Cytogenet Genome Res 101:206–211 (2003)
DOI: 10.1159/000074338

Cytogenetic and Genome Research

Structure, evolution and expression of the FOXL2 transcription unit

J. Cocquet,[a] E. De Baere,[b] M. Gareil,[a] M. Pannetier,[c] X. Xia,[d] M. Fellous,[a] and R.A. Veitia[a]

[a] INSERM E0021 and U361, Reproduction et Physiopathologie Obstétricale, Hôpital Cochin, Paris, Université Paris VII (France)
[b] Department of Medical Genetics, Ghent University Hospital, Ghent (Belgium)
[c] Biologie du Développement et Reproduction, INRA Bât J.Poly, Jouy-en-Josas (France)
[d] Department of Biology and the Center for Advanced Research in Environmental Genomics, University of Ottawa, Ottawa, Ontario (Canada)

Abstract. FOXL2 is a putative transcription factor involved in ovarian development and function. Its mutations in humans are responsible for the blepharophimosis syndrome, characterized by eyelid malformations and premature ovarian failure (POF). Here we have performed a comparative sequence analysis of FOXL2 sequences of ten vertebrate species. We demonstrate that the entire open reading frame (ORF) is under purifying selection leading to strong protein conservation. We also review recent data on FOXL2 transcript and protein expression. FOXL2 has been shown 1) to be the earliest known sex dimorphic marker of ovarian determination/differentiation in vertebrates, 2) to have, at least in mammals, an ovarian expression persisting until adulthood. The conservation of its sequence and pattern of expression suggests that FOXL2 might be a key factor in the early development of the vertebrate female gonad and involved later in adult ovarian function. Finally, we provide arguments for the existence of an alternative transcript in rodents, that may arise from a differential polyadenylation. Although it has only been demonstrated in rodents, its presence/absence in other species deserves further investigation.

FOXL2 is a putative winged helix/forkhead transcription factor gene (MIM 605597) involved in ovarian development and function. Its dominant mutations have recently been shown to be responsible for the blepharophimosis syndrome (or BPES for Blepharophimosis Ptosis Epicanthus inversus Syndrome, MIM 110100) a rare genetic disease characterized by eyelid malformations associated with premature ovarian failure (POF) (BPES type I) or occurring isolated (BPES type II) (Zlotogora et al., 1983; Crisponi et al., 2001; De Baere et al., 2001). Until now, no clear-cut pathogenic mutations have been found in patients affected with isolated POF (De Baere et al., 2002; Harris et al., 2002).

As a member of the large family of winged helix/forkhead transcription factors, FOXL2 contains a characteristic 110-amino acid DNA-binding domain, the forkhead. Many forkhead proteins are known to be involved in different developmental and metabolic processes (Kaufmann and Knöchel, 1996; Carlsson and Mahlapuu, 2002) and mutations in some of them are responsible for genetic developmental diseases (Carlsson and Mahlapuu, 2002). FOXL2 also contains a polyalanine tract (polyAla), whose role (if any) has not been identified yet. Expansions from 14 to 24 alanine residues in this region represent 30% of all FOXL2 mutations detected in the ORF and lead mainly to BPES type II (De Baere et al., 2003).

The nucleotide and amino acid sequences of FOXL2 in four vertebrates, ranging from fish to man, have previously been compared and shown to be highly conserved (Cocquet et al.,

J.C. is supported by a grant from the French Minister of Education, Research and Technology (MERT). E.D.B. is supported by the Fund for Scientific Research (FWO). M.F., M.G. and R.A.V. are funded by the University of Paris VII. X.X. is funded by the University of Ottawa and Natural Science and Engineering Research Council of Canada (NSERC).

Received 3 June 2003; manuscript accepted 30 July 2003.

Request reprints from Reiner A. Veitia
INSERM U361, Reproduction et Physiopathologie Obstétricale
Hôpital Cochin, 75014 Paris (France)
telephone: 33 1 43 26 28 26; fax: 33 1 43 26 44 08
e-mail: veitia@cochin.inserm.fr

