Regulation and function of Spalt proteins during animal development

JOSE F. DE CELIS\textsuperscript{1} and ROSA BARRIO\textsuperscript{2}

\textsuperscript{1}Centro de Bología Molecular “Severo Ochoa”. Universidad Autónoma de Madrid and Consejo Superior de Investigaciones Científicas, Madrid and \textsuperscript{2}Functional Genomics, CIC bioGUNE, Bizkaia Technology Park, Derio, Spain.

ABSTRACT  The genes of the \textit{spalt} (\textit{sal}) family play fundamental roles during animal development. The two members of this family in \textit{Drosophila}, \textit{spalt} (\textit{sal}) and \textit{spalt-related} (\textit{salr}) encode Zn-finger transcription factors that link the Decapentaplegic (Dpp)/BMP signalling pathway to the patterning of the wing. They are regulated by the Dpp pathway in the wing disc, and they were shown to mediate some of the morphogenetic activities of the Dpp/BMP4 secreted ligand. The \textit{sal} genes were initially found by virtue of mutations that produce homeotic transformations in the head and tail of the \textit{Drosophila} embryo. Since then, a number of other requirements have been associated to these genes in \textit{Drosophila}, including morphogenesis of the respiratory system, cell fate specification of sensory organs and the differentiation of several photoreceptor cells, among others. Vertebrate \textit{sal} orthologues (\textit{spalt-like/sall}) have also important developmental roles during neural development and organogenesis, and at least two human \textit{sal} genes are linked to the genetic diseases Townes Brocks Syndrome (TBS; \textit{SALL1}) and Okihiro Syndrome (OS; \textit{SALL4}). In this review, we will summarize the main characteristics of the \textit{sal} genes and proteins, pointing out to the similarities in their developmental roles during \textit{Drosophila} and vertebrate development.

KEY WORDS: \textit{spalt}, gene regulation, organogenesis, embryonic development

The Sall protein family

Sall proteins are zinc finger transcription factors present from \textit{C. elegans}, which harbours only one member of the family, to vertebrates, which generally present four \textit{spalt} genes (\textit{sall1-4}). The \textit{Drosophila} genome contains two paralogues, \textit{spalt} (\textit{sal}) and \textit{spalt-related} (\textit{salr}) which form part of a gene complex (Kuhnlein \textit{et al.}, 1994; Barrio \textit{et al.}, 1996). The more characteristic feature of Sall proteins is the presence of several zinc finger domains scattered along the protein (Fig. 1). Zinc finger domain 1 corresponds to the C2HC class, and it is only present in the vertebrate homologues. The rest of the domains (2-5) correspond to C2H2 zinc fingers arranged in pairs. The doublets are connected by a H/C link conserved throughout evolution, and the second finger from each pair contains a characteristic domain called Sal-box that is present in other zinc finger transcription factors. The third finger domain contains an associated finger, also highly conserved among orthologs. Another important domain characteristic of these proteins is a Glutamine rich region (polyQ), present from \textit{Drosophila} to humans, which might be involved in protein-protein interactions among members of the family and between Sall and other proteins. The four orthologues vertebrate proteins, Sall1-4, display differences in the finger distribution, being Sall2 the more distant member of the family (Fig. 1; Kohlhase \textit{et al.}, 1996; Hollemann \textit{et al.}, 1996; Kohlhase \textit{et al.}, 1999a; Onuma \textit{et al.}, 1999; Ott \textit{et al.}, 1996; Buck \textit{et al.}, 2000; Ott \textit{et al.}, 2001; Ma \textit{et al.}, 2001; Kohlhase \textit{et al.}, 2002a; Ma \textit{et al.}, 2006). The nematode Sall protein, named Sem-4, shares common features with their homologues, like the finger domains 3 and 5 (Fig. 1; Basson and Horvitz, 1996; Photos \textit{et al.}, 2006). For a recent phylogenetic analysis of the Sall family, and a comprehensive update on the nomenclature of vertebrate orthologues, see a recent review by Sweetman and Munsterberg (2006).

There are several similarities among Sall activities in different organisms, such as their functions during embryonic development in a variety of processes including organogenesis, limb formation and cell fate assignment during neural development. In

*Address correspondence to: Jose F. de Celis. Centro de Bología Molecular “Severo Ochoa”. Universidad Autónoma de Madrid and Consejo Superior de Investigaciones Científicas. Cantoblanco. Madrid 28049. Spain. e-mail: jfdecelis@cbm.uam.es or Rosa Barrio. Functional Genomics, CIC bioGUNE, Bizkaia Technology Park. Building 801A, E-48160 Derio. Spain. e-mail: rbarrio@cicbiogune.es

Final author-corrected PDF published online: 27 October 2008.
this review we will summarize different aspects of Sall proteins and genes biology, with emphasis in their modes of regulation, their functions in proliferation and transcription, their developmental roles in different organisms and their association with several human genetic diseases.

**Regulation of sall gene expression**

Most of what is known about the regulation of sall expression derives from studies in Drosophila sall and salr genes and in some vertebrate sall members. A common aspect is that the expression of sall genes depends on the activity of several signal transduction pathways (Table 1). In particular, the Wnt, FGF, Shh, EGFR and BMP pathways participate in the activation of sall expression in different tissues and, in some cases, it has been shown that Sall proteins are key mediators of the function of these pathways during organogenesis and cell differentiation. The regulation of sall and salr in Drosophila has been studied extensively, and a number of tissue specific enhancers have been characterized (Wagner-Bernholz et al., 1991; Kuhnlein et al., 1997; Chen et al., 1998; Barrio et al., 1999; de Celis et al., 1999; Guss et al., 2001; Barrio and de Celis, 2004). In this organism, the sall and salr transcription units are separated by 50kb of non-coding DNA containing regulatory sequences. sall is expressed during embryonic development in a variety of tissues, including the cellular blastoderm, posterior spiracles, trachea, oenocytes and cells in the central and peripheral nervous system (Fig. 2). The regions where salr is expressed overlap in all these tissues, except in the early blastoderm where sall is not expressed (Barrio et al., 1996). During larval development, sall and salr are expressed in the same cells in the wing, eye-antenna and haltere imaginal discs, as well as in the ring gland and central nervous system (Fig. 2). The structure of the sall and salr regulatory regions shows many similarities with those of other Drosophila gene complexes, such as the achaete-scute and iroquois complexes (Ruiz-Gomez and Modolell, 1987; Gomez-Skarmeta et al., 1996). Thus, tissuespecific enhancers are scattered in the 50 Kb intergenic region and also in the 5’ and intronic regions of both genes (Fig. 2). The expression of the sall and salr transcripts is regulated by separate and, in some cases, shared cis-regulatory elements (Fig. 2). Enhancers that direct the expression of sall in the blastoderm, wing and tracheae are some of the best characterized so far (Kuhnlein et al., 1997; Barrio and de Celis, 2004; Chen et al., 1998).

