Identification Of X Chromosome Genes
Causing Premature Ovarian Failure

Dr Rina M. Davison

Submitted for MD degree
ABSTRACT

The aetiology of premature ovarian failure (POF), defined as menopause before the age of 40, is unknown in the majority of cases but include radiation, chemotherapy, infections, autoimmunity and genetic disorders. Some genetic forms of ovarian failure are associated with single gene defects e.g. FSH receptor, or with known genetic diseases e.g. Fragile X syndrome. POF is also associated with deletions and translocations affecting the X chromosome, and studies of families with X chromosome rearrangements have indicated 2 loci on the long arm of the X at which this association is particularly strong. The POF1 locus is designated Xq26-28 and the POF2 locus is at Xq13-21 and it is presumed that these 2 loci relate to ovary determining genes.

One method of identifying ovarian determining genes is the study of naturally occurring defects of the X chromosome. Initial studies were conducted to determine the prevalence of abnormal karyotypes in seventy-nine POF patients. This identified a family in which POF had been transmitted through three generations in an X-linked dominant manner, in association with the karyotype 46XdelX(q26), a breakpoint interrupting the POF1 locus. The aims of the project were to determine possible POF candidate genes through characterisation of the breakpoint in this family, using the following methods:

- Polymerase chain reaction (PCR) using microsatellite markers
- Fluorescent in situ hybridisation (FISH) using P1 Artificial Chromosomes
- Computer based exon and gene searches

Four possible genes were identified, (see figure), two of which have reproductive roles in other species. Ovarian expression of these two genes was analysed using Northern Blotting.

Mutation screening of one candidate gene was carried out in idiopathic and familial POF patients.

Summary of POF1 locus candidate gene identification
Table of contents

ABSTRACT 2

1. INTRODUCTION 7
   1.1 Determinants of ovarian lifespan 10

1.2 CAUSES OF OVARIAN FAILURE 15
   1.2.1 Environmental factors - chemotherapy, radiation, viruses. 16
   1.2.2 Autoimmunity 18
   1.2.3 A genetic basis for premature ovarian failure (POF) 21
   1.2.4 Evidence for X chromosome genes causing POF 29
   1.2.5 Candidate genes for ovarian development on the X chromosome 32

2. PROJECT STRATEGY 41

STUDY 1 PREVALENCE OF CYTOGENETIC ABNORMALITIES IN POF PATIENTS 43

STUDY 2 IDENTIFICATION OF FAMILIAL BREAKPOINT GENE 45
   Microsatellites 45
   EST strategy 50
   Fluorescent in situ hybridisation (FISH) using cloned genomic sequence 51
   Computer Based Gene Searches 52

STUDY 3. CHARACTERISATION OF HYBRIDS MIMICKING BREAKPOINT 60

STUDY 4. CHARACTERISATION OF A CANDIDATE POF GENE 63

STUDY 5. MUTATION SCREENING OF A CANDIDATE POF GENE 63

3. SUBJECTS, METHODS & REAGENTS 64

3.1 RAPID DNA EXTRACTION FROM WHOLE BLOOD 64

3.2 PREVALENCE OF CYTOGENETIC ABNORMALITIES IN POF PATIENTS 64

3.3 POLYMERASE CHAIN REACTION 65

3.4 AGAROSE MINIGELS 69

3.5 FLUORESCENT IN SITU HYBRIDISATION (FISH) 69

3.6 PAC MINIPREP METHOD 72

3.7 CHARACTERISATION OF SOMATIC CELL HYBRID PANEL 73
List of tables and figures

Figure 1 - Decline in the number of germ cells in the ovary throughout life 9
Figure 2 - Examples of factors influencing follicular development 14
Figure 3 - FRAXA 27
Figure 4 - Putative ovary determining genes on the X chromosome 33
Figure 5 - Breakpoint identification strategy 44
Figure 6 - STS map of breakpoint region 47
Figure 7 - PACS spanning region between two microsatellites 53
Figure 8 - Hybrids mimicking breakpoint 62
Figure 9 - Family tree of index case 88
Figure 10 - ABI prism results of DXS 8071 and DXS 8033 91
Figure 11 - Semi-qualitative PCR results 95
Figure 12 - FISH results 98
Figure 13 - Microsatellite mapping onto PACS 108
Figure 14 - NIX analysis of PAC dA227L19 110
Figure 15 (a) - NIX analysis of PAC dJ97K10 0-7000 bp showing a predicted exon rich region from 2693 - 6238 bp 114
Figure 15 (b) - NIX analysis of PAC dJ97K10 48 000 - 63 000 bp showing exon rich region from 57 000 - 60 700 bp 115
Figure 16 (a) - NIX analysis of PAC dJ358H7 0-7000 bp, showing a predicted exon rich region from 510 -2906 bp 117
Figure 16 (b) - NIX analysis of PAC dJ358H7 95 000 -107 000 bp bp, showing a predicted exon rich region on the reverse strand from 104 241 -105454 bp 118
Figure 17 - NIX analysis of PAC dJ363L9 57 000 - 85 000 bp showing two predicted exon rich regions on the forward strand 120
Figure 18 - Mapping of DXS692 onto somatic cell hybrids 121
Figure 19 - Mapping of microsatellites onto somatic cell hybrids 122
Figure 20 - Mapping of HS1 onto somatic cell hybrids 123
Figure 21 - Mapping HS1 against monochromosomal hybrid panel DNA 124
Figure 22 - Northern blot of E2F 125
Figure 23 - Summary of candidate gene identification 131
Figure 24 - Map of the POF1 locus of the X chromosome 132
Table 1 - Causes of ovarian failure 15
Table 2 - Causes of POF in 323 women attending the Middlesex Hospital 16
Table 3 - Method of interpreting ABI results. 48
Table 4 - Primer pairs and their attributes 68
Table 5 - Primers for mutation detection 79
Table 6 - Results of PCR and ABI Prism 90
Table 7 - Example ratios of peak areas autosomal /X chromosome marker 96
Table 8 - Summary of FISH results 97
Table 9 - Predicted exon regions on PAC dJ97K10 with BLAST analysis results 113
Table 10 - Matches of HS6ST with PACS around breakpoint 113
Table 11 - Mapping of GPC3 onto PACS around breakpoint 113
Table 12 - Predicted exon regions on PAC dJ358H7 with BLAST analysis results 116
Table 13 - Mapping of E2F along PAC dJ358H7 116
Table 14 - Predicted exon regions on PAC dJ363L9 with BLAST analysis results 119
Table 15 - Genes predicted to lie in the breakpoint region 119
1. INTRODUCTION

Women are susceptible to ovarian failure throughout their life, in that it may be present from birth, develop during active reproductive life or occur naturally in the fourth to fifth decade of life. Premature ovarian failure (POF) is defined as the cessation of menses for a period of greater than six months associated with elevated gonadotrophin levels, before the age of 40 years (Coulam et al 1986). The estimated prevalence varies between 0.3% – 6%. The age of menopause has remained constant across generations and is known to be strongly inherited, but it is also affected by environmental factors such as smoking (Kaufman et al 1980).

In the normal situation, a number of factors or requirements contribute to ovarian development.

- Absence of genetically active Y material and presence of two intact and genetically active X-chromosomes in the germ line.

- Normal migration of germ cells in the genital ridge

- Normal mitosis of germ cells

- Development of normal granulosa cell layer

- Normal transformation of the oogonium into a primary oocyte with arrested meiotic prophase

- Physiological response of the primary follicle to fetal FSH.
Germ cell number is affected by four phases, about which little is known (figure 1).

- The control of mitosis in utero
- The apoptotic processes that bring about the destruction of two-thirds of germ cells before birth
- The apoptotic processes which cause both 90% destruction by 40 years
- Accelerated destruction after the age of 40.

Interference with any of these mechanisms may compromise ovarian development and result in failure to function or limit their functional lifespan. The ovary comprises four cell types: germ cells, granulosa cells, theca cells and support cells. The most vital functions of the ovary, producing gametes and oestrogen, are determined by germ cells and granulosa cells respectively, and hence ovarian failure can be considered in terms of these two components. While in some cases, POF is caused by a reduced complement of oocytes present in the ovaries at birth, either as a result of defective germ cell migration to the gonad or impaired multiplication, in most patients the cause is thought to be accelerated atresia of uncertain aetiology. This atresia of germ cells is thought to be due to apoptosis. Some mechanisms of ovarian failure primarily result in germ cell depletion (e.g. X chromosome abnormalities) whilst others target granulosa cells only (e.g. FSH receptor mutations).
Figure 1 - Decline in the number of germ cells in the ovary throughout life

![Graph showing the decline in the number of germ cells in the ovary throughout life.](image)
1.1 Determinants of ovarian lifespan

There follows a brief description of the determinants of ovarian lifespan. However for each step various signals come into play dependent on species, cell and tissue type. There is an ever increasing array of growth factors, cytokines, proteases and so on which have been shown to have a role in cell differentiation, growth and death, and hence just a few examples for illustration have been included.

Migration

The stem cells which give rise to oocytes are first visualized by alkaline phosphatase staining whilst they are still in the yolk sacs at approximately 4 weeks gestation. Subsequently, these primordial germ cells migrate to the genital ridge, continually multiplying, with some ending up in inappropriate locations and dying. Defects in gonadal development occurring as a consequence of an error in migration could be manifested by two events. Either complete lack of migration of sex cells to the urogenital ridge resulting in streak gonads, or migration of fewer than average sex cells resulting in gonadal hypoplasia or premature depletion of ovarian follicles. Survival and successful migration of these cells depend on the local production of cytokines, including kit-ligand and transforming growth factor beta1, (Gosden RG 1995) and hence in theory failure of certain cytokine production will affect germ cell migration. However in view of the fact that these cytokines have additional roles, it is difficult to isolate a pure migratory defect of germ cells.
Mitosis and meiosis

Following several rounds of mitotic amplification in the gonad, the germ cells or oogonia start to enter meiosis at the beginning of the second trimester. The mechanisms by which mitosis and meiosis are controlled are largely unknown, although it is clear that several proteins are involved, and there may be more than one signaling pathway. It is known that cyclic adenosine phosphate (cAMP), kit-ligand and growth differentiation factor 9 (GDF-9) can affect levels of meiosis, (Ismail et al 1996, Elvin et al 1999). The net effect of mitosis and meiosis is a germ cell peak of 6 million by 20 weeks gestation, followed by germ cell atresia or apoptosis.

Apoptosis

Physiological cell death or apoptosis occurs in proliferating and differentiating tissues. This term is derived from the Greek words apo and ptosis meaning to fall away from. It is used to describe a series of events resulting in developmental and homeostatic cell deaths controlled by the body and is distinguished from accidental cell death initiated by a noxious insult (Kerr et al 1972). In cellular terms, it is characterised by separation of the cell from its neighbouring cells, loss of cell volume, chromatin condensation and margination along the nuclear envelope (nuclear pyknosis) and budding and fragmentation of the cell into plasma membrane-bound vesicles called apoptotic bodies (Frisch 1994). The final outcome of the apoptotic body is phagocytosis by resident macrophages or neighbouring epithelial cells, which prevents secondary necrosis and inflammation (Bursch et al 1990).
Normal development of both female and male gonads is characterised by massive cell death. More than 99% of ovarian follicles endowed at early life are destined to undergo apoptosis and recent studies in other cell systems have shown that there exists a cohort of structurally and functionally conserved gene products that comprise the cell death machinery. In the ovary, gonadotrophins, oestrogen and several growth factors can act as survival factors to rescue follicles from apoptotic demise, where as androgens and tumour necrosis factors are atretogenic factors (Hsueh et al 1994). A particular example is the expanding family of Bcl-2 proteins that consists of different anti and proapoptotic members which allow the regulation of apoptosis in a tissue specific manner (Hsu et al 2000). The Bcl-2 proteins can be divided into three subgroups: 1) antiapoptotic proteins with multiple Bcl-2 homology (mainly Mcl-1) which when overexpressed in mice leads to diminished apoptosis, increased ovulation and favours cell survival. (2) & (3) - proapoptotic proteins such as Bok (Bcl-2 related ovarian killer) BOD (Bcl-2 related ovarian death agonist) and BAD (Bcl-2 associated death promoter) which differ in their transmembrane domains. Overexpression of BAD induces apoptosis in granulosa cells. In addition, the Bcl-2 members interact with various upstream and downstream follicle survival factors (e.g. 14-3-3 – a group of proteins involved in intracellular signalling and cell cycle progression (Hsu et al 1997) and proteases (e.g. caspases- proteases causing cell death through proteolytic cleavage Steller 1995). Future investigations focussing on the role of Bcl-2 members and their interacting proteins could provide insight into the normal physiology of ovarian follicle cell atresia.
**Primordial follicle maturation**

The oogonia, which have escaped apoptosis, are enveloped by pregranulosa cells to form primordial follicles. These then change to cuboidal cells and secrete mucopolysaccharides forming a translucent covering known as the zona pellucida. These primary follicles develop FSH, oestrogen and androgen receptors. The theca interna cells develop LH receptors, and these secondary follicles continue their growth becoming late antral follicles during the postnatal prepubertal period.

**Female gametogenesis**

The ability to produce knockout mice for various genes has resulted in a better understanding of follicular development. An example is the GDF-9 knockout female mouse which is infertile because of a block early in folliculogenesis (Carabatsos et al 1998, Elvin et al 1999). Most oocytes are capable of resuming meiosis but the follicles lack LH and 17 alpha hydroxylase receptors, i.e. the follicle are incompetent to emit a signal that recruits theca cell precursors to surround the follicle. The granulosa cells not only fail to proliferate but also do not undergo apoptosis implying the appropriate cell death signal is lacking.

Other knockout studies involving different growth factors and apoptotic proteins are slowly enlightening us to the normal steps of differentiation within ovarian follicles and follicular lifespan, with the future developments of new therapeutic approaches to POF.

Figure 2 depicts the ovarian life cycle with example of regulators.
Figure 2 - Examples of factors influencing follicular development

- **Kit-ligand**
  - migration
  - stem cell differentiation

- **GDF-9**
  - primordial follicle
  - meiosis
  - follicle maturation & ovulation

- **FSH**
  - apoptosis

**Pathways:**
- Yolk sac
- Genital ridge
- Ovary

**Stages:**
- 1st trimester
- 2nd trimester
- Birth
- Puberty → Menopause
1.2 CAUSES OF OVARIAN FAILURE

The aetiology of POF is unknown in the majority of cases but include radiation, chemotherapy, infections, autoimmunity and genetic disorders (Alper et al 1986) (Table 1). A study at the Middlesex Hospital showed that the cause of ovarian failure could only be identified in half the women, leaving a large *idiopathic* subgroup (Table 2). This so called *idiopathic* group encompasses patients in whom a suitable marker or diagnostic test is not as yet known to define the cause of POF.

Table 1 - Causes of ovarian failure

<table>
<thead>
<tr>
<th>Category</th>
<th>Causes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonadal dysgenesis</td>
<td>Turner's syndrome</td>
</tr>
<tr>
<td></td>
<td>46XX gonadal dysgenesis</td>
</tr>
<tr>
<td></td>
<td>46XY gonadal dysgenesis</td>
</tr>
<tr>
<td>Infection</td>
<td>Viral oophoritis</td>
</tr>
<tr>
<td>Iatrogenic</td>
<td>Chemotherapy</td>
</tr>
<tr>
<td></td>
<td>Radiotherapy</td>
</tr>
<tr>
<td>Autoimmune</td>
<td>Autoimmune Polyendocrinopathy Syndrome 1</td>
</tr>
<tr>
<td></td>
<td>Autoimmune Polyendocrinopathy Syndrome 2</td>
</tr>
<tr>
<td></td>
<td>Association with other autoimmune conditions</td>
</tr>
<tr>
<td>Genetic associations</td>
<td>Familial ovarian failure</td>
</tr>
<tr>
<td></td>
<td>FSH receptor mutations</td>
</tr>
<tr>
<td></td>
<td>BPES</td>
</tr>
<tr>
<td></td>
<td>FRAXA premutations</td>
</tr>
<tr>
<td></td>
<td>Galactosaemia</td>
</tr>
<tr>
<td></td>
<td>Small X chromosome defects</td>
</tr>
<tr>
<td>Idiopathic</td>
<td>? cause</td>
</tr>
</tbody>
</table>
Table 2 - Causes of POF in 323 women attending the Middlesex Hospital

<table>
<thead>
<tr>
<th>CAUSE</th>
<th>NUMBER</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiopathic</td>
<td>92</td>
<td>59</td>
</tr>
<tr>
<td>Turner's syndrome</td>
<td>73</td>
<td>23</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Familial POF</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Pelvic surgery</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Pelvic Irradiation</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Galactosaemia</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>46XY gonadal dysgenesis</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

1.2.1 Environmental factors - chemotherapy, radiation, viruses.

Extrinsic factors can interfere with ovarian function at any stage of development, from germ cell migration to follicular damage of the fully developed ovary. Chemotherapeutic drugs target rapidly dividing cells, both malignant and benign. In the ovary, these agents disrupt proliferating granulosa and theca cells of mature follicles causing amenorrhea during therapy (Gradishar et al 1989). Alkylating agents such as cyclophosphamide destroy cells by DNA damage and hence cause permanent damage to the ovary by affecting resting
oocytes and proliferating cells (Epstein 1990). The ovarian response to chemotherapy of individuals varies with age, dose and type of drug. It seems women under 40 are more resistant to the damaging effects of these drugs, with higher doses required to render amenorrhoea (Koyama et al 1977). However, with the increasing success of treatment for childhood malignancies more females are surviving to reproductive age, and the cytotoxic treatments used in treating leukaemia or Hodgkin’s disease cause ovarian failure in approximately 50% females (Clark et al 1995)

Mouse models have demonstrated that murine oocytes undergo rapid apoptosis when cultured in vitro with chemotherapeutic agents such as Doxyrubicin or Adriamycin (Perez et al 1997). From understanding the mechanisms of apoptosis, the BAX gene has been implicated in ovarian damage caused by chemotherapy. BAX knockout mice exhibit preserved primordial follicle numbers following chemotherapy, and also prolongation of ovarian lifespan into advanced chronological age (McCurrah et al 1997, Perez et al 1999). This is of great clinical significance concerning the pathological loss of oocytes in women undergoing chemotherapy which and may offer future benefits in the human ovary through gene therapy.

Irradiation to the pelvis can cause ovarian dysfunction through oocyte damage. Radiation induced ovarian failure, like chemotherapy is dependent on the age of the patient and the dose received. The most pronounced effect of radiation is in the early stages of follicle development, destroying oocytes and sparing mature follicles (Baker et al 1971). Doses of radiation above 600 rads is associated with ovarian failure in nearly all women over 40 years but has a more variable effect
on younger women. In addition, total body irradiation damages the uterus
making successful pregnancies after egg donation a rare event. In pregnancies
following chemotherapy there is a slightly increased rate of miscarriage but no
increased risk of congenital abnormalities (Averette et al 1990).

A prior record of viral infection (mumps) was originally described in POF. A
retrospective review of POF patients showed that 3.5% reported a preceding
viral infection including varicella, shigella, and malaria (Rebar et al 1990).
Although studies of patients during mumps epidemics have shown a 3%
incidence of oophoritis there is no long term follow up and documentation of
subsequent ovarian failure. Infections are a possible cause of POF but the
actual incidence will remain unknown in the absence of longitudinal prospective
studies.

1.2.2 Autoimmunity

There is an increasing body of evidence that suggests an autoimmune aetiology
in some POF patients. Data in support of an autoimmune aetiology include:

1. lymphocytic and plasma cell infiltration of the ovary and alteration of T cell
   subsets.

2. Circulating autoantibodies to ovarian antigens.

3. Association with other autoimmune disorders

4. Recovery of ovarian function after regression of autoimmune status.

The first cases of autoimmune disease and POF were described in 1954 (Guinet
et al). Initial reports focussed on the incidence of POF in patients with Addison's disease, suggesting that 10% of these patients have POF preceding hypoadrenalism by a number of years (Irvine et Barnes 1974). In addition, hypothyroidism develops concurrently or shortly after the Addison's. Hypothyroidism has subsequently been shown to be the most commonly associated autoimmune disease with POF with a prevalence of between 10 and 30% in POF women. Other autoimmune diseases coexisting with POF include Myasthenia Gravis, Crohn's, SLE and rheumatoid arthritis.

The specific polyendocrine autoimmune syndromes 1 and 2 exist in as many as 3% of POF patients (Kim et al 1995).

*Autoimmune polyendocrinopathy syndrome (APS) type 1* (Betterle et al 1998)

APS1 is a rare autosomal recessive syndrome comprising mucocutaneous candidiasis, hypoparathyroidism, Addison's disease. Hypothyroidism, IDDM, pernicious anaemia and ovarian failure are additional features of the syndrome. Ovarian failure occurs in 60% of affected females. The clinical picture of autoimmune polyendocrinopathy, candidiasis and ectodermal dystrophy is also known as the APCED syndrome. APS1 has no HLA association and is caused by mutations in the AIRE gene which codes for a transcription regulator.

*Autoimmune polyendocrinopathy syndrome (APS) type 2- Schmidt's syndrome* (Betterle et al 1996)

APS2 comprises Addison's disease, hypothyroidism, IDDM and ovarian failure. The sequence of glands is variable with up to 15 years spanning the first and
the last. APS2 is associated with the HLA haplotypes HLA-B8, DR3, and DR4, and the target antigens are P450 cytochromes from the steroidogenic pathway: P450c17, P450c21 and P450scc.

Autoimmunity in isolated ovarian failure

The frequency, pathophysiology and potential reversibility of autoimmune oophoritis needs clarification, but as yet there is no gold standard test for the diagnosis of ovarian autoimmune disease. Autoimmunity is clearly implicated in about 30% of women with ovarian failure (Conway et al 1996). The problem is however, that routine tests of ovarian autoimmunity are not sufficiently sensitive and they therefore have a high false negative rate. Prevalence of ovarian antibodies in POF patients varies from 0-67%, depending on the specific type of ovarian antigen, the detection technique employed and the population studied (Moncayo et al 1995). Many groups have tried to identify a specific autoantigen within the ovary in order to improve our ability to define an autoimmune subgroup of women with POF. Steroidogenic enzymes and the FSH receptor have all been candidate autoantigens (Anasti et al 1995, Chen et al 1996). Recently Arif et al (1997) have shown that 3-beta hydroxysteroid dehydrogenase is relatively specific for POF.

The most important feature of ovarian autoimmunity is whether some women with POF experience remission when treated with glucocorticoids as a result of dampening down the immune process. Uncontrolled studies and anecdotal cases of the return of ovarian function following the use of high dose steroid immunosuppression have been reported (Coulam et al 1981) but at present
there is no proven benefit.

1.2.3 A genetic basis for premature ovarian failure (POF)

Progress in the identification of ovarian determining genes has been slow. Ever since the experiments of Jost (1953) predicting the male determining role of testosterone and mullerian inhibiting factor, the dogma of sex differentiation has maintained that the female route is the default pathway from which male differentiation departs. This is evidenced by the ‘male switches’ at each point in this pathway:- genetic sex and the effect of the Y chromosome, gonadal sex and testis determination, and hormonal sex and the effects of androgens.

Clearly, this simplistic view obscures the fact that normal female development must be the end result of a cascade of genetic programming. First amongst these events must be the development of the ovary which has until recently not received the attention devoted to testis determination. Similarly, knowledge of the genes controlling spermatogenesis is far more advanced than that relating to oogenesis. Recently however, attention has turned to the search for ovary determining genes. In particular, what is the influence of the X chromosome in driving female development and what other genetic signals across the genome are vital for ovarian development?

Familial POF

Ovarian failure serves as a perfect model for the study of ovarian development and function. Several pedigrees have been described in the literature with more than one affected member with POF or ovarian dysgenesis (Coulam et al 1983,
Estimates vary as to the prevalence of familial POF from 5 to 30% depending on the referral base of each clinic and the detailed family history. Recently Vegetti et al (1998) studied the inheritance patterns in 71 cases of idiopathic POF and found that 71% had familial POF, suggesting a genetic cause. In particular their pedigree analysis suggests either an autosomal or X-linked dominant sex limited pattern of inheritance irrespective of maternal or paternal transmission. Some forms of ovarian failure are associated with single gene defects e.g. FSH receptor, or with other known genetic diseases e.g. Fragile X syndrome. These forms are discussed below, whilst the role of the X chromosome and POF will be presented in section 1.25. The list of genetic associations listed below is growing each year, but the fact remains that the genetic causes of familial POF are unknown in the majority of cases.

