An investigation into FOXE1 polyalanine tract length in premature ovarian failure

Wendy J. Watkins¹, Sarah E. Harris², Megan J. Craven¹, Andrea L. Vincent³, Ingrid M. Winship⁴, Ksenija Gersak⁵ and Andrew N. Shelling¹,⁶

¹Department of Obstetrics and Gynaecology, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand, ²Department of Psychology, University of Edinburgh, Edinburgh, UK, ³Department of Ophthalmology, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand, ⁴Genetic Health Services, Royal Children’s Hospital, Parkville, Victoria, Australia and ⁵Department of Obstetrics and Gynaecology, University Medical Centre, Ljubljana, Slovenia

Premature ovarian failure (POF) is a common condition affecting 1% of women worldwide. There is strong evidence for genetic involvement in POF as many cases are familial, and mutations in several genes have been associated with POF. We investigated variation in FOXE1 polyalanine tract length, following the observation that polyalanine tract deletions are seen in the closely related FOXL2 in patients with POF. In addition, polyalanine tract expansions in FOXL2 are often seen in patients with blepharophimosis–ptosis–epicanthus inversus syndrome (BPES), a rare eyelid disorder often associated with POF. The FOXE1 polyalanine tract shows marked variation in its length between POF patients and normal controls, existing as an allele of 12, 14, 16, 17 or 19 alanine residues. We found evidence to suggest that variation in FOXE1 polyalanine tract length predisposes to POF.

Key words: blepharophimosis–ptosis–epicanthus inversus syndrome/forkhead/FOXE1/polyalanine tract/premature ovarian failure

Introduction

Premature ovarian failure (POF) or premature menopause is defined as ovarian failure before the age of 40. It is a common disorder affecting 1% of women worldwide (Coulam et al., 1986). As well as the health risks related to reduced estrogen levels over many years, such as osteoporosis, the major concern for affected women is typically infertility.

The causes of POF are diverse and largely idiopathic. It has been predicted that up to 30% of POF cases may result from genetic causes, as many women have a family history of POF (Vegetti et al., 1998). POF is likely to be a heterogeneous disorder caused by mutations in several genes, with each mutation identified thus far affecting a small number of patients (<10%) (Shellings et al., 2000; Harris et al., 2002).

Blepharophimosis–ptosis–epicanthus inversus syndrome (BPES) is a rare syndrome which is characterized by a distinctive eyelid phenotype associated with POF in type I BPES. The eyelid phenotype occurs in isolation in type II BPES (De Baere et al., 2001). Mutations causing this syndrome are spread throughout the FOXL2 gene (detailed in the Human FOXL2 Mutation Database at http://medgen.ugent.be/foxl2) (Beyesen et al., 2004). However, many are concentrated in the polyalanine tract, a designated ‘hotspot’.

FOXL2 belongs to the large family of forkhead (FOX) transcription factors which are widely expressed in a range of tissues, have diverse roles in development and metabolism and have been implicated in various forms of cancer and disease. They contain a highly conserved 100 amino acid DNA-binding (forkhead) domain. The polyalanine tract of the FOXL2 gene contains 14 alanine residues and this number is highly conserved across several mammalian species, which suggests the existence of strong functional and structural constraints (Cocquet et al., 2002).

Expansions within the polyalanine tract account for 25–30% of the reported FOXL2 mutations and lead mainly to BPES type II (De Baere et al., 2003). We have found evidence for FOXL2 involvement in isolated POF (Harris et al., 2002), including a deletion of 10 alanine residues and a single alanine deletion within the FOXL2 polyalanine tract in our cohort of isolated POF patients (Harris et al., 2002; Watkins et al., unpublished data).

Polyalanine tracts are found in a variety of genes and their length is often highly conserved among orthologues. While the exact function of polyalanine tracts is unknown, they may act as flexible spacer elements between functional domains (Karlin and Burge, 1996). There is evidence that alanine repeats are associated with transcriptional repression domains (Licht et al., 1990; Han and Manley, 1993a,b; Licht et al., 1994). Alterations from wild-type length have been associated with several disease phenotypes. Polyalanine tract expansions in many genes have been reported, including; ARX in X-linked mental retardation (OMIM 300382), HOXA13 in hand–foot genital syndrome (OMIM 142959), HOXD13 in synpolydactyly (OMIM 142989), PABP2 in oculopharyngeal muscular dystrophy (OMIM 164300), PHOX2B in congenital central hypoventilation syndrome (OMIM 209880), RUNX2 in cleidocranial dysplasia (OMIM 600211), SOX3 in X-linked mental retardation with growth hormone deficiency (OMIM 313430) and ZIC2 in holoprosencephaly (OMIM 603073).

