Dev Genes Evol (2002) 212:186–195
DOI 10.1007/s00427-002-0228-1

ORIGINAL ARTICLE

Olivier Lespinet · Alexander J. Nederbragt
Michel Cassan · Wim J.A.G. Dictus
André E. van Loon · André Adoutte

Characterisation of two *snail* genes in the gastropod mollusc *Patella vulgata*. Implications for understanding the ancestral function of the *snail*-related genes in Bilateria

Received: 27 September 2001 / Accepted: 25 February 2002 / Published online: 18 April 2002
© Springer-Verlag 2002

Abstract *Snail* genes have been found to play a role in mesoderm formation in two of the three clades of bilaterians, deuterostomes (comprising the chordates) and ecdysozoans (comprising the arthropods). No clear data are available on the role these genes play in development of the mesoderm in the third clade, that of lophotrochozoans (comprising annelids and molluscs). We identified two new members of the *snail* gene family in the gastropod mollusc *Patella vulgata*. Phylogenetic analysis showed that the two genes clearly belong to the *snail* sub-family. Their expression patterns do not indicate a role during early mesoderm formation. In fact, contrary to expectations, the *snail* genes of *Patella* were mostly expressed in the ectoderm. In view of the location of their expression sites, we suggest that these genes could be involved in regulating epithelial-mesenchymal transitions (EMT) and cell motility, as has recently been demonstrated for *snail* genes in vertebrates. This may well correspond to the ancestral function of these genes. The results are discussed in the light of the evolutionary origin of the mesoderm. Electronic supplementary material to this paper can be obtained by using the Springer LINK server located at http://dx.doi.org/10.1007/s00427-002-0228-1.

Keywords Mesoderm formation · *snail/slug* and *scratch* gene family · Mollusc · *Patella vulgata* · Epithelial-mesenchymal transitions (EMT)

Edited by D. Tautz

O. Lespinet and A.J. Nederbragt contributed equally to this work

Electronic supplementary material to this paper can be obtained by using the Springer LINK server located at http://dx.doi.org/10.1007/s00427-002-0228-1

O. Lespinet (✉) · M. Cassan · A. Adoutte
Centre de Génétique Moléculaire, CNRS UPR 2167,
91198 Gif-sur-Yvette Cedex, France
e-mail: lespinet@ncbi.nlm.nih.gov
Tel.: +1-301-594-8079, Fax: +1-301-480-4311

A.J. Nederbragt · W.J.A.G. Dictus · A.E. van Loon
Department of Developmental Biology, Utrecht University,
Padualaan 8, 3584 Utrecht, The Netherlands

Present addresses:
O. Lespinet, National Center for Biotechnology Information,
National Library of Medicine, National Institutes of Health,
Bethesda, MD 20894, USA

A.J. Nederbragt, Kewalo Marine Lab,
Pacific Biomedical Research Center,
University of Hawaii, 41 Ahui Street, Honolulu, HI 96813, USA

Introduction

Existence of a third germ layer, the mesoderm, is a characteristic of all bilaterally symmetrical animals (Bilateria). However, on a purely descriptive basis, protostomes and deuterostomes appear to specify their mesoderm during early development through quite different mechanisms. This raises the question of mesoderm homology across Bilateria.

Recent molecular phylogenies (Halanych et al. 1995; Aguinaldo et al. 1997), as well as a re-analysis of morphological characters (Peterson and Eernisse 2001), have shown that bilaterians are subdivided into three major clades: one deuterostome clade and two protostome clades, the ecdysozoans and the lophotrochozoans (for a review see Adoutte et al. 2000). Of the three clades, only deuterostomes and ecdysozoans have been subjected to intensive molecular embryological investigations. We have set out to analyse the molecular mechanisms underlying mesoderm specification in the third clade, the lophotrochozoans. We focussed on the mollusc *Patella vulgata* since there is a wealth of descriptive and experimental embryological data about this genus (Verdonk and van den Biggelaar 1983).

Our approach consisted of identifying genes known or thought to be intimately involved in mesoderm specification in model organisms, especially *Drosophila melanogaster*, and to study their time and pattern of expression in early *Patella* embryogenesis. Here, we report the analysis of two *Patella* homologues of the *snail* gene, an ob-

vious candidate gene for the control of mesoderm formation.

The *snail* gene was first isolated as a zygotic lethal mutation in the *D. melanogaster* embryo (Simpson 1983; Nüsslein-Volhard et al. 1984). In *snail* mutants, no or very few mesodermal tissues are formed (Simpson 1983; Grau et al. 1984). This gene encodes a zinc-finger transcription factor (Boulay et al. 1987; Leptin 1991; Kasai et al. 1992) that possesses a characteristic domain composed of four canonical C-X_2-C-X_{12}-H-X_3-H motives and one variant C-X_2-C-X_{12}-H-X_4-C motif in its amino terminus, the so-called *snail*-box domain (Whiteley et al. 1992).

In *Drosophila*, *snail* is first expressed at gastrulation in the presumptive mesoderm (Alberga et al. 1991; Leptin 1991) where it mainly represses ventrolateral neurogenic genes (Kosman et al. 1991; Leptin 1991; Casal and Leptin 1996) but also participates in the regulation of mesoderm-specific genes (Hemavathy et al. 1997). Later in development, *snail* is also dynamically expressed in endodermal and ectodermal derivatives (Alberga et al. 1991; Leptin 1991; Ip et al. 1994) and thus its expression is not only restricted to mesoderm.

Subsequently, a role in mesoderm development for members of the *snail* family has been found in other arthropods (Sommer and Tautz 1994), in urochordates (Corbo et al. 1997; Wada and Saiga 1999), in cephalochordates (Langeland et al. 1998) and in vertebrates (Sargent and Bennett 1990; Nieto et al. 1992, 1994; Smith et al. 1992; Mayor et al. 1993; Thisse et al. 1993, 1995; Hammerschmidt and Nusslein-Volhard 1993). In the short-germ-band insect *Tribolium castaneum*, *snail* expression is very similar to that observed in *Drosophila* (Sommer and Tautz 1994). In chordates, expression was found in undifferentiated and migrating mesoderm (Sargent and Bennett 1990; Smith et al. 1992; Nieto et al. 1992; Mayor et al. 1993; Ros et al. 1997; Langeland et al. 1998; Linker et al. 2000).

Recent years have highlighted the diversification of the *snail* gene family, with the characterisation of multiple members in several species, and two subfamilies, the *snail/slug* and *scratch* subfamilies, have been identified (Manzanares et al. 2001). In concert with this, suspected additional roles for these genes have been confirmed. One of these is a role for *snail* genes in nervous system development in *Drosophila* (Ashraf et al. 1999; Cai et al. 2001) and vertebrates (Nakakura et al. 2001). In addition, in vertebrates, the *slug* and *snail* genes are also expressed in migrating neural crest cells and neural crest derivatives (Nieto et al. 1994; Mayor et al. 1995; Linker et al. 2000). Moreover, recent molecular analysis at the cellular level demonstrated that the *snail* genes control epithelial mesenchymal transitions (EMT) by repressing E-cadherin expression (Batlle et al. 2000; Cano et al. 2000).