**Gonadotrophin receptor gene mutations**

Both the LH and the FSH receptor genes map to the short arm of chromosome 2, and the receptors belong to the G protein coupled receptor superfamily. Inactivating LH receptor mutations in genetically male individuals leads to Leydig cell hypoplasia with varying phenotypic presentations ranging from pseudohermaphroditism with female external genitalia, to partial male sexual differentiation with micropenis depending on the degree of impairment of receptor function. Genetic females present with normal pubertal development but primary amenorrhoea in the absence of preovulatory follicles and anovulation. Histological analysis of their ovaries shows normal dimensions with follicles at all developmental stages up to the preovulation stage (Themmen et

Complete loss of function of the FSH receptor results in primary amenorrhoea, whereas partial loss of function results in secondary amenorrhoea. In Finland an Ala 189Val mutation affecting the extracellular domain of the FSH receptor gene was found to be associated with POF in 22 women from six families presenting with primary amenorrhoea (Aittomaki et al 1996). The histological appearance of the ovaries in these women showed hypoplasia with scant primordial follicles rather than ovarian dysgenesis. Conversely, male members of these families who shared the mutation had diminished testicular weight, minimal suppression of sperm count and in some cases were fertile (Tapanainen et al 1997). A compound heterozygote of this mutation, causing partial loss of function of the FSH receptor, presented with secondary amenorrhoea and her ovaries were of normal size containing small antral follicles (Beau et al 1998). The Finnish mutation appears to be restricted to that particular geographically isolated population since it has not been found in series from other countries (Whitney et al 1995, Conway et al 1999), and hence is a very rare cause of POF.

The evaluation of the phenotypes of individuals carrying mutations in their gonadotrophin receptor genes has helped establish the differing physiological roles of LH and FSH in male versus female reproductive function. It is apparent that the absence of LH action during development is more deleterious for the male than the female phenotype, although both sexes are infertile. However, in the case of FSH receptor inactivating mutations, females are infertile, but males retain their fertility, indicating that FSH is essential for female reproduction but
not an absolute requirement in the male.

*Blepharophimosis, epicanthus inversus and ptosis (BPES)*

The blepharophimosis, ptosis, and epicanthus inversus syndrome (BPES) was first described by Vignes in 1889. Townes and Muechler (1979) reported a family in which all affected females had primary ovarian failure, a normal female karyotype and a small uterus with small atrophic ovaries on laparoscopy. Since then, there are known to be 2 forms of BPES: type I with infertility of affected females; type II with transmission by both males and females (Zlotogora et al 1983). The BPES locus is on chromosome 3q23 as evidenced by deletions and translocation associated with the syndrome, which is inherited as an autosomal dominant sex-limited trait. Recently molecular cytogenetic evaluation of 2 patients with BPES type 1 has narrowed the BPES locus to within a 10.5 kb interval, and sequencing of this area will greatly facilitate efforts toward isolating the gene(s) involved in the condition (De Baere et al 1999, Praphanphoj et al 2000).

The ovarian failure is often intermittent with a clinical picture of “resistant ovaries” (Amati et al 1996) or can be a true premature menopause. Two pathophysiological mechanisms have been proposed: either that follicle atresia occurs at an early stage of development or that follicular function is disrupted later on in life. It is as yet unknown whether BPES types 1 and 2 result from microdeletion of genetic material containing at least two independent genes or whether BPES type 1 occurs as a consequence of the deletion of a single BPES gene and a nearby “ovarian function” gene. Female infertility in type I is a
predominant symptom, and the distinction between the two types is of importance for genetic counseling.

FRAXA

The fragile X syndrome is a relatively common cause of mental handicap, and is due to expansion of a polymorphic CGG trinucleotide repeat in the 5' untranslated region of the FMR1 gene at Xq27.3. Normally the number of repeats is 5 to 50. Premutations have 50 – 200 repeats, and do not affect the synthesis of the gene product FMRP. The full Fragile X (FRAXA) mutation of greater than 200 repeats is associated with methylation of the promoter and silencing of gene transcription. Absence of the gene product FMRP is associated with the clinical syndrome of mental retardation (Figure 3).

Studies have demonstrated that 13-25% of fragile X carriers experience POF (Allingham-Hawkins et al 1999, Uzielli et al 1999), and conversely 3-15% of women with idiopathic POF harbour a fragile X premutation, compared with an expected prevalence of 1:590 (Conway et al 1998, Uzielli et al 1999). Also, there were no full mutations in the POF population confirming previous reports of the association between FRAXA and POF being restricted to premutation carriers (Cronister et al 1991, Schwartz et al 1994).

Interpretation of this finding into a molecular mechanism is perplexing. When the FMR1 gene is inactivated by methylation as in the full FRAXA mutation, there appears to be no detrimental effect on ovarian function. Protein studies have been unable to detect a difference in the expression of the FMR1 protein from premutation alleles as compared with normal (Devys et al 1993, Feng et al
1995), and yet premutation alleles clearly have an adverse effect on ovarian function. One theory is that although the FMR1 protein is absent in the full mutation, its function is provided by a back up protein. In the case of premutations however, where FMR1 is expressed, the alternative mechanism fails to come into play and any adverse effect caused by the premutation is expressed. Zhong et al (1993) described a second mutable sequence within the FMR1 gene. AC1, a polymorphic marker flanks the unstable CGG repeat by 10 kb and is in linkage disequilibrium with it. The finding of a second mutable locus within FMR1 suggests that the target for tandem repeat instability may not be confined to the CGG repeat alone but may also involve microsatellites.

Another possibility is that a particular isoform of FMR1, crucial during oogenesis is less efficiently produced from premutation alleles. Murine studies have shown that FMR1 is particularly strongly expressed during the mitotic phases of oogenesis (Bachner et al 1993) and so any changes in expression at this critical time could dramatically reduce the number of oocytes. Against this hypothesis are the results of studies on at least 11 isoforms of FMR1 showing no differences between tissues of normal compared with premutation alleles (Devys et al 1993, Ashley et al 1993, Verheij et al 1995, Khandjian et al 1995). In conclusion, there is an association between FRAXA premutations and POF but the exact role of the FMR1 gene in reproduction is as yet undetermined.
Figure 3 - FRAXA

(A) promoter, 5' untranslated region in exon 1 with (AUG) start of transcription, FMR1 gene structure.
(B) normal trinucleotide repeat in exon 1. (C) FRAXA premutation. (D) FRAXA full mutation with methylation of the promoter and silencing of gene transcription.
Galactosaemia

Galactosaemia is inherited as an autosomal recessive disorder caused by mutations of the GALT gene mapped to locus 9p13 (Shih et al 1982, 1984). The cardinal features of galactosaemia are failure to thrive, hepatomegaly, cataracts, and mental retardation, which result from a deficiency of galactose-1-phosphate uridyltransferase (GALT). Gonadal dysfunction, specifically hypergonadotrophic hypogonadism, in female galactosaemics is an almost universal finding. Minimal transferase activity may modulate ovarian function and the time of menopause. In contrast, male galactosaemics have a relatively low risk of gonadal dysfunction. Current dietary restrictions can prevent liver and brain damage, but are inadequate to prevent ovarian failure.

Galactose metabolism was studied by Xu et al 1989, in human ovarian tissue obtained from 14 women controls between 21 and 72 years of age, and one 21-y-old galactosaemic patient with hypergonadotrophic hypogonadism. Activities of enzymes related to the galactose pathway were measured in ovarian homogenates using radioisotopic, spectrophotometric, and fluorometric techniques. In normal ovarian tissue, specific activities of galactokinase, transferase, epimerase, and UDPGlc pyrophosphorylase are much higher than those found in the red cells and in testes, whereas in tissue from the galactosemic individual, minimal enzyme activity occurred. It is hypothesized that the ovarian failure in patients with galactosemia is due to interference with nucleotide sugar metabolism and the synthesis of galactose containing glycoproteins and glycolipids consequent to the enzymatic defect in the major
pathway of galactose metabolism. Hence although galactosaemia is an inherited disorder, a metabolic rather than a genetic defect causes the ovarian failure.

Interestingly, a survey of 108 heterozygote women for the classic galactosaemia gene did not reveal that the carrier state was associated with POF (Kaufman et al 1993).

1.2.4 Evidence for X chromosome genes causing POF

The sex chromosomes evolved from autosomes (Rice 1994). The Y chromosome shed much of its genetic material and became a concentrated site of male determining genes which behave in a dominant manner i.e. only one copy is required for the testis determining role of the SRY gene and for the action of spermatogenesis gene families – AZF, DAZ and RBM (Saxena et al 1996). In fact, male determining genes comprise approximately half of the 20 genes on the Y chromosome. By comparison, the X chromosome has retained over 300 genes and therefore the sex-determining role of the X is much more dilute than that for the Y chromosome. In female mammals, dosage compensation of X-linked genes between males and females occurs by genetic inactivation of one of the two X chromosomes. Both the paternally and maternally inherited X chromosomes are active in preimplantation embryos after fertilization, and X inactivation occurs around the time of implantation. The majority of genes on the X chromosome are inactivated in females although those that have a homologue on the Y, escape inactivation – both sexes require two copies, and these genes are thought to have house keeping functions in
males and females (Lahn et al 1997). Other genes on the X, with no Y homologue, also escape inactivation and only females have two copies of these. This group of genes is a prime candidate for ovary determination.

Further evidence for an ovarian role for genes that escape X inactivation comes from an experiment of nature, Turners syndrome (45X0). Short stature, ovarian dysgenesis and primary amenorrhoea are features of Turners syndrome (TS). In fact, early ovarian development is normal in Turner foetuses but in the third trimester of gestation the ovary is rapidly destroyed so that few ova remain at birth. It is presumed that accelerated germ cell destruction in utero is the result of unstable meiosis which is disturbed by the absence of one X chromosome. Women with TS differ from normal women by the genes which escape X inactivation and are transcribed only in single copy in the 45X0 state. Hence candidate genes for ovarian failure in Turner’s syndrome are likely to be those which escape X inactivation (Stratakis et al 1994).

Study of naturally occurring defects of the X chromosome go some way to identifying ovarian determining genes. Deletions and translocations of the X chromosome are associated with ovarian dysgenesis or premature ovarian failure, and several theories have been postulated to explain this phenomenon.

- *Position effect*, in which the actions of specific genes are altered when the chromosomal segment is disrupted (Sarto et al 1973)
- Gene disruption as a consequence of deletion/translocation causing *haploinsufficiency* for critical X linked genes (Zin et al 1993)
- Gene rearrangements causing *meiotic arrest* during oogenesis and germ cell atresia (Speed 1988)
It appears that in general, abnormalities of the long arm of the X chromosome affect only ovarian function while those of the short arm affect stature as well (Sarto et al 1973, Therman et al 1990). In 1973 Sarto et al proposed the 'critical region' hypothesis that the region on Xq from band q21 – q25 must be intact for normal ovarian function. Since then the region has been extended to Xq13 – q21 with exclusion of Xq22 (Therman et al 1990). Deletions within this region as well as terminal deletions involving distal Xq have been associated with varying degrees of diminished reproductive capacity (Fitch et al 1982, Veneman et al 1991, Tharapel et al 1993, Powell et al 1994). Sala et al (1997) mapped 11 balanced X/autosome translocations associated with POF to a YAC contig, spanning most of Xq21 corresponding to a 15Mb region. A region of this size was estimated to contain at least 8 different genes in Xq21 involved in ovary development, and interruption of such genes could be the cause of POF. Furthermore, in the same study a breakpoint in a female with normal fertility mapped to this region, suggesting that specific ovarian function genes are disrupted at a breakpoint, rather than a general rearrangement per se in a POF locus causing ovarian dysfunction. Further evidence for the presence of specific POF genes comes from a report by Naguib et al (1988) claiming fertility in three women with Xq25 deletions within the critical region suggested by Sarto et al (1973).

Fewer than ten families have been reported in the literature, in which a stable X deletion was associated with POF (Fitch et al 1982, Kraus et al 1987, Veneman et al 1991, Tharapel et al 1993, Powell et al 1994.) In all but three instances however, (Kraus et al 1987, Tharapel et al 1993, Powell et al 1994) the
mapping of the X chromosome deletion has been imprecise. Nevertheless, several of the reported X chromosome deletions do not overlap, leading to the conclusion that several POF genes exist. Studies of these families have resulted in the allocation of a POF1 locus in the region of Xq26-q28 (Krauss et al 1987), and a POF2 locus at Xq13.3-q21.1 (Powell et al 1984).

1.2.5 Candidate genes for ovarian development on the X chromosome

As the fine structure of the human genome is mapped at an ever increasing pace, the opportunities arise to test various genes for their role in reproductive function. Various genes have been examined for their potential role as candidate ovary determining genes on the X chromosome (figure 4).

*Drosophila fat facets related X-linked gene (DFFRX)*

Jones et al (1996) reported that an expressed sequence tag (EST 221) derived from human adult testis shares homology with the Drosophila fat facets (faf) gene. They detected related sequences on both the human X and Y chromosomes. The human X-linked homologue was termed DFFRX and the corresponding Y-specific locus designated DFFRY. The Y homologue of DFFRX, DFFRY maps to Yq11.2 and the human DFFRY mRNA is expressed in a wide range of adult and embryonic tissues, including testis. The coding regions of the DFFRY and DFFRX genes show 89% identity at the nucleotide level. Three azoospermic males have been found to harbour deletions removing the entire coding sequence of DFFRY (Brown et al 1998). The mouse Dffry gene maps to the Sxr-b deletion interval on the short arm of the mouse Y chromosome. Expression of the Dffry gene in mouse testis can first be detected
between 7.5 and 10.5 days after birth confirming an evolutionary role for DFFRY in spermatogenesis.

Figure 4 - Putative ovary determining genes on the X chromosome

DFFRX maps to Xp11.4, escapes X-inactivation and is expressed in both human adult and embryonic tissues. In Drosophila, the faf gene has been shown to be important in eye function and oocyte development. The location of DFFRX on proximal Xp coincides with the region for the major stigmata associated with Turner's syndrome, as defined by partial X chromosome deletions. This raises the possibility that DFFRX is a candidate for the defects of oogonia proliferation and subsequent gonadal degeneration observed in
Turner's syndrome. However a recent study by James et al (1998) suggests that this is not the case. They tested eleven patients with breakpoints in proximal Xq for the presence of one or two copies of the DFFRX gene, and found that two patients with normal ovarian function had a single copy of DFFRX. Therefore from these small numbers it seems that haploinsufficiency for DFFRX may not be responsible for the ovarian failure in Turner syndrome. Alternatively the possibility that in these patients there was a 46,XX cell line in the ovaries cannot be excluded.

**Zinc finger protein ZFX**

The ZFX gene is the homologue of ZFY, which encodes a zinc finger protein formerly thought to represent a testis determining factor (Page et al 1987,1990). Schneider-Gadick et al (1989) showed that ZFX escapes X inactivation in humans, and Luoh et al (1995) provided evidence of profound evolutionary conservation across species using comparative nucleotide sequencing of human and mouse ZFX genes, and suggested a fundamental developmental role for this gene. A female ZFX knockout mouse was found to have a shortage of oocytes resulting in diminished fertility and shortened reproductive lifespan mimicking POF in humans (Louh et al 1997). Male mutant mice also had fewer germ cells and both sexes were smaller and less viable. In a recent study (Avey and Conway abstract), performed mutation screening in 52 women with familial and sporadic forms of POF. Only 3 sequence changes were identified, one of which was found to be an intronic polymorphism and the other 2 were predicted not to affect translation. Hence whilst alterations of the ZFX gene may contribute to POF in some women, they are unlikely to be important in the
development of an early menopause.

*X-inactivation-specific transcript; XIST*

One theory of the aetiology of POF and the X chromosome is that of "skewed inactivation" of the X chromosome. In humans many genes, particularly housekeeping genes are expressed from the inactive X. As previously mentioned in the case of Turner's syndrome, some genes required for ovarian function may be required in double dosage and hence escape X inactivation. XIST is a gene exclusively expressed from the inactive X, is located within the X-inactivation centre at band Xq13 and is thought to be intricately involved in X inactivation (Brown et al 1991). XIST shows significant conservation of sequence and gene structure with Xist, the murine homolog which is located at the mouse X inactivation center region and is expressed from the inactive X chromosome. Penny et al (1996) have provided evidence for the absolute requirement of Xist for X inactivation by showing that Xist knockout mice fail to inactivate an X chromosome.

Panning et al (1998) suggested that there are factors that firstly stabilize XIST transcripts at the inactive X, then block the stabilization at the active X, as well as a mechanism that silences low-level XIST expression from the active X, by demonstrating variable XIST expression in embryonic stem cells.

Although it is commonly believed that the initiation of X inactivation is random, there is significant variation in the proportion of cells with either X inactive both in mice and among normal human females in the population. Families in which multiple females demonstrate extremely skewed inactivation patterns that are
otherwise quite rare in the general population are thought to reflect possible genetic influences on the X-inactivation process. Plenge et al. (1997) reported a mutation in the XIST minimal promoter in 9 females from 2 unrelated families. All females demonstrated preferential inactivation of the X chromosome carrying the mutation, suggesting that there is an association between alterations in the regulation of XIST expression and X-chromosome inactivation. Mutations in human XIST might cause skew inactivation patterns resulting in haploinsufficiency of vital ovarian developmental genes. Plenge et al (1997) screened a further 1666 independent unrelated X chromosomes revealed only one more case of this particular XIST promoter mutation ruling it out as a common polymorphism.

**Angiotensin AT2 receptor**

Angiotensin II is a potent regulator of cardiovascular haemostasis, whose action is mediated through the type 1 receptor AT1. The angiotensin II (AT2) receptor is expressed abundantly in fetal tissues and decreases rapidly after birth (Grady et al 1991, Daud et al 1988). The mouse, rat and human AT2 receptor has been mapped to the X chromosome, the latter at Xq 22. (Koike et al 1995).

An ovarian role for the AT2 receptor was suggested by studies reporting that high levels of AT2 receptors were expressed in the granulosa cells of rat atretic ovarian follicles, whereas only AT1 receptors were present in other ovarian structures (Tanaka et al 1995). Further, stimulation of AT2 receptors may contribute to the physiological process of atresia of the ovary and indeed it has been demonstrated that AT2 receptor induces apoptosis in several cell lines.
(Yamada et al 1995,1996). Having cloned the human AT2 receptor gene, Katsuya et al (1997) searched for AT2 receptor mutations as a contributory factor to the early onset of atresia in two POF families, but no changes were found in nucleotide sequences. Since only four subjects were examined for mutations, the possibility remains that AT2 receptor abnormalities cause POF in other women.

Diaphanous

Bione et al (1998) demonstrated that a balanced X;12 translocation, t(X;12)q21;p1.3) in a POF family of Sala et al (1997) had a breakpoint in the last intron of the DIA gene. This gene is a human homologue of the *drosophila* gene diaphanous (*dia*), mapped to Xq22 by Banfi et al (1997). Dia is ubiquitously expressed and conserved across species from yeast upwards. The protein encoded by the human DIA gene was the first member of the FH1/FH2 protein family, which are involved in cytokinesis and other actin-mediated morphogenetic processes that are required in early stages of development. Mutant alleles of *drosophila* dia affect spermatogenesis and oogenesis and lead to sterility, with alteration in follicular cell division in the female. In humans mutations in DIA may interfere with mechanisms leading to follicle cell proliferation.

*FSH primary response rat homologue 1 FSHPRH1*

Ovarian development is dependent on FSH stimulation as evidenced by a loss of function mutation affecting the FSH receptor. However, FSH receptor gene deletions or mutations are a very rare cause of POF (Whitney et al 1995,
Conway et al. 1999). It may be however, that genes downstream of the FSH receptor take part in ovarian development. One such FSH response gene is found on the X chromosome.

Roberts et al. (1996) hypothesized that mutations in FSH response genes might be responsible for defect in female and male gonadal development. One such gene described in rats, leucine-rich primary response gene 1 (LRPR1) is transcriptionally activated in response to FSH stimulation of testicular Sertoli cells both in vitro and in vivo (Slegtenhorst-Eegdman et al. 1995). Roberts et al. (1996) characterised a human gene (FSHPRH1), which encodes a 756 amino acid polypeptide with a 72% identity to the rat LRPR1 at the amino acid level. This gene maps to Xq22 which is adjacent to areas critical for ovarian development and is therefore is a potential candidate for human X-linked disorders of gonadal development.

**SOX3**

The mammalian genome contains a family of genes that are related to SRY (sex determining region Y), the testes determining gene. The homology is restricted to the region of SRY that encodes a DNA-binding motif of the HMG-box class, and the various genes have been named SOX, for SRY related HMG-box. The SOX3 gene has been mapped to Xq26 – q27 (Stevanovic et al. 1993) by use of a panel of somatic cell hybrids. Foster et al. (1994) identified a sequence on the marsupial X chromosome that shares homology with SRY and shows near-identity with the mouse and human SOX3 gene. They suggested that the highly conserved X chromosome-linked SOX3 represents the ancestral
SOX gene from which the sex-determining SRY gene was derived. The close homology between SRY and SOX3 might suggest that each are responsible for their respective gonadal development: SRY for the testis and SOX3 for the ovary.

Analysis of the distribution of SOX3 RNA shows that its main site of expression is in the developing nervous system and also the urogenital ridge where SOX3 protein products bind the same DNA sequence motif as SRY in vitro (Collignon et al., 1996). A deletion of this gene was detected in a male patient with a contiguous gene syndrome of haemophilia, mental retardation and primary testicular failure (Rousseau et al., 1991), and since SOX3 is expressed in human fetal brain it is possible that its deletion causes mental retardation.

Graves (1998) has recently hypothesised on the interaction between SRY, SOX3 and SOX9. SOX9 is another member of the SRY related HMG-box genes. It is located on chromosome 17 and is essential for testis development, since mutations of this gene causes XY sex reversal in association with the skeletal malformation syndrome campomelic dysplasia. SOX9 appears to be intimately involved in gonad differentiation as mouse sox9 is expressed in the gonadal ridge of both sexes, with increasing expression in the testis and the converse in the ovary. Graves proposes that in females SOX3 inhibit SOX9 function, whereas in males SRY inhibits SOX3 thus allowing SOX9 to perform its testis-determining role. Therefore SOX3 may be a key regulator of sex determination in both sexes.

_Fragile X syndrome FRAXA (see section 1.23 )_
Conclusion

From studies of X chromosome deletions it seems likely that many ovarian determining genes will be found on the X but as yet none have been sufficiently well characterised to claim this status. In this chapter I have reviewed the evidence that POF genes are concentrated on the X chromosome and have presented the case for eight obvious candidates. It seems likely that these eight genes are only a start. In order to account for all of the breakpoints on the X which are associated with ovarian failure, many more ovary specific genes must emerge.
2. PROJECT STRATEGY

In chapter 1 I have explained the general mechanisms and causes of POF, and the specific role of the X chromosome in POF. In order to search for candidate POF genes one can use a variety of methods.

a) Linkage studies- this depends on frequency of recombination. Firstly in the case of humans, as compared with lower eukaryotes, the frequency of recombination is lower and the number of progeny from each mating is less. Crossovers are separated by about $10^7$ bp, so their distance apart is much greater than the length of the average gene (10,000 bp). In addition, POF pedigrees by the nature of their infertility have low numbers of family members, thus leading to a paucity of extended family, for linkage studies.

b) Predicting candidate genes- one can look at reproductive genes which are known to have a physiological role in the ovary. The gonadotrophin receptors have already been examined but one can target other possible candidates such as inhibin or activin which also have a role in ovarian function.

c) Disease association- the link with BPES has already been alluded to, and it would follow that there is an ovary function gene close to the BPES locus. However there are exceptions to the hypothesis that a genetic disease associated with POF will have an ovarian function gene close to that genetic locus e.g. galactosaemia. Originally the effect on the ovary was thought to be a result of defective germ cell migration, but now the detrimental effects...
on the ovary are thought to be metabolic in nature, making the locus of the GALT gene somewhat of a red herring with regard to ovary determination.

d) **Breakpoint identification**- the evidence that deletions and translocations of the X chromosome cause POF has been presented in section 1.25. There are 2 hypotheses- firstly an ovary-determining gene is interrupted at the breakpoint, and/or vital genes are deleted below the breakpoint. The evidence points to the first hypothesis being true since there are breakpoints in females with normal fertility, which map to POF loci (Naguib et al 1988). Sala et al (1998) reported a fertile female with a translocation at Xq21 in the region they had mapped as a POF locus, suggesting that specific ovarian function genes are disrupted to cause POF, rather than a general rearrangement per se in a POF locus causing ovarian dysfunction.

The course of my project is presented in outline here and discussed in more detail in the remainder of this chapter.

1. **Characterisation of POF patients & prevalence of X chromosome deletions**- I conducted a study to determine the frequency of abnormal karyotypes in POF patients. Having identified an X chromosome deletion in one family, I decided to pursue the search for POF candidate genes through characterisation of the breakpoint in this family.

(Figure 5) **Familial breakpoint identification using :-**

- Cytogenetic analysis  (completed by cytogenetics department, breakpoint at Xq 26).
• Microsatellites

• EST's

• Fluorescent in situ hybridisation (FISH) using cloned genomic sequence (PACs)

• Bioinformatics

2. Characterisation of hybrids mimicking breakpoint.


STUDY 1 PREVALENCE OF CYTOGENETIC ABNORMALITIES IN POF PATIENTS

Although the association between X chromosome deletions and premature ovarian failure (POF) is well established, previous reports have not documented the prevalence of X deletions in women with POF. Therefore a study to examine the prevalence of X chromosome deletions or any other cytogenetic aberrations was carried out in women attending a dedicated POF clinic at the Middlesex Hospital which is a tertiary referral centre for patients with POF.
Figure 5 - Breakpoint identification strategy

- Cytogenetics
- Microsatellites
- PACS & FISH

Candidate genes

Xq25
Xq27

DXS994
DXS8033
DXS8050
DXS1062

500 kb
200 kb

7 Mb
Secondly, a study to identify possible clinical differences between POF patients based on autoantibody status, ovarian ultrasound findings, family history and primary or secondary amenorrhea was carried out. In particular, measurement of ovarian volume has been shown to correlate with follicular density and hence ovarian reserve, (Lass et al 1997) and recent studies suggest that ovarian function in subjects whose POF is autoimmune in origin, may be salvageable with glucocorticoid treatment (Corenblum et al 1993). Therefore the utility of three markers of autoimmunity: thyroid microsomal, ovarian and 3-β HSD autoantibodies using ovarian volume as a surrogate, were tested for the potential to retrieve ovarian function in POF women. In theory, autoimmune status would argue against a single gene defect, and hence characterisation of an autoimmune group could be helpful.

