Developmental Biology 409 (2016) 277–287

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Evolution of Developmental Control Mechanisms

Msxb is a core component of the genetic circuitry specifying the dorsal and ventral neurogenic midlines in the ascidian embryo

Agnès Roure, Sébastien Darras *

Sorbonne Universités, UPMC Univ Paris 06, CNRS, Biologie Intégrative des Organismes Marins (BIOM), Observatoire Océanologique, F-66650 Banyuls/Mer, France

ARTICLE INFO

Article history:
Received 31 August 2015
Received in revised form
5 November 2015
Accepted 6 November 2015
Available online 18 November 2015

Keywords:
Ascidian
Gene regulatory network
Transcription factors
Peripheral nervous system
Neurogenic

ABSTRACT

The tail ascidian larval peripheral nervous system is made up of epidermal sensory neurons distributed more or less regularly in ventral and dorsal midlines. Their formation occurs in two-steps: the ventral and dorsal midlines are induced as neurogenic territories by Fgf9/16/20 and Admp respectively. The Delta2/Notch interaction then controls the number of neurons that form. The genetic machinery acting between the inductive processes taking place before gastrulation and neuron specification at tailbud stages are largely unknown. The analysis of seven transcription factors expressed in the forming midlines revealed an unexpected complexity and dynamic of gene expression. Their systematic overexpression confirmed that these genes do not interact following a linear cascade of activation. However, the integration of our data revealed the distinct key roles of the two upstream factors *Msxb* and *Nkx-C* that are the earliest expressed genes and the only ones able to induce neurogenic midline and ESN formation. Our data suggest that *Msxb* would be the primary midline gene integrating inputs from the ventral and dorsal inducers and launching a pan-midline transcriptional program. *Nkx-C* would be involved in tail tip specification, in maintenance of the pan-midline network and in a posterior to anterior wave controlling differentiation.

1. Introduction

The peripheral nervous system (PNS) is defined as the part of the nervous system outside of the central nervous system (CNS), the brain and the spinal cord in vertebrates. It enables animals to receive a large part of their information from the environment and is thus fundamental to adjusting their behavior based on external cues. In vertebrates, the PNS arises during embryonic development from discrete dorsally located structures, the neural crest and the placodes (Baker and Bronner-Fraser, 2001, 1997). These structures are unique to vertebrates and their emergence is thought to have been instrumental in the switch from passive filter feeding behavior of invertebrates to active predation typical of vertebrates (Glenn Northcutt, 2005; Simoes-Costa and Bronner, 2013). Interestingly, in ascidians that belong to the sister group of vertebrates, rudimentary cranial neural crest and neurogenic placodes have recently been identified (Abitua et al., 2015, 2012). We have previously proposed that the ascidian tail dorsal midline ectoderm could be homologous to the vertebrate caudal neural plate border (Pasini et al., 2006). This midline is a neurogenic territory that gives rise to epidermal sensory neurons (ESNs) that are thought to be mechanoreceptors. It is specified together with CNS early during development (32-cell stage) through an inductive process involving a Fgf ligand. We have recently characterized part of the genetic interactions that lead to the transcription of *Msxb*, a transcription factor essential for dorsal midline fate (Imai et al., 2006; Roure et al., 2014). Following midline fate acquisition, the Notch pathway controls the number of sensory neurons within this neurogenic domain by launching a regulatory cascade involving several proneural genes and a microRNA miR-124 (Akanuma et al., 2002; Chen et al., 2011; Joyce Tang et al., 2013; Pasini et al., 2006).

Interestingly, in the ascidian *Ciona intestinalis*, tail ESNs are not limited to the dorsal midline. They are present along the ventral tail ectodermal midline. Similarly to the dorsal ESNs, they are specified in two steps: a neurogenic midline is induced, then ESNs are selected through the Notch pathway (Pasini et al., 2006). The primary inducer is however different: this is Admp, a Bmp ligand, which acts at gastrula stages. While a ventral neurogenic territory does not exist in vertebrates, it has recently been shown that amphioxus (cephalochordate) possess a similar region that is specified like *Ciona* ventral midline (Bmp signaling defines a ventral neurogenic territory, then Notch controls the number of ESNs) (Lu et al., 2012). While dorsal and ventral midlines have

* Corresponding author.
E-mail address: sebastien.darras@obs-banyuls.fr (S. Darras).

http://dx.doi.org/10.1016/j.ydbio.2015.11.009
0012-1606/

different spatial origins and are induced by distinct ligands, they appear to form a uniform tissue as revealed by cell organization and ESNs differentiation. In particular they are characterized by a set of genes, the midline genes, that are uniformly expressed in the midline tissue with no detectable dorso-ventral differences (Pasini et al., 2006). While the genetic circuitry controlling ESNs specification downstream of Notch starts being elucidated (Chen et al., 2011; Joyce Tang et al., 2013), little is known about midline genes with the exception of *Msxb* that has been shown to be essential solely for dorsal midline formation (Imai et al., 2006). In order to better understand neurogenic midline fate acquisition we have analyzed the function and regulation of seven midline transcription factors.

2. Material and methods

2.1. Embryo obtention and manipulation

C. intestinalis type B were provided by the Centre de Ressources Biologiques Marines in Roscoff (EMBRC-France). Embryo obtention, injection and electroporation were performed as described in (Pasini et al., 2006). Staging was described according to (Hotta et al., 2007). Standard control-MO (5′-CCTCTTACCTCAGTTACAATTTATA 3′) and *Msxb*-MO (5′-GGATTCGTTTACTGTCATTTTTAAT-3′) were purchased from GeneTools LLC and were injected at 0.25 to 0.50 mM.

2.2. Gene model identifiers

The genes described in this study are represented by the following gene models in the KH2012 *C. intestinalis* assembly: *Msxb* (KH.C2.957), *Nkx-C* (KH.C1.922), *Klf1/2/4* (KH.C5.154), *Achaete-scute a-like2* (KH.L9.13), *Cagf9/tox* (KH.C3.330), *Orphan bHLH-1* (KH.C7.269), *Dll-C* (KH.C7.770), *Fog* (KH.C10.574), *SoxB2* (KH.S164.12), *ETR* (KH.C6.128) and *KH.C8.111* (KH.C8.111).

2.3. In situ hybridization

Whole mount *in situ* hybridization were performed as previously described (Bertrand et al., 2003). Dig-labeled probes were synthesized from the following clones obtained from *C. intestinalis* cDNA libraries (Gilchrist et al., 2015; Satou et al., 2002): *Msxb* (cien92596), *Nkx-C* (citb089b03), *Klf1/2/4* (ciem823l14), *Achaete-scute a-like2* (cien82323), *Cagf9/tox* (long isoform: cidg851f11, short isoform: cima838k22), *Orphan bHLH-1* (ciem814e10), *Dll-C* (cilv080b23), *ETR* (citb028e11) and *KH.C8.111* (citb002h08). Effects of overexpression were analyzed for each marker on 50 electroporated embryos for at least 2 independent electroporation experiments. Embryos electroporated with pFog > Venus were used as controls (Roure et al., 2007).

2.4. Generation of electroporation constructs

All the genes that we have analyzed are predicted to possess a single transcript encoding a unique protein, with the exception of *Cagf9/tox* where two transcripts are predicted each with a unique transcription start site and first exon. The two transcripts encode a short and a long isoform, the short isoform corresponding to the long isoform deleted of the first 225 amino acids. Overexpression of each isoform was analyzed. Since the short isoform had no effect or an effect similar to overexpression of the long isoform (Fig. S1), only results for the more active long isoform are presented.

