Differential expression of arid5b isoforms in Xenopus laevis pronephros
Ronan Le Bouffant, Anne-Claire Cunin, Isabelle Buisson, Jérôme Cartry, Jean-François Riou, Muriel Umbhauer

To cite this version:
Developmental Expression Pattern

Title: Differential expression of *arid5b* isoforms in *Xenopus laevis* pronephros

Authors: Ronan Le Bouffant, Anne-Claire Cunin, Isabelle Buisson, Jérôme Cartry, Jean-François Riou and Muriel Umbhauer.

Institutional affiliations:
Sorbonne Universités, UPMC Univ Paris 06, UMR7622 Developmental Biology, F-75005 Paris, France.
CNRS, UMR7622 Developmental Biology, F-75005 Paris, France.

Short running title: *arid5b* expression in *Xenopus* embryo

Key words: *arid5b*, *Xenopus*, pronephros

Abbreviations:
RA: retinoic acid
ARID: A-T Rich Interaction Domain
aa: amino acid
nt: nucleotide

Corresponding author: Muriel Umbhauer
Laboratoire de Biologie du Développement UMR7622, CNRS, UPMC, 9 quai Saint-Bernard, case 24, 75005, Paris France.
muriel.umbhauer@upmc.fr
phone: (33) 1 44 27 39 18 : fax: (33) 1 44 27 34 45

Author’s E-mail addresses:
Ronan Le Bouffant: ronan.le-bouffant@svn.jussieu.fr
Anne-Claire Cunin: aclairecunin@gmail.com
Isabelle Buisson: isabelle.buisson@upmc.fr
Jérôme Cartry: jeromecartry@hotmail.com
Jean-François Riou: jean-francois.riou@upmc.fr
Muriel Umbhauer: muriel.umbhauer@upmc.fr

GenBank accession numbers: HG518326, HG518327
Developmental Expression Pattern

Abstract
Arid5b belongs to the ARID family of transcription factors characterised by a helix-turn-helix motif-based DNA-binding domain called ARID (A-T Rich Interaction Domain). In human, alternative splicing leads to a long and a short isoforms (isoform1 and 2, respectively) that differ in their N-terminal part. In this study, we report the cloning and expression pattern of *Xenopus laevis* arid5b. We have isolated a full length cDNA that shows homology with the human arid5b isoform1. Furthermore, 5'RACE experiments revealed the presence of a shorter isoform equivalent to the human isoform2. Temporal expression analysis by RT-qPCR indicated that X. laevis arid5b isoform1 and isoform2 are differentially expressed during development. Isoform1 is strongly expressed maternally while isoform2 expression is essentially restricted to tailbud stages. Spatial expression analysis by whole mount *in situ* showed that arid5b is predominantly expressed in the developing pronephros. Arid5b mRNAs are detected in the antero-dorsal part of the pronephros anlage at the early tailbud stage and later on, in the proximal part of the pronephric tubule. RT-qPCR analyses with primers that allow to discriminate isoform1 from isoform2 showed that the latter is enriched in the pronephros anlage. In agreement with a specific pronephric signature of the isoform2, we also observed that isoform2 but not isoform1 is upregulated in animal caps induced to form pronephric tissue in response to activin A and retinoic acid. These results indicate that the two arid5b isoforms are differentially expressed and likely play different roles during early *Xenopus* development.

Introduction
The ARID (A-T Rich Interaction Domain) is a helix–turn–helix motif-based DNA-binding domain, conserved in eukaryotes, that defines the ARID family of transcription factors. The human ARID family can be divided into seven subfamilies (ARID1, ARID2, ARID3, ARID4, ARID5, JARID1 and JARID2) based both on degree of homology within the ARID domain, as well as similarity between highly variable non-ARID domain structures. The founding members, murine Bright (ARID3A) and Drosophila dead ringer (Dri), were independently cloned on the basis of their ability to selectively bind to AT-rich DNA sequences but this behaviour is not an intrinsic feature of the ARID domain since the majority of ARID subfamilies bind DNA without obvious sequence preference (Patsialou et al., 2005). ARID-encoding genes are involved in a variety of biological processes including regulation of cell cycle, gene expression, differentiation, embryonic development, transcriptional regulation and chromatin-remodeling. The ARID protein Osa has been shown to associate with the SWI/SNF
Developmental Expression Pattern

complex in both Drosophila and humans, suggesting that the function of ARID proteins in chromatin
remodelling has been conserved through evolution (Collins et al., 1999, Kozmik et al., 2001).

