

Review

# The vertebrate *spalt* genes in development and disease<sup>☆</sup>

Dylan Sweetman<sup>\*</sup>, Andrea Münsterberg

*School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK*

Received for publication 1 September 2005; revised 1 February 2006; accepted 5 February 2006

Available online 20 March 2006

## Abstract

The *spalt* proteins are encoded by a family of evolutionarily conserved genes found in species as diverse as *Drosophila*, *C. elegans* and vertebrates. In humans, mutations in some of these genes are associated with several congenital disorders which underscores the importance of *spalt* gene function in embryonic development. Recent studies have begun to cast light on the functions of this family of proteins with increasing understanding of the developmental processes regulated and the molecular mechanisms used. Here we review what is currently known about the role of *spalt* genes in vertebrate development and human disease.

© 2006 Elsevier Inc. All rights reserved.

**Keywords:** Spalt; SALL; Sall; Townes–Brocks; Holt–Oram; Okihiro

## Introduction

The *spalt* genes were originally identified in *Drosophila*, which has two members of the family, *spalt major* (*salm*) and *spalt related* (*salr*). These have important roles in processes as diverse as homeotic specification of the embryonic termini (Jürgens, 1988; Kuhnlein et al., 1994), wing patterning (de Celis and Barrio, 2000), sensory organ development (de Celis et al., 1999), tracheal system development (Kuhnlein and Schuh, 1996) and specification of photoreceptors (Domingos et al., 2004a,b). Vertebrate homologues of *spalt* have been shown to be involved in normal development and are implicated in several human genetic disorders. Here we will review the developmental roles of this gene family with reference to both animal model organisms and human disease. The *spalt* genes are known to be mutated in several human congenital syndromes and studies of the mutations associated with these syndromes have been critical in our current understanding of *spalt* gene function. Finally, we

will discuss how the developmental and disease studies relate to what is currently known about the structure/function of these proteins. We will also discuss the relationships between these genes and propose a unified nomenclature system for future use (see box, Fig. 1 and Table 1).

## Spalt proteins in development

Recent work has demonstrated that *spalt* genes are required for the normal development of the limbs and nervous system and several organs including the kidney and heart. In this section, we discuss which members of this gene family are necessary in these systems, the mechanisms by which they are thought to act and highlight some of the as yet unanswered questions about these genes in development.

### Limb development

The association of several syndromes affecting normal limb development with mutations in the *SALL* genes suggests a critical role of these proteins in the limb. Several members of the *spalt* family are known to be expressed during limb development including *Sall1*, *Sall3* and *Sall4* in mouse (Buck et al., 2001; Kohlhase et al., 2002a; Ott et al., 1996, 2001), *csall1*, *csall3* and *csall4* in chick (Barembaum and Bronner-Fraser, 2004; Farrell and Münsterberg, 2000; Farrell et al., 2001) and *Xsall3* and *Xsall4a* in *Xenopus* (Holleman et al., 1996; Neff

<sup>☆</sup> Genbank accession numbers for *spalt* mRNAs: *SALL1*: NM\_002968; *SALL2*: NM\_005407; *SALL3*: NM\_171999; *SALL4*: NM\_020436; *Sall1*: AAH62937; *Sall2*: NM\_015772; *Sall3*: NM\_178280; *Sall4*: NM\_201395; *csall1*: NM\_204707; *csall3*: NM\_204647; *csall4*: AAR01968; *sall1*: AJ293862; *sall2*: AJ293863; *sall3*: AJ293864; *Xsall1*: AF310007; *Xsall3*: L46583; *Xsall4a*: AB030827; *Xsall4b*: AAH82637; *Xsall2*: AY508953; *Medaka sal*: AAB51127; *salm*: NM\_164966; *salr*: Y07653.

<sup>\*</sup> Corresponding author. Fax: +44 1603 592250.

E-mail address: [d.sweetman@uea.ac.uk](mailto:d.sweetman@uea.ac.uk) (D. Sweetman).

## Nomenclature

Vertebrate spalt proteins fall into four groups, with four genes in human called *SALL1-4* (Kohlhase et al., 1996, 1999, 2002b). Four genes have been identified in mouse, *Sall1-4* (Buck et al., 2001; Kohlhase et al., 2000, 2002a; Ott et al., 1996, 2001), three in chick (Barembaum and Bronner-Fraser, 2004; Farrell and Münsterberg, 2000; Farrell et al., 2001), five in *Xenopus* (Hollemann et al., 1996; Neff et al., 2005; Onai et al., 2004; Onuma et al., 1999), three in zebrafish (Camp et al., 2003) and one in Medaka (Koster et al., 1997). A phylogenetic tree indicating the relationship of these homologues is shown in Fig. 1. The nomenclature for these genes is not entirely consistent and for the purposes of clarity in this review we will refer to all vertebrate *spalt* genes as *spalt like* (abbreviated *sall*). Table 1 shows our proposed nomenclature based on this phylogeny. Human and mouse genes retain their current names while those of other species have been changed to correspond to these genes. As Fig. 1 shows, the *spalt* genes fall into four distinct groups with the *SALL1* and *SALL3* groups being most closely related. This phylogeny suggests that an ancestral *spalt* diverged to produce the *SALL2* group with further duplications producing the *SALL4* group and then the *SALL1* and *SALL3* groups. The status of the zebrafish *sall1b* is least clear. Comparison of protein sequences shows that it has 55.2% identity with zebrafish *sall1* and 52.4% identity with zebrafish *sall3* and it is possible that this gene is of a subfamily unique to zebrafish.

et al., 2005). These genes are expressed in overlapping patterns and may well affect limb patterning and development in a co-ordinated way. Very little is known about the regulation of *spalt* genes in the developing limb; however, it has been shown that endogenous *csall1* expression in the distal part of the limb bud requires FGF and Wnt signaling (Farrell and Münsterberg, 2000). In the proximal limb, *csall1* expression can be induced ectopically by BMP-2 although this may reflect an induction of distal limb fate in this case rather than endogenous regulation (Capdevila et al., 1999). The range of possible interactions between different spalt proteins will vary across the limb bud depending on the precise combination of spalt proteins present and it is tempting to speculate that this could modulate function and produce different activities in different areas. For example in part of the chick limb bud, *csall1* is coexpressed with *csall3*. It is known that the chick protein *csall3* is cytoplasmic and can remove full-length spalt protein from the nucleus, providing another possible level of regulation of protein activity (Sweetman et al., 2003). In the mouse limb, no defects are observed in loss of function *Sall1* mutations but patterning is disrupted in

the presence of truncated Sall1 protein (McLeskey Kiefer et al., 2003; Nishinakamura et al., 2001). These findings are discussed further in the section on SALL1 in Townes–Brooks syndrome (TBS) below. In *Xenopus*, *Xsall4a* is expressed both during normal limb development and regeneration (Neff et al., 2005). Recent evidence has also shown that in zebrafish *sall1* and *sall4* are regulated by *tbx5* in pectoral fin development and are required for regulation of FGF signaling (Harvey and Logan, 2006). The control of these genes by *tbx5* may explain the clinical overlap between Holt–Oram syndrome and Okihiro syndrome with mutations in *TBX5* and *SALL4* giving remarkably similar phenotypes (see SALL4 in developmental disorders).

