**Int. J. Dev. Biol. 54**: 1375-1382 (2010)
doi: 10.1387/ijdb.103080cf

THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

# *Vestigial like* gene family expression in *Xenopus:* common and divergent features with other vertebrates

CORINNE FAUCHEUX, FRANÇOIS NAYE, KARINE TRÉGUER, SANDRINE FÉDOU, PIERRE THIÉBAUD[#] and NADINE THÉZÉ*,[#]

*Université de Bordeaux, UMR CNRS 5164, Bordeaux, France*

**ABSTRACT The *Drosophila* Vestigial and Scalloped proteins form heterodimers that control wing development and are involved in muscle differentiation. Four *vestigial like* genes have been described in mammals. Similar to the *Drosophila vestigial* gene, they encode a short conserved domain (TONDU) required for interaction with the mammalian paralogues of *Drosophila* Scalloped (i.e., TEAD proteins). We previously identified two *TEAD* genes in *Xenopus laevis* and we report here the expression of four distinct *vestigial like* genes in *Xenopus (vgll1-4)* that represent amphibian orthologs of the mammalian *vestigial like* genes. *Vgll1* has a unique expression pattern which is restricted to epidermal cells, both in the embryo and in the adult. *Vgll2* is expressed in the skeletal muscle lineage downstream of myogenic factors and in the embryonic brain similar to the avian and mammalian orthologues. *Vgll3* expression is transient, identifies embryonic hindbrain rhombomere 2, and is negatively regulated by en2, but not by egr2. *Vgll4* is mainly expressed in anterior neural structures. In summary, the four *Xenopus vgll* genes have unique/complex expression profiles and they are differently expressed during embryogenesis. Moreover, these amphibian *vestigial like* genes display distinct responses to the major signaling pathways (i.e., activin, FGF or BMP) that orchestrate pattern-formation during early development.**

KEY WORDS*: vestigial like, Xenopus, activin, FGF, BMP*

*Vestigial* (*vg*) belongs to the class of selector genes that were first described in *Drosophila* and whose functions are to govern the fates of groups of cells within embryos (Mann and Carroll, 2002). In *Drosophila*, *vg* is required for wing development but also in the specification/differentiation of embryonic somatic muscle as well as indirect flight muscle development (Bernard *et al*., 2003; Williams *et al*., 1991). Specifically, Vg forms a molecular complex with the protein Scalloped (Sd) to activate numerous target genes to mediate cell fate determination (Halder *et al*., 1998; Simmonds *et al*., 1998). The vertebrate paralogs of Sd belong to the TEA/ATTS domain (TEAD) transcription factor family that includes four genes in mammals (TEAD 1-4) (Kaneko and DePamphilis, 1998). TEAD proteins have emerged in vertebrates as transcription factors implicated in the specific activation of muscle genes through their binding to the M-CAT motif found in the regulatory region of numerous muscle-lineage-specific genes (Larkin and Ordahl, 1999). For example, the *Xenopus* α*-tropomyosin* gene is activated in the three embryonic muscle lineages through a highly conserved M-CAT motif that was shown to bind TEAD proteins (Pasquet *et al*., 2006). The downstream functionality of TEAD proteins depend largely on their interactions with cofactors such as vestigial like proteins. For instance TEAD1 dependent muscle gene activation *in vitro* has been shown to be regulated by Vgll2, one of four vestigial like proteins found in mammals (Maeda *et al*., 2002; Chen *et al*., 2004a; Gunther *et al*., 2004). These four distinct *vestigial like* genes (*Vgll1-4*) similar to the *Drosophila* vg gene encode proteins with a 26 amino acid conserved region (named TDU motif) that is located in the interaction domain between Vg and Sd (Chen *et al*., 2004b; Maeda *et al*., 2002; Mielcarek *et al*., 2002; Mielcarek *et al*., 2009; Vaudin *et al*., 1999).

The four mammalian *vestigial like* genes each display distinct patterns of expression in both embryonic and adult tissues. For example, *VGLL1* and *VGLL3* in humans are mainly expressed in placenta, *VGLL2* in skeletal muscle, whereas *VGLL4* is the only

*Abbreviations used in this paper:* BMP, bone morphogenetic protein; FGF, fibroblast growth factor; vgll, vestigial like.

***Address correspondence to:** Nadine Thézé. UMR CNRS 5164, Université Bordeaux 2, 146 rue Léo Saignat 33076 Bordeaux, France. Fax: +33-557-571-472. e-mail: nadine.theze@u-bordeaux2.fr - web: http://www.umr5164.u-bordeaux2.fr
**#Note:** Both authors share senior authorship.

*Accepted:* 23 June 2010. *Final author corrected PDF published online:* 9 July 2010.

ISSN: Online 1696-3547, Print 0214-6282

member of the gene family to be expressed in heart (Chen *et al.*, 2004b; Maeda *et al.*, 2002; Mielcarek *et al.*, 2002; Vaudin *et al.*, 1999). The developmental expression of *Vgll2* and *Vgll3* genes has also been investigated in mouse. During mouse development, *Vgll2* is expressed in somitic myotome from E8.75 mouse embryos onwards, in branchial arches and later on in skeletal muscle (Maeda *et al.*, 2002; Mielcarek *et al.*, 2002). This gene is also transiently expressed in the ventromedial hypothalamus (Kurrasch *et al.*, 2007). In contrast, the mouse *Vgll3* is expressed in the myogenic lineage and the nervous system of the embryo and in skeletal muscle, heart, kidney, liver and brain (Mielcarek *et al.*, 2009). Two orthologues of the mammalian *Vgll2* have also been identified in zebrafish and were shown to be expressed in the somites (Mann *et al.*, 2007). In addition, a recent report has described the expression of *Vgll2* in the skeletal myogenic lineages of the chicken embryo under the control of the myogenic factors (Bonnet *et al.*, 2009).

Previously, we have characterized *tead1 and tead2* genes in *Xenopus laevis* and defined their developmental pattern of expression and regulation (Naye *et al.*, 2007). In order to more completely understand the receptor-ligand interactions in *Xenopus* with regard to tead and its co-factor vestigial like, we have characterized the *vestigial like* genes in this amphibian and analyzed their expression pattern. We report that, similar to mammals, there are four *Xenopus vestigial like* genes (*vgll1-4)* that are highly conserved within their TDU motif. They are differently expressed in a complex pattern during development and adult tissues showing common but also distinct features with vertebrate orthologues form other species. The genes display distinct responses to the major signaling pathways controlling cell patterning during early development. In particular, we show here that *vgll2* is downstream of myogenic factors and that *vgll3* is repressed by en2 but not egr2.