Arid5b, also called Desrt or MRF-2 (modulator recognition factor-2), was originally cloned thanks to its ability
to bind sequences in the transcriptional modulator of the human cytomegalovirus major immediate-early
promoter (Lubon et al., 1989). Its ARID domain, whose three-dimensional structure has been solved, binds
preferentially to an AT-rich core sequence (Whitson et al., 1999). Arid5b has been found to associate with the
jmjC demethylase PHF2. Assembly of the PHF2–Arid5b complex, its recruitment to target promoters, and its
H3H9Me2 demethylase activity are dependent on protein kinase A activity (Baba et al., 2011). During mouse
organogenesis, arid5b displays a complex and highly dynamic pattern of expression. It is first expressed in the
intermediate mesoderm and subsequently in the nephrogenic cords of the urogenital ridges. Arid5b is also
detected in the limbs, the myotomes, the oro-naso-pharyngeal ectoderm and the underlying mesenchyme,
the otic vesicles, the gut and its derivatives, and transiently in the liver. Arid5b mutant mice generated by gene
targeting have reduced viability, pronounced growth retardation, and a high incidence of abnormalities of the
reproductive organs (Lahoud et al., 2001). They also show significant reductions in lipid accumulation and
weight gain in postnatal and adult life (Whitson et al., 2003). Arid5b is required for adipogenesis and to
maintain normal functions in mature adipocytes. Knockdown of Arid5b in mature 3T3-L1-derived adipocytes
activates both lipolysis and triglyceride synthesis, and causes a significant increase in the ratio of glycerol
release to free fatty acid release (Yamakawa et al., 2010, Yamakawa et al., 2008). Arid5b is highly expressed in
the cardiovascular system and is believed to play essential roles in smooth muscle cell differentiation and
proliferation (Watanabe et al., 2002). In homozygous arid5b mutant mice, kidneys are small showing often
degraded glomeruli with defects in smooth muscle cell number and location. Skeletal abnormalities, including
defects in the patterning of the ribs and sternum, have also been described (Schmahl et al., 2007).
Although the temporal and spatial pattern of expression of arid5b during embryogenesis have been described
in mouse (Ristevski et al., 2001), there is no detailed expression data available for non mammalian vertebrates.
We report the cloning of two arid5b isoforms in Xenopus laevis and have examined their expression patterns
during development.

Results and discussion

Molecular cloning of X. lævis arid5b

In order to clone a full coding sequence of arid5b in Xenopus, we started from a partial IMAGE clone sequence
(no 686,6480), and obtained the missing 5’ sequence by RACE PCR. A 3570 nucleotides (nt) clone was
amplified by end-to-end PCR (GenBank accession no. HG518326). Sequence analysis revealed an open reading
Developmental Expression Pattern

frame encoding a predicted 1187 amino acids (aa) protein (Fig.S1). This protein displays 87.6% identity with a predicted *X. tropicalis* protein sequence deduced from gene models (accession no. XM_002939542), 57.5% identity (82.2% similarity) with the long human Arid5b isoform1, and 56.7% identity (80.7% similarity) with the murine long isoform α. The ARID domain (aa 324-418) is highly conserved, with more than 90% identity observed with other vertebrate Arid5b sequences (Fig.1). The conserved sequence includes a lysine residue (lys-342) at a position homologous to the lysine of the long murine isoform α (lys-336) whose demethylation by PHF2 promotes recruitment of PHF2-Arid5b complex to promoters (Baba et al., 2011). Blast analysis on the *X. laevis* genome 6.0 scaffolds shows that the long isoform is encoded by two genes located on scaffolds 9729 and 48311, respectively, which probably represent pseudoalleles resulting from *X. laevis* allotetraploidy. In a similar way to human and murine *arid5b*, the long *X. laevis* isoform is encoded by ten different exons (table S1,S2).