The detailed analysis of sall/salr regulatory elements in the wing disc showed an even greater complex organization, in that independent enhancers control the expression in different territories such as the wing pouch, thorax, hinge and pleura (Fig. 2; Barrio et al., 1999; de Celis et al., 1999). Interestingly, the expression in the thorax is also controlled by multiple elements affecting specific sub-domains. The organization of modular regulatory regions implies that the territories of sall and salr expression are, from the regulatory point of view, a mosaic of cell populations where different combinations of factors are responsible for the activation of each gene in different groups of cells. The expression of sall genes in the wing pouch is directly regulated by the Dpp pathway, acting through sall and salr independent enhancers. The Dpp pathway activates sall expression in a central domain that is broader than the dpp expression territory through a genomic region of 453 bp localized 5’ of the sall transcript (Barrio and de Celis, 2004). This enhancer integrates positive inputs mediated by the Dpp effectors Mad/Medea with the repressor activity of Brinker. The mechanism of repression by Brinker does not rely on competition with Mad–Medea overlapping sites, but on the existence of adjacent binding sites for Brinker and Mad/Med (Barrio and de Celis, 2004). Additional factors such as the T-box transcription factor Optomotor blind, the trithorax protein Ash2, the activator complex Vestigial/Scalloped and the repressor Groucho are also involved in the regulation of sall in the wing blade (Guss et al., 2001; del Alamo Rodriguez et al., 2004; Angulo et al., 2004; Winter and Campbell, 2004; Hasson et al., 2005). The enhancer regulating salr expression in the wing blade has not yet been identified.

The regulation of sall genes expression in organisms other than Drosophila is less documented. However, some of the enhancers that direct tissue specific expression of human SALL1 have been identified by virtue of their sequence conservation, and have been tested experimentally in chicken and mice embryos (Table 1; Pennacchio et al., 2006; Izumi et al., 2007). A recurrent
aspect in the regulation of vertebrate sal genes is the involvement of signaling pathways in different developmental systems. For example, the expression of *Xenopus Xsal4* within the interdigital spaces suggests that BMP proteins are involved in regulating its expression in these territories (Neff et al., 2005). Similarly, the *Msall3* gene from Medaka fish is expressed in most places where Hedgehog signaling is active, and Hedgehog regulates the expression of the gene at the midbrain-hindbrain organizer region (Koster et al., 1997). In this territory, FGF signaling is required to activate *Msall3* expression in response to Shh during Medaka development, and this regulatory relationship is also observed during the growth of the optic vesicle (Carl and Wittbrodt, 1999). The FGF pathway, now in collaboration with Wnt signaling, is also required for the activation of *csall1* expression in chicken limb buds, where *csall1* is expressed in the apical ectodermal ridge and in the underlying distal mesenchyme (Farrell and Munsterberg, 2000). In these cells, a combination of Wnt3a and Wnt7a with FGF4 and FGF8, which are expressed in the apical ectodermal ridge, regulates *csall1* expression, whereas BMP function is also required to activate *csall1* in mesenchymal cells of the proximal limb (Capdevila et al., 1999; Farrell and Munsterberg, 2000). A recent analysis of the human *SALL4* promoter region identified 367 bp located upstream of the ATG which sequence is extremely conserved in several vertebrates *sal* genes. The observation that this region contains consensus-binding sites, which integrity is required for promoter activity in cell culture assays, for LEF/TCF, a transcription factor mediating the response to canonical Wnt signaling, implies a direct effect of TCF on TCF, a transcription factor mediating the response to canonical Wnt signaling, *in vivo* (Onai et al., 2004). The regulatory relationships between Wnt signaling and *sal* are also observed in Drosophila and *Xenopus*. Thus, *wingless*, a Drosophila Wnt homologue, induces sal expression during tracheal development in the fly (Chihara and Hayashi, 2000; Ribeiro et al., 2004), and TCF3 is required for *Xsal2* expression in the forebrain/midbrain at the early neural stage in the frog (Onai et al., 2004). Interestingly, *Xsal2* and human SALL1 modify the response to Wnt signaling, although *Xsal2* antagonises Wnt signaling *in vivo* (Onai et al., 2004), and human SALL1 promotes Wnt signaling in cell culture assays (Sato et al., 2004). The function of *Xsal2* is essential for the expression of the *Pax6, Otx2, and Bf-1* genes in the forebrain/midbrain region, and for the repression of the caudal genes *En2, Pax2, Wnt1, and Gbx2*. *Xsal2* is also required for anterior expressions of two antagonistic effectors of Wnt signaling, *GSK3* and *Tcf3* (Onai et al., 2004).

The expression patterns of *sal* family genes and the analysis of their regulation indicates that Sal function can not be universally assigned to specific signaling pathways, but rather that Sal has been adopted by different signaling pathways in different developmental contexts. Similarly, it appears that orthologues, as determined by degrees of conservation of *sal* coding sequences, do not imply similarities of expression patterns.

### Function of Sall proteins in gene regulation

The genetic approach to study *sal* function in *Drosophila* identified a number of developmental processes in which *sal* and *salr* are involved. In addition, this approach also allowed in some instances to place *sal* and *salr* into genetic hierarchies, in which both upstream and downstream elements to *sal/salr* were identified. Some of these aspects will be considered latter when addressing the specific roles of *sal/salr* in different tissues and limb development. However, very few data are available in flies about the molecular mechanisms of Sal function, and no comprehensive analysis of Sal/Salr partners and target genes has been carried out yet. Thus, a direct interaction with DNA has only been shown in the case of Salr, which is able to bind an AT-rich sequence in the chordin gene *s15* promoter with the central zinc finger domain 3 (Table 2; Shea et al., 1990; Barrio et al., 1996).

### Table 1

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene</th>
<th>Regulator</th>
<th>Enhancer</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>SALL1</td>
<td>SIX1</td>
<td>Tissue-specific enhancers (1)</td>
<td>Activation</td>
<td>Izumi et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WT1</td>
<td>Position -947</td>
<td>Repression</td>
<td>Chai et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>csall1</td>
<td>Position -2000 to +1</td>
<td>Repression</td>
<td>Chai et al., 2006</td>
</tr>
<tr>
<td>Mouse</td>
<td>SALL2</td>
<td>Wilms Tumor-1</td>
<td>Tissue-specific enhancers (2)</td>
<td>Repression</td>
<td>Pernacchio et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wnt (LEFTI)</td>
<td>Position -249 to -218</td>
<td>Activation</td>
<td>Boehm et al., 2006</td>
</tr>
<tr>
<td></td>
<td>Sall1</td>
<td>Shh (GLI3)</td>
<td>Position -1344 to -1137</td>
<td>Activation</td>
<td>Hu et al., 2006</td>
</tr>
<tr>
<td></td>
<td>Sall3</td>
<td>csall1</td>
<td>T-OMR region</td>
<td>Activation</td>
<td>Dhagane et al., 2004</td>
</tr>
<tr>
<td></td>
<td>Sall4</td>
<td>Nanog</td>
<td>ES enhancer</td>
<td>Activation</td>
<td>Wu et al., 2006</td>
</tr>
<tr>
<td></td>
<td>Sall/Salr</td>
<td>Tissue-specific enhancers (3)</td>
<td>?</td>
<td>?</td>
<td>Barrio et al., 1999</td>
</tr>
<tr>
<td>Drosophila</td>
<td>Bcd, Cad, TII, Hb</td>
<td>Blastoderm enhancer</td>
<td>Activation</td>
<td>Kuhlen et al., 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hkb, Kr</td>
<td>Blastoderm enhancer</td>
<td>Repression</td>
<td>Kuhlen et al., 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kni/Kni</td>
<td>Trachea enhancer</td>
<td>Repression</td>
<td>Kuhlen et al., 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ubx</td>
<td>sal 328 wing enhancer</td>
<td>Repression</td>
<td>Galant et al., 2002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sc</td>
<td>sal 328 wing enhancer</td>
<td>Activation</td>
<td>Guass et al., 2001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dpp (Mad/Med)</td>
<td>sal 328 wing enhancer</td>
<td>Activation</td>
<td>Guass et al., 2001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dpp (Med)</td>
<td>Ecoli-Ndel wing enhancer</td>
<td>Activation</td>
<td>Barrio and de Celis 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dpp (Bre)</td>
<td>Ecoli-Ndel wing enhancer</td>
<td>Repression</td>
<td>Barrio and de Celis 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Antp</td>
<td>Ecoli/Anthera enhancer</td>
<td>Repression</td>
<td>Wagner-Bernholz et al., 1991</td>
<td></td>
</tr>
</tbody>
</table>

