Genetic regulation of amphioxus somitogenesis informs the evolution of the vertebrate head mesoderm

Daniel Aldea[1], Lucie Subirana[1], Celine Keime[2], Lydvina Meister[1], Ignacio Maeso[4], Sylvain Marcellini[3], Jose Luis Gomez-Skarmeta[4], Stephanie Bertrand[1*], Hector Escriva[1*]

1 Sorbonne Université, CNRS, Biologie Intégrative des Organismes Marins, BIOM, Observatoire Océanologique, F-66650, Banyuls/Mer, France

2 IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), INSERM, U1258, CNRS, UMR7104, Université de Strasbourg, 67404 Illkirch, France.

3 Laboratory of Development and Evolution, Department of Cell Biology, Faculty of Biological Sciences, University of Concepcion, Concepción, Chile.

4 Centro Andaluz de Biología del Desarrollo (CABD), Consejo Superior de Investigaciones Científicas-Universidad Pablo de Olavide-Junta de Andalucía, Seville, Spain

* To whom correspondence should be addressed at: Laboratoire Arago, Avenue Pierre Fabre, 66650, Banyuls-sur-Mer, France. E-mail: stephanie.bertrand@obs-banyuls.fr, hescriva@obs-banyuls.fr. Phone: +33(0)468887390, Fax: +33(0)468887393

Abstract

The evolution of vertebrates from an ancestral chordate was accompanied by the acquisition of a predatory lifestyle closely associated to the origin of a novel anterior structure, the highly specialized head. While the vertebrate head mesoderm is unsegmented, the paraxial mesoderm of the earliest divergent chordate clade, the cephalochordates (amphioxus), is fully segmented in somites. We have previously shown that FGF signalling controls the formation of the most anterior somites in amphioxus and, therefore, unravelling the FGF-signalling downstream effectors is of crucial importance to shed light on the evolutionary origin of vertebrate head muscles. Here we show, by using a comparative RNA-seq approach and genetic functional analyses, that several transcription factors, such as *Six1/2*, *Pax3/7*, and *Zic*, act in combination to ensure the formation of three different somite populations. Interestingly, these proteins are orthologous to key regulators of trunk, and not head, muscle formation in vertebrates. Contrary to prevailing thinking, our results suggest that the vertebrate head mesoderm is of visceral and not paraxial origin and support a multi-step evolutionary scenario for the appearance of the unsegmented mesoderm of the vertebrates “new head”.

Introduction

Body segmentation is a morphological feature shared by several metazoan groups. In vertebrate embryos, the paraxial mesoderm is segmented and forms the so-called somites on both sides of the midline. However, this segmentation does not concern the whole body since the vertebrate mesoderm is not segmented anterior to the otic vesicle. Although the presence of a pseudo-segmentation of the head mesoderm was previously proposed [1-3], molecular and histological data on several vertebrate species do not support such a hypothesis [4-6], or at least do not support serial homology between head mesoderm and somitic mesoderm (see [7] for a review).

Vertebrate somites are formed in an antero-posterior progression by epithelialisation of the presomitic mesoderm, a process known as somitogenesis. This segmentation process is under the control of molecular signals (Retinoic Acid, FGF, Wnt and Notch pathways) that act through a clock and wavefront system [8,9]. Once somites are formed, they receive signals coming from the surrounding structures and get divided into a sclerotome region, that will give rise to the axial skeleton, and into a dermomyotome region that will further form muscles and dermis [10]. By contrast, the head mesoderm, which is in place at the mid-gastrula stage, does not form somites. Rather, it subdivides into the prechordal mesoderm in the most anterior axial region, and, laterally, into the cranial paraxial mesoderm and splanchnic mesoderm whose morphological delimitations are still unclear [11]. These territories participate to the formation of some head muscles, part of the neurocranium, and to the formation of the vertebrate heart [12]. However, head muscles develop from multiple origins and can derive from anterior somite migrating cells or head mesoderm.

Extant chordates include the vertebrates, their sister group the tunicates, and the cephalochordates (i.e. amphioxus) [13]. In cephalochordates as in vertebrates, the paraxial mesoderm forms segmented blocks on both sides of the notochord through the somitogenesis

64 process. However, this is not the case in tunicates that probably lost this chordate feature. In
65 addition, a major difference exists between vertebrates and amphioxus somitogenesis. While
66 in vertebrates somites are restricted to the trunk mesoderm, in amphioxus, somites form from
67 the most anterior to the most posterior region of the animal [14]. Moreover, all the amphioxus
68 mesodermal structures, except the axial notochord, derive from the somites. Indeed,
69 cephalochordates do not possess lateral plate mesoderm as do vertebrates. Thus, during
70 cephalochordate embryogenesis, once the somites are formed, they elongate ventrally, and
71 several dorso-ventral regions can be recognized. Along the antero-posterior axis, the
72 cephalochordate somites can be subdivided in several populations. The anterior somites (eight
73 to ten depending on the species) form by enterocoely from the paraxial dorsal roof of the
74 archenteron [15], and the most anterior of these somites form under the control of the FGF
75 signal [16]. The most posterior somites form by schizocoely, directly from the tailbud [17], and the
76 signals controlling their formation are still unknown, although a role for FGF and retinoic acid
77 has been discarded [18]. Most of the genes expressed in the enterocoelic somites during their
78 formation are also expressed in the tailbud during posterior elongation and schizocoelic
79 somite development [19], suggesting that the differences between these two populations could
80 result from the dissimilarity in the physical constraints imposed to the presomitic region
81 during these two somitogenesis phases. In support to this view, when schizocoelic somites
82 form, the tailbud is reduced to a small number of cells whereas the roof of the archenteron
83 from which form the enterocoelic somites is a much larger region. Overall, there are until now
84 no functional data arguing for a genetic difference in the control of enterocoelic and
85 schizocoelic somites formation [19].
86 The complete paraxial mesoderm segmentation observed in amphioxus has been proposed to
87 be an ancestral chordate feature [20], implying that the vertebrate unsegmented head mesoderm
88 is a derived character whose evolutionary origin has not yet been elucidated. Several

89 propositions to answer such a question were enunciated (for reviews see [21-26]) among which:
90 (1) the mesoderm of the head appeared through a loss of segmentation of the anterior paraxial
91 mesoderm, (2) the head mesoderm is a new structure that was added in the anterior region or
92 that appeared after the loss of the anterior segmented paraxial mesoderm. Key to bring new
93 insights into this subject is our understanding of the molecular mechanisms controlling
94 somitogenesis in cephalochordates. In order to identify the downstream cascade activated by
95 FGF for the formation of the amphioxus anterior somites [16], we undertook a comparative
96 RNA-seq approach followed by a genetic functional analysis of several transcription factors
97 putatively involved in the control of the formation of the different somitic populations. We
98 show here that ETV1/4/5, Six1/2, Pax3/7 and Zic play different roles during amphioxus
99 somitogenesis, demonstrating the existence of three genetically different somite populations.
100 Moreover, our data on anterior somites formation, together with known literature on tunicates
101 and vertebrates, allow us to propose an evolutionary scenario according to which the
102 vertebrate head mesoderm is of visceral and not paraxial origin as previously proposed, and
103 that reconciles the two main opposed hypotheses on the origin of head muscles.
104

Results

Comparative RNA-seq approach reveals putative transcription factors downstream of FGF for the control of anterior somitogenesis

To decipher the genetic regulation occurring downstream of FGF during anterior somitogenesis in amphioxus, we undertook a comparative RNA-seq approach (Supplementary Data). We analysed the transcriptomes of embryos treated with the FGF signalling pathway inhibitor SU5402 at stages at which the treatment induces the loss of the anterior somites, and of embryos treated at a later stage in which all the somites form. We focused our attention on genes whose expression profile shows a significant downregulation precisely at the time when anterior somites form in early treated embryos but whose expression is not downregulated otherwise. Indeed, these genes are putative downstream targets of the FGF signalling pathway specifically controlling the formation of anterior somites. These downregulated genes are enriched in GO terms associated with transcription factors and cranial skeletogenesis (Supplementary Fig. 1). To validate our RNA-seq approach, we analysed the expression of more than 80 of them by *in situ* hybridization (Fig. 1, Supplementary Fig. 1). Our results confirmed that the expression of the candidate genes coding for transcription factors and signalling pathway actors normally expressed in the presumptive anterior somites territory was lost specifically in this region when the FGF signal was inhibited.