Thus, although a role in early mesoderm formation among *Bilateria* seems to be a common aspect of *snail* genes, other functions, related to CNS development and morphogenetic movements, might also belong to the ancestral roles of this family. In order to better resolve this issue, we need to study the expression of the *snail*-related genes in the third metazoan group, namely the lophotrochozoans. Up to date, this has only been done in the annelid leech *Helobdella robusta*. There, the *snail* gene was found to be widely expressed in precursor cells of both the ectoderm and mesoderm, making a role in mesoderm specification unlikely, and leaving open what the function of this gene in leech development might be (Goldstein et al. 2001).

Here, we present data on the characterisation of two new *snail*-related genes from the gastropod mollusc *P. vulgata*. We cloned both genes, examined their sequence and analysed their spatio-temporal expression patterns during the early stages of development through whole-mount in situ hybridisation. Their expression patterns cause speculation about the ancestral function of the *snail* gene in the Bilateria. Where a mesodermal precursor function was previously suggested for this gene, the results presented here do not fit with this idea and instead suggest that the ancestral function of *snail* genes is probably related to EMT and neural specification.

Materials and methods

PCR amplification

PCR with degenerate primers Sna1W (GAAGATCTGGNGCNYTNAARATGCAYAT) and Sna4C (GAAGATCTRTGNGCNCKNARRTTNSWNCKRTC) was performed as follows: 50 ng sperm DNA from one individual were used as a template in a 30-µl final volume reaction with 1 µM each primer, 10 µM dNTP and 0.1 unit Taq polymerase with provided 1× buffer (Appligene). The PCR program was: 5 cycles [50 s at 94°C, 1 min at 55°C (switch to 72°C by an increase of 1°C every 9 s), 2 min at 72°C], followed by 30 cycles (50 s at 94°C, 1 min 30 s at 55°C, 1 min at 72°), and a final extension for 10 min at 72°C. Specificity of this first amplification was controlled in the same conditions by a nested PCR using primers Sna2W (GAAGATCTAARGCNTTYWSNMGNCCNTGG) and Sna3C (GAAGATCTRAANGGYTTYTCNCCNGTRTG) with the following program: 30 cycles (50 s at 94°C, 1 min 30 s at 55°C, 1 min at 72°C) and a final extension for 10 min at 72°C. Amplified fragments were cloned in TOPO-TA vector (Invitrogen).

cDNA and genomic library screening

Library construction has been described previously (van Loon et al. 1993). cDNA and genomic libraries were screened according to Sambrook et al. (1989). We screened 10^6 plaques for the cDNA library and 2×10^5 plaques for the genomic library.

Subcloning of the *Pv-Sna2* genomic clone by PCR

The genomic clone was sub-cloned by long range PCR according to the manufacturer's protocol (Pharmacia) with the following program: 1 min at 94°C, 35 cycles (45 s at 94°C, 45 s at 60°C, 10 min at 68°C). We used Nrl (TAAAAGCAGAAGTCCAACCCAGATAACGAT) as a vector-specific primer and Nslong (AGCTTTACCGCATAATTTACACTTACAGG) as a *Patella snail*-box-specific primer. The amplified fragment was cloned in TOPO-TA vector (Invitrogen).

3′ RACE on plasmid library

From the cDNA library, plasmids were prepared by mass-excision according to the Stratagene protocol. To obtain the 3′ part of the *Pv-sna2* gene, 1 ng of these plasmids was used in a PCR with a vector-specific primer pBS-E (TAACGCCAGGGTTTTCCCAGT) directed at the 3′ side of the insert and a primer Forw1 (AAGAACTCACACATTACCCTGT) specific for the 196-bp *snail*-like fragment. PCR conditions were: 10 mM Tris-HCl pH 9.0, 3.0 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100 and 0.01% (w/v) gelatin, 0.2 mM dNTPs (Amersham Pharmacia Biotech), 0.24 µM each primer and 0.005 units/µl SuperTaq (HT Biotechnology), with cycle parameters 3 min at 94°C; 20 cycles touchdown: 45 s at 94°C, 30 s at 72°C with a 1°C drop of temperature every other cycle, 1 min 30 s at 72°C; followed by 20 cycles 30 s at 94°C, 30 s at 62°C, 1 min 30 s at 72°C and a final extension for 10 min at 72°C. Nested PCR was done on the fragments of the first PCR by (following gel-electrophoresis) taking up part of the band with a sterile wooden toothpick, eluting it in 5 µl water, 1 µl of which was used as a template with the vector-specific primer T7 (GTAATACGACTCACTATAGGGC) and primer Forw2 (CTCGACCTTGGTTATTACAAGG) under similar conditions with annealing temperature starting at 64°C and ending at 54°C and extension of 3 min at 68°C. All resulting PCR fragments were isolated from gel using the Prep-a-gene kit (BioRad) and cloned into the pGEMT-Easy vector (Promega).

Pv-Sna2 exon/intron boundaries

The exon/intron boundaries were confirmed by sequencing a PCR fragment obtained from the cDNA library using primers T3 (AATTAACCCTCATAAAGGG) and Rev2 (AAGAAAATGGCTTCTCGCCAGT).

Fertilisation and embryo rearing

Adult animals were collected at the Roscoff marine station in September 1998, 1999 and 2000 and kept in natural seawater until the end of the breeding season (March). Embryos were obtained as described previously (van den Biggelaar 1977).

Embryo fixation and storage

Early stage embryos (until 24 hours post fertilisation, hpf) were dejellied by a brief treatment in millipore-filtered seawater (MPFSW) acidified with HCl to pH 3.9 and washed in normal MPFSW. Selected embryos were fixed in Eppendorf tubes for at least 1 h with MEMPFA-T (0.1 M MOPS, pH 7.4, 2 mM EGTA, 1 mM $MgSO_4$, 4% paraformaldehyde and 0.1% Tween-20) on a rotating wheel. Embryos were progressively dehydrated with 5-min washes with a gradual series of methanol/MEMPFA-T (25/75, 50/50, 75/25, 100/0) followed by a 15-min wash with 100% MeOH. At this point embryos were stored at –20°C until further use.

Probe synthesis

Sense and anti-sense probes were synthesised from PCR product using the Maxiscript T7 and SP6 RNA polymerase kit (Ambion) and the Dig RNA labelling kit (Roche Molecular Biochemicals). For the *Pv-Sna1* probe, we used the complete cDNA sequence (1.6 kb). For the *Pv-Sna2* probe we only used a 1-kb PCR fragment amplified by primers T3 and Rev2. This fragment corresponds to the region between the SNAG motive and the *Snail*-box (both included).

Whole-mount in situ hybridisation

The protocol used was modified after Houtzager (1998). All steps were performed at room temperature unless indicated otherwise. Embryos were progressively rehydrated in TBS-T (136 mM NaCl; 25 mM Tris, pH 7.4; 2.7 mM KCl and 0.1% Tween-20) by 5-min successive washes in methanol/TBST (75/25, 50/50, 25/75, 0/100). Then they were treated with proteinase K (50 ng/ml; Roche Molecular Biochemicals) in TBS-T for 20 min at 37°C. Embryos were then successively incubated in 0.1 M triethanolamine (TEA, Merck) pH 7.0–8.0, TEA with 2.5 µl acetic anhydride/ml (Baker Chemicals) and TEA with 5.0 µl acetic anhydride/ml, followed by two washes for 5 min each in TBS-T. Embryos were re-fixed for 20 min in MEMPFA-T, washed 5× 5 min in TBS-T, washed for 10 min in 500 µl 50/50 TBS-T/hybridisation buffer [HB: 750 mM NaCl, 75 mM Na citrate pH 7.0, 50% formamide, 100 µg/ml Torula RNA (Fluka), 1.5% blocking reagent (Roche Molecular Biochemicals), 5 mM EDTA and 0.1% Tween-20]. Prehybridisation was done in 100 µl HB in an Eppendorf tube for 6 h at 65°C. One microlitre probe was added and hybridisation was carried out overnight at 65°C.