We further investigated whether a shorter isoform homologous to the short human *arid5b* isoform was also detectable. Human *arid5b isoform2* is generated by alternative splicing resulting in the replacement of phe-244 by a start methionine. Using 5' RACE-PCR, we were able to clone a partial *X. laevis arid5b* sequence of 368nt containing an ORF encoding a 91 aa polypeptide where phe-244 is replaced by a methionine. The following aa are identical to those of the long isoform (ala-245-leu-334). The 95 nt sequence located upstream to this ATG codon does not contain any other in frame ATG codon, but three stop codons indicating that it probably encodes the 5’UTR of a shorter isoform mRNA. Blast analysis on the *X. laevis* genome revealed that this sequence is encoded by a novel exon (exon 4b, table S1,S2), while ala-245-leu334 are encoded by exons 5 and 6, supporting the idea of an isoform generated by alternative splicing. Nested 3’RACE-PCR was carried out to clone the full sequence encoding this putative short isoform. The first primer was located 61nt upstream of the putative start ATG, and the nested primer 1307nt downstream of this ATG. A 2163nt sequence was cloned containing 1529nt of putative coding sequence, and 634nt of 3’UTR. It is identical to the sequence of the long isoform, in line with human *arid5b isoform2* sequence data, which only differs from isoform1 at the start ATG. Using different sets of primers, we have then tried to amplify the entire short isoform by end-to-end PCR but failed to amplify a full cDNA. Using forward primers corresponding to exon 4b sequence, and reverse primers at different levels of the sequence obtained by 3’ RACE-PCR, we could only amplify a partial cDNAs encoding the first 476 aa of the short isoform. This cDNA corresponds to exons 4b, 5-9 and part of exon 10 (accession number HG518327). Whether the short isoform is ending at the same stop codon as the long one therefore remains unclear. Nonetheless, a 3’ sequence for the short isoform mRNA distinct from that of the long isoform would imply an alternative splicing within exon10, that is not occurring with human *arid5b isoform2*. Together, these data show that the short isoform1 lacks the first 243 N-terminal amino acids of the long isoform2 (Fig.S2). Using InterProSCan software and performing an extensive analysis of the literature on Arid5 family members, we could not find any known domain in this region, precluding to identify any functional difference
Developmental Expression Pattern

99 between the two isoforms.

Genomic synteny and molecular phylogeny

In order to further confirm the identity of the *Xenopus* ortholog of *arid5b*, we have examined the synteny maps from human, mouse and *X. tropicalis* genomes. Synteny maps were obtained from Ensembl genome browser (release 74, December 2013) and JGI genome browser (*X. tropicalis* version 7.1). Synteny maps are shown in Fig.1B. Flanking genes are partially conserved between *arid5b* genes in human, mouse and *X. tropicalis* genomes. The conserved syntenic region flanks the 3’ end of *arid5b*. It comprises *rkn2*, *znf365*, *ado* and *egr2* genes. Genes flanking the 5’ end of *arid5b* in *X. tropicalis* genome totally differ from those flanking human or mouse *arid5b*. Phylogenetic analysis indicates that *Xenopus arid5b* is relatively distant from the zebrafish ortholog. Tetrapod orthologs appear to be more closely related, with the chicken ortholog being the most closely related to *Xenopus arid5b* (Fig.1C).