Role of ETV1/4/5, Six1/2 and Pax3/7 during amphioxus somitogenesis

Transcription factors putatively involved in the control of anterior somitogenesis identified by the RNA-seq experiment were studied using a functional approach. We first chose the Ets family member *ETV1/4/5* because it had been shown to be a target of FGF signalling in amphioxus [16], and because its vertebrate orthologues are known FGF downstream effectors

and direct targets [27,28]. We also selected *Pax3/7* and *Six1/2* because (i) together with the non transcription factor partner *Eya* they were highly downregulated early on after FGF signal inhibition, and (ii) their orthologues are key transcription factors controlling somitogenesis and trunk muscle formation in vertebrates [29]. We assessed the function of ETV1/4/5, Six1/2 and Pax3/7 by constructing constitutive transcriptional activator or repressor chimeras through the fusion of the transcription factor sequence to the VP16 transcriptional activation domain [30] or to the Engrailed [31] repressor domain, respectively. The overexpression of VP16 chimeras for the three transcription factors had no obvious effect on amphioxus embryonic development (Supplementary Fig. 3). Nevertheless, when we injected the *ETV1/4/5-Engrailed* mRNA, the embryos presented a similar phenotype to embryos in which the FGF signalling pathway has been inhibited during early development (Fig. 2a, Supplementary Fig. 4). Thus, the injected embryos did not form anterior somites although the posterior somites were present, as shown by the expression of *FoxC* and *Myosin Light Chain* (*MLC*) (Fig. 2a), whereas the notochord was visible all along the embryo, as evidenced by the expression of *Brachyury2* (Fig. 2a). Interestingly, we also observed this phenotype when the *Six1/2-Engrailed* mRNA was injected (Fig. 2a, Supplementary Fig. 4), suggesting that both ETV1/4/5 and Six1/2 are required downstream of FGF for anterior somitogenesis. On the other hand, the embryos injected with the *Pax3/7-Engrailed* mRNA showed a different phenotype. The embryos were shortened and formed anterior somites as indicated by *FoxC* and *MLC* expression (Fig. 2a). In addition, posterior elongation was stopped and no *Brachyury2* expression was detected after gastrulation (Fig. 2a), advocating for a role of *Pax3/7* in the formation of posterior and not of anterior somites.

To better understand the epistatic relationships between these transcription factors, we analysed their expression in injected embryos. We showed that in *ETV1/4/5-Engrailed* mRNA injected embryos the expression of *Six1/2* and *Pax3/7* was lost in the presumptive anterior

paraxial mesoderm (Fig. 2b). When *Six1/2-Engrailed* mRNA was injected, *Pax3/7* anterior expression was also lost while *ETV1/4/5* expression was maintained (Fig. 2b). Finally, *ETV1/4/5* and *Six1/2* expression was maintained upon *Pax3/7-Engrailed* mRNA injection (Fig. 2b). Together these data suggest that ETV1/4/5 is at the top of the regulatory cascade downstream of the FGF signal followed by Six1/2, which is at a downstream position.

Zic is also an important actor in amphioxus somite formation

In order to discover *cis*-regulatory elements directly implicated in the regulation of anterior somitogenesis by FGF, whole genome ATAC-seq profiles for amphioxus embryos at gastrula and neurula (early and late) stages were generated. We specifically searched for peaks that (i) are located near the genes we found to be downregulated after early inhibition of the FGF signalling pathway (i.e. in intronic region or at less than 3 kb from the transcription start site), (ii) are present at the gastrula stage but absent at the beginning of neurulation, when the first somites are already specified, and (iii) contain putative Ets binding sites [32]. Only one peak fulfilling these criteria was identified in the first intron of *Zic* (Fig. 3a). The corresponding sequence was cloned in a GFP reporter plasmid upstream of the *Branchiostoma lanceolatum β-actin* minimal promoter [33]. Transient transgenic amphioxus injected with this construction showed *GFP* expression in the dorsal blastopore lip in gastrulae, precisely recapitulating the *Zic* expression pattern at this stage (Fig. 3b). This suggests that during anterior somite specification, *Zic* expression might be controlled through the binding of Ets family transcription factors (Fig. 3b). To evaluate the function of *Zic* during somitogenesis, we injected the *Zic-Engrailed* chimera mRNA into unfertilized amphioxus eggs. Injected embryos had no anterior somites as shown by the absence of anterior expression of *MLC* and *MRF2* thereby phenocopying SU5402-treated embryos (Fig. 3c, Supplementary Fig. 4).

To better understand the specific role of Pax3/7, Six1/2 and Zic downstream of the FGF signal, we co-injected the *Pax3/7-Engrailed* mRNA with the *Six1/2-Engrailed* or the *Zic-Engrailed* mRNAs. Hence, when we used constitutive repressor fusions to simultaneously interfere with the function of *Pax3/7* and *Six1/2*, we obtained embryos that had lost their anterior and posterior somitic structures and only presented some central somites (Fig. 3d). This phenotype is similar to embryos injected with *Pax3/7-Engrailed* mRNA and treated with SU5402 (Fig. 3d). By contrast, when *Pax3/7-Engrailed* and *Zic-Engrailed* mRNAs were co-injected, no somites formed (Fig. 3d).

Discussion

By employing a functional genetic approach we confirmed the presence of three different somite populations in amphioxus (anterior, intermediate and posterior) [16], and we further showed that the formation of each somite type is controlled by a specific set of transcription factors (Fig. 4a). The formation of the most anterior somites is under the control of FGF, probably through the Ets factor ETV1/4/5 which regulates the expression of *Six1/2* and *Pax3/7*, Six1/2 being indispensable for the establishment of this population of somites. On the other hand, Zic seems to be implicated in the development of all the anterior enterocoelic somites while Pax3/7 would be required for the formation of the posterior schizocoelic somites, although from our data we cannot exclude that *Pax3/7* is also controlling the formation of the posterior enterocoelic somites (Fig. 4a). Interestingly, in vertebrates, although the all somites form through a similar program, the signals controlling their formation are different in the anterior and posterior regions. Indeed, the clock and wavefront system applies well to the posterior somites whereas the formation of the occipital somites is often resistant to perturbations of the Notch signalling pathway, which is one of the main component of the clock (see [34] for a review). However, these differences in anterior and posterior somitogenesis can be hardly compared to what is known in amphioxus as there is no implication of retinoic acid nor FGF signals in the formation of the posterior somites [18], and Notch signal perturbation induces incomplete formation of the segmental boundaries of all the cephalochordate somites [35]. We thus propose that the differences observed between the formation of occipital somites and more posterior somites in vertebrates cannot be paralleled to the differences we observed in this study in the formation of the different amphioxus somite populations.

The functional results we obtained in this work have some important implications for our understanding of the evolutionary origin of the vertebrate unsegmented head mesoderm. In vertebrates, *Pax3*, together with *Six1* and *Six4* and their cofactors *Eya1* and *Eya2*, are important for somite formation and are the main actors activating the expression of the basic helix-loop-helix muscle regulatory factor genes (MRFs) that launch myogenesis in specific regions of the somites [29]. The formation of amphioxus muscles from all three somite populations would therefore be regulated by a similar set of transcription factors to the one controlling trunk muscle formation in vertebrates. On the other hand, the myogenic process in the vertebrate head mesoderm, which is delayed compared to trunk myogenesis and *Pax3* independent, is triggered by *Pitx2* [36] and *Tbx1* [37] upstream of MRFs [29]. Even if *Pax7* is expressed in head muscle stem cells, it is only at a late stage after *MRFs* expression [6,38], and although *Six1* has been shown to control the expression of some *MRFs* in vertebrate head mesoderm derived muscles, it is always acting downstream of, or in parallel with, *Tbx1* [39,40]. Interestingly, in contrast to what we observed in amphioxus, *Six1* has been shown to be upstream of FGF signal by regulating *Fgf8* expression in some pharyngeal arches and in the second heart field in mouse [39]. Altogether, these data suggest that the vertebrate *Pax* and *Six* genes are not major upstream regulators of head muscle formation even so they are implicated in some vertebrate species in the activation of some MRFs and/or the maintenance of some head muscle cell progenitors. Interestingly, in amphioxus, *Pitx* expression starts after somitogenesis on the left side of the embryo (Supplementary Fig. 5) [41,42], ruling out a possible role in somite and muscle formation in this species. Consistently, the injection of a *Pitx-Engrailed* (*Pitx-EN*) chimera mRNA in amphioxus embryos does not lead to somitogenesis defects [43]. In addition, the onset of amphioxus *Tbx1/10* expression occurs well after gastrulation, in the ventral region of the already formed somites (Supplementary Fig. 5) [44], and somites still form in *Tbx1/10* morpholino-injected *Branchiostoma floridae* embryos [45].