**STUDY 2 IDENTIFICATION OF FAMILIAL BREAKPOINT GENE**

**Microsatellites**

As a result of study 1, a family was identified in which POF had been transmitted through three generations in an X-linked dominant manner, in association with the karyotype 46Xdel(X)(q26). The deletion in this family may be the cause of POF through disruption of candidate ovarian genes. The breakpoint in the X chromosome of our index case with POF was reported by our cytogenetics department to be at Xq26 between polymorphic amplimerDXS994 & DXS 1062. A sequence tagged sites (STS) map of the pertinent region is available from HGMP Washington University, (Nagaraja et al 1998).
Using this map, a combination of eight polymorphic microsatellites and seven non-polymorphic (STS’s) which bridged this region were obtained (see figure 6). Polymerase chain reaction (PCR) with fluorescent labeling and an automated ABI system was used to map the deletion, and narrow down the breakpoint interval further (see table 4 in methods chapter).

Polymerase chain reaction (PCR) is a rapid and powerful technique for directly amplifying defined target DNA sequences present within the genome. Essentially the polymorphic microsatellites and STS’s are obtained as forward and reverse primers or amplimers sequences often 20-30 bases long. PCR is a chain reaction of about 30 cycles of successive steps of *denaturation*, *annealing* of amplimers and DNA synthesis or *extension*. The amplimers are designed so that in the presence of a suitably heat-stable DNA polymerase (often Taq polymerase) and DNA precursors (the four deoxynucleotide triphosphates, dATP, dCTP, dGTP and dTTP) they can initiate the synthesis of new DNA target strands which are complimentary to the individual DNA strands of the target DNA segment. The newly synthesised DNA strands will act as templates for further DNA synthesis in subsequent cycles, and hence the products of the PCR will include, in addition to the starting DNA, about $10^5$ - $10^8$ copies of the specific target DNA sequence. This can easily be visualised as a discrete band of a specific size according to the length in base pairs of the target region amplified, on agarose gel electrophoresis.

In order to compare allele sizes in the genome, one primer in each set was fluorescently labeled with FAM, so that when the PCR products were analysed using an automated laser DNA analyser, the specific alleles could be sized.
Figure 6 - STS map of breakpoint region

Mb

Xq25
  128-  DXS 994
  130-  DXS 8072
  131-  DXS 692
      - DXS 7596
      - DXS 6842
      - DXS 7314
      - DXS 7315
      - DXS 6847
      - DXS 7827
      - DXS 1114
      - DXS 1254
  131.4-  DXS 8041

Xq26
  132.4-  DXS 8033
  135-  DXS 8050

Xq27
  135.5-  DXS 1062

Polymorphic markers in red, non polymorphic markers in blue.
according to the position of fluorescent peaks. The polymorphic amplimers were di- or tri-nucleotide repeats and hence by comparing maternal, paternal and index case PCR products it was possible to ascertain whether the microsatellite lay above or below the breakpoint. For example, if there are three alleles (A, B, C) for a particular marker, the index case and mother can either have 2 copies i.e. this lies above the deletion, or one copy i.e. the microsatellite lies below the breakpoint and hence one copy is present on the normal X chromosome and the other copy has been deleted. Comparison of the index case result with the paternal allele size (single copy as only one X chromosome) differentiates between homozygosity and the deletion, as, in the latter case, the index case will share the paternal allele only (see below).

Table 3 - Method of interpreting ABI results.

<table>
<thead>
<tr>
<th>Position</th>
<th>Maternal alleles</th>
<th>Index alleles</th>
<th>Paternal alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Above breakpoint</td>
<td>A,B</td>
<td>A,C</td>
<td>C</td>
</tr>
<tr>
<td>Below breakpoint</td>
<td>B (1 peak PCR</td>
<td>C (paternal allele only)</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>product)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homozygous, above</td>
<td>B,B (1 peak PCR</td>
<td>B,C (2 peak PCR product)</td>
<td>C</td>
</tr>
<tr>
<td>breakpoint</td>
<td>product)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to narrow down the breakpoint further, several non-polymorphic STS's were utilized. However the non polymorphic nature of the STS's made it
impossible to differentiate between heterozygosity, homoygosity and the deletion, which relies upon different allele sizes.

Hence a semiquantitative PCR method was attempted. It has been shown that within the exponential phase of PCR amplification, the amount of specific DNA produced is proportional to the quantity of initial target sequence (Ferre 1992). This principle has been used for the detection of chromosome aneuploidies by quantitative fluorescent PCR (Pertl et al 1996,1997). A trisomic sample having 3 similarly sized alleles (i.e. a single PCR product) and a normal homozygote is indistinguishable, without the use of a separate non-polymorphic marker as a control (Pertl et al 1996,1997). Therefore a non-polymorphic primer set from the FSH receptor, on chromosome 2 was obtained to act as a control, (since the department had used it previously, and optimum PCR conditions were known.) The FSH receptor primer set (forward primer labeled with a fluorescent dye (HEX) different to the STS labeling) was added to the original PCR mix and the same methods as before were employed also using 4 male and female control DNA. The specific PCR products were sized and the amount of each PCR product evaluated by the extent of fluorescent activity, equal to the area of the fluorescent peak generated. A comparison of peak areas was made between the non-polymorphic autosomal primer pair (FSH receptor) results and the polymorphic X chromosome microsatellite results for control male, female, and family DNA. The PCR reaction was attenuated to 25 cycles, otherwise abnormal peaks can occur.
e.g. male  female

FSH chromosome 2

X chromosome marker

c d

Ratio areas \( \frac{a}{c} = 2 \times \frac{b}{d} \) since the male will only have a single copy of the X chromosome marker. Therefore in theory it should be possible to distinguish between male and female DNA just by comparing these ratios. It follows that index case /maternal DNA with non-polymorphic markers above the breakpoint would behave as normal female DNA, whereas below the breakpoint there would only be one copy like that of a male.

In this way PCR offers a powerful approach to distinguishing individual alleles in a genome, and thus for my purposes would enable me to narrow down the breakpoint to between two STS’s, a region of about 100kb.

**EST strategy**

The Human Genome Project aims to identify and map all genes in the genome, and is at a stage where huge numbers of Expressed sequence tags (EST) have been generated by sequencing the ends of cDNA clones. These EST’s are gradually being assigned to specific chromosomal regions mainly by PCR of panels of radiation hybrids and by fluorescent in situ hybridisation. At present about one million EST’s which represent about 40,000 genes have been
sequenced from a variety of tissue sources. It is possible to search an EST
database once a genomic region has been identified and in this way candidate
POF genes can be localised.

**Fluorescent in situ hybridisation (FISH) using cloned genomic sequence**

The Sanger Center, Cambridge, and the Saint Louis Center for Genetics in
Medicine, Washington University are currently sequencing the region of Xq25-
28. Approximately 90% of the sequence corresponding to our defined
breakpoint interval is currently available on databases for gene searches or in
the form of YACS, BACS and PACS. The breakpoint was narrowed using
microsatellites to about 1Mb, a much larger region than was originally planned.
The best method to narrow the breakpoint further is to use fluorescent in situ
hybridisation (FISH). Any purified DNA sequence can be assigned to its
chromosomal location by labeling it and hybridising it directly to the DNA of
intact chromosomes, in the form of an air-dried microscope slide preparation of
metaphase chromosomes, in which the chromosomal DNA has been denatured
by exposure to formamide. Traditionally in situ hybridisation has used $^3$H-
labeled DNA probes. Following autoradiography, positive signals are identified
by counting silver grains in the slide emulsion and using a statistical test to
discriminate genuine signal from background noise. Recently, the sensitivity
and resolution of in situ hybridisation has been significantly increased by the
development of (FISH). In this technique, the DNA probe is labelled by addition
of a reporter molecule. Following hybridisation and washing to remove excess
probe, the chromosome preparation is incubated in a solution containing a
fluorescently labeled affinity molecule that binds to the reporter on the
hybridised probe. FISH has the advantage of providing rapid results that can be conveniently scored by eye using a fluorescence microscope. In metaphase spreads of peripheral blood lymphocyte cultures positive signals show as double spots corresponding to probe hybridised to both sister chromatids.

The X chromosome breakpoint was defined as lying between two microsatellites corresponding to a genomic region of 1.2Mb. These microsatellites were then mapped onto PACS using the BLAST sequence alignment program. Ten PACS were acquired, each about 100kb from the Sanger Center, Cambridge, spanning this region, in order to perform FISH (figure 7). Essentially if a labeled PAC adhered to both the normal X chromosome and the X with the terminal deletion, this implied the genomic PAC sequence lay above the breakpoint. If however, the PAC only showed up on the normal X chromosome the genomic region was below the breakpoint. In this way it should be possible to narrow the breakpoint further to between 2 or 3 PACS.

**Computer Based Gene Searches**

Since the Human Genome Project have sequenced my region of interest, it is possible to use computer-based analysis of the sequence. Prokaryotic genes lack introns, are grouped together in operons and have well defined promoters and terminators of transcription, making them easy to predict. However in the human genome, most protein coding genes contain introns, and hence a combination of database searching and sophisticated gene prediction algorithms are required to reliably predict genes from human genomic sequence.
Figure 7 - PACS spanning region between two microsatellites

Cytogenetics

Microsatellites

PACS & FISH

7 Mb 500 kb 100 kb
Comparing a novel sequence to all sequences previously identified is obviously very important. Searching public databases using genomic sequences will identify:

1. Exact matches to previously unmapped human genes
2. High scoring matches to orthologues and close homologues
3. Weaker homologies to other genes, suggestive of coding regions rather than indicative of exactly which bases are coding
4. Near exact matches to human EST sequences.

Programs

The most commonly used sequence database searching algorithm is BLAST (best local alignment search tool, Altschul et al 1990), which has several different implementations (BLASTN, BLASTP, BLASTX, TBLASTN and TBLASTX) depending on whether a nucleotide or amino acid query sequence is used to search a nucleotide or amino acid database. A gapped BLAST search allows gaps (deletions and insertions) to be introduced into the alignments that are returned. Allowing gaps means that similar regions are not broken into several segments. The scoring of these gapped alignments tends to reflect biological relationships more closely. BLAST is a program which compares all ‘words’ of a query sequence with all words in each sequence in the target database. A word is a small subsequence (typically 12 bases for nucleic acid sequences) within the main sequence. The match between query and target words is scored and the total score compared with a user defined threshold score T. Attempts are then made to extend the word match into a longer
alignment or maximal segment pair (MSP) until extending the alignment further cannot raise the score. A query sequence may find a match in a large database by chance, and a statistical method has been developed to assess this probability (E). Hence a low E value reflects that a match is unlikely to have arisen by chance, and that match is biologically significant. The BLAST programs were tailored for sequence similarity searching - for example to identify homologues to a query sequence. The programs are not generally useful for motif-style searching, or for searching with short sequences.

There are many different programs available for predicting coding regions from genomic sequence. The methods for gene prediction fall into 2 types, reviewed by Gelfand 1995: gene search by signal where specific signals like splice junctions and promoter sequences are assessed, and gene search be content, based on the fact that protein-coding regions of DNA have several statistical features that are different from non-coding DNA (Staden 1990). Coding DNA is contained within open reading frames (ORF’s), regions containing no stop codons in at least one reading frame, and hence the codon usage of a region is used as a measure of coding potential.

GRAIL (gene recognition and analysis Internet link) is a suite of tools designed to provide analysis and putative annotation of DNA sequences both interactively and through the use of automated computation. The coding recognition portion of the system uses a neural network that combines a series of coding prediction algorithms. There are three basic versions of this neural network, GRAIL 1, GRAIL 1a and GRAIL 2. (Xu et al 1994, Lopez et al 1994, Uberbacher et al 1996).
GRAIL 1 has been in place for about three years. It uses a neural network which recognizes coding potential within a fixed size (100 base) window. It evaluates coding potential without looking for additional features (information such as splice junctions, etc).

GRAIL 1a is an updated version of GRAIL 1. It uses a fixed-length window to locate the potential coding regions and then evaluates a number of discrete candidates of different lengths around each potential coding region, using information from the two 60-base regions adjacent to that coding region, to find the "best" boundaries for that coding region. GRAIL 2 uses variable-length windows tailored to each potential exon candidate, defined as an open reading frame bounded by a pair of start/donor, acceptor/donor or acceptor/stop sites. This scheme facilitates the use of more genomic context information (splice junctions, translation starts, non-coding scores of 60-base regions on either side of a putative exon) in the exon recognition process. GRAIL 2 is therefore not appropriate for sequences without genomic context (when the regions adjacent to an exon are not present). These changes have improved the overall performance compared to GRAIL 1, particularly for short exons. All three systems have been trained to recognize coding regions in human DNA sequences, although they also work well on a number of other organisms, particularly other mammals.

GRAIL 1 typically finds about 90% of coding regions greater than 100 bases with performance falling off for shorter exons. GRAIL 1 has been tested on a set of human genes containing 102kb of sequence. This set contained 70 coding exons and the system identified 62 (89%) and assigned them all to the correct
strand. Of the eight missed 6 were less than 100 bases long. In a larger test set strand assignment was 90-95% correct. The preferred reading frame assignment was correct for 60 (95%) of these exons while the frame assignment for the other two had some ambiguity.

Of the predicted exons with a quality score of "excellent" all were actual coding exons. Of predicted exons scoring "good" 69% were real and of the predicted exons with a score of "marginal" only 16% were real. Though this is a rather limited test set, the results of this analysis give some guidance for interpreting GRAIL 1 output.

GRAIL 1a performs much better than GRAIL 1 in finding true exons and eliminating false positives. It is also better than GRAIL 1 in terms of finding the boundaries (edges) of coding regions. GRAIL 1a has been tested on a set of 137 sequences containing 954 exons. The system recognized 82% (787) of the exons in the set, with a false positive rate of 11%.

Of the 954 exons in the set, 711 exons were greater than 100 bases long. The system recognized 95 % (675) of these exons. The frame assignment was correct virtually always (greater than 98% of the time).

GRAIL 2 finds about 91% of all coding regions, with a performance that is close to being independent of exon size. Its false positive level is similar or even slightly better than GRAIL 1. GRAIL 2 has been tested on a set of 137 sequences containing 954 exons. The system recognized 91% (857) of the exons in the set, with an apparent false positive rate of 8.6% (most of these were outside the domain of the known genes and some may actually be real).
Of exons less than 100 bases long GRAIL 2 found 102 out of 122 or 84%. GRAIL 2 provides the best candidate for a given coding region in a manner which includes splice junctions (or translation start/stop) at the candidates edges, so the user will note that the edge of the candidates in the initial and summary tables correspond to putative edge signals. In the test set, about 61% of the recognized exons had both edges exactly correct (the right splice junctions picked) and about 96% had at least one edge correct. GRAIL 2 is perhaps better at estimating the true extent of an exon compared to GRAIL 1 and this additional accuracy may help in experimental protocols such as those involving PCR.

GENSCAN is a program designed to predict complete gene structures, including exons, introns, promoter and polyadenylation signals, in genomic sequences. It differs from the majority of existing gene finding algorithms in that it allows for partial genes as well as complete genes and for the occurrence of multiple genes in a single sequence, on either or both DNA strands. The program is based on a probabilistic model of gene structure / compositional properties and does not make use of protein sequence homology information. The text output of the program is a list of one or more (or possibly zero) predicted genes together with the corresponding peptide sequences. The graphical output (PostScript or gif) is a diagram of the locations of the predicted exons. Program versions suitable for vertebrate, maize and Arabidopsis sequences are currently available. Large sequences up to 500 kilobases in length may be submitted to this program.

Other gene prediction programs include HEXON, FEXH, Genemark and
Genefind. It is obvious that there are many different ways to detect interesting features of genomic sequences. Results from many different programs can be confusing – they may agree in some regions and disagree in others. Also the many different programs available can be time consuming to run and may require the input sequence in different formats.

A recent program available at the UK HGMP resource centre called NIX is a tool to view the results of running many DNA analysis programs on a DNA sequence in a single format. The analysis programs run include: GRAIL, Fex, Hexon, MZEF, Genemark, Genefinder, FGene, BLAST (against many databases) Polyah, RepeatMasker, tRNAscan. NIX is intended as a tool to aid the identification of interesting regions in Genomic or transcribed nucleic acid sequences. There are many useful computer tools to do this, none of these are 100% accurate and so it is useful to be able to compare the results of many programs that use different methods and so will fail with different things. Viewing the results of many such programs side by side makes it easy to see when many programs have a consensus about a feature. Rather than give the user total control over how the programs are run by providing them with innumerable poorly-understood choices of arguments for all of the programs, the NIX system selects reasonable defaults for the programs based on whether the sequence was genomic or transcribed, its species of origin and its size. This makes a vastly simplified interface for running a dozen or so programs on the specified sequence. However there is no way for the users to play around with different parameters to programs to adjust them to give results under more or less stringent conditions.
Databases and choice of server

There are many different nucleotide and amino acid sequence databases available to be searched by programs like BLAST. There are three main nucleotide sequence databases, EMBL, maintained at the European Bioinformatics Institute, Genbank, maintained at the US National Center for Biotechnology Information and the DNA Data Bank of Japan. These databases overlap considerably as there are procedures in place to ensure that all new and updated database entries are exchanged between the databases on a daily basis.

Many different computing sites, including NCBI in the USA (http:www.ncbi.nlm.nih.gov), and the UK HGMP resource centre (http:www.hgmp.mrc.ac.uk) offer sequence database searching facilities by a variety of means. The choice of site for any search at a given time depends on a number of factors such as accessibility (HGMP requires a user account), speed of network connection (NCBI slow in the afternoon), regularity of database updates (NCBI daily) and ease of links to other databases.

STUDY 3. CHARACTERISATION OF HYBRIDS MIMICKING BREAKPOINT

Somatic cell hybrids derive from experimentally induced fusion of cultured cells from different species. Hybrids used in human genetic mapping are made by fusion of a human cell and a rodent cell. The human chromosomes can be distinguished by their different morphology and differential staining with DNA.
binding dyes, or they can be identified by screening for human DNA sequences
or gene products which are known to map to specific chromosomes. Conventional somatic cell hybrids permit the construction of synteny maps, where panels of markers are mapped to particular chromosomes using a PCR based screen. However special hybrids which contain only part of a specific chromosome are required in order to permit assignment of DNA probes to subchromosomal segments. This occurs by one of three ways: translocation hybrids made from cells containing a translocation, deletion hybrids from cells with terminal or interstitial deletions and radiation hybrids. This method uses an existing somatic cell hybrid which contains a single chromosome, to generate a panel of deletion hybrids containing different small fragments of the chromosome by controlled X-ray irradiation cleavage of the chromosome into small pieces at random. When DNA samples from such a panel of radiation hybrids are screened by PCR against a series of DNA clones it is possible to produce a linear map order for the clones.

Radiation hybrids were obtained with fragments of the X chromosome. Their breakpoints were characterised to allow mapping of POF genes identified in the project (figure 8).
Figure 8 - Hybrids mimicking breakpoint

Xp

Xq

GM 09142

GM 11099

GM 11000

GM 10482
STUDY 4. CHARACTERISATION OF A CANDIDATE POF GENE

The fact that a gene might be interrupted by a breakpoint does not make it a POF gene. Other evidence is required:

a) it is expressed in the ovary

b) its function is plausible

c) it is mutated in another individual with POF.

The specificity for ovarian tissue can be ascertained using tissue specific Northern blots, and function in other species gives a clue to possible gene function in humans.

STUDY 5. MUTATION SCREENING OF A CANDIDATE POF GENE

If a candidate gene is found, a mutation screen using 40 POF patients will be performed. PCR will be used to amplify exons of the gene, and single stranded conformational polymorphism carried out to look for changes. Any base pair changes will be sent for sequencing to ascertain whether they are genuine mutations.
3. SUBJECTS, METHODS & REAGENTS

3.1 RAPID DNA EXTRACTION FROM WHOLE BLOOD

A rapid procedure was used to prepare DNA for PCR. 0.5ml blood was mixed with 0.5ml sucrose lysis buffer in 1.5ml eppendorf tube, which was spun at 13,000rpm 20 seconds at room temperature. The supernatant was decanted into chloros and the pellet resuspended, using a vortex, in 1ml lysis buffer. The centrifugation and resuspension was repeated twice more or until the pellet was free of haemoglobin contamination.

The sample was centrifuged at 13,000rpm for 20 secs, the supernatant removed and the pellet resuspended in 0.5ml PCR buffer using a vortex. 3ul proteinase K was added, and the resuspended pellets digested at 60 °C for 1hr. Further proteinase K was added and reincubated at 60°C for 1hr. Then the samples were incubated at 90 °C for 10mins to inactivate proteinase K, and stored at -20°C.

3.2 PREVALENCE OF CYTOGENETIC ABNORMALITIES IN POF PATIENTS

Seventy nine women were recruited from a dedicated POF clinic at the Middlesex Hospital. The criteria for the diagnosis of POF was amenorrhoea greater than 6 months duration, FSH measurement greater than 20iu on at least two occasions, and age of onset of less than 40 years. Patients with amenorrhoea due to known causes such as surgery, chemotherapy,
radiotherapy or Turner's syndrome were excluded from the study.

Thirty metaphase spreads were routinely analysed from each patient using conventional cytogenetic techniques including the giemsa/trypsin banding method (Seabright, 1971) and late replication analysis using bromodeoxyuridine and acridine orange. Fluorescent in situ hybridisation (FISH) was then performed on the spreads. Because of our experience that women with familial POF frequently carry FRAXA premutations, (Conway et al 1998), PCR analysis of the FRAXA repeat was performed in all cases.

Data was collected on autoantibody status (thyroid microsomal, adrenal, ovarian, 3-β HSD) ovarian ultrasound findings, family history and primary or secondary amenorrhoea in order to identify possible clinical differences between POF patients. Ovarian volumes were available in 43 of 79 patients and compared with antibody status.

Clinical characteristics of patients were compared with the Fisher's exact test and ANOVA for differences in proportion.

3.3 POLYMERASE CHAIN REACTION

Polymerase chain reaction (PCR) with fluorescent labeling and an automated ABI system was used to map the deletion narrow down the breakpoint interval further, using eight polymorphic microsatellites and seven non polymorphic amplimers (figure 5, table 4). Reactions were carried out in 25 ul volumes unless otherwise stated. In each case the forward primer from each set was labeled with a fluorescent dye (FAM (HEX in the case of FSH receptor
primer)Perkin Elmer ABI). The final concentration of primer pair master mix was as shown below, but magnesium concentrations were titrated from 1.0μM – 2.5Mm for each primer set.

