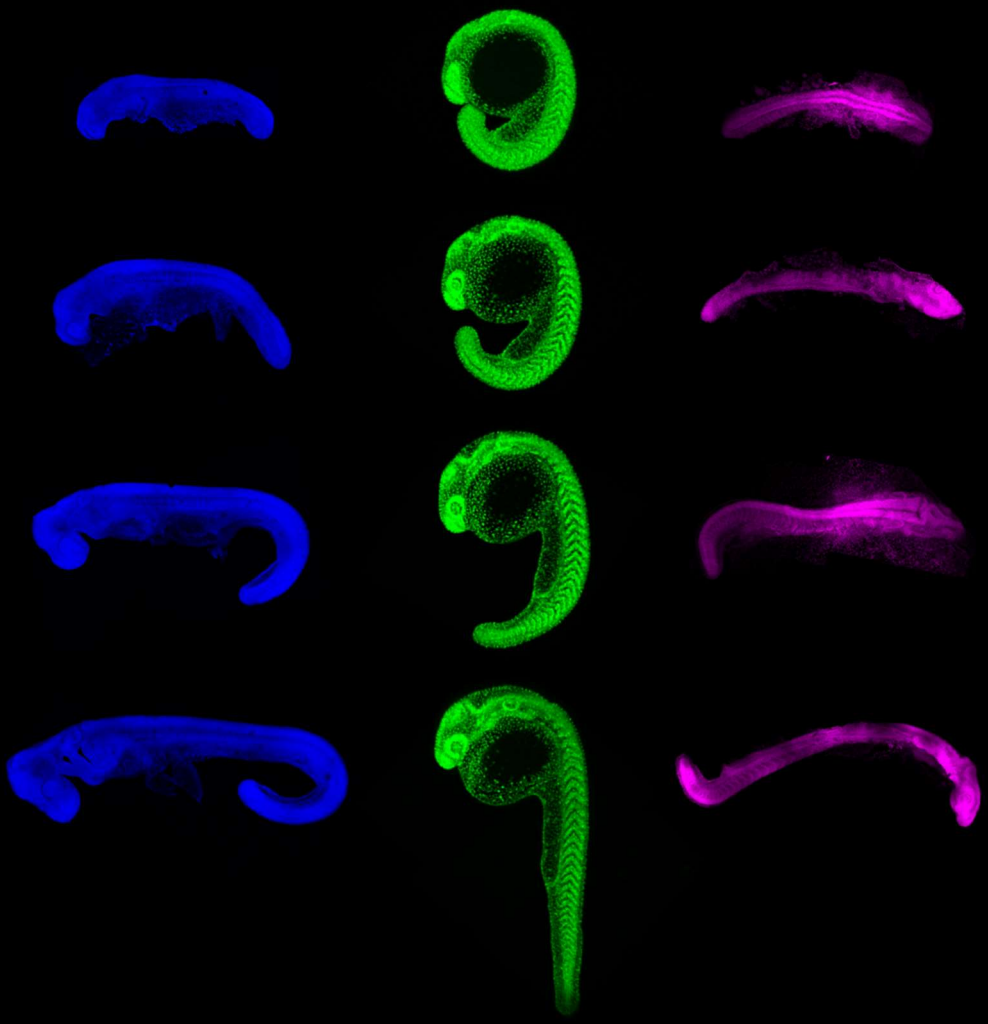
UNIVERSITY OF CAMBRIDGE

The evolution of somitogenesis:

Mechanisms of paraxial mesoderm elongation in zebrafish and other vertebrates



Lewis David Thomson, Darwin College, March 2021

This thesis is submitted for the degree of Doctor of Philosophy

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

It does not exceed the prescribed word limit for the Biology Degree Committee.

Lewis David Thomson

March 30, 2021

Title page image: Vertebrate embryos at different stages of somitogenesis. From left to right: catsharks fixed and stained with YO-PRO (blue); a live zebrafish injected with mRNA for KikGR (green), and Eastern Happy cichlids fixed and stained with DAPI (magenta).

