Characterisation of the coding sequence and fine mapping of the human **DFFRY** gene and comparative expression analysis and mapping to the **Sxr**b interval of the mouse Y chromosome of the **Dffry** gene

Graeme M. Brown¹, Robert A. Furlong¹, Carole A. Sargent¹, Robert P. Erickson¹, Guy Longepied², Michael Mitchell², Michael H. Jones³, Timothy H. Hargreave⁴, Howard J. Cooke⁵ and Nabeel A. Affara¹,*

¹Human Molecular Genetics Research Group, University of Cambridge Department of Pathology, Tennis Court Road, Cambridge CB2 1QP, UK, ²INSERM U406, Genetique Medicale et Developpement, Faculte de Medecine de la Timone, 27, Bd Jean Moulin 13385 Marseille Cedex 05, France, ³Chugai Research Institute for Molecular Medicine153-2 Nagai, Niihari, Ibaraki 300-41, Japan and ⁴Department of Urology and University of Edinburgh Department of Surgery (Urology), and ⁵MRC Human Genetics Unit Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK

Received September 4, 1997; Revised and Accepted October 20, 1997

**DFFRY** (the Y-linked homologue of the **DFFRX** *Drosophila* fat-facets related X gene) maps to proximal Yq11.2 within the interval defining the AZFa spermatogenic phenotype. The complete coding region of **DFFRY** has been sequenced and shows 89% identity to the X-linked gene at the nucleotide level. In common with **DFFRX**, the potential amino acid sequence contains the conserved Cys and His domains characteristic of ubiquitin C-terminal hydrolases. The human **DFFRY** mRNA is expressed in a wide range of adult and embryonic tissues, including testis, whereas the homologous mouse **Dffry** gene is expressed specifically in the testis. Analysis of three azoospermic male patients has shown that **DFFRY** is deleted from the Y chromosome in these individuals. Two patients have a testicular phenotype which resembles Sertoli cell-only syndrome, and the third diminished spermatogenesis. In all three patients, the deletions extend from close to the 3′ end into the gene, removing the entire coding sequence of **DFFRY**. The mouse **Dffry** gene maps to the **Sxr**b deletion interval on the short arm of the mouse Y chromosome and its expression in mouse testis can first be detected between 7.5 and 10.5 days after birth when type A and B spermatogonia and pre-leptotene and leptotene spermatocytes are present.

**INTRODUCTION**

In eutherian mammals, the homologue of the *Drosophila* developmental gene fat facets (**faf**) is located on the sex chromosomes (¹). There is an X-linked copy (**DFFRX**) mapping to Xp11.4 which escapes X-inactivation, and related sequences (**DFFRY**) have been mapped to Yq11.2 (¹). The sequence of the equivalent mouse X-linked gene (**Dffrx**) has been deposited in GenBank (S.A. Wood, 1996; accession number U67874). Both the human X- and Y-linked loci are transcribed (8–9 kb transcripts), and the X-linked transcript encodes a protein of 2547 amino acids which shows 44% identity and 88% similarity to the *Drosophila* gene product. The mapping interval in Yq11.2 containing **DFFRY** has been shown by deletion analysis to contain at least three genetic functions: (i) the AZFa phenotype associated with male infertility characterised by a histology similar to Sertoli cell-only syndrome (²); (ii) short stature ascribed to a defective GCY locus (³–⁵); and (iii) the skeletal anomalies found in the Turner phenotype (⁶). Thus **DFFRY** is potentially a candidate for these genetic functions.

Recently, it has been shown that **faf** is a member of a family of deubiquitinating genes whose products remove ubiquitin from protein–ubiquitin conjugates (⁷,⁸). The *Drosophila*, mouse and human X and Y (see below) genes contain the conserved Cys and His domains that are characteristic of ubiquitin-specific hydrolases. The high degree of conservation of these domains found in the human genes argues strongly that they perform a homologous biochemical function. Thus, these genes may play an important regulatory role at the level of protein turnover by preventing degradation of proteins through the removal of conjugated ubiquitin.

The biological function(s) of **DFFRX** and **DFFRY** may be indicated by the two phenotypes that arise in *Drosophila* from mutations in the **faf** gene. First, the gene is important for normal eye development where it influences the fate of cells within the ommatidia destined to become photoreceptors. In mutant flies, there is disorganised ectopic differentiation of photoreceptor cells in the facets of the compound eye. In humans, **DFFRX** can be
Figure 1. (a) The amino acid sequence of the \textit{DFFRY} gene product shown compared with the amino acid sequence of the products of \textit{DFFRX} and \textit{Dffrx}. Only amino acid differences are shown. (b) The partial sequence of the mouse \textit{Dffry} gene (aligned against the \textit{DFFRY} sequence) used to design Y-specific primers for RT-PCR analysis of mRNA isolated from mouse tissues. The primer sequences are underlined, with an arrowhead indicating the orientation. \textit{Forward} pointing primer, \textit{Dffry-2F}; \textit{reverse} pointing primer, \textit{Dffry-2R}. A – indicates the deletion of an amino acid. (c) The alignment of the conserved Cys and His domains of \textit{DFFRY} (EMBL accession Nos Y13618 and Y13619); \textit{DFFRX} (X98296) (1); \textit{fam} (\textit{Dffrx}, U67874, unpublished); \textit{faf} (L04959) (7); \textit{TRE-2} (P35125) (31); \textit{UHX-1} (U44839) (32); \textit{UBPX} (P40818, unpublished); \textit{Unph} (U20657) (33); \textit{Mus-Ubp} (P35123) (34); \textit{DOA4} (L08070) (35); \textit{UPB1} (M63484) (36); \textit{UPB4} (U02518) (35). The filled boxes mark the highly conserved residues and the shaded boxes the weakly conserved residues.
### CYS DOMAIN