Electroporation constructs for overexpression were generated using the Gateway technology (Roure et al., 2007) with the promoter of *Fog* driving expression throughout ectoderm from the 16-cell stage (Pasini et al., 2006; Rothbacher et al., 2007) or with the promoter of *SoxB2* driving expression throughout b-line ectoderm except the b6.5 lineage from the 64-cell stage (Fig. S2). A 1363 bp fragment upstream of *SoxB2* was amplified by PCR from genomic DNA using the primers ataaagtaggctACTAATCTGGTGACACCCACGTTC and caaaagttgggtTGGGGTCGTCGAGTAGATTATACAG to generate the pSoxB2_upstream > nlsLacZ construct. pENTRY clones containing the full coding sequence were obtained from a full-ORF gene collection (Gilchrist et al., 2015) for *Msxb* (cien92596), *Klf1/2/4* (ciem823l14), *Achaete-scute a-like2* (cien82323), *Cagf9/tox* (long isoform: cidg851f11, short isoform: cima838k22) and *Orphan bHLH-1* (ciem814e10). For Nkx-C, the pENTRY clone was generated by amplification of the coding sequence from the clone citb089b03 using the primers aaaaagcaggctcagaaaaaATGGCGCATGCTTGCAATG and agaaagctgggtTTAACGTCGTAAGGTAGTGGG followed by a BP reaction. No cDNA clone was found to contain the entire coding sequence of Dll-C. Two overlapping clones (cilv069c15 and cilv008b23) were used to reconstitute a full ORF using two consecutive PCR amplifications. The first PCRs were done using the clone cilv069c15 and the primers (B1-DllC 5′-ggggacaagtttgtacaaaaaagcaggctcagaaaaaATGAGCGCCTACGGTTACAAC-3′ and DllC-rev 5′-ACGTCACTTCCCATCACGTG-3′) and the clone cilv008b23 and the primers (B2-DllC 5′-ggggaccactttgtacaagaaagctgggtTTATTTTAAGTAACCCGAGTCCC-3′ and DllC-fw 5′-GGCATCTCTGCATCCGTGG-3′). The final PCR was done on the first PCR products using the primers B1-DllC and B2-DllC.

3. Results and discussion

3.1. Dynamic expression of tail epidermis midline transcription factors

By searching the *Ciona* expression databases (Ghost and Aniseed) (Satou et al., 2005; Tassy et al., 2010), we identified seven transcription factors (TF) expressed in tail dorsal and ventral midlines at tailbud stages: the homeobox genes *Msxb*, *Nkx-C* and *Dll-C*, the bHLH factors *Achaete-scute a-like2* and *Orphan bHLH-1*, the HMG encoding gene *Cagf9/tox* and the Zn finger factor *Klf1/2/4*. None of these genes was exclusively expressed in tail midlines. For example, several of them were expressed in trunk epidermis, in regions that will also give rise to ESNs like the palps. *Msxb* was also expressed in the pigment cell forming region of the CNS (Abitua et al., 2012). *Orphan bHLH-1* was expressed in the notochord. Here we will solely describe their expression in forming midlines from gastrula to late tailbud stages when ESNs are post-mitotic and about to differentiate (Figs. 1, 2A, S3, and S4).

3.1.1. Initiation of expression

Careful analysis of the onset of expression of these seven genes led us to distinguish three different behaviors. The first category includes a single gene *Msxb* whose expression in the dorsal midline precursors was initiated very early, at the 64-cell stage (Roure et al., 2014). This expression occurred with no interruption during gastrulation and neurulation (Figs. 1, and S3). Expression in the ventral midline was initiated posteriorly at late gastrula/neurula stages and extended rapidly anteriorly (Figs. 1, and 2B–D). The second category includes *Achaete-scute a-like2*, *Klf1/2/4* and *Cagf9/tox*. Their dynamic of gene activation was similar to *Msxb* but started later: dorsal expression was initiated throughout the dorsal midline at late neurula stages and preceded the posterior to anterior initiation of expression in the ventral midline (Fig. 1). The sequence of gene activation is as follows: *Achaete-scute a-like2* and *Klf1/2/4*, followed by *Cagf9/tox*. The third category includes *Nkx-C*, *Dll-C* and *Orphan bHLH-1* that were sequentially activated at the tip of the tail epidermis at late gastrula/neurula stages. The

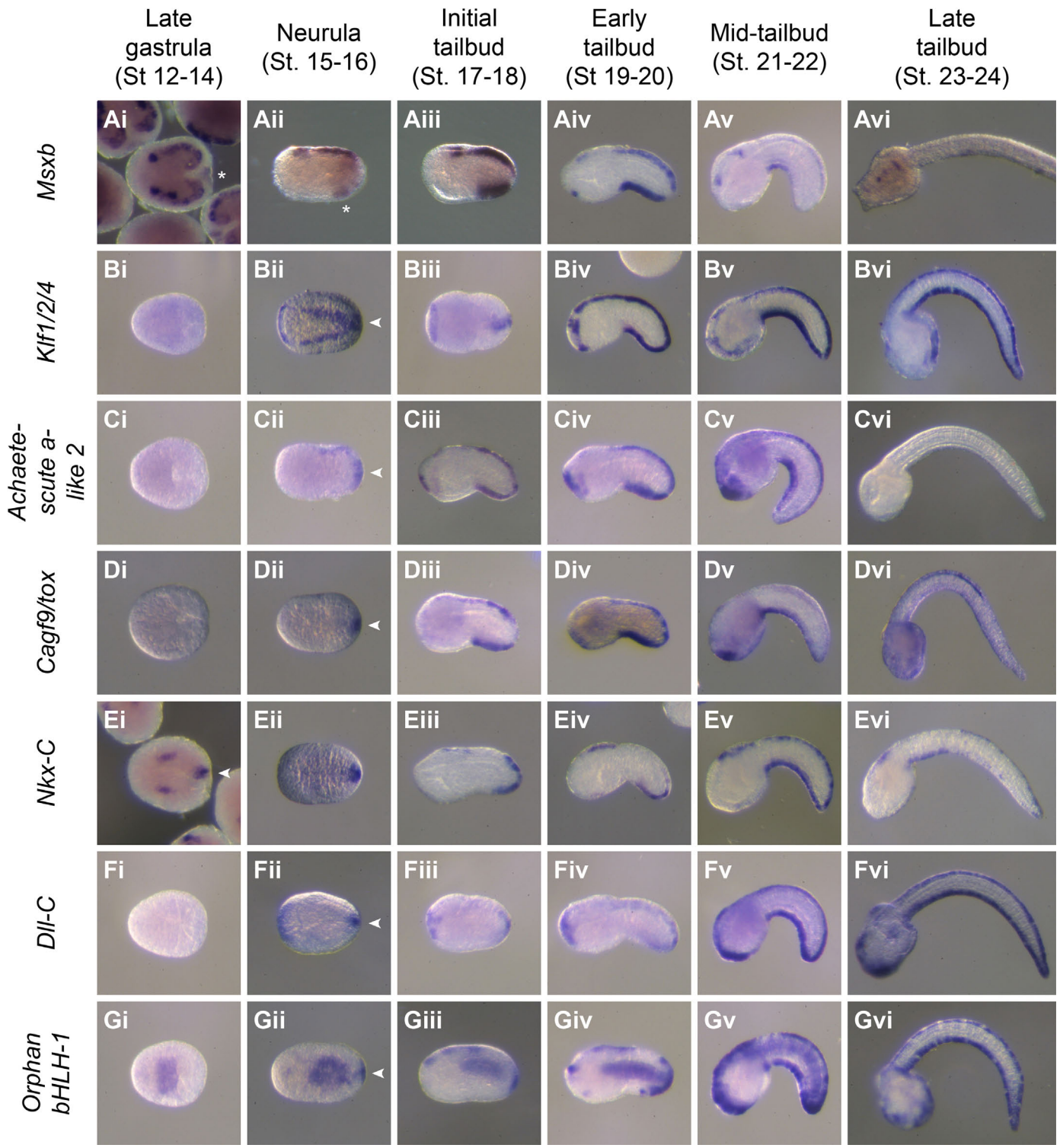


Fig. 1. Expression patterns of seven midline transcription factors. *In situ* hybridization for *Msxb* (A), *Klf1/2/4* (B), *Achaete-scute a like-2* (C), *Cagf9/tox* (D), *Nkx-C* (E), *Dll-C* (F) and *Orphan bHLH-1* (G) at late gastrula (i), neurula (ii), initial tailbud (iii), early tailbud (iv), mid-tailbud (v) and late tailbud (vi) stages. In addition to tail midline expression, all genes were expressed in other territories. *Msxb* was expressed in the CNS, dorsal nerve cord (Ai–ii) and various domains of the sensory vesicle (A), and around the palp forming region (Aiv). *Klf1/2/4* was expressed in specific domains of the trunk epidermis (Bii–vi). *Achaete-scute a-like 2* was expressed in the palp forming region (Ciii–v). *Cagf9/tox* was expressed in a portion of the tail nerve cord (Dvi). *Nkx-C* was expressed in the dorsal trunk epidermis midline (E). *Dll-C* was expressed in anterior epidermis including the palp forming region (Fii–vi). *Orphan bHLH-1* was expressed in anterior trunk ectoderm (Giii–vi) and the notochord (Gi–v). Embryos are oriented with anterior to the left. Lateral views and dorsal to the top for all, except for i, Bii and Gii that are dorsal views, and Dii, Eii, Fii and Biii that are ventral views. The asterisk indicates the first detectable expression of *Msxb* in the ventral midline. The arrowheads show the onset of expression at the tip of the tail.

expression in the midlines then followed a posterior to anterior wave of activation in both dorsal and ventral midlines at the early to mid-tailbud stages.