Spatial and temporal expression of *arid5b* during *X. laevis* development

Temporal expression of *ari5b* during early *X. laevis* development was examined by RT-qPCR. Specific primers were designed in order to amplify either both *arid5b* isoforms (total *arid5b*), *isoform1* or *isoform2* (Fig. S3). *Isoform1* was strongly expressed during cleavage stages; its expression declined during gastrulation and neurulation, then it increased at tailbud stages (fig.2B). In contrast, *isoform2* transcripts were scarcely detected at the pre-tailbud stages, but became detectable during organogenesis and persisted at least up to the late tailbud stage (stage 28) (fig.2C). Notably, the temporal profile of *isoform1* is almost identical to the profile obtained for total *arid5b*, suggesting that *isoform1* is the main isoform to be expressed during embryonic development (fig.2A,B). In agreement with this idea, when normalized to total *arid5b*, relative expression of *isoform1* was constant during embryonic development while relative mRNA expression of *isoform2* increased during tailbud stages (fig.2D,E). Thus, *isoform1* and *isoform2* are differentially expressed during development. *Isoform1* is strongly expressed maternally and is the main isoform expressed during embryonic development. *Isoform2* expression is essentially restricted to tailbud stages.

Next, we studied the spatial expression pattern of *arid5b* by whole mount *in situ* hybridization using two probes: the first encompassing nucleotide 775 to 1612 of *isoform1* (which corresponds to nucleotides 46 to 853 of *isoform2*); the second corresponding to nucleotide 529 to 3564 of *isoform1*. Both probes gave the same expression pattern (fig.4 and data not shown). At cleavage, gastrula and neurula stages, embryos were uniformly stained indicating that *arid5b* mRNAs were ubiquitously distributed (not shown). RT-qPCR analysis of dissected explants from early gastrula confirmed this observation and further showed that none of the two isoforms showed a regionalized expression (fig.3). At the early tailbud stage, a specific signal was detected in
the pronephric region by in situ hybridization (fig.4A). At stage 25, arid5b was strongly expressed in the antero-
dorsal part of the pronephric anlage. In comparison with delta1 expression which is restricted to the most antero-dorsal part of the anlage ((Rones et al., 2000) et fig.4C,F), arid5b is expressed in a larger domain that extends more ventrally and medially (fig.4B,E). Expression of arid5b remained restricted to the anterior part of the developing pronephros at tailbud stages (fig.4D). At tadpole stage 35/36, arid5b mRNAs were localized in the proximal part of the pronephric tubule (fig.4H,J,I,L). Arid5b expression domain is adjoining to that of scl12a1 which marks the intermediate and the first distal segments (Raciti et al., 2008) (fig.4H,I,L). Thus, from the early tailbud stage onward, arid5b expression is restricted to the developing proximal part of the pronephros. Since our in situ hybridization experiments do not allow to distinguish between the two isoforms, we performed RT-qPCR to analyse the expression levels of each isoform in the pronephric anlage in comparison to the whole embryo and to different embryonic regions. The results clearly showed that isoform2 is more strongly expressed in the pronephric anlage than in the other tissues at tailbud stage (fig.5). No such enrichment was observed for isoform1 although it is slightly more expressed in the head in comparison to other tissues. As expected, the kidney marker pax8 was found to be strongly expressed in the pronephric anlage (fig.5). It has been previously shown that treatment of blastula animal cap ectoderm with activin A and retinoic acid (RA) results in the formation of pronephric tubules at high frequency (Ariizumi and Asashima, 2001). We studied whether this treatment could upregulate arid5b isoforms expression. We showed that isoform2 but not isoform1 expression is upregulated in response to activin A and RA (fig.6). This result is in agreement with a specific pronephric signature of the isoform2.

In summary, we have cloned the X. laevis ortholog of arid5b, identified two isoforms and examined their expression pattern during embryonic development. The two isoforms are differentially expressed: isoform1 is strongly expressed maternally, while isoform2 is specifically expressed in the pronephric anlage at tailbud stage. These results indicate that the two arid5b isoforms likely play different roles during early Xenopus development.