KARGER
Fax + 41 61 306 12 34
E-mail karger@karger.ch
www.karger.com

0301–0171/03/1014–0206$19.50/0

Accessible online at:
www.karger.com/cgr

Table 1. Estimated ratio of synonymous (dS) and non-synonymous (dN) substitution rates obtained using the CODEML program in the PAML package (lower matrix, Yang, 2002). *t*: Genetic distances based on dS and dN (upper matrix). Pairwise comparisons result in very low dN/dS ratios confirming FOXL2 protein is under a strong purifying selection throughout evolution

dN/dS, *t*	Human	Goat	Cow	Pig	Rabbit	Mouse	Rat	Pufferfish	Zebrafish	Tetraodon
Human		*0.4811*	*2.1789*	*0.0271*	*0.3775*	*0.3826*	*0.2809*	*17.9228*	*39.0000*	*32.1378*
Goat	0.0039		*0.0812*	*2.9036*	*0.3764*	*0.4427*	*2.8280*	*39.0000*	*39.0000*	*39.0000*
Cow	0.0010	0.0168		*3.0764*	*0.3681*	*0.5533*	*2.9948*	*39.0000*	*39.0000*	*39.0000*
Pig	0.0306	0.0010	0.0010		*0.4167*	*0.3388*	*0.3026*	*26.0215*	*39.0000*	*39.0000*
Rabbit	0.0053	0.0040	0.0045	0.0068		*0.2130*	*0.5005*	*22.1978*	*39.0000*	*32.8754*
Mouse	0.0051	0.0029	0.0024	0.0086	0.0081		*0.4178*	*39.0000*	*39.0000*	*39.0000*
Rat	0.0028	0.0010	0.0010	0.0055	0.0064	0.0077		*24.9313*	*39.0000*	*39.0000*
Pufferfish	0.0026	0.0012	0.0012	0.0019	0.0019	0.0012	0.0020		*33.1133*	*0.2717*
Zebrafish	0.0022	0.0021	0.0023	0.0023	0.0021	0.0022	0.0023	0.0011		*30.2038*
Tetraodon	0.0014	0.0011	0.0012	0.0012	0.0012	0.0012	0.0012	0.0034	0.0011	

2002). Here, we have performed a broader comparative analysis including seven mammalian and three non-mammalian vertebrate species. We confirm and extend our findings showing the conservation of the FOXL2 open reading frame (ORF) and protein sequences, especially at the level of the forkhead domain and the C-terminal region, while the homopolymeric runs of amino acids, such as the polyalanine tract, are less conserved.

The expression of FOXL2 has already been studied extensively by us and others, in mammals (Crisponi et al., 2001; Cocquet et al., 2002) as well as in non-mammalian vertebrate species (Loffler et al., 2003). Here, we review recent data on FOXL2 mRNA and protein pattern of expression and also provide novel data on protein expression in the human developing eyelid and peri-ocular muscles.

In addition, we provide evidence for the existence of a second alternative transcript in rodents, which is likely to result from a differential polyadenylation. Our results strongly suggest that this second transcript contains the entire *Foxl2* ORF.

FOXL2 protein sequence is highly conserved

In a previous study, we have demonstrated that the FOXL2 protein is highly conserved in four species (human, goat, mouse and pufferfish) and is under a purifying selection (Cocquet et al., 2002). We aimed to extend our study by performing an analysis in other mammalian and non-mammalian sequences. FOXL2 sequences for *Homo sapiens* (human), *Capra hircus* (goat), *Mus musculus* (mouse), *Rattus norvegicus* (rat) and *Fugu rubripes* (pufferfish) were collected or assembled from GenBank (AI: AF301906, AY112725, AF522275, AC105826 and Scaffold–8165/ProtJGI–24134, respectively). The *Danio rerio* (zebrafish) and *Tetraodon nigroviridis* (tetraodon) sequences were obtained from public genome resources by interrogating them with the pufferfish *Foxl2* sequence. *Sus scrofa* (pig), *Oryctolagus cuniculus* (rabbit), *Bos taurus* (cow) and *Macropus eugenii* (tammar wallaby) sequences were directly obtained by PCR on genomic DNA using the primers described previously (pBAD-FOXL2F and pBAD-FOXL2R; Cocquet et al., 2002) and the primer FOXL2-R420: 5′ CTC GAA CAT GTC PyTC GCA GGC CGG GTC 3′ (AI: AY340971 for *S. scrofa*, AY 340972 for *O. cuniculus*, AY340970 for *B. taurus* and AY 340969 for *M. eugenii*).

We have previously used Li's method to compare the FOXL2 coding sequence among different species (Li, 1993; Cocquet et al., 2002). However, a statistically more justifiable method to calculate the rate of synonymous substitutions (dS) (i.e. when a change in the nucleotide sequence does not lead to a change in the corresponding amino acid) and non-synonymous (dN) substitutions between a pair of homologous sequences is the maximum likelihood method implemented in the CODEML program of the PAML package (Yang and Bielawski, 2000; Yang, 2002; Xia 2001). We used therefore this method to compare FOXL2 nucleotide and protein sequences in all the above mentioned species. The ratio dN/dS is a good indicator of the selective pressure at the protein level: dN/dS = 1 implies no selection; dN/dS < 1, a purifying selection; and dN/dS > 1, a positive selection (Yang and Bielawski, 2000).

