sequence is shown to be present in the DNA of X-Y interchange males at a very high frequency and, based on these studies, it is categorised with the sequence defined by the probe, GMGY3, as the closest known Y chromosome derived marker to TDF. In contrast to GMGY3, however, this locus shares no homology with any other human chromosome. Southern blot analysis also reveals specific hybridization to the Y chromosome of other primates. It therefore defines, for the first time, a conserved and Y chromosome unique locus that is near to TDF.

INTRODUCTION

The mammalian Y chromosome encodes a gene, the testis determining factor (TDF), that is essential for normal male sex determination (1,2,3). Under the influence of TDF, the embryonic gonad differentiates as a testis and most subsequent male sexual differentiation is in response to the soluble products of this organ (4,5). Cloning TDF will allow the molecular investigation of sex determination and will provide a model system for studying the regulation of complex phenotypes. Unfortunately, cloning strategies based on expression are limited by the lack of knowledge of the temporal and tissue distribution of the TDF function. An alternative approach is to define the chromosomal position of TDF and to isolate all the sequences from that part of the genome.

The human Y chromosome is composed of two distinct regions (6). The pseudoautosomal region is present at the telomeric end of the Y chromosome short arm and is shared with the X chromosome. Sequences and genes in this region recombine between the X and Y chromosome and this has facilitated the construction of a meiotic map (7,8). The second region of the Y chromosome encodes TDF and fails to recombine with the X chromosome. Maps of this

region have been constructed by studying structurally abnormal Y chromosomes and the Y-derived DNA present in XX males (9,10). The majority of XX males are generated by the terminal transfer of Y chromosomal sequences, including TDF, from the Y chromosome to the X chromosome during male meiosis (11,12,13).

The amount of Y-derived material varies in different XX males and this forms the basis of deletion maps around TDF. These maps agree in placing TDF on the distal portion of the short arm of the Y chromosome, but proximal to the pseudoautosomal region. The closest markers to TDF are MIC2, DXYS5 (defined by probe 47z) and the sequence recognised by probe GMGY3. MIC2 is the most proximal of the known pseudoautosomal markers and is the distal flanking marker for TDF (8). Similarly, DXYS5 is the closest proximal marker to TDF (10). The GMGY3 locus maps between MIC2 and DXYS5, but its relative order with respect to TDF is not known (9). The distance between MIC2 and DXYS5 is also unclear.

A striking feature of the sequences that map close to TDF is their cross-homology with sequences elsewhere in the genome. For example, several Y-derived sequences, including probe 47z, hybridize to sequences on the long arm of the X chromosome (14) and it has been demonstrated that this homology arose by duplication and transposition of material from the X to the Y chromosome in human evolution (15). Similarly, probe GMGY3 hybridizes to autosomal as well as Y chromosome sequences (16).

Genetic markers for the Y chromosome have been obtained by cloning from sources enriched for Y-derived sequences, either by using the fluorescence activated cell sorter or by using human-rodent hybrid cells containing the human Y chromosome (17,18,19). Isolation of sequences that can be used for the precise mapping of TDF has then required further systematic screening. In a previous study we demonstrated that chromosome mediated gene transfer (CMGT), combined with cotransfer of selectable markers and cell sorting for MIC2 antigen positive cells, could be used to isolate fragments of the human Y chromosome in a mouse cell background (20). In this report we describe the use of this approach to produce a hybrid cell containing a Y-derived fragment that spans the region around TDF. A genomic library in bacteriophage lambda has been constructed with DNA from this hybrid and the library has been used to directly clone a sequence that maps close to TDF. This sequence shows no cross homology to other human chromosomes.

MATERIALS AND METHODS

Chromosome transfer

Chromosomes from the human cell line Oxen (49;YYYY) were transfected to mouse Ltk- cells by the method of Weis *et al* (21). The conditions for cotransformation and selection by fluorescence activated cell sorting have been described in detail elsewhere (20).