There is less information on genes whereby deletions in the polyalanine tract have occurred, and even less correlating polyalanine tract deletions to human disease phenotypes. For RUNX2, a shorter allele with 11 alanine residues was found compared to the wild-type 17; however, this is thought to be an uncommon but normal variant, as it is also found in control samples (Mundlos et al., 1997). A loss of 9 of...
the 17 alanines in ZNF358 was found in 20 independent cases of Hirschsprung disease and 2 of 60 controls (Andrew et al., 2002). An allelic variant of transforming growth factor β receptor I (TGFβR1) with a shortened polyalanine tract (6A) has been linked to increased tumour susceptibility (Pasche et al., 1999; Baxter et al., 2002). However, recently a large population-based case–control study found that the 6A allele of the TGFβR1 does not appear to increase ovarian cancer risk (Spillman et al., 2005).

Polyalanine tracts are present in many transcription factors including two members of the FOX family of transcription factors, FOXE1 (OMIM *602617) and FOXL2 (OMIM *605597). Following the discovery of FOXL2 polyalanine tract deletions in POF patients, it was important to study other forkhead genes with a polyalanine tract in order to determine whether changes in this region might also be associated with POF. FOXE1, unlike FOXL2, has a polyalanine tract length which is polymorphic and has been reported to range from 11 to 17 alanine residues, although the most frequently occurring allele has 14 residues (Macchia et al., 1999; Hishinuma et al., 2001). Previously referred to as thyroid transcription factor 2 (TTF-2), FOXE1 is located at 9q22 and this single exone gene encodes a protein of 376 amino acids. Its expression has been detected in the human thyroid as well as the oropharyngeal epithelium and thymus (Trueba et al., 2005) and FOXE1 is thought to be involved in thyroid morphogenesis. Mutations in FOXE1 cause Bambforth–Lazarus Syndrome (OMIM #241850), which is characterized by athyroidal hypothyroidism, with spiky hair and cleft palate. FOXE1 polyalanine tract length has not been specifically correlated with these thyroid dysfunctions (Hishinuma et al., 2001). However, it is possible that certain alleles, or certain alleles in combination, are linked to a higher risk of thyroid complications. It was noted that between 10 and 20% of women with POF have an autoimmune disease, most commonly hypothyroidism (Conway, 1997).

FOXE1 is therefore an appropriate gene to investigate, in order to identify more genetic causes of POF. Results from this screen may provide clues about its role in autoimmune or a novel genetic replication mechanism which results in polyalanine tract errors, hence contributing to the pathogenesis of POF. We have screened this region of the FOXE1 gene in our cohort of POF patients from New Zealand and Slovenia as well as controls from both countries.

With the increased understanding of the genes and pathways involved in POF, it is hoped that screening will enable early detection of those at risk for developing POF, and in the future, treatments may be designed to return fertility to these women.

Materials and methods

Nomenclature

Nucleotide numbers refer to GenBank accession number NM_004473 for FOXE1, where A of the ATG, of the initiator methionine codon begins at 661. The polyalanine tract for this entry spans nucleotides 1150–1197.

Study group

POF patients from New Zealand and Slovenia were recruited for this study by the Departments of Obstetrics and Gynaecology in Auckland, New Zealand and Ljubljana, Slovenia, following institutional ethics guidelines to obtain informed consent. POF was defined as the cessation of menstruation for a period of >40 IU/l (two measurements taken at least a month apart). A complete medical and gynaecological history was taken from each patient as previously described (Shelling et al., 2000). Normal control samples were obtained from each of the general populations of New Zealand and Slovenia.

PCR

Genomic DNA was extracted from blood (Shelling et al., 2000) and 100ng was used as a template for PCR. Primers were designed to flank the polyalanine tract region of FOXE1 and produce a PCR product of 206bp. Primers were designed using the Primer Select module in the DNASTAR computer program from Lasergene® 1994 (DNASTAR Inc., Madison, WI, USA) and were as follows. FOXE1F: 5′-CTTCAAGGCTCGAAGCTCTC-3′ (nucleotides 1104–1124, accession no. NM_004473); FOXE1R: 5′-AGCCGCGGTTAGTACTG-3′ (nucleotides 1309–1289, accession no. NM_004473). Reaction conditions were 0.4μmol/l of each primer, 0.2μmol/l of each dNTP, 0.625 U Taq polymerase, 1 × Q solution and 1 × PCR buffer (Qiagen, GmbH, Hilden, Germany) in a 25μl total volume. Thirty cycles of PCR were performed, consisting of 1 min at 94°C, 1 min at 65°C and 1 min at 72°C, with a final 10 min extension at 72°C.