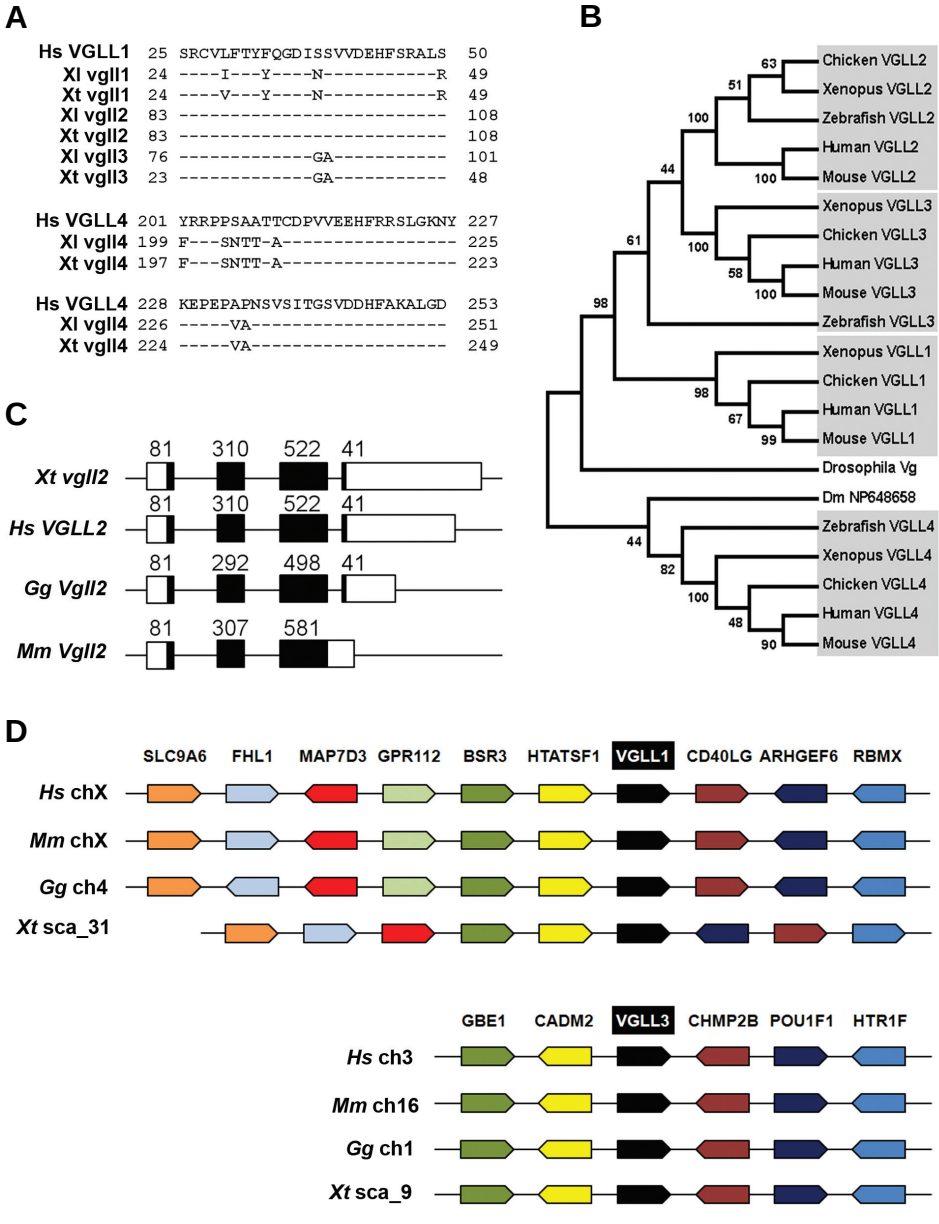


**Fig. 1. Identification of *Xenopus vestigial like* genes by sequence comparison, phylogeny and synteny analysis. (A)** *Sequence comparison between human (Hs) VGLL (VGLL1 and VGLL4) proteins and* Xenopus laevis *(Xl) and* Xenopus tropicalis *(Xt) vgll (vgll1-4) proteins over the TONDU motif. Numbers indicate the position of the relevant sequences relatively to the complete sequences. Dashes indicate identical amino acids between sequences.* **(B)** *Evolutionary relationships between vertebrate VGLL proteins and* Drosophila *Vg proteins as recovered from neighbor-joining tree analysis. The consensus tree was based on amino acid sequence alignment. Numbers below branches show percent bootstrap support for each node.* **(C)** *Schematic representation of* vestigial like 2 *genes from human* (Hs), *mouse* (Mm), *chicken* (Gg) *and* Xenopus tropicalis (Xt). *Coding exons are figured in black boxes and untranslated region in open boxes. Introns are figured as a solid line. Numbers above the sequences correspond to the size in nucleotides of the coding sequences. The genes are not on scale for simplicity.* **(D)** *Conserved syntenic regions between human* (Hs), *mouse* (Mm), *chicken* (Gg) *and* Xenopus tropicalis (Xt) *chromosome regions containing VGLL1 and VGLL3 loci. Genes are represented as colored boxes with the arrow indicating the orientation of the transcription unit. Boxes with the same color correspond to ortholog genes. The drawing is not on scale to avoid complexity.*

## Results

Several cDNA sequences from both *Xenopus laevis* and *Xenopus tropicalis* that were identifiable from available databanks were found to encode proteins containing a TDU motif. These sequences are classified into four groups with each group representing one of the four known mammalian *vestigial like* genes. As a consequence, we have named the corresponding *Xenopus* genes *vgll1*, *vgll2*, *vgll3* and *vgll4* according to the Xenbase nomenclature guide (http://www.xenbase.org/common/). *Vgll2*, *vgll3* and *vgll4* full length cDNA sequences are present in available databases for both *X. laevis* and *X. tropicalis*. However, only a *X. tropicalis vgll1* cDNA sequence was identified (see Experimental Procedures section for accession numbers). In order to characterize the *X. laevis* vgll1 sequence we used the *X. tropicalis* sequence to design oligonucleotides and clone (by RT-PCR from embryo RNA) a complete sequence encoding the *X. laevis* vgll1. The *X. laevis* vgll1 is 228 amino acids long while the *X. tropicalis* is composed of 227 amino acids. The two proteins have 90% of identity and only one amino acid of difference over their TDU motif (data not shown and Fig. 1A). Although the two *Xenopus* proteins show 25 to 28% identity with the avian or the mammalian orthologous proteins, when considering only the TDU domain, the amphibian protein now displays ~85% sequence identity (Fig. 1A).