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFFRX</td>
<td>YIVLYTV</td>
<td>Human Yq1L.2</td>
</tr>
<tr>
<td>DFFRX</td>
<td>YIVLYTV</td>
<td>Human Xp11.4</td>
</tr>
<tr>
<td>FAM</td>
<td>YIVLYTV</td>
<td>Mouse X</td>
</tr>
<tr>
<td>TRX-2</td>
<td>YIVLYTV</td>
<td>Human 17q21-24</td>
</tr>
<tr>
<td>UBX-1</td>
<td>YIVLYTV</td>
<td>Human 3p21.3</td>
</tr>
<tr>
<td>FAF</td>
<td>YIVLYTV</td>
<td>Mouse</td>
</tr>
<tr>
<td>MUS-B</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
<tr>
<td>UBP1</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
<tr>
<td>UBP4</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
</tbody>
</table>

### HIS DOMAIN

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DFFRX</td>
<td>YIVLYTV</td>
<td>Human Yq1L.2</td>
</tr>
<tr>
<td>DFFRX</td>
<td>YIVLYTV</td>
<td>Human Xp11.4</td>
</tr>
<tr>
<td>FAM</td>
<td>YIVLYTV</td>
<td>Mouse X</td>
</tr>
<tr>
<td>TRX-2</td>
<td>YIVLYTV</td>
<td>Human 17q21-24</td>
</tr>
<tr>
<td>UBX-1</td>
<td>YIVLYTV</td>
<td>Human 3p21.3</td>
</tr>
<tr>
<td>FAF</td>
<td>YIVLYTV</td>
<td>Mouse</td>
</tr>
<tr>
<td>MUS-B</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
<tr>
<td>UBP1</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
<tr>
<td>UBP4</td>
<td>YIVLYTV</td>
<td>Yeast</td>
</tr>
</tbody>
</table>

### Consensus

<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus</td>
<td>YLVYTV</td>
</tr>
</tbody>
</table>

---

**Note:** The sequences are presented in the context of human and mouse genes, with specific domains (CYS and HIS) and variations across different proteins.
excluded as a serious candidate for a number of X-linked eye disorders by current mapping data (9). The second function relates to the role that faf plays in oocyte development. In Drosophila, mutations in the gene lead to abnormal oocyte content and the inability of the fertilized egg to undergo normal embryogenesis. The embryo fails to undergo both nuclear migration and subsequent cellularisation. This is a maternal effect requiring the product of faf (expressed in follicular cells) for normal development of the oocyte. The map location of DFFRX coincides with the region of the X defined by partial deletions in females (resulting in monosity for part of proximal Xp) as being critical for the major stigmata associated with the Turner phenotype. Thus DFFRX may be implicated in the gonadal phenotype observed in Turner syndrome where there is a failure of oocytes to pass through the first meiotic prophase, with massive oocyte loss leading to the degeneration of the developing ovary into a streak gonad (10,11). A second mammalian gonadal phenotype, one more similar to the faf phenotype, is shown by XO mice where functional oocytes are produced, but those that escape elimination at pachytene (probably because of pairing failure) appear to be compromised (12), mimicking the situation in Drosophila faf mutants. The resulting embryos are retarded during pre-implantation development, many failing to form normal blastocysts, and there is increased embryonic death around the time of implantation (12). The most likely cause of this phenotype is a reduced accumulation (because of the presence of only a single X chromosome) of an X-linked gene(s) product(s) in the maternal germ line during oogenesis critical for the early stages of embryogenesis (12–14). Both X chromosomes are active during oogenesis, and thus Dffrx may well be subject to X-inactivation in other tissues. Clearly, a role for DFFRX/Dffrx in oocyte function suggests that the Y homologue may have a different function, perhaps a role associated with the development and maintenance of male germ cells mediated through Sertoli cell (the equivalent in males of the ovary granulosa cells) function.

Here we describe the sequence of the DFFRY gene, the mapping of the Dffry gene into the Sxrαβ (15) interval on the mouse Y chromosome, the expression of the human and mouse genes and analysis of infertile patients deleted for the AZFa interval. The data show that DFFRY occupies approximately half of the AZFa interval and is expressed ubiquitously. In contrast, the mouse Dffry gene is expressed specifically in the testis and can first be detected between 7.5 and 10.5 days after birth when type A and B spermatogonia and pre-leptotene and leptotene spermatocytes are present.