Name Lewis David Thomson

Title The evolution of somitogenesis: Mechanisms of paraxial mesoderm elongation in zebrafish and other vertebrates

Abstract

In vertebrate embryos, a process called somitogenesis lays the foundations of the adult spine. This process involves elongation and segmentation of the paraxial mesoderm to form somites. Although the segmentation aspect of this has been widely studied, the elongation aspect is not well understood. Posterior growth is widely assumed to be the main driver, but there is very little evidence for this – particularly in fast-developing species like zebrafish. In this thesis, I present the first long term, multi-scale, 3D characterisation of the zebrafish paraxial mesoderm, and show that this tissue elongates through some form of convergent extension, not through growth. In fact, the tissue is compressed over time, and so decreases in volume. I suggest that these processes may be functionally linked, and thus propose a novel mechanism of "compression-extension". Cell tracking, agent-based modelling, and perturbations show that this form of convergent extension does not involve PCP-dependent directional intercalation but, instead, involves convergent flows of cells towards the midline and non-directional intercalation. The cause of compression is not clear, but perturbation experiments suggest that extrinsic forces from the neural tube and TGFβ signalling may be involved. Comparative work in cichlids, chickens, and catsharks suggests that tissue convergence is not unique to zebrafish, and instead is a conserved feature of paraxial mesoderm elongation – even in species that undergo high levels of growth during somitogenesis. This suggests that the relative contributions of growth and tissue convergence to the process of paraxial mesoderm elongation have evolved differently across vertebrate lineages, resulting in a spectrum of elongation strategies.

Acknowledgements

This thesis, and indeed this entire project, would not have been possible without the following people, who I would like to acknowledge.

Ben Steventon, my supervisor, has taught me a great deal over the last four years, and many of the ideas presented here are the direct results of my discussions with him. He has also been an incredible source of support and encouragement, and always had faith in myself and my work – even when I did not. I consider myself extremely fortunate to have had such a brilliant supervisor.

The BBSRC funded this project, and the Cambridge BBSRC DTP provided training and support throughout my PhD.

Leila Muresan wrote the MATLAB scripts that allowed me to measure neighbour mixing and neighbour angle changes (Sections 2.2 & 2.5). She was incredibly patient with my various requests for additional analysis options, and always sought to make the analysis tools as useful and as user-friendly as possible.

Berta Verd designed the agent-based model that provided a comparison with embryo cell tracks and allowed me to test the idea that convergent extension does not require directional intercalation (Section 2.6). Working with Berta was one of the highlights of this PhD project, and I have learned a great deal from her.

Dillan Saunders performed the neural tube laser ablations that allowed me to test a possible cause of presomitic mesoderm compression (Section 3.2). I am very grateful for his selflessness and willingness to help me with my project.

Toby Andrews provided countless helpful suggestions and ideas, including the idea of taking indirect measurements of the paraxial mesoderm (Section 1.2), and he explained how to apply cross-sectional area reductions to fixed volumes to calculate convergent extension length increases (Section 1.7). He and Lara Busby also discovered

how to validate Imaris spot reconstructions (through “masking” signal from the spots), which was incredibly useful for my analyses (Section 1.1).

Tim Fulton and Meagan Hennessy, in their roles as research assistants/lab managers, did all the most important (and often thankless) jobs that kept the lab running, and ensured all of us had the resources we needed. Tim Fulton trained me in various lab techniques when I first joined the lab, including zebrafish microinjections and basic molecular biology techniques. They also provided a great deal of moral support, primarily through the enforcement of regular tea breaks.

The rest of the Steventon lab, and others who joined our lab meetings, provided many helpful suggestions and feedback throughout the project. John Stuart also helped with molecular biology sample preparation.

The Fish Facility staff in the PDN department looked after the various fish lines and provided a huge amount of support and assistance.

Emilia Santos and her lab trained me in harvesting cichlid eggs, and also shared their own cichlid eggs with me, which allowed me to carry out the comparative work in cichlids (Section 4.2). Aleksandra Marconi, as well as these things, provided the staging images in Figure 34. Bethan Clark, during her time as a rotation student in our lab, also helped me fix, dissect, stain and image some of the cichlid HCRs (Section 4.2).

Sylvie Mazan and her lab hosted my research trip to Banyuls-sur-Mer, giving me the chance to carry out comparative work in catsharks (Section 4.4).

Various staff members in the Genetics Department, including my advisor Felipe Karam Teixeira, supported me throughout my PhD.

Jane Wishart, my tutor during my undergraduate degree at St Andrews, gave me my first real experience of scientific research; inviting me to be part of a fieldwork ecology research project in Trinidad. She helped me with all of my PhD applications, and has continued to provide support and guidance. I expect that, without her, I would not have secured a place on this PhD programme.

My parents, Martin and Lorna Thomson, have been an incredible source of love and encouragement during my PhD, and throughout my life. I am incredibly privileged to be their son. Martin Thomson, being a former maths teacher, also helped me translate some of my incomprehensible R code into somewhat comprehensible mathematical equations, so that it could be presented and explained here (Section 1.2).