Pairwise comparisons resulted in low dN/dS ratios confirming that FOXL2 is under a strong purifying selection throughout evolution (Table 1). We also calculated the genetic distances based on dN and dS results (*t*, Table 1). Interestingly, the multiple alignment of the protein sequences shows that homopolymeric runs of amino acids are not conserved among species, in contrast with the high conservation of the rest of the protein (Fig. 1). Indeed, in the seven eutherian mammals studied, the length of the polyalanine region is strictly conserved (i.e. 14 residues) whereas in the three fish species this domain is absent. This suggests that the polyAla tract length in mammals is under strong constraints limiting its maximum length to 14 residues. This view is supported by the deleterious effect of its expansion in the blepharophimosis syndrome (30% of reported mutations, De Baere et al., 2003). The aforementioned constraints are probably structural rather than functional, as several studies have shown that alanine runs present in mammalian genes are absent or shortened in non-mammalian homologues (Nakachi et al., 1997; Cocquet et al., 2003). However, the polyAla tract might tune transcription factor activity (Nakachi et al., 1997; Cocquet et al., 2003) with 14 Ala residues being the optimum length for FOXL2 function in mammals. This would explain why the polyAla tract, when present, is always (as far as observed) composed of 14 residues. Interestingly, an isolated POF case carrying a FOXL2 allele with a shorter polyAla

```
            10        20        30        40        50        60        70        80
    ----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---
M   MMASYPEPEDTAGTLLAPESGRAVKEAEASPP-SPGKGGG----------TTPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
R   MMASYPEPEDTAGTLLSPESGRAVKEAEASPP-SPGKGGG----------TAPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
B   MMASYPEPENASGALLAPETGRAAKEPEAPPPPSPGKGGGGGTG------TAPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
C   MMASYPEPEDASGALLAPETGRAAKEPEAPAPPSPGKGGGGGAG------AAPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
S   MMASYPEPEDAAGALLAPETGRTAKEPEAPPPLSPGKGGGGGAS------TAPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
H   MMASYPEPEDAAGALLAPETGRTVKEPEGPPP-SPGKGGGGGGG------TAPEKPDPAQKPPYSYVALIAMAIRESAEKRLT
O   MMASYPEPEEAAGALLAPESGRAAKEPEAPPP-SPGKGGGGGGGGGS----AAEKPDPAQKPPYSYVALIAMAIRESAEKRLT
F   MMATYQNPEDDAMALMVHDTNTTKEKERPKEEPVQDK--------------VEEKPDPSQKPPYSYVALIAMAIRESSEKRLT
T   MMATYQNPEDDAMALMIHDTNTTKEKERPKEEPVQDK--------------VEEKPDPSQKPPYSYVALIAMAIRESSEKRLT
D   MMATYPGHEDNGMILMDT-TSSSAEKDRTKDEAPPEK--------------GPDKSDPTQKPPYSYVALIAMAIRESSEKRLT
Ma  MMASYPEPEDDAGALLAPEPATGRPGKDSEPRGPRDCKEELSPEKSGGGGGVAEKPDPSQKPPYSYVALIAMAIRESAEKRLT
    *** *   *     *                                       * ** ****************** *****
          90        100       110       120       130       140       150       160
     -|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|--
M    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
R    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
B    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
C    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
S    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
H    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
O    LSGIYQYIIAKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
F    LSGIYQYIISKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
T    LSGIYQYIISKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
D    LSGIYQYIISKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFEKGNYRRRRRMKRPFRP
Ma   LSGIYQYIIGKFPFYEKNKKGWQNSIRHNLSLNECFIKVPREGGGERKGNYWTLDPACEDMFE
     ********* *********************************************************************

           170       180       190       200       210       220       230       240
     --|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-
M    PPAHFQPGKGLFGSGGAAGGCGVPGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
R    PPAHFQPGKGLFGSGGGAGGCGVPGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
B    PPAHFQPGKGLFGAGGAAGGCGVAGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
C    PPAHFQPGKGLFGAGGAAGGCGVAGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
S    PPAHFQPGKGLFGAGGAAGGCGVAGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
H    PPAHFQPGKGLFGAGGAAGGCGVAGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
O    PPAHFQPGKGLFGAAGAAGGCGVAGAGADGYGYLAPPKYLQSGFLNNSWPLPQPPSPMPYASCQMAAAAAAAAAAAAAA
F    PPTHFQPGKSLFG--------------GDGYGYLSPPKYLQSSFMNNSWSLGQPPAPMSYTSCQMASGNVSPVN-----
T    PPTHFQPGKSLFG--------------GDGYGYLSPPKYLQSSFMNNSWSLGQPPPPMSYTSCQMASGNVSPVN-----
D    PPTHFQPGKSLFG--------------GEGYGYLSPPKYLQSGFINNSWS----PAPMSYTSCQVSSGSVSPVN-----
     ** **********                ***** ******* * ****     * ** * ***

            250       260       270       280       290       300       310       320
     ---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
M    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPPPPP---HPHPHPHAHHLHAAAAPPP
R    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPPPR----HPHPHPHAHRLHAAAAPPP
B    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPP------HPHSHPHAHHLHAAAAPPP
C    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPP------HPHSHPHAHHLHAAAAPPP
S    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPP------HPHSHPHAHHLHAAAAPPP
H    GPGSPGAAAVVKGLAGPAASYGPYTRVQSMALPPGVVNSYNGLGGPPAAPPPPP------HPHPHPHAHHLHAAAAPPP
O    GPGSPGAAAVVKGLAGPAASYGPYSRVQSMALPPGVVNSYNGLGGPPAAPPPPPPPPPPPHPHPHPHAHHLHAAAAPPP
F    ----------VKGLSAP-TSYNPYSRVQSMALPS-MVNSYNGMS-------------------HHHHPAHPH-------
T    ----------VKGLSAP-TSYNPYSRVQSMALPS-MVNSYNGMS-------------------HHHHPAHPH-------
D    ----------MKGLSAP-SSYNPYSRVQSIGLPS-MVNSYNGIS-------------------HHHHHHHTHPHALP--
                ***  *  ** ** ****  **   ******                      * *    *

             330       340       350       360       370       380       390
     ----|----|----|----|----|----|----|----|----|----|----|----|----|----|
M    APPHHGAAAPPPGQLSPASPATAAPPAPAPTSAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
R    APPHHGAAAPPPGQLSPASPATAAPPAPAPTSAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
B    APPHHGAAAPPPGQLSPASPATAAPPAPAPTNAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
C    APPHHGAAAPPPGQLSPASPATAAPPAPAPTNAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
S    APPHHGAAAPPPGQLSPASPATAAPPAPAPTNAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
H    APPHHGAAAPPPGQLSPASPATAAPPAPAPTSAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
O    APPHHGAAAPPPGQLSPASPATAAPPAPAPTSAPGLQFACARQP-ELAMMHCSYWDHDSKTGALHSRLDL
F    -------------HAQQLSPATAAPPPVSSSNGAGLQFACSRQPAELSMMHCSYWEHETKHSALHTRIDI
T    -------------HAQQLSPATAAPPPVSSSNGAGLQFACSRQPSEFSMMHCSYWEHETKHSPYHTRIDI
D    -------------HAQQLSPATAAAPPVTTGNGTGLQFACSRQPAELSMMHCSYWDHESKHSALHARIDI
                       ****** *        ****** *** *  ******* *  *    * * *
```