<table>
<thead>
<tr>
<th>Master mix</th>
<th>Final conc in 25 ul PCR reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Taq polymerase buffer (Promega)</td>
<td>2.5x 1x</td>
</tr>
<tr>
<td>2.5 mM dNTPs (Promega)</td>
<td>500 μM 200μM</td>
</tr>
<tr>
<td>25 mM MgCl₂</td>
<td>3.75 μM 1.5μM</td>
</tr>
<tr>
<td>5’ primer</td>
<td>2.5 μM 1.0 μM</td>
</tr>
<tr>
<td>3’ primer</td>
<td>2.5 μM 1.0 μM</td>
</tr>
</tbody>
</table>

Ten ul of the above master mix was added to 5 ul ddH₂O. The samples were overlaid with two drops of sterile mineral oil and 5 ul of rapid DNA was added under the oil to give a final volume of 20 ul. All PCR assays underwent Hotstart at 98° C for 5 minutes on a thermal cycler (Hybaid Ltd.) to denature DNA. As the thermal cycler cooled to below 90°C, 0.25 u of Taq polymerase enzyme (Promega, UK) in 5 ul of 1x Taq polymerase buffer was added to each sample and the thermal cycler cooled to the annealing temperature. The samples then underwent 35 repeat cycles of 72° C (extension) for 30 seconds, 94° C (denaturing) for 1 min, and annealing temp for 30 sec (shown in table 4). For each primer set, annealing temperatures were titrated from 42 – 62 degrees.
Initially, 5μl of each PCR product were run on an agarose gel to assess the success of the reaction and the approximate product size. Then PCR products were separated and analysed using automated laser DNA analysers (Perkin Elmer ABI Prism 310) and the appropriate software (Genescan version 2.02 and ALF fragment manager). 1μl of each fluorescent PCR product mixed with 12μl of formamide and 0.5μl of size standard (Genescan 500-TAMRA) were run through a capillary (15kVolts, 24 min at 60°C). The specific PCR products were sized according to the position of fluorescent peaks. Different peaks represented alleles at the same locus.
Table 4 - Primer pairs and their attributes

<table>
<thead>
<tr>
<th>marker</th>
<th>Sequence of primers 5’-3’</th>
<th>Published heterozygosity</th>
<th>Conc pM</th>
<th>label</th>
<th>Anneal temp °C</th>
<th>Size of product (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS 8072 (F)</td>
<td>GTAAAAATTTACGGTTGTNCCAA</td>
<td>0.69</td>
<td>15.8</td>
<td>FAM</td>
<td>46</td>
<td>215 - 239</td>
</tr>
<tr>
<td>DXS 8072 (R)</td>
<td>TCTCCCTATCCAACCTCAGC</td>
<td>14.5</td>
<td>None</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 8071 (F)</td>
<td>CACATAACCAAGATGTGGA</td>
<td>0.64</td>
<td>16.3</td>
<td>FAM</td>
<td>48</td>
<td>185 - 201</td>
</tr>
<tr>
<td>DXS 8071 (R)</td>
<td>CATATGCCATCAAGTTTCA</td>
<td>16.3</td>
<td>None</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 692 (F)</td>
<td>CTAGAAGGCTCAAAGTCGAGC</td>
<td>0.55</td>
<td>16.6</td>
<td>FAM</td>
<td>54</td>
<td>122-132</td>
</tr>
<tr>
<td>DXS 692 (R)</td>
<td>CAGGTTATGTGATGATGCA</td>
<td>11.1</td>
<td>None</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 1254 (F)</td>
<td>GATCTTCTGGTCTTAATTTTAAAT</td>
<td>0.79</td>
<td>11.1</td>
<td>FAM</td>
<td>44</td>
<td>107 - 113</td>
</tr>
<tr>
<td>DXS 1254 (R)</td>
<td>CCGAGAGGACTGTTAAGAG</td>
<td>11.2</td>
<td>None</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 8041 (F)</td>
<td>GCAAGACTCGTGCTCAATAAACC</td>
<td>0.68</td>
<td>13.2</td>
<td>FAM</td>
<td>44</td>
<td>144 - 164</td>
</tr>
<tr>
<td>DXS 8041 (R)</td>
<td>TTCCGCTACCTCGCAATCC</td>
<td>9.1</td>
<td>None</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 8033 (F)</td>
<td>GAGACAAACCCCATGAGA</td>
<td>0.67</td>
<td>16.5</td>
<td>FAM</td>
<td>46</td>
<td>254 - 266</td>
</tr>
<tr>
<td>DXS 8033 (R)</td>
<td>ATGCCTTGAATGAGTCGAG</td>
<td>14.8</td>
<td>None</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 8050 (F)</td>
<td>CAGTTCTTGAACCTAAACC</td>
<td>0.5</td>
<td>17.5</td>
<td>FAM</td>
<td>46</td>
<td>194 - 200</td>
</tr>
<tr>
<td>DXS 8050 (R)</td>
<td>CTCCAGATTTGACATAAATACCT</td>
<td>14.1</td>
<td>None</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DXS 1062 (F)</td>
<td>GAGATGTGTTGACCTGAGC</td>
<td>0.74</td>
<td>17.3</td>
<td>FAM</td>
<td>48</td>
<td>232 - 248</td>
</tr>
<tr>
<td>DXS 1062 (R)</td>
<td>GTGGCCGTATGAAAGCATTTGAACT</td>
<td>10.7</td>
<td>None</td>
<td>48</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SEQUENCE TAGGED SITES

| DXS 7586 (F) | TGCTTTAGTGCTCTG          | 16.0                     | FAM  | 52   |               | 151                 |
| DXS 7586 (R) | GTGCTTCTGTTTTACCTG       | 6.1                      | None  | 52   |               |                     |
| DXS 6842 (F) | TTCTTTTTAGGATGCTGGATG    | 15.4                     | FAM  | 54   |               | 353                 |
| DXS 6842 (R) | ATCCAGCACAACCTAGCTCTTCC  | 6.8                      | None  | 54   |               |                     |
| DXS 7314 (F) | ATGTCTTCCAGCAGTGTTG      | 19.9                     | FAM  | 54   |               | 121                 |
| DXS 7314 (R) | GGGAAAGGTAAACTGTAAG      | 5.0                      | None  | 54   |               |                     |
| DXS 7315 (F) | CTGGTCAAGTGAGTCTGGG      | 14.8                     | FAM  | 54   |               | 60                  |
| DXS 7315 (R) | GTCTCTTTCGTTTTTTTGCC     | 11.6                     | None  | 54   |               |                     |
| DXS 6847 (F) | ACAACATACGTGAAAGCTGGG    | 7.3                      | FAM  | 52   |               | 277                 |
| DXS 6847 (R) | GGAATGTCAGTGTTCTGACAGGA  | 9.2                      | None  | 52   |               |                     |
| DXS 7827 (F) | CCCCCTCCTCTATTACCATTTTCC | 13.4                     | FAM  | 54   |               | 105                 |
| DXS 7827 (R) | CTGTTTTGCAAAACTATAACTG   | 11.0                     | None  | 54   |               |                     |
| DXS 1114 (F) | TGACTACATTTAAAAGGACAAATGC| 15.3                     | FAM  | 52   |               | 109                 |
| DXS 1114 (R) | ACTAAAAAAGAGTTGACTCTTC   | 5.7                      | None  | 52   |               |                     |
| FSH receptor (F) | AACCCCTTGCCTCTATGCGCATCT | 12.5                     | HEX   | 52 - 54 |               |                     |
| FSH receptor (R) | GGCGTAAATGACTTAAAGGGACA | 9.6                      | None  | 52 - 54 |               |                     |
3.4 AGAROSE MINIGELS

To determine the presence and size of DNA fragments after PCR, a 2% agarose gel was prepared by dissolving 1g agarose (Sigma A9539) in 50ml 1x TBE buffer through stirring on a hot plate. When the solution became clear after about ten minutes, 2.5ul ethidium bromide was added, the gel poured into a tray, two combs with 20 wells each were inserted and it was left to set. 5ul of PCR product was mixed with 2ul agarose gel loading buffer and electrophoresed in 1x TBE for 20 minutes at 135 Ma and 150V. After electrophoresis, DNA was visualised using an ultraviolet transilluminator and gels were photographed using a Polaroid camera.

3.5 FLUORESCENT IN SITU HYBRIDISATION (FISH)

FISH was carried out by modifications of the technique first described by Pinkel et al (1986). PAC DNA was labeled with biotin-14-dATP by nick translation (Bionick kit, Gibco-BRL 18247-014) Life Technologies) and the labeled probe was precipitated together with Cot-1-DNA and herring sperm DNA. Cot-1-DNA is a highly repetitive fraction of repeat DNA, which is used as a competitor DNA to suppress hybridization of repeat signals that are present in the probe, and Herring sperm DNA also reduces background and acts as carrier DNA.

*Labeling of DNA by enzymatic modification using NICK-TRANSLATION*

The mechanism of nick-translation involves the combined activities of DNase I and the 5'->3' polymerase and 5'->3' exonuclease activities of *E. coli* DNA...
polymerase I (Rigby et al., 1977). DNase I randomly introduces a single-stranded break (nick) in double-stranded DNA to create a free 3'-hydroxyl group. The 5'→3' exonuclease activity of DNA polymerase removes one or more bases at the 5' phosphoryl side of this nick. Simultaneously, DNA polymerase I catalyzes the incorporation of a nucleotide to the 3'-hydroxyl termini thus filling the gap. The nick will have been shifted along by one nucleotide in a 3' direction so that this 3' shift or translation of the nick will result in the sequential addition of new nucleotides. If radiolabeled dNTPs are incorporated in this reaction, the DNA fragment becomes labeled throughout to a uniform specific activity. The reagents and methods of labeling DNA with biotin, probe precipitation and hybridisation mix preparation are discussed in FISH reagents 3.11.

The next steps commencing with resuspending the probe in hybridisation mixture is described below. Since FISH and its interpretation involves technical experience and expertise, I performed the following experiments under the constant guidance of Dr M Fox, Dept of Human Biology, UCL. The results were advised by Dr Fox.

*Metaphase spreads*

Human metaphase chromosomes were obtained from lymphocyte cultures synchronised by addition of thymidine to block DNA synthesis. After removal of the block, 5, bromo-deoxyuridine (BrdU) was incorporated prior to harvest. Standard cytogenetic techniques were used for colcemid arrest, hypotonic treatment, fixing of the cells in methanol:acetic acid, and slide preparation.
Slide pretreatment and hybridization

The slides were pretreated with RNAse (Sigma R5125 10μg/ml) (37°C 1 hour) and proteinase K (Boehringer 109766 0.035μg/ml) (37°C 10 mins) and prefixed in 1% formaldehyde (in PBS/5% MgCl₂) at room temperature for 10 minutes. This was then dehydrated through an ascending ethanol series (70%, 90%, 100%) before denaturing in 70% formamide in 2xSSC for 5 minutes at 75°C, and dehydrating immediately afterwards in an ice-cold ethanol series. The separately denatured probe (75°C for 3.5 minutes exactly) which had been preannealed for 30 minutes at 37°C, was placed on the slide, covered with a coverslip and hybridised overnight by incubation at 37°C. Denaturation temperature and timing are crucial in maintaining good chromosome morphology. Longer denaturation will result in good hybridization but chromosomes will give dull fluorescence.

Washes and detection

Slides were washed in 50% formamide in 2xSSC at 42°C (three times for 5 minutes each), and additionally in stringency washes of 2xSSC (two times for 2.5 minutes), and in 0.1xSSC (two times for 2.5 minutes), all at 42°C. The remaining steps took place at room temperature. Slides were then blocked in 5% Marvel non-fat milk/4xSSC for 20 minutes. The fluorescent signal was detected by incubation in fluorescein isothiocyanate (FITC)-conjugated avidin (Vector laboratories A3101) (5μg/ml in Marvel/4xSSC) and amplified in
biotinylated anti-avidin (Vector laboratories B A0300) (5µg/ml in Marvel/4xSSC), followed by a second round of FITC-avidin. Incubations lasted for at least 20 minutes each and were carried out in the dark because of light sensitivity. Between these steps the slides were washed in detergent mixture (0.05% Tween 20 detergent (Sigma# P1379) in 4xSSC). Chromosomes were identified by banding made visible by fluorescent counterstains diamidino-2-phenylindole (DAPI) and propidium iodide which were added to the antifade mounting medium (Vectashield H1200 Vector Laboratories).

**Analysis and microscopy**

The slides were viewed using a fluorescent microscope. Filters with two different wavelengths are required because 2 colours (green and blue) must be made visible i.e. the DAPI counterstained chromosomes and nuclei as well as the signals and FITC.

### 3.6 PAC MINIPREP METHOD

PAC DNA samples were extracted by an alkaline lysis miniprep method, as follows (for solutions see 3.11 reagents). Five to ten milliliters of 2 LB medium contained 25µ/ml of Kanamycin were inoculated with 1-2ul of PAC library bacterial stock and cultured overnight at 37°C in s shaking incubator (120RPM). After culture the tubes were centrifuged at 3600g for 10 minutes and the supernatant discarded. The cell pellets were then tapped to loosen them and 2ml of Solution I (10x) added, the tubes mixed gently and left at room temperature for 5 minutes. Two millilitres of Solution II (freshly made) were added to the tubes, mixed very gently by inverting two or three times, and left at
room temp for 5 minutes. Two millilitres of cold Solution III were added to the tubes and mixed gently, and left for 10 minutes on ice. The tubes were centrifuged at 3600g for 10 minutes, the supernatant was then transferred to a new tube containing 4ml Isopropanol and left at -20°C for 30min. The tubes were centrifuged at 3600g for 10 minutes, the supernatant was then discarded, the tube drained, the pellet resuspended in 2ml water and mixed. Once the pellet was resuspended, 2ml of 4.4M LiCl was added and mixed by inverting the tubes which were then placed at 4°C for at least an hour. The tubes were centrifuged at 3600g for 10 minutes, the supernatant was then transferred to a new tube containing 4ml Isopropanol and left at -20°C for 1 hour or overnight. The tubes were centrifuged at 3600g for 10 minutes, the supernatant discarded and the tube allowed to drain well before being placed in an incubator at 37°C for 20 minutes to an hour, until the ethanol had evaporated completely. The dried pellet was resuspended in 100μl water.

3.7 CHARACTERISATION OF SOMATIC CELL HYBRID PANEL

Somatic cell hybrids GM11100, GM10482, 11099 and GM 09142 were obtained from the NIGMS Human Genetic Mutant Cell Repository (figure 8). They contain human X chromosomal DNA the limits of which have been tested by southern hybridisation techniques with known probes. These cell lines were cultured, DNA extracted and the remainder of the cells frozen. Using the somatic cell hybrid DNA, the same polymorphic satellites originally used to determine the breakpoint were mapped onto the hybrids using PCR.

Trypsinization and Division of cells
When cells were confluent the media was removed. 5ml warm (37 degrees) 1x Trypsin was added and then washed in 5ml phosphate buffered saline. After incubating for 5 mins, check cells are floating with microscope. 25ml media was added to dilute Trypsin and the cells divided into two flasks which had another 15ml media.

**Freezing cells**

The cells were Trypsinized, and then spun 1000rpm at 4 degrees for 5 mins. Pipette off supernatant leaving pellet. 3ml of 10% DMSO & media mixture were added and then aliquoted into three 1ml chilled cryotubes. Pack well, -20 4hrs, then -80 degrees.

**Thawing cells**

Frozen cryotubes were warmed quickly to 37 degrees and when fully defrosted, cells pipetted slowly in a fresh flask. Then 5ml warmed media was dripped in slowly. This was incubated overnight, and the cells fed the next day to remove DMSO.

**DNA isolation from somatic cell hybrids**

The cells were trypsinized and centrifuged for 10 mins at 3500rpm. Pour off supernatant, leaving behind a pellet of cells. 6ml of cell lysis solution was added and then incubated at 37°C for 15mins until solution is homogeneous. This was cooled to room temperature. 2ml of protein precipitation solution was mixed and then the sample was centrifuged at 3500rpm for 10 mins. The precipitated proteins will form a tight brown pellet. The supernatant containing
the DNA was poured into a clean tube containing 6ml isopropanol. This was mixed by gently inverting 50 times until threads of DNA were seen, then centrifuged at 3500rpm for 10 mins; the DNA will be visible as a white pellet. The supernatant was poured off and 3ml of 70% ethanol added, and centrifuged at 3500rpm for 5mins. The ethanol was poured off and allowed to air dry for 20 mins. 300-500μl DNA hydration solution was added, and the DNA allowed to rehydrate overnight at room temp.

3.8 CHARACTERISATION OF HEPARAN SULPHATE 6 SULPHOTRANSFERASE (HS6ST)

The HS6ST gene (cDNA 2051bp) mapped to the breakpoint (see results). Using the oligonucleotide design program at the HGMP, “PRIMER”, two primer sets were obtained.

HS1 Bases 379 Forward 5’ TTCCTGCACATCCAGAAGAC
583 Reverse 5’ AGTTGGTGAGCTCGGTCCA

HS2 Bases 131 Forward 5’ GCAGGACCATGGTTGAGC
308 Reverse 5’ TTCTTCTCGTAGTGGGGGTC

Firstly PCR analysis with both primer sets was carried out using control and the X chromosomal somatic cell hybrid DNA. Secondly a human monochromosomal somatic cell hybrid DNA panel containing each chromosome was obtained from the HGMP Resource Centre. PCR analysis
with the HS6ST primer sets was carried out using this panel to map the gene onto any human chromosome.

3.9 NORTHERN HYBRIDISATION

The ability of restriction enzymes to cut double stranded DNA at specific sequences was used in this project to isolate the cDNA of the E2F gene from its vector, (see results) and also to verify that this was the correct insert. The web cutter program http:\www.medkem.gu.se\cutter was used to select enzymes which cut E2F. The enzymes were obtained from several suppliers (Promega, U.K, Boehringer Mannheim, U.K, New England Biolabs, UK.) 10μl DNA, 1.5μl 10X buffer (as supplied with the enzyme), 1μl endonuclease and 2.5μl ddH2O were incubated overnight at 37°C.

Having cut the DNA fragments containing E2F DNA out of the agarose gel, they were purified using the GeneClean Spin kit (Bio 101, Anachem, UK), following the manufacturer's instructions. Samples were added to a spin filter containing GeneClean Spin Glassmilk (a supplied solution of small glass particles in a sodium iodide solution). After allowing DNA to bind to glass particles for 5 mins at room temperature, salt and proteins are eliminated by spinning the tubes for 30 secs at 13,000 rpm. After a washing step with an ethanol containing wash solution, DNA was eluted in 50μl of sterile ddH2O into a fresh tube.

Human multiple tissue Northern blots carrying poly A and mRNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leucocyte were obtained, and the recommended protocol adhered to. In summary the
blots were preybridised at $68^\circ C$ with shaking for 30 mins. The E2F DNA probe was labeled using Redeprime 2 and purified using Probe Quant G-50 microcolumns (both from Amersham Pharmacia Biotech), according to manufacturer’s instructions. The denatured probe was then added to the prehybridisation solution and hybridisation was carried out at $68^\circ C$ for 60 mins. The probe was discarded, and the blots were then washed at room temp in 2XSCC,0.05% SDS for 40 mins followed by a wash at $50^\circ C$ in 0.1XSCC,0.05% SDS for 20 mins. Blots were then exposed to film at $-70^\circ C$ for 1 week.

3.10 MUTATION SCREEN & SINGLE STRAND CONFORMATION POLYMORPHISM ANALYSIS

Screening of forty patients, thirty with idiopathic POF, and ten with familial POF, for mutations in the entire coding region of the HS6ST gene was carried out using PCR based SSCP. PCR primers were designed using the oligonucleotide design program at the HGMP, “PRIMER”. Specific primers and PCR conditions are shown in table 5. Due to the large sizes of some exons, overlapping regions were amplified.

Single strand conformation polymorphism (SSCP) was performed on PCR products using a Multiphor Electrophoresis unit with a refrigerated tank and precast gels (10% polyacrylamide, CleanGel 48s). All components and the DNA silver staining kit were purchased from Pharmacia Biotech (St Albans, UK) and procedures were carried out following the instructions of the supplier. In short, 1-2 ul of PCR product was added to 5ul of formamide SSCP loading dye, heated for 3 minutes to $98^\circ C$ to form single stranded DNA and chilled on ice.
water. Samples were loaded onto the rehydrated gels and electrophoresis was performed for 60-90 minutes at 600V, 30mA and 18W. Under these conditions the faster running double stranded DNA band was not eluted from the gel and was used to check for relative amount of DNA loaded. Standard temperature was 15°C, but if variant patterns were not detected, a second run at 50°C was undertaken. Gels were then fixed, silver stained, developed and conserved using the supplied solutions. The whole staining procedure took about 2 hours. Samples were then analysed for variant patterns by comparison to controls on a light box and gels were preserved by covering with cling film.
Table 5 Primers for mutation detection

Primers (forward and reverse) flanking each exon of the HS6ST gene. Due to the size of exons 1 and 4 overlapping primers were used. For reference, the location of the primers is shown in the appendix.

<table>
<thead>
<tr>
<th>FRAGMENT</th>
<th>SIZE (bp)</th>
<th>PRIMERS (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Forward reverse</td>
</tr>
<tr>
<td>Exon 1a</td>
<td>261</td>
<td>CGAAGTCTACCTTGCAGCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CGTAACGAGCTTGTTTC</td>
</tr>
<tr>
<td>Exon 1b</td>
<td>227</td>
<td>GGAGAACCTGGAGAACAGCC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCTACAATTCACCCGCG</td>
</tr>
<tr>
<td>Exon 1c</td>
<td>261</td>
<td>GTCAGGCTGCACGCTAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GAAGAAATGCACCTGACC</td>
</tr>
<tr>
<td>Exon 2</td>
<td>123</td>
<td>CTAGCTTACTGGGTGTAGTC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TACCAAGCTCTGTACTAGC</td>
</tr>
<tr>
<td>Exon 3</td>
<td>166</td>
<td>GTCAGATGTACAGGACAC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTGTGGCTCTGTTGTGTA</td>
</tr>
<tr>
<td>Exon 4a</td>
<td>250</td>
<td>GGCTAGATTGTAGGGACAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCATATCTCTGACTGTGTTG</td>
</tr>
<tr>
<td>Exon 4b</td>
<td>272</td>
<td>GGCCCTAGTGTTATTATTACTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCCCTCAAGAGTTATGG</td>
</tr>
<tr>
<td>Exon 4c</td>
<td>281</td>
<td>ATTCGGATTCTGGCTCTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ACATGAACCTTTATTTGCC</td>
</tr>
<tr>
<td>Exon 4d</td>
<td>286</td>
<td>CATCTTTTCAGTGCGCT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CATTTCCAGAGCCAGGGT</td>
</tr>
<tr>
<td>Exon 4e</td>
<td>309</td>
<td>GTTTTCACTTTGTCTGCCAA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AAGGCGTGACATACTTCTC</td>
</tr>
</tbody>
</table>
3.11 REAGENTS

RAPID DNA EXTRACTION FROM WHOLE BLOOD REAGENTS

Lysis buffer  Weigh out 10.94g sucrose. Add 1ml 1M Tris.HCl, 0.5 ml 1M MgCl₂, 1ml triton X100. Make up to 100ml with dd water.

PCR buffer with nonionic detergents and proteinase K  Mix 5 ml 1 M KCl, 1 ml 1M Tris.HCl (pH8.3), 0.25 ml 1 M MgCl₂, 1ml 1% gelatin, 0.45 ml NP40 and 0.45 ml Tween 20. Make up to 100ml with dd water. Autoclave and store frozen. When ready to use thaw and add 3 ul of 10mg/ml proteinase K (Promega, UK) per 100 ul of solution required.

ELECTROPHORESIS SOLUTIONS

10x TBE- Tris-borate buffer: Dissolve 121.1 g Tris base, 51.4 g boric acid and 20 ml 0.5 M EDTA (pH 8.0) in 800 ml dd H₂O on a magnetic stirrer. Adjust volume to 1000 ml with dd H₂O. Dilute 1 in 10 for working solution.

Ethidium bromide: Add 0.25 g ethidium bromide to 25 ml dd H₂O. Stir on a magnetic stirrer for several hours to ensure that the dye has dissolved. Wrap the container in aluminium foil and store at room temperature. The ethidium bromide solution is used in agarose gels at a concentration of 0.5 µl/ml, i.e. 2.5 µl stock solution in 50 ml agarose.

FISH SOLUTIONS

Labelling DNA with biotin
The Bionick labelling system (Gibco BPL Life Technologies catalogue no. 18247-014) is used, containing: **10X DNTP Mix** (0.2 mM each dCTP, dGTP, dTTP, 0.1 mM dATP, 0.1 mM biotin-14-dATP, 500 mM Tris-HCL (ph 7.8), 50 mM MgCl₂, 100 mM beta-metcaptoethanol, 100 ug/ml nuclease-free BSA), **10X Enzyme Mix** (0.5 units/ml DNA Polymerase I, 0.0075 units/ul DNAase I, 50 mM Tris-HCL (pH 7.5), 5 mM magnesium acetate, 1 mM beta-mercaptoethanol, 0.1 mM phenyimethylsulfonyl fluoride, 50% (v/v) glycerol, 100 ug/ml nuclease-free BSA) and **stop buffer** (0.3M EDTA, pH 8).

400ng DNA, 5ul l0x DNTP mix, 5ul l0x enzyme mix and a volume of supplied H₂O to make 40 ul were pipetted into a 1.5ml Eppendorf tube on ice and spun briefly. This was then incubated at 15 °C for 90 minutes, 5ul stop buffer added, and then stored at -20 C before use.

**Probe precipitation**

200ng of labeled probe DNA, 10ul mouse Cot-1-DNA (Gibco B RL Life Technol. catalogue no. 8440 SA 1mg/ml) and 2ul Herring sperm DNA (Sigma catalogue no. D7290 10mg/ml) were mixed together. A 10⁻⁷ volume Na acetate 3M and 2 volumes of ice-cold absolute ethanol were added. This was then placed in -70 C freezer for 30 minutes minimum, or -20 C overnight. The sample was spun for 5 minutes in a microfuge at 15,000 rpm, the supernatant poured off and left to dry. The hybridization mix was added and dissolved, and now the probe was ready for FISH.

*Hybridization mix preparation*

**Deionized formamide** (analar, BDH 10326) (To deionize: take 100 ml and
add 5g Amberlite mixed bed resin. Agitate gently for an hour then filter through two thicknesses of Whatman no. 1 filter paper).

20x SSPE (0. 15M NaCl, 0.01 M NaH2PO4, 0.001 M EDTA, pH 7.4) (Sigma S8140).

5ml deionized formamide, 1ml 20x SSPE, and about 3 ml H2O were mixed with 0.1g dextran sulfate (BDH). Incubating at 70 C, shaking occasionally for 3 hours dissolved the dextran sulfate. The pH was adjusted to 7.2 using concentrated HCl or 10N NaOH, and the volume adjusted to 10ml with purified H2O. This was then aliquoted into 1.5ml microcentrifuge tubes and stored at -20° C. Each aliquot can be thawed and re-frozen between use.

Washes and detection

1x SSC: 0.15 M NaCl, 0.015 M trisodium citrate.