And finally, my wife, Naushin Thomson. She is the greatest friend I could have asked for. She has walked alongside me every step of the way, supporting me when my experiments failed, and helping me to keep going. I can never thank her enough.

Contents

List of Abbreviations

Abbreviation	Meaning	Page
AIC	Akaike Information Criterion	53
AM	Autoregressive Motion [tracking algorithm]	79
AME	Autoregressive Motion Expert [tracking algorithm]	79
AP	anterior-posterior/anteroposterior	23
BM	Brownian Motion [tracking algorithm]	78
CC	Connected Components [tracking algorithm]	78
CSA	cross-sectional area	62
DEP+	Dishevelled protein with deleted DEP domain	38
DMSO	dimethyl sulfoxide	38
DV	dorsal-ventral/dorsoventral	47
ECM	extracellular matrix	31
GFP	Green Fluorescent Protein	18
GS	Gap Size [tracking algorithm parameter]	80
HH	Hamburger & Hamilton [chicken embryo stages]	138
HCR	hybridization chain reaction	39
INT	internal [somite measurement]	150
LI	Lineage [tracking algorithm]	79
LMP	low melting point [agarose]	39
MABT	maleic acid buffer (MAB) with 0.1% Tween-20 (T)	39
MD	Maximum Distance [tracking algorithm parameter]	80
ML	medial-lateral/mediolateral	47
MO	morpholino	30
NMp	neuromesodermal progenitor	27
NT	neural tube	110
OH	overhang [somite measurement]	150
PBS	phosphate-buffered saline	39
PCP	planar cell polarity	105
PFA	paraformaldehyde	39
PH3	phospho-histone H3	39
PM	paraxial mesoderm	84
PSM	presomitic mesoderm	15
RF	reference frame	77
SB50	SB-505124	38
SSCT	saline-sodium citrate (SSC) with 0.1% Tween-20 (T)	39
TL	Tüpfel Long Fin	38
WT	wild-type	30

INTRODUCTION

Vertebrates and other segmented animals

Vertebrates are an incredibly diverse group of animals; comprising over 70,000 described species (IUCN, 2020) spread across all major ecosystems. This group originated approximately 520 million years ago (Holland et al., 2008), and is made up of all jawless fishes; cartilaginous fishes; bony fishes; amphibians; reptiles; mammals; and birds. The defining feature of vertebrates is that they have a backbone/spine, made up of "vertebrae", which supports and protects the spinal cord. Vertebrates (and the wider group to which they belong, called "chordates") are thus classified as "segmented" animals, meaning their bodies are primarily formed of similar, repeated units/segments. Chordates are one of only three groups of truly segmented animals – the other groups being arthropods and annelids (Davis and Patel, 1999). Vertebrate segmentation is subtle, as each segment (of bone/cartilage, nerves, and muscle) is covered by skin/epidermis which is not segmented. Most arthropods (the phylum which includes insects; chelicerates; crustaceans; and myriapods), on the other hand, are visibly segmented. Their exoskeleton reflects their internal segmentation, and some segments have a pair of external appendages (e.g. legs, wings, pleopods). Annelids (the phylum which includes earthworms and leeches) are even more clearly segmented. Each annelid segment is almost identical, having the same set of internal organs and, in many cases, external appendages. Although vertebrate segmentation is more subtle than that of arthropods and annelids, it is still a highly important and conserved aspect of their anatomy.

However, despite this conservation in vertebrate segmentation, the number of vertebrae (and therefore body segments) varies greatly: from frogs with as few as 10 vertebrae to snakes with over 300 (Gomez et al., 2008). Additionally, the relative proportions of types of vertebrae varies between groups. Most mammals (including