RESULTS

Sequence of the DFFRY gene

The complete sequence of the coding region of the DFFRY gene (EMBL accession Nos Y13618 and Y13619) was determined by sequencing cDNA clones isolated from testis, foetal brain and retinal cDNA libraries and 3' and 5' extension from a series of Y-derived cDNA clones first identified during the course of defining the sequence of the X-linked gene (1). In addition, part of the DFFRY cDNA sequence was obtained from the RT-PCR product derived from the 7631 Y-only somatic cell hybrid. The partial coding sequence of the mouse Dffry gene was obtained from a Y-specific genomic clone isolated using probes derived from the DFFRX gene in order to design Y-specific PCR primers for RT-PCR expression analysis of the mouse Y-linked gene (see below for identification of the partial Dffry coding sequence). Figure 1a shows the complete DFFRY-encoded amino acid (2555 residues) sequence compared with that of DFFRX. Only amino acid differences have been entered into the alignment, which show a high degree of conservation between the X and Y genes. Figure 1b shows the partial sequence of Dffry aligned against the human DFFRY sequence, and it too shows a high degree of conservation. Comparisons between the amino acid sequences of DFFRX, DFFRX and the mouse X-linked gene Dffrx are as follows: (i) DFFRX/DFFRX, 89% identity and 98% similarity; (ii) DFFRY/Dffrx, 89% identity and 98% similarity; and (iii) DFFRX/Dffrx, 97% identity and 99% similarity. From these comparisons, it can be seen that there is greater conservation between DFFRX and Dffrx than between the two human genes. Alignment of the conserved Cys and His domains that define the family of ubiquitin C-terminal hydrolases reveals that faf, DFFRX and DFFRY are almost identical at the key positions of these motifs, strongly suggesting that the DFFRX and DFFRY genes function as deubiquitinating enzymes (Fig. 1c). PCR analysis of male and Y-only somatic cell hybrid genomic DNA with primers specific for the Y coding sequence has revealed the presence of numerous introns in the structure of the Y gene, indicating that it cannot have arisen as a result of retrotransposition.

Expression of the DFFRY mRNA in human adult and embryonic tissues

The expression of the DFFRY mRNA in a series of human adult and embryonic tissues was examined by RT-PCR analysis. PCR primers B17cs2 and Y5 were designed from the DFFRX mRNA, one of which (Y5) bridges the junction between two exons; thus PCR amplification will only occur from an mRNA-derived template and not from genomic DNA. The primers generate a fragment of 301 bp from the DFFRX mRNA, and no product from genomic DNA (see Fig. 2). It can be seen that DFFRY is expressed ubiquitously in adult and embryonic (≈10 weeks of gestation) tissues, giving a pattern similar to the DFFRX gene (1).

Dffry is expressed specifically in mouse testis

From the partial sequence of the Dffry coding region and the sequence of the X-linked gene Dffrx, mouse Y-specific and X-specific primers covering the same region were designed to perform RT-PCR analysis on a variety of adult mouse tissues. The Y-specific primers Dffry-2F and Dffry-2R amplify a 272 bp fragment from the reverse transcription product and a 400 bp fragment from genomic DNA, whereas the X-specific primers Fam-F and Fam-R amplify a 224 bp and 3 kb product respectively. Thus for both genes, the forward and reverse primers are derived from different exons and amplify across intronic sequence; this acts as an internal control for genomic contamination of the RT-PCR. In Figure 3a, it can be seen that Dffrx is expressed ubiquitously, mirroring the findings for DFFRX. In contrast, Dffry is not detectable by RT-PCR in several somatic tissues. Actin primers (β-act5′ and β-act3′ producing a 600 bp product) were used as a control to demonstrate the presence of actin transcripts in each mRNA preparation. Figure 3b shows that Dffry is expressed specifically in mouse testis and can be detected in RNA from 14 day, 21 day and adult mouse testis. The same samples were analysed withPgk2 (expressed primarily in round spermatids at 21 days) primers and Dffrx primers.
A detailed assessment of $Dffry$ expression in RNA from staged mouse testis was carried out by RT-PCR analysis with the same primers in order to determine whether expression varied with changes in the germ cell population. Total RNA from 17 and 18 dpc (days post-coitum), 0.5, 3.5, 7.5, 10.5 and 21 dpp (days post-partum) and adult mouse testis was isolated and, in addition to $Dffry$, was analysed with primer pairs for the mouse $Dffrx$, $Rbm$, $Prm1$ (protamine 1) and $Pgk2$ genes. Expression of $Dffry$ (see Fig. 4) is first detected between 7.5 and 10.5 dpp when type A and B spermatogonia and pre-leptotene and leptotene spermatocytes are present, whereas $Dffrx$ is expressed at all stages tested. As expected, the expression of $Pgk2$ correlated well with changes in the germ cell population, showing greatly increased expression at 21 days when post-meiotic round spermatids first appear.