The probe was removed by washes at 65°C in 50/50 HB/2× SSC for 20 min, 25/75 HB/2× SSC for 20 min, 2× SSC for 30 min and two washes in 0.2× SSC with 0.3% CHAPS (Sigma) for 30 min. Embryos were washed twice in MaNaT (100 mM maleic acid/150 mM NaCl pH 7.5 and 0.1% Tween-20) for 10 min at room temperature and then blocked for 1 h with blocking buffer [BB: MaNaT containing 1% blocking reagent (Roche Molecular Biochemicals)]. Antibody incubation was performed for 1 h in a fresh solution of BB containing a 1:5,000 dilution of the affinity-purified anti-digoxigenin antibody coupled to alkaline phosphatase (Roche Molecular Biochemicals). This was followed by extensive washing in MaNaT for 2×5 min and 6×15 min.

Embryos were washed with alkaline phosphatase buffer (APB: 100 mM Tris, pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl) after which the embryos were stained in a fresh tube with APB containing PVA (10% polyvinyl alcohol, 100 kDa), 3.5 µl NBT (nitro blue tetrazolium; 75 mg/ml; Roche Molecular Biochemicals) and 3.5 µl BCIP (5-bromo-4-chloro-3-indolylphosphate; 50 mg/ml; Roche Molecular Biochemicals) per ml of buffer. After reaching an adequate signal compared to the background, the reaction was stopped by washing twice with TBS-T. Subsequently, the embryos were refixed with MEMPFA-T for 20 min, washed twice with TBS-T, once with 70% EtOH and stored in 70% EtOH at 4°C.

For photography, embryos were dehydrated with 2×5-min washes in DMP (di-methoxy propane; Fluka) that was activated with HCl (5 µl 10 N HCl – 37% – per 10 ml DMP), 2×5 min in 100% Histoclear (National Diagnostics) and mounted in Canada balsam (BDH). Pictures were taken with an Olympus BX-50 microscope with DIC optics (Nomarski) on 50 ASA Fuji Velvia films.

For a more accurate analysis of the expression pattern, some embryos were embedded in paraffin, cut at 5 µm thickness, mounted and photographed as described in Joly et al. (1997).

Results

Cloning of *Pv-sna1* and *Pv-sna2*, two *snail* orthologues in *P. vulgata*

PCR amplification with primers Sna1W and Sna4C allowed us to isolate a 196-bp fragment (Fig. 1), that was cloned in the appropriate vector (see Materials and methods). Twenty clones from two independent PCR amplifications were sequenced. All of them share the same sequence which displayed 88.3% and 96.7% amino acid sequence identity with the *snail-box* of the *snail* and *escargot* genes of *D. melanogaster*. This PCR fragment was used as a probe to screen a genomic library and a 16-h-old trochophore larva cDNA library from *P. vulgata*. From the cDNA library, we isolated a cDNA with a

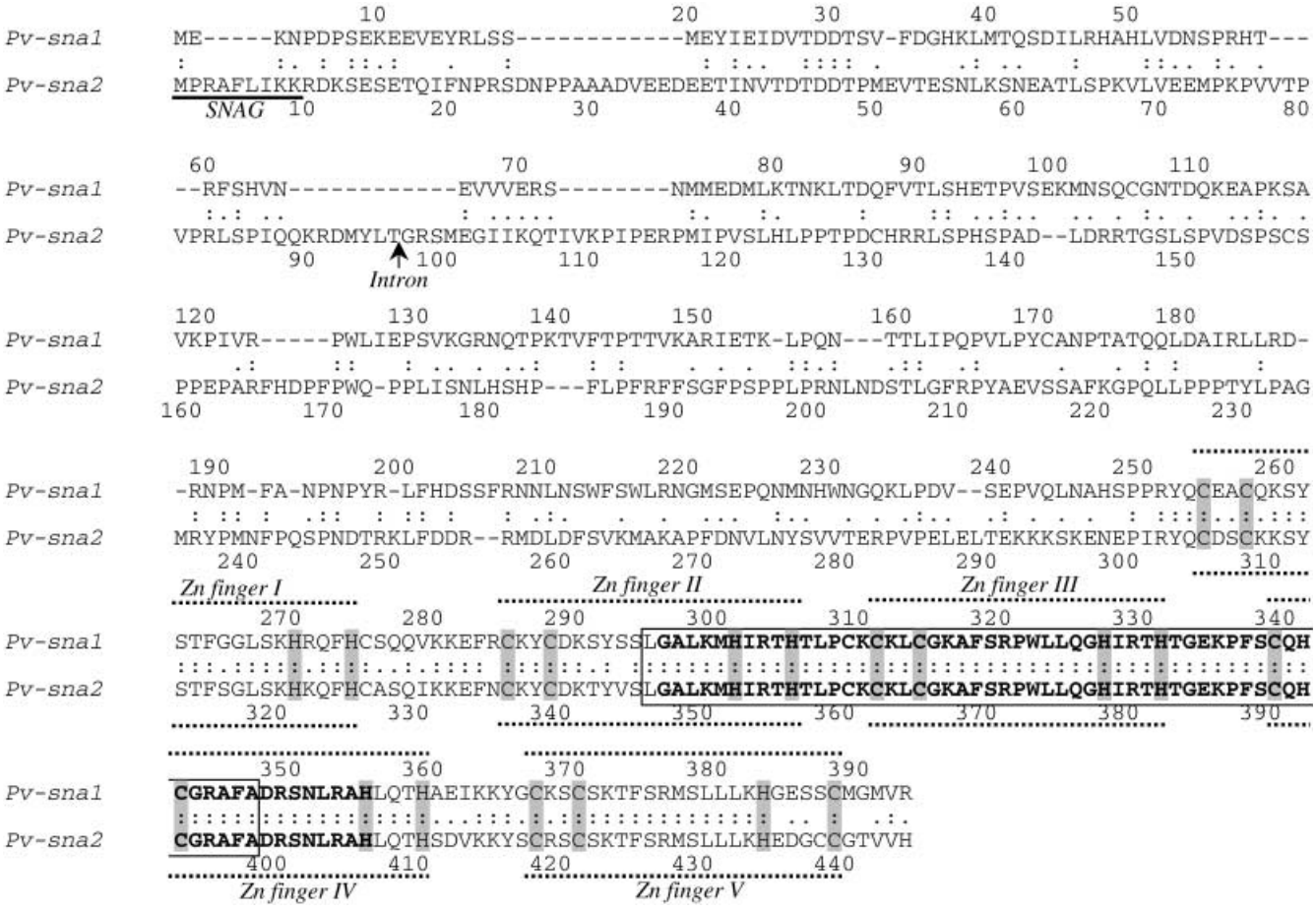

Fig. 1 *Patella vulgata snail* genes. Alignment of the deduced amino acids sequences of the two *Pv-sna* genes. Identities between residues is indicated by a *colon*, and similarities by a *point*. Complete cDNA sequences are available in GenBank under accession numbers AY049727 (*Pv-Sna1*) and AY049791 (*Pv-Sna2*). Position of the 2 kb intron in *Pv-sna2* is indicated on the corresponding sequence by an *arrow* (partial genomic sequence for this gene is available in GenBank under accession number AY049790). The repressor SNAG motive (Grimes et al. 1996) located at the amino terminus of *Pv-Sna2* is *underlined*. All zinc fingers (I, II, III, IV, V) of the "*snail*-box" are indicated by a *dotted line* on each sequence, and the cystein and histidine residues essential to the five zinc fingers are *highlighted in grey*. The *box* indicates the area of high sequence identity at the nucleotide level. The region corresponding to the 196-bp fragment amplified by PCR using primers Sna1W and Sna4C is in bold

1.6-kb fragment that possessed a 1,179-bp open reading frame (ORF) surrounded by a 58-bp 5′ untranslated region (UTR) and a 328-bp 3′ UTR. This ORF coded for a 393-amino-acid (45.1 kDa) putative protein that possessed a classical 5 zinc finger *snail*-box (Fig. 1). We called it *Pv-sna1*.