The details of how spalt proteins affect limb patterning are unclear but one possibility is that they interact genetically with members of the *Iroquois* gene family. These genes are expressed during vertebrate limb development (Houweling et al., 2001; Zulch et al., 2001) and in *Drosophila* it is known that interactions between the *spalt* and *iroquois* gene families are required in wing patterning (de Celis and Barrio, 2000).

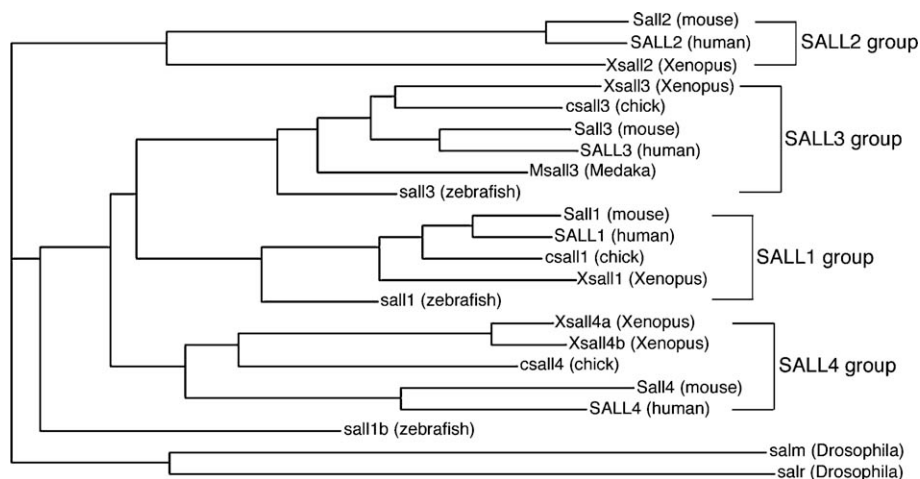


Fig. 1. Clustal alignment of vertebrate spalt protein sequences. Four orthologous groups can be identified based on sequence similarity, which are termed SALL1, SALL2, SALL3 and SALL4 groups based on the human gene.

Table 1  
Proposed new nomenclature of the vertebrate *spalt* like genes

Current name	Species	Genbank accession no.	Reference	Proposed name
<i>csall1</i>	Chick	NM_204707	(Capdevila et al., 1999; Farrell and Münsterberg, 2000)	<i>csall1</i>
<i>csal3</i>	Chick	NM_204647	(Farrell et al., 2001; Sweetman et al., 2003)	<i>csall3</i>
<i>csal4</i>	Chick	AAR01968	(Barembaum and Bronner-Fraser, 2004)	<i>csall4</i>
<i>Xsal1</i>	<i>Xenopus leavis</i>	L46583	(Hollemann et al., 1996)	<i>Xsall3</i>
<i>Xsal1</i>	<i>Xenopus leavis</i>	AF310007	Unpublished	<i>Xsall1</i>
<i>Xsal3</i>	<i>Xenopus leavis</i>	AB030827	(Onuma et al., 1999)	<i>Xsall4b</i>
<i>Xlsall4</i>	<i>Xenopus leavis</i>	AAH82637	(Neff et al., 2005)	<i>Xsall4a</i>
<i>XsalF</i>	<i>Xenopus leavis</i>	AY508953	(Onai et al., 2004)	<i>Xsall2</i>
<i>sall1a</i>	Zebrafish	AJ293862	(Camp et al., 2003)	<i>sall1</i>
<i>sall1b</i>	Zebrafish	AJ293863	(Camp et al., 2003)	<i>sall1b</i>
<i>sall3</i>	Zebrafish	AJ293864	(Camp et al., 2003)	<i>sall3</i>
<i>sal</i>	Medaka	AAB51127	(Koster et al., 1997)	<i>Msall3</i>

### Heart development

In the developing heart, expression has been reported of *Sall1* in mouse (Buck et al., 2001; Ott et al., 2001), *sall1* in zebrafish (Camp et al., 2003) and *csall1* in chick (Sweetman et al., 2005). Together with the presence of heart defects in TBS (Kohlhase et al., 2003a) (see *SALL1* in Townes–Brocks syndrome), this is strongly suggestive of a role in heart development. Interestingly in the developing chick embryo, *csall1* expression is observed not only in the sinus venosus of the heart itself but also in the pharynx at the same axial level (Sweetman et al., 2005). The pharynx is known to be involved in the inductive events required for cardiac neural crest specification and disruptions to this population of cells can result in ventral septal defect and Tetralogy of Fallot (Creazzo et al., 1998; Farrell et al., 1999). These defects are also occasionally seen in TBS. This would be consistent with an additional indirect role of *csall1* in heart development via the regulation of inductive events mediated by the pharynx.

### Neural development

Many *spalt* genes are expressed in the developing nervous system; however, only recently have studies begun to cast light on the function of these genes in neural development. In *Xenopus*, *Xsall2* has been shown to be required for the determination of the forebrain/midbrain and for repressing the acquisition of a more caudal midbrain/hindbrain fate. It is interesting to note that in this context *Xsall2* acts by repressing canonical Wnt signaling which acts to posteriorize the neural tissue in the absence of *Xsall2* (Onai et al., 2004). This is in contrast to the reported ability of *Sall1* to enhance canonical Wnt signaling (Sato et al., 2004) and it remains to be determined

whether this difference is due to cellular context or to different properties of these *spalt* proteins. Another report shows that in the chick embryo *csall4* expression is required for the correct migration of neural crest cells into the sensory ganglia. Neural crest cells overexpressing *csall4* do not migrate normally into the ganglia and do not express differentiation markers. The authors suggest that *csall4* is required for these cells to remain undifferentiated and that this could affect the ability of these cells to form neurons (Barembaum and Bronner-Fraser, 2004). Cranial nerve development is also disrupted in a *Sall3* null mouse model. In these mice, the nerves that innervate the oral cavity do not form normally, particularly the glossopharyngeal nerve (Parrish et al., 2004). Together, these studies demonstrate important and varied roles for the *spalt* genes in the normal development of the nervous system.