Therefore, from the genetic point of view, vertebrate head mesoderm seems at a first glance not to be homologous to the amphioxus anterior somites, at least in their entirety.

While the aforementioned data might seem to support the hypothesis stating that vertebrate head muscles are a vertebrate novelty that was "added" to an amphioxus-like body plan [46,47], we would like to propose a distinct interpretation. Indeed, one problem of the "head addition" scenario is that it supposes that amphioxus somites are exclusively homologous to vertebrate somites. However, we and others have provided a sound argumentation against this view [48,49], and we suggest that amphioxus somites are homologous to three vertebrate mesodermal compartments: the somites, the cardiopharyngeal mesoderm, and the lateral plate mesoderm. In support of this, the ventral region of amphioxus somites express lateral plate mesoderm markers such as *Ets1/2*[50] , *FoxF* [51], *GATA1/2/3* [52], *Hand* [52], and *Twist* [50]. Moreover, this ventral somitic region also expresses heart markers such as *Nk2-tin* [53], *Tbx20* [54] together with head mesoderm genes such as *Alx* [50] or *Tbx1/10* [44]. In vertebrates, it has been shown that some progenitors of the vertebrate second heart field derive from the head mesoderm [55,56] and, in their sister group the tunicates, the anterior muscles (cardiac muscle cells and muscles of the atrial siphon) have a common origin and depend upon *Tbx1/10* for their formation [57,58]. Thus, the comparison of cell fates and gene expression patterns with amphioxus strongly suggest that the ventral part of amphioxus somites, which is therefore segmented at early embryonic stages, would be homologous to both the cardiopharyngeal field of vertebrates and tunicates [12,59] and to the lateral plate mesoderm of vertebrates. Hence, the head mesoderm of vertebrates, at least the pharyngeal mesoderm, would be of visceral and not paraxial origin as already proposed [12]. It would therefore not be a completely novel structure, but a structure homologous to the ventral part of the amphioxus anterior somites.

Altogether, if we assume that the complete anteroposterior segmentation of the paraxial mesoderm, as found in cephalochordates, represents the ancestral state within

262 chordates, our functional data in amphioxus and known gene expression patterns in chordates
263 allow us to propose an evolutionary scenario that reconciles aspects of the two traditional and
264 seemingly conflictual hypotheses in the field (*i.e.* "head addition" *vs* "segmentation loss").
265 Hence, we propose a series of evolutionary steps explaining how the vertebrate head
266 mesoderm might have derived from the ventral part of an ancestrally fully segmented anterior
267 paraxial mesoderm (Fig. 4b). The first two phases, the order of which we cannot define,
268 would have involved: (i) the segregation of the paraxial mesoderm from the lateral plate
269 mesoderm during gastrulation and the loss of segmentation of the lateral plate mesoderm
270 along the whole antero-posterior axis, and (ii) the regionalisation of the lateral plate
271 mesoderm, in an anterior and a posterior zone, as previously proposed by others [48,60].
272 Subsequently, the paraxial mesoderm of the anterior region would have been lost, probably
273 through a functional modification of the role of the FGF signalling pathway, as supported by
274 our data. This last step would have played a crucial role by relaxing the developmental
275 constraints imposed by a segmented paraxial mesoderm and allowing the lateral plate
276 mesoderm to occupy this evolutionary "old" territory for the formation of a "novel" muscular
277 system.
278

Methods

Embryo manipulation

Ripe adults from the Mediterranean amphioxus species (*Branchiostoma lanceolatum*) were collected at the Racou beach near Argelès-sur-Mer, France, (latitude 42° 32' 53" N and longitude 3° 03' 27" E) with a specific permission delivered by the Prefect of Region Provence Alpes Côte d'Azur. *Branchiostoma lanceolatum* is not a protected species. Gametes were collected by heat stimulation as previously described [61,62]. Prior to pharmacological treatments, and before hatching, embryos were transferred to new Petri dishes with a known final volume of seawater. SU5402 (Calbiochem 572631) was dissolved in dimethyl sulfoxide (DMSO) at 10^{-2} M and added to cultures of embryos at a final concentration of 25 µM at the blastula stage (5 hours post fertilization (hpf) at 19°C) or at the gastrula stage (15,5 hpf at 19°C). Control embryos were raised simultaneously with equivalent concentrations of DMSO in filtered seawater. Embryos were either fixed in PFA4%-MOPS as previously described [63] or frozen in liquid nitrogen.

RNA-seq experiment

Total RNA was extracted from embryos 3, 6 or 9 hours post treatment (hpt) using the RNeasy Plus Mini Kit (Qiagen) after disrupting and homogenizing the sample with the TissueLyser (Qiagen). Library preparation and sequencing were performed at the GenomEast Platform, IGBMC, Illkirch, France. Illumina "Truseq RNA sample preparation low throughput" protocol was followed for cDNA synthesis (using 2 µg total RNA), then SPRIworks Fragment Library System I kit (ref A84801, Beckman Coulter, Inc) with the SPRI-TE instrument was used to prepare the libraries, afterwards libraries were purified using AMPure XP beads (Agencourt Biosciences Corporation). Single-end sequencing was performed on Illumina GAIIx platform (54 or 72 bp, 6 hpt libraries) or on Hiseq2000 system (50 bp, 3 and 9 hpt

libraries). The first 50 bases of pass-filter reads were retained, in order to be comparable between the different samples. These reads were mapped onto a reference transcriptome constructed using the data obtained by Oulion et al., [64] and the data obtained in this study, using bwa v0.6.1 [65] and the following set of parameters : -l 27 -n 4 -e 4. Only uniquely aligned reads were then retained. Subsequent analysis was performed using R v2.15.2 : the number of reads aligned to each contig was computed and normalization and differential expression analysis was performed using DESeq v1.10.1 [66]. GO terms enrichment analysis was undertaken using Blast2GO [67].

In situ hybridization

For *B. lanceolatum* genes not previously published, specific primers were designed for RT-PCR amplification of partial coding regions. Total RNA of *B. lanceolatum* extracted from a mix of embryos at different developmental stages was used as a template for retro-transcription. Amplification was performed using Advantage 2 Polymerase kit (Clontech) and a touchdown PCR program with annealing temperature ranging from 65 to 40°C. Amplified fragments were cloned using the pGEM-T Easy system (Promega) and sub-cloned in pBluescript II KS+ for probe synthesis. For *GFP*, probe was synthesized from a pcDNA3-spacer-GFP-NX plasmid (gift from Angela Nieto and Jose Manuel Mingot). Whole mount *in situ* hybridizations were performed as described in [68]. After in situ hybridization some embryos were washed several times in PBS and labelled using DAPI for further confocal microscopy imaging at the BIOPIC platform.

Plasmid constructions

All the vectors for mRNA synthesis were constructed using the pCS2+ expression vector backbone. Constitutive activator forms of Pax3/7 (VP16-Pax3/7), Six1/2 (VP16-Six1/2) and

ETV1/4/5 (VP16-ETV1/4/5) were created by fusing the coding sequence of the 81 aa activation domain of VP16 protein to the N-terminal side of the DNA binding domain coding sequence of *Pax3/7* or *ETV1/4/5* and to the N-terminal side of the full-length coding sequence of *Six1/2*. Constitutive repressor forms of Pax3/7 (Pax3/7-Engrailed), Six1/2 (Six1/2-Engrailed) and ETV1/4/5 (ETV1/4/5-Engrailed) were created by fusing the coding sequence of the repressor domain of the engrailed protein [31] to the N-terminal side of the DNA binding domain coding sequence of *Pax3/7* or *ETV1/4/5* and to the N-terminal side of the full-length sequence of *Six1/2*. The vectors were linearized and *in vitro* transcription was performed using the mMESSAGE mMACHINE® SP6 Transcription Kit. Microinjections of plasmids and mRNA were carried out as described in [69,70].