Denaturing high-performance liquid chromatography

Denaturing high-performance liquid chromatography (dHPLC) was used to accurately size PCR amplicons. For the FOXE1 amplicon, partial-denaturing conditions were used. After optimization, a single injection at one temperature (65.1°C) was required for sizing the range of variants. dHPLC analysis was carried out on the dHPLC 2100 WAVE DNA® Fragment Analysis System and DNAsep® column (Transgenic, Santa Clara, USA).

DNA sequencing

PCR products to be sequenced were purified using Roche’s High Pure PCR Product Purification Kit (Roche Diagnostics GmbH, Mannheim, Germany). Sequencing reactions were performed using the ABI PRISM®BIG DYETM Terminator Sequencing Kit version 3.1 under standard conditions and then separated on an ABI PRISM® 3100 Genetic Analyzer (PE Biosystems, Foster City, CA, USA). All sequencing reactions were performed at the DNA Sequencing Facility, The Centre for Genomics and Proteomics at The University of Auckland.

Results

Fifty-four New Zealand POF patients, 71 Slovenian POF patients, 120 New Zealand controls and 37 Slovenian controls were screened to ascertain polyalanine tract length for each allele of their FOXE1 gene. Eight different length variants were identified consisting of various combinations of five alleles, predicted to encode a polyalanine tract consisting of 12, 14, 16, 17 or 19 alanine residues in length (Table I). The preferred screening approach was to use dHPLC to size the different alleles. Eight samples showing distinct traces (Figure 1) were chosen for confirmatory sequencing, and other samples were then typed by comparison to these. Additional sequencing was employed for a limited number of samples which produced ambiguous dHPLC chromatograms.

The most common genotypes were the homozygous 14/14 and the heterozygous 14/16, whereby one allele of the FOXE1 gene consists of 14 alanine residues within the polyalanine tract, and the other allele consists of 16 alanine residues. These genotypes were each identified in a total of 107 subjects (37.9%) across all subject groups mentioned above. The rarest genotypes were 12/16 and 16/17, found in only three and a single subject, respectively. The 12/16 genotype was found in two New Zealand and one Slovenian POF patient and the 16/17 genotype in a Slovenian POF patient. Neither 12/16 nor 16/17 genotypes were identified in either of the control groups. The most common allele overall was 14, which accounted for 59.8% of total alleles across all groups screened (Table I). The rarest allele found was 17, at a frequency of 0.2% (Table I).

Upon close examination of the sequencing electropherograms, it was discovered that simple GCC repeat contraction and expansion alone do not account for the various alanine alleles (Figure 2), as there were minor permutations of the repeat length variants. For example, there were two possibilities (denoted 12a and b and 19a and b), which could not all have arisen as a direct contraction or expansion of an intermediate wild-type allele. Only one nucleotide sequence was found for the 14, 16 and 17 alleles, but as only a limited number of samples were sequenced, the existence of other nucleotide combinations in these cohorts and the wider population cannot be excluded.
Further studies are required to determine the true level of variation of this coding allele, in normal and disease populations. Fisher’s exact test was used to compare individual allele frequencies between Total POF and Total controls. The $FOXE1_{14}$ alanine allele was significantly less common in the Total POF patient group (126/250) than the Total controls (211/314) ($p$ value = 0.0001). The $FOXE1_{16}$ alanine allele was significantly more common in the Total POF patient group (112/250) than the Total controls (89/314) ($p$ value = 0.0001). This was also confirmed in the Slovenian subpopulation after Fisher’s exact test, which revealed the $FOXE1_{16}$ alanine allele was significantly more common in the Slovenian POF patient group (73/142) than the Slovenian controls (17/74) ($p$ value = 0.0001). After Bonferroni’s adjustment for multiple statistical analyses, with the number of possible comparisons as 15, the alpha level is reduced from 0.05 to 0.003 and the above findings remain significant.