giraffes) have exactly 7 neck vertebrae (Asher et al., 2011), whereas extinct plesiosaurs had up to 75 (Sachs, Kear and Everhart, 2013). Similarly, extinct sauropods had up to 80 tail vertebrae (Wilson, 2005), whereas humans have only 3-5, which are partially fused together to form the coccyx. Clearly, the vertebral column is a highly evolvable system. Indeed, it has been argued that segmented animals are more evolvable than unsegmented animals. This is related to the concept of modularity: if an animal is made up of discrete units/modules, a mutation which alters one unit is less likely to negatively affect the rest of the organism (Wagner and Altenberg, 1996). In the case of segmented animals, where the modules are individual segments, this means each segment can adapt more freely. This idea has some ecological support: chordates, arthropods and annelids are only 3 of the 40 animal phyla, but they represent over 80% of all animal species (Zhang, 2013) and approximately 90% of all animal biomass (Bar-On, Phillips and Milo, 2018). However, segmentation alone cannot account for the diversity and ecological success of these groups. The chordate phylum includes two other groups (as well as vertebrates): tunicates (sea squirts) and cephalochordates (lancelets). While the tunicates have secondarily lost their segmented body plan in evolution, cephalochordates have not – yet there are only ~20 extant species, compared with over 70,000 extant vertebrate species. Additionally, the second-most speciose animal phylum (after arthropods) is the molluscs, which are not segmented. Clearly, factors other than a segmented body plan have contributed to vertebrate diversity, and a segmented body plan is not a necessity for diversification. Nevertheless, the fact that all vertebrates share this basic body plan, despite great morphological diversity, shows that this feature has been important in facilitating their evolutionary success. Therefore, to understand vertebrate evolution, we must understand how vertebrate segmentation has evolved and diversified. And to understand this, we must understand the developmental basis of vertebrate segmentation.

Somitogenesis is the foundation of vertebrate segmentation

During vertebrate embryonic development there is a stage called the "phylotypic stage", which is loosely defined as the stage at which all members of the phylum show the most morphological similarity (Slack, Holland and Graham, 1993). One of the defining morphological features of this stage in chordates is the somites. These are paired epithelial blocks of mesoderm which run the length of the neural tube, on either side (Figure 1). In vertebrates, the somites are the precursor tissues of the vertebrae (and ribs), as well as the skeletal muscles and dermis associated with each vertebra (Dequéant and Pourquié, 2008). As such, the somites are the developmental foundation that the segmented adult body is built upon.

Somitogenesis is the process of somite formation from unsegmented paraxial mesoderm, also called the presomitic mesoderm (PSM). This segmentation process occurs in a bilaterally symmetrical fashion from anterior to posterior until all paraxial mesoderm is segmented. After a somite has formed, it is patterned into different portions which will give rise to different structures. Roughly speaking, the ventral half of each somite will become sclerotome (giving rise to cartilage/bone) and the dorsal half will become dermomyotome (giving rise to dermis and muscles) (Dequéant and Pourquié, 2008). However, the specific derivative that each part forms also depends on mediolateral position within the somite (Iimura et al., 2007; Brent and Tabin, 2002; Dequéant and Pourquié, 2008). The ventromedial portion will give rise to vertebrae, as cells undergo epithelial-to-mesenchymal transition (EMT), migrate around the notochord and neural tube, and condense. This portion will also give rise to the proximal part of the ribs. The ventrolateral portion will give rise to the distal part of the ribs and the hypaxial muscles (including limb and intercostal muscles). The dorsomedial portion will give rise to the dermis (deep skin layer) of the back and the epaxial muscles (including back muscles). The dorsolateral portion will only give rise to dermis of the back. The anteroposterior positional identity of somite derivatives along the body axis is determined by the

expression of *Hox* genes (Wellik, 2007), and this determines which somites will produce structures such as ribs and limb muscles.

The development of somite-sclerotome to vertebra is not a simple 1:1 correspondence. Rather, the cells from the anterior part of one somite-sclerotome combine with the cells from the posterior part of the somite-sclerotome anterior to it. In other words, a vertebra is formed not from a single somite, but from two halves of adjacent somites. This process is called resegmentation (Aoyama and Asamoto, 2000), as the segmental pattern established by the somites is shifted by half a segment to form the adult segmented pattern of the vertebrae. The anteroposterior patterning of the somites that drives this also directs the segmentation of the nervous system, to form a spinal nerve in each segment (Kelly Kuan et al., 2004).

The resegmentation process is not as restricted in teleost fishes, in that cells from one part of a somite-sclerotome can contribute to more than just one vertebra – a phenomenon described as "leaky" resegmentation by Morin-Kensicki, Melancon and Eisen (2002). However, given that all other studied vertebrate groups (including cartilaginous fishes) show clear resegmentation, it is likely that resegmentation is an ancestral feature of jawed vertebrates (Criswell and Gillis, 2020).

An additional feature which affects the correspondence of somites to vertebrae is that of diplospondyly, which means that for every segment, there are two vertebrae rather than one. Rather than a single somite giving rise to two halves of two adjacent vertebrae, in this case a single somite gives rise to four halves of three adjacent vertebrae. In other words, for every somite present, two vertebrae form. This process has only been observed in the tail vertebrae of cartilaginous fishes and not in any other vertebrate groups (Criswell and Gillis, 2020).