ATAC-seq experiment

ATAC-seq was undertaken as previously described in [71,72]. Embryos were grown at 19°C until 8 hpf, 15 hpf, 36 hpf and 100, 30, 13 embryos were centrifuged at 13,000 rpm to remove seawater. Embryos were resuspended in 50 µl of cold lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% Igepal). Half of the lysate was centrifuged at 500 g for 10 minutes at 4°C, whereas the other half was used to count nuclei. Supernatant was removed and the nuclei were resuspended in the following transposition mix (25 µl 2x TD buffer (Illumina), 2.5 µL Tn5 transposase (Illumina), 22.5 µL nuclease free H2O) and incubated at 37°C for 30 minutes. To adjust the pH, 3 µl of 3M AcoNa (pH5.3) were added to the reaction mix and the DNA was purified using the MinElute PCR purification Kit (Qiagen), following the manufacturer's instructions using 10 µL of elution buffer preheated at 37°C. The following components were combined for amplification: 10 µL of transposed DNA, 10 µL nuclease free H2O, 2.5 µL Nextera PCR primer 1 (25 µM) [73], 2.5 µL Nextera PCR primer 2 (25 µM) [73] and 25 µL NEBNext® high-fidelity 2x PCR master mix (NEB). The PCR program

is as follows: 72°C for 5 minutes, 98°C for 30 seconds, followed by 13 cycles at 98°C for 10 seconds, 63°C for 30 seconds and 72°C for 1 minute. Following PCR amplification, we added 3 µl of 3M AcoNa (pH5.3) and the library was purified using the MinElute PCR purification Kit (Qiagen), following the manufacturer's instructions with 20 µL of elution buffer (37°C).

Data Availability

Sequences for probe synthesis are available in Genbank (see Supplementary Table 1). RNA-seq data are available under Gene Expression Omnibus (GEO) accession GSE122245. ATAC-seq data sets presented in this study were previously used in [71,72] and are available under Gene Expression Omnibus (GEO) accession GSE68737.

References

1 Goethe, J. v. Das Schadelgrut aus sechs Wirbelknochen aufgebaut. *Zur Naturwissenschaft uberhaupt, besonders zur Morphologie* (1790).

2 Goodrich, E. S. *Studies on the structure & development of vertebrates*. (1986).

3 Owen, R. *The principal forms of the skeleton and of the teeth*. (Blanchard and Lea, 1854).

4 Kuratani, S., Horigome, N. & Hirano, S. Developmental morphology of the head mesoderm and reevaluation of segmental theories of the vertebrate head: evidence from embryos of an agnathan vertebrate, Lampetra japonica. *Developmental biology* **210**, 381-400 (1999).

5 Freund, R., Dorfler, D., Popp, W. & Wachtler, F. The metameric pattern of the head mesoderm--does it exist? *Anat Embryol (Berl)* **193**, 73-80 (1996).

6 Bothe, I. & Dietrich, S. The molecular setup of the avian head mesoderm and its implication for craniofacial myogenesis. *Dev Dyn* **235**, 2845-2860, doi:10.1002/dvdy.20903 (2006).

7 Kuratani, S. & Adachi, N. What are head cavities?—a history of studies on vertebrate head segmentation. *Zoological science* **33**, 213-228 (2016).

8 Dequeant, M. L. & Pourquie, O. Segmental patterning of the vertebrate embryonic axis. *Nat Rev Genet* **9**, 370-382 (2008).

9 Hubaud, A. & Pourquie, O. Signalling dynamics in vertebrate segmentation. *Nat Rev Mol Cell Biol* **15**, 709-721 (2014).

10 Yusuf, F. & Brand-Saberi, B. The eventful somite: patterning, fate determination and cell division in the somite. *Anatomy and Embryology* **211**, 21-30, doi:10.1007/s00429-006-0119-8 (2006).

11 Tzahor, E. Heart and craniofacial muscle development: a new developmental theme of distinct myogenic fields. *Dev Biol* **327**, 273-279, doi:10.1016/j.ydbio.2008.12.035 (2009).

12 Sambasivan, R., Kuratani, S. & Tajbakhsh, S. An eye on the head: the development and evolution of craniofacial muscles. *Development* **138**, 2401-2415, doi:10.1242/dev.040972 (2011).

13 Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965-968 (2006).

14 Bertrand, S. & Escriva, H. Evolutionary crossroads in developmental biology: amphioxus. *Development* **138**, 4819-4830, doi:10.1242/dev.066720 (2011).

15 Holland, L. Z., Kene, M., Williams, N. A. & Holland, N. D. Sequence and embryonic expression of the amphioxus engrailed gene (AmphiEn): the metameric pattern of transcription resembles that of its segment-polarity homolog in Drosophila. *Development* **124**, 1723-1732. (1997).

16 Bertrand, S. *et al.* Amphioxus FGF signaling predicts the acquisition of vertebrate morphological traits. *Proc Natl Acad Sci U S A* **108**, 9160-9165 (2011).

17 Schubert, M., Holland, L. Z., Stokes, M. D. & Holland, N. D. Three amphioxus Wnt genes (AmphiWnt3, AmphiWnt5, and AmphiWnt6) associated with the tail bud: the evolution of somitogenesis in chordates. *Dev Biol* **240**, 262-273 (2001).

18 Bertrand, S. *et al.* Evolution of the Role of RA and FGF Signals in the Control of Somitogenesis in Chordates. *PLoS One* **10**, e0136587, doi:10.1371/journal.pone.0136587 (2015).

19 Beaster-Jones, L. *et al.* Expression of somite segmentation genes in amphioxus: a clock without a wavefront? *Dev Genes Evol* **218**, 599-611 (2008).

20 Holland, P. W. H. Embryonic development of heads, skeletons and amphioxus: Edwin S. Goodrich revisited. *Int J Dev Biol* **44**, 29-34 (2000).

21 Gans, C. & Northcutt, R. G. Neural crest and the origin of vertebrates: a new head. *Science* **220**, 268-274 (1983).

22 Holland, L. Z., Holland, N. D. & Gilland, E. Amphioxus and the evolution of head segmentation. *Integr Comp Biol* **48**, 630-646, doi:10.1093/icb/icn060 (2008).

23 Kuratani, S. Is the vertebrate head segmented?-evolutionary and developmental considerations. *Integr Comp Biol* **48**, 647-657, doi:10.1093/icb/icn015 (2008).

24 Kuratani, S. & Schilling, T. Head segmentation in vertebrates. *Integr Comp Biol* **48**, 604-610, doi:10.1093/icb/icn036 (2008).

25 Northcutt, R. G. Historical hypotheses regarding segmentation of the vertebrate head. *Integr Comp Biol* **48**, 611-619, doi:10.1093/icb/icn065 (2008).

26 Onai, T., Adachi, N. & Kuratani, S. Metamerism in cephalochordates and the problem of the vertebrate head. *Int J Dev Biol* **61**, 621-632, doi:10.1387/ijdb.170121to (2017).

27 Raible, F. & Brand, M. Tight transcriptional control of the ETS domain factors Erm and Pea3 by Fgf signaling during early zebrafish development. *Mech Dev* **107**, 105-117 (2001).

28 Roehl, H. & Nusslein-Volhard, C. Zebrafish pea3 and erm are general targets of FGF8 signaling. *Curr Biol* **11**, 503-507 (2001).

29 Buckingham, M. Gene regulatory networks and cell lineages that underlie the formation of skeletal muscle. *Proc Natl Acad Sci U S A* **114**, 5830-5837, doi:10.1073/pnas.1610605114 (2017).

30 Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. GAL4-VP16 is an unusually potent transcriptional activator. *Nature* **335**, 563-564, doi:10.1038/335563a0 (1988).

31 Jaynes, J. B. & O'Farrell, P. H. Active repression of transcription by the engrailed homeodomain protein. *EMBO J* **10**, 1427-1433 (1991).

32 Khan, A. *et al.* JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* **46**, D1284, doi:10.1093/nar/gkx1188 (2018).

33 Feng, J., Li, G., Liu, X., Wang, J. & Wang, Y. Q. Functional analysis of the promoter region of amphioxus beta-actin gene: a useful tool for driving gene expression in vivo. *Mol Biol Rep* **41**, 6817-6826, doi:10.1007/s11033-014-3567-x (2014).