TNB: Use a 1x TN made up from 10x TN solution (10x TN: IM Tris HCl, 1.5M NaCl, pH7.5) and add 0.5% blocking agent (Boehringer cat no. 1096176). This was heated slowly to 60 °C and incubated for 2-3 hours or until dissolved. Then centrifuge in microfuge at 15,000g for 5 minutes, use supernatant. Store -20 C.

BACTERIAL GROWTH MEDIA

Luria bertani (LB) broth: Dissolve 10 g bactotryptone, 5 g yeast extract and 10 g NaCl in 800 mg ddH2O on a magnetic stirrer. Adjust pH to 7.5 with approximately 10 drops of 10 M NaOH solution. Adjust volume to 1 l and
autoclave to sterilise.

**LB-agar for bacterial plates:** to 500 ml LB-broth add 7.5 g bactoagar (Difco, detroit USA). Autoclave, swirl to mix and allow to cool to below 50 °C. Add kanamycin, and pour 25 ml per 90 mm diameter plate. Allow to set, store light protected at 4°C.

**COSMID MINIPREP SOLUTIONS**

**10x Solution I:** Contained 0.5M Glucose, 0.25M Tris-HCl (pH 8.0) and 0.1M EDTA (pH 8.0). This 10x stock was autoclaved and could be stored and could be stored at 4°C for 6 months or more.

**Solution II:** Contained 0.2N NaOH (made freshly from a 10N stock solution) and 1% SDS. This solution needed to be freshly prepared for each experiment.

**Solution III:** was prepared by mixing 60ml of 5M Potassium Acetate, 11.5ml of Glacial Acetic acid and 28.5 ml of H₂O. This solution could be stored at 4°C for up to 6 months.1x SSC: 0.15 M NaCl, 0.015 M trisodium citrate.

**SOMATIC CELL HYBRIDS**

Cell culture growth medium

<table>
<thead>
<tr>
<th></th>
<th>GM11099</th>
<th>GM10482</th>
<th>GM09142</th>
<th>GM11100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEDIA</strong></td>
<td>Ham's F12</td>
<td>Modified Dubisco's Eagles</td>
<td>Minimum essential Eagle modification</td>
<td>Ham's F12</td>
</tr>
<tr>
<td>Fetal bovine serum</td>
<td>10% uninactivated</td>
<td>5% uninactivated</td>
<td>15% uninactivated</td>
<td>10% uninactivated</td>
</tr>
<tr>
<td>Additive</td>
<td>AAT medium</td>
<td>HAT medium</td>
<td>NONE</td>
<td>HAT medium</td>
</tr>
</tbody>
</table>
Puregene DNA isolation kit (Flowgen) containing: Cell lysis solution (25mM EDTA, 2% sodium dodecyl sulphate). Protein precipitation solution (10M ammonium acetate). DNA hydration solution (TE) (10mM Tris HCl, 1mM EDTA). Isopropanol and 70% ethanol.
4. RESULTS

4.1 PREVALENCE OF X DELETIONS AND CLINICAL CHARACTERISATION OF POF

Of the 79 women with POF, 17 women had primary amenorrhoea, and 62 presented with secondary amenorrhoea of median onset 26 years (range 11-39). 23 patients reported a positive family history of POF. Thyroid microsomal antibodies showed the highest prevalence (30%), with ovarian and adrenal antibodies at 5% and 4% respectively. There were no statistical differences between the prevalence of positive antibodies, in terms of proportion with primary amenorrhoea or positive family history. A comparison was made between ovarian volume and particular antibody status.

Comparison of ovarian volume against antibody status

<table>
<thead>
<tr>
<th>autoantibody</th>
<th>ANTIBODY -VE</th>
<th>ANTIBODY +VE</th>
<th>P value ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroid microsomal</td>
<td>2.79, (0.58)</td>
<td>3.22, (0.57)</td>
<td>0.72</td>
</tr>
<tr>
<td>3-beta HSD</td>
<td>2.21, (0.38)</td>
<td>4.79, (1.38)</td>
<td><strong>0.014</strong></td>
</tr>
<tr>
<td>ovarian</td>
<td>2.49, (0.41)</td>
<td>3.53, (1.12)</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Statistical significance was achieved in the case of 3-beta HSD antibodies suggesting that possibly positivity for 3-beta HSD antibodies predicts larger
ovarian volume in women presenting with POF. Whether this means women with 3-beta HSD antibodies are more amenable to treatment is as yet unknown.

With regards to cytogenetic analysis a normal karyotype was found in 77 of 79 women. Of the women with primary amenorrhoea, one was found to have an XY karyotype. In the secondary amenorrhoea group one woman had a deletion of the long arm of the X chromosome.

Case 1

This patient had already presented to 2 tertiary centres with primary amenorrhoea and was subsequently found to have become pregnant through ovum donation. This individual with XY gonadal dysgenesis required counselling for gonadal neoplasms, and subsequently underwent gonadectomy.

Case 2

AA, aged 26 years and her mother were aware of a family history of POF that had been extensively investigated in the 1960’s. The age range of onset of amenorrhoea in this family was 28-34 years (figure 9) and had been transmitted through three generations in an X-linked dominant manner, in association with the karyotype 46Xdel(X)(q26). Based on the estimate of dominant inheritance, advice was given of a 50% chance of POF and regular FSH monitoring was accepted. Her serum LH concentrations ranged from 1.7 - 4.3 IU/l, FSH 6.6 - 14.9 IU/l and oestradiol from 89 - 336 pmol/l. Her ovarian ultrasound was normal and she maintained a regular menstrual cycle.
The index case was found not to harbour an excess of FRAXA premutations. Cytogenetic analysis showed a 46,Xdel(X)(q26) karyotype in a total of 127 of 130 cells examined. The remaining 3 cells had a 45,X karyotype. The YAC probe yWXD2107, specific to the subtelomeric region of the long arm of the X chromosome was shown to be absent from the deleted X indicating that the deletion was terminal. Late replication analysis showed that the deleted X was late replicating in 49 of 50 cells examined in which it was present.

The mother of the index case, who had experienced POF at 28 yrs, was the only surviving affected female. She had the same 46,X,del(X)(q26) karyotype in a total of 118 of 138 cells and a 45,X karyotype in the remaining 20. The deleted X was again late replicating in all 46 out of 55 cells in which it was present. The mother's two half sisters and the father of the index case had normal cytogenetic analyses. The genetic advice was revised in the light of these findings whereby the onset of POF might occur as early as one year hence and almost certainly by age 34. After considering various options the index case chose to undergo in vitro fertilization, with embryo freezing. In fact only three embryos were produced- an indication of the already poor ovarian function. Subsequent spontaneous ovulations did not achieve conception, but her first superovulation cycle with intra-uterine insemination was successful and she had a spontaneous full term normal vaginal delivery of a girl, who on cytogenetic analysis does not share the deletion. Eighteen months later, using the same fertility treatment, she has had a baby boy.
Figure 9 - Family tree of index case

X chromosome showing deletion at Xq26.
4.2 - PCR AND ABI PRISM

The results of narrowing the breakpoint with eight polymorphic STS's is shown in figures 10 a-f and in table 6. The allele size was determined by the position of fluorescent peaks when maternal, paternal and index case PCR products were analysed using an automated DNA analyser. Once the STS lay below the breakpoint, the mother and index case have a single allele on the normal X chromosome as the other copy has been deleted.

Examples of results from the DNA laser analyser for microsatellite markers DXS8071 and DXS 8033 are shown in figures 10 a-f. Firstly 3 graphs (figure 10a-c) displaying maternal DNA showing 2 peaks at 183 &194 bp, index DNA showing 2 peaks at 183 &194 bp and paternal DNA with a single peak at 194 bp, indicating that DXS8071 is above the deletion. The next three graphs (figure 10 d-f) display maternal DNA showing a peak at 263 bp, index DNA showing a peak at 252 bp and paternal DNA with a single peak at 252 bp, i.e index case has inherited paternal X chromosome and hence DXS 8033 lies below breakpoint.

The results in table 6 indicate that the breakpoint lies between DXS 692 and DXS 1254 as two alleles are present in the index case at DXS 692, but a single allele at DXS 1254 (for explanation see chapter2 ).
### Table 6 - Results of PCR and ABI Prism

<table>
<thead>
<tr>
<th>microsatellite</th>
<th>mother</th>
<th>index</th>
<th>father</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS 8072</td>
<td>212,228</td>
<td>218,228</td>
<td>218, 220</td>
</tr>
<tr>
<td>DXS 8071</td>
<td>183, 194</td>
<td>183,194</td>
<td>194</td>
</tr>
<tr>
<td>DXS 692</td>
<td>122, 130</td>
<td>122,132</td>
<td>132</td>
</tr>
<tr>
<td>DXS 1254</td>
<td>87.4</td>
<td>90.9</td>
<td>90.8</td>
</tr>
<tr>
<td>DXS 8041</td>
<td>154</td>
<td>158</td>
<td>158</td>
</tr>
<tr>
<td>DXS 8033</td>
<td>263</td>
<td>252</td>
<td>252</td>
</tr>
<tr>
<td>DXS 8050</td>
<td>195</td>
<td>193</td>
<td>193</td>
</tr>
<tr>
<td>DXS 1062</td>
<td>235</td>
<td>240</td>
<td>240</td>
</tr>
</tbody>
</table>

The results indicate that the breakpoint lies between DXS 692 and DXS 1254 as two alleles are present in the index case at DXS 692, but a single allele at DXS 1254.
Figure 10 - ABI prism results of DXS 8071 and DXS 8033

(a) index case 2 peaks 183, 194 bp

(b) maternal 2 peaks 183, 194 bp

(c) paternal single peak 194 bp
(d) DXS 8033 – maternal single peak at 263 bp

(e) DXS 8033 – index case single peak at 252 bp

(f) DXS 8033 paternal peak at 252 bp
4.3 QUANTITATIVE PCR RESULTS

The above method depends on a polymorphism being present in order to distinguish the heterozygous state from a single allele below the breakpoint. Since all STS's between DXS692 and DXS1254 were nonpolymorphic, a quantitative PCR method was instituted. Using the same DNA analyser and software as previously, the area of the fluorescent peaks generated for the X marker and autosomal FSH marker in males compared to females should be 2. It follows that index case /maternal DNA with non-polymorphic markers above the breakpoint would behave as normal female DNA, whereas below the breakpoint there would only be one copy like that of a male.

However from table 7 it can be seen the ratios were not fixed at all. Despite titrating temperatures, magnesium concentrations, and PCR conditions, the results were not reproducible and appeared to be completely unpredictable.

The two agarose gels (figure 11) show firstly the nonpolymorphic amplimer PCR reactions with maternal, paternal and index case DNA; and secondly the same nonpolymorphic amplimer reaction with the FSH primers added in the master mix, using 2 pairs of control male and female DNA. In the second gel the FSH receptor sequence is amplified in addition to the nonpolymorphic amplimers and hence 2 bands are present in each lane.
Figure 11 - Semi-qualititative PCR results

**DXS6842,7314,7315,7827.**
Maternal:paternal:index DNA.

**DXS7596,6847,1114.**
Maternal:paternal:index:control DNA.

In this gel the amplified FSH receptor sequence is amplified in addition to the nonpolymorphic amplimers and hence 2 bands are present in each lane.
Table 7 - Example ratios of peak areas autosomal / X chromosome marker

<table>
<thead>
<tr>
<th></th>
<th>DXS 7596</th>
<th>DXS 6842</th>
<th>DXS 7314</th>
<th>DXS 1114</th>
<th>7827</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>4.23</td>
<td>2.2</td>
<td>14</td>
<td>4.4</td>
<td>0.75</td>
</tr>
<tr>
<td>Female</td>
<td>6.8</td>
<td>1.9</td>
<td>20</td>
<td>4.4</td>
<td>0.75</td>
</tr>
<tr>
<td>Female</td>
<td>5.4</td>
<td>1.74</td>
<td>64</td>
<td>1.4</td>
<td>0.57</td>
</tr>
<tr>
<td>Index case</td>
<td>16.0</td>
<td>2.4</td>
<td>13</td>
<td>3.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Index case</td>
<td>8.83</td>
<td>2.8</td>
<td>54</td>
<td>9.2</td>
<td>0.65</td>
</tr>
<tr>
<td>Mother</td>
<td>6.4</td>
<td>4.7</td>
<td>13</td>
<td>5.8</td>
<td>0.6</td>
</tr>
<tr>
<td>Mother</td>
<td>18</td>
<td>3.8</td>
<td>23</td>
<td>3.3</td>
<td>0.64</td>
</tr>
<tr>
<td>Male</td>
<td>13</td>
<td>2.6</td>
<td>13</td>
<td>2.5</td>
<td>0.84</td>
</tr>
<tr>
<td>Male</td>
<td>19</td>
<td>3.8</td>
<td>24</td>
<td>2.39</td>
<td>0.84</td>
</tr>
<tr>
<td>Male</td>
<td>27</td>
<td>2.6</td>
<td>28</td>
<td>3.5</td>
<td>0.84</td>
</tr>
<tr>
<td>Father</td>
<td>4.3</td>
<td>4.3</td>
<td>2.0</td>
<td>5.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Father</td>
<td>4.3</td>
<td>5.0</td>
<td>11</td>
<td>27.6</td>
<td>0.37</td>
</tr>
</tbody>
</table>

The ratio for males should be twice that for females but experimentally this was not the case.

4.4 INTERPRETATION OF FISH RESULTS

Ten PACs spanning the region between DXS 692 and DXS 1254 were used to narrow the breakpoint further. If a labeled PAC adhered to both the normal X chromosome and the X with the terminal deletion, this implied the genomic region lay above the breakpoint. If however, the PAC only showed up only on the normal X chromosome the genomic region was below the breakpoint (figure 12a-j, table 8).

The FISH figures around the breakpoint are figures 12c-f. Figure 12c shows the
PAC363H7 signal at Xq26 on the normal X chromosome with the rest of the X visible beyond this point, whereas on the X(del) the signal is terminal. Figure 12d showing PACdJ97K10 shows that the signal is less intense on X(del) i.e. only partially hybridizing as if spanning the breakpoint. Although strictly speaking FISH is not quantifiable, this differential PAC dJ97K10 signal as compared to that on the normal X was reproducible in further metaphase spreads. Figure 12f shows the PAC dJ358H7 signal at Xq26 on the normal X whereas on the X(del) the signal is absent. (see discussion for details on figure 12e PAC227L19.)

Table 8 - Summary of FISH results

<table>
<thead>
<tr>
<th>PAC</th>
<th>Normal X chromosome</th>
<th>X chromosome with deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>297</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>435D1</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>363L9</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>97K10</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>227L19</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>358H7</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>84F12</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>417G15</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>448E20</td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>119E23</td>
<td>Positive</td>
<td>Negative</td>
</tr>
</tbody>
</table>

The table summarises fluorescent signals from PACs on index case metaphase spread. The signal changes on (X)del from positive to absent between PAC dJ358H7 & dJ97K10 indicating the breakpoint lies in between these PACs.
Figure 12 - FISH results

Figure 12 (a) FISH results of dJ297

297 on Xq25-q26
Figure 12 (b) FISH results of dJ435D1  
435 on Xq25 - q26

Figure 12 (c) FISH results of dJ363L9  
363 no. 2 on Xq25 - q26
Figure 12 (d) FISH results of dJ97K10
Figure 12 (e) FISH results of dJ227L19
Figure 12 (f) FISH results of dJ358H7
Figure 12 (g) FISH results of dJ84F12
Figure 12 (h) FISH results of dJ417G15
Figure 12 (i) FISH results of dJ448E20
Figure 12 (j) FISH results of dJ119E23
4.5 GENOMIC SEQUENCE ANALYSIS RESULTS

Figure 13 indicates that the PCR results and FISH results agree in that the breakpoint was predicted to lie between polymorphic markers DXS692 and DXS1254, and as previously mentioned all markers in between were nonpolymorphic and the results from quantitative PCR were uninterpretable. These markers map onto PACS dJ363L9 and dJ119E23 respectively. The two breakpoint PACS dJ97K10 and dJ358H7 lie between these PACS, thus providing a corroborative computer mapping method of the breakpoint (figure 13).

'NIX' analysis was performed on contigs dJ97K10 (156000bp) and dA227L19 (3377bp), as well as neighbouring contigs dJ363L9 (125971bp) and dJ358H7 (148883bp). In each case, areas which appeared 'exon rich', i.e. exons predicted by 2 or more gene prediction programs on NIX such as GENESGAN and FGENES, were then subjected to BLAST analysis. Both BLAST2 N (nucleotide versus nucleotide database) and BLASTX (nucleotide versus protein database) were used, as this produced optimal cross species gene matches.

Nix analysis of PAC dA227L19

From figure 14 it can be seen that an exon rich region from bases 1996-2910 on the forward strand was predicted with FGENES, GENESSCAN and GENEMARK. BLAST2N analysis of embl databases using the predicted exonic sequence
Figure 13 - Microsatellite mapping onto PACS
revealed seven plausible matches. Firstly the sequence found itself i.e. PAC dA229L19 as would be expected. The next best matches of equally high scores (E value 0.0) spanning a distance of 500bp from 2000 –2900 were to five human line retrotransposable L1 elements 8,14,15,20 and 39 (see discussion).

The final match, again of high score (E value 0.0) from bases 2320-2800 was sequence AF003529. This corresponds to the partial cDNA sequence of the glypican 3 (GPC3) gene (see discussion).

The reverse strand of dA227L19 did not predict any exons that agreed across 2 programs or more.
The green line in the centre signifies base pairs and separates forward strand (above) from reverse strand (below). Exon prediction programs are down the left-hand column. High density bands from different programs signifies an exon rich region (see arrow)
NIX analysis of PAC dJ97k10

NIX analysis (figure 15 a-b, table 9) revealed 9 possible exon rich regions, of which BLAST analysis determined gene matches in four of them. BLAST analysis revealed 6 matches with bases 2693-6238 of dJ97k10. Firstly a 100% match E value 0.0 with AL049679 (1748 BP), from bases 3405-3912 with 1-508 on AL049679. Secondly a 92% match E value 5e-22 with AB006179 (2051bp), from bases 3302- 3595 with 752 to 995 on AB006179. Thirdly a 88% match, E value 1e-10 with AB006180 (710 BP) from 3302- 3595 with 130-373 on AB006180. AL049679 corresponds to the human gene on chromosome X, similar to heparan-sulphate 6-sulfotransferase i.e X chromosome isoform. AB006179 corresponds to homo sapiens mRNA for heparan sulphate 6 sulphotransferase (HS6ST) complete cDNA sequence, and AB006180 is cricetulus griseus mRNA for heparan sulphate 6 sulphotransferase (HS6ST) complete cDNA sequence.

The other 2 pairs of matches C83619, C82863, AA808759, AI225069 turned out to be EST’s from cDNA clones from cricetulus griseus and human respectively. BLAST analysis of these EST’s revealed strong matches with dJ97k10 as expected, and also the two above complete HS6ST sequences.

The PAC genomic sequence had an 82% nucleotide sequence homology and 65% aminoacid homology with HS6ST. BLAST analysis of HS6ST with Drosophila genes revealed a nucleotide match with egl (59% identity, 59% similarity) (see discussion).

In terms of mapping out the entire HS6ST gene X chromosome isoform which
yielded 100% match with the predicted exon region on PACdJ97K10, a technique of PAC walking was used. The gene has four exons and BLAST2 analysis of the human sequence AL0049679 revealed hits at PACdJ435D1 (184974BP) which is centromeric to PAC 97K10 and had a 75% nucleotide sequence match and a 89% aminoacid match with human HS6ST (table 10).

BLAST analysis of predicted exons between bases 57476-57829 and 58800-60742 revealed significant matches (Evalue 1e-10) with L1 elements 3,4,6,8,12,19 and 20 (see discussion).

Finally BLAST analysis of bases 86300-86819 on the reverse strand (predicted as an exon region) revealed a 92% match over 200bp with AF003529 (previously described), a partial sequence of the human GPC3 gene. As with the previously described method of PAC walking to map the HS6ST gene, the GPC3 gene was mapped, with the complete mRNA sequence of GPC3, 2130bp (Table 11).

Polyadenylation sites and CPG islands

The predicted exon region from bases 2693-6238 on PAC dJ97k10 was associated with a CPG island and polyadenylation tails. The predicted exon region from bases 57476-60742 was associated with polyadenylation tails.
Table 9 - Predicted exon regions on PAC dJ97K10 with BLAST analysis results

<table>
<thead>
<tr>
<th>forward strand (bp)</th>
<th>BLAST result</th>
<th>reverse strand (bp)</th>
<th>BLAST result</th>
</tr>
</thead>
<tbody>
<tr>
<td>2693-6238 fig 15a</td>
<td>AL049679/AB006179/6180 (HS6ST)</td>
<td>2545-4315</td>
<td>NIL</td>
</tr>
<tr>
<td>57476-57829 fig 15b</td>
<td>L1 elements</td>
<td>29553-32676</td>
<td>NIL</td>
</tr>
<tr>
<td>58800-60742 fig 15b</td>
<td>L1 elements</td>
<td>55800-56300</td>
<td>NIL</td>
</tr>
<tr>
<td>91028-97825</td>
<td>NIL</td>
<td>72063-74763</td>
<td>NIL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>86300-86819</td>
<td>AF003529 (GPC3)</td>
</tr>
</tbody>
</table>

Table 10 - Matches of HS6ST with PACS around breakpoint

<table>
<thead>
<tr>
<th>PAC</th>
<th>sequence on PAC</th>
<th>sequence on HS6ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>dJ97K10</td>
<td>3405-3912</td>
<td>1-508 EXON1</td>
</tr>
<tr>
<td>dJ435D1</td>
<td>113818 -113783</td>
<td>503-538 EXON2</td>
</tr>
<tr>
<td>matches range 98%-100%</td>
<td>74522 -74433</td>
<td>537-626 EXON3</td>
</tr>
<tr>
<td></td>
<td>34414-33292</td>
<td>626-1748 EXON4</td>
</tr>
<tr>
<td>119E23</td>
<td>76464-76138</td>
<td>1-327</td>
</tr>
<tr>
<td></td>
<td>44079-43915</td>
<td>324-488</td>
</tr>
<tr>
<td>417G15</td>
<td>88149-87453</td>
<td>487-1183</td>
</tr>
<tr>
<td></td>
<td>34000-33866</td>
<td>1184-1318</td>
</tr>
<tr>
<td></td>
<td>26468-26341</td>
<td>1316-1443</td>
</tr>
<tr>
<td>84F12</td>
<td>73730-73609</td>
<td>1443-1564</td>
</tr>
<tr>
<td></td>
<td>84810-83190</td>
<td>1562-1742</td>
</tr>
<tr>
<td>97K10</td>
<td>matches with AF003529 partial cds.</td>
<td>1742 onwards</td>
</tr>
<tr>
<td>435D1</td>
<td>matches with AF003529 partial cds.</td>
<td>1800 onwards</td>
</tr>
</tbody>
</table>

Table 11- Mapping of GPC3 onto PACS around breakpoint

<table>
<thead>
<tr>
<th>PAC</th>
<th>sequence on PAC</th>
<th>Sequence on GPC3</th>
</tr>
</thead>
<tbody>
<tr>
<td>119E23</td>
<td>76464-76138</td>
<td>1-327</td>
</tr>
<tr>
<td></td>
<td>44079-43915</td>
<td>324-488</td>
</tr>
<tr>
<td>417G15</td>
<td>88149-87453</td>
<td>487-1183</td>
</tr>
<tr>
<td></td>
<td>34000-33866</td>
<td>1184-1318</td>
</tr>
<tr>
<td></td>
<td>26468-26341</td>
<td>1316-1443</td>
</tr>
<tr>
<td>84F12</td>
<td>73730-73609</td>
<td>1443-1564</td>
</tr>
<tr>
<td></td>
<td>84810-83190</td>
<td>1562-1742</td>
</tr>
<tr>
<td>97K10</td>
<td>matches with AF003529 partial cds.</td>
<td>1742 onwards</td>
</tr>
<tr>
<td>435D1</td>
<td>matches with AF003529 partial cds.</td>
<td>1800 onwards</td>
</tr>
</tbody>
</table>
Figure 15 (a) - NIX analysis of PAC dJ97K10 0-7000 bp showing a predicted exon rich region from 2693 - 6238 bp
NIX analysis of PAC dj358H7

PAC dj358H7 is the PAC neighbouring the breakpoint in the direction of the telomere. NIX analysis predicted two major exon rich regions, one on the forward strand 510-2906 bp and one on the reverse strand 104241-105454 bp
(figure 16 a-b, Table 12). Both predicted exon regions were associated with polyadenylation tails.

Table 12 - Predicted exon regions on PAC dJ358H7 with BLAST analysis results

<table>
<thead>
<tr>
<th>forward strand (bp)</th>
<th>BLAST result</th>
<th>reverse strand (bp)</th>
<th>BLAST result</th>
</tr>
</thead>
<tbody>
<tr>
<td>510-2906 <strong>fig 16a</strong></td>
<td>L1 elements</td>
<td>104241-105454 <strong>fig 16b</strong></td>
<td>L23959 (E2F)</td>
</tr>
</tbody>
</table>

BLAST analysis of bases 510-2906 revealed matches of 2000bp with L1 elements 2,5,6,8,12,14,15,19,20,25,33 and 39 (see discussion).