Only the regulators shown to interact directly with *sal* promoters or enhancers are included in this Table. In some cases, the enhancers have been isolated, but the regulators are unknown and they are indicated by question marks. (1) Tested in chicken; Prospero enhancer and anterior neural ridge. (2) Tested in mouse; Forebrain, midbrain, hindbrain, neural tube, limb, eye, dorsal root ganglia, somites, nose, branchial arch, genital tubercle, trigeminal nerve, heart, neural crest mesenchyme, melanocytes and cranial nerve. (3) Tested by reporter activity. (4) Embryonic (central nervous system, peripheral nervous system, oenocytes, trachea, gut, epidermis and larval (wing, haltere, eye, CNS, leg, ring gland) enhancers. (5) Wing blade, hinge and thorax. Data were compiled from the references indicated in the right-hand column.
First, the N-terminal part of the protein contains a 12 amino acids sequence that is able by itself to confer repression capacity and to interact with the Histone Deacetylase Complex NuRD (Kiefer et al., 2002; Lauberth and Rauchman, 2006). This interaction can be modified by phosphorylation of Sall1 (Lauberth et al., 2007). The NuRD-interaction domain is also found in other Sall homologues, including C. elegans Sem-4, and in transcription factors not related to the Sall family, but it is not present in the Drosophila Sal homologues. In the cases of human and murine SALL2/Sall2 and SALL4/Sall4, alternative spliced forms have been described that lack this repression domain that would function independently of the NuRD repression complex, although the functional role of these alternative forms is still unexplored. The N-terminal part of the Sall1 shows localization to heterochromatin foci when fused to a nuclear localization signal, suggesting an association between transcriptional repression and protein location (Kiefer et al., 2002; Sato et al., 2004).

Similarly, CK2 kinase is the only protein reported to interact with Drosophila Sal (Trott et al., 2001). However, the biological relevance of these interactions has not yet been explored.

In contrast to the paucity of data concerning Sal molecular function in Drosophila, a wealth of data identifying Sall protein-protein interactions, Sall subcellular localization and Sall transcriptional effects are stemming from the analysis of vertebrate sall genes (Table 2). In what follows we will summarise some of the interactions identified for the sall genes and proteins 1, 2 and 4, which taken together suggest that the variety of processes requiring Sall function can be accounted by the diversity of protein-protein and protein-DNA interactions in which Sall proteins are engaged (see Fig. 1).

Human SALL1 has been described as a transcriptional repressor in a number of experimental settings, most of them involving the regulation of heterologous promoters fused to reporter genes, and presents two possible mechanisms of repression (Nishinakamura et al., 2001; Netzer et al., 2001; Kiefer et al., 2002; Sweetman et al., 2003; Netzer et al., 2006). First, the N-terminal part of the protein contains a 12 amino acids sequence that is able by itself to confer repression capacity and to interact with the Histone Deacetylase Complex NuRD (Kiefer et al., 2002; Lauberth and Rauchman, 2006). This interaction can be modified by phosphorylation of Sall1 (Lauberth et al., 2007). The NuRD-interaction domain is also found in other Sall homologues, including C. elegans Sem-4, and in transcription factors not related to the Sall family, but it is not present in the Drosophila Sal homologues. In the cases of human and murine SALL2/Sall2 and SALL4/Sall4, alternative spliced forms have been described that lack this repression domain that would function independently of the NuRD repression complex, although the functional role of these alternative forms is still unexplored. The N-terminal part of the Sall1 shows localization to heterochromatin foci when fused to a nuclear localization signal, suggesting an association between transcriptional repression and protein location (Kiefer et al., 2002; Sato et al., 2004).
The second repression mechanism is independent of the Histone Deacetylase Complex and requires the central region of the protein including the finger domains 2 and 3 (Netzer et al., 2001; Netzer et al., 2006). This region also shows localization in heterochromatin foci in murine cells. In addition, SALL1 can interact with PIN2, an isoform of telomere-repeat binding factor 1 (TRF1) (Netzer et al., 2001). TRF1/PIN2 binds to telomeres, suggesting a mechanism of repression for SALL1 by association to pericentromeric heterochromatin. Yet another region of the protein located in the C-terminal fingers has been described as important for the interaction with heterochromatin. This domain is particularly well conserved from Drosophila humans and it has been reported to bind the major satellite DNA (Table 2; Yamashita et al., 2007).

Even though controversy exists about the identity of the domain involved in SALL-DNA interactions and the existence of different repression domains, it is interesting to speculate that SALL proteins might recruit remodelling factors to heterochromatin. In this context, SALL1 is able to bind to β-catenin and activate synergistically a reporter construct responding to the Wnt signalling pathway (Sato et al., 2004). However, the domain of SALL1 that co-activates this reporter does not coincide with the β-catenin binding domain, but with the heterochromatin localization domain, indicating that SALL1 localization, and not its interaction with β-catenin, is the mediator of the interactions between SALL1 and the Wnt signalling pathway. In vivo, the role of human SALL1 as a transcriptional repressor has been shown during stereogoni- sesis in adrenal gland, where SALL1 represses the expression of the enzymes 11-hydroxylase and aldosterone synthase, involved in the glucocorticoid and mineralocorticoid biosynthetic pathways under the modulation of Angiotensin II (Romero et al., 2007). In contrast, murine SALL1 is necessary for the activation of some kidney mesenchymal markers, consistent with its role in ureteric bud invasion (Nishinakamura et al., 2001). As in the case of the activation of Wnt signalling, the up-regulation of these genes might not be direct.

Sall protein interactions

The subcellular localization and transcriptional capacity of Sall proteins might be conditioned by posttranslational modifications. Thus, human SALL1 interacts with UBE2I, the homologue to ubiquitin conjugating enzyme 9, which promotes the binding of ubiquitin-like SUMO to target proteins. SALL1 is indeed sumoylated in vitro although the biological relevance of this modification remains to be explored (Netzer et al., 2002). Recently, it has been reported the capacity of protein kinase C to phosphorylate SALL1 at its repression motif, leading to the modification of its activity (Lauberth et al., 2007). No other posttranslational modifications have been described for other members of the Sall family. Most of the Sall proteins accumulate in the nucleus, with the exception of murine and chicken SALL3 and human SALL1 in certain cell types (Ma et al., 2001; Ma et al., 2002; Sweetman et al., 2003; Yamashita et al., 2007). SALL1 is engaged in interactions with other SALL family members and this could have dramatic functional consequences. Thus, cSALL3 promotes changes in the subcellular localization of cSALL1, which is retained in the cytoplasm in presence of cSALL3 through protein-protein interactions via the conserved polyQ domains (Sweetman et al., 2003). The conservation of the polyQ region in Sall proteins opens the possibility of interactions among all the paralogues, which could be important for the biological activity of the proteins.