Expression of $Dffry$ in mouse adult tissues. (a) The mouse primers $Dffry$-2F and $Dffry$-2R (see Fig. 1) were used to examine the expression of the mouse Y-linked gene and the primers Fam-F and Fam-R the expression of the mouse X-linked gene in various adult tissues. Primers ($\beta$-act5′ and $\beta$-act3′), which amplify $\beta$-actin transcripts, were used as a positive control to show the presence of actin transcripts in each mRNA preparation. The fragments produced by each primer pair are: $Dffry$-2F/$Dffry$-2R = 272 bp; Fam-F/Fam-R = 224 bp; $\beta$-act5′/$\beta$-act3′ = 600 bp. The low molecular weight band is primer dimer. Epid. + v.d. = epididymis and vas deferens. (b) The same X- and Y-specific primer pairs were used to examine the expression of $Dffry$ and $Dffrx$ in mRNA from adult, 14 and 21 day mouse testis. Primers (Pgk2F and Pgk2R) from the testis-specific $Pgk2$ gene were used as a positive control. This gene first increases most dramatically in post-meiotic round spermatids (21 days).

**Figure 2.** RT-PCR analysis of RNA extracted from human adult and embryonic tissues. Primers Y5 and B17cs2 were designed from the $DFFRY$ sequence to permit specific amplification of PCR product (301 bp) from Y gene transcripts.

**Figure 3.** Expression of $Dffry$ and $Dffrx$ in mouse adult tissues. (a) The mouse primers $Dffry$-2F and $Dffry$-2R (see Fig. 1) were used to examine the expression of the mouse Y-linked gene and the primers Fam-F and Fam-R the expression of the mouse X-linked gene in various adult tissues. Primers ($\beta$-act5′ and $\beta$-act3′), which amplify $\beta$-actin transcripts, were used as a positive control to show the presence of actin transcripts in each mRNA preparation. The fragments produced by each primer pair are: $Dffry$-2F/$Dffry$-2R = 272 bp; Fam-F/Fam-R = 224 bp; $\beta$-act5′/$\beta$-act3′ = 600 bp. The low molecular weight band is primer dimer. Epid. + v.d. = epididymis and vas deferens. (b) The same X- and Y-specific primer pairs were used to examine the expression of $Dffry$ and $Dffrx$ in mRNA from adult, 14 and 21 day mouse testis. Primers (Pgk2F and Pgk2R) from the testis-specific $Pgk2$ gene were used as a positive control. This gene first increases most dramatically in post-meiotic round spermatids (21 days).

**Interstitial deletions in patients with the AZFa phenotype remove the DFFRY gene**

Previous mapping studies had placed the $DFFRY$ gene sequences in proximal Yq11.2 between the markers $DYS1$ and $DYS246$ (1). This interval of the Y chromosome carries at least three genetic functions as determined by deletion analysis of affected individuals: short stature (4,5), skeletal anomalies associated with Turner syndrome (6) and male infertility as represented by the AZFa phenotype which in some patients histologically bears close
resemblance to Sertoli cell-only syndrome (2,16). Three patients (JOLAR, ELTOR and SAYER) defining the AZFa deletion interval were selected for analysis with markers from the PAC clones (11d8 and 229d16) as the markers Sy87, DFFRY -4.1 (coding position 4089–4542), DFFRY -JD (coding position 7211–7493) and DFFRY -3 (coding position 7634–7887). Since DFFRY -3 defines the very end of the DFFRY transcript, this result shows that the breakpoint in patient SAYER is close to the 3′ end of the DFFRY gene, within the span of one PAC clone. Thus clones 265f2, 229d16 and 11d8 cover approximately half of the DFFRY coding sequence.

The markers DFFRY -5′ (coding position 351–472, close to the 5′ end of the DFFRY transcript) and DFFRY -1.5 (coding position 1502–1835) are deleted from all three patients and not located on the PAC clones covering the 3′ half of the mRNA or on the PAC clones containing Sy82, Sy83, Sy84, Sy85 or Sy86. Four 30mer oligonucleotides from positions 1 to 30 (5pol-1), 150 to 180 (5pol-2), 800 to 830 (5pol-3) and 1200 to 1230 (5pol-4) of the open reading frame (ORF) were labelled and hybridised to the series of PAC clones shown in Figure 5. The two oligonucleotides (5pol-1 and 5pol-2) marking the extreme 5′ end of the ORF hybridised to the PAC clone 196h4, thus anchoring the 5′ end of the transcript to the marker Sy84. The other two oligonucleotides (5pol-3 and 5pol-4) lying 3′ to DFFRY -5′ failed to hybridise to any PAC clones. This demonstrates that the markers DFFRY -5′, 5pol-3, 5pol-4 and DFFRY -1.5 must lie in the gap between the PAC clones 196h4 and 229d16, but the relative order of Sy84 with respect to 5pol-1 and 5pol-2 is not known. As Gy6 is telomeric to Sy82, Sy83, Sy84, Sy85 and Sy86, these results demonstrate (i) that the DFFRY gene is orientated 5′ centromeric and 3′ telomeric and (ii) that the deletions in all three patients extend into DFFRY towards the 5′ end, resulting in the complete loss of the coding region of the gene.