In summary, the process of somitogenesis is a fundamental process in vertebrates for establishing the segmented adult body. As well as ensuring the correct segmentation of somite derivatives, other tissues like the spinal cord are also segmented by the

patterning of the somites. Although processes such as resegmentation and diplospondyly mean there is not always an exact correspondence between somites and adult segments, somitogenesis is the foundation for vertebrate segmentation.

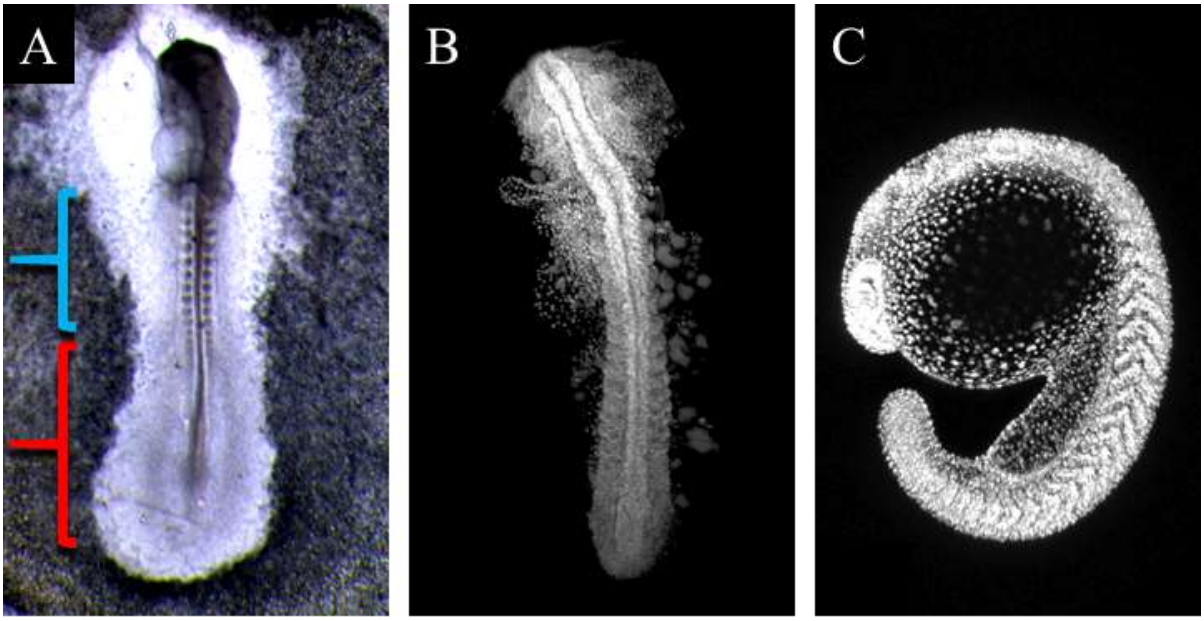


Figure 1: Vertebrate embryos during the process of somitogenesis. The presomitic mesoderm is sequentially segmented into somites from anterior to posterior. Somites are the embryonic precursors of the vertebrae, muscles, and dermis of the spine. Collectively, the presomitic mesoderm and the somites comprise the paraxial mesoderm. (A) A chick embryo (dorsal view, anterior top). The somites are in the region highlighted by the blue bar, and the presomitic mesoderm is in the region highlighted by the red bar. Image is a brightfield image of a live embryo. (B) A cichlid embryo (dorsal view, anterior top). Image is a confocal stack of a fixed embryo. Nuclei stained with DAPI (grey). (C) A zebrafish embryo (lateral view, anterior top left). Image is a confocal stack of a live embryo. Nuclei are expressing GFP (grey).

The process of somitogenesis

In order, then, to understand the evolution of vertebrate diversity (particularly the diversity of the vertebral column), we must understand the process of somitogenesis. The general process is quite simple: somite pairs pinch off from the anterior PSM progressively, from anterior to posterior, in a regular manner until roughly all of the paraxial mesoderm has been segmented. However, the details of how this is controlled and regulated have long been of interest to developmental and mathematical biologists. What determines somite number within a species, and how is this kept relatively constant? What determines the rate of somitogenesis, and how does this change in evolution? What determines the sizes, and relative sizes, of the somites? Cooke and Zeeman (1976) proposed "The clock and wavefront model" to answer these questions.