Cell Culture conditions

All CMGT transfectants were grown in Dulbecco-modified Eagle's medium supplemented with 10% foetal calf serum, penicillin and streptomycin. To retain the presence of the appropriate Y chromosomal fragment, hypoxanthine (10^{-4} M), thymidine (1.6×10^{-6} M) and aminopterin (10^{-5} M) were added to the culture medium. The XX male lymphoblastoid and fibroblast cell lines were cultured in RPMI medium with 10% foetal calf serum.

Genomic cloning and preparation of unique sequence probes

High molecular weight DNA was prepared from the IP2.2F secondary CMGT transfectant and partially digested with *Mbo*I. After size fractionation, DNA fragments in the range 15-20Kb were cloned into the *Bam*HI site of EMBL3 (22). The resulting library (1×10^6 recombinant clones) was screened with nick-translated total human DNA and the positive phage identified. Those still positive after a second screen were purified from cleared plate lysates by a method described by Maniatis *et al* (23).

Complete *Eco*RI and *Hind*III digests of recombinant phage DNA were subcloned into the plasmid vector, pUC8. The ligation mix was transformed into the bacterial strain JM101 by standard methods (23). Plasmid subclones containing phage inserts and giving no reaction with human repeat DNA on Southern blot hybridization were chosen for further analysis. A full description of all recombinant clones obtained by this procedure is provided in ref.36.

Southern analysis

Genomic DNA was prepared from the CMGT transfectants, EBV-transformed lymphoblastoid cell lines, cultured skin fibroblasts and peripheral blood cells by standard methods that have been published elsewhere (24). In each case, the DNA was digested with the appropriate restriction enzyme (see figure legends), electrophoresed through a 0.8% agarose gel and transferred to a nylon membrane (Gene Screen Plus; New England Nuclear) under conditions recommended by the manufacturer. DNA inserts used as probes were labelled by

the random priming method (25). Filters were prehybridized and hybridized in 10% dextran sulphate, 1M sodium chloride and 1% sodium dodecyl sulphate (SDS) at 65°C. They were washed the next day at room temperature in 2 x SSC then at 65°C in 0.2 x SSC, 0.2% SDS. Autoradiography was carried out at -70°C with intensification. For conditions of reduced stringencies, filters were hybridized at 55°C and washed to 0.2 x SSC at 55°C.

XX male patients

The patients we have studied are listed in Table 2. All individuals were diagnosed clinically as true XX males. The chromosomes of each patient were analysed cytogenetically at the time of medical examination and, for most cases, karyotypic tests were repeated during the course of this work.

RESULTS

Derivation of chromosomal fragments spanning the human sex determining gene

Expression of the MIC2 gene product (the 12E7 antigen) on the cell surface provides an endogenous selectable marker for the human Y chromosome (26). MIC2 also lies within close proximity of the sex determining gene (8). Chromosome mediated gene transfer (CMGT) followed by fluorescence activated cell sorter selection for 12E7 antigen positive cells was used to derive a series of cell lines containing fragments of the human Y chromosome in a mouse background. The strategy for creating such cell lines has been described in detail elsewhere (27). We have also reported previously the use of cotransformation methods to enrich for CMGT-competent cells (20). The frequent association of DNA plasmid with donor chromosomal fragments in cotransfection has provided an approach for introducing selectable genes into specific regions of the Y chromosome. Selection for these genes after a second chromosome transfer has facilitated further fragmentation of the Y chromosome.

Nine primary and twenty nine secondary CMGT transfectants were analysed in detail for the cotransfer of loci syntenic with MIC2. The cell line IP2.2 contains very little Y chromosome DNA; it possesses the most centromeric marker for TDF, DXYS5, and the two most proximal pseudoautosomal loci, DXYS17 and MIC2 (Table 1). The locus defined by GMGY3 is also present. Despite segregating the closest markers known to flank TDF, analysis of the total human DNA content indicates the presence of asyntenic chromosomal fragments in IP2.2. In an effort to reduce the proportion of contaminating sequences, chromosomes from IP2.2 were used as donors for a secondary chromosome transfer. After selection for the exogenously added thymidine kinase (tk) gene, ten cell lines were derived. The Y chromosome single copy sequences