The protein-protein interactions of SALL4 during embryonic development have also been studied in mouse and zebrafish limb development. In mice, SALL4 interacts with Tbx5, a T-box transcription factor involved in limb development, regulating the formation of the forelimb through the activation of FGF10 in a feed-forward mechanism (Koshiba-Takeuchi et al., 2006). In the hindlimb, an analogous interaction occurs with Tbx4, a factor necessary for hindlimb development. The interaction with Tbx5 seems to be important for the activation of Gia5 in the heart where, at the same time, SALL4 interferes with the capacity of Tbx5 to activate Nppa. How SALL4 can achieve its role as transcriptional activator and repressor, and how this is related to its capacity to bind heterochromatin and promote the methylation of histones remains unclear.

Sall proteins in stem cell and cancer biology

Murine SALL1 has a role in maintaining cellular pluripotency and proliferation. Thus, renal primordial cells in the ureteric bud epithelium and metanephric mesenchyme are able to produce nephrons and collecting ducts when induced from pluripotent embryonic stem cells. Only cells expressing high levels of SALL1 can reconstitute a three-dimensional kidney structure in an organ culture setting, indicating that renal progenitors with multipotent capacity require SALL1 (Osafune et al., 2006; Yamamoto et al., 2006). In these cells, SALL1 is not required for generation or differentiation of renal progenitors but for their proliferation or survival (Osafune et al., 2006). SALL1, expressed in embryonic stem cells, seems to contribute to the activation of Oct4 (Zhang et al., 2006) and SALL1a is necessary for the activation of FGFR2 downstream of Tbx5 during zebrafish pectoral fin development (Harvey and Logan, 2006). Whether this activation capacity is direct or indirect remains to be investigated.

Mouse and human SALL2 and SALL2 genes have been reported as tumour suppressors in several conditions. Thus, SALL2 was

**TABLE 2**

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene</th>
<th>Target Gene</th>
<th>Regulatory Region</th>
<th>Sall Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
<td>SALL2</td>
<td>p21</td>
<td>-2610 to -151 promoter</td>
<td>Activation</td>
<td>Li et al., 2004</td>
</tr>
<tr>
<td></td>
<td>SALL4</td>
<td>Bmi-1</td>
<td>-270 to -168 from enhancer</td>
<td>Activation</td>
<td>Yang et al., 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ATAA A/T A/T A/T</td>
<td>Repression</td>
<td>Yamashita et al., 2007</td>
</tr>
<tr>
<td>Drosophila</td>
<td>Sall4</td>
<td>Nanog</td>
<td>ES Enhancer; TTAGATTCCCTTTCC</td>
<td>Activation</td>
<td>Wu et al., 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Activation</td>
<td>Wu et al., 2006</td>
</tr>
<tr>
<td>C. elegans</td>
<td>Sall4</td>
<td>egl-5</td>
<td>e5-1; TTAGGT e5-2; TTAGCT</td>
<td>Repression</td>
<td>Toker et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>e5-3; ACACAA</td>
<td>Repression</td>
<td>Toker et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m3-1; ACACAA</td>
<td>Repression</td>
<td>Toker et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>m3-3; ACACAA</td>
<td>Repression</td>
<td>Toker et al., 2003</td>
</tr>
</tbody>
</table>

Only the sequences bound by Sall proteins are included. (1) Repression is inferred but not proved. (2) Repression is inferred, as SALL is not expressed at the same time than S15. Data were compiled from the references indicated in the right-hand column.
identified in a large screen looking for targets of the Large T antigen from the highly oncogenic mouse polyoma virus (Li et al., 2001). The interaction with Sall2 is important to suppress viral DNA replication and the growth of the virus (Li et al., 2001). Moreover, the presence of Sall2 in ovarian cancer cells inhibits their growth rate and their capacity to form colonies in soft agar. Some human ovarian carcinoma cell lines express low levels of SALL2 which, when re-introduced, results in a substantial reduction in the capacity of these cells to grow as tumours in nude mice. The control of cell growth and proliferation by SALL2 could be determined by its direct activation of p21 and Bax (Table 2; Li et al., 2004).

Human SALL2 is also necessary for the activation of a number of genes expressed after serum deprivation, a situation in which there is inhibition of cell growth. These genes are repressed in many types of prostate, blood and lung cancers, and their repression can predict the increased risk of cancer progression and death in human breast cancers (Table 3; Liu et al., 2001; Li et al., 2002; Nielsen et al., 2007). The molecular mechanisms underlying the roles of SALL2 are still unknown. The complex Sall4-Nanog could regulate the transcription of genes necessary for self-renewal, such as Sox2 and Oct4, in addition to their own transcription, constituting a regulatory circuit (Table 2 and Fig. 3: Wu et al., 2006). Similarly to Oct4, the reduction in Sall4 expression results in re-specification of ESCs to the trophoblast lineage, and this change is related to the expansion of Cdx2 expression (essential to the trophoblast lineage) into the inner cell mass of the blastocyst (Zhang et al., 2006; Elling et al., 2006). The co-occupancy of Nanog binding sites by the complex Nanog-Sall4 results in the activation of Nanog downstream genes by the over-expression of Sall4 (Wu et al., 2006). In this experimental setting, the up-regulation of the trophoblast lineage markers CDX2, HAND1 and GATA6 observed in addition,Embryonic Stem Cells (ESC) derived from Sall4-null embryos proliferate poorly with no aberrant differentiation, and no embryonic nor extraembryonic endoderm stem cell lines can be established from Sall4 mutant blastocysts (Elling et al., 2006; Sakaki-Yumoto et al., 2006). The role of Sall4 on ESC maintenance can be achieved through its interaction with Nanog, a homeodomain transcription factor identified as a protein able to sustain pluripotency in murine ESCs. The complex Sall4-Nanog could regulate the transcription of genes necessary for self-renewal, such as Sox2 and Oct4, in addition to their own transcription, constituting a regulatory circuit (Table 2 and Fig. 3: Wu et al., 2006). Similarly to Oct4, the reduction in Sall4 expression results in re-specification of ESCs to the trophoblast lineage, and this change is related to the expansion of Cdx2 expression (essential to the trophoblast lineage) into the inner cell mass of the blastocyst (Zhang et al., 2006; Elling et al., 2006). The co-occupancy of Nanog binding sites by the complex Nanog-Sall4 results in the activation of Nanog downstream genes by the over-expression of Sall4 (Wu et al., 2006). In this experimental setting, the up-regulation of the trophoblast lineage markers CDX2, HAND1 and GATA6 observed in the absence of human SALL4 could be indirect, occurring through the loss of POU5F1 expression (Zhang et al., 2006).