From the genomic library, a second *snail* gene was isolated and was sub-cloned by a PCR approach (see Materials and methods). This led us to isolate a 3.6-kb fragment that possessed a 307-bp 5′ UTR and a 1,104-bp ORF interrupted by a 2-kb intron. The 3′ part of the coding region, the exact intron/exon boundaries, and the 3′ UTR of this gene have been determined by PCR amplification on the 16-h cDNA library (see Materials and methods). This gave us a final 1,332-bp ORF that coded for a putative 444-amino-acid (50.4 kDa) protein, and which also possessed a classical 5 zinc fingers *snail*-box. It also possessed a SNAG repressor domain (Grimes et al. 1996) which is not present in *Pv-sna1* (Fig 1). This gene was the second member of the *snail*-family identified in *P. vulgata* and was called *Pv-sna2*. Neither of these two genes show the presence of a so-called *scratch* or a *slug* domain (Manzanares et al. 2001).

Sequence comparison of the two *Pv-sna* genes

The two *Pv-sna* genes share 39.8% amino acid sequence identity with 86% identity for the *snail*-box itself. The sequence conservation outside of this putative DNA binding site is not relevant as we tested it by a statistical reshuffling alignment analysis using PRSS (Pearson 1997). Indeed, the reshuffled proteins produce a better alignment score in the majority of the cases (145 times on 200 tries). From this, we could easily conclude that only the *snail*-box is conserved between these two proteins. Surprisingly, we identified a 53-amino-acid fragment inside the *snail-box*, which shared 100% identity between the two proteins (Fig. 1). Moreover, this fragment presents a 98.7% identity when we look at the nucleotide level (2 silent substitutions on 159 nucleotides;

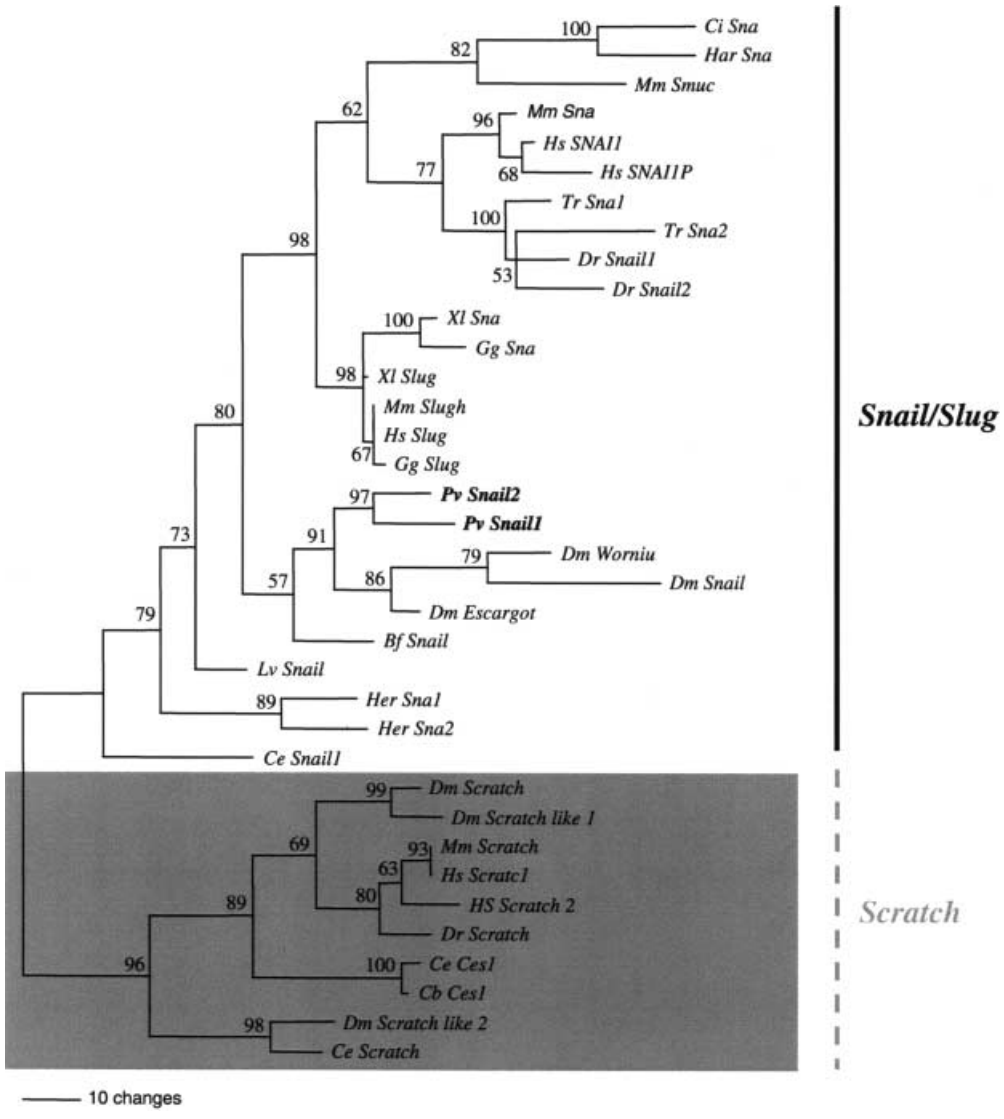

Fig. 2 Maximum likelihood tree of the *snail* family. To perform our analysis, we used the ProtML program in the MOLPHY package (Adachi and Hasegawa 1996) and chose the JTT-F model for nuclear proteins (Jones et al. 1992) with local rearrangement of the NJ tree obtained by PAUP (Swofford 1998). This tree has an Ln L=–4227.01+–290.09 and the bootstrap probabilities (RELL – BP, 10,000 replications; Hasegawa and Kishino 1994) are *indicated on each node*. Branches with BP lower than 50% were collapsed. BP-RELL value for the (Hs_Slug, Mm_Slug) cluster is 60%. Sequence alignments, and complete references of the sequences are available on the journal's web site (electronic supplementary material Figs A and B). Species abbreviations used: *Bf Branchiostoma floridae, Cb Caenorhabditis briggsae, Ce Caenorhabditis elegans, Ci Ciona intestinalis, Dm Drosophila melanogaster, Dr Danio rerio, Gg Gallus gallus, Har Halocynthia roretzi, Her Helobdella robusta, Hs Homo sapiens, Pv Patella vulgata, Lv Lytechinus variegatus, Mm Mus musculus, Tr Takifugu rubripes, Xl Xenopus laevis*

data not shown). This contributes to explaining why we have only identified one fragment after PCR amplification, although there are at least two paralogous *snail* genes in *P. vulgata* (Fig. 1).