STS content analysis of CEPH mega YAC and St. Louis YAC library clones mapping to the region covered by the deletions has shown that the clones 787G2, 711F4 and 900E7 must contain almost the entire DFFRY gene. These clones extend from Gy6 to Sy84 and do not contain Sy85 and Sy86, indicating that the PAC clone (143g20) linking Sy85 and Sy86 must be located in the gap between Sy84 and Sy83. The YAC clone SL147h5 contains Sy81, Sy82, Sy83 and Sy86, thus confirming the order of these STSs. A detailed long range restriction map of these YAC clones (19) has shown that the interval Gy6 to Sy84 covers 200 kb, the distance over which the DFFRY gene extends. The size of the interval from Sy84 to Sy83 has not been measured accurately but is estimated at a further 200 kb, making it possible that an additional gene(s) could have been deleted in these patients.

The mouse Dffry gene maps to the Sxr b deletion interval

In the mouse, the Sxr b deletion of the Y chromosome is associated with an early post-natal blockage of spermatogonial proliferation and differentiation which results in the adult testis being almost totally devoid of germ cells (20,21). As this phenotype presents
Figure 5. Summary of the results of the STS content analysis of three infertile patients with STS markers (Sy82–Sy89, Sy165, GY6 and GY8) from the AZFa interval on the human Y chromosome. Additionally, the patients have been analysed with various STS markers (DFFRY-5′, DFFRY 1.5, DFFRY-JD and DFFRY-3′). + denotes the marker is present in genomic DNA from the patient; N denotes that the marker is deleted from patient DNA; n.t. = not tested. The figure also displays the overlaps detected between PAC clones (determined by STS content analysis) isolated with markers from the AZFa interval. Also, PACs were tested with labelled oligonucleotide probes (5pol-1–5pol-4) from the DFFRY gene. The vertical lines intersecting PAC/YAC clones indicate the presence of the marker in the clone. The clones are not drawn to scale and reflect only the overlaps that have been determined.

similarities to that seen in AZFa-deleted men, the possibility that a functional DFFRY homologue exists in the Sxr<sup>b</sup> region of the mouse Y chromosome was investigated. Southern analysis of HindIII-digested genomic DNA from male and female C57BL/6 mice using a probe from the 3′ end of the human DFFRX cDNA revealed seven male–female common bands and a single faint male-specific band at 5 kb (data not shown). Screening of a library of size-selected HindIII-digested male genomic DNA with a DFFRX cDNA 3′ probe led to the isolation of a DFFRX homologous 5 kb DNA fragment. This fragment was subcloned, and subfragments hybridising to DFFRX cDNA were partially sequenced. Comparison of these partial sequences with DFFRX/DFFRY/Dffrx cDNA sequences revealed four ‘exons’, and virtual splicing using consensus splice sites bordering the regions of homology generated a ‘cDNA’ covering 658 bp of the putative coding region, with one gap where a further exon is predicted to lie. These sequences are shown in Figure 1b aligned against DFFRY. Figure 6a shows the restriction map of the 5 kb fragment and the location of the four potential exons of the mouse Dffry gene. All available spliced ‘exon’ sequence produces an ORF and has 84% nucleotide and 83% amino acid identity with the Dffrx gene. The preservation of an ORF despite divergence from Dffrx suggests that Dffry has been selectively maintained on the Y chromosome and remains functional.

DISCUSSION

Previous analysis of the X-linked gene DFFRX (1) had clearly established that it was the mammalian homologue of the Drosophila fat facets gene. The conserved Cys and His domains pointed to a biochemical function as a C-terminal hydrolase which could target specific proteins and remove conjugated ubiquitin. For both ubiquitin conjugation and its removal by hydrolases, there are two growing gene families which appear to function as antagonists in the regulation of protein degradation. Thus the activities of these genes have the potential to superimpose a further layer of fine control over the phenotypic contribution of a gene.

The DFFRY gene contains intact Cys and His domains which are almost identical to those found in fly Dffrx and DFFRX. Thus, it too is likely to function as a C-terminal ubiquitin hydrolase. The question arises as to whether the DFFRX and DFFRY genes are used as a probe in Southern analysis of HindIII-digested genomic DNA from C57BL/6 male, female, XXSxr<sup>a</sup> and XXSxr<sup>b</sup> mice. In XXSxr<sup>a</sup> mice, only the Sxr<sup>a</sup> region of the Y chromosome is present, whereas in the Sxr<sup>b</sup> variant of the Sxr<sup>a</sup> region there is a further deletion which removes the spermatogenesis factor Spy (see schematic in Fig. 6b) and several characterised genes. Hybridisation was seen to a 5 kb fragment present only in male and XXSxr<sup>a</sup> tracks, but not in female or XXSxr<sup>b</sup> tracks (Fig. 6c). Thus it can be concluded from the partial sequence and mapping data that Dffry, the mouse equivalent of the human DFFRY gene, has been identified on the mouse Y chromosome and that Dffry is located in the Sxr<sup>b</sup> deletion interval.
genes studied to date escape X-inactivation (1, 22–25) and thus the same biochemical and biological function. This is consistent with the fact that their encoded amino acid sequences are very similar mapping to the non-recombining portion of the Y chromosome. This is a general question that confronts all studies of X–Y homologous genes interchangeable with respect to function. This is consistent with the observation that in humans most of the X–Y homologous genes studied to date escape X-inactivation (1, 22–25) and thus dosage parity is established between males and females by providing functional interchangeability. Ubiquitous expression of DFFRX and DFFRY would appear to support this view. However, it remains possible that the differences in amino acid sequence between the X- and Y-linked genes may lead to subtle differentiation of biological function. For example, amino acid differences may alter substrate specificity, leading to the targeting of the deubiquitinating role to alternative proteins, thus resulting in a different biological function.