Fig. 1. Multiple alignment of FoxL2 protein of ten species. M: *Mus musculus*, mouse (GenBank AI: AF522275); R: *Rattus norvegicus*, rat (AI: AC105826); B: *Bos taurus*, cow (AI: AY340970); C: *Capra hircus*, goat (AI: AY112725); S: *Sus scrofa*, pig (AI: AY340971); H: *Homo sapiens*, human (AI: AF301906); O: *Oryctolagus cuniculus*, rabbit (AI: AY340972); F: *Fugu rubripes*, pufferfish (AI: Scaffold_8165/Prot JGI_24134); T: *Tetraodon nigroviridis*, tetraodon; D: *Danio rerio*, zebrafish and of FoxL2 partial protein of the *Macropus eugenii* (Ma), tammar wallaby (AI: AY340969). The regions of the FOXL2 primers used for PCR amplification are represented in italics. The forkhead domain and the polyalanine region are indicated in bold. Asterisks indicate fully conserved sites. Notice that many of the changes are highly conservative.

stretch has been reported (Harris et al., 2002) but further studies are required to establish a formal link between polyAla contraction and isolated POF. Other amino acid repeats in FOXL2 (such as proline [Pro], glycine [Gly]) have unequal lengths in mammals and are absent in fishes. In a recent genomic study, we have shown that human proteins containing polyAla tracts have a tendency to carry runs of other amino acids, encoded, as alanine, by GC-rich codons (i.e. Gly, Pro and histidine [His]). Besides, the Ala, Gly, Pro and His contents of these proteins are correlated with the GC contents of the third codon base, suggesting that their amino acid composition reflects the constraints operating on the genomic compartments in which their ORFs lie (i.e. isochores) (Cocquet et al., 2003). Compositional constraints acting on the FOXL2 ORF may be different between mammals and fishes and could explain the differences of amino acid runs that we have observed.