Materials and Methods

Molecular cloning and Bioinformatic analyses

The partial IMAGE clone 696,6480 was obtained from RZPD ImaGenes. RACE-PCR was performed with the SMARTer™ RACE cDNA amplification kit (Clontech). End-to-end PCR was carried out according to manufacturer instructions with Advantage 2 polymerase (Clontech), and stage 35/36 embryo cDNA prepared as described (Le Bouffant et al., 2012). Amino acid sequence comparison were performed with MultAlin software (http://multalin.toulouse.inra.fr/multalin/) (Corpet, 1988) and CLUSTAL W (version 1.83).
Developmental Expression Pattern

Embryos, explants dissection, animal cap assay

*Xenopus laevis* were purchased from the CNRS Xenopus breeding Center (Rennes, France). Embryos were obtained after artificial fertilization, and were raised in modified Barth’s solution (MBS). Stages were according to the normal table of *Xenopus laevis* (Nieuwkoop and Faber, 1967). Dissections were all performed in 1X MBS on 1% agar-coated dishes. Presumptive ectoderms (animal cap) were isolated from blastula stage embryos (stages 8–9) and immediately transferred into 1X MBS, 0.1% BSA in the presence of recombinant human activin A (10 ng/mL, Sigma) and all-trans retinoic acid (10^{-4} M, Sigma) or DMSO alone (1/250). The animal caps were incubated for 3 hrs, after which they were washed twice in 1X MBS and further cultured for 48 hrs.

Explants comprising the pronephric anlage were dissected from early tailbud embryos (stage 21). Using platinum loop and wire, stage 25 embryos were dissected into several pieces: the head, the tail, the truncal dorsal part (essentially somites, neural tube and notochord) and the truncal ventral part (mainly endoderm, ventral and lateral mesoderm). Pronephric anlagen were isolated from somitic and lateral plate mesoderm and separated from the underlying endoderm. The overlying ectoderm was kept. Explants were immediately processed for RT-qPCR.

Real-time quantitative PCR

RT-PCR analyses were carried out as previously reported in (Le Bouffant et al., 2012). Sequences of oligonucleotides used are the following: *arid5b*: forward: 5’TATGTTTCAAGCTGCGCAAAA3’, reverse: 5’CCATTGCCCTCGTGCAATA3’; *arid5b isoform1*: forward: 5’CCCAGAAGATACCCCCAAGG3’, reverse: 5’ACTTCATGCTCTCCGTGGCT3’; *arid5b isoform2*: forward: 5’TGCTCTGTGGCCTCATGAG3’ reverse: 5’TGGCTCTGTGGCCTCATGAG3’; *pax8*: forward: 5’CAGCAATTATATAGGTACGG3’, reverse: 5’TGGGAAAGGAGCTTTGAACG3’; *ODC*: forward: 5’GGGCAAAGGAGCTTAATGG3’, reverse: 5’TGCCAACATGGAAACTCAG3’. The Comparative Ct method was used to determine the relative quantities of mRNA, using *ODC* mRNA as the endogenous reporter except for figure 2D,E for which *arid5b* was used. Same results were obtained using *β-actin* mRNA as the endogenous reporter instead of *ODC* (data not shown). Each RNA sample was analysed in duplicate. Each data point represents the mean ± SEM of at least three independent experiments. Data were analysed using R Commander (R software) by paired Student’s t-test.

In situ hybridization

Whole mount in situ hybridization for *arid5b*, *delta1* (Rones et al., 2000), and *slc12a1* (Raciti et al., 2008) were carried out as previously reported (Cartry et al., 2006). The antisense and control sense RNA probes for *arid5b* were generated from linearized plasmids containing cDNA sequences from nucleotide 529 to 3564 of
Developmental Expression Pattern

isoform1 and from nucleotide 775 to 1612 of isoform1 (which corresponds to nucleotides 46 to 853 of isoform2). The arid5b RNA probes were subjected to limited alkaline hydrolysis in two volumes of carbonate buffer (60 mM Na₂CO₃, 40 mM NaHCO₃, pH 10.2) for 5 min at 60°C to reduce its size and increase its access to tissues. The hydrolysis was terminated by adding an equal volume of neutralizing solution (1 M Tris-HCl, pH 8.0, containing 1.5 M NaCl). Hydrolyzed fragments were precipitated with ethanol.

Acknowledgements:
We thank S. Autier and E. Manzoni for animal care, E. Jones and E. Bellefroid for plasmids. This work was supported by grants from CNRS and from University Pierre et Marie Curie. We acknowledge funding from Emergence-UPMC-2009 research program.