34 Pais-de-Azevedo, T., Magno, R., Duarte, I. & Palmeirim, I. Recent advances in understanding vertebrate segmentation. *F1000Research* **7**, 97-97, doi:10.12688/f1000research.12369.1 (2018).

35 Onai, T., Aramaki, T., Inomata, H., Hirai, T. & Kuratani, S. On the origin of vertebrate somites. *Zoological letters* **1**, 33-33, doi:10.1186/s40851-015-0033-0 (2015).

36 Shih, H. P., Gross, M. K. & Kioussi, C. Expression pattern of the homeodomain transcription factor Pitx2 during muscle development. *Gene Expr Patterns* **7**, 441-451, doi:10.1016/j.modgep.2006.11.004 (2007).

37 Dastjerdi, A. *et al.* Tbx1 regulation of myogenic differentiation in the limb and cranial mesoderm. *Dev Dyn* **236**, 353-363, doi:10.1002/dvdy.21010 (2007).

38 Nogueira, J. M. *et al.* The emergence of Pax7-expressing muscle stem cells during vertebrate head muscle development. *Frontiers in Aging Neuroscience* **7**, doi:10.3389/fnagi.2015.00062 (2015).

39 Guo, C. *et al.* A Tbx1-Six1/Eya1-Fgf8 genetic pathway controls mammalian cardiovascular and craniofacial morphogenesis. *J Clin Invest* **121**, 1585-1595, doi:10.1172/JCI44630 (2011).

40 Lin, C. Y. *et al.* The transcription factor Six1a plays an essential role in the craniofacial myogenesis of zebrafish. *Dev Biol* **331**, 152-166, doi:10.1016/j.ydbio.2009.04.029 (2009).

41 Boorman, C. J. & Shimeld, S. M. Pitx homeobox genes in Ciona and amphioxus show left-right asymmetry is a conserved chordate character and define the ascidian adenohypophysis. *Evol Dev* **4**, 354-365 (2002).

42 Yasui, K., Zhang, S., Uemura, M. & Saiga, H. Left-right asymmetric expression of BbPtx, a Ptx-related gene, in a lancelet species and the developmental left-sidedness in deuterostomes. *Development* **127**, 187-195 (2000).

43 Li, G. *et al.* Cerberus-Nodal-Lefty-Pitx signaling cascade controls left-right asymmetry in amphioxus. *Proc Natl Acad Sci U S A* **114**, 3684-3689, doi:10.1073/pnas.1620519114 (2017).

44 Mahadevan, N. R., Horton, A. C. & Gibson-Brown, J. J. Developmental expression of the amphioxus Tbx1/ 10 gene illuminates the evolution of vertebrate branchial arches and sclerotome. *Dev Genes Evol* **214**, 559-566, doi:10.1007/s00427-004-0433-1 (2004).

45 Koop, D. *et al.* Roles of retinoic acid and Tbx1/10 in pharyngeal segmentation: amphioxus and the ancestral chordate condition. *Evodevo* **5**, 36, doi:10.1186/2041-9139-5-36 (2014).

46 Nandkishore, N., Vyas, B., Javali, A., Ghosh, S. & Sambasivan, R. Correction: Divergent early mesoderm specification underlies distinct head and trunk muscle programmes in vertebrates (doi: 10.1242/dev.160945). *Development* **145**, doi:10.1242/dev.173187 (2018).

47 Nandkishore, N., Vyas, B., Javali, A., Ghosh, S. & Sambasivan, R. Divergent early mesoderm specification underlies distinct head and trunk muscle programmes in vertebrates. *Development* **145**, doi:10.1242/dev.160945 (2018).

48 Onimaru, K., Shoguchi, E., Kuratani, S. & Tanaka, M. Development and evolution of the lateral plate mesoderm: comparative analysis of amphioxus and lamprey with implications for the acquisition of paired fins. *Dev Biol* **359**, 124-136, doi:10.1016/j.ydbio.2011.08.003 (2011).

49 Holland, N. D. Formation of the initial kidney and mouth opening in larval amphioxus studied with serial blockface scanning electron microscopy (SBSEM). *Evodevo* **9**, 16, doi:10.1186/s13227-018-0104-3 (2018).

50 Meulemans, D. & Bronner-Fraser, M. Insights from amphioxus into the evolution of vertebrate cartilage. *PLoS One* **2**, e787, doi:10.1371/journal.pone.0000787 (2007).

51 Mazet, F., Amemiya, C. T. & Shimeld, S. M. An ancient Fox gene cluster in bilaterian animals. *Curr Biol* **16**, R314-316, doi:10.1016/j.cub.2006.03.088 (2006).

52 Pascual-Anaya, J. *et al.* The evolutionary origins of chordate hematopoiesis and vertebrate endothelia. *Dev Biol* **375**, 182-192, doi:10.1016/j.ydbio.2012.11.015 (2013).

53 Holland, N. D., Venkatesh, T. V., Holland, L. Z., Jacobs, D. K. & Bodmer, R. AmphiNk2-tin, an amphioxus homeobox gene expressed in myocardial progenitors: insights into evolution of the vertebrate heart. *Dev Biol* **255**, 128-137 (2003).

54 Belgacem, M. R., Escande, M. L., Escriva, H. & Bertrand, S. Amphioxus Tbx6/16 and Tbx20 embryonic expression patterns reveal ancestral functions in chordates. *Gene Expr Patterns* **11**, 239-243, doi:10.1016/j.gep.2010.12.006 (2011).

55 Lescroart, F. *et al.* Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo. *Development* **137**, 3269-3279, doi:10.1242/dev.050674 (2010).

56 Tirosh-Finkel, L., Elhanany, H., Rinon, A. & Tzahor, E. Mesoderm progenitor cells of common origin contribute to the head musculature and the cardiac outflow tract. *Development* **133**, 1943-1953, doi:10.1242/dev.02365 (2006).

57 Stolfi, A. *et al.* Early chordate origins of the vertebrate second heart field. *Science* **329**, 565-568, doi:10.1126/science.1190181 (2010).

58 Wang, W., Razy-Krajka, F., Siu, E., Ketcham, A. & Christiaen, L. NK4 antagonizes Tbx1/10 to promote cardiac versus pharyngeal muscle fate in the ascidian second heart field. *PLoS Biol* **11**, e1001725, doi:10.1371/journal.pbio.1001725 (2013).

59 Diogo, R. *et al.* A new heart for a new head in vertebrate cardiopharyngeal evolution. *Nature* **520**, 466-473, doi:10.1038/nature14435 (2015).

60 Tanaka, M. Developmental Mechanism of Limb Field Specification along the Anterior-Posterior Axis during Vertebrate Evolution. *J Dev Biol* **4**, doi:10.3390/jdb4020018 (2016).

61 Fuentes, M. *et al.* Insights into spawning behavior and development of the European amphioxus (Branchiostoma lanceolatum). *J Exp Zool B Mol Dev Evol* **308**, 484-493, doi:10.1002/jez.b.21179 (2007).

62 Fuentes, M. *et al.* Preliminary observations on the spawning conditions of the European amphioxus (Branchiostoma lanceolatum) in captivity. *J Exp Zool B Mol Dev Evol* **302**, 384-391, doi:10.1002/jez.b.20025 (2004).

63 Holland, L. Z. & Holland, N. D. Expression of AmphiHox-1 and AmphiPax-1 in amphioxus embryos treated with retinoic acid: insights into evolution and patterning of the chordate nerve cord and pharynx. *Development* **122**, 1829-1838 (1996).