### Kidney development

Mice homozygous for a null mutation in *Sall1* die perinatally due to renal agenesis. In these animals, the uterine bud does not grow out, tubules do not form and the metanephric mesenchyme undergoes apoptosis. Although these defects occur in the nephrogenic mesenchyme, the mesenchyme itself can be rescued by wild type tissue. Therefore, it is the inductive signals leading to uterine bud invasion that fail under these conditions (Nishinakamura et al., 2001). In another mouse model where a truncated *Sall1* protein is expressed, homozygous mice have renal agenesis, while in heterozygous mice with one wild type *Sall1* allele renal cystic hypoplasia results (McLeskey Kiefer et al., 2003). Another link between *SALL1* and kidney development comes from studies of patients with Townes–Brocks syndrome which is associated with kidney defects (Powell and Michaelis, 1999). In other species, *spalt* genes are also expressed in the developing kidney with *sall1a* in zebrafish and *Xsall4b* in *Xenopus* expressed in the pronephric ducts (Camp et al., 2003; Onuma et al., 1999) while in the chick embryo *csall3* is expressed in the mesonephros (Farrell et al., 2001). The range of different *spalt* genes expressed in kidney development in different species suggests some divergence in function with other *spalt* genes substituting for *SALL1* in both *Xenopus* and chick embryos.

### Future directions

The regulation of *spalt* gene expression has not yet been examined in detail in most of the organ systems known to require these genes. The regulation of *Drosophila spalt* has been studied and shown to respond to a gradient of the TGF $\beta$  homologue *dpp* (Barrio et al., 1999; Smith, 1996) and the T box gene *omb* (del Alamo Rodriguez et al., 2004); in vertebrates the only known regulators are Shh in the Medaka midbrain/hindbrain boundary (Koster et al., 1997) and FGF and Wnt signaling in the distal limb bud (Farrell and Münsterberg, 2000) although BMP signaling can induce ectopic *csall1* expression in the proximal limb (Capdevila et al., 1999). Further elucidation of the spatio-temporal regulation will be critical in the understanding of how these proteins are able to affect the

patterning of different tissues as downstream effectors of developmental signaling cascades.

In *Drosophila* wing development, *sal* regulates expression of the *Iroquois* complex and *knirps* (de Celis and Barrio, 2000) and is itself controlled by the T box gene *optomotor blind* (del Alamo Rodriguez et al., 2004). In vertebrate limb development, homologues of these genes are known to be required for normal patterning and mutations in one vertebrate T box gene, *TBX5*, also lead to Holt–Oram syndrome, as can mutations in *SALL4*. The discovery that *tbx5* regulates *sall1* and *sall4* in zebrafish pectoral fin development (Harvey and Logan, 2006) and that *Tbx5* and *Sall4* interact in mouse heart and limb development (Koshiba-Takeuchi et al., 2006) underscores the conservation of *spalt* gene function across species. In the future, it will be fascinating to determine if other functional relationships known from *Drosophila* are also conserved in vertebrates, such as the interaction between *spalt* and *Iroquis* gene families, and what implications this has for normal limb development.

Another area that is currently not well understood is the role of *spalt* genes in ear development. This is of particular interest as mutations in some human *spalt* genes lead to outer ear and auditory defects. Unpublished data from chick embryos indicate that *csall4* is required for the invagination of sensory placodes (M. Bronner Fraser, personal communication) and this will be an exciting area for further investigation.

### Spalt genes in human disease

One of the striking aspects of *spalt* biology is the clear association of mutations within this gene family with congenital syndromes affecting limb, ear, kidney and heart development. Much of the work on developmental aspects of *spalt* gene function has been based on information from studies of these patients. In this section, we describe the relationships between these syndromes and mutations in the human *spalt* genes and how these have provided valuable insights into the function of these proteins in normal development.

#### *SALL1* in Townes–Brocks syndrome

Townes–Brocks syndrome (TBS) is an autosomal dominant disorder characterized by preaxial polydactyly and triphalangeal thumbs, external ear defects and sensorineural hearing loss, imperforate anus, kidney defects and heart defects including Tetralogy of Fallot and ventricular septal defects (Powell and Michaelis, 1999) with some rare cases of patients also showing a variety of other phenotypic anomalies (Botzenhart et al., 2005). Mutations in *SALL1* have been shown to result in TBS (Kohlhase et al., 1998). Most of these mutations cause premature stop codons in the 5' region of the gene which, if transcribed, would lead to the production of truncated forms of the *SALL1* protein that terminate after the glutamine rich region and around the region of the first set of C2H2 zinc fingers leading to a predicted protein containing either none or only one of the C2H2 domains (Botzenhart et al., 2005; Kohlhase, 2000) (see Fig. 2). The question of whether or not these mutations actually lead to the expression of truncated *SALL1* proteins has

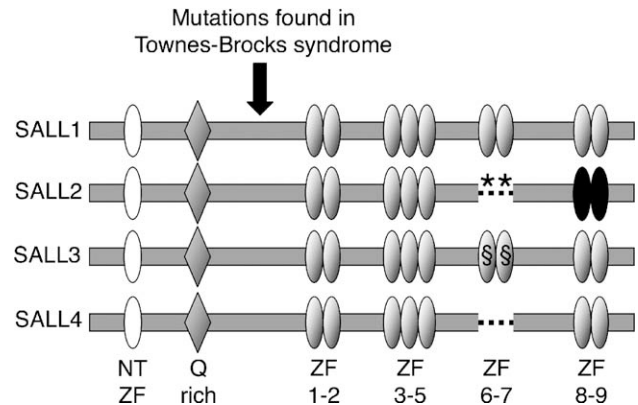


Fig. 2. Schematic representation of the predicted structure of human *spalt* proteins. The N terminal zinc finger (NT ZF) is of the C2HC type while the other zinc fingers (ZF 1–9) are of the C2H2 type. The Q rich region mediates interactions between members of this protein family. The mutations found in TBS are clustered between the Q rich region and the first set of C2H2 zinc fingers (ZF 1–2). The double asterisks (\*\*) indicate the position of the zinc fingers present in *Xsall2* but not in *SALL2*. The double sign (§§) indicates the zinc fingers missing in the orthologous gene *csall3*. The black circles indicate the C terminal zinc fingers in *SALL2* which are not homologous to those in other *spalt* proteins.