CMGT TRANSFECTANT	DXYS17	SEQUENCE			HUMAN DNA CONTENT (Kb)
		MIC2	GMGY3	DXYS5	
IP2.2	+	+	+	+	30,000
IP2.2A	+	+	+	+	30,000
IP2.2B	+	+	-	-	10,800
IP2.2C	+	+	-	-	7,200
IP2.2D	+	+	+	+	15,000
IP2.2E	+	+	-	-	5,400
IP2.2F	+	+	+	+	10,200
IP2.2G	+	+	-	-	5,400
IP2.2H	+	+	+	+	10,800
IP2.2J	+	+	+	+	15,000
IP2.2K	+	-	-	+	1,800

Table 1. Characterisation of secondary CMGT transfectants A-K derived from the cell line IP2.2. Fragment sizes were calculated by the whole cell dot blot method described previously (29). The human Y chromosome is estimated to contain approximately 30,000 Kb of DNA. Therefore, based on the dot blot quantitations, IP2.2 would be expected to contain the equivalent of a whole Y chromosome. As previous studies have shown this cell line to contain only a few Y chromosome-derived markers (27), it must possess contaminating fragments from other human chromosomes.

contained within these secondary transfectants are indicated in Table 1. The approximate total human DNA content, as calculated by the whole cell dot blot method (28), is also shown. As DXYS17 is cotransferred in all cases, the selected tk gene must be associated closely with this locus in the primary transfectant IP2.2. An interstitial deletion has occurred in the generation of the cell line IP2.2K. This is not an unexpected result as chromosomal rearrangements are known to form frequently in the CMGT process (29,30). The cell line IP2.2F contains the bracketed sex determining region and the least amount of extraneous human DNA.

Isolation and chromosomal location of the sequence 27a

A genomic library was constructed in the bacteriophage lambda vector, EMBL3 using DNA from IP2.2F. To identify recombinant phage containing human sequences, the library was screened with human DNA labelled by nick-translation. After isolation of the positive phage 27, EcoRI/HindIII doubly restricted subfragments were cloned in the plasmid vector, pUC8. A unique subclone, 27a, of 600 bp in size was identified by the method of Gusella *et al* (31).

DNA from two females and one male was digested with a series of restriction enzymes and hybridized to probe 27a. The results presented in Figure 1 show that 27a only recognises restriction fragments in male DNA. No hybridization to female DNA is observed. Hence, unlike many other Y chromosome-derived sequences, the locus detected by 27a must be unique to the

Figure 1. Male specificity of probe 27a. DNA from one male (lane 1) and two females (lanes 2 and 3) was restricted with the enzymes HindIII (H), PstI (Ps), PvuI (Pv), BamHI (B) and BglII (Bg). After transfer to a nylon membrane, the DNA was hybridized with the probe 27a at 65°C in 1M NaCl. The filter was washed to a final stringency of 0.2 x SSC at 65°C. HindIII restriction fragments of phage lambda were used as sized markers.

Y chromosome. As no homology to the X chromosome is observed, this locus cannot be pseudoautosomal nor part of the chromosomal material transferred from the X to the Y chromosome in recent human evolution. Furthermore, by virtue of its mode of derivation, this locus most probably maps between MIC2 and DXYS5 on the Y chromosome short arm.

Deletion mapping in XX males

To map the locus defined by 27a with respect to the male determining gene, the DNA of individuals with abnormal sex chromosome constitutions was analysed. A 2.2Kb fragment is detected in TaqI-digested male DNA by 27a (Fig.2, lane 10). The same restriction fragment is present in six out of the eight XX males tested in Figure 2 (lanes 2,3,5,7,8,9). 27a was hybridized additionally to a series of XX males whose chromosomal rearrangements had been characterized previously by the probes GMGY3 and 47z. The results presented in Table 2 show that the 27a locus is not present in the DNA of all XX males. Despite this observation, 27a does hybridize to the DNA of XX males as frequently as GMGY3 (89% of cases) and more so than 47z (78% of cases). As the loci defined by GMGY3 and 27a cannot be separated by their form of deletion mapping, we cannot deduce whether they lie on the proximal or distal side of TDF.