534 64 Oulion, S., Bertrand, S., Belgacem, M. R., Le Petillon, Y. & Escriva, H. Sequencing
535 and analysis of the Mediterranean amphioxus (Branchiostoma lanceolatum)
536 transcriptome. *PLoS One* **7**, e36554, doi:10.1371/journal.pone.0036554 (2012).
537 65 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler
538 transform. *Bioinformatics* **26**, 589-595, doi:10.1093/bioinformatics/btp698 (2010).
539 66 Anders, S. & Huber, W. Differential expression analysis for sequence count data.
540 *Genome Biol* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).
541 67 Gotz, S. *et al.* High-throughput functional annotation and data mining with the
542 Blast2GO suite. *Nucleic Acids Res* **36**, 3420-3435, doi:10.1093/nar/gkn176 (2008).
543 68 Somorjai, I., Bertrand, S., Camasses, A., Haguenauer, A. & Escriva, H. Evidence for
544 stasis and not genetic piracy in developmental expression patterns of Branchiostoma
545 lanceolatum and Branchiostoma floridae, two amphioxus species that have evolved
546 independently over the course of 200 Myr. *Dev Genes Evol* **218**, 703-713,
547 doi:10.1007/s00427-008-0256-6 (2008).
548 69 Hirsinger, E. *et al.* Expression of fluorescent proteins in Branchiostoma lanceolatum
549 by mRNA injection into unfertilized oocytes. *J Vis Exp*, 52042, doi:10.3791/52042
550 (2015).
551 70 Yu, J. K., Holland, N. D. & Holland, L. Z. Tissue-specific expression of FoxD
552 reporter constructs in amphioxus embryos. *Dev Biol* **274**, 452-461,
553 doi:10.1016/j.ydbio.2004.07.010 (2004).
554 71 Acemel, R. D. *et al.* A single three-dimensional chromatin compartment in amphioxus
555 indicates a stepwise evolution of vertebrate Hox bimodal regulation. *Nat Genet* **48**,
556 336-341, doi:10.1038/ng.3497 (2016).
557 72 Marletaz, F. *et al.* Amphioxus functional genomics and the origins of vertebrate gene
558 regulation. *Nature* (Accepted).

73 Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**, 21 29 21-29, doi:10.1002/0471142727.mb2129s109 (2015).

Acknowledgements

The laboratory of H.E. was supported by the CNRS and the ANR-16-CE12-0008-01 and S.B. by the Institut Universitaire de France. D.A. holds a fellowship from CONICYT "Becas Chile". J-L.G-S. was supported by the Spanish government (grants BFU2016-74961-P, the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 740041) and the institutional grant Unidad de Excelencia María de Maeztu (MDM-2016-0687). H.E. and S.M. were supported by the ECOS-CONICYT C15S02 grant. We would like to acknowledge the Mediterranean Amphioxus Genome Consortium for providing access to genomic and transcriptomic data before publication. We would like to thank Tokiharu Takahashi for providing the Tol2 mRNA production plasmid. The confocal imaging was undertaken at the BIOPIC platform which belongs to the EMBRC-France infrastructure supported by the ANR (ANR-10-INBS-02).

Author Contributions

Conceptualization: H.E., S.B.

Methodology: H.E., S.B., J-L. G-S., S.M.

Investigation: D.A., C.K., L.M., I.M., L.S., S.B., H.E.

Resources: C.K., J-L. G-S., S.M.

Writing–Original Draft: D.A., S.B., H.E.

Writing –Review & Editing: D.A., L.S., C.K., I.M., S.M., J-L. G-S., H.E., S.B.

Supervision: H.E., S.B., J-L. G-S., S.M.

Competing interests

The authors declare no competing interests.

Figure legend

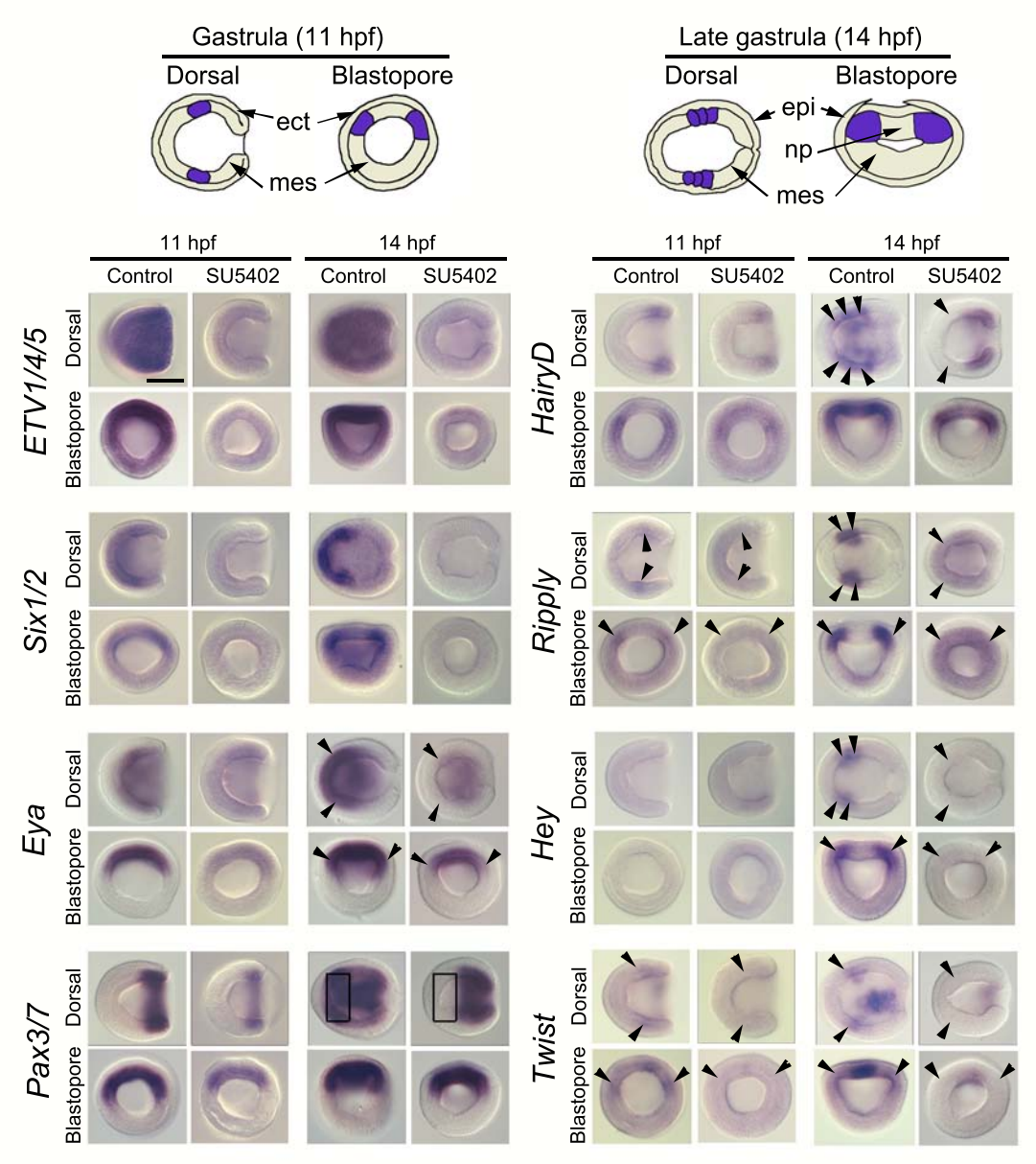
Figure 1. Expression of some candidate genes defined after RNA-seq data analysis in control and SU5402 treated embryos. On the top, schemes of embryos in dorsal view (anterior to the left) and blastopore view (dorsal to the top) are presented with the presumptive anterior somitic mesoderm coloured in purple. Ect: ectoderm; mes: mesendoderm; epi: presumptive epidermis; np: presumptive neural plate. The expression of candidate genes that are expressed in the presumptive anterior somite region in control embryos at 11 hpf and/or 14 hpf is lost in this territory after SU5402 treatment. For *ETV1/4/5* and *Six1/2*, the expression is completely lost. For *Eya* the expression in the mesendoderm is lost (arrowheads). *Pax3/7* anterior expression is lost after treatment (black rectangles). *HairyD*, *Ripply*, *Hey* and *Twist* expression is lost in the presumptive anterior somites territory (arrowheads). Dorsal views with anterior to the left and blastopore views with dorsal to the top. Scale bar, 50 µm. The expression was analyzed for at least 10 embryos for each stage, condition and gene.