not been resolved and initial reports suggested that TBS is caused by haploinsufficiency at the *SALL1* locus caused by nonsense-mediated decay of the mutant transcript (Kohlhase et al., 1998). This hypothesis did not explain the clustering of mutations and several studies have suggested an alternative possibility, namely that truncated proteins produced from the *SALL1* locus can act in a dominant negative manner, interfering with the function of full-length *SALL1* protein. Intriguingly transfection experiments have shown that truncated proteins similar to those potentially present in TBS patients are expressed throughout the cell. They can interact with full-length *spalt* proteins and can cause full-length protein to be relocalized from the nucleus to the cytoplasm providing a potential mechanism of dominant negative interference (McLeskey Kiefer et al., 2003; Sweetman et al., 2003). Studies in transgenic mice also support the idea of dominant negative activity by truncated *spalt* protein. Mice completely lacking *Sall1* show renal agenesis but lack the features associated with TBS (Nishinakamura et al., 2001). If haploinsufficiency is indeed the underlying cause of TBS in human, then mice carrying null alleles of *Sall1* might be expected to show a phenotype similar to that seen in TBS even though transgenic and knock out mice do not always completely recapitulate human disease phenotypes; however, heterozygous *Sall1* null mice have no apparent defects. In contrast, mice carrying a knock-in into the *Sall1* locus of a cDNA encoding a protein with a truncation similar to that predicted to be present in TBS have a range of defects much more reminiscent of TBS including sensorineural hearing loss, renal cystic hypoplasia and limb abnormalities (McLeskey Kiefer et al., 2003). That truncating mutations of *SALL1* orthologues can affect the function of full-length protein has been well established in vitro; however, so far truncated proteins have not been detected in TBS patients. Recently, further evidence has shown that some TBS patients carry deletions which entirely remove one *SALL1* allele and in

these cases haploinsufficiency can safely be assumed to lead to TBS; however, it is interesting to note that these cases show a milder phenotype than those associated with truncating mutations in *SALL1* (Borozdin et al., 2006). One plausible interpretation of these data is that while loss of one *SALL1* allele can cause TBS, a truncated protein has a stronger effect and produces a more severe phenotype. The potential dominant negative activity of truncated spalt proteins also raises the possibility that truncated *SALL1* protein could interact with and affect the function of other spalt proteins as these truncations are able to interact with all full-length spalt genes so far tested (McLeskey Kiefer et al., 2003; Sweetman et al., 2003). For example in limb development at least three spalt proteins (1, 3 and 4) are known to be expressed in partially overlapping domains and the presence of a truncated *SALL1* protein could therefore potentially affect the functioning of all of them. Thus, it may be the case that TBS is caused by interfering with the function of multiple *SALL* proteins during embryonic development. Interestingly, in *Drosophila*, flies carrying mutations in both *spalt* and *spalt related* show defects in antennal, genital and auditory defects, remarkable evidence of functional conservation of *spalt* gene function across vertebrate and invertebrate development (Dong et al., 2003).

While TBS is the best understood of the diseases caused by *spalt* gene defects, there remain some unanswered questions. In Townes–Brooks syndrome, the truncating mutations in *SALL1* are potentially able to interfere with the activity of any other *spalt* gene coexpressed in the same cell. Based on mouse models, it appears that the kidney defects observed may well be due to reduced function of *SALL1* protein while the patterning defects in the ear and limb are likely to be caused by dominant negative effects on other members of the family expressed in these tissues as well as *SALL1* (McLeskey Kiefer et al., 2003; Nishinakamura et al., 2001). The requirement for *csall4* in the invagination of sensory placodes (M. Bronner Fraser, personal communication) is consistent with the idea that the ear defects seen in TBS may result from interference with *SALL4* as well as *SALL1*.

#### *SALL4* in developmental disorders

Mutations in *SALL4* have been found in a range of clinically overlapping syndromes including Okihiro syndrome (also known as acro-renal-ocular syndrome or Duane radial ray syndrome), Holt–Oram syndrome and some cases previously thought to be thalidomide embryopathy (Borozdin et al., 2004a,b; Kohlhase et al., 2002b, 2003b). These syndromes present with defects in limb and heart development with Okihiro syndrome also showing defects in eye and kidney development. Although there is phenotypic variability in patients identified to have *SALL4* mutations, there is no clear correlation between specific mutations and phenotypes (Kohlhase et al., 2005). Unlike *SALL1* mutations in TBS, the mutations observed in these cases do not cluster around a critical region (Kohlhase et al., 2005) and haploinsufficiency has been shown to be the cause of these disorders (Borozdin et al., 2004a). The clinical overlap between these syndromes and TBS reinforces the idea that a

dominant negative effect of a *SALL1* truncation on *SALL4* function may be involved in TBS.

#### *Spalt* proteins in cancer

A possible role for *SALL1* in Wilms tumor was reported when high levels of expression were observed in these cancers (Ma et al., 2001b). Another spalt gene, *SALL2*, was shown to be repressed by the Wilms tumor protein, WT1 (Ma et al., 2001a). *SALL2* is perhaps the most likely candidate for a tumor suppressor gene as its mouse orthologue binds to the oncogenic polyoma large T antigen and has been shown to upregulate p21 expression (Li et al., 2001, 2004). *SALL1* may also have a role in tumorigenesis via its ability to enhance Wnt signaling (Sato et al., 2004). To date, there has been no firm evidence to support the idea that spalt proteins can act as tumor suppressors. However, if this were the case, one would predict that syndromes with either null or dominant negative mutations would correlate with an increased risk of cancer but this has not been reported and therefore the case for *spalt* genes in cancer aetiology remains speculative.

#### *SALL3* in 18q deletion syndrome

Very little is known about the role of *spalt* genes in 18q deletion syndrome which comprises a wide range of developmental defects including mental retardation, hearing loss and digit abnormalities. Although *SALL3* is likely to be lost in these deletions, it is one of several genes within this region and so cannot be regarded as the sole cause of this syndrome. Nevertheless, it remains possible that hemizygoty of *SALL3* is responsible for at least some of the features observed in this syndrome (Kohlhase et al., 1999).

### Structure and function of the spalt proteins

Developmental, clinical and biochemical studies have all contributed to our understanding of how the structure of the spalt proteins relates to their function. The combination of these approaches has provided critical insights into how this gene family operates during development and how mutation can lead to clinical disorders. In this section, we discuss what is known about the structure–function relationship of the spalt proteins and how this knowledge relates to the developmental and disease aspects of *spalt* gene activity.

#### Structure

The spalt proteins range from 105 to 140 kDa and have several conserved features including an N-terminal zinc finger domain of the C2HC type (which is not present in *Drosophila* spalt proteins), a glutamine-rich region and several double or triple zinc fingers of the C2H2 type throughout the protein with the precise number and spacing of these varying across the family. The predicted structures of the four human spalt proteins are shown in Fig. 2. In *SALL1* group proteins characterized to date, there are 9 of these C2H2 zinc fingers. However, in the

other groups, there seems to be variability with respect to zinc fingers 6 and 7. For example, in the predicted protein structures of SALL2 and SALL4, zinc fingers 6 and 7 are absent, while Xsall2, although closest to SALL2 in terms of its protein sequence, contains additional zinc fingers (Onai et al., 2004). It is also notable that the C terminal zinc fingers in SALL2 are not homologous to those seen in other spalt proteins lacking, for example, the distinctive ‘SAL box’ seen in other zinc fingers in these proteins. This suggests that these zinc fingers in SALL2 may have a separate origin to those in other spalt proteins. In the chick, csall3 is missing zinc fingers 6 and 7 compared to SALL3 (Farrell et al., 2001). However, the genomic sequence of *csall3* contains a predicted exon containing these missing zinc fingers and RT-PCR experiments suggest that an alternative splice variant of *csall3*, containing these additional zinc fingers, is expressed during development (DS and AM, unpublished data). As yet, little is known about the prevalence and extent of alternative splice forms of the spalt proteins and the possible functional consequences of this, and this may well prove to be an important area in the study of these genes.