The testis-specific expression of Dffry in mouse suggests that the Y-linked gene has a different function to Dffrx, which, like DFFRX, is also expressed ubiquitously. The more restricted testis-specific expression of the mouse Dffry gene is reminiscent of another mouse X–Y homologous gene (Ube1y) where the X-linked homologue is subject to X-inactivation (26). Under these circumstances, parity of dosage for the X-linked gene exists between males and females, and selection may act on the Y-linked gene for a male-specific function. The X-inactivation status of Dffry is not known, but a similar situation to Ube1y/x may have arisen for the Dffry/Dffrx gene pair. The location of the DFFRY and Dffry genes in deleted regions of the human and mouse Y chromosomes that are associated in both species with an early failure of spermatogenesis is an intriguing finding, but only a circumstantial correlation with these spermatogenetic phenotypes.

There are three main areas of uncertainty that need to be resolved in order to determine whether DFFRY/Dffry have a role in spermatogenesis and infertility.

(i) The Sxr<sup>b</sup> deletion removes a number of other genes in addition to Dffry; these include Zfy-2 (20, 21), Ube1y (26), Smcy (27) and Uty (28). It could, therefore, be argued that loss of one or a combination of these genes, with or without Dffry, is responsible for the sterile phenotype. However, it is worth making the point that there is no human Y homologue of Ube1y and that (in contrast to DFFRY) ZFY, UTY and SMCY do not map to the AZFa deletion interval. Clearly, it will also need to be established that the mouse Dffry gene is functional and not a pseudogene by completing the sequence defining the full ORF.

(ii) The question arises as to whether the AZFa deletion interval carries other genes in addition to DFFRY. Analysis of this interval through the isolation of PAC and YAC clones has demonstrated that approximately half is occupied by the DFFRY gene. The telomeric breakpoint in patient SAYER (defining the telomeric limit of the interval) has been anchored to the 3′ end of DFFRY (between GY6 and DFFRY-3′) by being mapped to the PAC clones 1108 and 229d16. As these PAC clones are ∼120 kb in size, less than this amount of DNA is available for encoding a gene 3′ to DFFRY. This distance may be very much less once the precise location of GY6 and the 3′ end of DFFRY have been established on PAC clones 1108 and 229d6. However, one cannot exclude the possibility that there is a gene close to the 3′ end of DFFRY which is involved in the AZFa phenotype. The deletions in patients JOLAR and ELTOR are more extensive at the 3′ end of DFFRY by one STS marker than patient SAYER and have a more severe spermatogenetic phenotype. This may reflect the involvement of a further gene, possibly in combination with DFFRY or whose expression has been affected by the deletion of DFFRY. On the other hand, the milder phenotype in patient SAYER could reflect the action of particular alleles at loci that can modify the infertile phenotype. A gap in the contig exists between Sy84 and Sy83 where it has been difficult to close the contig covering the AZFa interval at the 5′ end of the DFFRY gene. This may be due to this part of the Y chromosome being unstable. From physical

Figure 6. Mapping of the Dffry gene to the Sxr<sup>b</sup> interval. (a) A restriction map of the 5 kb DNA fragment derived from the murine Dffry gene. Comparison of the sequence covering this genomic DNA clone with the cDNA sequence of the mouse Dffry gene has identified the position of six exons of Dffry (black boxes). Various primers for RT-PCR analysis, mapping and to generate PCR hybridisation probes were designed from the Y sequence. (b) A schematic representation of the Sxr<sup>a</sup> and Sxr<sup>b</sup> regions of the mouse Y chromosome showing the genes mapped to the Sxr<sup>b</sup> deletion. (c) Southern analysis of HindIII-digested genomic DNA from male, female, XXSxr<sup>a</sup> and XXSxr<sup>b</sup> mice using a PCR probe derived from the 5 kb DNA fragment shown in (a).
mapping of YAC clones in this region (19), it is estimated that this gap may be ~200 kb. It is probable the 5’ end of the gene (the 5’ UTR) extends further into this gap. Nevertheless, another gene(s) could lie in this gap and be responsible either singly or in combination with DFFRY for the AZFa phenotype. From the physical mapping of YAC clones containing most of DFFRY, the gene appears to cover ~200 kb of genomic DNA, which amounts to about half of the deletion interval. This is a sizeable gene, and raises the further possibility of additional genes contained within the introns of DFFRY. Complete sequencing of PAC clones covering this interval would be the most effective way of resolving these issues.