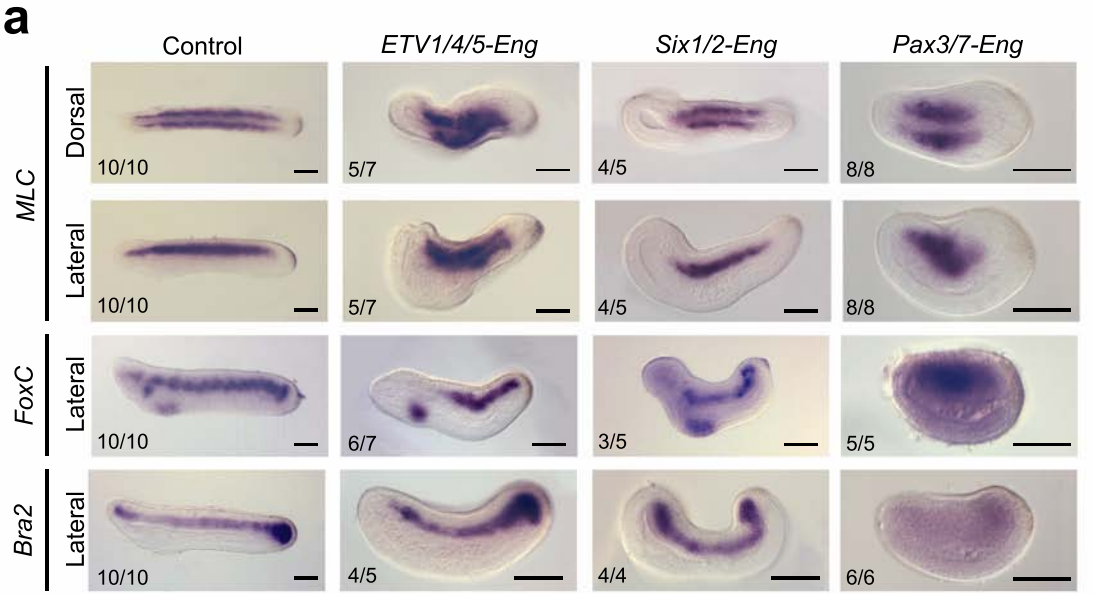
Figure 2. *ETV1/4/5, Six1/2* and *Pax3/7* are key factors for the formation of amphioxus somites. (a) Expression at 36 hpf of the mesodermal markers *MLC*, *FoxC* and *Bra2* in control embryos and embryos injected with the mRNA coding for the chimeras ETV1/4/5-Eng, Six1/2-Eng and Pax3/7-Eng. Anterior is to the left and dorsal is to the top in lateral views. Scale bars, 50µm. (b) Expression at 14 hpf of *ETV1/4/5*, *Six1/2*, *Pax3/7* and the myogenic factor gene *MRF2* in control embryos and in embryos injected with the mRNA coding for the chimeras ETV1/4/5-Eng, Six1/2-Eng and Pax3/7-Eng. Dorsal views with anterior to the left. Scale bar, 50 µm. The number of embryos showing the presented expression pattern is indicated on each panel.

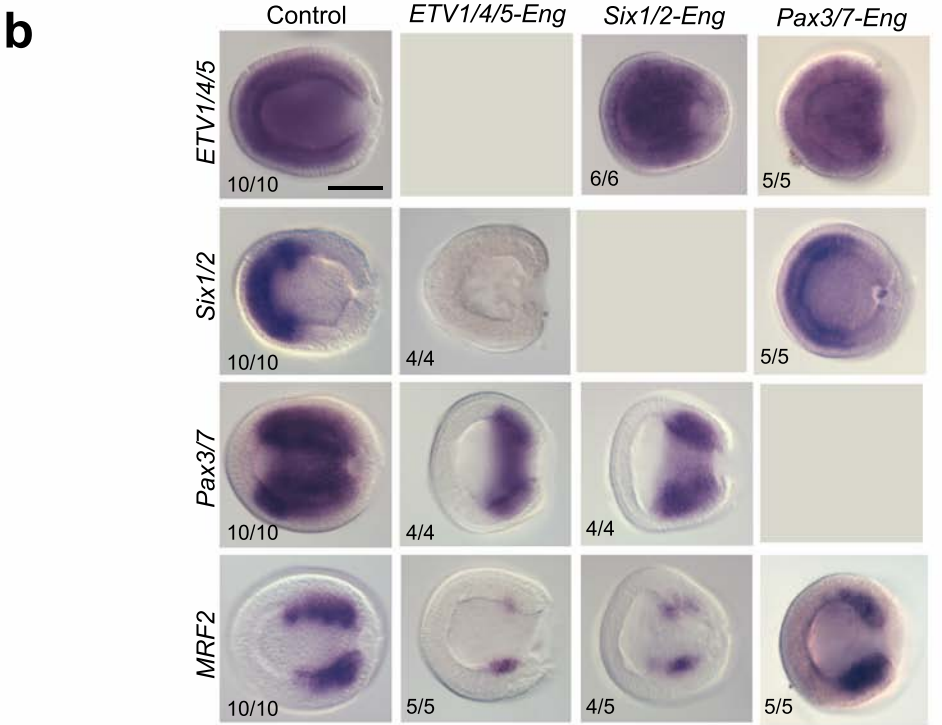
Figure 3. Zic is a major actor for somite formation in amphioxus. (a) Snapshot of the genomic browser at the *Zic* locus showing the RNA-seq data at 8 hpf and the ATAC-seq peaks at 8 hpf, 15 hpf and 36 hpf. The region with a peak in the first intron that was used for functional study is indicated. This region bears an Ets family putative binding site (orange box). (b) *Zic* and *GFP in situ* hybridization at 12 hpf in control embryos and in embryos injected with the reporter construct bearing the putative enhancer shown in (a), respectively. Dorsal is to the top. Scale bar, 50 µm. Ten injected embryos were analyzed and they all show expression in part of the dorsal blastopore lip. The injected embryo presented here is the one with the larger GFP expression. (c) Expression of *MLC*, *MRF2* and *Bra2* at 36 hpf in control embryos and in embryos after injection of *Zic-Eng* mRNA. Anterior is to the left and dorsal to the top for side views. Brackets indicate the anterior region where the notochord still forms whereas somites are absent. Scale bars, 50µm. The number of embryos showing the presented expression pattern is indicated on each panel. (d) Expression of *MLC* at 36 hpf in embryos after injection of *Pax3/7-Eng* and *Six1/2-Eng* mRNAs, after injection of *Pax3/7-Eng* mRNA and treatment with SU5402, or after injection of *Pax3/7-Eng* and *Zic-Eng* mRNAs. Dorsal views with anterior to the left. Scale bar, 50 µm. The number of embryos showing the presented expression pattern is indicated on each panel.

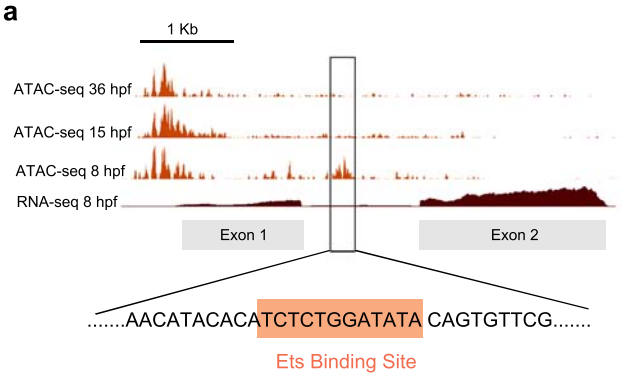
Figure 4. Gene regulatory logic for somites formation in amphioxus and hypothesis for the evolutionary scenario underlying vertebrate head mesoderm origin. (a) Scheme representing the relationships between FGF signal and genes studied in this work for the formation of the three populations of somites in amphioxus. Anterior and intermediate somites form by enterocoely. The most anterior somites formation is dependent upon the FGF signal that activates expression of target genes through the *ETV1/4/5* transcription factor. The

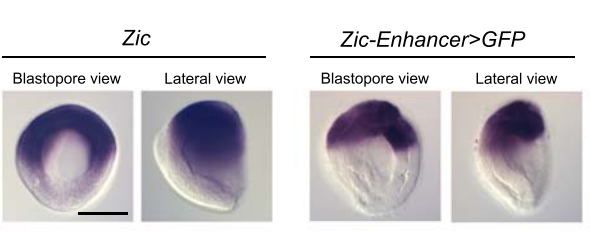
638 formation of intermediate somites is under the control of *Zic*. On the other hand, the posterior
639 somites that form by schizocoely are dependent upon *Pax3/7*. Dotted lines indicates
640 expression of the genes that are not required for somites formation. (b) Hypothesis concerning
641 the nature of the vertebrate head mesoderm and its evolutionary history. Lateral views of
642 putative embryos are schematized with anterior to the left and dorsal to the top. Genes
643 expressed in cardiac and pharyngeal mesoderm are in light red, genes expressed in lateral
644 plate mesoderm are in yellow, genes expressed in ventral amphioxus somites or in both
645 cardiopharyngeal and lateral plate mesoderm are in orange. See text for details.
646
647