The *X. laevis* and the *X. tropicalis vgll2* genes encode proteins that are each 317 amino acids long and show only 9 amino acids of difference. Moreover, both proteins have 61 to 64% sequence identity with the avian and mammalian VGLL2 orthologues and 100% sequence identity within the TDU domain (data not shown and Fig. 1A).

The *X. laevis* vgll3 protein is 295 amino acids long while the *X. tropicalis* protein is 244 amino acids long. The two proteins display 91% sequence identity over their common region (data not shown). In addition, the overall identity between the *X. laevis* and *X. tropicalis* vgll3 and the avian or mammalian proteins ranges from 61 to 63%, respectively. This sequence identity increases to 92% when comparing the TDU domains of the two *Xenopus* vgll3 proteins and the human VGLL3 (Fig. 1A).

The *X. laevis* vgll4 and the *X. tropicalis* vgll4 are 293 and 291 amino acids long, respectively, and display 92% sequence identity (data not shown). The overall identity of the *X. laevis* and *X. tropicalis* vgll4 *Xenopus* proteins compared with the avian or mammalian protein ranges from 71 to 82% and >85% when considering only the TDU motifs of these proteins (Fig. 1A).

A phylogenetic tree analysis was constructed using either Neighbor-joining (Fig. 1B) or Maximum-likelihood (data not shown). In both cases, the vertebrates VGLL genes constitute four distinct groups where VGLL1-2-3 groups have evolved from the vestigial *Drosophila* gene while VGLL4 has evolved from the *Drosophila* related vestigial gene. When comparing the structure of the *vestigial like* genes between known vertebrates, only *VGLL2* has retained an evolutionary conserved structure (Fig. 1C). For example, while the amphibian and the human *VGLL2* genes differ in size (11kb versus 17kb), they have an identical structure. Their coding sequence is split into four exons with intron/exon junctions completely conserved between the two genes (Fig. 1C). The chicken *vgll2* gene is only 6kb long but is nonetheless divided into four exons. In contrast, the mouse *Vgll2* gene, as well as the rat orthologue (data not shown), have lost their third intron (Fig. 1C).

We confirmed the evolution of the *vestigial like* genes in the vertebrate lineage by synteny analysis. We have identified in databanks sufficient information on the genomic regions containing the *X. tropicalis vgll1* and *vgll3* to perform this synteny analysis. As shown in Fig. 1D, *vgll1* and *vgll3* present a remarkable degree of conservation between human, mouse or chicken genomes. We have also found a similar synteny for *Vgll1* and *Vgll3* of the puffer fish *Takifugu rubripes* indicating that the genomic regions containing *VGLL1* and *VGLL3* genes have been conserved for at least 400-450 million years (i.e., the expected divergence time for bony-fish and tetrapod lineages). In the case of *vgll2* and *vgll4* genes there is not enough information available in databanks to establish their synteny.

We used semi quantitative RT-PCR analysis to assess the expression of *vestigial like* genes in embryo and adult tissues. *Vgll1*, *vgll2* and *vgll3* are not expressed maternally and their expression is first detected, although faintly, in stage 11 embryos (Fig. 2A). *Vgll1* and *vgll2* expression remains unchanged from stage 15 to tadpole stage whereas *vgll3* expression peaks in stage 15-20 embryos before decreasing to lower levels by stage 41 (Fig. 2A). In stage 45 embryo, *vgll3* expression is limited to the brain, stomach and somites but also, albeit at a lower level, in

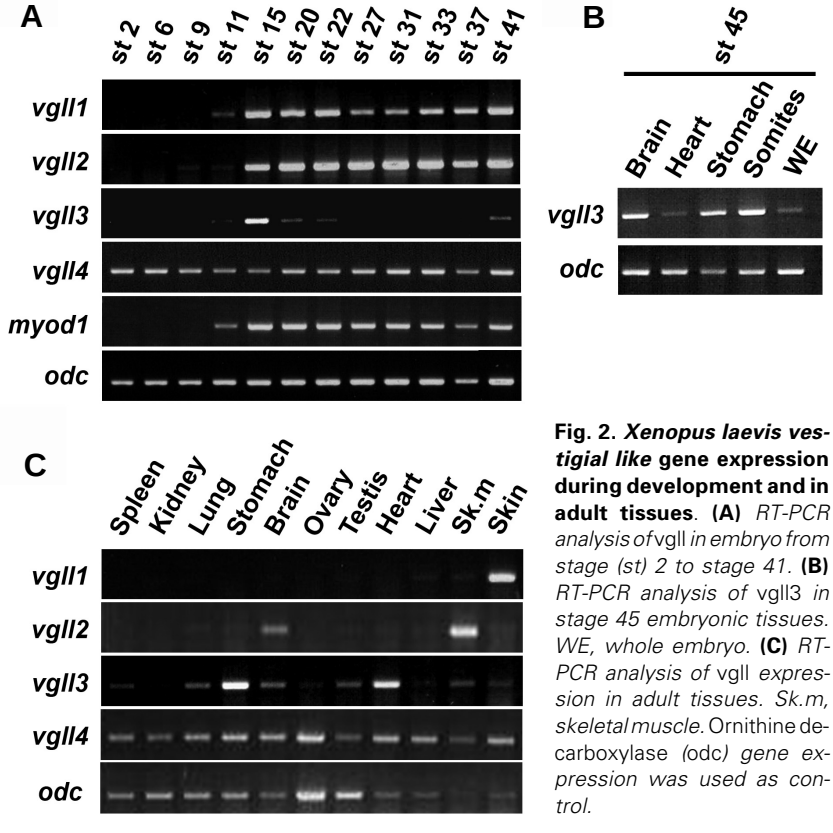


**Fig. 2. *Xenopus laevis vestigial like* gene expression during development and in adult tissues**. **(A)** *RT-PCR analysis of* vgll *in embryo from stage (st) 2 to stage 41.* **(B)** *RT-PCR analysis of* vgll3 *in stage 45 embryonic tissues. WE, whole embryo.* **(C)** *RT-PCR analysis of* vgll *expression in adult tissues. Sk.m, skeletal muscle.* Ornithine decarboxylase *(*odc*) gene expression was used as control.*

embryonic heart (Fig. 2B). *Vgll4* expression is detected in stage 2 embryos and its mRNA level remains at a constant level throughout early development (Fig. 2A).

In adult tissues, the expression of *vgll1* is strictly restricted to skin (Fig. 2C). The expression of *Xenopus vgll2* is detected solely in skeletal muscle and in brain, albeit at a much lower level (Fig. 2C). The *X. laevis vgll3* is predominantly expressed in the adult stomach and heart and at a lower level in lung, brain, testis, and skeletal muscle. The amphibian *vgll4* is expressed at a roughly similar level in all tissues analyzed (Fig. 2C).