Our analysis also reveals that the N-terminal region of FOXL2 (the first 45 amino acids) is less conserved than the C-terminal region, when excluding low complexity repeats (i.e. homopolymeric runs). Note that, in the latter region, most of amino acid changes are highly conservative. Thus, the N-terminal region might be responsible for functional differences among species or has a less important role (if any) in FOXL2 function. On the contrary, the well-conserved C-terminal region might contain yet unidentified functional domain(s).

We focused then on the forkhead domain, including the sequences of the horse *(Equus caballus)*, obtained by PCR as previously described, the chicken *(Gallus gallus)* and the turtle

(Trachemys scripta) (respective AI: AY155534 and AY155535). We found that the amino acid sequence is identical among eutherian mammals and that only a few minor changes, at three sites, exist when we extend the comparison to other vertebrates (alignment available from authors upon request). These changes concern amino acids that are not charged and thus do not disturb the global isoelectric point of the forkhead domain, estimated to be 10.87 in all the studied species. This highly basic isoelectric point has to be correlated with the DNA-binding activity of this domain.

The sequence and properties of FOXL2 forkhead domain are highly conserved, which is a general characteristic of forkhead transcription factors (Carlsson and Mahlapuu, 2002).

FOXL2 pattern of expression

The pattern of expression of FOXL2 has already been studied at both the mRNA and protein levels in different species and has provided interesting insights. In human, mouse and goat FOXL2 has only been detected in developing eyelids and in fetal and adult ovaries while it has not been detected in testis at any stage. Its protein localization has been shown to be nuclear, which is in line with its putative function as a transcription factor (Crisponi et al., 2001; Cocquet et al., 2002).

Here, we provide data on FOXL2 protein expression in human developing eyelids: as shown in Fig. 2, it is located in a wide region of the primordial mesenchyma of the developing eyelids. Magnetic resonance imaging (MRI) performed on the orbit of BPES patients has revealed the absence or hypotrophy of the eyelid superior levator muscle (Dollfus et al., 2003). The authors have suggested that FOXL2 may be involved in the development of this muscle. However, the much wider expression domain of the protein (in the bulging and surrounding primordial mesenchyma, Fig. 2) also suggests a role in the development of other peri-ocular muscles, which is supported by their same mesodermal origin and the relatively frequent observation of strabismus in BPES patients (Oley and Baraitser, 1988; Barishak et al., 1992).

FOXL2 ovarian expression in mammals begins early in development, before the onset of folliculogenesis, and persists until adulthood. It is restricted to the somatic compartment: the follicular (granulosa) cells display a strong protein expression, the stromal cells a more diffuse one, whereas in the oocytes no signal has been observed (Cocquet et al., 2002; Pannetier et al., 2003). The staining of the mouse stromal cells is almost undetectable, in line with the RNA in situ hybridization results of Crisponi et al. (2001). In a recent study, *Foxl2* mRNA has been observed in both granulosa cells and some oocytes of fetal and adult mouse ovaries (Loffler et al., 2003). Low levels of FOXL2 transcript have been detected in goat and mouse testis, at fetal and adult stages, but no protein could be found in any of those tissues (Pailhoux et al., 2001; Cocquet et al., 2002; Pannetier et al., 2003). Thus, either the levels of transcript are too weak to be significant, or its expression undergoes a post-transcriptional or translational regulation. One of these hypotheses may also explain the presence of *Foxl2* transcript in some mouse oocytes where the protein remains undetectable.

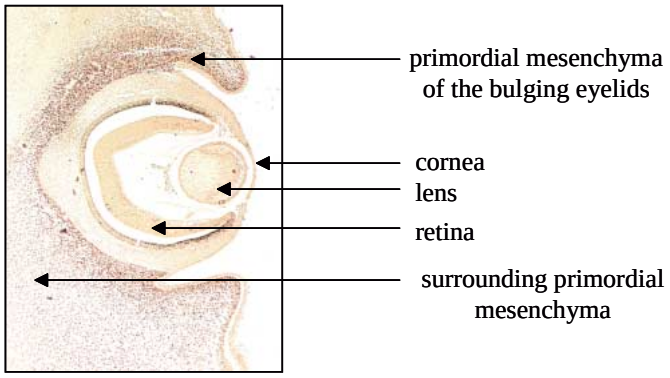

Fig. 2. Immunohistochemistry performed with anti-C-terminal FOXL2 antibody on human developing eyelid section (7th week, Carnegie stage 19) as previously described (Cocquet et al., 2002). FOXL2 is expressed in the bulging and surrounding primordial mesenchyma suggesting a role in the development of extra-ocular muscles consistent with the BPES phenotype (Oley and Baraitser, 1988).