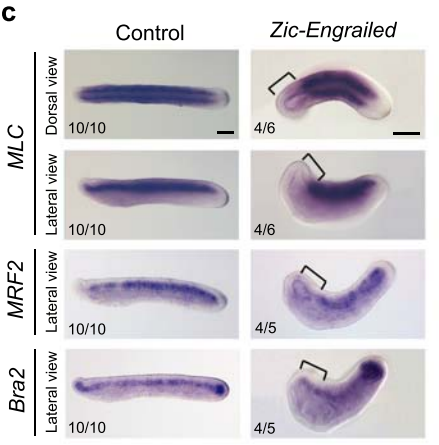

Gastrula (11 hpf)
Dorsal
Blastopore
ect
mes
Late gastrula (14 hpf)
Dorsal
Blastopore
epi
np
mes
11 hpf
14 hpf
Control
SU5402
Control
SU5402
ETV1/4/5
Six1/2
Eya
Pax3/7
HairyD
Ripply
Hey
Twist
Dorsal
Blastopore

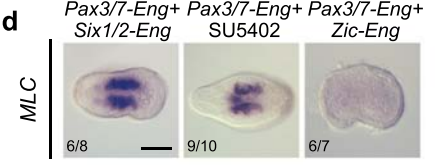

a
Control
ETV1/4/5-Eng
Six1/2-Eng
Pax3/7-Eng
MLC
Dorsal
Lateral
FoxC
Lateral
Bra2
Lateral
10/10
5/7
4/5
8/8
10/10
5/7
4/5
8/8
10/10
6/7
3/5
5/5
10/10
4/5
4/4
6/6

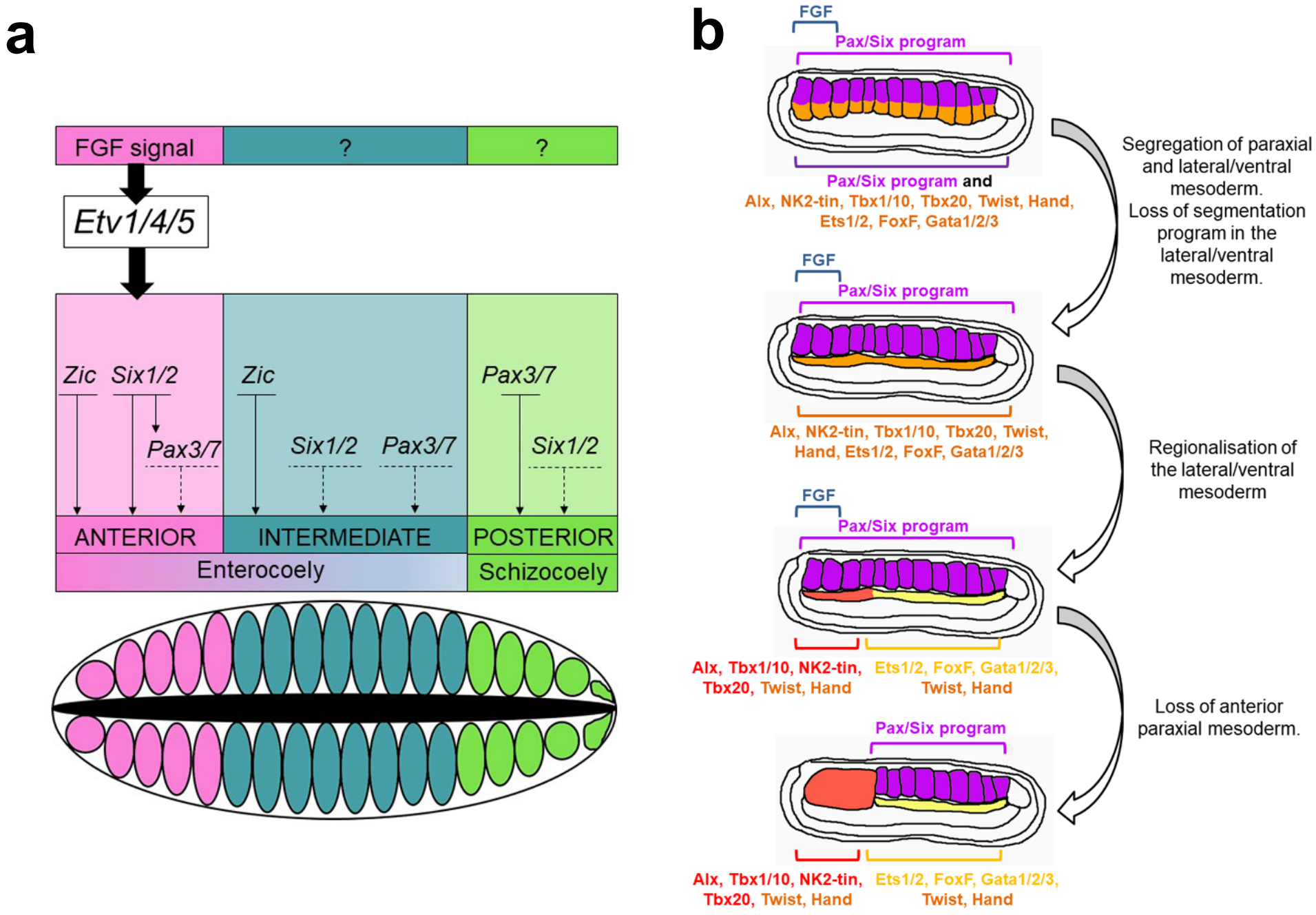

b
Control
ETV1/4/5-Eng
Six1/2-Eng
Pax3/7-Eng
ETV1/4/5
Six1/2
Pax3/7
MRF2
10/10
6/6
5/5
10/10
4/4
5/5
10/10
4/4
4/4
10/10
5/5
4/5
5/5


a
1 Kb
ATAC-seq 36 hpf
ATAC-seq 15 hpf
ATAC-seq 8 hpf
RNA-seq 8 hpf
Exon 1
Exon 2
.....AACATACACATCTCTGGATATA CAGTGTTCG.....
Ets Binding Site


b
Zic
Zic-Enhancer>GFP
Blastopore view
Lateral view
Blastopore view
Lateral view


c
Control
Zic-Engrailed
MLC
Dorsal view
10/10
4/6
Lateral view
10/10
4/6
MRF2
Lateral view
10/10
4/5
Bra2
Lateral view
10/10
4/5


d
Pax3/7-Eng+ Six1/2-Eng
Pax3/7-Eng+ SU5402
Pax3/7-Eng+ Zic-Eng
MLC
6/8
9/10
6/7


a
FGF signal
?
?
Etv1/4/5
Zic
Six1/2
Pax3/7
Zic
Six1/2
Pax3/7
Pax3/7
Six1/2
ANTERIOR
INTERMEDIATE
POSTERIOR
Enterocoely
Schizocoely
b
FGF
Pax/Six program
Pax/Six program and
Alx, NK2-tin, Tbx1/10, Tbx20, Twist, Hand,
Ets1/2, FoxF, Gata1/2/3
Segregation of paraxial and lateral/ventral mesoderm.
Loss of segmentation program in the lateral/ventral mesoderm.
FGF
Pax/Six program
Alx, NK2-tin, Tbx1/10, Tbx20, Twist,
Hand, Ets1/2, FoxF, Gata1/2/3
Regionalisation of the lateral/ventral mesoderm
FGF
Pax/Six program
Alx, Tbx1/10, NK2-tin,
Tbx20, Twist, Hand
Ets1/2, FoxF, Gata1/2/3,
Twist, Hand
Loss of anterior paraxial mesoderm.
Pax/Six program
Alx, Tbx1/10, NK2-tin,
Tbx20, Twist, Hand
Ets1/2, FoxF, Gata1/2/3,
Twist, Hand