The spatial expression of *vestigial like* genes in the early embryo was also investigated by whole mount *in situ* hybridization. *Vgll2* expression appears in *neurula* stage embryo (stage 19) as one stripe on each side of the dorsal midline, corresponding to the labeling of somites (Fig. 3A, a). As development proceeds, this expression extends along the dorsal side as more somites are formed. In addition, *vgll2* expression occurs in the branchial arches and the stomodeal-hypohyseal anlage of stage 26 embryos (Fig. 3A, b). In stage 30 embryos, this expression persists in the same territories as stage 26 embryos. However, in a transverse section the staining is more pronounced in the somitic dermomyotome region (Fig. 3 A, c,d). In stage 38 embryos, *vgll2* expression is found in the organizing somites, developing pituitary region, and in hypaxial and head muscle (Fig. 3A, e). In a transverse section at the head level *vgll2* staining is detected in the floor plate of the mesencephalon (Fig. 3A, f). Because the restricted expression of *vgll2* in the somite could reflect a possible probe penetration, we have performed double *in situ* hybridization and our data clearly indicate this is not case as *myod1* and *myl1* expression is detected in the whole somite (Fig. 3A, g,h).

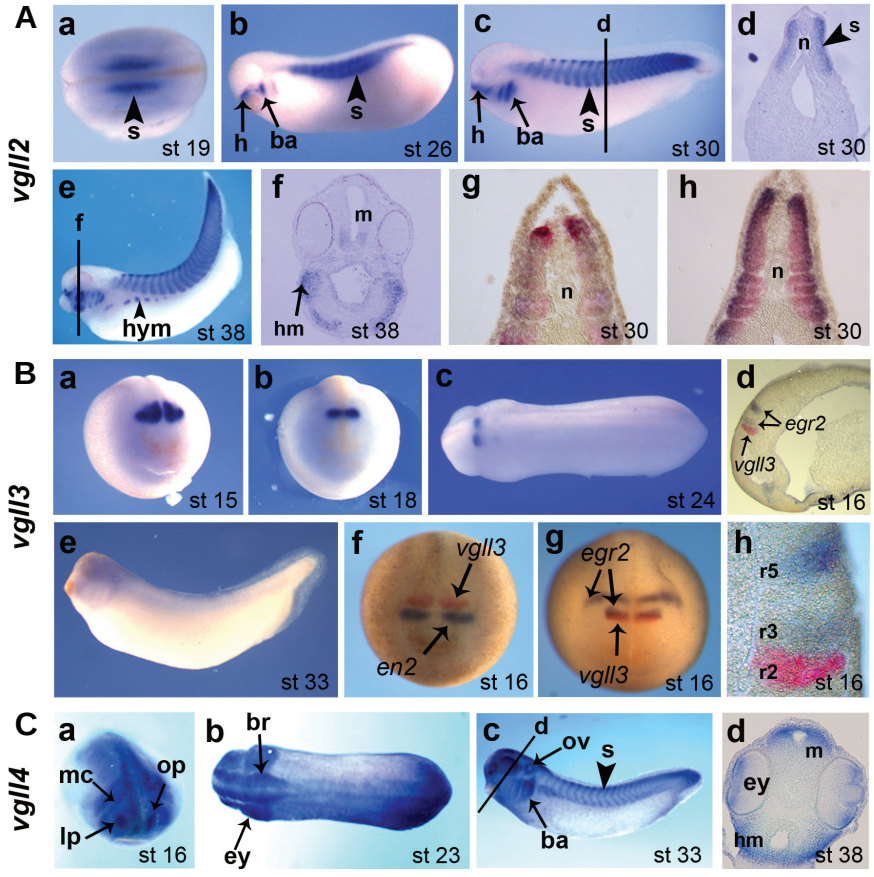


**Fig. 3. Spatial expression analysis of *Xenopus laevis vestigial like* genes in embryo by *in situ* hybridization. (A)** vgll2 *expression.* (a) *dorsal view;* (b,c,e) *lateral views;* (d,f,g,h) *transversal sections. Double* in situ *hybridization of stage 30 embryo with vgll2 and myod1 (g) or vgll2 and myl1 (h).* **(B)** vgll3 *expression. (a,b,f,g), frontal views; (c) dorsal view; (e) lateral view. (f) double* in situ *hybridization of stage 16 embryo with* vgll3 *and* en2 *(f) or vgll3 and egr2 (g). (d) parasagittal section of the embryo shown in g; (h) higher magnification of d. The position of rhombomeres is indicated (r2, r3 and r5).* **(C)** vgll4 *expression. (a) anterior view; (b) dorsal view; (c) lateral view; (d) transversal section. The positions of the section were indicated. Embryo staging according to Nieuwkoop and Faber (1967) is indicated. ba, branchial arches; br, brain; ey, eye; h, hypophyseal anlage; hm, head mesenchyme; hym, hypaxial muscle; lp, lens placode; m, mesencephalon; mc, mandibular neural crest; n, notochord; op, olfactory placode; ov, otic vesicle; r, rhombomere; s, somites.*

The expression of the *Xenopus vgll3* is most clearly detectable from stage 15 *neurula* embryo and is restricted to a prominent wide stripe on each side of the dorsal midline at the level of the anterior neural plate (Fig. 3B, a). In stage 18, the stripe of expression has become much narrower due to the closure of the neural tube (Fig. 3B, b). The expression of *vgll3* is still detected, albeit faintly, at the hindbrain level in stage 24 embryo (Fig. 3B, c) but is not detected anymore in later stages (Fig. 3B, e). In order to define more precisely the expression of *vgll3*, we carried out double *in situ* hybridization with either engrailed homeobox gene *2 (en2)* or *egr2* gene (formerly *krox-20)* that mark the midbrain/hindbrain boundary and rhombomeres 3 and 5 respectively. *En2* expression is anterior and separated by a gap from *vgll3* expression which lies just anterior to rhombomere 3 labeled by *egr2* (Fig. 3B, f,g). This is more conspicuous in cross sections of the embryos (Fig. 3B, d,h). Together, we can conclude that *vgll3* expression is restricted to hindbrain rhombomere 2 and this expression is transiently regulated during development. *Vgll4* expression is first detected in stage 16 embryos in epidermis, lens and olfactory placodes together with mandibular neural crest cells (Fig. 3C, a). In stage 23 embryos, the expression is found in the whole brain (Fig. 3C, b) and at later stages, *vgll4* staining persists in the eye, olfactory placode, otic vesicle, and brain as well as the branchial arches and somites (Fig. 3C, c). In transverse section, *vgll4* staining is conspicuous in the dorsal mesencephalon, eye and head mesenchyme (Fig. 3C, d).