In concordance with its role in preserving the pluripotency of

**Fig. 3.** Schematic representation of genetic regulatory circuits in which Sall proteins and genes are involved during development. (A) Regulation by Sal and Senseless (Sens) of rhodopsin gene expression (Rh6 and Rh1) during the differentiation of the photoreceptor cell R8 (modified from Domingos et al., 2004a). (B) Regulation of sal expression by Distal-less, and requirement of Sal and Atonal in the formation of the Drosophila auditory organ, the Johnston’s organ (modified from Si Dong et al., 2003). (C) Regulation of Oct4 by Sall4, and requirement of Nanog, Oct4 and Sall4 during Embryonic stem cell maintenance and Epiblast development (modified from Zhang et al. (2006) and Pereira et al. (2006)). (D) Regulation of sal expression by Wg, EGFR and Dpp signalling during trachea development and its function in the specification of the dorsal trunk (modified from Kühnlein and Schuh (1996); Chen et al. (1998) and Chihara and Hayashi (2000)). (E) Regulatory interactions between the Dpp (BMP) downstream transcription factors Sal, Kni, Omb and Iro during Drosophila wing blade development (de Celis and Barrio, 2000; del Alamo Rodriguez et al., 2004; Cook et al., 2004). (F) Regulatory interactions occurring during Butterfly eye spot formation involving Sal, Engrailed and the candidate eyespot signalling molecules BMP and Wnt homologues. The dashed line indicates that Sal only represses En expression in some species, but not in others, generating either concentric rings or nested domains of Sal and En expression. Modified from Brunetti et al. (2001) and Monteiro et al. (2006).
stem cells in mice, the *Xenopus* homologue, *Xisall4*, was identified in a subtracted limb regeneration screen (King *et al.*, 2003). *Xisall4* transcripts are expressed during the early and middle phases of limb development and also in the fore- and hindlimb during regeneration-competent stages, suggesting that its activity could maintain blastema cells in an undifferentiated state (Neff *et al.*, 2005). Similarly, the chicken homologue *csall4* seems to keep neural crest cells in an undifferentiated stage (Barembaum and Bronner-Fraser, 2004). All these vertebrate homologues are expressed in the growing tail tip region rich in undifferentiated cells (Kohlhase *et al.*, 2002a; Barembaum and Bronner-Fraser, 2004; Neff *et al.*, 2005). The expression of human and murine *SALL4* during adulthood is restricted to testis and ovaries (Kohlhase *et al.*, 2002a; Kohlhase *et al.*, 2002b). Furthermore, microarray analysis shows that in the ovaries of newborn mice mutant for *Nbox*, a homeobox gene expressed in oocytes and required during oogenesis, *Sall4* is drastically downregulated, coinciding with a rapid loss of postnatal oocytes (Choi *et al.*, 2007).

The lack of proliferation observed in *Sall4* null mutant mouse cultured blastocysts and embryos *in vivo* (Sakaki-Yumoto *et al.*, 2006) might be related to the inefficient G1/S transition observed in ESCs, which could be linked to the interaction of Sal4 with CyclinD1 (Bohm *et al.*, 2007). A possible role of Sal4 in promoting cell proliferation could also be related to the expression of human *SALL4* in certain type of tumours. Accordingly, *SALL4* is upregulated in acute myeloid leukaemia (Table 3). The constitutive expression of *SALL4* may enable leukaemic blasts to acquire stem cell properties, such as self-renewal and/or lack of differentiation, and become leukaemia stem cells (Ma *et al.*, 2006; Cui *et al.*, 2006). This is probably achieved through the activation of the Wnt/β-catenin signalling pathway, as shown by the up-regulation of the Wnt targets c-Myc and CyclinD1 (Yang *et al.*, 2007). The activation of BMI-1, which plays an essential role in regulating adult, self-renewing hematopoietic stem cells and leukaemia stem cells (Yang *et al.*, 2007). The activation of *BMI-1* is associated to increased levels of histone methylation in the *BMI-1* promoter, but the mechanism relating the over-expression of *SALL4* and the hypermethylation of histones is still unknown (Yang *et al.*, 2007). A different role for SALL4 during tumourigenesis might be achieved through its role as a "caretaker" of chromosomal stability, which could be related to the capacity of SALL4 to bind to heterochromatic regions through its most C-terminal finger pair (Sakaki-Yumoto *et al.*, 2006; Bohm *et al.*, 2007). Human *SALL4* is epigenetically silenced in colorectal cancer aneuploid cells where *SALL4* promoter is more frequently hypermethylated than in diploid cells (Habano *et al.*, 2007). Thus, the absence of SALL4 might influence tumourigenesis by destabilization of chromosomes, but its upregulation might influence tumourigenesis by promoting proliferation.

**Sal proteins in cell specification and morphogenesis**

*Sal* genes are required for multiple developmental processes, suggesting that they engage in a variety of interactions and modify the expression of target genes in a context-dependent manner. We have attempted to classify these processes into several categories that include *sal*/invertebrate and vertebrate members, and will discuss in more detail some representative examples.

**Cell fate assignment**

The *Drosophila sal* and *sall* genes, and also several members of the *sal* family in other organisms, participate in a variety of cell-fate decisions during development, controlling the distinction between alternative cell fates or the implementation of a particular program of cell differentiation. Examples of the former are the function of the *sem-4* orthologue in *C. elegans* during the specification of touch receptor neurons (Miti *et al.*, 1993), and the function of *Drosophila sal* genes in the formation of the oocytes and strech receptors (Rusten *et al.*, 2001; Elstob *et al.*, 2001). In the first case Sem-4 regulates, by repression, the expression of the Hox gene *egl-5* and the LIM homeobox gene *mec-3*. These interactions are direct, because *Sem-4* binds to a common motif present in the *mec-3* and *egl-5* promoters (Table 2; Toker *et al.*, 2003). *Sem-4* also regulates the expression of the Hox genes *lin-39* and, in the absence of *sem-4*, the secondary vulval cell lineage is not correctly specified (Grant *et al.*, 2000). The relationships between Sal and Hox functions in the specification of cell identities is a common aspect of Sal proteins also observed in *Artemia* and *Drosophila*, although the interactions between *sal* and Hox genes vary in different developmental systems. Thus, the *Artemia sal*orthologue is expressed in the pre-segmental growth zone and in the segments that emerge from this zone (Cof et al., 2006). The loss of *sal* function, caused by RNA interference, results in a variety of homeotic transformations associated with the de-repression of different *Hox* genes in the corresponding segments, indicating that Sal regulates Hox gene expression (Cof et al., 2006). Because *Artemia sal* is expressed in all segments, and the observed homeotic transformations in knockdown animals are variable and stochastic, it was suggested that Sal function is related to the maintenance of spatial domains of Hox expression acting in transcriptional repression by chromatin modifications (Cof et al., 2006).