3.1.2. Extinction of gene expression

Three genes were downregulated in the midlines during the window of development that we have analyzed. Genes whose

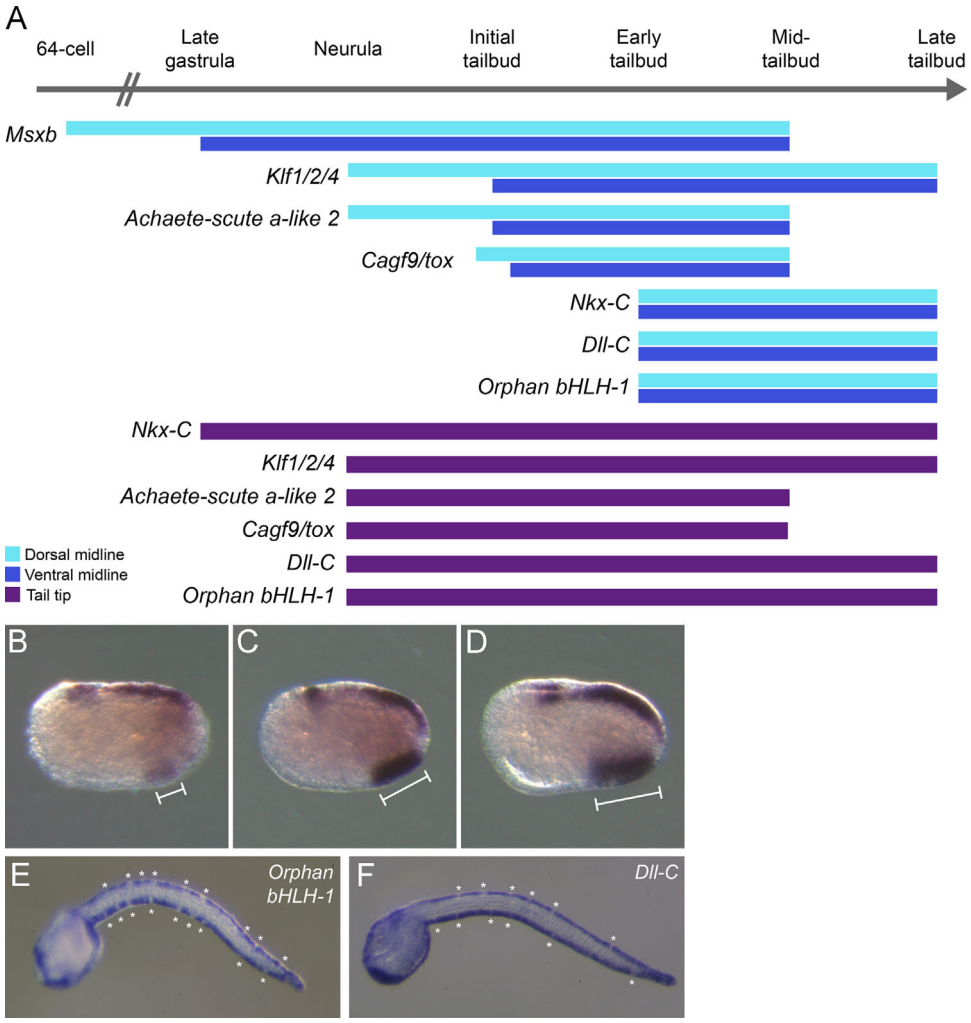


Fig. 2. Expression summary and highlights of some specific expression patterns. (A) Summary chart for the temporal expression of the seven transcription factors in the various midline compartments: dorsal midline (light blue), ventral midline (dark blue) and tail tip (purple). (B–D) At its onset in the ventral midline, *Msxb* was detected in an expanding domain following a posterior to anterior wave. Note that *Msxb* was not expressed at the tip of the tail. At late tailbud stages, *Orphan bHLH-1* (E) and *Dll-C* (F) expression was downregulated in irregularly spaced cells that are presumably ESNs (white asterisks). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

expression was initiated early were switched off first, in the following order: *Msxb*, *Achaete-scute a-like2* and *Cagf9/tox* (Figs. 1, and S3). While *Nkx-C* started to be downregulated at late tailbud stages, the last three genes were still robustly detected in ventral and dorsal midlines at late tailbud stages.

3.1.3. Spatial domains

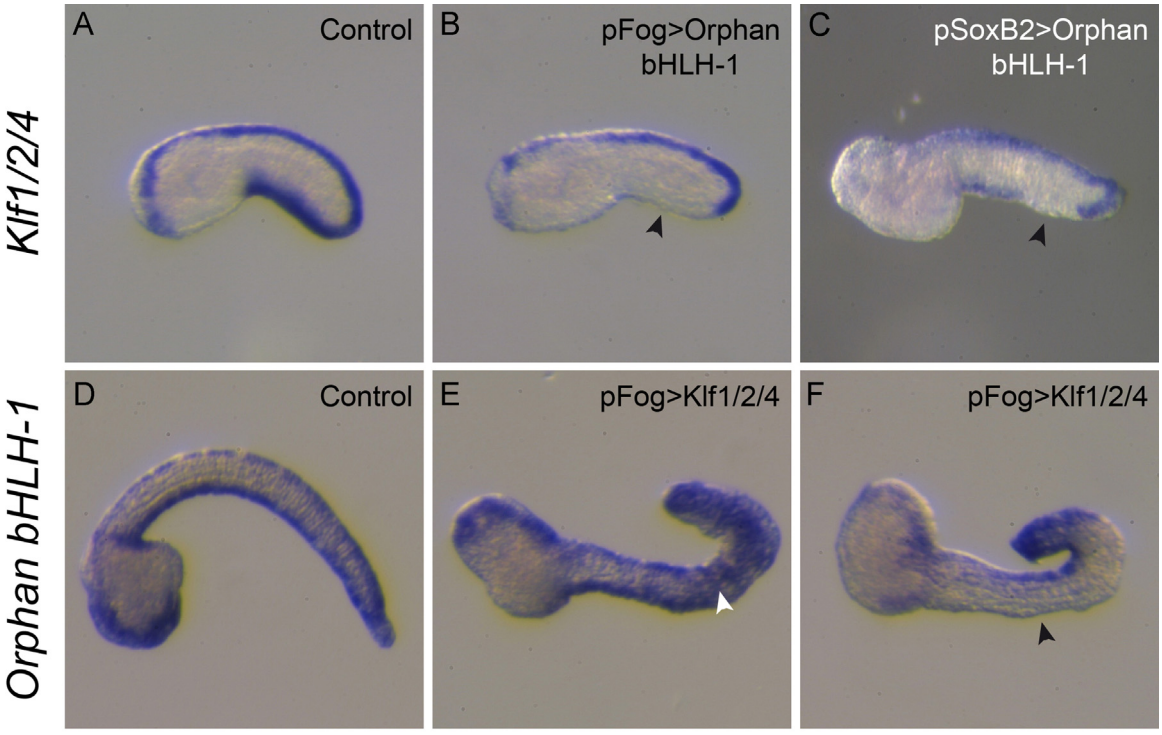
Msxb appeared to a have unique expression pattern since its expression was never detected in the tip of the tail epidermis while the other genes were detected there from neurula stages (Figs. 1, and 2).

Two genes that maintained strong expression at tailbud stages, *Dll-C* and *Orphan bHLH-1*, displayed a downregulation of expression in discrete cells within the midlines (Fig. 2E, and F). The random location of these non-expressing cells suggests that they correspond to the ESNs.