#### *Transcriptional regulation by spalt genes*

Spalt proteins are thought to act as transcription factors although few target genes have been identified. Support for a role in transcriptional regulation has come from studies using fusion proteins of SALL1, Sall1, csall1 and csall3 to the GAL4 DNA binding domain. These fusion proteins are able to repress transcription from a GAL4 responsive promoter. Furthermore, in the case of Sall1, this ability was mapped to the N-terminal zinc finger which has been shown to interact with histone deacetylase to mediate repression (McLeskey Kiefer et al., 2002; Netzer et al., 2001; Sweetman et al., 2003). To date, the only demonstration of transcriptional regulation by a vertebrate spalt on a native promoter has come from studies of SALL2 which has been shown to bind to and regulate the activity of the CDK inhibitor p21 promoter (Li et al., 2004). Strikingly, SALL2 binding to the p21 promoter increases transcription and so provides evidence of positive regulation of transcription by this family of proteins. SALL1 has also been shown to act as a positive regulator of  $\beta$ -catenin-dependent Wnt signaling and this activity correlates with its ability to bind heterochromatin within the nucleus (Sato et al., 2004). In this study, SALL1 was shown to interact with  $\beta$ -catenin in coimmunoprecipitation assays but did not completely colocalize with  $\beta$ -catenin within the nucleus. The precise role of the spalt proteins in transcriptional regulation remains to be elucidated although evidence so far suggests that interactions with other proteins are likely to be critical in this process (see below). More spalt responsive promoters will have to be examined in order to obtain a more complete understanding of the mechanisms by which spalt proteins affect target gene expression.

#### *Interactions of spalt proteins*

The function of the spalt proteins is likely to be modulated through interactions with other proteins in multi protein

complexes. The spalt proteins themselves can form both homo- and heteromeric complexes and to date all members of the family that have been tested show the ability to interact with each other in coimmunoprecipitation assays (McLeskey Kiefer et al., 2003; Sweetman et al., 2003). These interactions are mediated by the N-terminal region of the protein and require the glutamine-rich region, and proteins lacking this domain do not coprecipitate. It has been demonstrated that spalt protein interactions can influence subcellular localization and thereby potentially modify function. For example, full-length SALL1, Sall1 and csall1 have all been shown to localize to the nucleus as expected for a transcriptional regulator (McLeskey Kiefer et al., 2002; Netzer et al., 2001; Sweetman et al., 2003). However, csall3 is found in the cytoplasm and, when coexpressed with csall1, can cause the relocalization of csall1 from the nucleus to the cytoplasm (Sweetman et al., 2003). The expression pattern of *csall1* and *csall3* during limb development is partially overlapping. This finding, together with the cellular localization studies, suggests a mechanism whereby the activity of csall1 in the nucleus can be modulated across the limb bud by coexpression of *csall3*, and this may be important for normal limb development.

Other proteins that have been shown to interact with spalt proteins include HDAC and  $\beta$ -catenin. In addition, SALL1 has been shown to interact with TRF1/PIN2 (Netzer et al., 2001), UBE2I and SUMO-1 (Netzer et al., 2002) and Sall2 with polyoma virus large T antigen (Li et al., 2001). It will be important to dissect the implications of these interactions for spalt protein function in both development and disease.

While it is known that spalt proteins can interact in vitro, it is not clear if this happens in vivo and, if so, what the consequences might be. It is possible that target gene expression may be modified by different combinations of spalt proteins and this could explain why discrete but overlapping expression patterns are seen in different regions of the embryo, for example the developing limb and kidney. There is good evidence for both repressive and activating functions on target gene expression, presumably dependent on cellular context. Although GAL4 fusions have uncovered transcriptional repressor activity and this has been linked to interaction with HDAC, the only known target gene in vertebrates so far, p21, is upregulated in the presence of SALL2 (Li et al., 2004). Understanding the details of these of these processes, including the identification of further *spalt* targets, will be critical in our attempts to unravel the function of this family of proteins in development.

#### *Future directions*

One area of interest that requires further elucidation is the subcellular localization of the spalt proteins. SALL1 and its orthologues are nuclear, as would be expected of transcription factors, but at least one family member, csall3 in chick, has been shown to be cytoplasmic. The ability of csall3 to remove csall1 from the nucleus when both proteins are coexpressed suggests a role in negative regulation of other spalt proteins. However, this may not be the whole story. Preliminary data from our lab show that csall3 undergoes tyrosine phosphorylation while csall1

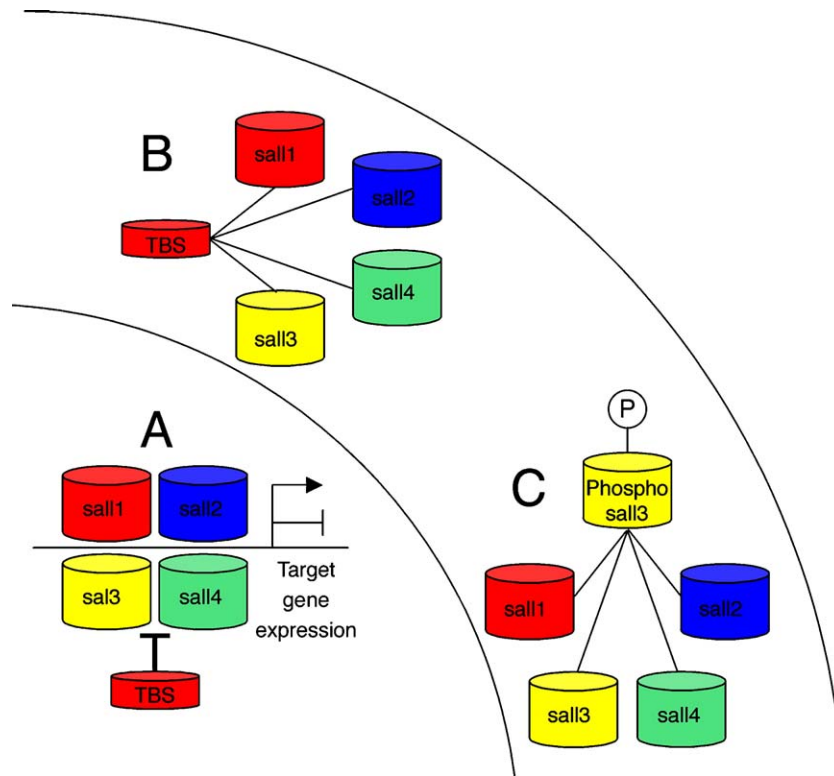


Fig. 3. Model of possible spalt protein activities. (A) Combinatorial control of target gene expression within the nucleus by homo- and heteromeric spalt protein complexes. 'TBS like' truncations could interfere with this activity. (B) Cytoplasmic 'TBS like' truncations can relocate full-length protein from the nucleus. (C) Phosphorylated csal3 protein is cytoplasmic and can relocate spalt proteins from the nucleus, this may apply to other SALL3 group members.