Interestingly, mouse *Foxl2* transcript (previously named *P-Frk* for Pituitary forkhead factor) is also expressed in the pituitary Rathke's pouch (Treier et al., 1998) and has thus been suggested to be involved in pituitary organogenesis. Recently, protein expression has been shown to be consistent with these results (Buffy Ellsworth, personal communication).

Many forkhead proteins are responsible for differentiation processes during development and are later recycled to control metabolism in the adult (Carlsson and Mahlapuu, 2002). The ovarian somatic expression of FOXL2 in mammals beginning at an early stage of development and persisting in adult life suggests a similar dual role: at first, in the ovarian somatic cell determination/differentiation and later, in maintenance of the adult ovarian function.

In non-mammalian vertebrates, *Foxl2* mRNA has also been shown to be expressed at an early stage in the developing female gonad and to have a sex-dimorphic expression. In the chicken, it is expressed in the two ZW (female) gonads at early stages of development (at least from day 5 to 8) whereas no expression was found in male gonads at the same stages. In the turtle, a species undergoing temperature-dependent sex determination, *Foxl2* is expressed in the developing gonads at a higher level at female than male promoting temperatures (Loffler et al., 2003).

The conservation of FOXL2 sequence and pattern of expression throughout evolution leads to the conclusion that it is, by now, the earliest known sex-dimorphic marker of ovarian determination/differentiation in vertebrates and might be a key gene in early development and maintenance of the vertebrate female gonad.

Rodents possess alternative *Foxl2* transcripts

The previous discussion demonstrates a strong conservation of FOXL2 sequence and pattern of expression among different species. Nevertheless, when exploring the GenBank expressed sequence tag (EST) division, we found a potentially interesting

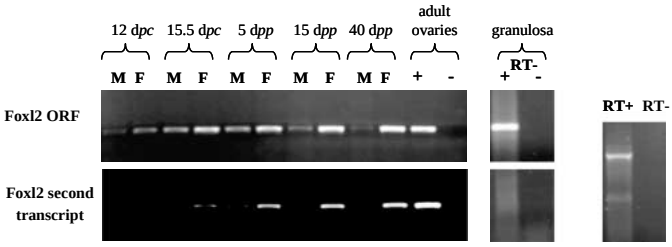

Fig. 3. RT-PCR analysis of mouse *Foxl2* at different developmental stages (12, 15.5 days post coitum and 5, 15, 40 days post partum) and in the mouse granulosa cell line AT29C. cDNAs are normalized with respect to β actin (as in Pannetier et al., 2003). (**A**) Upper panel: RT-PCR that amplifies 370 bp of *Foxl2* ORF using primers pBAD-FOXL2F, described above, and FOXL2-B, described by Crisponi et al. (2001). Lower panel: amplification with primers specific of the *Foxl2* second transcript (amplicon length: 415 bp). (M) and (F) stand for male and female, respectively. (RT+): adult ovary cDNA. (RT–): adult ovary mRNA, no reverse transcriptase (negative control). Granulosa: mouse granulosa cell line AT29C with (RT+) or without (RT–) reverse transcriptase. (**B**) RT-PCR with primers that amplify from the beginning of the *Foxl2* ORF to the end of the second transcript (pBAD-FOXL2F and mFoxl2polyAde-R, described above; amplicon length: 2.7 kb). (RT+): adult ovary cDNA. (RT–): adult ovary mRNA, no reverse transcriptase (negative control).

difference between rodents and other species: at least in mouse and rat there may be two alternative *Foxl2* transcripts. Several ESTs, carrying a standard polyadenylation signal (AAUAAA), were found to end as expected from the comparison with the human sequence (for the mouse BG071987, AU046041, AU045768, AU045128, BB334136, for the rat AA850108). However we also detected other ESTs that run beyond this point in both rodents. In mouse ESTs BB556269, AW558663, BG076281, BB441670 are representative of this "longer" transcript, carrying a putative polyadenylation signal (AATATAAA) while for the rat ESTs, BU758851 and AI556368 would carry another non-canonic AAGAAT signal. These potential polyadenylation sites are located about 430 bp downstream of the known one. The perfect alignment of these ESTs with the genomic sequences suggests that these alternative transcripts would result from a differential polyadenylation.