BLAST analysis of the exon region on the reverse strand revealed a 96% nucleotide match and a 73% similarity with the protein sequence of the human single exon gene (1440 bp) human E2F related transcription factor (see discussion).

E2F mapped solely to this PAC (Table 13).

Table 13 Mapping of E2F along PAC dJ358H7

<table>
<thead>
<tr>
<th>E2F</th>
<th>E2F</th>
<th>dJ358H7</th>
<th>dJ358H7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start (bp)</td>
<td>End (bp)</td>
<td>Start (bp)</td>
<td>End (bp)</td>
</tr>
<tr>
<td>19</td>
<td>111</td>
<td>105477</td>
<td>105385</td>
</tr>
<tr>
<td>130</td>
<td>1200</td>
<td>105381</td>
<td>104314</td>
</tr>
<tr>
<td>1250</td>
<td>1322</td>
<td>104261</td>
<td>104181</td>
</tr>
<tr>
<td>1338</td>
<td>1436</td>
<td>104166</td>
<td>104072</td>
</tr>
</tbody>
</table>
Figure 16 (a) - NIX analysis of PAC dJ358H7 0-7000 bp, showing a predicted exon rich region from 510-2906 bp
Figure 16 (b) - NIX analysis of PAC dJ358H7 95 000 - 107 000 bp bp, showing a predicted exon rich region on the reverse strand from 104 241 - 105454 bp
NIX analysis of PAC dJ363L9 (Table 14)

Table 14 - Predicted exon regions on PAC dJ363L9 with BLAST analysis results

<table>
<thead>
<tr>
<th>Forward strand (bp)</th>
<th>BLAST results</th>
</tr>
</thead>
<tbody>
<tr>
<td>64460-65963 (fig 17)</td>
<td>L1 elements 5,6,8,12,14,20,25,33,39.</td>
</tr>
<tr>
<td>79851-81245 (fig 17)</td>
<td>L1 elements 5,6,8,12,14,15,19,20,21,24,25,33,39.</td>
</tr>
</tbody>
</table>

PAC dJ363L9 is the PAC neighbouring the breakpoint in the direction of the centromere. Again two major exon rich regions were predicted by NIX analysis, both on the forward strand between 64460 and 81245bp (figure 17). Both predicted exon regions were associated with polyadenylation tails.

Summary of PAC sequence analysis

Analysis of four cosmids, two spanning the breakpoint, and the two adjoining PACS, using NIX gene prediction program and BLAST searching revealed three genes, HS6ST, E2F GPG3 and LINE elements (table 15).

Table 15 - Genes predicted to lie in the breakpoint region

<table>
<thead>
<tr>
<th>PAC</th>
<th>GENE</th>
</tr>
</thead>
<tbody>
<tr>
<td>dJ363L9</td>
<td>LINE 1 ELEMENTS</td>
</tr>
<tr>
<td>dJ97K10</td>
<td>LINE 1 ELEMENTS, HS6ST, and GPC3.</td>
</tr>
<tr>
<td>dJ227L19</td>
<td>LINE 1 ELEMENTS, GPC3.</td>
</tr>
<tr>
<td>dJ358H7</td>
<td>LINE 1 ELEMENTS, E2F.</td>
</tr>
</tbody>
</table>
Figure 17 - NIX analysis of PAC dJ363L9 57 000 - 85 000 bp showing two predicted exon rich regions on the forward strand
4.6 CHARACTERISATION OF SOMATIC CELL HYBRID PANEL

The polymorphic markers were mapped onto the somatic cell hybrids. The gel photograph below (figure 18) shows the mapping of microsatellite DXS692 onto the deletion hybrid DNA, and indicates that this marker lies on hybrids GM11099 and GM10482.

GM11100 and GM10482 contain X chromosomal material which closely matches the deleted region, whilst GM11099 contains the rest of the X chromosome above the breakpoint and GM09142 contains only the short arm of X (see figure 19).

**Figure 18 - Mapping of DXS692 onto somatic cell hybrids**

![Image of gel photograph](image)

Bands are present in the control lane and GM11099 and GM10482, indicating DXS692 maps onto these 2 hybrids.
Figure 19 - Mapping of microsatellites onto somatic cell hybrids

X chromosome

- GM 11100
- DXS 1062
- DXS 8003
- DXS 8050
- DXS 692
- DXS 8041
- DXS 8071
- DXS 8072

- GM 11099
- GM 10482
- GM 109142

Breakpoint
4.7 CHARACTERISATION OF HS6ST

The gel in figure 20 shows the mapping of HS1 and HS2 onto the X chromosome deletion somatic hybrid DNA. It can be seen that HS1 appears to map to all segments of the X which is nonsensical, whereas HS2 did not amplify despite all efforts at titrating temperature and magnesium concentration.

Figure 20 - Mapping of HS1 onto somatic cell hybrids

The gel reads HS1, HS2 alternately with control, 11099, 10482, 09142 & 11100 DNA respectively.

From figure 21 it is apparent that HS1 to be present throughout all chromosomes when using the monochromosomal hybrid panel DNA, but again HS2 did not amplify except on the control DNA despite rigorous alteration in
PARC conditions.

**Figure 21 - Mapping HS1 against monochromosomal hybrid panel DNA**

4.8 NORTHERN BLOT OF E2F

Image clone 4319 was cloned into vector PSPORT1 using SAI and NotI, and hence these restriction enzymes were used to isolate the E2F cDNA from the plasmid. This gave a fragment of approximately 1000bp, which after purification from the agarose gel slices was then cut using Hind II (gtt/rac) and sty1 (c/cwwgg). The Northern blot results indicated that E2F is expressed in the ovary as well as other tissues (figure 22).

4.9 MUTATION SCREENING OF THE HS6ST GENE

All the exons of the gene were successfully amplified by PCR in all patients, suggesting that no deletions were present. The PCR products were screened in
duplicate for mutations by SSCP analysis. There was no mobility shift in any of the PCR products, suggesting the absence of polymorphisms/mutations in the exons.

**Figure 22 - Northern blot of E2F**

![Northern blot of E2F](image)
5. DISCUSSION

STUDY 1 PREVALENCE OF CYTOGENETIC ABNORMALITIES IN POF PATIENTS

Studies of patients with X chromosome deletions have resulted in the allocation of a POF1 locus in the region of Xq26-q28 (Krauss et al 1987), and a POF2 locus at Xq13.3-q21.1 (Powell et al 1984). There are patients with apparently similar breakpoints who frequently have quite different degrees of reduced reproductive capacity, ranging from primary amenorrhoea with gonadal dysgenesis to POF. It is not clear why this variation exists.

Loss of deleted X chromosomes was found in the index case and her mother and this is not unusual in patients with del(X) karyotypes. It is therefore possible that POF in the mother was attributable to her 45,X cell line. However the 45 XO karyotype was present in less than 15% of cells examined and hence concerns a limited number of cells detected in somatic tissue; it is far more likely that the deleted X is responsible for the phenotype seen. Furthermore, the deleted X has been transmitted from the parent and must have been present in the zygote. The inactivated X chromosome is late-replicating, genetically unexpressed, and has a tendency to breakage and fragility. Hence the fact that the deleted X was late-replicating indicates that its genes are usually inactive unless they escape X inactivation and two copies are required, as has been suggested in the case of ovarian function genes.

Veneman et al (1991) recognized that it may be possible to predict POF in
daughters of mothers who underwent POF, and the family described is an example of putting this advice into practice. The family are evidence that both the opportunity to freeze embryos and accelerate conception can be made available prior to the onset of amenorrhoea. It might have been expected that cytogenetic analysis was only useful in women with primary amenorrhoea, but the importance of X deletions causing POF later in life is illustrated.

Clinical characterisation of POF patients

The autoimmune status of POF patients was ascertained in order to firstly characterise the presence of an autoimmune group which might argue against a genetic defect, and secondly to find a possible correlation between ovarian volume and a particular autoimmune marker. Measurement of ovarian volume has been shown to correlate with follicular density and hence ovarian reserve, (Lass et al 1997) and recent studies suggest that ovarian function in subjects whose POF is autoimmune in origin, may be salvageable with glucocorticoid treatment (Corenblum et al 1993).

Statistical significance was achieved in the case of 3-beta HSD antibodies suggesting that possibly positivity for 3-beta HSD antibodies predicts larger ovarian volume in women presenting with POF. Whether this means 3-β HSD positivity might be a marker for preserved ovarian apparatus as determined by ovarian volume is unknown. Currently, in vitro fertilization with donated oocytes is the best therapeutic option for fertility although ovulation induction and immunosuppressive treatment might be used in selected cases. It would warrant further investigation as to whether 3-β HSD positivity is a useful guide
for targeting immunosuppressive therapy in women with POF.

STUDY 2 FAMILIAL BREAKPOINT IDENTIFICATION

PCR and microsatellites

This method using polymorphic microsatellites was originally instituted in the hope that the breakpoint might lie in between two relatively close STS's. However it turned out that DXS692 and DXS1254 were in fact around 1kb apart with non polymorphic microsatellites in between them. In principle semiquantitative PCR should have been able to distinguish the heterozygous state from a single allele (i.e. below the breakpoint) by comparing with fluorescent peak areas of male versus female controls. This method has inherent pitfalls. Firstly the autosomal amplimers have to have similar product size and PCR conditions as those on the X chromosome. I initially attempted the reaction with primers from the cystic fibrosis gene on chromosome 7 but it was impossible to coamplify both sets of primers. Secondly, different primer set PCR reactions plateau at varying numbers of cycles which can affect the reliability of peak areas. The result was that fluorescent peak areas were completely random despite titrating temperatures, magnesium concentrations, and PCR conditions. The results were not reproducible and appeared to be completely unpredictable and hence uninterpretable.

FISH

I was fortunate that within the first six months of my project (the time taken to realize that semiquantitative PCR was not behaving as I might wish) the HGMP
had sequenced Xq26 and PACs had become available for this region. In summary PAC dJ363L9 hybridised to both X chromosomes, dJ97K10 faintly to X(del) as opposed to the normal X, dA227L19 only weakly hybridised to both X chromosomes and dJ358H7 only hybridised to the normal X chromosome.

The FISH results were unequivocal for three main reasons. Firstly they all hybridised to the X chromosome at the expected position in the same order as presented by the Sanger center mapping data. Secondly, the results correlated with the STS mapping predicted by the BLAST ALIGNMENT program. Thirdly, although strictly speaking FISH is not quantifiable, the fact that in the case of PAC dJ97K10 the signal on del(X) as compared to that on the normal X was fainter, and reproducible in further metaphase spreads, implied that the PAC was only partially hybridizing as if spanning the breakpoint. A probe length of at least 3000bp is recommended for adequate hybridisation, but PAC DA227L19 was shorter and hence only a faint signal was elicited on repeated samples.

**GENOMIC SEQUENCE ANALYSIS**

Figure 23 shows a summary of the methods used to identify POF candidate genes, and figure 24 shows the map of the POF1 locus at Xq26. There follows a description and possible significance of each gene identified.

*Long interspersed nuclear elements (LINES or L1’s)*

(Review article by Kazazian et al 1998)

Approximately 3% of the human genome is comprised of exonic sequences. The remainder so called junk DNA, is composed largely of introns, simple
repeated sequences and mobile elements. L1 elements are mobile elements that occupy roughly 15% of the genome and are concentrated in AT rich regions. These elements are known as retrotransposons, which are sequences that can be transcribed into RNA, reverse transcribed into cDNA and then reintegrated as such into the genome at a new location. Although over 100,000 L1’s exist, the vast majority are incapable of retrotransposition because they contain deleterious mutations.

L1 insertions were first recognised as potential causal agents of human disease in 1988 when two separate insertions of truncated L1’s were found to disrupt the factor V111 gene, resulting in haemophilia. Since then six other retrotransposed L1 insertions were discovered: - another in the factor 8 gene, three in the Duchenne muscular dystrophy gene, one in APC and the other in the beta globin gene. Bar one, the insertions occurred in the germ line or during early development, with six of the eight L1 insertions appearing in genes located on the X chromosome. This may represent a predilection for the X, or may merely be due to the hemizygousity of X linked genes in males.
Figure 23 - Summary of candidate gene identification

- Cytogenetics
  - Xq25
  - Xq27
  - 7 Mb

- Microsatellites
  - DXS994
  - DXS8033
  - DXS8050
  - DXS1062

- PACS & FISH
  - 500 kb

- NIX & BLAST
  - 200 kb
  - GPC3
  - HS6ST
  - E2F
  - LINES
Figure 24 - Map of the POF1 locus of the X chromosome

PACs above the breakpoint are shaded in black, PACs below in white, and PAC 97K10 which spans the breakpoint is hatched. The first 5 of 8 exons of GPC3 and the single exon gene E2F are on the forward strand. LINE1 elements and all four exons of HS6ST are shown on the reverse strand.
The frequency of mutations in humans that can be ascribed to retrotransposition events is 1/670, compared to the mouse where it approaches 4/17. ALU sequences are also mobile elements in the genome, but their movement requires a cellular source of reverse transcriptase most likely encoded by L1’s. Insertions of ALU sequences have been identified as the cause of mutations in Neurofibromatosis, Haemophilia B, Fabry disease and Tay-Sachs disease.

L1’s contain two open reading frames (ORF’s). The product of ORF1 is P40, a sequence specific RNA binding protein which specifically binds to the L1 transcript near the 5 prime end of ORF2. ORF2 encodes a protein with reverse transcriptase activity, and it is thought that the P40/ORF2 protein complex activates target primed reverse transcription. From then on the exact mechanism of retrotransposition, and insertion of cDNA into the genome remains to be explained.
L1's may have a direct role in the ovary. The Drosophila melanogaster L1 factor belongs to the class of LINEs and actively transpose in the germline of the female progeny from crosses between reactive females and inducer males. This process of L1 insertion is known as I-R hybrid dysgenesis and gives rise to reduced fertility, increased frequency of mutations and X chromosome loss (Dimitri et al 1997).

L1's may be expressed in female human primordial germ cells at the time of entry into meiosis. Goto et al (1999) compared profiles of gene expression in male and female germ cells and found that a female germ cell specific transcript showing homology with the 3' end of an L1 containing sequence. This has implications for a role in female ovarian development.

In summary L1's are the 'master human mobile elements' which, either by self-insertion or mobilising ALU sequences, are responsible for disease associated mutations. It is interesting to note that many L1's were found to show significant homology with exon rich regions in both the breakpoint and both neighbouring PACS. This heavy density of L1's around the breakpoint raises the possibility that the mechanism of POF is due to L1 insertion into a candidate gene.

_E2F related transcription factor_

This single exon gene E2F(DP1) mapped entirely to PACdJ358H7 adjacent to
the breakpoint PAC dJ97K10. The E2F family of mammalian transcription factors plays a critical role in the expression of genes that are required for cell cycle progression. The first cDNA from the E2F family was designated E2F-1, and mapped to chromosome 20q11. Since then multiple E2F members have been identified. They bind to DNA as heterodimers, interacting with proteins known as DP. The DP/E2F complex is capable of activating transcription and initiating apoptosis, as shown in fibroblasts, myeloid cells and glioma cells. In fact, the massive apoptosis that occurs when E2F expression is induced in glioma cells is now being considered as a therapeutic strategy for cancer gene therapy (Gomez et al 1999).

Recently Royzman et al (1999) have shown a specific role for E2F in the drosophila ovary. E2F regulates the expression of ORC2 protein (Origin Recognition Complex) in follicle cells, such that mutations in E2F cause female sterility by affecting cell cycle transition and chorion gene amplification.

With regards to the ovary it is tempting to postulate that this apoptosis gene which lies just distal to the breakpoint, is disrupted directly or through a proximal promoter region, thus affecting its regulatory function. After all, a possible mechanism of POF is due to accelerated apoptosis of germ cells. On the other hand, E2F factors are neither ovary nor X specific, and their effects are ubiquitous. Also, E2F is a single exon gene, but we already know that disruption at various points in the POF1 locus affects ovarian function thus making this gene an unlikely candidate for POF.

_Heparan-sulphate-6-sulphotransferase (HS6ST)_
HS6ST catalyses the transfer of sulphate from 3'-phophoadenosine 5'-phosphosulphate to position 6 of the N-sulphoglucosamine residue of heparan sulphate. The complete cDNA has been cloned in humans and partially cloned in Chinese hamster ovary cells (Habuchi et al. 1998). Heparan sulphate proteoglycans (HSPG) are ubiquitously present on the cell surface and in the extracellular matrix. They are known to interact with a variety of proteins such as heparin binding growth factors, extracellular matrix components, protease inhibitors and lipoprotein lipase. HSPG's are implicated in morphological regulation during development, cell proliferation, differentiation, adhesion, migration, and also physiological phenomena such as viral infection and blood coagulation.

The interactions of HS6ST with various ligands seem to be mediated by specific regions of heparan sulphate each of a different structure. The ability of cell proteins to bind to these structural regions depends on the specific mechanism of heparan sulphate biosynthesis. In particular, sulphation plays a key role and several kinds of sulphotransferases in addition to HS6ST have been described. Heparan sulphate is required for high affinity binding of fibroblast growth factors to their receptors, as well as binding to other cell growth factors. It is postulated that different kinds of sulphotransferases affect the fine structure of heparan sulphate, in turn affecting morphological regulation and cell growth factors. With reference to the ovary, it is possible that vital follicular growth factors are inhibited through disruption of the HS6ST gene and its effect on heparan sulphate ligand binding.

A particular role for sulphation of proteoglycans has been shown in the
Drosophila oocyte (Sen et al 1998). An enzyme similar to heparan sulphate -2 - sulphotransferase has been expressed in follicle cells, and plays a pivotal role in the mechanism that establishes oocyte and furthermore embryonic polarity. A parallel role for HS6ST may occur within the human oocyte. The match of HS6ST with the drosophila gene egl is intriguing, as it appears this gene is required for oocyte microtubule polarity. Immunoprecipitation experiments show that egl protein localises to the oocyte at all stages of oogenesis from oocyte differentiation to its requirement for axis determination of the embryo (Mach et al 1997). The findings in drosophila may have a parallel in human oocyte function.

The fact that Habuchi et al (2000) have recently cloned and characterised three mouse isoforms of HS6ST present on different chromosomes corroborates the theory that the human X isoform has a particular ovarian role. The three mouse isoforms share 50-57% identical aminoacid sequences, but exhibit different sulphation activity and substrate preference. Northern analysis showed that tissue expression was characteristic to the respective isoform, and it is known that HS6ST is expressed in the ovary (unpublished observation Habuchi et al 1999). My hypothesis is that different isoforms (e.g. X chromosome) may be involved in the synthesis of heparan sulphates with tissue (e.g. ovary) specific structures and functions. Generation of isoform specific knockout mice or targeted gene mutations in Drosophila may provide evidence of the biological role of each isoform.
Glypican 3 (GPC3)

Glypican 3 is a 8 exon gene spanning 500kb, which maps to human Xq26, and rat Xq36 by fluorescent in situ hybridisation. The 2130 bp cDNA encodes a putative protein chain of 580 amino acids beginning 151bp from the start of the sequence. GPC3 is a member of the glypican family, which is a cell surface proteoglycan, anchored to the peripheral membrane through glycosylphosphatidylinositol linkage.

GPC3 is expressed ubiquitously in the embryo but is restricted to the colon and ovary in the adult. Lin et al (1999) have shown that GPC3 induces apoptosis in MCF-7 breast cancer cells in vitro, and it may also function as a tumour suppressor in the ovary. However there is no evidence that GPC3 is involved the regulation of apoptosis in the normal ovary.

Mutations of the GPC3 gene is associated with Simpson Dysmorphism syndrome (SYDS). Simpson et al (1975) observed two male cousins with a broad stocky appearance, large protruding jaw, widened nasal bridge, enlarged tongue, short hands and fingers together with normal intelligence. Since then other families have been observed with the same disorder, and the pedigree pattern is consistent with X-linked recessive inheritance. In addition more severe cases can have cleft palate, cardiac anomalies and mental retardation. Using extensive family linkage analysis Xuan et al mapped the SYDS gene to Xq26. To isolate the gene carrying the mutation the disease, Pilia et al (1996) identified breakpoints in two female patients with X/autosome translocations. The breakpoints occurred near the 5 prime and 3 prime ends of the gene.
GPC3. Since then deletions along this gene have been found in affected patients, but the clinical spectrum varies with the type of deletion. For instance, a patient with an exon 4 and 5 deletion lacked facial dysmorphism.

Glypican 3 is thought to play an important role in control of embryonic mesoderm tissue in which it is selectively expressed. With regards to ovarian effects, a case report of affected females did not mention any adverse effects (Pilia et al 1996). Similarly, a carrier mother, i.e. a heterozygote did not have ovarian dysgenesis, although whether an early menopause ensued in unknown (Golabi et Rosen 1984). Although GPC3 fills the criteria for being a large multiexon gene spanning a region known to be in the POF1 locus, in view of the severe phenotypic effects resulting from GPC3 disruption, it would be hard to imagine a scenario where GPC3 deletions would cause isolated POF.

**STUDY 3 RADIATION HYBRID MAPPING**

Radiation hybrids were characterised and found to closely map the breakpoint. However, as valuable as these resources have been to human genetics, their utility is rapidly passing. As my region of interest was sequenced within 6 months of starting the project, the use of radiation hybrids for candidate gene mapping was superceded by a computer based approach.

**STUDY 4 CHARACTERISATION OF HS6ST AND E2F**

The fact that HS6ST turned out to have highly homologous isoforms, and also the fact that there was a high degree of homology across species (>90%) meant that the results of mapping against a monochromosomal panel were unhelpful.
The first primer set amplified across many chromosomes and all species, whereas the second primer set did not amplify at all. Repeated experiments using two further primer sets again failed to yield meaningful results.

E2F1 is part of a family of factors, again highly homologous and present on at least chromosomes 20 and 22. Hence attempts at mapping would have caused the same problems as those encountered with HS6ST.

I already had reliable information regarding the presence of HS6ST in the ovary based on Northern blotting, and my own results regarding the tissue distribution of E2F confirmed its presence in the ovary as well as other tissues. In the case of both HS6ST and E2F the various individual isoforms were highly homologous. My hypothesis is that the X chromosome isoform of each gene may be more ovary specific but the cDNA's available for Northern blotting did not distinguish the subtleties of various isoforms.

STUDY 5 MUTATION SCREENING OF HS6ST GENE

This study did not detect any mutation in the entire coding region of the HS6ST gene in 40 patients with POF. There are several possible reasons for this. Firstly this may be due to type2 error, in that the number of patients examined was too small, especially as only 10 familial POF patients were examined and much larger numbers would be required in order to rule out mutations in the HS6ST gene as one cause of inherited POF. Secondly SSCP acts as a screening tool looking for a change in conformation or pattern with a sensitivity of 80% per PCR sample. Therefore there is the possibility that SSCP
missed a possible mutation, but there were no changes in the 40 samples which, even with a high false negative rate of 20% would be unlikely. DNA sequencing would have been an option but due to the exon size, would have been a laborious task without some guidance from SSCP.

Another possibility is that a polymorphism is present in the promoter region of the HS6ST gene, which was not looked at. Examples of promoter polymorphisms include the cytotoxic T lymphocyte antigen associated with autoimmune thyroid disease (Park et al 2000) and corticotrophin releasing hormone (CRH) promoter associated with rheumatoid arthritis (Baerwald et al 2000). Finally HS6ST may not be a POF gene but just happens to be interrupted in the index family and lie at the POF1 locus.
CONCLUSION

The strategy used in this project of X chromosome breakpoint analysis in order to identify ovary-determining genes is being increasingly utilized. The ability to analyze X chromosome breakpoints has been vastly accelerated by three advances. 1) A high-resolution polymorphic map of the X. 2) The fact that much of the X chromosome has been sequenced. 3) Efficient computer programs for the analysis of genomic sequence for coding exons. Once a candidate gene has been identified the next step lies in ascertaining its expression in the ovary, and proving the gene is disrupted in patients with POF. In the case of HS6ST, had the mutation screen shown this, further studies would have been carried out to ascertain function of the gene in the ovary using in vitro models, and finally the gold standard knock out mouse systems.

Increasingly, women choose to postpone reproduction until the fourth decade of life. The risk of age related infertility is presumably deemed to be sufficiently small to be acceptable to this group of women. POF is a distressing condition, which can affect young women at a time in her life before fertility has even been considered. As yet, there is no way to reverse the process or predict when a particular family member will be affected.

Detailed knowledge of POF genes would enable us to understand the mechanisms of ovarian development and aid in the diagnosis and treatment of the condition. In practice, premature ovarian failure (POF) can only be predicted if a suitable genetic marker is identified within a pedigree. However mutation screening might reveal smaller defects in a POF1 gene which would be equally
informative in women with familial or sporadic POF who are cytogenetically normal. Coping psychologically with infertility can be facilitated by a specific diagnosis, and on a practical level, the ability to provide an explanation to the mechanism of POF is therapeutic. Within a short time it should be possible to have a better understanding of POF and its prevention or treatment.
PUBLICATIONS ARISING FROM THESIS


REFERENCES


Katsuya, T., Horiuchi, M., Minami, S., Koike, G., Santoro, N.F., Hsueh,


Science 24, 675-80


152


Speed , R.M. (1988) The possible role of meiotic pairing anomalies in the


Uzielli, M.L., Guarducci, S. Lapi E., Cecconi, A., Ricci, U., Riccoti, G., Biondi,


ACKNOWLEDGEMENTS

This study would not have come to fruition without the help and encouragement of Dr Gerard Conway. I would like to thank him for the time and effort he has taken not only to encourage me to execute the work but also for providing the sole financial support during the latter part of this project.