The clock and wavefront model

The clock and wavefront model, in its original form, involves two interacting processes. The "clock" represents cellular oscillations (between active and inactive phases), and the "wavefront" represents a sharp boundary (between permissive and non-permissive states) which gradually moves from anterior to posterior. Although all PSM cells are oscillating, only those anterior to the wavefront (i.e. in the permissive state) are patterned to form a somite when they are in their active phase. The wavefront then moves posteriorly, and the next group of cells are patterned to form the next somite. The process repeats until the wavefront reaches the posterior end of the PSM, and the last group of cells are patterned into the posterior-most somite. The model, therefore, suggests that somite number, somitogenesis rate, and somite sizes, are all controlled and regulated by the relative speeds of the clock and the wavefront. A faster clock, relative to the wavefront, will result in more somites which will be smaller in size. A faster wavefront, relative to the clock, will result in fewer, larger somites. A faster clock, regardless of wavefront velocity, will also result in a faster rate of somitogenesis.

Experimental support for the model

The first experimental support for this model came over 20 years later, when Palmeirim et al. (1997) showed that *hes1* (a member of the Notch signalling pathway) is cyclically expressed in chick PSM cells. These oscillations do not occur in phase along the PSM, but rather occur as oscillatory waves from posterior to anterior PSM (Figure 2). Cells in the posterior PSM express *hes1*, then cells anterior to these also express *hes1*, at which point the first group of cells stop their expression. This can be thought of as a repeating "Mexican wave" of expression. Importantly, the oscillation/clock period matches the rate of somitogenesis: one oscillation period for every somite produced - as predicted by the clock and wavefront model. The oscillatory wave has an added dynamic element in that the wave of expression slows down and shortens as it moves from posterior to anterior (Oates, Morelli and Ares, 2012). Similar cyclical expression patterns of Notch signalling genes were later reported in mouse, zebrafish, and frog embryos (Forsberg, Crozet and Brown, 1998; Holley, Geisler and Nüsslein-Volhard, 2000; Li et al., 2003), confirming a conserved role of Notch signalling in vertebrate somitogenesis. In chick and mouse embryos, genes from the Wnt and FGF signalling pathways (including *axin2* and *snail* homologs) also show similar cyclic expression (Dale et al., 2006; Dequéant et al., 2006), but there is no evidence of this in zebrafish or frog embryos (Mara and Holley, 2007; Dequéant and Pourquié, 2008). Importantly, those genes which show cyclic expression have been shown to be necessary for somitogenesis: mutant embryos for these genes have various segmentation defects (although there is some redundancy between certain genes) (Oates, Morelli and Ares, 2012). It is important to note that somite specification happens several cycles before the formation of a morphological somite. The anterior band of Notch activity is followed by a stripe of *mesp2* expression, which prevents further Notch cycling in this region, and activates downstream somite-specific genes like *MyoD* (Dequéant and Pourquié, 2008). This stripe of expression is not immediately posterior to the most recently formed somite, but 2-3

somite-lengths posterior to it (Sawada et al., 2001; Dequéant and Pourquié, 2008), indicating that the next 2-3 somites have already been specified.

Support for the existence of a wavefront was provided by the discovery of an *fgf8* expression gradient (from posterior to anterior) in the PSM of chick, zebrafish, and mouse embryos (Dubrulle, McGrew and Pourquié, 2001; Sawada et al., 2001; Dubrulle and Pourquié, 2004). These authors suggested that this *fgf8* gradient could create a threshold response; only cells below a certain level of *fgf8* expression would be able to respond to the clock signals to form a somite. They hypothesised that because the vertebrate embryo is elongating during somitogenesis, this threshold point would move posteriorly, acting as a wavefront. Dubrulle, McGrew and Pourquié (2001) and Sawada et al. (2001) showed that experimental inhibition of FGF signalling led to the formation of larger somites, while overactivation led to the formation of smaller somites, thus supporting the idea of an FGF-based wavefront in chick and zebrafish. Wnt signalling may also be involved in this wavefront; a similar posterior-to-anterior gradient of β-catenin protein (which is involved in Wnt signalling) has been reported in the mouse PSM by Aulehla et al., (2008). These authors overactivated β-catenin and found that somite specification occurred more anteriorly. However, these mutants did not form morphological somites, and so these results are not easily comparable with the FGF experiments in chick and zebrafish. It has also been suggested that retinoic acid (RA) may be involved in the wavefront. RA signalling is active in the formed somites and anterior PSM, but not in the posterior PSM, of zebrafish, frog, chick and mouse embryos (Mueller, Huang and Ho, 2010; Moreno and Kintner, 2004; Berggren et al., 1999; Vermot, 2005). This, combined with experiments showing mutual inhibition between RA and FGF signalling in the PSM of frog, chick and mouse embryos, led to the suggestion that the wavefront may be positioned by opposing gradients of RA and FGF (Moreno and Kintner, 2004; Diez del Corral et al., 2003; Abu-Abed et al., 2001). Inhibiting RA synthesis leads to the formation of smaller somites in chick and mouse embryos (Vermot