In contrast to this role in the maintenance of Hox expression, the *Drosophila sal* gene acts downstream of different Hox genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Disease</th>
<th>Expression</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sal1</td>
<td>Townes-Brooks Syndrome</td>
<td>Mutated, Deleted</td>
<td>Kohlhase <em>et al.</em>, 1996</td>
</tr>
<tr>
<td></td>
<td>Trophoblast tumours</td>
<td>Uregulated</td>
<td>Ma <em>et al.</em>, 2002</td>
</tr>
<tr>
<td></td>
<td>Sex hormone-producing tumours</td>
<td>Uregulated</td>
<td>Ma <em>et al.</em>, 2002</td>
</tr>
<tr>
<td></td>
<td>Wnt's Tumour</td>
<td>Downregulated</td>
<td>Cui <em>et al.</em>, 2006</td>
</tr>
<tr>
<td></td>
<td>Testicular carcinoma</td>
<td>Uregulated</td>
<td>Hoel-Hansen <em>et al.</em>, 2004</td>
</tr>
<tr>
<td></td>
<td>Congential Renal Dysplasia</td>
<td>Downregulated</td>
<td>Jain <em>et al.</em>, 2007</td>
</tr>
<tr>
<td></td>
<td>Congential Obstructive Nephropathy</td>
<td>Downregulated</td>
<td>Lianis, 2003</td>
</tr>
<tr>
<td>Sal2</td>
<td>Wnt’s Tumour</td>
<td>Uregulated</td>
<td>Niesen <em>et al.</em>, 2003; Subramaniam <em>et al.</em>, 2006</td>
</tr>
<tr>
<td></td>
<td>Prostate and Breast cancer</td>
<td>Downregulated</td>
<td>Liu <em>et al.</em>, 2007</td>
</tr>
<tr>
<td></td>
<td>Lung cancers</td>
<td>Downregulated</td>
<td>Ma <em>et al.</em>, 2001b</td>
</tr>
<tr>
<td></td>
<td>Colon and Prostate adenocarcinoma</td>
<td>Downregulated</td>
<td>Ma <em>et al.</em>, 2004</td>
</tr>
<tr>
<td></td>
<td>Ovarian carcinoma</td>
<td>Downregulated</td>
<td>Li <em>et al.</em>, 2004</td>
</tr>
<tr>
<td></td>
<td>Okihoro Syndrome</td>
<td>Mutated, Deleted</td>
<td>Al-Blaise <em>et al.</em>, 2002</td>
</tr>
<tr>
<td></td>
<td>Acute Myeloid Leukemia</td>
<td>Uregulated</td>
<td>Ma <em>et al.</em>, 2006</td>
</tr>
<tr>
<td></td>
<td>Synovial sarcomas</td>
<td>Uregulated</td>
<td>Cui <em>et al.</em>, 2006</td>
</tr>
<tr>
<td></td>
<td>Lymphoblastic leukaemia/lymphomas</td>
<td>Uregulated</td>
<td>Cui <em>et al.</em>, 2006</td>
</tr>
<tr>
<td></td>
<td>Aneuploid sporadic colorectal cancer</td>
<td>Downregulated</td>
<td>Habano <em>et al.</em>, 2007</td>
</tr>
</tbody>
</table>

Only the first references to TBS and OS are included. Data were compiled from the references indicated in the right-hand column.
in the haltere, labial and antennal imaginal discs. The distinction between wing and haltere relies in the function of the Ultrabithorax (Ubx) Hox gene. Among several other target genes, Ubx directly represses sal expression in the haltere, suppressing the positive input of Dpp on sal and contributing to the differences between these two structures (Weatherbee et al., 1998). Similarly, the Hox proteins Proboscipedia and Sex combs reduced directly the development of the proboscis by repressing sal expression in the labial disc (Abzhanov et al., 2001). In the antennal disc sal also acts downstream of genes specifying segmental identity, but its expression is activated rather than repressed by the combination of Distal-less and Homothorax (Dong et al., 2000). Interestingly, reminiscent to the loss of hearing associated to human SALL1 mutations (see below), loss of sal and salr in the antennal disc causes a severe reduction in the major Drosophila auditory organ, the Johnston's organ, and is associated with deafness (Dong et al., 2003). Finally, Drosophila Sal proteins also have homeotic functions independent of Hox genes during embryogenesis, acting to promote head versus trunk development (Jurgens, 1988).

The function of sal genes in specifying cell types does not always rely in their relationships with Hox genes. A clear example of a direct role of Sal proteins in cell differentiation occurs during Drosophila eye development, where Sal influences the formation of the R3, R4, R7 and R8 photoreceptors (Fig. 3A). Thus, Sal is required for the specification of R7 and the expression of R7 specific markers, the terminal differentiation of R8 and the regulation of photoreceptor specific rhodopsins, the correct specification of the R3/R4 pair of cells and establishment of planar cell polarity. Finally, Sal expression needs to be repressed later in these cells to inhibit their transformation to R7 fate (Mollereau et al., 2001, Domingos et al., 2004a; Domingos et al., 2004b). A similar function in cell fate specification can be operative in many cell populations during neural system development, because sal and sall genes are expressed predominantly in the developing nervous system in a variety of organisms. An interesting example of Sal functions in cell fate decisions is the formation of a particular type of sensory organs in Drosophila, where Sal operates as a switch between two cell types induced by EGFR activity, the oenocytes and the precursors of the pentsaculopodial sensory organ. This organ is formed by five sensory units derived from five chordotonal organ precursors (COPs; Gould et al., 2001). The oenocytes form around the most dorsal COP and express high levels of Sal. The absence of Sal results in the lack of oenocytes form around the most dorsal COP and express high levels of sal. (Abzhanov et al., 2001). This role of Sal is reminiscent of the role of Sall4 in the decision between inner cell mass and trophoblasts in the mouse (Elling et al., 2006).

Regional specification

Another common aspect of sal functional requirements in different organisms occurs during the subdivision of a cell population into smaller developmental units, which we refer to as “regional specification”. This feature of sal function was first identified for the Drosophila sal and salr genes during the growth and patterning of the wing imaginal disc, an epithelial tissue that differentiate during metamorphosis the fly wing and thorax. The Sal/Salr proteins act in the wing blade as transcription factors conferring regional identity to the central part of the wing, linking the activity of the secreted molecule Dpp to pattern formation (de Celis et al., 1996). Thus, sal and salr are expressed in a central domain of cells in the wing region of the disc, where they participate to the patterning of the wing blade (Fig. 3E). The Dpp pathway directly regulates the expression of sal and salr in this territory, and they direct the localisation of characteristic wing pattern elements, the veins, by regulating the expression of the vein-specific genes of the kniops and Iroquois gene complexes (Fig. 3E; de Celis and Barrio, 2000). In the case of the Iroquois genes, Sal/Salr repress their expression in all cells not exposed to Hedgehog signalling, confining Iroquois expression to the posterior L5 provein territory. The relationship between Sal/Salr and the kniops genes is more complex, because their expression is activated in the domain where Sal/Salr levels are lower in anterior cells, and repressed by higher levels of Sal/Salr in the rest of the wing (de Celis and Barrio, 2000). In addition to its pattern-promoting function, Sal and Salr are also required for cell viability, cell proliferation and epithelial integrity of the cell population where they are expressed (de Celis et al., 1996; Milan et al., 2002).