Embryonic cell patterning of the early *Xenopus* embryo relies on limited signaling pathways that influence cell fate and ultimately regulate the expression of transcription factors in a cell type dependent manner. Of major importance are the FGF and activin signaling pathways (Heasman, 2006). We have tested in the animal cap assay whether those pathways could modulate *vestigial like* gene expression. As shown in Figure 4A, there is a decrease in *vgll1* expression concomitantly with increased activin

concentration; *vgll1* mRNA being undetectable at high dose of activin. The expression of *vgll1* was not significantly modified by bFGF treatment at either concentration used in these studies. In contrast, *Vgll2* expression was stimulated at all concentrations of activin used. This activin-dependent induction of *vgll2* is consistent with the expression of the gene in embryonic somites and adult skeletal muscle. In bFGF treated animal caps, *vgll2* is expressed at high concentration but not at low concentration like the somitic gene specific marker *myl1* (formerly *mlc1f/3f*) (Fig. 4A). Neither activin nor bFGF treatments modify significantly *vgll3* expression whereas the expression of *vgll4* is stimulated in a dose dependent manner by activin, but not by bFGF treatment (Fig. 4A).

We have not been able to detect the expression of *vgll1* by *in situ* hybridization directly in embryos. Because *vgll1* expression in adult tissues is restricted to epidermis we asked whether it was also expressed in the prospective epiderm of the embryo. We have isolated superficial (non neuronal) outer layer cells and deep (neuronal) inner layer cells from stage 10 embryo and analyzed in both cell types *vgll1* mRNA expression by RT-PCR. *Vgll1* mRNA, like *epidermal keratin* mRNA, is exclusively found in the epithelial outer cell layer (Fig. 4B). As expected, *hey1* gene is expressed solely in the inner layer cells of the *gastrula* embryo (Fig. 4B). *Tead1* and *tead2* are expressed in both layers of the animal cap explants and thus they can both act in cooperation with *vgll* genes in the distinct layers (Fig. 4B). We found that *vgll1* expression is markedly stimulated in dissociated cells after BMP treatment, similar to *vent1* which is a direct target of BMP signaling pathway (Fig. 4C). In complementary experiments, the expression of a dominant negative truncated BMP4-specific receptor (tBR4) in animal cap cells induced a down regulation of *vgll1* mRNA coincidentally with a reduction of both *epidermal keratin* and the immediate early responsive *msx1* genes (Fig. 4D). Similarly, *vgll1* expression is decreased in animal caps in the presence of noggin a BMP inhibitor (data not shown). Together these data indicate that unlike *vgll2*, *vgll3* or *vgll4* expression, which is not affected by BMP signals (data not shown), *vgll1* expression is activated by BMP4 in early embryos; a pattern consistent with its restricted expression pattern in the ectodermal lineage. During *Xenopus* early development, *vgll2* is mainly expressed in the somites but its expression is clearly detected after that of *myod1* (see Fig. 2A). Moreover a recent report indicates that *vgll2* is downstream of myogenic factors in chicken embryo (Bonnet *et al.*, 2009). We have analyzed *vgll2* expression in animal cap explants derived from embryos injected by mRNA encoding myogenic factors. As shown in Fig. 4E, *vgll2* expression is induced by the three myogenic factors tested myod1, myf5 and myf6. All three myogenic factors also induced the expression of the somite specific *myl1* gene as expected from our previous work (Thézé *et al.*, 1995). In contrast, overexpression of *vgll2* in animal cap cells never induces myogenic factor gene expression.

The restriction of *vgll3* expression in early embryo to rhombomere 2 prompted us to test whether this expression was regulated by en2 or egr2 which are found in midbrain/hindbrain