and Pourquié, 2005; Vermot, 2005), as would be predicted by a role of RA in the wavefront. However, it has been argued that RA inhibition-caused somite defects in mouse, at least, are caused by left-right symmetry defects, rather than any wavefront shift (Sirbu and Duester, 2006; Niederreither et al., 2002). And in zebrafish, inhibition of RA synthesis has no effect on somite formation or size (Linville et al., 2004; Berenguer et al., 2018).

Taken together, these results support the idea that somitogenesis is controlled by a clock-and-wavefront process (Figure 2). However, the details of which molecules are involved, and how they contribute, are not clear. It is difficult to determine whether the differences in results among model organisms are due to genuine evolutionary differences, or whether they are simply due to differences in experimental design – as it is not feasible to perform inhibition/overactivation experiments in exactly the same way for each organism. Nevertheless, it is clear that cyclic waves of signalling, combined with some form of positional identity/wavefront, specifies somite boundaries.

The importance of axis elongation

As previously mentioned, the posterior movement of the wavefront is, to some degree, dependent on the elongation of the embryo. The posterior to anterior gradient of *fgf8* expression is not the result of diffusion (i.e. a morphogen gradient), but of mRNA decay. Dubrulle and Pourquié (2004) showed that, in chick, only the most posterior cells are actively expressing *fgf8*, but that a clear gradient persists for several hours in the anterior PSM when the posterior region is surgically removed. If the gradient was the result of diffusion, then removing the source would lead to equilibrium across the tissue. These results, combined with the knowledge that posterior PSM cells are continuously moving into the anterior PSM (Knezevic, De Santo and Mackem, 1998; Lawton et al., 2013), show that the *fgf8* gradient is caused by mRNA decay as cells move anteriorly, out of the expressing region. As this anterior movement of cells is part of the elongation

process, there is a clear link between elongation and the wavefront. As the paraxial mesoderm elongates, the source of *fgf8* moves posteriorly, which causes the wavefront to move posteriorly. This process of elongation during the patterning and segmentation of the anterior-posterior (AP) axis is termed "axis elongation"; and understanding the mechanisms that drive this process is essential to understanding somitogenesis. However, this seemingly simple process is still not well understood.

The end of somitogenesis

This clock and wavefront model, in its simplest form, explains a "steady-state" form of somitogenesis in which the length of the PSM remains constant – as the length of each somite specified is equal to the amount of paraxial mesoderm elongation since the previous somite was specified. This raises the question of how this process ends: how is all of the paraxial mesoderm segmented into somites by the end of somitogenesis?

Measurements of PSM length over time in zebrafish, snake, chick, and mouse embryos show that the PSM length is not constant, but dynamic (Gomez et al., 2008). In zebrafish, the PSM decreases in length for the entirety of somitogenesis, suggesting that the length of each somite specified is *greater than* the amount of paraxial mesoderm elongation between somite stages. In other words, the rate of depletion (to form somites) is greater than the rate of elongation, and so there is a net length decrease over time. In the other three species, PSM length increases then decreases, suggesting that, initially, the length of each somite specified is *less than* the amount of elongation, but at later stages the length of each somite specified is *greater than* the amount of elongation. Given that the length of each somite at the time of formation is relatively constant (when compared with PSM length), it is likely that this dynamic is due to a slowing of elongation, rather than changes in somite lengths. Importantly, there is no steady state phase in any of the species – the PSM is always either increasing or decreasing in length.

It would seem then, that somite size is not simply specified by the amount of elongation between stages.