Several vertebrate Sall proteins are also expressed in the growing limbs, where they could also function to provide territorial identities to mesenchymal cell populations. Xenopus Xsall4 is expressed in developing hind- and forelimbs in a dynamic temporal and spatial pattern that first is confined to the distal half of the limb bud, later is excluded from proximal-posterior and anterior regions of the bud, and finally becomes restricted in the future autopod to six interdigital domains (Neff et al., 2005). In chicken, csall1 and csall2 are also expressed in developing limbs (Farrell and Munsterberg, 2000; Farrell et al., 2001). The expression of csall1 is observed continuously through the distal limb mesenchyme and the apical ectodermal ridge (Capdevila et al., 1999; Farrell and Munsterberg, 2000). In contrast, csall2 displays a dynamic temporal and spatial pattern of expression that is differentially regulated in wing and leg primordia, being in both cases detected mainly in the posterior-distal mesenchyme (Farrell et al., 2001). In zebrafish, sal1a and sal4 are expressed in developing limb-like structures, the pectoral fins (Camp et al., 2003). The expression of sal4 is first detected through the fin bud mesenchyme, and as its development proceeds, sal4 transcripts are accumulated at the distal tip of the fin. Loss-of-function experiments using sal4 morpholinos showed that this gene is required for the outgrowth of pectoral fins and the formation of its distal structures (Harvey and Logan, 2006). The gene sal1a is expressed in both the mesenchyme and the ectoderm (Camp et al., 2003; Harvey and Logan, 2006), with highest levels in the distal fin bud in a pattern comparable with the observed for sal1a in limb buds in mouse and chick (Farrell and Munsterberg, 2000; Buck et al., 2001). Similar to sal4, sal1a morphants develop truncated and often absent pectoral fins, indicating a requirement for fin bud outgrowth. In the double sal4/sal1a morphant embryo the fin bud is initially formed, but it fails to develop further due in part to the absence of FGF10 expression (Harvey and Logan, 2006). Similarly to other vertebrate orthologs, mouse Sal1a is also expressed in the developing limb, in a pattern that evolves during limb development from most of the mesenchyme and ectoderm to the tips of the digits and interdigital territories (Buck et al., 2001). Interestingly, distal limb defects, such as bifid thumbs and loss of
thumbs, as well as polydactyly are characteristic abnormalities of TBS and OS (Kohlhase et al., 1998; Kohlhase et al., 2002b; Al-Baradie et al., 2002 and see below).

A conceptually similar function of Sall proteins during regional specification is observed during the development of eyespots in the wings of butterflies. Eyespots are pigmentation patterns characteristic of many butterflies and moth wings. The formation of the eyespot is controlled from its centre, the focus, which induces surrounding cells to acquire different colour fates. In Bicyclus anynana, the Sal homolog is expressed in the focus from its onset, and later in several concentric rings outside the focal region (Brunetti et al., 2001; Monteiro et al., 2006). Interestingly, the Engrailed homolog is expressed in an outer ring outside the domain of Sal expression, suggesting that regulatory interaction between Sal and Engrailed orthologs participate in the elaboration of gene expression domains. This interaction is reminiscent to the repression of Iroquois expression by Sal observed in the Drosophilawing, and in both case leads to the creation of adjacent domains of gene expression (Fig. 3E-F).

Organogenesis

During organogenesis, cells from distinct origins, or with different developmental programs, must be integrated to form functional structures. The activity of sall genes is required in several internal organs such as the heart and kidney in vertebrates and the tracheae (respiratory tubes) in Drosophila. A conserved feature among vertebrates is the expression of sall genes during the development of the kidney. Thus Xenopus Xsall4b and zebratfish sal1a, are expressed in the pronephric ducts, and chicken csall3 is expressed in the mesonephros (Onuma et al., 1999; Farrell et al., 2001; Camp et al., 2003). The function of sall during kidney development has been mainly studied using Sall1 knockout mice. The development of the vertebrate metanephros implies mutual inductive interactions between the ureteric bud and the metanephric mesenchyme. In this manner, the invasion of the mesenchyme by the ureteric bud epithelia, and its accompanying branching morphogenesis to form the collecting ducts and urethra, is induced by the mesenchyme, and reciprocally, the ureteric bud induces mesenchymal aggregation around the bud tip and mesenchyme-to-epithelial conversion to form the renal vesicle (Dressler, 2006). The Sall1 mice gene is exclusively expressed in the metanephric mesenchyme prior to bud invasion, and this expression is maintained in the mesenchyme condensing around the ureteric bud tips. The function of Sall1 is required to promote ureteric bud invasion, which failure causes a subsequent collapse of tubule differentiation by the mesenchyme. In this manner, in Sall1-null mice the metanephric mesenchyme and the ureteric bud are formed, but the bud fails to invade the mesenchyme (Nishinakamura et al., 2001). FGFR signalling could regulate the expression of Sall1 in the early metanephric mesenchyme, as double mutant FGFR1/FGFR2 mice display renal aplasia and the expression of Sall1 is absent from the rudimentary mutant metanephric mesenchyme (Poladia et al., 2006). It is not clear what is the exact role of Sall1 in the mesenchyme, because direct targets activated or repressed by Sall1 in this tissue have not yet been identified. In contrast to the requirement of Sall1 during vertebrate kidney development, the function of Drosophila sal genes is not operative in the fly kidney equivalent, the Malpighian tubules, even though the formation of this structure also includes interactions between ectodermal epithelial buds and mesenchymal mesodermal cells (Denholm et al., 2003).

The formation of the Drosophila tracheal system involves a number of cellular activities similar to vertebrate kidney formation, such as oriented cell migration, branching morphogenesis and inductive signalling from independent tissues (Metzger and Krasnow, 1999; Affolter and Shilo, 2000). Trachea formation is initiated from ectodermic placodes that invaginate into the underlying mesoderm and undertake a complex branching pattern to form a three-dimensional network of tubes. Loss of sal functions results in a variety of phenotypes including the formation of ectopic placodes and the lack of the dorsal trunk (Kühnlein and Schuh, 1996). The first phenotype suggest an early role of Sal in suppressing tracheal fate, whereas the loss of the dorsal trunk is due to faulty cell specification within the tracheal placodes (Kühnlein and Schuh, 1996; Franch-Marro and Casanova, 2002). The failure to form the dorsal trunk in sal mutants, caused by the lack of antero-posterior migration and fusion into a trunk of the dorsal trunk primordia, is reminiscent of the requirement of Sall1 in promoting ureteric bud invasion, although during tracheal development the requirement of Sal is cell autonomous in the migrating cells. Wnt and EGFR signalling induce and maintain, respectively, the expression of Sal in the dorsal part of all tracheal placodes, in a region that initially encompasses the primordia of the dorsal branch and the dorsal trunk (Fig. 3D; Chihara and Hayashi, 2000). Latter, Sal expression is restricted to the dorsal trunk primordia, where it is present after the connection between the posterior and anterior dorsal trunk branches from adjacent placodes (Kühnlein and Schuh, 1996; Wappner et al., 1997; Chen et al., 1998). The downregulation of Sal in the dorsal branch primordia is mediated by repression of Knirps, acting directly on a sal regulatory element (Chen et al., 1998). The repression of sal expression by Knirps is a requisite for normal dorsal branch morphogenesis. In this manner, Sal and Knirps became expressed to adjacent territories, the primordia of the dorsal branch and the dorsal trunk, which will follow different developmental fates (Fig. 3; Chen et al., 1998; Franch-Marro and Casanova, 2002).