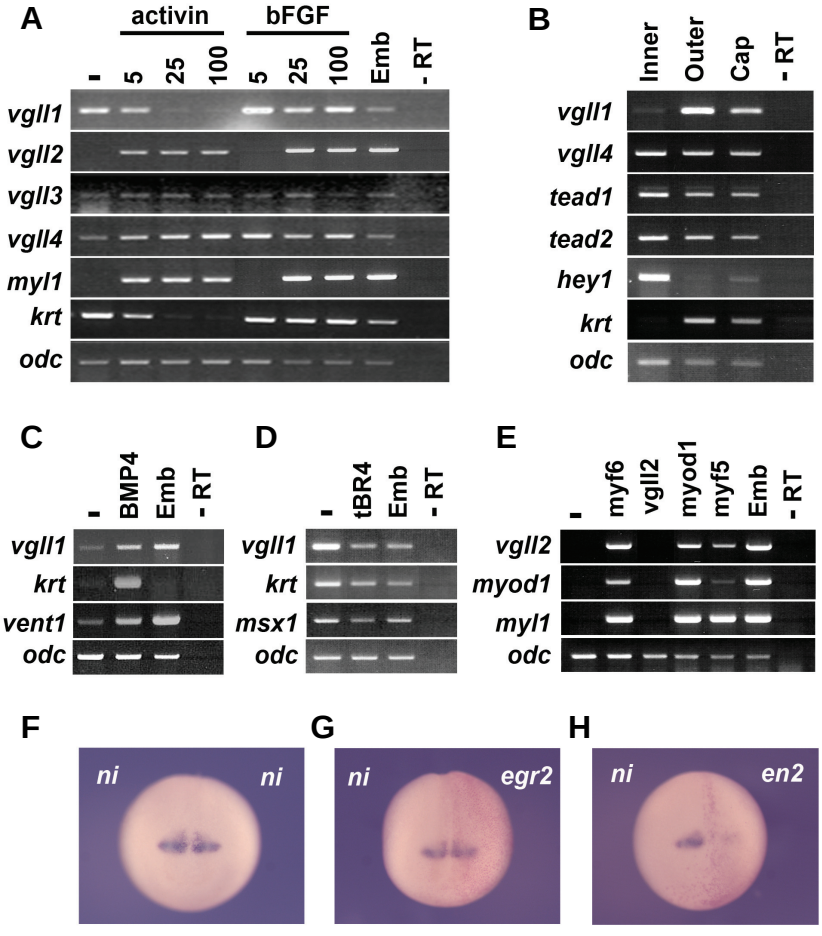


**Fig. 4. *Xenopus laevis vestigial like* gene regulation.** **(A)** *Animal caps treated with different amounts of activin or bFGF (5, 25 or 100 ng/ml) were analyzed by RT-PCR for the expression of the different* vgll *genes. For comparison the expression of* fast skeletal myosin light chain *(*myl1*) and* epidermal keratin *(*krt*) genes was assayed.* **(B)** *Animal cap explants from stage 9 embryos were dissociated into deep layer (Inner) cells and superficial layer (Outer) and assayed for* vgll1, vgll4, tead1 *and* tead2 *genes expression. Expression of* hey1 *and* epidermal keratin (krt) *genes was assayed in control. RNA from intact caps is used as control (Cap).* **(C)** *Dissociated inner cells from animal caps were treated with 1µg/ml of BMP4 (BMP4) for one hour or not (-) and assayed for the expression of* vgll1, epidermal keratin (krt) *and* vent1. **(D)** *250 pg of* tBR4 *mRNA were injected into 2-cell stage embryos and animal cap explants were assayed by RT-PCR for the expression of* vgll1, epidermal keratin (krt) *and* msx1. **(E)** *2-cell stage embryos were injected with 1ng of myogenic factors encoding mRNA (myod1, myf5 or myf6) plus 1ng of* tcf3 *(formerly E12) mRNA or 1ng of* vgll2 *mRNA alone. Animal cap explants from injected embryos or uninjected embryos (-) were assayed by RT-PCR for the expression of* vgll2, myod1 or myl1. *In (A,C,D,E), analyses were performed when control embryo (Emb) reaches stage 20. (-)* Odc *was used as control of loading and a reaction was performed in absence of reverse transcriptase to check for genomic DNA contamination (-RT).* **(F-H)** In situ *analysis of* vgll3 *expression in* neurula *embryo (frontal view): F, control embryo; G, H embryo injected with the indicated mRNA (*egr2, en2*). LacZ mRNA was used as a tracer in G, H. Uninjected side (*ni*) was used as control.*

junction or rhombomere 3 respectively. Overexpression of either egr2 or en2 in embryos was used to determine corresponding effects on the expression of *vgll3*. We have found that egr2 expression does not modify *vgll3* expression in rhombomere 2 (88.5% of embryos, n=52) while en2 clearly inhibits *vgll3* expression (84% of embryos, n=56) (Fig. 4, F,G,H).

## Discussion

In this paper we report the identification of the four members of the *vgll* gene family in *Xenopus* and their expression during development and adult tissues. This work is the first to comprehensively compare the expression pattern of the four vestigial like genes during the development of any vertebrate.

Like their vertebrate counterparts, the *Xenopus vgll* genes encode a protein containing a highly conserved TDU domain that mediates interaction with TEAD proteins. Phylogenetic analyses suggest that the vertebrate *vgll* are subdivided into four paralogous groups with VGLL1, VGLL2 and VGLL3 deriving from *Drosophila Vg* while the VGLL4 group having derived from the *Drosophila* NP_648658 gene which contains two TDU motifs (Chen *et al.*, 2004b). The synteny analysis has confirmed, in the case of *vgll1* and *vgll3,* a strong conservation between the amphibian genes and their mammalian or avian counterparts. In term of gene structure, the amphibian *vgll2* is structurally identical to the human and the avian genes and contains four exons. Moreover the exon/intron junctions are totally conserved between the three species. This is in contrast with the situation in mouse or rat where *Vgll2* contains only three exons. However, this is not unexpected because the loss of introns in the rodent lineage during mammalian evolution has been observed for several genes (Coulombe-Huntington and Majewski, 2007). Together, these data indicate that the genomic context of the vertebrate *vestigial like* genes has been conserved from amphibians to mammals. Moreover, each member of the amphibian *vgll* gene family is more related to its orthologs than to the other family members in the same species, suggesting that any function identified in *Xenopus* may well be conserved in other vertebrates.

We showed that among the four amphibian *vgll* genes, only *vgll4* is maternally expressed while the three others start to be expressed in the early embryo post mid-blastula transition. In adult tissues, the expression of *vgll1* is restricted to skin and this is in agreement with its embryonic expression where it is exclusively found in the epithelial outer cell layer of the *blastula*. This is in contrast with the situation observed in mammals where in human, *VGLL1* mRNA is mainly detected in placenta and at a trace level in kidney in adult and in kidney and lung during gestation (Maeda *et al.*, 2002; Vaudin *et al.*, 1999). *Vgll2* is expressed in adult skeletal muscle and brain while in mammals *VGLL2* expression is restricted to skeletal muscle (Maeda *et al.*, 2002; Mielcarek *et al.*, 2002). In adult *Xenopus*, *Vgll3* expression is found in stomach and heart and at a lower level in lung, brain, testis, and skeletal muscle. This expression pattern expression is quite similar to the one observed in mouse but distinct from the situation in human where *VGLL3* expression has been shown only in placenta (Maeda *et al.*, 2002; Mielcarek *et al.*, 2009). The ubiquitous expression of the amphibian *vgll4* closely resembles to the one described in adult human tissues. During *Xenopus* embryonic development, *vgll2* is expressed in the somitic dermomyotome and later in the hypaxial and head muscles. In mouse and zebrafish embryos, *Vgll2* showed a similar pattern of expression in the somites throughout embryonic development. However, there are some differences between vertebrate species as the zebrafish *Vgll2* gene is also expressed in notochord (Mann *et al.*, 2007).The expression in chicken embryo of *Vgll2* is also very similar to the one observed in *Xenopus* and marks all sites of skeletal muscle formation and the ventral region of the diencephalon (Bonnet *et al.*, 2009). However a major difference between the two species is that in *Xenopus*, *vgll2* is expressed in the dermomyotome while in chicken it is only observed in the central myotome and not in dermomyotome (Bonnet *et al.*, 2009). Regarding *vgll2* regulation, we have found that *vgll2* is downstream of the myogenic factor signaling pathway and seems not to act as a positive feedback regulator of the expression of myogenic factors; consistent with what was shown in chicken embryonic development (Bonnet *et al.*, 2009). The only site of expression of *vgll3* in the early *Xenopus* embryo is hindbrain rhombomere 2. This is in contrast with the situation observed in mouse embryo where *Vgll3* is mainly expressed in the myogenic lineage starting at E9.5 (Mielcarek *et al.*, 2009). However, *Vgll3* is also detected in the dorsal root ganglion at E11.5, suggesting a transient expression in hindbrain that should need to be confirmed. In a first approach to understand the rhombomere boundary 2 regulation, we have shown that en2 but not egr2 repress *vgll3* expression. This suggests that *vgll3* expression is regulated in its anterior border by *en2,* the genes involved in the posterior limit of expression at the r2/3 boundary awaiting characterization.