Figure Legends

Fig. 1 : Characterization of a X. laevis ortholog of arid5b.

(A) Predicted ARID domain amino acid sequence comparison: G. gallus (ac Q5ZJ69); H. sapiens (ac Q14865); X. tropicalis (F6QQ73); D. rerio (E7F888). (B) Synteny blocks containing arid5b genes in H. sapiens; M. musculus and X. tropicalis genomes. Genes organization in the human arid5b gene region was used as basis for comparison. Chromosomal localization is indicated. The scaffold number is given for X. tropicalis. Relative spacing between the genes is not shown. (C) Phylogenetic tree of arid5b genes from various vertebrate species constructed using the neighbour-joining method. Accession numbers used are shown.

Fig. 2 : Temporal expression of arid5b during Xenopus embryonic development

Expression of arid5b analysed by RT-qPCR at cleavage (stages 5 and 9), gastrula (stage 11), neurula (stages 14) and tailbud stages (stages 22 and 28). Primers were designed in order to amplified either both isoforms (total arid5b) (A), isoform1 (B,D) or isoform2 (C,E). The relative quantities of mRNA were determined with ODC (A,B,C) or total arid5b (D,E) mRNA as the endogenous reporter. Isoform1 is strongly maternally expressed; isoform2 is mainly expressed during tailbud stages. Average values from three independent experiments.

Fig. 3 : Expression of arid5b at the early gastrula stage

RT-qPCR analyses for total arid5b, isoform1 and isoform2 were performed on dissected explants from early gastrula stage embryo (stage10.5). Embryos were dissected either into ventral and dorsal halves or into dorsal
marginal zone (DMZ), ventral marginal zone (VMZ), lateral marginal zone (LMZ) ectoderm and endoderm. Both isoforms are ubiquitously expressed at the early gastrula stage. Average values from three independent experiments.

**Fig. 4** : Spatial expression of *arid5b* during *Xenopus* development

*In situ* hybridization of whole (A-D, G-L) or transverse fractured embryos (E, F) at the indicated stages of development with antisense probe for *arid5b* (nucleotide 775 to 1612 of *isoform1*) (A,B,D,E,H,J,L), *delta1* (C,F), *slc12a1* (I,K,L) and with control sense probe for *arid5b* (G). Lateral views with anterior to the right (A-D,G,H, J-L). Transverse section at the level of the proximal pronephric tubule. In I and L, *arid5b* is revealed in light blue and *slc12a1* in purple. *Arid5b* mRNAs are detected in the anterior pronephric anlage during tailbud stages in a broader domain than *delta1*. At tadpole stage 35/36, *arid5b* expression is restricted to the proximal part of the tubule and does not overlap with *slc12a1* expression which is specific for the intermediate and the first segment of the distal tubule. Scale bars are 0.3 mm.

**Fig. 5** : *Arid5b isoform2* is specifically expressed in the pronephros

Expression of *isoform1*, *isoform2* and *pax8* analysed by RT-qPCR in different embryonic regions (see materials and methods) dissected at tailbud stage 25. *Isoform2* as well as *pax8* are strongly expressed in pronephric anlage in comparison to other embryonic tissues (statistically significant for all). *Isoform1* is slightly more expressed in the head in comparison to other tissues. Average values from three independent experiments. *P*<0.05

**Fig. 6** : *Arid5b isoform2* is specifically induced in animal caps treated with activin A and RA

RT-qPCR analysis of *arid5b isoform1* and *isoform2*, as well as *pax8* expression in induced blastula animal caps. Animal caps were dissected at blastula stage 9, incubated for 3 hours in presence of RA and activin A, or mock solution for the control group. Animal caps were further cultured in 1X MBS for 48 hours and processed for RT-qPCR. A significant increase of *isoform2* expression is observed in response to RA and activin A. Average values from three independent experiments. *P*<0.05, **P** < 0.005

**Legends to Supplementary materials**

**Fig. S1** : Comparison of *Arid5b isoform1* amino acid sequence between vertebrates

Predicted amino acid sequence comparison of *Arid5b*: *X. tropicalis* (F6QQ73) isoform1; *G. gallus* (Q5ZJ69); *H. sapiens* (Q14865); *M. musculus* (Q8BM75); *D. rerio* (E7F888). Red boxes indicate amino acid residues
Developmental Expression Pattern

conserved in all species. The green arrow indicates phe-244 of isoform1 that is replaced by a start methionine in isoform2.