**Discussion**

We were interested to look at the $FOXE1$ polyalanine tract length in our cohort of POF patients and controls. $FOXE1$ polyalanine tract length has already been shown to be polymorphic (Macchia et al., 1999; Hishinuma et al., 2001); however, there exists the possibility that the presence of certain alleles might predispose to POF. The polyalanine tract length of $FOXE1$ is of interest as $FOXL2$, a gene commonly mutated in BPES and more recently found to be altered in patients with isolated POF, contains a highly conserved polyalanine tract (De Baere et al., 2001; Harris et al., 2002). The polyalanine tract is frequently the site of mutations, with expansions being a common occurrence in BPES (De Baere et al., 2001), while a large deletion in the polyalanine tract has been found in an isolated POF case (Harris et al., 2002; Gersak et al., 2004). Alterations from wild-type polyalanine tract length have also been associated with several disease phenotypes in many other genes (Brown and Brown, 2004).

A comparison of individual allele frequencies between total POF and total controls using Fisher’s exact test with Bonferroni’s adjustment revealed two significant associations (Table I). The $FOXE1_{14}$ alanine allele is significantly less common in the Total POF patient group than in the Total controls. The $FOXE1_{16}$ alanine allele is significantly more common in the Total POF patient group than in the Total controls. The $FOXE1_{16}$ alanine allele is significantly more common in the Slovenian POF patient group than in the Slovenian controls.

<table>
<thead>
<tr>
<th>Allele frequency [n (%)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

$p$ Value remains significant after Fisher’s exact test with Bonferroni’s correction for multiple testing: $p$ value = 0.003.

*The $FOXE1_{14}$ alanine allele is significantly less common in the Total POF patient group than in the Total controls.

†The $FOXE1_{16}$ alanine allele is significantly more common in the Total POF patient group than in the Total controls.

‡The $FOXE1_{16}$ alanine allele is significantly more common in the Slovenian POF patient group than in the Slovenian controls.
controls. In addition, this is confirmed within the Slovenian subpopulation. The FOXE1 16 alanine allele is significantly more common in the Slovenian POF patient group than in the Slovenian controls. This suggests that alteration from the most frequently occurring 14 alanine allele to the 16 allele increases the risk of developing POF.

After Bonferroni’s correction, significant differences are not seen for the 12, 17 and 19 alanine alleles but these alleles occur at much lower frequencies than the 14 and 16 alleles. Larger population numbers would therefore be required to investigate this further.

It is interesting to note that there are population differences in allele frequencies in the FOXE1 polyalanine tract as shown in Table II. The FOXE1 12 alanine allele is absent in Japanese controls (Hishinuma et al., 2001). The most common allele in all populations is the 14 alanine allele which is seen at almost equal frequencies in the New Zealand (67.5%) and Slovenian controls (66.2%) and slightly lower in the Italian study (54%). In the Japanese controls, this allele is overrepresented at 97% (Table II). This may suggest that the Japanese population is more homogeneous than the other studied populations. The 16 alanine allele is found in a narrow range of frequencies in the New Zealand (30%), Slovenian (23%) and Italian (40%) controls, whilst only 3% of the Japanese study population have this allele. The 17 alanine allele is only found in controls from the Italian study at a low frequency of 4%, although it is also found at low frequencies in the Slovenian POF (0.7%) group (Table I). The 19 alanine allele was not described by either of the previous studies, and was found in both our control groups at a frequency of 9.5% for Slovenians and 2.1% for New Zealand controls. It is apparent from this data that there may be population-specific differences in the allele frequencies for polyalanine tract length of the FOXE1 gene. Comprehensive screening with greater subject numbers across more populations is needed to determine what these population-specific alleles are. A FOXE1 polyalanine tract with 13, 15 or 18 alanine residues has not yet been described.

Different mechanisms have been proposed for the occurrence of expansions and/or deletions which include replication slippage and unequal crossing-over (Brown and Brown, 2004; Chen et al., 2005; Robinson et al., 2005). Replication slippage is thought to account for the expansion at mutation loci composed of perfect trinucleotide repeats. The fundamental feature of this model is that hairpins may cause the DNA polymerase to pause while replicating, causing the nascent and template strands to transiently dissociate and then reassociate in a misaligned configuration (Nag, 2003). If slipping forward occurs, this will lead to deletion, and if slipping backward occurs, this will lead to insertion (Chen et al., 2005). Diseases associated with expansions of a perfect trinucleotide repeat show that the unstable allele tends to expand further in successive generations, leading to a reduced age of onset and/or increased severity (Cummings and Zoghbi, 2000).