In summary (Fig. 2A), expression in the tail epidermis midlines of the seven transcription factors that we have analyzed has revealed unexpected complexity and high dynamics. We have observed sequential activation of the seven midline genes. While *Nkx-C* is the first gene expressed at the tip of the tail, *Msxb* is the first gene expressed in the midlines proper. These observations suggest that these two genes could initiate a regulatory cascade leading to midline fate acquisition and ESN formation.

3.2. Epistatic relationships among the seven midline transcription factors

To characterize regulatory interactions among these TFs and to determine a possible cascade of activation we have overexpressed each factor individually and analyzed by *in situ* hybridization the effect on the expression of the other six genes. We have used electroporation-mediated overexpression using two independent ectodermal drivers. The promoter of *Fog* has pan-ectodermal activity from the 16-cell stage (Pasini et al., 2006; Rothbacher et al., 2007). The promoter of *SoxB2* drives transcription in the posterior ectoderm lineage, b-line, with the exclusion of the b6.5 lineage (dorsal midline precursors) from the 64-cell stage (Fig. S2). These drivers allow assessing at the same time ectopic gene activation in the flanks of the tail epidermis and gene inhibition at the midlines. While overexpression often led to strong embryo malformations, we could determine gene expression changes. Interestingly, both drivers led to qualitatively identical gene expression modifications (examples shown in Figs. 3 and 4; data not shown,) but the pFog driver led usually to stronger phenotypes. However, the morphology of embryos overexpressing Dll-C using the pFog driver



G

Overexpressed transcription factor	Effect on gene expression: *Msxb*	*Nkx-C*	*Klf1/2/4*	*Achaete-scute a-like2*	*Cagf9/tox*	*Dll-C*	*Orphan bHLH-1*
Msxb							
Nkx-C							
Klf1/2/4							
Achaete-scute a-like2							
Cagf9/tox							
Dll-C							
Orphan bHLH-1							

Fig. 3. Effects of midline TF overexpression on the expression of midline TFs. (A–F) Representative examples of gene interactions analysis. Overexpressing Orphan bHLH-1 using the pFog driver (B) or the pSoxB2 driver (C) led to *Klf1/2/4* inhibition at the early tailbud stage in the ventral midline compared to control (A). Overexpression of Klf1/2/4 using the pFog driver led to activation (E) or inhibition (F) of *Orphan bHLH-1* at the late tailbud stage compared to control (D). Embryos are in lateral views with anterior to the left and dorsal to the top. (G) Summary table of gene interactions following midline transcription factor overexpression. Color code: green=activation, red=inhibition, gray=no effect, and black=not tested. In four cases a dual activity, as depicted in E and F, is represented by the presence of green and red. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

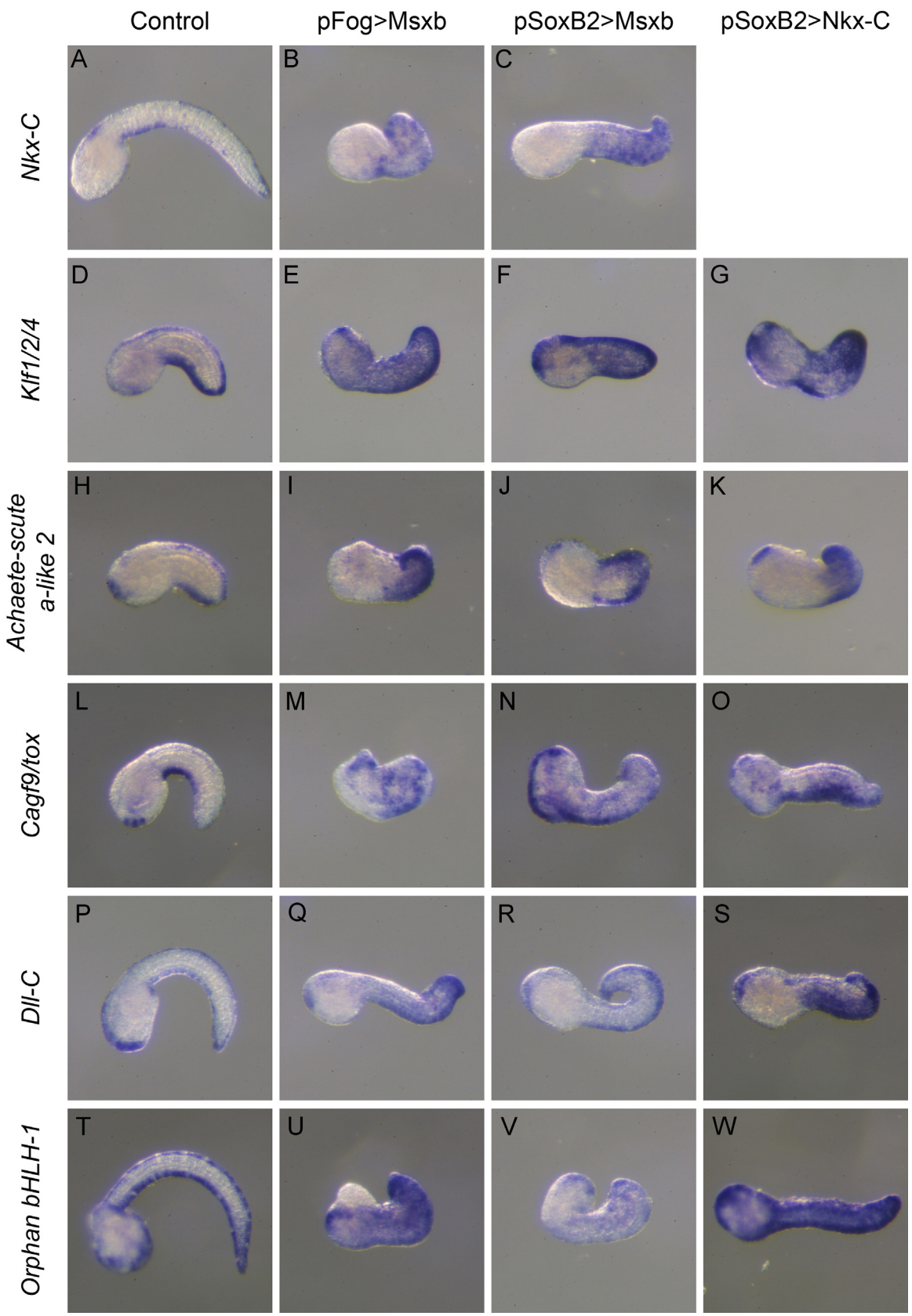


Fig. 4. Msxb and Nkx-C can activate the other midline transcription factors. Overexpression of Msxb with either pFog (B, E, I, M, Q and U) or pSoxB2 (C, F, J, N, R, and V) can activate the other six genes in the lateral epidermis. Similarly, overexpressing Nkx-C using the pSoxB2 driver ectopically activated *Klf1/2/4* (G), *Achaete-scute a like-2* (K), *Cagf9/tox* (O), *Dll-C* (S) and *Orphan bHLH-1* (W). *In situ* hybridization for *Nkx-C* (A–C), *Klf1/2/4* (D-G), *Achaete-scute a like-2* (H–K), *Cagf9/tox* (L-O), *Dll-C* (P–S) and *Orphan bHLH-1* (T–W). Embryos are in lateral views with anterior to the left and dorsal to the top.

was dramatically altered and we were unable to interpret gene expression pattern modifications in these embryos. The overexpression results for Dll-C were thus only obtained using the pSoxB2 driver.

The gene interactions are summarized in Fig. 3G and representative examples are shown in Fig. S5. For each of the 42 pairs of overexpressed TF/detected TF expression, we could determine either an absence of effect (8 cases), an activation (13 cases), an inhibition (17 cases) or a mixed effect (activation on the flanks and inhibition at the midlines) (4 cases). This latter observation was rather unexpected (an example is shown in Fig. 3E and F) and could correspond to a combinatorial effect (at the midlines, due to the presence of other midline TFs, a given TF would have a different effect than on the flanks). This observation could also result from incoherent feedforward motifs. For example, we observed both activation and inhibition of *Achaete-scute a-like 2* upon Klf1/2/4 overexpression. This could be the result of the following interactions: Klf1/2/4 activates *Achaete-scute a-like 2* but also *Dll-C* that in turn inhibits *Achaete-scute a-like 2* (incoherent feedforward motif visible in Fig. S6). We further considered the fact that a given TF would have the potential to both activate and inhibit a target gene. In the table of Fig. 3G, we have organized our results according to the number and types (activation *vs* inhibition) of potential interaction received (inputs) or sent (outputs). Interestingly, the resulting potential functional hierarchy was very much in agreement with the temporal hierarchy that we have described above.