Fig. S2: Amino acid comparison of the two Arid5b isoforms. The ARID domain is highlighted in yellow. Regions specific to isoform1 or isoform2 are indicated in red letters. The sequence in blue corresponds to the predicted C-terminal sequence of isoform2 observed in 3’RACE PCR experiment but that we could not amplify by end-to-end PCR.

Fig. S3: Nucleotide position of PCR primers used for RT-qPCR. Nucleotide sequences of the N-terminal region of arid5b isoform1 and isoform2 are aligned. Arrows indicate the primer sequences that were used to amplify both isoforms (in blue), isoform1 (green) and isoform2 (brown). Conserved nucleotides are in red.

Table S1: Exon-intron organization of the arid5b gene in Xenopus laevis on scaffold 9729. Exon sequences are indicated by uppercase letters and intron sequences by lowercase letters. Splice donor and acceptor sites are underlined. Exon and intron size are reported as base pairs.

Table S2: Exon-intron organization of the arid5b gene in Xenopus laevis on scaffold 48311. Exon sequences are indicated by uppercase letters and intron sequences by lowercase letters. Splice donor and acceptor sites are underlined. Exon and intron size are reported as base pairs.
References


Developmental Expression Pattern

Figure 1
Developmental Expression Pattern

Figure 2

- **A** total arid5b (normalized to ODC)
- **B** arid5b isoform1 (normalized to ODC)
- **C** arid5b isoform2 (normalized to ODC)
- **D** arid5b isoform1 (normalized to arid5b)
- **E** arid5b isoform2 (normalized to arid5b)
Developmental Expression Pattern

**figure 3**

- **total arid5b**

- **arid5b isoform1**

- **arid5b isoform2**
Figure 4
Developmental Expression Pattern

**Figure 5**

*arid5b isoform 1*

![Expression levels of arid5b isoform 1](image)

*arid5b isoform 2*

![Expression levels of arid5b isoform 2](image)

*pax8*

![Expression levels of pax8](image)
Developmental Expression Pattern

*arid5b isoform 1*  
*arid5b isoform 2*  
*pax8*

![Graphs showing mRNA relative expression levels for *arid5b isoform 1*, *arid5b isoform 2*, and *pax8* under control and activin + RA conditions.](image)

Figure 6
Developmental Expression Pattern

Figure S2
### Developmental Expression Pattern

**Figure S3**

<table>
<thead>
<tr>
<th>Gene 1</th>
<th>Gene 2</th>
<th>Gene 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
<tr>
<td>aridSb</td>
<td>aridSbisoform2</td>
<td>aridSbisoform1</td>
</tr>
</tbody>
</table>