Chromosomal location of the 27a locus in other species

Given the conserved nature of mammalian sex determination, candidate

Figure 2. Presence of the locus defined by probe 27a in the XX male genome. In each case the DNA was restricted with TaqI. Lanes: (1) XX female; (2)-(9) represent XX males from Table 2 in the following order: LGL163, GORE, PP, WB, GM1889, WER, GIM, DP; (10) XY male.

sequences for TDF are expected to be cross-species homologous. To investigate the Y-linked conservation of the 27a locus, its chromosomal location in a series of other species was examined. The hybridization of 27a to male and female human, gorilla, chimpanzee and mouse DNA is shown in Figure 3. In this case, moderate stringencies of hybridization (55°C; 1M NaCl) and wash

PATIENT	SEQUENCE		
	27a	GMGY3	47z
RP	+	+	+
PP	-	-	-
GA	+	+	+
WB	+	+	+
DP	+	+	+
GIM	+	+	+
WER	+	+	+
RE(1)	+	+	+
RE(2)	+	+	+
RE(3)	+	+	+
RE(4)	+	+	+
RE(5)	+	+	+
RE(6)	+	+	+
RE(7)	+	+	+
GM1889	-	-	-
LGL163	+	+	+
CON101	+	+	-
CHM006	+	+	-

Table 2. Characterisation of XX males. The presence (+) or absence (-) of the loci defined by probes 47z, 27a and GMGY3 is indicated in each case. Patients CON101 and CHM006 are referenced in Vergnaud *et al* (10). GIM and WER are patients 1 and 3 respectively in Guellan *et al* (35).

Figure 3. Species conservation of the 27a locus. Lanes: (1) human male; (2) human female; (3) gorilla male; (4) gorilla female; (5) chimpanzee male; (6) chimpanzee female; (7) mouse male; (8) mouse female. The DNA was restricted with EcoRI and hybridized to probe 27a at moderate stringency (55°C; 1M NaCl).

conditions (55°C; 0.2 x SSC) were used. Two male-specific EcoRI fragments are observed in human DNA; the cognate 9.4Kb band and a weaker band at 1.65Kb. This smaller fragment has been mapped to the long arm of the Y chromosome and is not observed when higher stringencies of hybridization are used (same conditions but at 65°C; data not shown). Male-specific signals in gorilla and chimpanzee are observed (Fig.3, lanes 3 and 5), and the hybridization pattern of chimpanzee DNA is identical to that of human DNA. Hence, despite the rearrangements that took place between the X and Y chromosomes in recent human evolution (15), the locus defined by this probe has remained specific to the Y chromosome. No hybridization to male or female mouse DNA is seen (Fig.3, lanes 6 and 7), which suggests that this sequence is not widely conserved.

Figure 4. Arrangement of sequences around the human testis determining gene (TDF) on the Y chromosome short arm.

DISCUSSION

We have demonstrated the directed isolation of a sequence from a predefined region of the Y chromosome by exploiting CMGT. A similar approach has been used by others for studying the murine MHC complex (32) and for isolating probes adjacent to the Wilm's tumour-aniridia locus (33). CMGT has also made a vital contribution to the molecular cloning of a candidate sequence for the cystic fibrosis gene (34). Taken together, these results suggest that chromosome transfer can play a valuable role in reverse genetics strategies designed to clone genes defined only by chromosomal position. One potential disadvantage of this technique, however, is the frequent occurrence of interstitial deletions generated during the chromosome transfer process. The cotransfection of asyntenic chromosomal fragments also reduces the efficiency with which clones from the selected region can be isolated. Despite these drawbacks, L1 repeat fingerprinting and long-range restriction mapping show that extensive stretches of donor chromosomal DNA can be transferred intact (C.P., unpublished observation, 29).