To explore this observation experimentally, we performed RT-PCR experiments using primers specific for this putative alternative transcript: mFoxl2polyAde-F (5′ GAACTAGAGCACTTTTGTTGT 3′) located 59 bp upstream of the first polyadenylation signal and mFoxl2polyAde-R (5′ CACCTAGTACAATTATGTAAGAG 3′) located 350 bp downstream. These experiments were conducted on cDNA from mouse gonads at different developmental stages (normalized with respect to β actin, as in Pannetier et al., 2003, in this issue) and on a mouse granulosa cell line (AT29C-U493, described by Dutertre et al., 2001). The *Foxl2* second transcript was detected in female gonads, from 15.5 days post coitum (dpc) with an increasing level and persisting at adulthood. In male gonads, very faint levels were observed (Fig. 3A). The results obtained for the *Foxl2* ORF amplification were very similar, with the exception that it could be faintly detected in female gonads at 12 dpc (Fig. 3A, in line with the results obtained by Pannetier et al., 2003). This correlation between the profiles of the ORF and the second transcript suggests that the latter encompasses the *Foxl2* ORF. This is further suggested by the existence of an amplicon running from the beginning of the *Foxl2* ORF to the end of the second transcript, obtained with adult mouse ovary cDNA (Fig. 3B). This implies that the *Foxl2* ORF amplifications represent, in fact, both transcripts (Fig. 3A).

The experiments described previously have shown the presence of *Foxl2* transcript and protein in the follicular cells (i.e. granulosa) of adult mouse ovaries. Interestingly, by RT-PCR we have not detected the second transcript in the mouse granulosa cell line AT29C, while the first one was detected (Fig. 3A). Either this may be a specificity of this cell line or may be due to the fact that the second transcript is not expressed in the granulosa but in another ovarian compartment. Further studies are required to evaluate the potential existence of other alternative transcripts. For instance, rat EST AW920539 would represent an even shorter transcript, if artifactual RT-priming is ruled out (since there is a polyadenylate stretch immediately downstream of the AAUAAA site). Indeed this polyadenylation site (~770 bp downstream of the stop codon) is the optimum recognized by the Genscan program in human, rat and mouse. ESTs BG088806 and BF544943, among others, seem to start downstream of the *Foxl2* stop codon (more details in a manuscript in preparation). It would be interesting to assess whether the existence of alternative transcripts has a functional significance or not.

Besides, although a similar region exists in human, there is no EST-based evidence of the existence of a longer human transcript. This is consistent with Northern blot experiments on human ovary RNA where only one FOXL2 transcript has been detected (Crisponi et al., 2001). However, the presence/absence of an alternative transcript in human and other species deserves more in-depth studies. This might be relevant to understanding FOXL2 function. Further studies in animal models, such as a conditional knock-out mouse model for *Foxl2* might allow us to shed light on potential species-specific differences, and also on the implication of FOXL2 in early ovarian development and later during adult female life.

Acknowledgements

The authors wish to thank Francis Jaubert (Hôpital Necker, Paris) for help with immunohistochemistry, Eric Pailhoux and Corinne Cotinot (INRA, Jouy en Josas) for providing the mouse ovary cDNA, Sandrine Caburet (Hôpital Cochin, Paris) for reading the manuscript and Jenny Marshal-Graves for the tammar wallaby DNA. The nomenclature used is based on the guidelines found at the Web site http://www.biology.pomona.edu/fox.html.

References

Barishak YR: Embryology of the eye and its adnexae. Dev Ophthalmol 24:1–142 (1992).

Carlsson P, Mahlapuu M: Forkhead transcription factors: key players in development and metabolism. Dev Biol 250:1–23 (2002).

Cocquet J, Pailhoux E, Jaubert F, Servel N, Xia X, Pannetier M, De Baere E, Messiaen L, Cotinot C, Fellous M, Veitia RA: Evolution and expression of FOXL2. J med Genet 39:916–921 (2002).

Cocquet J, De Baere E, Caburet S, Veitia RA: Of compositional biases and polyalanine runs in man. Genetics, in press (2003).

Crisponi L, Deiana M, Loi A, Chiappe F, Uda M, Amati P, Bisceglia L, Zelante L, Nagaraja R, Porcu S, Ristaldi MS, Marzella R, Rocchi M, Nicolino M, Lienhardt-Roussie A, Nivelon A, Verloes A, Schlessinger D, Gasparini P, Bonneau D, Cao A, Pilia G: The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. Nature Genet 27:159–166 (2001).