does not. Furthermore, upon withdrawal of serum, csall3 can enter the nucleus (DS and AM, unpublished). This implies a further level of control of spalt activity where phosphorylation,

induced by growth factors, can alter subcellular localization and thus potentially transcriptional activity of at least one member of the spalt family. The possible existence of a splice variant of

Table 2  
Spalt protein function and regulation

Spalt group	Associated disease	Interacts with	Regulated by	Developmental roles	Target genes
SALL1	Townes–Brocks syndrome (Kohlhase et al., 1998)	sall1, sall3, N terminal truncations of sall1 (McLeskey Kiefer et al., 2003; Sweetman et al., 2003) HDAC (McLeskey Kiefer et al., 2002) UBE21, SUMO-1 (Netzer et al., 2002)	FGFs/Wnts (Farrell and Münsterberg, 2000) BMP (Capdevila et al., 1999)	Limb patterning (McLeskey Kiefer et al., 2003), kidney development (Nishinakamura et al., 2001)	
SALL2		N terminal truncations of sall1 (McLeskey Kiefer et al., 2003) Polyoma virus large T (Li et al., 2001)	Wilms tumor protein (Ma et al., 2001a)	None known (Sato et al., 2003)	p21 (Li et al., 2004)
SALL3	18q23 deletion syndrome (?) (Kohlhase et al., 1999)	sall1 full-length and N terminal truncations (McLeskey Kiefer et al., 2003; Sweetman et al., 2003)		Palate formation, cranial nerve development (Parrish et al., 2004)	
SALL4	Okiihiro syndrome (Kohlhase et al., 2002b) Holt–Oram and acro-renal-ocular syndrome (Kohlhase et al., 2003a,b)	N terminal truncations of sall1 (McLeskey Kiefer et al., 2003)	Tbx5 (Harvey and Logan, 2006; Koshiba-Takeuchi et al., 2006)	Neural crest development (Barembaum and Bronner-Fraser, 2004), ear development (M Bronner Fraser, personal communication), limb development (Harvey and Logan, 2006; Koshiba-Takeuchi et al., 2006)	

*csall3* which could alter the localization and therefore function is also intriguing. Another related question concerns the mode of action of truncated spalt proteins, such as those potentially involved in TBS. N terminal truncations of SALL1 group proteins can remove full-length spalt proteins from the nucleus; however, a fusion of such an N terminal truncation to the glucocorticoid receptor shows enhanced repressor activity upon addition of dexamethasone (Onai et al., 2004). So it appears that these truncations can act to antagonize full-length protein function in multiple ways within the cell interfering with protein function both within and outside the nucleus.

The model shown in Fig. 3 summarizes our current understanding of spalt protein function. In the nucleus, spalt proteins may act in combination to activate or repress expression of target genes while ‘TBS like’ truncations can interfere with this activity (Fig. 3A). ‘TBS like’ truncations may also abrogate the function of full-length spalt proteins by relocating them into the cytoplasm (Fig. 3B). Sall3 proteins may also act as negative regulators of other spalts by removing them from the nucleus although this may depend on the phosphorylation state of the *sall3* protein (Fig. 3C).

## Conclusions

This review illustrates that the spalt family of proteins are involved in many fundamental processes during development. Their roles and mechanisms of action are beginning to be elucidated by developmental biologists and clinical scientists; however, many exciting questions remain unanswered. Some of these aspects of spalt protein biology are summarized in Table 2 and we can look forward to learning more about the diverse developmental, cellular and molecular roles of this protein family.

## Acknowledgments

We thank Malcolm Logan and Marianne Bronner-Fraser for sharing unpublished data and Grant Wheeler and Laura Wagstaff for their comments on the manuscript. D.S. was funded by a Wellcome Trust Project grant and research in A. M.’s laboratory is funded by the Wellcome Trust, the BBSRC and the British Heart Foundation.