Many single copy probes for the Y chromosome short arm have been isolated and mapped with respect to one another and the human sex determining gene (9,10). Most of these sequences are known to recognise homologous loci on other human chromosomes. In this report, we have described the isolation of a sequence that is unique to the Y chromosome and is closely associated with TDF. The map location of this sequence with respect to two other close markers for the male determinant and the human pseudoautosomal region is illustrated in Figure 4. At present, it is not clear whether the locus for 27a is positioned above or below the sex determining gene, however, given its

Y-specificity, the locus may be closer to TDF than either that defined by MGY3 or 47z. Recent studies using pulsed field gel electrophoresis suggest that the 27a locus is more distal on the Y chromosome than the other two loci (C.P., unpublished observations). Although the 27a sequence is found on the Y chromosome of other primates, it fails to hybridize to the male or female mouse genome. We therefore suggest that it is a relatively new sequence that has 'hitch-hiked' with the sex determining gene since primate evolution.

In most cases it is clear that XX males are generated by a terminal exchange of a Y chromosome fragment including TDF (11,12,13). If this is the mechanism by which all XX males are generated, then the 27a and MGY3 loci would map proximal to TDF because three XX males fail to inherit these sequences. However, a prediction of this hypothesis is that the paternal pseudoautosomal region inherited by XX males should be derived from the Y chromosome. This prediction has been tested recently for three XX males containing no detectable Yp sequences: none of the three were found to inherit the Y chromosome pseudoautosomal region (13). A fourth Yp negative XX male (patient PP in this report) is also known not to possess sequences from his paternal Y pseudoautosomal region (unpublished observation). These males must have either an undetected Y chromosome mosaicism or a mutation at an X-linked, autosomal or pseudoautosomal locus.

We are currently isolating further sequences from the TDF region and are constructing a long-range restriction map across the region defined by MIC2, 27a, MGY3 and DXYS5.

ACKNOWLEDGEMENTS

We are grateful to Drs.R.Erickson, N.Dennis, M.Hulten, J.Insley, K.Madan, R.Quaife, M.Crawford, M.Taylor, A.de la Chapelle, M.Fellous, C.Petit and J.Weissenbach for providing the XX male patients studied in this report. We also thank Kirby Smith for the primate DNA. Our colleagues G.Banting, S.Darling, B.Pym, B.Williams, G.Knott and M.V.Wiles provided a great deal of encouragement and support.

REFERENCES

1. Ford,C.E., Jones,K.W., Polani,P.E., de Almeida,J.C. and Briggs,J.H. (1959) *Lancet* **1**, 711-713
2. Jacobs,P.A. and Strong,J.A. (1959) *Nature* **183**, 302-303
3. Miller,O.J., Drayna,D. and Goodfellow,P. (1983) *Human Gene Mapping 7*, Cytogenet.Cell.Genet. **37**, 176-204
4. Jost,A. (1953) *Recent Prog.Horm.Res.* **8**, 379-418
5. Jost,A., Vigier,B., Prepin,J. and Perchellet,J.P. (1973) *Recent Prog.Horm.Res.* **29**, 1-41
6. Burgoyne,P.S. (1982) *Hum.Genet.* **61**, 85-90