The four *Xenopus vestigial like* genes have a very distinct pattern of expression during development and in adult tissues that yet display similarities with the expression patterns observed among the mammalian and non-mammalian vertebrate orthologues. Thus, our data suggest that the amphibian *Xenopus*, represents a tractable vertebrate model to investigate the functions of *vgll* during early development and whether the diversity in both gene number and expression pattern of *vgll* genes is related to distinct functions.

## Materials and Methods

### *Phylogenetic and synteny analysis*

Sequences were extracted from NCBI and aligned with the Clustal W program Phylogenetic analyses were performed with the following sequences: *Xenopus laevis* (Xl) vgll1: GU327662; Xl vgll2: BC056001; Xl vgll3: BP689609; Xl vgll4: BC123267; *Xenopus tropicalis* (Xt) vgll1: CT030526; Xt vgll2: NM_203847; Xt vgll3: BC118899; Xt vgll4: AAI23006; *Homo sapiens* (Hs) VGLL1: NM_016267; Hs VGLL2: NP_872586; Hs VGLL3: NM_ 016206; Hs VGLL4: NP_001121691; *Mus musculus* (Mm) Vgll1: NP_573514; Mm Vgll2: AAN37898; Mm Vgll3:†NP_082848; Mm Vgll4: NP_808351; *Gallus gallus* (Gg) Vgll1: XP_001234166; Gg Vgll2: NP_001139462; Gg Vgll3: XP_416671; Gg Vgll4: NP_001025764; *Danio rerio* (Dr) Vgll2: NP_001020657; Dr Vgll3: XP_001332050; Dr Vgll4: NP_001073467; *Drosophila Melanogaster* (Dm) Vg: NP_523723; *Drosophila Melanogaster* (Dm) NP_648658. A phylogenetic tree was generated by the neighbor joining method using Mega 4 software The map location of the orthologous genes in human, mouse, chicken and *Xenopus tropicalis* were obtained from Ensembl (http://www.ensembl.org). The *X. tropicalis vgll3* was identified and located on scaffold_9 between igsf4d (*cadm2*) and *chmp2b* genes. Alignment of the cDNA and genomic sequences allowed us to identify the splice junctions.

***Plasmids and probes***

Plasmids containing cDNAs encoding *X. laevis vgll2* (IMAGE clone 4930090, accession number BC056001) and *vgll4* (IMAGE clone 8320681, accession number BC123267) were obtained from Geneservice (http://www.geneservice.co.uk). cDNA encoding *X. laevis vgll3* (XL405a05ex, accession number BP689606) was obtained from the National BioResource Project (http://www.nbrp.jp). Clone identity was confirmed by sequencing both strands. The cDNA encoding the open reading frame of *X. laevis* vgll1 was cloned by RT-PCR from stage 35 total RNA embryo using the following primers designed from the *X. tropicalis vgll1* sequence (Accession number CT030526): Forward primer 5'-ATGGAAGACCTTCATAAA-3' and Reverse primer 5'-ATCTGTTTGAGGATTCATC-3'. PCR fragment was cloned with PCR®II-TOPO® TA cloning kit according to manufacturer instructions (Invitrogen) followed by sequencing.

***Xenopus embryo collection and whole mount* in situ *hybridization***

*Xenopus laevis* embryos were obtained using current protocols (Sive *et al.*, 2000). Whole-mount *in situ* hybridization was carried out with full length anti-sense probes labeled with digoxigenin corresponding to indicated gene sequences and visualized with BM purple staining (Boehringer) (Harland, 1991). For double *in situ* hybridization, probes were visualized sequentially with the corresponding specific alkaline phosphatase conjugated antibody using first Fast Red (Boehringer) for DIG labeled anti-sense *vgll3* probe and then BMP purple (Boehringer) for the fluorescein-labeled anti-sense *en2* and *egr2* probes. For serial sections, embryos were post fixed in MEMFA for 1hr at RT and embedded in agarose before sectioning.

***Microinjection of embryos and animal cap assay***

Microinjection of embryo, animal cap assay and RT-CR analysis were performed as previously described (Naye *et al.*, 2007). For microinjection experiments, capped mRNAs were synthesized *in vitro* by using Ambion mMessage mMachine SP6 and T7 kits (Austin, TX). The *Xenopus laevis vgll2* complete coding sequence was subcloned into *pCS2+* vector and mRNA was synthesized by SP6 RNA polymerase. Embryos were injected in the animal cap with a known amount of the mRNA solution at the two-cell stage into both blastomeres using a Nanoject system (Drummond Scientific). Animal caps were then dissected from stage 8-9 injected embryos and cultured until appropriate stages before RNA extraction. In egr2 and en2 misexpression experiments, 50 pg or 100 pg of mRNA respectively were injected into one blastomere at the two cell stage with 250pg of *LacZ* mRNA as a lineage tracer. For animal cap cells dissociation, superficial and deep layers cells were isolated from early stage 10 embryos and cultured in calcium magnesium free medium (Sive *et al.*, 2000). In BMP experiments, the dissociated deep layer cells were cultured for one hour in presence of 1µg/ml of recombinant human BMP4 (R&D system) before RNA extraction and RT-PCR analysis.

***RT-PCR analysis***

RT-PCR analysis was performed as previously described (Naye *et al.*, 2007) with the following primer pairs:

| | |
|---|---|
| *vgll1* | 5'-TAAGGATGCCTGCATCTGTC-3' |
| | 5'-ATCCCTGTTGCTGTAGATGC-3' |
| *vgll2* | 5'-GAAGATGCAGGAAGCACCAG-3' |
| | 5'-CCAAAATGATGGTGGGAAGC-3' |
| *vgll3* | 5'-ACATGAGACCATACCACCTC-3' |
| | 5'-CACATACCTGTCTAAAGCC-3' |
| *vgll4* | 5'-TGTGTCTATCACCGGTTCGG-3' |
| | 5'-TCTTTGGTAACAGAGGCAGG-3' |
| *tead1* | 5'-ATTCCAGCCTGCAAACTTCC-3' |
| | 5'-GTGACTAGACACCTGTTTCC-3' |
| *tead2* | 5'-TTCTGCGATCGCACTAAGAAGC-3' |
| | 5'-GCCATACATTTTGCCTTC-3' |
| *msx1* | 5'-ACTGGTGTGAAGCCGTCCCT-3' |
| | 5'-TTCTCTCGGGACTCTCAGGC-3' |
| *vent1* | 5'-TTCCCTTCAGCATGGTTCAAC-3' |
| | 5'-GCATCTCCTTGGCATATTTGG-3' |
| *krt* | 5'-CACCAGAACACAGAGTAC-3' |
| | 5'-CAACCTTCCCATCAACCA-3' |
| *hey1* | 5'-GGATTACAAGCAAGGGTTC-3' |
| | 5'-TCCCATAGGATAACGTTCAT-3' |
| *myl1* | 5'-TTTGACAAGGAAGGCAATGG-3' |
| | 5'-CATTCTGCTGACAGTTCTTG-3' |
| *odc* | 5'-GTCAATGATGGAGTGTATGGATC-3' |
| | 5'-TCCATTCCGCTCTCCTGACCAC-3' |