Sequence comparison of the *snail*-related proteins among Bilateria and phylogenetic analysis of the *snail*-box

To date, about 35 members of the *snail* family have been isolated and sequenced. In order to compare these proteins all together and to avoid comparison biases due to different protein lengths, we used SSPA score identity (Brocchieri and Karlin 1998) instead of classical percentage identity. This analysis showed that protein pairwise identity scores of the sequences outside of the *snail* motive are quite low (below 40%) except for the vertebrate *slug* proteins (above 84%). The conservation of the *snail*-box is much higher and even much higher in vertebrates (score identity >77%) than in ecdysozoans (score identity >51%; see supplementary material available at http://link.springer-ny.com, supplementary Fig. A).

These results led us to use only the *snail* box in order to try to resolve the orthology relationships between members of the *snail* family. By distance analysis [through the Neighbour-Joining (NJ) algorithm] we retrieved results quite similar to those of Manzanares et al. (2001), namely that the sequences distribute themselves into two broad groups, respectively the *snail/slug* and the *scratch* sub-families (not shown). Since species belonging to the whole range of Bilateria are found in each of these two sub-families, this is indicative of a pre-bilaterian duplication of the ancestral *snail* gene to yield the two main paralogy groups. The NJ tree also shows that the *slug* genes originate within the *snail* sub-family. However, the tree also provides strong suspicions of a number of artefacts related to the fact that the various sequences appear to evolve at different rates. We therefore performed a maximum likelihood (ML) analysis using the JTT-F model (Fig. 2). This new analysis gave us comparable results and did not fully resolve some long-branch

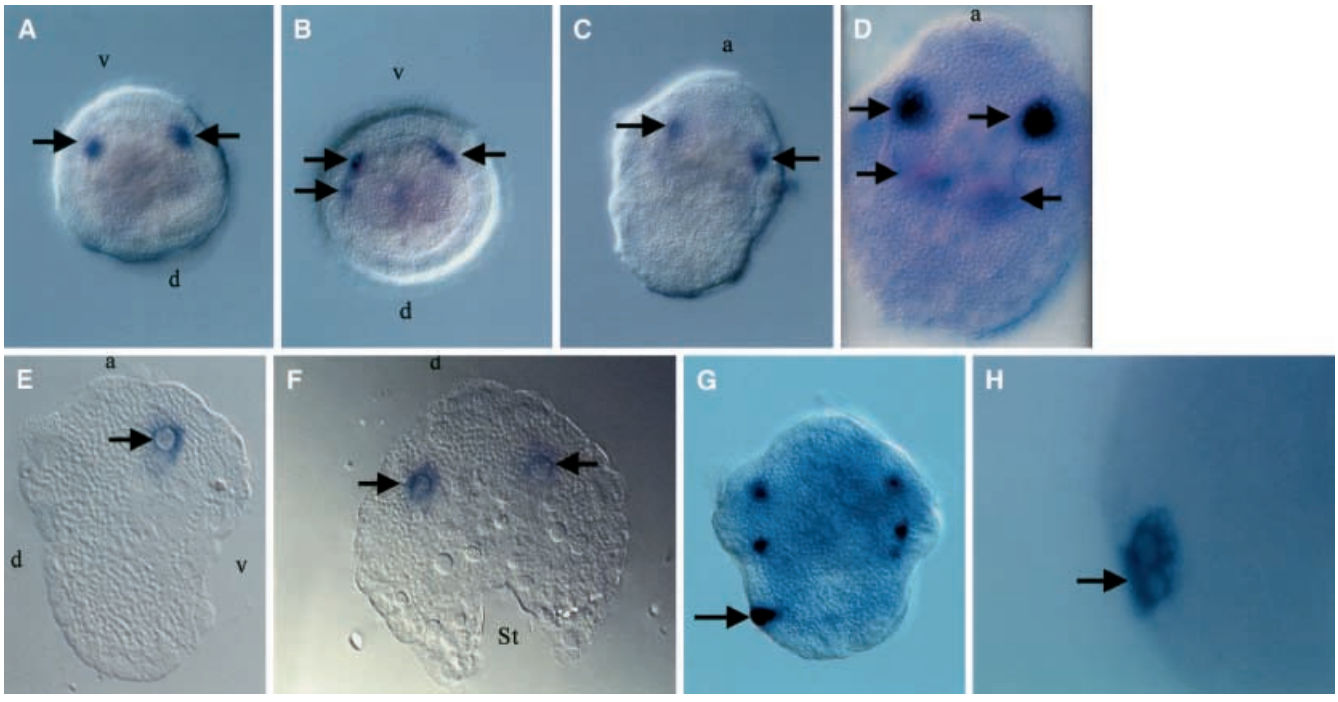

Fig. 3 A–H *Pv-sna1* expression. Full length cDNA was used as a probe (see Materials and methods). *Pv-Sna1* expression is indicated by *arrows*. Trochophore larvae size is 200×160 microns. **A**–**F** Sixteen-hour-old trochophores. The expression is first restricted to two pretrochal regions. Shortly after two posttrochal regions also express *Pv-Sna1*. **A**, **B** Apical view. **C**, **D** Ventral view. **E** Lateral view, cross section. **F** Apical view, cross section. **G**, **H** In 24-h-old trochophores, a new site of *Pv-Sna1* expression appears on the right part of the foot. **G** Ventral view, **H** magnification of the foot staining (*a* apical side, *d* dorsal side, *v* ventral side, *St* stomodeum)

attraction artefacts (see Discussion). Some improvement was however observed. The *snail/slug* sub-family was again clearly split from the *scratch* sub-family. Within the *snail/slug* portion of the tree, the three *Drosophila* genes, *snail*, *escargot* and *worniu*, form a monophyletic group suggesting that they result from a relatively recent set of duplications that could be specific to arthropods. All the deuterostomes *snail/slug* genes, with the notable exception of the sea urchin and amphioxus *snail*, now form a monophyletic group. Surprisingly, however, two vertebrate *snail* sequences, that of chick and *Xenopus*, do not emerge with the other *snail* genes. Instead they form an independent clade also comprising the *slug* genes. The latter point clearly shows that *slug* genes originate from a duplication that occurred early on in the vertebrate lineage. Upon examination of the tree, it is immediately noted that the echinoderm, ascidian, cephalochordate, and protostome *snail* sequences are long branched while the vertebrate (with the exception of the mouse *smuc* gene) and *slug* ones are short. This suggests that the sea urchin, *Amphioxus*, as well as the protostome sequences, were pulled together and at a distance from the other sequences by the well-known long branch artefact. Similarly, the basal position of the nematode sequences within the *snail* portion of this tree are most likely due to the same artefact, which is known to plague most of *Caenorhabditis elegans* genes. Finally the lack of clustering of all the vertebrate sequences may also be accounted for on the same basis, the fish and mammalian sequences having apparently undergone more substitutions than the chick and amphibian ones. In spite of these difficulties, it can be seen that the two *P. vulgata* genes emerge strongly within the protostome sequences. Thus, they are clearly *snail* orthologues.