Sall genes in disease

Human SALL1 mutations are associated to TBS, an autosomal dominant group of malformations characterized by imperforate anus, triphalangeal and supernumerary thumbs, dysplastic ears and sensorineural hearing loss (Kohlhase et al., 1998; Surka et al., 2001; reviewed by Powell and Michaelis, 1999). So far, 56 family mutations in SALL1 associated to TBS disorders are characterized (Botzenhart et al., 2007 and references therein), most of them located between the polyQ domain and the C-terminal part of the zinc finger domain 2. Therefore, it is likely that the TBS patients express a truncated protein able to interact to other SALL proteins via the polyQ region and block their function. Only two of the reported cases would produce truncated proteins lacking the glutamine-rich domain, although in both cases the mutant proteins contain the initial C2HC zinc finger motif and the N-terminal repressor domain, indicating that the glutamine domain is not absolutely required for typical TBS symptoms (Kohlhase et al., 1999b; Botzenhart et al., 2007). Some patients present deletions of the whole SALL1 gene and, in fact, they show a rather mild TBS phenotype, reinforcing the idea that the haploinsuffi-
ciency is not enough to cause the severe classical TBS symptoms (Borozdin et al., 2006). Confirming the role of Sall1 in kidney formation, SALL4 expression is reduced in patients with congenital dysplastic kidneys, a major cause of renal failure in infants (Jain et al., 2007), as well as in congenital obstructive nephropathy, a common disease affecting foetuses and young children (Table 3; Liapis, 2003). Mice homozygous for Sall1 show kidney agenesis and die in the perinatal period. The abnormal kidneys result from incomplete ureteric bud outgrowth, deficient mesenchyme tubule formation and apoptosis of the mesenchyme (Nishinakamura et al., 2001; reviewed by Nishinakamura and Osafune, 2006). However, in contraposition to the dominant effect shown in human TBS patients, heterozygous Sall1 mutants do not show any phenotype. Interestingly, the expression in mice of truncated Sall1 lacking all the double zinc fingers but preserving the N-terminal part of the protein, recapitulate remarkably all the abnormalities found in human TBS, supporting the idea of TBS being caused by the dominant negative effect of truncated SALL1 proteins (Kiefer et al., 2003).

Mutations in SALL4 are involved in the autosomal dominantly inherited human OS (Al-Baradie et al., 2002; Kohlhase et al., 2002b). This malformation syndrome is characterized by radial defects of the upper limbs and by Duane anomaly, a rare form of strabismus, also associated with hearing loss. There is large intra- and interfamilial variability in the clinical features of patients with SALL4 mutations and patients can be miss-diagnosed, being the mutational analysis of SALL4 important for the interpretation of the symptoms (Kohlhase et al., 2002b; Brassington et al., 2003; Borozdin et al., 2004; Kohlhase and Holmes, 2004; Kohlhase et al., 2005). In contrast to SALL1 mutants causing TBS, the mutations found in SALL4-related syndromes do scatter along the gene, indicating that the clinical features are caused by loss-of-function and haploinsufficiency, rather than by a dominant negative effect of truncated proteins (Borozdin et al., 2004).

Some OS patients also show severe growth retardation, also seen in patients affected by TBS that might indicate pituitary dysfunctions associated with SALL4 mutations (Kohlhase et al., 2005; Miertus et al., 2006). A plausible explanation for the features shared by OS and TBS, is the high rate of spontaneous miscarriages in TBS patients. The C-terminally truncated SALL1 protein produced by removing the N-terminal part of the protein, recapitulate remarkably all the abnormalities found in human TBS, supporting the idea of TBS being caused by the dominant negative effect of truncated SALL1 proteins (Kiefer et al., 2003).

The most prominent expression domain of Sall2 is the brain, raising the possibility for a function in this organ (Kohlhase et al., 2000). However, no behavioural defects or any other anomalies were reported.

Sall3 deficient mutant mice present malformation in organs necessary for normal feeding behaviour, such as the palate, the epiglottis, the tongue, and the corresponding cranial nerves. Homozygous animals die shortly after birth because their inability to feed properly, but the heterozygotes are fertile and indistinguishable from wild type (Parrish et al., 2004). In a similar way to Sall4, Sall3 could also be required during the specification of embryonic versus throphoblast stem cells (Oghane et al., 2004).

Concluding remarks

The understanding of Sall proteins function and sall genes regulation is still incomplete, but the use of different experimental models and the combination of biochemical and genetic approaches is unravelling many significant aspects of their biology. The existence of many Sall interacting proteins and the likely variety of Sall mechanisms of transcriptional regulation confers a great versatility to Sall function. Similarly, it is expected that the existence of multiple cis-regulatory regions in sall genes is a general trend, contributing to the deployment of sall expression in multiple developmental contexts under the regulation of a diversity of transcriptional regulators. These two characteristics most likely determine the multiple requirements identified for Sall function during multicellular development and the variety of tissues where they are expressed. Future research avenues into Sall biology will certainly include the identification of additional Sall-interacting proteins, the analysis of Sall posttranscriptional modifications and their functional consequences, and the study of the molecular mechanism of transcriptional regulation. The identification of Sall downstream genes, and the characterisation of their mode of regulation are expected to contribute fundamentally to the understanding of the biological requirements of Sall during animal development.

Acknowledgements

We are grateful to F. C. Kafatos, A. García-Bellido and M. Ashburner in which laboratories we carried out a great part of our research on the Drosophila spalt gene complex. We thank J. D. Sutherland, A. Talamillo, J. Sánchez, R. Cantera and T. E. Rusten for the critical reading of this manuscript, and the criticism from two anonymous reviewers that helped to improve the paper. R.B. would like also to acknowledge the continuous contribution and inspiration from her colleagues R. Cantera and T. E. Rusten and the help with the protein analysis from L. Sánchez. J.F.dC. is supported by an institutional grant from the Ramón Areces Foundation to the Centro de Biología Molecular “Severo Ochoa” and by the Ministerio de Educación y Ciencia (BFU2006-06501). R.B. belongs to the Cajal
References


ELSTOB, P. R., V. BRODU and A. P. GOULD (2001) *spalt*-dependent switching between two cell fates that are induced by the *Drosophila* EGF receptor. *Development* 128: 723-732.


YAMAMOTO, M., L. CUI, K. JOHKA, K. ASANUMA, Y. OKOUCHI et al. (2006) Branching ducts similar to mesonephric ducts or ureteric buds in teratomas
Further Related Reading, published previously in the *Int. J. Dev. Biol.*


See our recent Special Issue *Limb Development* edited by Juan Hurlé and Juan Carlos Izpisua-Belmonte at: http://www.ijdb.ehu.es/web/contents.php?vol=46&issue=7

**Development of Johnston’s organ in Drosophila**
Daniel F. Eberl and Grace Boekhoff-Falk
*Int. J. Dev. Biol.* (2007) 51: 679-687

**Expression of csal1 in pre limb-bud chick embryos**
Dylan Sweetman, Terence G. Smith, Elizabeth R. Farrell and Andrea Münsterberg
*Int. J. Dev. Biol.* (2005) 49: 427-430

**Developmental basis for vein pattern variations in insect wings.**
José F De Celis and Fernando J Diaz-Benjumea
*Int. J. Dev. Biol.* (2003) 47: 653-663

**Positioning and differentiation of veins in the Drosophila wing.**
J F De Celis
*Int. J. Dev. Biol.* (1998) 42: 335-343

**Regulation of cell migration during tracheal development in Drosophila melanogaster.**
Valérie Petit, Carlos Ribeiro, Andreas Ebner and Markus Affolter

**The ash2 gene is involved in Drosophila wing development.**
Montserrat Amorós, Montserrat Corominas, Peter Deák and Florenci Serras

---