*Acknowledgements*

*This work has been supported by the University of Bordeaux, the CNRS and a grant from the AFM "Association Française contre les Myopathies". We thank Dr. J. Lee for critical reading of the manuscript and Dr. P. Charnay and Dr. M. Gradl for providing us the egr2 and en2 expression plasmids respectively.*

## References

BERNARD, F., LALOUETTE, A., GULLAUD, M., JEANTET, A. Y., COSSARD, R., ZIDER, A., FERVEUR, J. F., and SILBER, J. (2003). Control of apterous by vestigial drives indirect flight muscle development in *Drosophila*. *Dev Biol* 260: 391-403.

BONNET, A., DAI, F., BRAND-SABERI, B., and DUPREZ, D. (2009). Vestigial-like 2 acts downstream of MyoD activation and is associated with skeletal muscle differentiation in chick myogenesis. *Mech Dev* 127: 120-136.

CHEN, H. H., MAEDA, T., MULLETT, S. J., and STEWART, A. F. (2004a). Transcription cofactor Vgl-2 is required for skeletal muscle differentiation. *Genesis* 39: 273-279.

CHEN, H. H., MULLETT, S. J., and STEWART, A. F. (2004b). Vgl-4, a novel member of the vestigial-like family of transcription cofactors, regulates alpha1-adrenergic activation of gene expression in cardiac myocytes. *J Biol Chem* 279: 30800-30806.

COULOMBE-HUNTINGTON, J., and MAJEWSKI, J. (2007). Characterization of intron loss events in mammals. *Genome Res* 17: 23-32.

GUNTHER, S., MIELCAREK, M., KRUGER, M., and BRAUN, T. (2004). VITO-1 is an essential cofactor of TEF1-dependent muscle-specific gene regulation. *Nucleic Acids Res* 32: 791-802.

HALDER, G., POLACZYK, P., KRAUS, M. E., HUDSON, A., KIM, J., LAUGHON, A., and CARROLL, S. (1998). The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in *Drosophila*. *Genes Dev* 12: 3900-3909.

HARLAND, R. M. (1991). *In situ* hybridization: an improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol* 36: 685-695.

HEASMAN, J. (2006). Patterning the early *Xenopus* embryo. *Development* 133: 1205-1217.

KANEKO, K. J., and DEPAMPHILIS, M. L. (1998). Regulation of gene expression at the beginning of mammalian development and the TEAD family of transcription factors. *Dev Genet* 22: 43-55.

KURRASCH, D. M., CHEUNG, C. C., LEE, F. Y., TRAN, P. V., HATA, K., and INGRAHAM, H. A. (2007). The neonatal ventromedial hypothalamus transcriptome reveals novel markers with spatially distinct patterning. *J Neurosci* 27: 13624-13634.

LARKIN, S. B., and ORDAHL, C. P. (1999). Multiple layers of control in transcriptional regulation by MCAT elements and the TEF-1 protein family. In *Heart development* (Eds. R. P. Harvey and N. Rosenthal), pp. 307-329. Academic Press, San Diego.

MAEDA, T., CHAPMAN, D. L., and STEWART, A. F. (2002). Mammalian vestigial-like 2, a cofactor of TEF-1 and MEF2 transcription factors that promotes skeletal muscle differentiation. *J Biol Chem* 277: 48889-48898.

MANN, C. J., OSBORN, D. P., and HUGHES, S. M. (2007). Vestigial-like-2b (VITO-1b) and Tead-3a (Tef-5a) expression in zebrafish skeletal muscle, brain and notochord. *Gene Expr Patterns* 7: 827-836.

MANN, R. S., and CARROLL, S. B. (2002). Molecular mechanisms of selector gene function and evolution. *Curr Opin Genet Dev* 12: 592-600.

MIELCAREK, M., GUNTHER, S., KRUGER, M., and BRAUN, T. (2002). VITO-1, a novel vestigial related protein is predominantly expressed in the skeletal muscle lineage. *Gene Expr Patterns* 2: 305-310.

MIELCAREK, M., PIOTROWSKA, I., SCHNEIDER, A., GUNTHER, S., and BRAUN, T. (2009). VITO-2, a new SID domain protein, is expressed in the myogenic lineage during early mouse embryonic development. *Gene Expr Patterns* 9: 129-137.

NAYE, F., TRÉGUER, K., SOULET, F., FAUCHEUX, C., FÉDOU, S., THÉZÉ, N., and THIÉBAUD, P. (2007). Differential expression of two TEF-1 (TEAD) genes during *Xenopus laevis* development and in response to inducing factors. *Int J Dev Biol* 51: 745-752.

PASQUET, S., NAYE, F., FAUCHEUX, C., BRONCHAIN, O., CHESNEAU, A., THIÉBAUD, P., and THÉZÉ, N. (2006). Transcription enhancer factor-1-dependent expression of the alpha-tropomyosin gene in the three muscle cell types. *J Biol Chem* 281: 34406-34420.

SIMMONDS, A. J., LIU, X., SOANES, K. H., KRAUSE, H. M., IRVINE, K. D., and BELL, J. B. (1998). Molecular interactions between Vestigial and Scalloped promote wing formation in *Drosophila*. *Genes Dev* 12: 3815-3820.

SIVE, H. L., GRAINGER, R. M., and HARLAND, R. M. (2000). «Early Development of *Xenopus laevis*.» Cold Spring Harbor Laboratory Press,

THÉZÉ, N., HARDY, S., WILSON, R., ALLO, M.-R., MOHUN, T., and THIÉBAUD, P. (1995). The MLC1f/3f gene is an early marker of somitic muscle differentiation in *Xenopus laevis* embryo. *Dev. Biol* 171: 352-362.

VAUDIN, P., DELANOUE, R., DAVIDSON, I., SILBER, J., and ZIDER, A. (1999). TONDU (TDU), a novel human protein related to the product of vestigial (vg) gene of *Drosophila Melanogaster* interacts with vertebrate TEF factors and substitutes for Vg function in wing formation. *Development* 126: 4807-4816.

WILLIAMS, J. A., BELL, J. B., and CARROLL, S. B. (1991). Control of *Drosophila* wing and haltere development by the nuclear vestigial gene product. *Genes Dev* 5: 2481-2495.