## References

- Barembaum, M., Bronner-Fraser, M., 2004. A novel spalt gene expressed in branchial arches affects the ability of cranial neural crest cells to populate sensory ganglia. *Neuron Glia Biol.* 1, 57–63.
- Barrio, R., de Celis, J.F., Bolshakov, S., Kafatos, F.C., 1999. Identification of regulatory regions driving the expression of the *Drosophila* spalt complex at different developmental stages. *Dev. Biol.* 215, 33–47.
- Borozdin, W., Boehm, D., Leipoldt, M., Wilhelm, C., Reardon, W., Clayton-Smith, J., Becker, K., Muhlendyck, H., Winter, R., Giray, O., Silan, F., Kohlhasse, J., 2004a. SALL4 deletions are a common cause of Okhiro and acro-renal-ocular syndromes and confirm haploinsufficiency as the pathogenic mechanism. *J. Med. Genet.* 41, E113.
- Borozdin, W., Wright, M.J., Hennekam, R.C., Hannibal, M.C., Crow, Y.J., Neumann, T.E., Kohlhasse, J., 2004b. Novel mutations in the gene SALL4 provide further evidence for acro-renal-ocular and Okhiro syndromes being allelic entities, and extend the phenotypic spectrum. *J. Med. Genet.* 41, e102.
- Botzenhart, E.M., Green, A., Ilyina, H., Konig, R., Lowry, R.B., Lo, I.F., Shohat, M., Burke, L., McGaughan, J., Chafai, R., Pierquin, G., Michaelis, R.C., Whiteford, M.L., Simola, K.O., Rosler, B., Kohlhasse, J., 2005. SALL1 mutation analysis in Townes–Brocks syndrome: twelve novel mutations and expansion of the phenotype. *Hum. Mutat.* 26, 282.
- Borozdin, W., Steinmann, K., Albrecht, B., Bottani, A., Devriendt, K., Leipoldt, M., Kohlhasse, J., 2006. Detection of heterozygous SALL1 deletions by quantitative real time PCR proves the contribution of a SALL1 dosage effect in the pathogenesis of Townes–Brocks syndrome. *Hum. Mutat.* 27, 211–212.
- Buck, A., Kispert, A., Kohlhasse, J., 2001. Embryonic expression of the murine homologue of SALL1, the gene mutated in Townes–Brocks syndrome. *Mech. Dev.* 104, 143–146.
- Camp, E., Hope, R., Kortschak, R.D., Cox, T.C., Lardelli, M., 2003. Expression of three spalt (sal) gene homologues in zebrafish embryos. *Dev. Genes Evol.* 213, 35–43.
- Capdevila, J., Tsukui, T., Rodriguez Esteban, C., Zappavigna, V., Izpisua Belmonte, J.C., 1999. Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. *Mol. Cell* 4, 839–849.
- Creazzo, T.L., Godt, R.E., Leatherbury, L., Conway, S.J., Kirby, M.L., 1998. Role of cardiac neural crest cells in cardiovascular development. *Annu. Rev. Physiol.* 60, 267–286.
- de Celis, J.F., Barrio, R., 2000. Function of the spalt/spalt-related gene complex in positioning the veins in the *Drosophila* wing. *Mech. Dev.* 91, 31–41.
- de Celis, J.F., Barrio, R., Kafatos, F.C., 1999. Regulation of the spalt/spalt-related gene complex and its function during sensory organ development in the *Drosophila* thorax. *Development* 126, 2653–2662.
- del Alamo Rodriguez, D., Terriente Felix, J., Diaz-Benjumea, F.J., 2004. The role of the T-box gene *optomotor-blind* in patterning the *Drosophila* wing. *Dev. Biol.* 268, 481–492.
- Domingos, P.M., Brown, S., Barrio, R., Ratnakumar, K., Frankfort, B.J., Mardon, G., Steller, H., Mollereau, B., 2004a. Regulation of R7 and R8 differentiation by the spalt genes. *Dev. Biol.* 273, 121–133.
- Domingos, P.M., Mlodzik, M., Mendes, C.S., Brown, S., Steller, H., Mollereau, B., 2004b. Spalt transcription factors are required for R3/R4 specification and establishment of planar cell polarity in the *Drosophila* eye. *Development* 131, 5695–5702.
- Dong, P.D., Todi, S.V., Eberl, D.F., Boekhoff-Falk, G., 2003. *Drosophila* spalt/spalt-related mutants exhibit Townes–Brocks’ syndrome phenotypes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10293–10298.
- Farrell, E.R., Münsterberg, A.E., 2000. *csal1* is controlled by a combination of FGF and Wnt signals in developing limb buds. *Dev. Biol.* 225, 447–458.
- Farrell, M., Waldo, K., Li, Y.X., Kirby, M.L., 1999. A novel role for cardiac neural crest in heart development. *Trends Cardiovasc. Med.* 9, 214–220.
- Farrell, E.R., Tosh, G., Church, E., Münsterberg, A.E., 2001. Cloning and expression of CSAL2, a new member of the spalt gene family in chick. *Mech. Dev.* 102, 227–230.
- Harvey, S.A., Logan, M.P.O., 2006. *sall4* acts downstream of *tbx5* and is required for pectoral fin development. *Development* 133, 1165–1173.
- Holleman, T., Schuh, R., Pieler, T., Stick, R., 1996. *Xenopus* Xsal-1, a vertebrate homologue of the region specific homeotic gene spalt of *Drosophila*. *Mech. Dev.* 55, 19–32.
- Houweling, A.C., Dildrop, R., Peters, T., Mummenhoff, J., Moorman, A.F., Ruther, U., Christoffels, V.M., 2001. Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mech. Dev.* 107, 169–174.
- Jürgens, G., 1988. Head and tail development of the *Drosophila* embryo involves spalt, a novel homeotic gene. *EMBO J.* 7, 189–196.
- Kohlhasse, J., 2000. SALL1 mutations in Townes–Brocks syndrome and related disorders. *Hum. Mutat.* 16, 460–466.
- Kohlhasse, J., Schuh, R., Dowe, G., Kuhnlein, R.P., Jackle, H., Schroeder, B., Schulz-Schaefer, W., Kretzschmar, H.A., Kohler, A., Muller, U., Raab-Vetter, M., Burkhardt, E., Engel, W., Stick, R., 1996. Isolation, characterization, and organ-specific expression of two novel human zinc finger genes related to the *Drosophila* gene spalt. *Genomics* 38, 291–298.