(iii) The pattern of expression of Dffry in RNA from staged mouse testis shows that it is first detectable between 7.5 and 10.5 dpp when type A and B spermatogonia and pre-leptotene and leptotene spermatocytes are present and increases considerably by day 14 (see Figures 3b and 4). This contrasts with the expression of Dfrrx. The correlation of Dfrrx expression with the expansion of the germ cell compartment would be consistent with the suggested cell-autonomous action of the Spy locus which demands that the gene(s) is expressed in germ cells. It is possible that Dfrrx is expressed in a stage-specific manner in the Sertoli cell and Leydig cell compartments, but it is more likely that changes in Dfrrx mRNA reflect differentiating germ cells. Clearly, it will be important to confirm the testicular compartment(s) within which the DFFRY/Dfrrx gene is expressed, either by RNA in situ hybridisation or RT-PCR analysis of RNA from fractionated testicular cells types. In man, the block in spermatogenesis manifested in the AZFa phenotype occurs at an early stage, yet DFFRY is not expressed in a stage-specific manner. This is not necessarily inconsistent with a function at a defined stage of spermatogenesis. A debuquinitating function for a constitutively expressed Dfrrx gene could act to prevent degradation of a protein product from a critical gene expressed specifically in spermatogonia or early spermatocytes. A stage- and testis-specific pattern of expression is not obligatory.

In summary, DFFRY is the first major gene to be located in the deletion interval defining the human AZFa phenotype and to have homologous sequences in the mouse Sry2 interval associated with the Spy spermatogenic phenotype. The mouse Dfrrx gene is expressed in testis in a stage-specific manner, but whether DFFRX/Dfrrx are involved in spermatogenesis and the AZFa or Spy phenotypes awaits analysis of non-deletion patients with Sertoli cell-only syndrome for evidence of mutation or micro-deletions in DFFRY, further analysis of the AZFa interval for the presence of other genes and more detailed expression studies of mouse testis mRNA.

**MATERIALS AND METHODS**

**PCR amplification**

All PCRs were performed typically in 20 µl volumes containing 10 mM Tris–HCl, pH 9.0, 1.5 mM Mg²⁺, 50 mM KCl, 0.1% Triton X-100, 0.01% (w/v) gelatin, 0.2 U Taq polymerase, 200 mM dNTPs, 20 pmol of each primer, 10–50 ng of template on a Biometra Trio thermocycler at 94°C for 4 min 30 s; 30–35 cycles of 94°C for 30 s, 55–60°C for 30 s, 72°C for 1 min; final extension of 72°C for 10 min. Unless otherwise stated, products were gel electrophoresed on 1% agarose at 100 V and visualised under UV illumination by staining with ethidium bromide.

**PCR screening for DFFRY cDNA**

Initially, 5’ and 3’ DFFRY clones were isolated by hybridisation of λgt10 retina and λgt11 testis cdNA libraries (Clontech) with DFFRX probes, which identified clones with sequence similar but not identical to the DFFRX sequence (1). Further 5’ and 3’ Y-specific cDNA clones were isolated by PCR screening of testis, retina and foetal brain (Clontech) cdNA libraries. Two rounds of PCR screening employed firstly a Y gene-specific primer, designed to include mismatches to DFFRX and directed into the gene, with a vector primer λgt11F (5’-CGGTTTCCATATGGGATGC-3’) or λgt11R (5’-GGCCTGCGCCGTATTTATAT-3’). PCR of the retina cdNA library employed NM1149F (5’-CCTTGTGGAATGGTCCAGCT-3’) or NM1149R (5’-AGAGGTGCTTATGAATTTCTTCTT-3’) vector primers. These reactions were performed at 55°C annealing for 30 cycles.

Single band PCR products were purified using the Wizard PCR Purification kit (Promega) for sequencing. Alternatively, when multiple bands were present, ~50 ng of purified PCR product was used in a second round of PCR employing a Y gene-specific primer and an appropriate nested vector primer: λgt11F (5’-ATTTGTGCGGAGCAGCT-3’), λgt11R (5’-TTTGTGACACGAGCCACACTG-3’), NM1149F (5’-CAGCGTTGATGTCCAGAG-3’) or NM1149R (5’-CCAGGTAAACGAGCAAG-3’). To enable cloning of the PCR products using the Cloneamp kit (Gibco BRL), the primers in each second round reaction were modified with CAU/CUA tails according to the manufacturer’s protocol. The PCR product (50 ng) was cloned into XL1-blue using the pAMP vector, single colonies were picked into 150 µl of L-broth containing ampicillin and 1 µl of this suspension was amplified using the primers pSPORTF (5’-GTAAAGACGACGCCAGTGA-3’) and pSPORTR (5’-CTATGACCATTGATACCAGCAAG-3’). The PCR product was purified for sequencing.

**Sequencing of DFFRY cDNA PCR products**

Purified PC products was quantified by agarose gel electrophoresis and sequenced using ABI Prism Dye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems). First round PCR products were sequenced using a nested gene-specific primer and the appropriate nested vector primer. Second round cloned PCR products were sequenced using pSPORTFcs (5’-CGGTACCTAGTGAATGATC-3’) and pSPORTRcs (5’-TAAGGGAAAGCTGTGACG-3’). Sequences were analysed on a 373A DNA Sequencer (Applied Biosystems).