3.2.1. *Msxb and Nkx-C are potent activators of midline TFs*

We first analyzed the positive regulatory interactions that could support a sequential cascade of activation suggested by TF gene expression during development. Interestingly, *Msxb*, the first gene expressed in the midlines, was able to activate the other 6 TFs (Fig. 4). A similar situation was observed for the earliest gene expressed at the tip of the tail, *Nkx-C*. It could activate all TFs with the exception of *Msxb*. By contrast, the number of activations was much more reduced for downstream factors (three for Klf1/2/4 and zero or one for the other TFs). Consequently, we are not in the presence of a simple linear network of gene as it was observed for proneural genes downstream of Notch (Joyce Tang et al., 2013). If it were the case, we would expect that the top right part of the Table 3G would be filled with activations (green). We could obtained such linear cascade if two genes, *Achaete-scute a-like2* and *Cagf9/tox*, were removed from the analysis (Fig. S7). Indeed, these TF have different behaviors: Achaete-scute a-like2 is able to regulate only 3/6 TFs while Cagf9/tox acts mainly as an inhibitor.

We next scored the number of positive inputs for each TF. *Msxb* cannot be activated by any TF and *Nkx-C* can be activated only by Msxb. This observation suggests that these two early genes are initially activated by factors that are absent from the network that we have analyzed. Intermediate genes, *Klf1/2/4*, *Achaete-scute a-like2* and *Cagf9/tox*, can be activated by 2 or 3 TFs, while the latest genes, *Dll-C* and *Orphan bHLH-1* can be activated by 4 and 5 TFs respectively.

In conclusion, scoring the positive outputs and inputs for each TF allowed us to draw three hierarchical classes of genes whose position in the hierarchy actually corresponds to their onset of expression during normal development. The first class includes the upstream factors *Msxb* and *Nkx-C*, the second class includes the intermediate TFs *Klf1/2/4*, *Achaete-scute a-like2* and *Cagf9/tox*, and the third class includes the latest expressed genes *Dll-C* and *Orphan bHLH-1*.

3.2.2. *Extensive negative feedbacks*

Our analysis has revealed an extensive number of inhibitions. The most striking case is *Msxb* that is inhibited by all other TFs. This is an agreement with negative feedback loops of downstream targets towards the upstream activator and with the early extinction of *Msxb* expression observed during normal development, suggesting that *Msxb* expression is actively turned off by downstream components. The same observation applies to the other TFs according to the hierarchy that we have described above with two exceptions. The highly ranked factor *Nkx-C* is only inhibited by two TFs. This might correlate with the fact that this gene is expressed during later stages compared to *Msxb*. *Cagf9/tox* is only inhibited by Dll-C, yet we have no obvious interpretation for this observation. When these two genes were excluded from the analysis, we could nicely observe a reciprocal distribution of positive activations from upstream factors and of negative feedbacks loops from downstream factors (Fig. S6).

3.3. *Gene regulatory network reconstruction*

We then aimed at integrating expression patterns and epistatic relationships of the seven TFs that we have analyzed. Based solely on overexpression data, we could propose a sequential activation of *Msxb* and *Nkx-C* that are both potent activators of midline TFs. However, these two genes have globally complementary expression patterns and *Msxb* cannot activate *Nkx-C* expression at the tail tip since is not expressed there. Moreover, at the time *Nkx-C* gets expressed at the midlines (early to mid-tailbud stages), *Klf1/2/4*, *Achaete-scute a-like2* and *Cagf9/tox* are already expressed. While *Nkx-C* could account for the activation of *Dll-C* and *Orphan bHLH-1* at the midlines, we propose that Nkx-C function is to maintain midline TF expression when *Msxb* is no longer expressed. We consequently considered two territories (the tail tip and the midlines) on one hand, and two events (initiation and maintenance) on the other hand. The tentative networks of interactions that we have built are shown in Fig. 5. Our overexpression approach cannot distinguish between direct and indirect interactions. We reasoned that activation of the most downstream genes could result from two sequential activations. We built the provisional gene regulatory networks (GRNs) by taking into account the sequence of gene expression and by reducing the number of interactions. The gene expression initiation network (Fig. 5A) is controlled by *Msxb* and does not need the action of *Nkx-C*. This network is characterized by a high density of negative feedback loops. A similar GRN can explain gene expression initiation at the tip of the tail and gene expression maintenance at the midlines (Fig. 5B). This network is identical to the previous one with *Nkx-C* replacing *Msxb*. Finally, we could draw a simplified network describing gene expression maintenance at late tailbud stages (Fig. 5C).

3.4. *Transcriptional control of midline and ESN fate acquisition*

We next determined the effect of overexpressing each midline TF on the final fate of midline cells. As presented above, midlines are single rows of epidermal cells interspersed with ESNs at irregular intervals. ESNs can be identified using the neuronal marker *Etr* (Joyce Tang et al., 2013; Pasini et al., 2006). Very little is known about epidermal cell differentiation, and for example it is not known whether the epidermal cells of the midlines differentiate differently from the other epidermal cells that cover the embryo. Given the lack of a specific midline epidermal cell marker, we selected a gene specifically expressed in the midlines but that would not be a TF. *KH.C8.111* is an excellent midline marker and it encodes a protein of 182 amino acids with no homology.

As expected from the key positions of *Msxb* and *Nkx-C* in the midline networks, both could activate expression of both *Etr* and *KH.C8.111* in the lateral epidermis of the tail (Fig. 6, and S4). Interestingly, *KH.C8.111* ectopic expression was detected as a patch while *Etr* as spots suggesting that Msxb and Nkx-C are sufficient to transform lateral epidermis into neurogenic midline territory

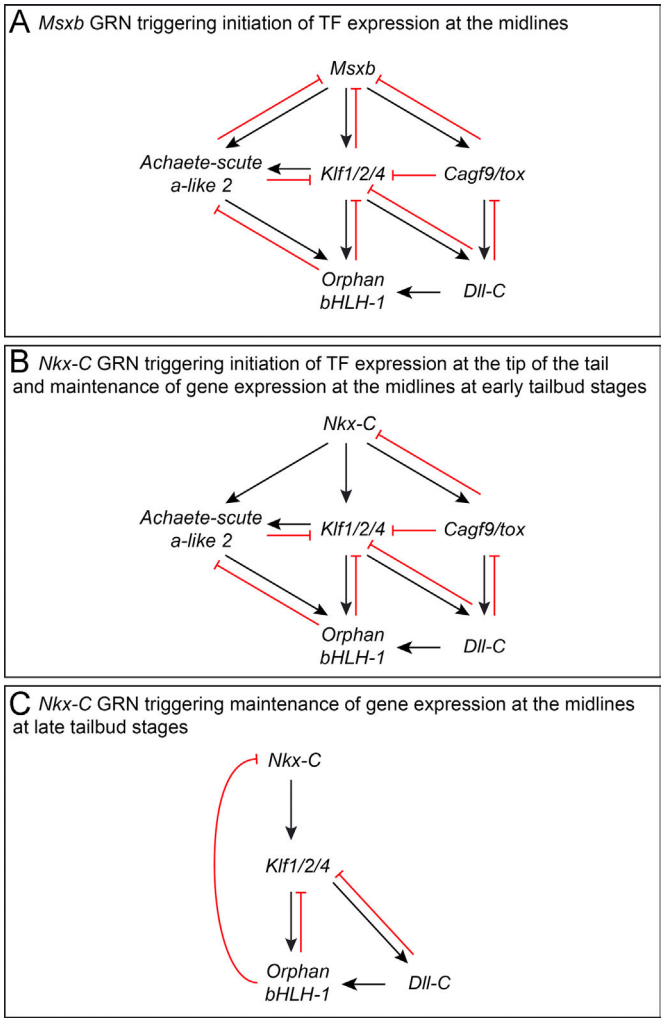


Fig. 5. Tentative gene regulatory networks for midline transcription factors. The networks were built by combining expression data and gain-of-function results. (A) *Msxb* is the highest rank factor initiating midline gene expression at neurula/initial tailbud stages. Note that most activations are coupled with a reciprocal negative feedback. (B) The same network, but with *Nkx-C* instead of *Msxb* at the top, can explain initiation of gene expression at the tip of the tail and maintenance of gene expression in midlines. (C) A simplified and rather linear network may represent the interactions between midline TFs at late tailbud stages.