- Kohlhase, J., Wischermann, A., Reichenbach, H., Froster, U., Engel, W., 1998. Mutations in the SALL1 putative transcription factor gene cause Townes–Brocks syndrome. *Nat. Genet.* 18, 81–83.
- Kohlhase, J., Hausmann, S., Stojmenovic, G., Dixkens, C., Bink, K., Schulz-Schaeffer, W., Altmann, M., Engel, W., 1999. SALL3, a new member of the human spalt-like gene family, maps to 18q23. *Genomics* 62, 216–222.
- Kohlhase, J., Altmann, M., Archangelo, L., Dixkens, C., Engel, W., 2000. Genomic cloning, chromosomal mapping, and expression analysis of msal-2. *Mamm. Genome* 11, 64–68.
- Kohlhase, J., Heinrich, M., Liebers, M., Frohlich Archangelo, L., Reardon, W., Kispert, A., 2002a. Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. *Cytogenet. Genome Res.* 98, 274–277.
- Kohlhase, J., Heinrich, M., Schubert, L., Liebers, M., Kispert, A., Laccone, F., Turmpenny, P., Winter, R.M., Reardon, W., 2002b. Okihiro syndrome is caused by SALL4 mutations. *Hum. Mol. Genet.* 11, 2979–2987.
- Kohlhase, J., Liebers, M., Backe, J., Baumann-Muller, A., Bembea, M., Destree, A., Gattas, M., Grussner, S., Muller, T., Mortier, G., Skrypnik, C., Yano, S., Wirbelauer, J., Michaelis, R.C., 2003a. High incidence of the R276X SALL1 mutation in sporadic but not familial Townes–Brocks syndrome and report of the first familial case. *J. Med. Genet.* 40, e127.
- Kohlhase, J., Schubert, L., Liebers, M., Rauch, A., Becker, K., Mohammed, S.N., Newbury-Ecob, R., Reardon, W., 2003b. Mutations at the SALL4 locus on chromosome 20 result in a range of clinically overlapping phenotypes, including Okihiro syndrome, Holt–Oram syndrome, acro-renal-ocular syndrome, and patients previously reported to represent thalidomide embryopathy. *J. Med. Genet.* 40, 473–478.
- Kohlhase, J., Chitayat, D., Kotzot, D., Ceylaner, S., Froster, U.G., Fuchs, S., Montgomery, T., Rosler, B., 2005. SALL4 mutations in Okihiro syndrome (Duane-radial ray syndrome), acro-renal-ocular syndrome, and related disorders. *Hum. Mutat.* 26, 176–183.
- Koshiba-Takeuchi, K., Takeuchi, J.K., Arruda, E.P., Kathiriya, I.S., Mo, R., Hui, C.C., Srivastava, D., Bruneau, B.G., 2006. Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. *Nat. Genet.* 38, 175–183.
- Koster, R., Stick, R., Loosli, F., Wittbrodt, J., 1997. Medaka spalt acts as a target gene of hedgehog signaling. *Development* 124, 3147–3156.
- Kuhnlein, R.P., Schuh, R., 1996. Dual function of the region-specific homeotic gene spalt during *Drosophila* tracheal system development. *Development* 122, 2215–2223.
- Kuhnlein, R.P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J.F., Gehring, W.J., Jackle, H., Schuh, R., 1994. spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the *Drosophila* embryo. *EMBO J.* 13, 168–179.
- Li, D., Dower, K., Ma, Y., Tian, Y., Benjamin, T.L., 2001. A tumor host range selection procedure identifies p150(sal2) as a target of polyoma virus large T antigen. *Proc. Natl. Acad. Sci. U. S. A.* 98, 14619–14624.
- Li, D., Tian, Y., Ma, Y., Benjamin, T., 2004. p150(Sal2) is a p53-independent regulator of p21(WAF1/CIP). *Mol. Cell. Biol.* 24, 3885–3893.
- Ma, Y., Li, D., Chai, L., Luciani, A.M., Ford, D., Morgan, J., Maizel, A.L., 2001a. Cloning and characterization of two promoters for the human HSAL2 gene and their transcriptional repression by the Wilms tumor suppressor gene product. *J. Biol. Chem.* 276, 48223–48230.
- Ma, Y., Singer, D.B., Gozman, A., Ford, D., Chai, L., Steinhoff, M.M., Hansen, K., Maizel, A.L., 2001b. Hsal 1 is related to kidney and gonad development and is expressed in Wilms tumor. *Pediatr. Nephrol.* 16, 701–709.
- McLeskey Kiefer, S., McDill, B.W., Yang, J., Rauchman, M., 2002. Murine Sall1 represses transcription by recruiting a histone deacetylase complex. *J. Biol. Chem.* 277, 14869–14876.
- McLeskey Kiefer, S., Ohlemiller, K.K., Yang, J., McDill, B.W., Kohlhase, J., Rauchman, M., 2003. Expression of a truncated Sall1 transcriptional repressor is responsible for Townes–Brocks syndrome birth defects. *Hum. Mol. Genet.* 12, 2221–2227.
- Neff, A.W., King, M.W., Harty, M.W., Nguyen, T., Calley, J., Smith, R.C., Mescher, A.L., 2005. Expression of *Xenopus* XISALL4 during limb development and regeneration. *Dev. Dyn.* 233, 356–367.
- Netzer, C., Bohlander, S.K., Rieger, L., Muller, S., Kohlhase, J., 2002. Interaction of the developmental regulator SALL1 with UBE2I and SUMO-1. *Biochem. Biophys. Res. Commun.* 296, 870–876.
- Netzer, C., Rieger, L., Brero, A., Zhang, C.D., Hinzke, M., Kohlhase, J., Bohlander, S.K., 2001. SALL1, the gene mutated in Townes–Brocks syndrome, encodes a transcriptional repressor which interacts with TRF1/PIN2 and localizes to pericentromeric heterochromatin. *Hum. Mol. Genet.* 10, 3017–3024.
- Nishinakamura, R., Matsumoto, Y., Nakao, K., Nakamura, K., Sato, A., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Scully, S., Lacey, D.L., Katsuki, M., Asashima, M., Yokota, T., 2001. Murine homolog of SALL1 is essential for ureteric bud invasion in kidney development. *Development* 128, 3105–3115.
- Onai, T., Sasai, N., Matsui, M., Sasai, Y., 2004. *Xenopus* XsalF: anterior neuroectodermal specification by attenuating cellular responsiveness to Wnt signaling. *Dev. Cell* 7, 95–106.
- Onuma, Y., Nishinakamura, R., Takahashi, S., Yokota, T., Asashima, M., 1999. Molecular cloning of a novel *Xenopus* spalt gene (Xsal-3). *Biochem. Biophys. Res. Commun.* 264, 151–156.
- Ott, T., Kaestner, K.H., Monaghan, A.P., Schutz, G., 1996. The mouse homolog of the region specific homeotic gene spalt of *Drosophila* is expressed in the developing nervous system and in mesoderm-derived structures. *Mech. Dev.* 56, 117–128.
- Ott, T., Parrish, M., Bond, K., Schwaeger-Nickolenko, A., Monaghan, A.P., 2001. A new member of the spalt like zinc finger protein family, Msal-3, is expressed in the CNS and sites of epithelial/mesenchymal interaction. *Mech. Dev.* 101, 203–207.
- Parrish, M., Ott, T., Lance-Jones, C., Schuetz, G., Schwaeger-Nickolenko, A., Monaghan, A.P., 2004. Loss of the Sall3 gene leads to palate deficiency, abnormalities in cranial nerves, and perinatal lethality. *Mol. Cell. Biol.* 24, 7102–7112.
- Powell, C.M., Michaelis, R.C., 1999. Townes–Brocks syndrome. *J. Med. Genet.* 36, 89–93.
- Sato, A., Matsumoto, Y., Koide, U., Kataoka, Y., Yoshida, N., Yokota, T., Asashima, M., Nishinakamura, R., 2003. Zinc finger protein sall2 is not essential for embryonic and kidney development. *Mol. Cell. Biol.* 23, 62–69.
- Sato, A., Kishida, S., Tanaka, T., Kikuchi, A., Kodama, T., Asashima, M., Nishinakamura, R., 2004. Sall1, a causative gene for Townes–Brocks syndrome, enhances the canonical Wnt signaling by localizing to heterochromatin. *Biochem. Biophys. Res. Commun.* 319, 103–113.
- Smith, J., 1996. How to tell a cell where it is. *Nature* 381, 367–368.
- Sweetman, D., Smith, T., Farrell, E.R., Chantry, A., Münsterberg, A., 2003. The conserved glutamine-rich region of chick csal1 and csal3 mediates protein interactions with other spalt family members. Implications for Townes–Brocks syndrome. *J. Biol. Chem.* 278, 6560–6566.
- Sweetman, D., Smith, T.G., Farrell, E., Münsterberg, A., 2005. Expression of csal1 in pre limb-bud chick embryos. *Int. J. Dev. Biol.* 49, 427–430.
- Zulch, A., Becker, M.B., Gruss, P., 2001. Expression pattern of Irx1 and Irx2 during mouse digit development. *Mech. Dev.* 106, 159–162.