Pv-sna1 expression pattern

The expression pattern of *Pv-sna1* was determined from oocyte stage to 24-h-old trochophore larvae (24 h after first cleavage). No maternal expression has been detected so far for this gene and the first sign of expression was seen in 16-h trochophore larvae (Fig. 3). At this stage the staining was first restricted to two pretrochal regions (Fig. 3A). Shortly after, two new areas located in the posttrochal region also appeared to express *Pv-Sna1* (Fig. 3B–D). Cross sections showed that these four areas corresponded to four unique large cells (Fig. 3E, F). Two of these cells were located ventrally in the pretrochal area and the two others were situated more dorsally in the posttrochal area (Fig. 3D). This expression persisted at least until the 24th hour of development. At this stage, a second domain of expression appeared in several cells on the right side of the larval foot (Fig. 3G, H).

Pv-sna2 expression pattern

Expression of *Pv-sna2* was not detected before the 8th hour after first cleavage (Fig. 4). The staining was

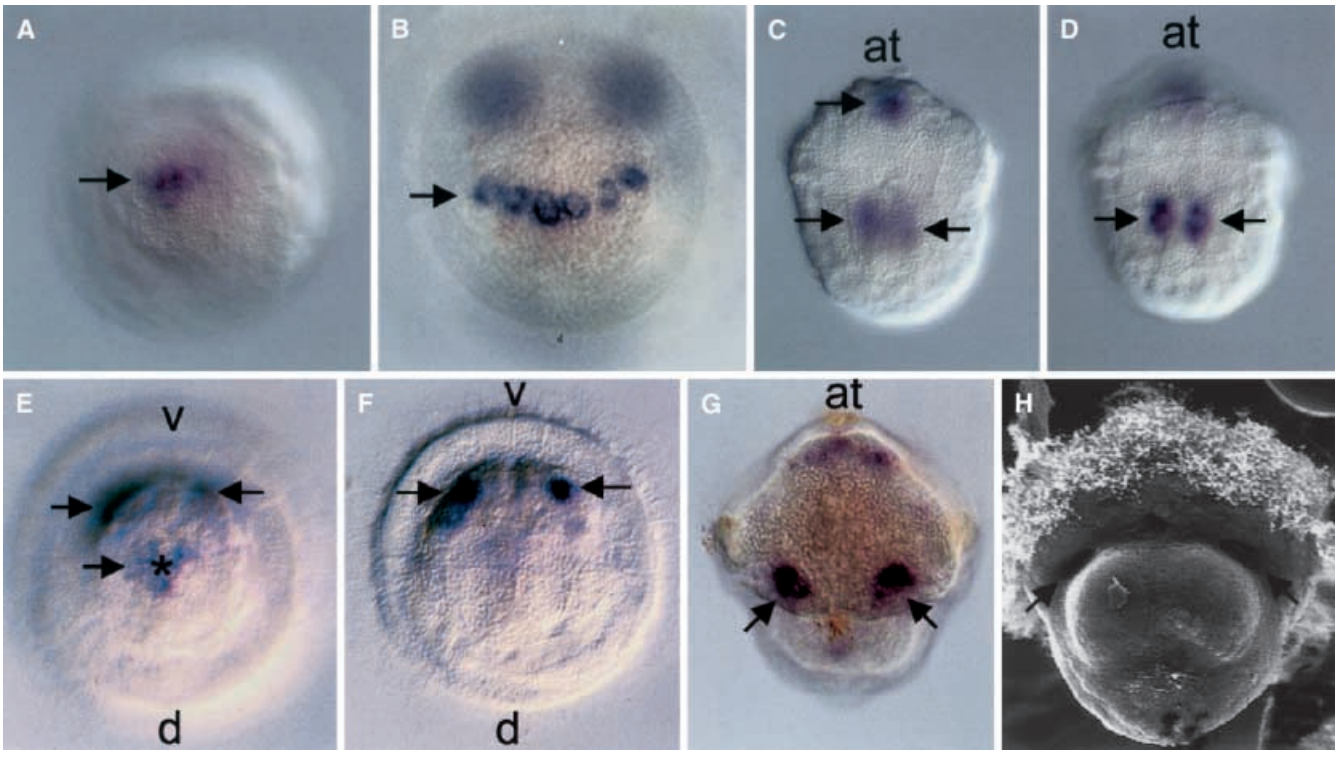

Fig. 4 A–H *Pv-sna2* expression. PCR fragment amplified by primers T3 and Rev2 on a cDNA template was used as a probe (see Materials and methods). *Pv-Sna2* expression is indicated by *arrows*. Trochophore larvae size is 200×160 microns. **A** Eight-hour-old embryos, first expression in seen in the apical plate; apical view. **B** Twenty-four hour-old trochophores, expression persists in the apical plate and begins in the invaginating mantle fold; apical view. **C**, **D** Twenty-hour-old trochophores, ventral view of the same trochophore with two different focus planes. **E–G** Twenty-four hour-old trochophores, **E**, **F** apical view of the same trochophore with different focus planes; **G** ventral view. **H** Scanning electron micrograph, ventral view, of a 24-h-old trochophore larva; *arrows* indicate the mantle invaginations (*at* apical tuft, *d* dorsal side, *v* ventral side)

first restricted to a few cells in the apical plate located at the base of the apical tuft (Fig. 4A). This apical expression was maintained until at least 24 h after first cleavage (Fig. 4B–F). Later, two new areas became stained with the *Pv-sna2* probe (Fig. 4C–G). At the 11th hour of development, they were located in the posttrochal ectoderm (Fig. 4C, D), and at 24 h they were located in the ectoderm "behind" the prototroch on the ventral side, on either side of the foot with the blastopore in between (Fig. 4E–G). This expression was at the base of two ectodermal invaginations that become the mantle folds (Fig. 4H).

Discussion

Sequence analysis of the *P. vulgata snail* genes

We have identified two new novel lophotrochozoan members of the *snail* family gene family, *Pv-sna1* and *Pv-sna2* (Fig. 1), that represent an independent duplication within the *snail* sub-family (Fig. 2). The level of sequence conservation between *Pv-sna1* and *Pv-sna2* is low outside but high inside the *snail*-box, at the protein as well as the nucleotide level (Fig. 2). This high level of nucleotide sequence identity could be explained by supposing that a gene conversion occurred between the two limpet genes in this region. If this was the case, the sequence identity of the first and fifth zinc fingers was apparently not high enough for these fingers to be included in the conversion. This could mean that selective forces on the excluded fingers (I and V) were less important than on the included fingers (II, III and IV). This is corroborated by the fact that the first zinc finger is sometimes missing in some of the *snail* orthologs (in particular in vertebrate *snail* genes and two *C. elegans* sequences; Manzanares et al. 2001), and this might indicate that this zinc finger is perhaps not as important for the function of the *Snail* proteins as the other fingers.

Expression of the *snail* genes in *P. vulgata*

We analysed the expression patterns of *Pv-sna1* and *Pv-sna2* during early embryonic and larval development of *Patella*. We set out to clone *snail* orthologs because we were interested in their possible role in mesoderm development. We thus expected these genes to be expressed around the time of mesoderm induction, and/or during the development of the mesodermal tissue. The mesoderm is induced between the 32- and 40-cell stage (van den Biggelaar 1977) but we did not find any expression of either *Patella snail* gene at this stage. During

gastrulation, the mesoderm moves to the inside of the embryo; we therefore would expect the *snail* orthologs to be expressed internally. We only found four *Pv-sna1*-expressing cells located inside the larva but if they are probably not of ectodermal origin, we were not able to assign these cells to any germ layers (Fig. 3E and F). As they do not fit with any well known structures (Dictus and Damen 1997), the fate of these four *Pv-sna1*-expressing cells remains unknown.