where the Notch pathway is active. Surprising, none of the 5 downstream TFs were able to do so. Only Klf1/2/4 was capable of activating *KH.C8.111*. These data suggest that a combination of downstream midline TFs is required to activate midline fate. Surprisingly, overexpression of several of the downstream TFs inhibited expression of both markers. It is possible that these observations are not relevant to the normal function of the TFs and result from an artificial prolonged presence of the TFs after overexpression. For example, *Achaete-scute a-like 2* which appears to act upstream of the proneural network for ESN differentiation (Joyce Tang et al., 2013) is turned off at mid-tailbud stages (Fig. 1). We would like to propose that the expression of *Achaete-scute a-like 2* is incompatible with ESN formation and that its extinction is required for proper ESN differentiation. Interestingly, the incoherent feedforward motif involving *Klf1/2/4*, *Achaete-scute a-like 2* and *Dll-C* (Fig. S6) could explain how *Achaete-scute a-like 2* is turned off at mid-tailbud stages. In the case of *Dll-C* and *Orphan bHLH-1*, we would like to propose a functionally relevant interpretation of their overexpression phenotype. We have shown above that the expression of these two genes was downregulated in ESNs at late tailbud stages. The result of their overexpression suggests that their continuous expression is incompatible with ESN formation. Interestingly, Chen and colleagues have identified *Dll-C* and *Orphan bHLH-1* as predicted targets of miR-124, an essential microRNA downstream of Notch for ESN specification, their 3′ UTR harboring several canonical miR-124 binding sites (Chen et al., 2011). They have also provided *in vivo* evidence for *Orphan bHLH-1* downregulation.

3.5. *Msxb is essential for midline formation*

Msxb was the first gene to be expressed in both dorsal and ventral midlines. Moreover, its overexpression was sufficient to induce the formation a neurogenic territory expressing six midline TFs where ESNs formed. Consequently, *Msxb* is likely the gene at the top of the hierarchy of the GRN defining midlines (Fig. 5A). However, a previous report showed that its loss-of-function using antisense morpholino (MO) injection only affected dorsal midline formation (Imai et al., 2006). We reinvestigated *Msxb* function using the same approach. We found that injection of a MO against *Msxb* abolished both dorsal and ventral expression of *Klf1/2/4* in the midlines (Fig. 7A, and B). While the MO used here was slightly different from the one used by Imai and colleagues (it targets a region with a shift of 2 bp in 3′), we believe that the discrepancy between the two phenotypes results mainly from a strength of phenotype. We observed in most cases a loss of dorsal midline expression (94%, $n=17$). The effect on the ventral midline was more variable: complete loss (65%), small posterior domain (24%) and normal expression (12%). Importantly, expression of *Klf1/2/4* at the tip of the tail was always present in agreement with our proposal of a *Msxb*-independent specific GRN for this part of the tail epidermis (Fig. 5B).

3.6. *Signaling pathways controlling midline GRN*

The existence of a specific GRN for the tip of the tail prompted us to test whether this territory could be specified independently of the dorsal and ventral midline inducers, Fgf9/16/20 and Admp. Since abolition of Fgf signaling has a dramatic impact on gastrulation and neurulation making gene expression analysis very difficult, we abolished Nodal signaling that acts downstream of Fgf9/16/20 by overexpressing the secreted inhibitor Lefty (Pasini et al., 2006). As expected dorsal expression of *Msxb* at initial tailbud stages was abolished dorsally except in the posterior-most region (1 or 2 cells on each side), a territory specified independently of Nodal (Pasini et al., 2006) (83% of embryos, $n=81$). *Nkx-C* expression at this stage had weak and variable anterior expression in the dorsal midline. Following Lefty overexpression, only 7% of the embryos ($n=72$) had such an expression, the other 93% had only a posterior dorsal expression, in the Nodal independent territory. Abolition of Bmp signaling by overexpression of the secreted inhibitor Noggin inhibited ventral expression of both *Msxb* and *Nkx-C* (*Msxb*: 79% of embryos with no ventral expression, $n=78$; *Nkx-C*: 90% of embryos with no ventral expression, $n=78$). When both Nodal and Bmp signalings were inhibited, *Mxb* and *Nkx-C* expression was only detected in the dorsal posterior-most midline (*Msxb*: 69%, $n=81$; *Nkx-C*: 77%, $n=68$). These results demonstrate that the expression of the tail tip marker *Nkxc-C* is regulated by the dorsal and ventral inducers and suggest that the tail tip, despite having a specific GRN, is also dependent on early midline inducers.

We finally would like to discuss the primary inputs that launch the midline and tail tip GRNs (Fig. 7K). Our analysis has revealed that these inputs should regulate the initiation of *Msxb* and *Nkx-C* expression respectively. We have recently shown that the earliest activation, at the 64-cell stage, of *Msxb* in the dorsal midline precursors is controlled, downstream of Fgf9/16/20, by the