Interestingly, we have discovered two different nucleotide sequences which make up the 12 and 19 alanine FOXE1 alleles leading us to speculate that different mechanisms have operated to bring about these contractions and expansions, but resulting in the same protein length. It has previously been reported that the FOXE1 polyalanine tract length is determined by the number of GCC repeats, being 7, 9, 11 or 12, corresponding to alanines stretches of 12, 14, 16 and 17, respectively (Macchia et al., 1999). This is based on a preceding sequence of GCG GCC GCT GCC GCA; however, we have discovered a 12 alanine allele consisting of GCG GCC GCT (GCC)₂ (12a, Figure 2). In addition, the 19 alanine allele does not appear to exist as predicted by the pattern previously reported being: GCG GCC GCT GCC GCA (GCC)₄ and instead exists as either GCG GCC (GCT GCC GCA GCC GCC)₂ (GCC)₂ (denoted 19a in Figure 2) or GCG GCC GCT GCC GCA GCC GCC)₂ (GCC)₂ (denoted 19b in Figure 2). It is therefore likely that simple expansions are not the only mechanism which has resulted in the variation in FOXE1 polyalanine tract length observed. Population studies would be required to investigate this further.

The exact function of polyalanine tracts is unknown, but when expansions take place within polyalanine tracts of genes encoding transcription factors, this may result in altered protein–DNA or protein–protein interactions and either loss of function or gain of function (Brown and Brown, 2004). It is unclear whether the activity of FOXE1 as a transcription factor is different when the polyalanine tract length (screened previously) was investigated to see if there was any correlation to their FOXE1 polyalanine tract length. We postulated that certain individuals may be more prone to sporadic expansions/deletions in more than one gene due to some sort of faulty mechanism.

### Table II. Frequency of FOXE1 polyalanine tract alleles in normal populations (%)

<table>
<thead>
<tr>
<th>Study</th>
<th>Nationality of controls</th>
<th>Allele</th>
<th>Watkins et al. (This study)*</th>
<th>Watkins et al. (This study)$†$</th>
<th>Macchia et al. (1999)‡</th>
<th>Hishinuma et al. (2001)$§$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>New Zealand</td>
<td>12</td>
<td>0.4</td>
<td>1.3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>67.5</td>
<td>66.2</td>
<td>54</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>30.0</td>
<td>23.0</td>
<td>40</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>2.1</td>
<td>9.5</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- *240 alleles screened.
- †74 alleles screened.
- ‡196 alleles screened.
- §202 alleles screened.

Figure 2. An alignment of FOXE1 alleles showing the different combinations of nucleotides which make up polyalanine tract length of 12, 14, 16, 17 or 19 alanine residues.
of DNA replication such as replication slippage or unequal crossing over. However, this does not seem to be the case. In POF patients with either a FOXL2 deletion of 10 alanine residues (A221_230del), or a single alanine residue, (A224del) we observed the most frequently occurring FOXL2 genotypes of either 14/14 or 14/16. Similarly, two BPES patients with 10 alanine expansions (A221–231dup and A224–234dup) in FOXL2 presented with the most common FOXL2 genotypes found in this study of 14/16 and 14/14, respectively. A total of 18 BPES patients were screened (data not shown).

Conclusion

The FOXE1 polyalanine tract shows marked variation in length between POF patients and controls, existing as alleles of 12, 14, 16, 17 or 19 alanine residues, in the populations we have screened. In addition, there are at least two different nucleotide sequences which make up the 12 and 19 alanine alleles. There may also be more variations in alanine codons for the 14, 16 and 17 alleles, and further studies are required to investigate this further. The most frequently occurring allele consists of 14 alanines in all populations screened. The frequency of the 14 alanine allele was significantly less in POF patients than in controls and the frequency of the 16 alanine was significantly more in POF patients than in controls. Variation in FOXE1 polyalanine length predisposes to POF. Larger population-based studies and functional studies are required to support this.

Acknowledgements

We would like to thank the POF and BPES patients for their involvement in this study. We would also like to thank the many clinicians who referred these patients.

References


Submitted on December 15, 2005; accepted on January 13, 2006

Downloaded from https://academic.oup.com/molehr/article-abstract/12/3/145/999476 by guest on 28 January 2019