All other sites of expression we have found for the two *snail* genes are ectodermal. The ectodermal *Pv-sna1*-expressing cells are located asymmetrically at the right side of the foot (Fig. 3G and H). Although the significance of this staining remains cryptic it is very tentatively suggested that, as in the chick and mouse (Isaac et al. 1997; Sefton et al. 1998), this expression could be related to the control of left-right asymmetry in this region.

The ectodermal *Pv-sna2* expression in the apical plate is at the base of the apical tuft (Fig. 4A and B). The tuft develops into a neural sensory structure suggesting that *Pv-sna2* could be involved in the neural differentiation of this structure. This situation can be readily compared to the neural expression and role in neural differentiation of many of the *snail* family members (Roark et al. 1995; Ashraf et al. 1999; Metzstein and Horvitz 1999).

The other site of *Pv-sna2* expression might correspond to the base of the mantle folds (Fig. 4G, H). Cells in this region are forming pockets that will expand inwards and become the epithelial covering of the cavity enclosed by the shell. Although there is no direct evidence, an EMT process could be involved in this pocketing of the ectoderm. A function in EMT and cell-motility control has been found for vertebrate members of the *snail* gene family (Batlle et al. 2000; Cano et al. 2000).

Recent results on mouse (Cano et al. 2000) and human (Batlle et al. 2000) have shown that *snail* is able to bind the E-cadherin promoter and thus to repress the expression of this molecule that plays a central role in the control of cell adhesiveness and migration. Also, *snail* genes have repeatedly been linked to EMT, e.g. in neural crest formation (Carl et al. 1999) and heart development (Romano and Runyan 1999) in vertebrates, and in tumour formation (Cano et al. 2000). This has led to the suggestion that EMT (Manzanares et al. 2001) and the control of cell motility (Hemavathy et al. 2000) are ancient functions of this gene family.

Ancestral function of the *Snail* gene

Although the *snail* gene is conserved in lophotrochozoans (Goldstein et al. 2001; this work), the patterns of expression of the two *Patella snail* homologues we have characterised appear to be dynamic and different from what might be expected if these genes were involved in mesoderm formation. Based on the timing of expression, and special localisation, it seems more likely that this expression is not linked to a role in mesoderm formation. The other expression sites of the *Patella snail* genes (in the foot, apical tuft and mantle cavities) are ectodermal.

The probable expression of the *snail Patella* orthologs in involuting cells of the mantle tips is compatible with the control of EMT and cell motility at this site during the early development of *Patella*. Thus, based on data presented here and on data from the literature, instead of being a key gene ancestrally involved in mesoderm formation, we suggest that the ancestral role of the *snail* gene could be in the regulation of E-cadherin and the control of epithelio-mesenchymal transitions.

With this idea in mind, many of the expression patterns of members of this gene family in the fly as well as in vertebrates can be related by their roles in cell movement during gastrulation (Oda et al. 1998), mesoderm formation, neuroblast delamination, and neural crest cell migration.

The ancestral role that we propose for the *snail* genes allows us to speculate on how the mesoderm evolved. In order for this third germ layer to arise, cells would have had to leave the epithelium of the ectoderm and/or the endoderm and move in between the two cell sheets. For this to occur, a process similar to EMT was required. Therefore, a prerequisite for the evolution of the mesoderm was the ability for cells to undergo an epithelial-mesenchymal transition. The connection between *snail* and E-cadherin could perhaps be considered as a kind of pre-adaptation for the evolution of the mesoderm.

Acknowledgements This work was supported by the Centre National de la Recherche Scientifique, Université Paris-Sud, and "Programme Génome" from the Centre National de la Recherche Scientifique. O.L. was supported by a fellowship from the Ministère de l'Education Nationale, de la Recherche et de l'Enseignement Supérieur. A.J.N. was supported by Earth and Life Sciences (ALW), which are subsidised by the Netherlands Organisation for Scientific Research (NWO). We thank Jo van den Biggelaar for his inspiration and constant interest, Nicolas Lartillot for his participation in this work and many fruitful discussions, and Angela Nieto for stimulating discussions and communicating results prior to publication.

References

Adachi J, Hasegawa M (1996) MOLPHY: programs for molecular phylogenetics. Institute of Statistical Mathematics, Tokyo

Adoutte A, Balavoine G, Lartillot N, Lespinet O, Prud'homme B, de Rosa R (2000) Special feature: the new animal phylogeny: reliability and implications. Proc Natl Acad Sci USA 97:4453–4456

Aguinaldo AM, Turbeville JM, Linford LS, Rivera MC, Garey JR, Raff RA, Lake JA (1997) Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387:489–493

Alberga A, Boulay JL, Kempe E, Dennefeld C, Haenlin M (1991) The *snail* gene required for mesoderm formation in *Drosophila* is expressed dynamically in derivatives of all three germ layers. Development 111:983–992

Ashraf SI, Hu X, Roote J, Ip YT (1999) The mesoderm determinant *snail* collaborates with related zinc-finger proteins to control *Drosophila* neurogenesis. EMBO J 18:6426–6438

Batlle E, Sancho E, Franci C, Dominguez D, Monfar M, Baulida J, Garcia De Herreros A (2000) The transcription factor *snail* is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol 2:84–89

Biggelaar JAM van den (1977) Development of dorsoventral polarity and mesentoblast determination in *Patella vulgata*. J Morphol 154:157–186

Boulay JL, Dennefeld C, Alberga A (1987) The *Drosophila* developmental gene *snail* encodes a protein with nucleic acid binding fingers. Nature 330:395–398

Brocchieri L, Karlin S (1998) A symmetric-iterated multiple alignment of protein sequences. J Mol Biol 276:249–264

Cai Y, Chia W, Xiaohang Y (2001) A family of Snail-related zinc finger proteins regulates two distinct and parallel mechanisms that mediate *Drosophila* neuroblast asymetric divisions. EMBO J 20:1704–1714

Cano A, Perez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, del Barrio MG, Portillo F, Nieto MA (2000) The transcription factor *snail* controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol 2:76–83

Carl TF, Dufton C, Hanken J, Klymkowsky MW (1999) Inhibition of neural crest migration in *Xenopus* using antisense *Slug* RNA. Dev Biol 213:101–115

Casal J, Leptin M (1996) Identification of novel genes in *Drosophila* reveals the complex regulation of early gene activity in the mesoderm. Proc Natl Acad Sci USA 93:10327–10332

Corbo JC, Erives A, Di Gregorio A, Chang A, Levine M (1997) Dorsoventral patterning of the vertebrate neural tube is conserved in a protochordate. Development 124:2335–2344

Dictus WJAG, Damen P (1997) Cell-lineage and clonal-contribution map of the trochophore larva of *Patella vulgata* (Mollusca). Mech Dev 62:213–226

Goldstein B, Leviten MW, Weisblat DA (2001) *Dorsal* and *Snail* homologs in leech development. Dev Genes Evol 211:329–337

Grau Y, Carteret C, Simpson P (1984) Mutations and chromosomal rearrangement affecting the expression of *Snail*, a gene involved in embryonic patterning in *Drosophila melanogaster*. Genetics 108:347–360

Grimes HL, Chan TO, Zweidler-McKay PA, Tong B, Tsichlis PN (1996) The Gfi-1 proto-oncoprotein contains a novel transcriptional repressor domain, SNAG, and inhibits G1 arrest induced by interleukin-2 withdrawal. Mol Cell Biol 16:6263–6272