I am grateful to Professor S Povey and Dr Margaret Fox for their academic guidance and stimulation in the design and analysis of FISH studies. I would also like to thank Dr Gill Rumsby for her scientific advice and direction, and Mrs. Nadia Payne for her friendship, patience and technical support in the lab.

I would like to thank the Wellcome Trust for funding my research, and Professor Brook for the use of the Cobbold laboratories.

Finally and most importantly I thank my parents who have been supportive and encouraging throughout my career, and have been the source of my ambition to succeed, for which I am deeply indebted and grateful.
APPENDIX

Exon distribution of HS6ST on X chromosome PACs

*primer position underlined and in italics.*
*exons shown in bold*

EXON 1 HS6ST 1-508 ON 97K10 3912 3405 +/-100 BP

```
GGGCgtcGgGgtGCGGcTAgTcCCGGGCtTGGGCCacCCGCGCtCCagCCTCCGC
3305 ANGLES---------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3364

CCCGcAgTCCaCgCcCGaCTAgGggCCACCCGCCTGGGCCACCGGGCCACGCTGC
3365 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3424

GgcGgATCCGGcCagGCGGCAGCTAGGCtCCGCCTGGGCCACCGGGCCACGCTGC
3425 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3484

CCCGACGCgGCAaCCGAcGGAcGgtGgCCTCCGcTCCGCCTGGGCCACCGGGCCACGCTGC
3545 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3604

GTGcAcCCGcGACGCTCCGCCTGGGCCACCGGGCCACGCTGC
3545 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3664

GATGttACgCcAccaAGtGCCGCCGAaAGtGgGCTCCGCCTGGGAcGAgGTCgGCAAGAA
3664 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3724

CtAcAATGcGgtGGtCTcAgCcGCcGcCtTgCAGCtCCGCCTGGGCCACCGGGCCACGCTGC
3724 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3784

gtgctAgTGccAgTgACGgGgaAcTcAaGcCTcAgTGcAgGgGccCTCCGCCTGGGCCACCGGGCCACGCTGC
3785 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3844

GAAATttGTAcCCgGgCACCACACCTGcCAGGAcGCTCtCCGCCTGGGCCACCGGGCCACGCTGC
3844 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3904

AACCCtCcTGcAgTGCcGCTTAttCTCtCCGCCTGGGCCACCGGGCCACGCTGC
3905 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------3964

TTtctCAtCCATtgCCGCtCCGCCTGGGCCACCGGGCCACGCTGC
3964 ANGLES---------------+---------------------+---------------------+---------------------+---------------------+---------------------+-------------4012

AgTcGgGgCcGcAGCcAcCGGgGgGCTGGGCcGGGapGgGAcGCTCCGCCTGGGCCACCGGGCCACGCTGC
```

EXON 2 HS6ST 503-538 ON 435D1 113818 - 113783 +/- 100BP.
exon 1a
FORWARD primer: 387 18
REVERSE primer: 647 19

<table>
<thead>
<tr>
<th>Self</th>
<th>Tm</th>
<th>GC%</th>
<th>any 3'</th>
</tr>
</thead>
</table>
? base primer 5'->3' |    |     |        |
> 387 CGAAGTCCTACCTGCA | 58.7 | 56% | 6 6 | ok |
< 647 GTCACCAGCCTTGGC | 58.4 | 53% | 5 3 | ok |

PCR product length: 261, GC = 61%, pairwise match (any 3' 3'): 4 4 2

1b
FORWARD primer: 205 20
REVERSE primer: 431 18

<table>
<thead>
<tr>
<th>Self</th>
<th>Tm</th>
<th>GC%</th>
<th>any 3'</th>
</tr>
</thead>
</table>
? base primer 5'->3' |    |     |        |
> 205 GGAGAACCTGGAGGAC | 60.3 | 60% | 3 2 | ok |
< 431 GTCAGCATTTCCCGTC | 60.2 | 56% | 4 4 | ok |

PCR product length: 227, GC = 61%, pairwise match (any 3' 3'): 5 5 1

1c
FORWARD primer: 5 17
REVERSE primer: 265 20

<table>
<thead>
<tr>
<th>Self</th>
<th>Tm</th>
<th>GC%</th>
<th>any 3'</th>
</tr>
</thead>
</table>
? base primer 5'->3' |    |     |        |
> 5 GTCAAGGCGGCAGTAG | 61.8 | 71% | 6 4 | ok |
< 265 GAAGAATGACTGCCCAC | 60.6 | 50% | 6 2 | ok |

PCR product length: 261, GC = 70%, pairwise match (any 3' 3'): 4 3 3

EXON 2
FORWARD primer: 55 20
REVERSE primer: 177 18

<table>
<thead>
<tr>
<th>Self</th>
<th>Tm</th>
<th>GC%</th>
<th>any 3'</th>
</tr>
</thead>
</table>
? base primer 5'->3' |    |     |        |
> 55 CTAAGGCTTGGGTCT | 51.4 | 45% | 4 3 | ok |
< 177 TACAAACTCGCTGCTAC | 51.7 | 50% | 8 6 | ok |

PCR product length: 123, GC = 39%, pairwise match (any 3' 3'): 5 5 2

EXON 3
FORWARD primer: 71 19
REVERSE primer: 236 19
EXON 4A

FORWARD primer: 1013 19
REVERSE primer: 1262 20

PCR product length: 250, GC = 54%, pairwise match (any 3' 3'): 4 2 4

4b
FORWARD primer: 780 21
REVERSE primer: 1051 19

PCR product length: 272, GC = 48%, pairwise match (any 3' 3'): 4 2 3

4c
FORWARD primer: 546 18
REVERSE primer: 826 20

PCR product length: 281, GC = 46%, pairwise match (any 3' 3'): 4 2 2

4d
FORWARD primer: 302 18
REVERSE primer: 587 19

PCR product length: 286, GC = 49%, pairwise match (any 3' 3'): 5 2 5

4e
FORWARD primer: 35 22
REVERSE primer: 343 20

PCR product length: 309, GC = 35%, pairwise match (any 3' 3'): 3 2 2
A familial case of X chromosome deletion ascertained by cytogenetic screening of women with premature ovarian failure


1Cobbold Laboratories, Division of Endocrinology, Department of Medicine, University College London, School of Medicine, London, 2Clinical Cytogenetics, The Galton Laboratory, University College London, Wolfson House, 4 Stephenson Way, London, NW1 2HE, 3Wessex Regional Genetics Laboratory, Salisbury Healthcare Trust, Salisbury District Hospital, Salisbury, Wiltshire, SP2 8BJ, 4Reproductive Medicine Unit, University College Hospital, 25 Grafton Way, London WC1E 6DB and 5The Assisted Conception Unit, University College Hospital, Private Patients Wing, 25 Grafton Way, London WC1E 6DB, UK

The association between X chromosome deletions and premature ovarian failure is well established. Previous anecdotal reports however, have not documented the prevalence of X deletions in women with premature ovarian failure. We therefore performed cytogenetic analyses on 79 women with primary or secondary amenorrhoea to assess the utility of screening for a genetic marker for familial premature ovarian failure. A normal karyotype was found in 77 women. One woman with primary amenorrhoea had an XY karyotype and a woman with secondary amenorrhoea had a deletion at Xq 26.1. This second case had a family history of premature ovarian failure, and her mother who underwent premature ovarian failure at 28 years shared this deletion. The early diagnosis of familial X deletions causing premature ovarian failure allowed for the prediction of impending menopause and the implementation of manoeuvres to advance conception. Although cytogenetic aberrations are rare in secondary amenorrhoea, the ability to predict premature ovarian failure can be vital. Key words: cytogenetic screening/familial premature ovarian failure/X deletions

Introduction
Increasingly, women choose to postpone reproduction until the 4th decade of life. The risk of age related infertility is presumably deemed to be sufficiently small to be acceptable to this group of women. The prediction of premature menopause based on family history has been recognized as a theoretical determinant of family planning. In practice, premature ovarian failure can only be predicted if a suitable genetic marker is identified within a pedigree. We therefore undertook cytogenetic screening of tertiary referrals attending a clinic specializing in premature ovarian failure to assess the utility of this service in defining the course of early menopause.

Here, we present the events surrounding the prediction of premature ovarian failure using cytogenetic analysis, in a woman who was able successfully to advance her planned timing for conception.

Materials and methods
Women were recruited from a dedicated premature ovarian failure clinic at the Middlesex Hospital. The study was authorized by the Ethics Committee of the hospital. The criteria for the diagnosis of premature ovarian failure was amenorrhoea >6 months duration, follicle stimulating hormone (FSH) measurement >20 IU on at least two occasions, and age of onset of <40 years. We performed cytogenetic analysis on 79 women: 17 with primary amenorrhoea, and 62 with secondary amenorrhoea of median onset 26 years (range 11–39). Thirty metaphase spreads were routinely analysed from each patient using conventional cytogenetic techniques including Giemsa/trypsin banding (Seabright, 1971) and late replication analysis using bromodeoxyuridine and acridine orange. Fluorescent in-situ hybridization (FISH) was performed using modifications of the method of Pinkel et al. (1988). Because of our experience that women with familial premature ovarian failure frequently carry FRAXA premutations (Conway et al., 1995, see Discussion), polymerase chain reaction (PCR) analysis of the FRAXA repeat was performed in all cases (Conway et al., 1998).

Results
A normal karyotype was found in 77 women. Of the women with primary amenorrhoea, one was found to have an XY karyotype. In the secondary amenorrhoea group one was found to have a terminal deletion of the long arm of the X chromosome. The clinical details of these two cases are outlined below.

Case 1
One woman with no family history of premature ovarian failure was found to have a 46 XY karyotype. She had presented to two tertiary centres previously with primary amenorrhoea and was subsequently found to have become pregnant through ovum donation. This individual with XY gonadal dysgenesis will require counselling for gonadal neoplasms.

Case 2
The index case, aged 26 years, and her mother were aware of a family history of premature ovarian failure which had been extensively investigated in the 1960s. The age range of onset
Figure 1. Family history of premature ovarian failure in case 2 (A) and karyogram of deleted Xq 26 (B).

of amenorrhoea in this family was 28–34 years (Figure 1A). Based on the estimate of dominant inheritance, advice was given of a 50% chance of premature ovarian failure and regular FSH monitoring was accepted. Her serum luteinizing hormone (LH) concentrations ranged from 1.7 to 4.3 IU/l, FSH from 6.6 to 14.9 IU/l and oestradiol from 89 to 336 pmol/l. Her ovarian ultrasound was normal and she maintained a regular menstrual cycle.

The index case was found to be hemizygous for FRAXA. Cytogenetic analysis showed a 46,Xdel(X) karyotype in a total of 127 of 130 cells examined (Figure 1b). The remaining three cells had a 45,X karyotype. The YAC probe yWXD2107, specific to the subtelomeric region of the long arm of the X chromosome, was shown to be absent from the deleted X, indicating that the deletion was terminal. Late replication analysis showed that the deleted X was late replicating in 49 of 50 cells in which it was present.

The mother of the index case, who had experienced premature ovarian failure at 28 years, was the only surviving affected female. She had the same 46,X.del(X)(q26) karyotype in a total of 118 of 138 cells and a 45,X karyotype in the remaining 20. The deleted X was again late replicating in all 46 out of 55 cells in which it was present. The genetic advice was revised in the light of these findings, whereby the onset of premature ovarian failure might occur as early as 1 year hence, and almost certainly by age 34 years. After considering various options, the index case chose to undergo in-vitro fertilization, with embryo freezing. In fact, only three embryos were produced, an indication of the already poor ovarian function. Subsequent spontaneous ovulations did not achieve conception, but her first superovulation cycle with intrauterine insemination was successful and a pregnancy is ongoing.

Discussion

This prospective cytogenetic study in women with premature ovarian failure has identified two out of 77 women in whom the karyotype was abnormal. The first, an individual with XY gonadal dysgenesis, and the second had a terminal deletion affecting the long arm of the X chromosome. In the latter cases, the study of the genetic basis for premature ovarian failure has led to the prediction of impending menopause being given to a 26 year old woman who was able to pursue both embryo storage and assisted conception prior to the predicted onset of premature menopause.

All patients in the study group underwent PCR analysis of the FRAXA repeat. Conister et al. (1991) suggested that females carrying FRAXA premutations had an increased frequency of premature ovarian failure, and Schwartz et al. (1994) reported 24 women who were typed for the FRAXA CGG repeat and 24% of women with premutations had premature ovarian failure compared with 8% of controls. Further evidence for such a non random association between FRAXA premutations and premature ovarian failure is reported in families where both occur (Conway et al., 1995; Vianna-Morgante et al., 1996). In the converse experiment, a recent study screened 132 women with idiopathic premature ovarian failure for the number of trinucleotide repeats at the FRAXA loci (Conway et al., 1998). In all, 3% of sporadic cases and 15% of kindreds with familial premature ovarian failure carried FRAXA premutations compared with an expected prevalence of 1:590. Also, there were no full mutations in the premature ovarian failure population confirming previous reports of the association between FRAXA and premature ovarian failure being restricted to premutation carriers (Cronister et al., 1991; Schwartz et al., 1994).

A large variety of deletions and X; autosome translocations of the long arm of the X chromosome are associated with gonadal dysgenesis, but patients with apparently similar breakpoints frequently have quite different degrees of reduced reproductive capacity, ranging from primary amenorrhea with gonadal dysgenesis to premature ovarian failure. It is not clear why this variation exists.

Loss of deleted X chromosomes is not unusual in patients with del(X) karyotypes and was found in index case 2 and her
mother. It is therefore possible that premature ovarian failure in the mother was attributable to her 45,X cell line. However, this concerns a limited number of cells detected in somatic tissue. It is far more likely that the deleted X, including the premature ovarian failure region, is responsible for the phenotype seen, especially as the deleted X in case 2 has been transmitted from her parent and must have been present in the zygote.

Cytogenetic analysis of Xq rearrangements has suggested that individuals with deletions or translocations localized to Xq 21.3-Xq27 (described as a ‘critical region’), experience premature ovarian failure (Fitch et al., 1982; Krauss et al., 1987; Therman et al., 1990; Veneman et al., 1991; Tharapel et al., 1993; Powell et al., 1994). Very few instances of inherited deletions of Xq causing familial premature ovarian failure have been described. Krauss et al. (1987) reported ovarian failure in three generations of a single family, found to have an interstitial deletion 46,XX,del(X)(pter→q21.3::q27→qter), and Veneman et al. (1991) reported a mother and daughter, both of whom experienced a premature menopause and shared the karyotype 46,X,del(X), (q25→qter).

In the most detailed study to date, Sala et al. (1997) mapped 11 X-autosome translocations associated with premature ovarian failure to a 15 Mb region between the polymorphic amplimers DXS223 and DXS1171. They propose that as many as eight different genes in this area may be important for ovarian function, challenging the hypothesis that one or two gene loci might be responsible for premature ovarian failure (Powell et al., 1994).

Contrary to the notion of ‘critical regions’, Naguib et al. (1988) have reported a mother and two daughters who, despite sharing a deletion at Xq25, are fertile. In fact, the literature contains several instances where X chromosome deletions within the critical region are not associated with premature ovarian failure (Therman et al., 1990).

Of notable interest, Bione et al. (1998) demonstrated that a balanced X;12 translocation, t(X;12)(q21;p1.3), in a premature ovarian failure family of Sala et al. (1997) had a breakpoint in the last intron of the DIA gene, a human homologue of the Drosophila gene diaphanous, mapping to Xq22 (Bafni et al., 1997). Bione et al. (1998) showed that the DIA gene is expressed in developing ovaries and testis of the mouse, as well as in all other tissues. Mutant alleles of Drosophila DIA affect spermatogenesis and oogenesis and lead to sterility, with alteration in follicular cell division in the female. In humans, mutations in DIA may interfere with mechanisms leading to follicle cell proliferation.

Veneman et al. (1991) recognized that it may be possible to predict premature ovarian failure in daughters of mothers who underwent a premature menopause, and in the family presented here, we have an example of putting this advice into practice. We have shown that both the opportunity to freeze embryos and to accelerate natural conception can be made available prior to the onset of amenorrhoea. It might have been expected that cytogenetic analysis was only useful in women with primary amenorrhoea, but we see here the importance of X deletions causing premature ovarian failure with onset through the third and fourth decades. However rare cytogenetic aberrations may be in secondary amenorrhoea, the vital ability to predict premature ovarian failure pays dividends.

In our experience, forewarned is forearmed!

Acknowledgements

We thank Dr David Schlessinger for the gift of YAC probe yWXD2107; also Christine Joyce and Nikki Savage for help with the FISH. Dr Davison is supported by the Wellcome Trust.

References


Received March 11, 1998; accepted August 4, 1998

X chromosome deletion and premature ovarian failure
Clinical Endocrinology (1999) 51, 673–679

Review

The X chromosome and ovarian failure

Rina M. Davison, Colin J. Davis and Gerard S. Conway
Cobbold Laboratories, Division of Endocrinology,
Department of Medicine, University College London
School of Medicine, London, UK

The X and Y chromosomes have taken very different paths in evolution. Starting as autosomes the Y chromosome shed 90% of its genetic content while the X retains over 200 genes. Half of the Y chromosome genes have homologues on the X chromosome and are thought to be housekeeping genes. Of the 20 or so genes remaining on the Y, half have a clear male determining role as befits the male chromosome (Lahn & Page, 1997). To what degree then can we consider the X chromosome to be a ‘female sex chromosome’?

The clinical models for studying ovary-determining genes are gonadal dysgenesis and premature ovarian failure (POF). Several pedigrees have been described in the literature with more than one affected member with ovarian dysgenesis or POF (Coulam et al., 1983; Mattison et al., 1984; Portuondo et al., 1987; Aittomaki, 1994; Vegetti et al., 1998). Clinically, there are dividends in being able to predict POF prior to its onset in these families. A particular example involves a family with POF in association with the karyotype 46X, Xdel(X)(q26) (Davison et al., 1998). The proband was found to share the deletion but managed to conceive before her early menopause. In this case the gross deletion was easily detectable. Thus a genetic marker for POF enables clinicians to predict whether a cytogenetically normal person was susceptible to POF. In this review we present various proposed candidate genes for ovarian failure on the X chromosome and discuss their relative merits (Fig. 1).

Evidence for X chromosome genes causing ovarian failure

Complete absence of one X chromosome, as in Turner’s syndrome, results in short stature, ovarian dysgenesis and primary amenorrhoea. In female mammals, dosage compensation of X-linked genes between males and females occurs by inactivation of one of the two X chromosomes. Both the paternally and maternally inherited X chromosomes are active in embryos prior to implantation, and X inactivation occurs around the time of implantation. The choice of X chromosome to be inactivated is random and once established the inactivity is clonally maintained. Several X-linked genes have been shown to escape X inactivation and the fact that an absent X chromosome is so deleterious to ovarian function, as compared to a largely inactivated X chromosome in normal females, suggests that two intact alleles are required for the normal function of some genes on the X chromosome. Hence candidate genes for ovarian failure in Turner’s syndrome are likely to be those which escape X inactivation (Stratakis & Rennet, 1994).

Study of naturally occurring defects of the X chromosome provides further evidence of an association between ovarian failure and deletions and translocations of the X chromosome. It appears in general that abnormalities of the long arm of the X chromosome affect only ovarian function while those of the short arm affect stature as well, resulting in a typical ‘Turner’ phenotype (Sarto et al., 1973; Thereman et al., 1990). Sarto et al. (1973) proposed the ‘critical region’ hypothesis whereby the region Xq21–q25 must be intact for normal ovarian function. Since then the region has been refined to two loci; POF 1 at Xq26–q27 (Krauss et al., 1987) and POF 2 at Xq13–q21 (Powell et al., 1994). Deletions within this region as well as terminal deletions involving distal Xq have been associated with varying degrees of diminished reproductive capacity (Fitch et al., 1982; Veneman et al., 1991; Tharapel et al., 1993; Powell et al., 1994). Yet further detail was reported by Sala et al. (1997) who mapped 11 balanced X/autosome translocations associated with POF to a Yeast Artificial chromosome (YAC) contig, spanning most of Xq21 corresponding to a 15Mb region. A region of this size was estimated to contain at least 8 different genes in Xq21 involved in ovary development, and interruption of such genes could be the cause of POF.

Candidate genes for premature ovarian failure on the X chromosome

Drosophila fat facets related X-linked gene (DFFRX)

Jones et al. (1996) reported that an expressed sequence tag derived from human adult testis shares homology with the Drosophila fat facets (faf) gene, and related sequences are on both the human X and Y chromosomes. The human X-linked homologue was termed DFFRX and the corresponding Y-specific locus designated DFFRY. DFFRY maps to Yq11.2 and is expressed in a wide range of adult and embryonic tissues, including testis. Three azoospermic males have been found to harbour deletions removing the entire coding sequence of...
DFFRY (Brown et al., 1998), confirming a role for DFFRY in spermatogenesis.

DFFRX maps to Xp11.4, escapes X-inactivation and is expressed in both human adult and embryonic tissues. In Drosophila, the $f$ gene has been shown to be important in eye function and oocyte development. The location of DFFRX on proximal Xp coincides with the region for the major stigmata associated with Turner’s syndrome, as defined by partial X chromosome deletions. This raises the possibility that DFFRX is a candidate for the defects of oogonia proliferation and subsequent gonadal degeneration observed in Turner’s syndrome. However a recent study by James et al. (1998) suggests that this is not the case. They tested 11 patients with breakpoints in proximal Xq for the presence of one or two copies of the DFFRX gene, and found that two patients with normal ovarian function had a single copy of DFFRX. Therefore from these small numbers it seems that haplo-insufficiency for DFFRX may not be responsible for the ovarian failure in Turner syndrome. The alternative possibility that these women possessed an occult 46,XX cell line within the ovaries cannot be excluded.

**Zinc finger protein ZFX**

The ZFX gene is the homologue of ZFY, which encodes a zinc finger protein formerly thought to represent a testis-determining factor (Page et al., 1987, 1990). Schneider Gadicke et al. (1989) showed that ZFX escapes X inactivation in humans, and Luoh et al. (1995) provided evidence of profound evolutionary conservation across species using comparative nucleotide sequencing of human and mouse ZFX genes, and suggested a fundamental developmental role for this gene. A female ZFX knockout mouse was found to have a reduced number of oocytes resulting in diminished fertility and shortened reproductive lifespan mimicking POF in humans (Luoh et al., 1997). Male mutant mice also had fewer germ cells and both sexes were smaller and less viable. In a recent study by Avey and Conway (unpublished observations), mutation screening in 52 women with familial and sporadic forms of POF revealed only 3 sequence changes none of which were predicted to affect translation. Hence whilst alterations of the ZFX gene may contribute to ovarian failure in some women, they are unlikely to be important in the development of an early menopause.

**X-inactivation-specific transcript; XIST**

As previously mentioned in the case of Turner’s syndrome, some genes required for ovarian function may be required in double dosage and hence escape X inactivation. It may be therefore, that an abnormality in the mechanism of X inactivation causes POF. XIST is a gene exclusively expressed from the inactive X, is located within the X-inactivation centre at band Xq13 and is thought to be intricately involved in X inactivation (Brown et al., 1991). XIST shows significant homology with Xist, the murine homologue that is located at the mouse X inactivation centre region and is also expressed from the inactive X chromosome. It has been shown that Xist knockout mice failed to inactivate an X chromosome (Penny et al., 1996). Panning & Jaenisch (1998) suggested that there are factors that firstly stabilize XIST transcripts at the inactive X, then block the stabilization at the active X, as well as a mechanism that silences low-level XIST expression from the active X, by demonstrating variable XIST expression in embryonic stem cells.

Although it is commonly believed that the initiation of X inactivation is random, there is significant variation in the proportion of cells with either X inactive both in mice and...
among normal human females in the population. Families in which multiple females demonstrate extremely skewed inactivation patterns that are otherwise quite rare in the general population are thought to reflect possible genetic influences on the X-inactivation process. Plenge et al. (1997) reported a mutation in the XIST minimal promoter in 9 females from 2 unrelated families. All females demonstrated preferential inactivation of the X chromosome carrying the mutation, suggesting that there is an association between alterations in the regulation of XIST expression and X-chromosome inactivation. Mutations in human XIST might cause skew inactivation patterns resulting in haploinsufficiency of vital ovarian developmental genes. Plenge et al. (1997) screened a further 1666 independent unrelated X chromosomes and revealed only one more case of this particular XIST promoter mutation ruling it out as a common polymorphism.

Angiotensin AT2 receptor

Angiotensin II is a potent regulator of cardiovascular haemostasis, whose action is mediated through the type 1 receptor AT1. The angiotensin II type 2 (AT2) receptor is expressed abundantly in fetal tissues and decreases rapidly after birth (Daud et al., 1988; Grady et al., 1991). The mouse, rat and human AT2 receptor has been mapped to the X chromosome, the latter at Xq22 (Koike et al., 1995).