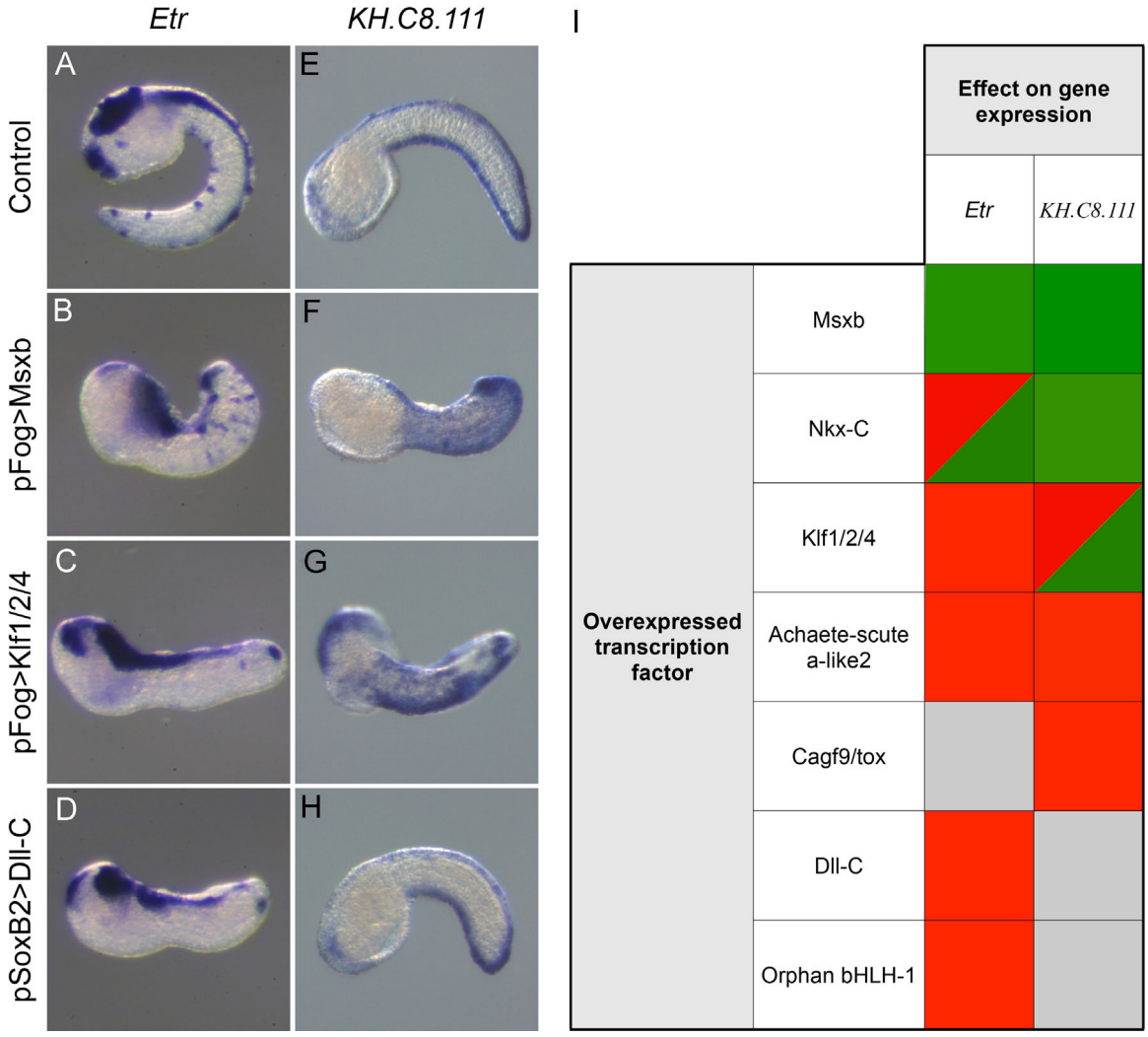


Fig. 6. Effects of midline TF overexpression on the expression of midline markers. (A–H) Representative examples of gene interactions analysis. Overexpressing Msxb using the pFog driver led to *Etr* (B) and *KH.C8.111* (F) ectopic expression in lateral epidermis at late tailbud stages compared to controls (A, E). Overexpression of Klf1/2/4 using the pFog driver led to inhibition of *Etr* (C) and ectopic expression of *KH.C8.111* (G). Dll-C overexpression inhibited *Etr* (D) but did not change *KH.C8.111* expression (H). *In situ* hybridizations for *Etr* (A–D) and *KH.C8.111* (E–H). Embryos are in lateral views with anterior to the left and dorsal to the top. (I) Summary table of midline marker expression following midline transcription factor overexpression. Color code: green=activation, red=inhibition, gray=no effect, black=not tested. In two cases, a dual activity is represented by the presence of green and red. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

combined action of the TF Otx and the signaling molecule Nodal (Roure et al., 2014). These precursors give rise not only to the dorsal midline but also to the dorsal nerve cord. Since *Msxb* expression disappears from this latter tissue during neurulation, further work will be needed to determine whether additional factors are required to maintain *Msxb* expression in dorsal midline precursors. *Msxb* expression in ventral midline was robustly detected at neurula stages, but it might even start earlier, at late gastrula stages (Fig. 1). While ventral *Msxb* expression might be directly activated by the ventral midline inducer Admp, we would rather favor a model involving intermediate factors such as the TFs *Tbx2/3*, *Nk4* and *Nkx-A* that are expressed ventrally (Imai et al., 2004). These three genes are expressed during gastrula and neurula stages in the ventral midline with a posterior to anterior wave similar to early *Msxb* expression, and their expression is regulated by the Bmp pathway (data not shown).

Nkx-C had a biphasic expression: onset at the tip of the tail at late gastrula stages and posterior to anterior extension in both midlines at the early to mid-tailbud stage transition. Its unique pattern of expression was strikingly reminiscent of previous reports (Ikuta et al., 2010; Imai et al., 2004; Pasini et al., 2012). The tail tip is characterized by the expression of a TF *Hox12* and two signaling molecules, *Wnt5* and *Fgf8/17/18*, that are involved in tail elongation and cell shape control (Ikuta et al., 2010). Moreover, these two signaling pathways appear to control the number and position of ESNs along the midlines (Pasini et al., 2012). It is tempting to speculate that Wnt and Fgf signaling pathways control *Nkx-C* onset and posterior to anterior wave of expression. Since *Nkx-C* expression at the tip of the tail is also regulated by Bmp and Nodal signalings, we would like to propose that *Nkx-C* is activated in the region where active Wnt and/or Fgf signaling and midline identity overlap. *Nkx-C* would thus presumably control ESN differentiation that proceeds from posterior to anterior as suggested by the dynamic expression of *Delta2* and *Etr* (Pasini et al., 2012, 2006).

4. Conclusions

The combination of precise expression pattern analyses and systematic gain-of-functions has allowed the building of provisional gene circuits for tail midline epidermis and PNS formation and the identification of *Msxb* and *Nkx-C* as upstream nodes. Downstream of these two factors, our analyses have revealed a

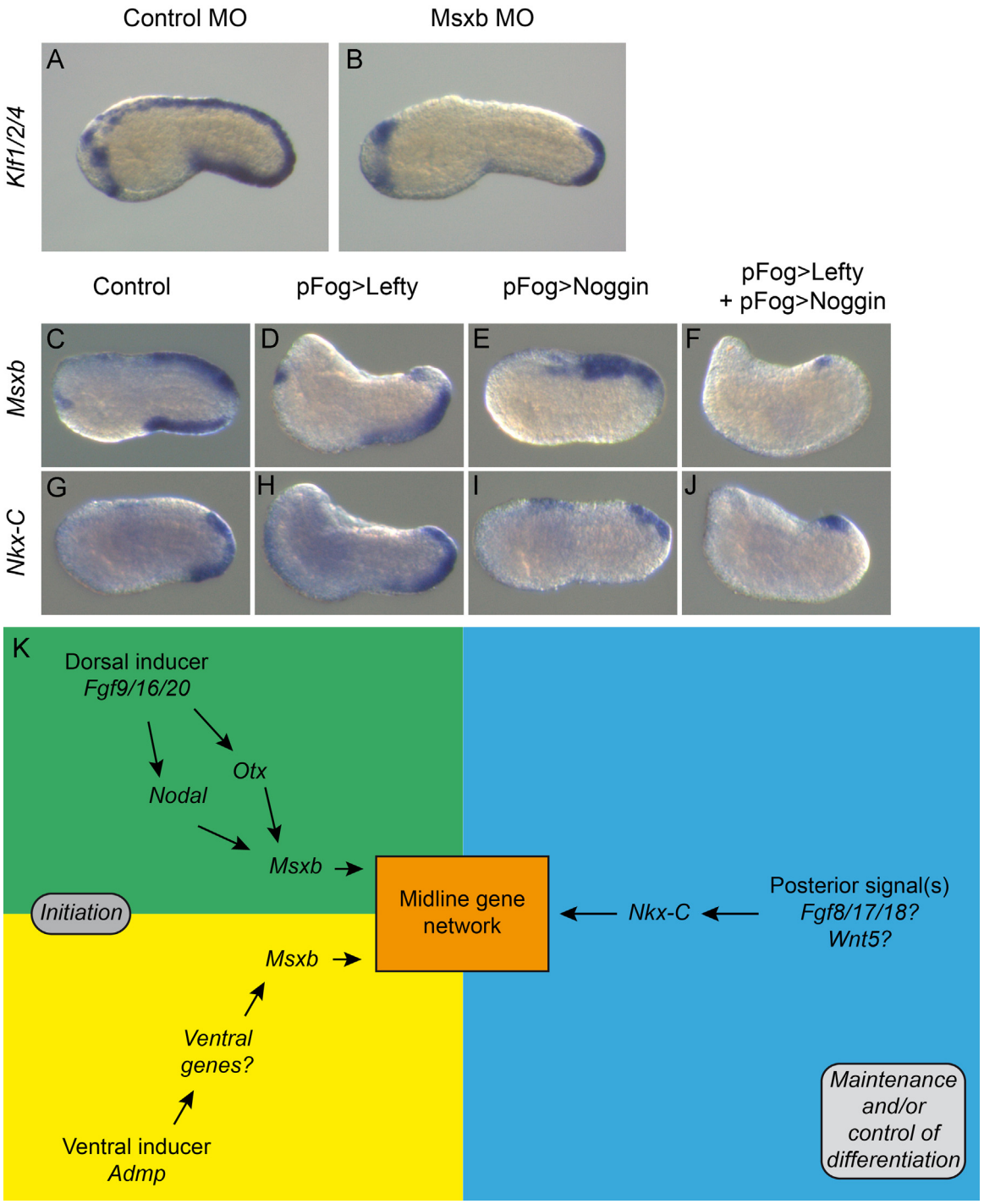


Fig. 7. *Msxb* is required for midline specification, *Nkx-C* regulation and summary model. Knockdown of *Msxb* using MO injection led to abolition of *Klf1/2/4* expression in both ventral and dorsal midlines (B) while injection of a control MO did not change *Klf1/2/4* expression (A). (C–J) Expression of *Msxb* (C–F) and *Nkx-C* (G–J) following Lefty (D, H), Noggin (E, I) and the combined Lefty and Noggin (F, J) overexpression. Embryos are in lateral views with anterior to the left and dorsal to the top. (K) Summary model proposing how different primary inputs control distinct biological functions through the regulation of the expression of *Msxb* and *Nkx-C*, the two top-ranked TF of the midline network.

higher degree of complexity with extensive levels of negative interactions. In particular, none of the downstream factor was able to trigger neurogenic midline formation and ESN formation. These results might be the consequences of our over/ectopic expression approach, and a better control of the timing of ectopic expression, including a precise extinction of the transgene, may provide more relevant or coherent results. However, the results may simply reveal the complexity of the midline GRN where downstream TFs may control midline fate and ESN formation in a combinatorial manner. Additionally, it is possible that other midline TFs are missing in our analysis and previous expression screens (Imai et al., 2004). Further validation and edition of the GRNs we have proposed will require gene loss-of-function and *cis*-regulatory DNA analyses.