Halanych KM, Bacheller JD, Aguinaldo AM, Liva SM, Hillis DM, Lake JA (1995) 18S rDNA evidence that the lophophorates are protostome animals. Science 267:1641–1643

Hammerschmidt M, Nusslein-Volhard C (1993) The expression of a zebrafish gene homologous to *Drosophila snail* suggests a conserved function in invertebrate and vertebrate gastrulation. Development 119:1107–1118

Hasegawa M, Kishino H (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. Mol Biol Evol 11:142–145

Hemavathy K, Meng X, Ip YT (1997) Differential regulation of gastrulation and neuroectodermal gene expression by *Snail* in the *Drosophila* embryo. Development 124:3683–3691

Hemavathy K, Ashraf SI, Ip YT (2000) *Snail/Slug* family of repressors: slowly going into the fast lane of development and cancer. Gene 257:1–12

Houtzager E (1998) The role of connexins and the function of the dorsal endoderm in embryogenesis of *Xenopus laevis*. Utrecht University, Utrecht

Ip YT, Levine M, Bier E (1994) Neurogenic expression of *snail* is controlled by separable CNS and PNS promoter elements. Development 120:199–207

Isaac A, Sargent MG, Cooke J (1997) Control of vertebrate left-right asymmetry by a *snail*-related zinc finger gene. Science 275:1301–1304

Joly JS, Bourrat F, Nguyen V, Chourrout D (1997) *Ol-Prx 3*, a member of an additional class of homeobox genes, is unimodally expressed in several domains of the developing and adult central nervous system of the medaka (*Oryzias latipes*). Proc Natl Acad Sci USA 94:12987–12992

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Kasai J, Nambu JR, Lieberman PM, Crews ST (1992) Dorsal-ventral patterning in *Drosophila*: DNA binding of snail protein to the single-minded gene. Proc Natl Acad Sci USA 89:3414–3418

Kosman D, Ip YT, Levine M, Arora K (1991) Establishment of the mesoderm-neuroectoderm boundary in the *Drosophila* embryo. Science 254:118–122

Langeland JA, Tomsa JM, Jackman WR Jr, Kimmel CB (1998) An amphioxus *snail* gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. Dev Genes Evol 208:569–577

Leptin M (1991) *twist* and *snail* as positive and negative regulators during *Drosophila* mesoderm development. Genes Dev 5:1568–1576

Linker C, Bronner-Fraser M, Mayor R (2000) Relationship between gene expression domains of *Xsnail*, *Xslug*, and *Xtwist* and cell movement in the prospective neural crest of *Xenopus*. Dev Biol 224:215–225

Loon AE van, Goedemans H, Daemen A, van der Kamp A, van den Biggelaar J (1993) Actin genes expressed during early development of *Patella vulgata*. Roux's Arch Dev Biol 203:77–84

Manzanares M, Locascio A, Nieto MA (2001) The increasing complexity of the *Snail* gene superfamily in metazoan evolution. Trends Genet 17:178–181

Mayor R, Essex LJ, Bennett MF, Sargent MG (1993) Distinct elements of the *xsna* promoter are required for mesodermal and ectodermal expression. Development 119:661–671

Mayor R, Morgan R, Sargent MG (1995) Induction of the prospective neural crest of *Xenopus*. Development 121:767–777

Metzstein MM, Horvitz HR (1999) The *C. elegans* cell death specification gene *ces-1* encodes a snail family zinc finger protein. Mol Cell 4:309–319

Nakakura RK, Watkins DN, Schuebel KE, Sriuranpong V, Borges MW, Nelkin BD, Ball DW (2001) Mammalian *Scratch*: a neural-specific *Snail* family transcriptional repressor. Proc Natl Acad Sci USA 98:4010–4015

Nieto MA, Bennett MF, Sargent MG, Wilkinson DG (1992) Cloning and developmental expression of *Sna*, a murine homologue of the *Drosophila snail* gene. Development 116:227–237

Nieto MA, Sargent MG, Wilkinson DG, Cooke J (1994) Control of cell behavior during vertebrate development by *Slug*, a zinc finger gene. Science 264:835–839

Nüsslein-Volhard C, Wieschaus E, Kluding H (1984) Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster.* I. Zygotic loci on the second chromosome. Roux's Arch Dev Biol 193:267–282

Oda H, Tsukita S, Takeichi M (1998) Dynamic behavior of the cadherin-based cell-cell adhesion system during Drosophila gastrulation. Dev Biol 203:435–450

Pearson WR (1997) FASTA. University of Virginia, Charlottesville, Va.

Peterson KJ, Eernisse DJ (2001) Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18 rDNA gene sequences. Evol Dev 3:170–205

Roark M, Sturtevant MA, Emery J, Vaessin H, Grell E, Bier E (1995) *scratch*, a pan-neural gene encoding a zinc finger protein related to *snail*, promotes neuronal development. Genes Dev 9:2384–2398

Romano LA, Runyan RB (1999) Slug is a mediator of epithelial-mesenchymal cell transformation in the developing chicken heart. Dev Biol 212:243–254

Ros MA, Sefton M, Nieto MA (1997) *Slug*, a zinc finger gene previously implicated in the early patterning of the mesoderm and the neural crest, is also involved in chick limb development. Development 124:1821–1829

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sargent MG, Bennett MF (1990) Identification in *Xenopus* of a structural homologue of the *Drosophila* gene *snail*. Development 109:967–973

Sefton M, Sanchez S, Nieto MA (1998) Conserved and divergent roles for members of the Snail family of transcription factors in the chick and mouse embryo. Development 125:3111–3121

Simpson P (1983) Maternal-zygotic gene interactions during formation of the dorsoventral pattern in *Drosophila* embryos. Genetics 105:615–632

Smith DE, Franco del Amo F, Gridley T (1992) Isolation of *Sna*, a mouse gene homologous to the *Drosophila* genes *snail* and *escargot*: its expression pattern suggests multiple roles during postimplantation development. Development 116:1033–1039

Sommer RJ, Tautz D (1994) Expression patterns of *twist* and *snail* in *Tribolium* (Coleoptera) suggest a homologous formation of mesoderm in long and short germ band insects. Dev Genet 15:32–37

Swofford DL (1998) PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods), version 4. Sinauer, Sunderland, Mass.

Thisse C, Thisse B, Schilling TF, Postlethwait JH (1993) Structure of the zebrafish *snail1* gene and its expression in wild-type, *spadetail* and *no tail* mutant embryos. Development 119:1203–1215

Thisse C, Thisse B, Postlethwait JH (1995) Expression of *snail2*, a second member of the zebrafish *snail* family, in cephalic mesendoderm and presumptive neural crest of wild-type and *spadetail* mutant embryos. Dev Biol 172:86–99

Verdonk NH, van den Biggelaar JAM (1983) Early development and the formation of germ layers. In: Verdonk NH, van den Biggelaar JAM (eds) The Mollusca: development, vol 3. Academic Press, New York, pp 91–122

Wada S, Saiga H (1999) Cloning and embryonic expression of *Hrsna*, a *snail* family gene of the ascidian *Halocynthia roretzi*: implication in the origins of mechanisms for mesoderm specification and body axis formation in chordates. Dev Growth Differ 41:9–18

Whiteley M, Noguchi PD, Sensabaugh SM, Odenwald WF, Kassis JA (1992) The *Drosophila* gene *escargot* encodes a zinc finger motif found in *snail*- related genes. Mech Dev 36:117–127