De Baere E, Dixon MJ, Small KW, Jabs EW, Leroy BP, Devriendt K, Gillerot Y, Mortier G, Meire F, Van Maldergem L, Courtens W, Hjalgrim H, Huang S, Liebaers I, Van Regemorter N, Touraine P, Praphanphoj V, Verloes A, Udar N, Yellore V, Chalukya M, Yelchits S, De Paepe A, Kuttenn F, Fellous M, Veitia R, Messiaen L: Spectrum of FOXL2 gene mutations in blepharophimosis-ptosis-epicanthus inversus (BPES) families demonstrates a genotype-phenotype correlation. Hum molec Genet 10: 1591–1600 (2001).

De Baere E, Lemercier B, Christin-Maitre S, Durval D, Messiaen L, Goderis N, Fellous M, Veitia: FOXL2 mutation screening in a large panel of POF patients and XX males. J med Genet 39:e43 (2002).

De Baere E, Beysen D, Oley C, Lorenz B, Cocquet J, De Sutter P, Devriendt K, Dixon M, Fellous M, Fryns JP, Garza A, Jonsrud C, Koivisto PA, Krause A, Leroy BP, Meire F, Plomp A, Van Maldergem L, De Paepe A, Veitia R, Messiaen L: FOXL2 and BPES: mutational hotspots, phenotypic variability, and revision of the genotype-phenotype correlation. Am J hum Genet 72:478–487 (2003).

Dollfus H, Stoetzel C, Riehm S, Lahlou Boukoffa W, Bediard Boulaneb F, Quillet R, Abu-Eid M, Speeg-Schatz C, Francfort JJ, Flament J, Veillon F, Perrin-Schmitt F: Sporadic and familial blepharophimosis-ptosis-epicanthus inversus syndrome: FOXL2 mutation screen and MRI study of the superior levator eyelid muscle. Clin Genet 63:117–120 (2003).

Dutertre M, Gouedard L, Xavier F, Long WQ, di Clemente N, Picard JY, Rey R: Ovarian granulosa cell tumors express a functional membrane receptor for anti-Mullerian hormone in transgenic mice. Endocrinology 142:4040–4046 (2001).

Harris SE, Chand AL, Winship IM, Gersak K, Aittomaki K, Shelling AN: Identification of novel mutations in FOXL2 associated with premature ovarian failure. Mol Hum Reprod 8:729–733 (2002).

Kaufmann E, Knöchel W: Five years on the wings of fork head. Mech Dev 57:3–20 (1996).

Li WH: Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J molec Evol 36:96–99 (1993).

Loffler KA, Zarkower D, Koopman P: Etiology of ovarian failure in blepharophimosis-ptosis-epicanthus-inversus syndrome (BPES): FOXL2 is a conserved, early-acting gene in vertebrate ovarian development. Endocrinology 144:3237–3243 (2003).

Nakachi Y, Hayakawa T, Oota H, Sumiyama K, Wang L, Ueda S: Nucleotide compositional constraints on genomes generate alanine-, glycine-, and proline-rich structures in transcription factors. Mol Biol Evol 14:1042–1049 (1997).

Oley C, Baraitser M: Blepharophimosis, ptosis, epicanthus inversus syndrome (BPES syndrome). J med Genet 25:47–51 (1988).

Pailhoux E, Vigier B, Chaffaux S, Servel N, Taourit S, Furet JP, Fellous M, Grosclaude F, Cribiu EP, Cotinot C, Vaiman D: An 11.7-kb deletion triggers intersexuality and polledness in goats. Nature Genet 29:453–458 (2001).

Pannetier M, Servel N, Cocquet J, Besnard N, Cotinot C, Pailhoux E: Expression studies of the PIS-regulated genes suggest different mechanisms of sex-determination within mammals. Cytogenet Genome Res 101:199–205 (2003).

Treier M, Gleiberman AS, O'Connell SM, Szeto DP, McMahon JA, McMahon AP, Rosenfeld MG: Multistep signaling requirements for pituitary organogenesis in vivo. Genes Dev 12:1691–1704 (1998).

Xia X: Data analysis in molecular biology and evolution. Second printing, pp 188–191 (Kluwer Academic Publishers, Boston 2001).

Yang Z, Bielawski JP: Statistical methods for detecting molecular adaptation. Trends Ecol Evol 15:496–503 (2000).

Yang Z: Phylogenetic analysis by maximum likelihood (PAML). Version 3.12. University College, London (2002).

Zlotogora J, Sagi M, Cohen T: The blepharophimosis, ptosis, and epicanthus inversus syndrome: delineation of two types. Am J hum Genet 35:1020–1027 (1983).