**Expression analysis of DFFRY mRNA**

Human adult poly(A)+ cDNA libraries (in λgt10 or λgt11 vector) were obtained from Clontech. RNA was extracted from 9–10-week foetal tissues using the Tri-Reagent kit (Molecular Research Centre Inc.) as instructed by the manufacturer. cDNA was synthesised from 1 µg of total foetal tissue RNA using the Promega reverse transcription system and an oligo(dT) primer. The primer Y5 (5’-TGGCAAGTGTCACTCTGATGTC-3’) was designed from Y-only hybrid-derived sequence and paired with primer B17c2 (5’-CCCCTCACGAGATTTATCTT-3’) in PCR amplification of 1 µl of each of the above cDNAs at 55°C annealing for 35 cycles. Products were gel electrophoresed on 4% NuSieve 3:1 (FMC) at 100 V and visualised by ethidium staining.
Isolation of PAC clones
The RPCI1 library was obtained from the UK HGMP resource centre. The pools were screened by PCR using the published set of Y primers, Sy82, Sy83, Sy84, Sy85, Sy86 and Sy87 (29), GY6 (18) and the DFRFY-derived primer pairs DFRFY-5′F (5′-CC-AACGTTTTTCGAGATG-3′) and DFRFY-5′R (5′-CCAC-TAGGCGATGATTTG-3′); DFRFY-1.5F (5′-ATGTTGATGGCAGACCAAG-3′) and DFRFY-1.5R (5′-GCTCATTTTGTGTTGCGGACATG-3′); DFRFY-D (5′-GGGACATCTCCTGAGG-3′) and DFRFY-3.1F (5′-GAAAGATGATCTCCTACCTCAGATG-3′) and DFRFY-3.1R (5′-TTTATCTCCTGTTGTTTG-3′). The primer pair DFRFY-4.1F (5′-GGGACACGACGAAGAGAG-3′) and DFRFY-4.1R (5′-ACA-GCCAAACGCTAATGTCAAGGACATG-3′) were designed from the DFRFX sequence and found to cross-hybridise with the Y gene. PACs were obtained for all STS pairs tested with the DFRFY-derived primer pairs and found to cross-hybridise with the Y gene.

Dissection to dot-blot of primary PAC library

The overlaps between the genomic inserts were determined by a combination of PCR-based STS content analysis, primer hybridisation to dot-blots of fractionated library preparation protocols described in Mitchell et al. (30). Southern analysis was carried out using standard procedures. The primers oMJ324 (5′-CAAGAAGAACTCCATCTCTG-3′) and oMJ326 (5′-CAAGAAGAACTCCATCTCTG-3′), respectively, were used to amplify products from genomic DNA using the Reverse Transcription System (Promega) following the manufacturer's recommended protocol. One µg of mRNA was reverse transcribed in a final volume of 20 µl using the Reverse Transcription System (Promega) following the recommended protocol. One µl of these reactions was used in subsequent PCR. The primers Dfrfy-2F (5′-ATGCGAGTTGCA-CATCATCC-3′) and Dfrfy-2R (5′-GTCTTCTAACCCTGGAAGAGTTAGGATC-3′) were designed from the DFRFY probe for Southern analysis. The primer pair RbmlF (5′-CAAGAAGAACTCCATCTCTG-3′) and RbmlR (5′-CTCCCCAGAGATCTTATTAC-3′) amplified product from the Rbm gene. The primer pair Pm1F (5′-CAAGACTGATCTGCTTCTCAG-3′) and Pm1R (5′-GTAAAACAAAAACACCTTACTG-3′) amplified product from the Pml gene. The primer pair β-act5′F (5′-CTGATTTGCGTCAAGAAGATTAC-3′) and β-act3′ (5′-ACGGAGGCTGTAACCTTTCT-3′) amplified the product from the mouse β-actin gene. These PCR products were performed as described above and optimised to a final concentration of 1.25 mM Mg2+, annealing temperature of 60°C, extensions of 68°C for 3 min for 35 cycles.

Sequencing and mapping of DFRFY

Standard cloning and sequencing procedures were used to isolate the 5 kb HindIII Y-derived DNA fragment applying the fractionated library preparation protocols described in Mitchell et al. (30). Southern analysis was carried out using standard procedures. The primers oMJ324 (5′-CAAAATCTTGTGTGACAC-3′) and oMJ325 (5′-TGGCGAGTGCTCAACAGATG-3′) were used for PCR on genomic DNA to confirm localisation of the 5 kb DNA fragment to the Y chromosome. The primers oMJ324 and oMJ326 (5′-TTTTATTGTGGAGAGACGACATG-3′) were used for RT-PCR with testis RNA to produce a cDNA probe for Southern analysis.

ACKNOWLEDGEMENTS

We thank Dr Paul Burgoyne for the gift of cDNA prepared from RNA extracted from XXYxr and XXYxr testis and for the primer sequences for the Rbm gene. This work was supported by grants from the Wellcome Trust, the Medical Research Council and the European Union, R.P.E. was supported by Fellowship #42 USC 242-1 from the Fogarty International Center of the NIH.

REFERENCES