7. Rouyer, F., Simmler, M.-C., Johnsson, C., Vergnaud, G., Cooke, H.J. and Weissenbach, J. (1986) *Nature* 319, 291-295
8. Goodfellow, P.J., Darling, S.M., Thomas, N.S. and Goodfellow, P.N. (1986) *Science* 234, 740-743
9. Affara, N.A., Ferguson-Smith, M.A., Tolmie, J., Kwok, K., Mitchell, M., Jamieson, D., Cooke, A. and Florentin, L. (1986) *Nucl. Acids Res.* 14, 5375-5387
10. Vergnaud, G., Page, D.C., Simmler, M.-C., Brown, L., Rouyer, F., Noel, B., Botstein, D., de la Chapelle, A. and Weissenbach, J. (1986) *Am. J. Hum. Genet.* 38, 109-124
11. Ferguson-Smith, M.A. (1966) *Lancet* ii, 475-476
12. de la Chapelle, A. (1981) *Hum. Genet.* 58, 105-116
13. Petit, C., de la Chapelle, A., Levilliers, J., Catillo, S., Noel, B. and Weissenbach, J. (1987) *Cell* (in press)
14. Geldwerth, D., Bishop, C., Guellan, G., Koenig, M., Vergnaud, G., Mandel, J.L. and Weissenbach, J. (1985) *EMBO J.* 4, 1739-1743
15. Page, D.C., Harper, M., Love, J. and Botstein, D. (1984) *Nature* 311, 119-123
16. Affara, N.A., Florentin, L., Morrisson, N., Kwok, K., Mitchell, M., Cook, A., Jamieson, D., Glasgow, L., Meredith, L., Boyd, E. and Ferguson-Smith, M. (1986) *Nucl. Acids Res.* 14, 5352-5373
17. Bishop, C.E., Guellan, G., Geldwerth, D., Voss, R., Fellous, M. and Weissenbach, J. (1983) *Nature* 303, 831-832
18. Wolfe, J., Erickson, R.P., Rigby, P.W.J. and Goodfellow, P.N. (1984) *EMBO J.* 3, 1997-2003
19. Burk, R.D., Ma, P. and Smith, K. (1985) *Mol. Cell. Biol.* 5, 576-581
20. Pritchard, C. and Goodfellow, P.N. (1986) *EMBO J.* 5, 979-985
21. Weis, J.H., Nelson, D.L., Przyborski, M.J., Chaplin, D.D., Mulligan, R.C., Houseman, D.E. and Seidman, J.G. (1984) *Proc. Natl. Acad. Sci. USA.* 81, 4879-4883
22. Frischauf, A.-M., Lehrach, H., Poustka, A.-M. and Murray, N. (1983) *J. Mol. Biol.* 170, 827-842
23. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. Hofker, M.H., Wapenaar, M.C., Goor, N., Bakker, E. and Pearson, P.L. (1985) *Hum. Genet.* 70, 148-156
25. Feinberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.* 137, 266-267
26. Goodfellow, P.N., Banting, G., Sheer, D., Ropers, H.-H., Caine, A., Ferguson-Smith, M.A., Povey, S. and Voss, R. (1983) *Nature* 302, 346-349
27. Darling, S.M., Goodfellow, P.J., Pym, B., Banting, G.S., Pritchard, C. and Goodfellow, P.N. (1987) Cold Spring Harbor Symp. Quant. Biol. 51, (in press)
28. Porteus, D.J. (1985) *Som. Cell. Mol. Genet.* 11, 445-454
29. Pritchard, C.A. and Goodfellow, P.N. (1987) *Genes and Develop.* 1, 172-178
30. Klobutcher, L.A. and Ruddle, F.H. (1979) *Nature* 280, 657-660
31. Gusella, J.F., Keys, C., Varysanyi-Breiner, A., Kao, F.-T., Jones, C., Puck, T.T. and Houseman, D. (1980) *Proc. Natl. Acad. Sci. USA.* 77, 2829-2833
32. Weis, J.H., Seidman, J.G., Houseman, D.E. and Nelson, D.L. (1986) *Mol. Cell. Biol.* 6, 441-451
33. Porteus, D.J., Boyd, P.A., Christie, S., Cranston, G., Fletcher, J.M., Gosden, J.R., van Heyningen, V., Rout, D., Seawright, A., Simola, K.O.J. and Hastie, N. (1987) *Proc. Natl. Acad. Sci. USA.* (in press)
34. Estivill, X., Farrall, M., Scambler, P.J., Bell, G.M., Hawley, K.M.F., Lench, N.J., Bates, G.P., Kruger, H.C., Frederick, P.A., Stanier, P., Watson, E.K., Williamson, R. and Wainwright, B.J. (1987) *Nature* 326, 840-845
35. Guellan, G., Casanova, M., Bishop, C., Geldwerth, D., Andre, G., Fellous, M. and Weissenbach, J. (1984) *Nature* 307, 172-173
36. Pritchard, C. Ph.D. Thesis, University of London (1987)