**Legend:**
- Blue arrow: aridSb
- Red arrow: aridSbisoform2
- Green arrow: aridSbisoform1
### Table S1

<table>
<thead>
<tr>
<th>Exon n°</th>
<th>Exon size (bp)</th>
<th>5' splice donor</th>
<th>3' splice acceptor</th>
<th>Intron size (bp)</th>
<th>Amino acid interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&gt;27</td>
<td>ACTCAAGgttatttagctg</td>
<td>tcttgcagTGGGTCGG</td>
<td>328</td>
<td>Trp-8</td>
</tr>
<tr>
<td>2*</td>
<td>&gt;68</td>
<td>?</td>
<td>attttcagCATGAAGT</td>
<td>&gt;39437</td>
<td>Glu-92?</td>
</tr>
<tr>
<td>3</td>
<td>226</td>
<td>GAAATTAGgttaacctggttc</td>
<td>aaccacagTGAAGGCG</td>
<td>40399</td>
<td>His-93</td>
</tr>
<tr>
<td>4</td>
<td>231</td>
<td>AATTGGGgtatgctcaaccccttttt</td>
<td>tttaacagCACCAGATT</td>
<td>66835</td>
<td>Gly-168</td>
</tr>
<tr>
<td>5</td>
<td>113</td>
<td>GGCACAGgtaatccttttttt</td>
<td>tcttttcagCTCAATTG</td>
<td>1316</td>
<td>Lys-282</td>
</tr>
<tr>
<td>6</td>
<td>220</td>
<td>AAGCAGAGtaagtagacagcttccoacag</td>
<td>TCACAATTCTTTT</td>
<td>4458</td>
<td>Ile-356</td>
</tr>
<tr>
<td>7</td>
<td>53</td>
<td>TGAACAGgtatgctttta</td>
<td>tataacagATTACTGC</td>
<td>3545</td>
<td>Ile-374</td>
</tr>
<tr>
<td>8</td>
<td>98</td>
<td>ACGAAAGgtagataatctca</td>
<td>ccccaagATTAAACTCC</td>
<td>33496</td>
<td>Arg-406</td>
</tr>
<tr>
<td>9</td>
<td>199</td>
<td>GACTGAGgttaaatttgggga</td>
<td>ttgtttagCTCCACC</td>
<td>5264</td>
<td>Val473</td>
</tr>
<tr>
<td>10</td>
<td>&gt;2712</td>
<td>GGAATGGgttaaccgttcag</td>
<td>tcttttcagCTCAATTG</td>
<td>6493</td>
<td>Ala-2 (short isoform)</td>
</tr>
</tbody>
</table>

Scaffold 9729. * sequence nt89-275 mRNA ORF missing in genomic sequence

### Table S2

<table>
<thead>
<tr>
<th>Exon n°</th>
<th>Exon size (bp)</th>
<th>5' splice donor</th>
<th>3' splice acceptor</th>
<th>Intron size (bp)</th>
<th>Amino acid interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>&gt;27</td>
<td>ACTCAAGgttatttagctg</td>
<td>?</td>
<td>?</td>
<td>Trp-8</td>
</tr>
<tr>
<td>2</td>
<td>236</td>
<td>CCGAGAGgttacattttctctac</td>
<td>ttttcttcagCATGAAGT</td>
<td>53173</td>
<td>His-93</td>
</tr>
<tr>
<td>3</td>
<td>226</td>
<td>GAACAGAGgttaaccatgattc</td>
<td>aaccacagTGAAGGCG</td>
<td>53008</td>
<td>Gly-168</td>
</tr>
<tr>
<td>4</td>
<td>231</td>
<td>GAATTTGGgtatgtccaattc</td>
<td>cctaacagTTAACCTT</td>
<td>77336</td>
<td>Ile-356</td>
</tr>
<tr>
<td>5</td>
<td>113</td>
<td>TGCCACAGgttaatgcttttttt</td>
<td>tttttagCTCAATTG</td>
<td>1264</td>
<td>Val-283</td>
</tr>
<tr>
<td>6</td>
<td>220</td>
<td>AAGCAGAGgttaagtagacattc</td>
<td>tctaacagTTAACCTT</td>
<td>4033</td>
<td>Ile-356</td>
</tr>
<tr>
<td>7</td>
<td>53</td>
<td>TGAACAGAGgtatgcttttttt</td>
<td>ttttcttcagCTCAATTG</td>
<td>4690</td>
<td>Ile-374</td>
</tr>
<tr>
<td>8</td>
<td>98</td>
<td>ACGAAAGgtatgcttttttt</td>
<td>cccccacagATTAACTCC</td>
<td>33865</td>
<td>Arg-406</td>
</tr>
<tr>
<td>9</td>
<td>199</td>
<td>GCTGAGAGgttatatttgggga</td>
<td>ttgtttagCTCCACC</td>
<td>4311</td>
<td>Val-473</td>
</tr>
</tbody>
</table>

Scaffold 48311 * sequence nt21-39 mRNA ORF missing in genomic sequence