An ovarian role for the AT2 receptor was suggested by studies reporting that high levels of angiotensin II AT2 receptors were expressed in the granulosa cells of rat atretic ovarian follicles, whereas only AT1 receptors were present in other ovarian structures (Tanaka et al., 1995). Further, stimulation of AT2 receptors may contribute to the physiological process of atresia of the ovary and indeed it has been demonstrated that AT2 receptor induces apoptosis in several cell lines (Yamada et al., 1995, 1996). Greater than 99-9% of the follicles present at birth are destined to die by apoptosis during reproductive life. Accelerated ovarian follicular apoptosis may therefore cause POF (Hsueh et al., 1994). Having cloned the human AT2 receptor gene, Katsuya et al. (1997) searched for AT2 receptor mutations as a contributory factor to the early onset of atresia in two POF families, but no changes were found in nucleotide sequences. Since only four subjects were examined for mutations, the possibility remains that AT2 receptor abnormalities cause POF in other women.

Diaphanous

Bione et al. (1998) demonstrated that a balanced X; 12 translocation, t(X; 12)(q21; p1.3) in a family suffering from premature ovarian failure (Sala et al., 1997) produced a breakpoint in the last intron of the DIA gene. This gene is a human homologue of the drosophila gene diaphanous (dia), mapped to Xq22 by Banfi et al. (1997). Dia is ubiquitously expressed and conserved across species from yeast upwards. The protein encoded by the human DIA gene was the first member of the FH1/FH2 family of proteins, which are involved in cytokinesis and other actin-mediated morphogenetic processes that are required in early stages of development. Mutant alleles of drosophila dia affect spermatogenesis and oogenesis and lead to sterility, with alteration in follicular cell division in the female. In humans mutations in DIA may interfere with mechanisms leading to follicle cell proliferation.

FSH primary response rat homologue 1 FSHPRH1

A mutation in the FSH receptor gene (FSHR), has been shown to cause hereditary hypergonadotrophic ovarian failure (Aittomaki et al., 1995). The histological appearance of the ovaries of women with this mutation showed hypoplasia with scant primordial follicles; none had the appearance of complete ovarian dysgenesis with streak ovaries suggesting that early ovarian development is not FSH dependent (Aittomaki et al., 1996). Two subsequent screening studies of women with both sporadic and familial POF have failed to identify any FSH receptor gene deletions or mutations and hence this is a very rare cause of POF (Whitney et al., 1995; Conway et al., 1998a; Layman et al., 1998). It may be however, that genes downstream of the FSH receptor take part in ovarian development. One such FSH response gene is found on the X chromosome.

Roberts et al. (1996) hypothesized that mutations in FSH response genes might be responsible for defect in female and male gonadal development. One such gene described in rats, leucine-rich primary response gene 1 (LRPR1) is transcriptionally activated in response to FSH stimulation of testicular Sertoli cells both in vitro and in vivo (Slegtenhorst-Eegdeman et al., 1995). Furthermore LRPR1 mRNA is expressed in the ovary even before FSHR mRNA, suggesting that LRPR1 may play an FSH independent role during ovarian development (Slegtenhorst-Eegdeman et al., 1998). Roberts et al. (1996) characterized a human gene (FSHPRH1), which encodes a 756 amino acid polypeptide with a 72% identity to the rat LRPR1 at the amino acid level. This gene maps to Xq22 which is adjacent to areas critical for ovarian development and is therefore a potential candidate for human X-linked disorders of gonadal function.

SOX3

The mammalian genome contains a family of genes that are related to SRY (sex determining region Y), the putative testis determining gene. The homology is restricted to the region of
SRY that encodes a DNA-binding motif of the HMG-box class, and the various genes have been named SOX, for SRY related HMG-box. The SOX3 gene has been mapped to Xq26 – q27 (Stevanovic et al., 1993) by use of a panel of somatic cell hybrids. Foster & Graves (1994) identified a sequence on the marsupial X chromosome that shares homology with SRY and shows near-identity with the mouse and human SOX3 gene. They suggested that the highly conserved X chromosome-linked SOX3 represents the ancestral SOX gene from which the sex-determining SRY gene was derived. The close homology between SRY and SOX3 might suggest that each is responsible for its respective gonadal development: SRY for the testis and SOX3 for the ovary.

Analysis of the distribution of SOX3 RNA shows that its main site of expression is in the developing nervous system and the urogenital ridge where SOX3 protein products bind the same DNA sequence motif as SRY in vitro (Collignon et al., 1996). A deletion of this gene was detected in a male patient with a contiguous gene syndrome of haemophilia, mental retardation and primary testicular failure (Rousseau et al., 1991), and since SOX3 is expressed in human fetal brain it is possible that its deletion causes mental retardation. SOX3 deletions in females however, have yet to be reported. Screening of 164 women with POF did not reveal any mutations in this gene (Conway 1999 unpublished observations).

**Fragile X syndrome FRAXA**

The fragile X syndrome is a relatively common cause of mental handicap, and is due to expansion of a polymorphic CGG trinucleotide repeat in the 5′ untranslated region of the FMR1 gene at Xq27.3. Normally the number of repeats is 5–50. Premutations have 50–200 repeats, and do not affect the synthesis of the gene product FMRP. The full Fragile X (FRAXA) mutation of greater than 200 repeats is associated with methylation of the promoter and silencing of gene transcription (Fig. 2). Absence of the gene product FMRP is associated with the clinical syndrome of mental retardation.

Studies have demonstrated that 13–25% of fragile X carriers experience POF (Allingham-Hawkins et al., 1999; Uzielli et al., 1999), and conversely 3–15% of women with idiopathic POF harbour a fragile X premutation, compared with an expected prevalence of 1:590 (Conway et al., 1998b; Uzielli et al., 1999). Also, there were no full mutations in the POF population confirming previous reports of the association between FRAXA and POF being restricted to premutation carriers (Cronister et al., 1991; Schwartz et al., 1994).

Interpretation of this finding into a molecular mechanism is perplexing. When the FMR1 gene is inactivated by methylation as in the full FRAXA mutation, there appears to be no detrimental effect on ovarian function. Protein studies have
been unable to detect a difference in the expression of the FMR1 protein from premutation alleles as compared with normal (Devys et al., 1993; Feng et al., 1995), and yet premutation alleles clearly have an adverse effect on ovarian function. One theory is that although the FMR1 protein is absent in the full mutation, a back up protein provides its function. In the case of premutations however, where FMR1 is expressed, the alternative mechanism fails to come into play and any adverse effect caused by the premutation is expressed.

Another possibility is that a particular isoform of FMR1, crucial during oogenesis is less efficiently produced from premutation alleles. Murine studies have shown that FMR1 is particularly strongly expressed during the mitotic phases of oogenesis (Bachner et al., 1993) and so any changes in expression at this critical time could dramatically reduce the number of oocytes. Against this hypothesis are the results of studies on at least 11 isoforms of FMR1 showing no differences between tissues of normal compared with premutation alleles (Ashley et al., 1993; Devys et al., 1993; Khandjian et al., 1995; Verheij et al., 1995). In conclusion, there is an association between FRAXA premutations and POF but the exact role of the FMR1 gene in reproduction is as yet undetermined.

Conclusions
From studies of X chromosome deletions it seems likely that many ovarian determining genes will be found on the X but as yet none has been sufficiently well characterized to claim this status. Here, we have reviewed the evidence that POF genes are concentrated on the X chromosome and have presented the case for eight obvious candidates. It seems likely that these eight genes are only a start. As the full complement of genes on the X chromosome emerges, the list of ovarian determining genes will define the ‘female sex chromosome’.

References


Slegtenhorst-Egedeman, K.E., Post, M., Baarends, W.M., Themmen, A.P. & Grootegoed, J.A. (1995) Regulation of gene expression in Sertoli cells by follicle-stimulating hormone (FSH): cloning and


Mapping of the *POF1* locus and identification of putative genes for premature ovarian failure

Rina M. Davison\(^1,3\), M. Fox\(^2\) and G. S. Conway\(^1\)

\(^1\)Cobbold Laboratories, Division of Endocrinology, Department of Medicine, University College London, WC1N 8AA, and \(^2\)Department of Biology, The Galton Laboratory, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK

\(^3\)To whom correspondence should be addressed

We have identified a breakpoint on the X chromosome which is associated with premature ovarian failure (POF). Using polymerase chain reaction (PCR) probes of polymorphic microsatellites and fluorescent in-situ hybridization (FISH), this breakpoint has been narrowed to a region of 300 kb spanned by two P1 artificial chromosomes (PAC). Computer exon prediction and gene homology programs revealed three genes in this area. Our results suggest that two of these genes, *HS6ST* and *E2F*, and *LINE* 1 elements may be involved in ovarian development. Interruption of these genes could be the cause of POF. This study demonstrates how various molecular techniques and bioinformatic searches can complement each other in order to solve a clinical problem.

**Key words:** E2F, *HS6ST*, long interspersed nuclear elements, POF, premature ovarian failure

Introduction

Premature ovarian failure (POF), defined as the cessation of menses for >6 months and associated with elevated gonadotrophin concentrations, before the age of 40 years (Coulam et al., 1986) affects 1% of women. The aetiology of POF is usually obscure. A genetic contribution is suggested by the occurrence of families with many affected women and an association with X chromosome defects (Therman et al., 1990). Two specific regions on Xq have been defined as POF loci: *POF1* Xq26-qter (Tharapel et al., 1993), and *POF2* Xq13.3–Xq21.1 (Powell et al., 1994). It appears that distal deletions involving the *POF1* locus results in POF at ages 24–39 years, (Krauss et al., 1987, 1990, Tharapel et al., 1993), whereas translocations involving the *POF2* locus cause POF at an earlier age of 16–21 years (Powell et al., 1994).

Complete absence of one X chromosome, as in the case of Turner’s syndrome, karyotype 45XO, results in short stature and ovarian dysgenesis. Ovarian dysfunction in Turner’s syndrome may be the result of a lack of diploid dosage of one or more vital genes, both alleles of which are active in oogenesis (Zinn et al., 1990). Haploinsufficiency of *POF* genes may explain why deletions and translocations causing disruption of genes at *POF1* and *POF2* loci affects ovarian function.

At the *POF1* locus, there is an association between POF and the fragile X premutation (50–200 trinucleotide repeats in the 5’ prime untranslated region of the *FMR1* gene at Xq27.3). Studies have demonstrated that 13–25% of fragile X carriers experience POF (Allingham-Hawkins et al., 1999, Uzielli et al., 1999), and conversely 3–15% of women with idiopathic POF harbour a fragile X premutation (Conway et al., 1998; Uzielli et al., 1999). However the exact molecular role of the *FMR1* gene in reproduction is as yet undetermined.

At the *POF2* locus, 11 breakpoints were mapped to 15 Mb region at Xq21 (Sala et al., 1997), and came to the conclusion that defects in as many as eight genes would be required to explain the association with POF over such a large area. A gene (DIA) interrupted by a breakpoint in a family with POF was identified as a possible candidate for a *POF2* gene (Bione et al., 1998). In the current study, the characterization of a breakpoint at the *POF1* locus is presented.

We recently identified a family (Davison et al., 1998) in which POF was transmitted through three generations in association with the karyotype 46X, delX(q26), between microsatellites DXS994 and DXS1062. The age range of onset of amenorrhoea in six affected females in this family was 28–34 years. The index case was 26 years at the time of presentation and her mother, who experienced POF at the age of 28 years, shared the deletion. In view of this genetic evidence, the index case was advised that menopause may occur imminently and she chose to undergo IVF with embryo freezing. We postulated that a putative *POF1* gene was disrupted at the breakpoint in this family. It might be that genes below the breakpoint are also deleted in this family and hence may be the cause of POF. However, evidence against this hypothesis (Therman et al., 1990), showed that nine out of 12 patients with a breakpoint at Xq26-28 had normal ovarian function, implying that specific genes are interrupted at breakpoints to cause POF.

Another author (Sala et al., 1997) reported a normal female with a breakpoint in the *POF2* locus suggesting that not all breakpoints in the region would produce POF, only those where POF genes are disrupted.

Materials and methods

Eight polymorphic markers were identified which lie within the breakpoint region identified by cytogenetics. Polymerase chain reac-
Slides were then blocked in 5% Marvel non-fat milk/4x SSC. The fluorescent signal was detected by incubation in fluorescein isothiocyanate (FITC)-conjugated avidin (Vector Laboratories) (5 μg/ml in Marvel/4x SSC) and amplified in biotinylated anti-avidin (Vector Laboratories, CA, USA) (5 μg/ml in Marvel/4x SSC), followed by a second round of FITC-avidin. Chromosomes were identified by banding made visible by fluorescent counterstains diaminido-2-phenylindole (DAPI) and propidium iodide which were added to the antifade mounting medium (Vectashield; Vector Laboratories).

Eight PACs were obtained from the Sanger Centre, Cambridge, UK: 435D1, 363L9, 97K10, 358H7, 84F12, 417G15, 448E12, 119E23 (sizes 88 000–184 000 bp). These clones came from the RPC11 and RPC13 library segments. DNA was extracted from PACs by an alkaline lysis miniprep method.

**Bioinformatics**

Computer prediction of exon sequences was performed using the NIX analysis program at the UK HGMP resource centre (http://www.hgmp.mrc.ac.uk). Exons predicted by three or more gene prediction programs on NIX were then submitted for BLAST (basic local alignment search tool) analysis, either at the National Center for Biotechnology Information NCBI and National Library of Medicine in the USA (http://www.ncbi.nlm.nih.gov) or the UK HGMP. Both BLAST2 N (nucleotide versus nucleotide database) and BLASTX (nucleotide versus protein database) were used, as this produced optimal cross species gene matches. Also homology between PACs and an Expressed Sequence Tagged Site (EST) database http://www.ncbi.nlm.nih.gov/dbEST/ were sought.

The BLAST program at the European Drosophila Genome Project http://edgp.ebi.ac.uk was used to search for sequence homology with *Drosophila* genes.

**Results**

The PCR results are shown in Table I. Two alleles are present above the breakpoint, whereas below this point is a single allele. Hence it is evident that the breakpoint was between microsatellite DXS692 and DXS1254, and that the female index case inherited the paternal intact X chromosome and maternal X(del).

Eight PACs span this interval, and each was used to perform FISH. The breakpoint lies at the point where a double hybridization signal from both normal and deleted X chromosomes, switches to a single signal only on the normal X and this was found to be between PACs 97K10 (156000 bp) and 358H7 (184974 bp). Representative metaphase spreads are shown in Figure 1. Although FISH is not quantifiable, the signal of PAC97K10 on the X(del) was always less intense when compared with that on the normal X, (Figure 1b). Separate batches of this PAC were used and the phenomenon was reproducible. This might imply that PAC 97K10 was bridging the breakpoint and so only part of it was able to hybridize.

The results of NIX and BLAST analysis of the two breakpoint PACs revealed four genes. Two exons immediately adjacent to the breakpoint corresponded to two genes, heparan sulphate-6-sulphotransferase (*HS6ST*) and an E2F-related transcription factor. The EST database searches for homology to our genomic sequence also revealed matches with human and *Cricetulus griseus* *HS6ST* (2051 bp) sequences. The PAC
R.M. Davison, M. Fox and G. S. Conway

Figure 1. Results of fluorescent in-situ hybridization (FISH) showing that the breakpoint lies between PI artificial chromosomes (PAC) 97 and 358. This is the point where a double hybridization signal switches to a single signal only on the normal X chromosome, (a) 363L; (b) 97K10; and (c) 358H7. del(X) = deleted X; NX = normal X.

The genomic sequence had an 82% nucleotide sequence homology and 65% aminoacid homology with HS6ST. BLAST analysis of HS6ST with Drosophila genes revealed a nucleotide match with egl (59% identity, 59% similarity). Also the PAC 435D1, which is centromeric to PAC 97K10 had a 75% nucleotide sequence match and a 89% amino acid match with human HS6ST. BLAST analysis of the second exon region revealed a 95% nucleotide match and a 73% similarity with the protein sequence of the human single exon gene (1440 bp) human E2F related transcription factor. Between exons 2 and 3 of HS6ST, BLAST analysis of another predicted exon region corresponds to long interspersed nuclear elements (LINE 1) elements with a 97% match over 2000 bp with LINE 1 elements 2,5,6,8,12,14,15,19,20,25,33 and 39 (which are homologous) – a LINE 1 element cluster. Finally BLAST analysis of an exon on PAC 97K10 centromeric to the HS6ST exon revealed a 92% match over 200 bp with a partial sequence of the human GPC3 gene.

Discussion

We have characterized the organization of genes at the breakpoint in this family with POF, and ascertained that there are three genes: HS6ST, E2F, GPC3 that may be interrupted at the breakpoint (Figure 2). In addition LINE 1 elements are inserted within this region.

Heparan-sulphate-6-sulphotransferase (HS6ST)

HS6ST catalyses the transfer of sulphate from 3’-phophoadenosine 5’-phosphosulphate to position 6 of the N-sulphoglucosamine residue of heparan sulphate. The complete human cDNA has been cloned and partially cloned in Chinese hamster ovary cells (Habuchi et al., 1998). Heparan sulphate proteoglycans (HSPG) are ubiquitously present on the cell surface and in the extracellular matrix. They are known to interact with a variety of proteins such as heparin binding growth factors, extracellular matrix components, protease inhibitors and lipoprotein lipase. HSPG are implicated in morphological regulation during development, cell proliferation, differentiation, adhesion, migration, and also physiological phenomena such as viral infection and blood coagulation.

Heparan sulphate is required for high affinity binding of fibroblast growth factors to their receptors, as well as binding to other cell growth factors. It is postulated that different kinds of sulphotransferases affect the fine structure of heparan sulphate, in turn affecting morphological regulation and cell growth factors. With respect to the ovary, it is possible that full follicular growth factors are inhibited through disruption of the HS6ST gene and its effect on heparan sulphate ligand binding.

A particular role for sulphation of proteoglycans has been shown in the Drosophila oocyte (Sen et al., 1998). An enzyme similar to HS6ST has been expressed in follicle cells, and plays a pivotal role in the mechanism that establishes oocyte and furthermore embryonic polarity. A parallel role for HS6ST may occur within the human oocyte. The match of HS6ST with the Drosophila gene, egl, is intriguing, as it appears this gene is required for oocyte microtubule polarity. Immunoprecipitation experiments show that egl protein localizes to the oocyte at all stages of oogenesis from oocyte differentiation to its requirement for axis determination of the embryo (Mach et al., 1997).

E2F-related transcription factor

The 1440 bp mRNA sequence of E2F(DP1) showed a 75% match with PACdJ358H7 and a high similarity with the 408 amino acid sequence (70% identity, 73% similarity). The E2F family of mammalian transcription factors plays a critical role in the expression of genes that are required for cell cycle progression. They interact with proteins and bind to DNA as heterodimers. The E2F complex is capable of activating transcription and initiating apoptosis, as shown in fibroblasts, myeloid cells and glioma cells. In fact, the massive apoptosis that occurs when E2F expression is induced in glioma cells is now being considered as a therapeutic strategy for cancer gene therapy (Gomez et al., 1997).
Recently, a specific role for E2F in the Drosophila ovary has been described (Royzman et al., 1999). E2F regulates the expression of ORC2 protein (Origin Recognition Complex) in follicle cells, such that mutations in E2F cause female sterility by affecting cell cycle transition and chorion gene amplification. E2F is a single exon gene and it may be that the breakpoint interrupts its coding sequence or promoter. With regards to the ovary, POF may be caused by accelerated apoptosis of germ cells. On the other hand, E2F factors are neither ovary nor X-specific, and their effects are ubiquitous.

**Long interspersed nuclear elements (LINES or L1)**

L1 elements are mobile elements that occupy ~15% of the genome and are concentrated in AT rich regions (Kazazian et al., 1998). These retrotransposon elements are sequences that can be transcribed into RNA, reverse transcribed into cDNA and then reintegrated as such into the genome. LINE 1 insertions were first recognized as potential causal agents of human disease in 1988 when two separate insertions in the germ line or during early development, with six of the eight L1 insertions appearing in genes located on the X chromosome. This apparent predilection for the X chromosome may be due to the hemizygosity of X-linked genes in males. Since L1 insertions occur in the germline and are scattered on the X, the clustering of L1s around the breakpoint may have a similar significance in the human, by inducing mutations and diminishing fertility.

Alternatively, LINE 1 elements may have a direct role in the ovary. Drosophila melanogaster L1 factor belongs to the class of LINES and actively transposes in the germline of the female progeny from crosses between reactive females and inducer males. This process of L1 insertion is known as I-R hybrid dysgenesis and gives rise to reduced fertility, increased frequency of mutations and X chromosome loss (Dimitti et al., 1997). Recent studies (Goto et al., 1999) have demonstrated that LINE elements are differentially expressed in human female primordial germ cells (as compared to male) at the time of entry into meiosis, suggesting a role in ovarian development.

**Glypicans (GPC3)**

GPC3, a member of the glypican family, is an 8 exon gene spanning 500 kb, which has been mapped to human Xq26 by FISH. GPC3 is expressed ubiquitously in the embryo but is restricted to the colon and ovary in the adult. It has been shown (Lin et al., 1999) that GPC3 induces apoptosis in MCF-7 breast cancer cells in vitro, and it may also function as a tumour suppressor in the ovary. However there is no evidence that GPC3 is involved in the regulation of apoptosis in the normal ovary.

Mutations in the GPC3 gene are responsible for Simpson dysmorphia syndrome (Simpson et al., 1975) and are inherited as X-linked recessive. The phenotypic effects of this syndrome are a broad stocky appearance, large protruding jaw, widened nasal bridge, enlarged tongue, short hands and fingers together with normal intelligence, or, in more severe cases cleft palate, cardiac anomalies and mental retardation. With regards to ovarian effects, a case report of affected females did not mention any adverse effects (Pilia et al., 1996). A mother who was heterozygous for a GPC3 mutation did not have ovarian dysgenesis, although whether an early menopause ensued is unknown (Golabi and Rosen 1984). Although GPC3 fills the criteria for being a large multi-exon gene spanning a region known to be in the POF1 locus, women with POF do not show the phenotype of GPC3 mutations.

**Conclusions**

The aim of this study was to demonstrate the various techniques available to isolate candidate genes at a time when the sequence of the human genome rapidly unfolds. In summary we have identified three genes adjacent to a breakpoint at the POF1 locus, as well as a LINE 1 element cluster. Amongst these, we currently feel that HS6ST is the strongest candidate gene for POF. It is a gene which spans across four exons and hence is likely to be interrupted by a number of translocations in addition to deletions, as in this index family. In addition there
is a strong similarity with the Drosophila gene egl which has a role in fertility as well as the evidence for sulphotransferase activity in Drosophila oocytes.

Identification of POF genes would enable us to understand the mechanisms of ovarian development. On a practical level, the ability to provide an explanation to the mechanism of POF is therapeutic. Further, in our particular family the gross deletion at Xq26 makes it possible to predict POF prior to the onset of infertility. However mutation screening might reveal smaller defects in a POF1 gene which would be equally informative in women with familial or sporadic POF who are cytogenetically normal.

Acknowledgements
We thank the Sanger Centre for the supply of PACs. Dr Davison is supported by the Wellcome Trust.

References


ERRATUM

Page 2, paragraph 1, line 2: "include" should read "includes".
Page 12, paragraph 1, line 7: "where as" should be a single word.
Page 16, paragraph 2, line 7: "Gradishar et al 1989" should read "Gradishar & Schilsky, 1989".
Page 18, paragraph 2, line 7: "Rebar et al 1990" should read "Rebar & Connolly 1990".
Page 18, paragraph 4, line 21: "Guinet et al 1954" should read "Guinet & Pommatau 1954".
Page 20, paragraph 2, line 13: "Moncayo et al 1995" should read "Moncayo & Moncayo 1995".
Page 30, paragraph 1, line 1: "Lahn et al 1997" should read "Lahn & Page 1997".
Page 30, paragraph 2, line 11: "Stratakis et al 1994" should read "Stratakis & Rennet 1994".
Page 38, paragraph 3, line 6: "Foster et al 1994" should read "Foster & Graves 1994".
Page 42, paragraph 2, line 7: "Sala et al 1998" should read "Sala et al 1997".
Page 46, paragraph 2, line 10: "complimentary" should read "complementary".
Page 54, paragraph 2, line 1: "Programs" should read "Programmes".
Page 66, paragraph 3, line 10: "42-62 degrees" should read "42-62° C".
Page 71, paragraph 2, line 2: "two times" should read "twice".
Page 73, paragraph 1, line 1: "tempt" should read "temp".
Page 85: the mean ovarian volumes shown in the table were measured in mls.
Page 90, table 6: the father's allele size for DXS 8072 should read "218".
Page 95, paragraph 1, line 1: "Semi-qualitative" should read "semiquantitative".
Page 129, paragraph 4, line 2: "Kazazian et al 1998" should read "Kazazian & Moran 1996".
Page 130, paragraph 2, line 9: "hemizygousity" should read "hemizygosity".
Page 133, paragraph 1, line 2: "events in 1/670" should read "events is 1/670".
Page 135, paragraph 1, line 10: "Gomez et al 1999" should read "Gomez-Manzano et al 1999".
Page 137, paragraph 1, line 8: "Mach et al 1997" should read "Mach & Lehmann 1997".
Page 138, paragraph 3, line 1: "is associated" should read "are associated".

The following should have appeared in the reference section.