Funding

The authors are CNRS staff. The work was supported by CNRS,

UPMC and a grant from the Agence Nationale de la Recherche (ANR-11-JSV2-007-01). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceived and designed the experiments: AR SD. Performed most of the experiments: AR. Analyzed the data: AR SD. Wrote the paper: SD.

Acknowledgments

We are grateful to L. Kodjabachian and P. Lemaire for support and hosting us in their laboratories at an early stage of the project, and to J.M. Donnay and N. Trouillard for ascidian care. We thank Y. Satou, N. Satoh, Y. Kohara, and U. Rothbächer for sharing plasmids. We thank the CRBM (Station biologique de Roscoff) for collecting ascidians.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ydbio.2015.11.009.

References

Abitua, P.B., Wagner, E., Navarrete, I.A., Levine, M., 2012. Identification of a rudimentary neural crest in a non-vertebrate chordate. Nature 492, 104–107. http://dx.doi.org/10.1038/nature11589.

Abitua, P.B., Gainous, T.B., Kaczmarczyk, A.N., Winchell, C.J., Hudson, C., Kamata, K., Nakagawa, M., Tsuda, M., Kusakabe, T.G., Levine, M., 2015. The pre-vertebrate origins of neurogenic placodes. Nature 524, 462–465. http://dx.doi.org/10.1038/nature14657.

Akanuma, T., Hori, S., Darras, S., Nishida, H., 2002. Notch signaling is involved in nervous system formation in ascidian embryos. Dev. Genes Evol. 212, 459–472.

Baker, C.V., Bronner-Fraser, M., 1997. The origins of the neural crest. Part I: embryonic induction. Mech. Dev. 69, 3–11.

Baker, C.V., Bronner-Fraser, M., 2001. Vertebrate cranial placodes I. Embryonic induction. Dev. Biol. 232, 1–61.

Bertrand, V., Hudson, C., Caillol, D., Popovici, C., Lemaire, P., 2003. Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. Cell 115, 615–627.

Chen, J.S., Pedro, M.S., Zeller, R.W., 2011. miR-124 function during *Ciona intestinalis* neuronal development includes extensive interaction with the Notch signaling pathway. Development 138, 4943–4953. http://dx.doi.org/10.1242/dev.068049.

Gilchrist, M.J., Sobral, D., Khoueiry, P., Daian, F., Laporte, B., Patrushev, I., Matsumoto, J., Dewar, K., Hastings, K.E.M., Satou, Y., Lemaire, P., Rothbächer, U., 2015. A pipeline for the systematic identification of non-redundant full-ORF cDNAs for polymorphic and evolutionary divergent genomes: application to the ascidian *Ciona intestinalis*. Dev. Biol. 404, 149–163. http://dx.doi.org/10.1016/j.ydbio.2015.05.014.

Glenn Northcutt, R., 2005. The new head hypothesis revisited. J. Exp. Zool. B Mol. Dev. Evol. 304, 274–297.

Hotta, K., Mitsuhara, K., Takahashi, H., Inaba, K., Oka, K., Gojobori, T., Ikeo, K., 2007. A web-based interactive developmental table for the ascidian *Ciona intestinalis*, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. Dev. Dyn. 236, 1790–1805.

Ikuta, T., Satoh, N., Saiga, H., 2010. Limited functions of Hox genes in the larval development of the ascidian *Ciona intestinalis*. Development 137, 1505–1513. http://dx.doi.org/10.1242/dev.046938.

Imai, K.S., Levine, M., Satoh, N., Satou, Y., 2006. Regulatory blueprint for a chordate embryo. Science 312, 1183–1187.

Imai, K.S., Hino, K., Yagi, K., Satoh, N., Satou, Y., 2004. Gene expression profiles of transcription factors and signaling molecules in the ascidian embryo: towards a comprehensive understanding of gene networks. Development 131, 4047–4058.

Joyce Tang, W., Chen, J.S., Zeller, R.W., 2013. Transcriptional regulation of the peripheral nervous system in *Ciona intestinalis*. Dev. Biol. 378, 183–193. http://dx.doi.org/10.1016/j.ydbio.2013.03.016.

Lu, T.-M., Luo, Y.-J., Yu, J.-K., 2012. BMP and Delta/Notch signaling control the development of amphioxus epidermal sensory neurons: insights into the evolution of the peripheral sensory system. Development 139, 2020–2030. http://dx.doi.org/10.1242/dev.073833.

Pasini, A., Manenti, R., Rothbacher, U., Lemaire, P., 2012. Antagonizing retinoic acid and FGF/MAPK pathways control posterior body patterning in the invertebrate chordate *Ciona intestinalis*. PLoS One 7, e46193. http://dx.doi.org/10.1371/journal.pone.0046193.

Pasini, A., Amiel, A., Rothbacher, U., Roure, A., Lemaire, P., Darras, S., 2006. Formation of the ascidian epidermal sensory neurons: insights into the origin of the chordate peripheral nervous system. PLoS Biol. 4, e225.

Rothbacher, U., Bertrand, V., Lamy, C., Lemaire, P., 2007. A combinatorial code of maternal GATA, Ets and beta-catenin-TCF transcription factors specifies and patterns the early ascidian ectoderm. Development 134, 4023–4032.

Roure, A., Lemaire, P., Darras, S., 2014. An Otx/Nodal regulatory signature for posterior neural development in ascidians. PLoS Genet. 10, e1004548. http://dx.doi.org/10.1371/journal.pgen.1004548.

Roure, A., Rothbacher, U., Robin, F., Kalmar, E., Ferone, G., Lamy, C., Missero, C., Mueller, F., Lemaire, P., 2007. A multicassette Gateway vector set for high throughput and comparative analyses in ciona and vertebrate embryos. PLoS One 2, e916.

Satou, Y., Kawashima, T., Shoguchi, E., Nakayama, A., Satoh, N., 2005. An integrated database of the ascidian, *Ciona intestinalis*: towards functional genomics. Zool. Sci. 22, 837–843.

Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasakura, Y., Nakayama, A., Ishikawa, H., Inaba, K., Satoh, N., 2002. A cDNA resource from the basal chordate *Ciona intestinalis*. Genesis 33, 153–154.

Simoes-Costa, M., Bronner, M.E., 2013. Insights into neural crest development and evolution from genomic analysis. Genome Res. 23, 1069–1080. http://dx.doi.org/10.1101/gr.157586.113.

Tassy, O., Dauga, D., Daian, F., Sobral, D., Robin, F., Khoueiry, P., Salgado, D., Fox, V., Caillol, D., Schiappa, R., Laporte, B., Rios, A., Luxardi, G., Kusakabe, T., Joly, J.S., Darras, S., Christiaen, L., Contensin, M., Auger, H., Lamy, C., Hudson, C., Rothbacher, U., Gilchrist, M.J., Makabe, K.W., Hotta, K., Fujiwara, S., Satoh, N., Satou, Y., Lemaire, P., 2010. The ANISEED database: digital representation, formalization, and elucidation of a chordate developmental program. Genome Res. 20, 1459–1468. http://dx.doi.org/10.1101/gr.108175.110.