One possible explanation to this is that the expression levels (or expression domain sizes) of wavefront genes (e.g. *fgf8*, Wnt ligands) are dynamic rather than constant. During the initial net increase phase in snake, chick, and mouse embryos, perhaps the expression levels/domain sizes are increasing. This would cause the wavefront movement (and therefore somite size) between stages to be less than the amount of elongation between stages. Conversely, during the net decrease phase (which is not a phase but is the norm in zebrafish), expression levels/domain sizes may be decreasing, which would cause wavefront movement between stages to be greater than the amount of elongation between stages. This explanation is consistent with the suggestion that axis elongation slows towards the end of somitogenesis, given that wavefront genes are involved in elongation (Dubrulle and Pourquié, 2004; Steventon et al., 2016) – and so a correlation between expression level/domain size changes and elongation rate would be expected. Indeed, it has been qualitatively shown (using *in situ* hybridisation) in mouse embryos that *wnt3a* and *fgf8* expression levels decrease towards the end of somitogenesis, at which point expression ceases (Cambray and Wilson, 2007; Young et al., 2009). The same has been also shown for *fgf8* expression levels in chick and zebrafish embryos (Olivera-Martinez et al., 2012; Zhang et al., 2018).

Interestingly, there is evidence in some animals that not all of the PSM is segmented into somites. In chick, PSM progenitors undergo cell death towards the end of somitogenesis, and so these cells do not contribute to somites. This process appears to be triggered by RA signalling, as this pathway is activated in the tailbud at the same stages, and inhibition of RA signalling leads to a reduction in the number of apoptotic cells (Olivera-Martinez et al., 2012). PSM cell death and tailbud RA signalling are not observed in mouse embryos, however, which suggests that this process may be unique to chick embryos – or at least to animals with reduced tails.

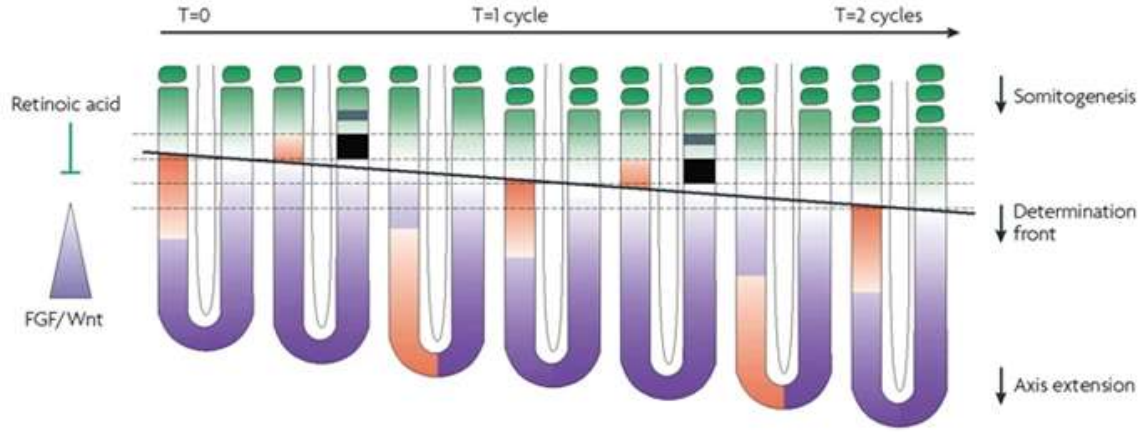


Figure 2: A clock and wavefront model of somitogenesis. The paraxial mesoderm (dorsal view) is shown at seven successive timepoints (left to right), which spans two clock cycles i.e. the formation of two new somites (top). Activity of clock signals are shown (orange) only on the left side of the presomitic mesoderm at each timepoint. These clock signals, which include Notch signalling, move in decreasing waves from posterior to anterior. Meanwhile, a determination front (or "wavefront") is moving from anterior to posterior over time. Here, the wavefront is depicted as established by opposing gradients of FGF/Wnt and Retinoic acid, although there is some debate over this. The distance travelled by the wavefront during one cycle specifies the length of the future somite, as only cells anterior to the wavefront are able to respond to clock signals to specify a new somite. This specification is marked by a stripe of *mesp2* expression (black, right side of relevant timepoints), which later becomes restricted to the anterior part (grey) of the future somite. The posterior movement of the wavefront is partly driven by axis extension ("axis elongation"). Figure is from Dequéant and Pourquié (2008).

The assumption of posterior growth

Early research on segmentation was carried out on *Drosophila melanogaster* embryos. Here, body segments are specified simultaneously rather than sequentially, without any elongation occurring. Later studies in *Tribolium castaneum* (flour beetles) and *Oncopeltus fasciatus* (milkweed bugs) showed a different form of segmentation; that of sequential segmentation, which coincides with elongation. These two types of segmentation are referred to as long germ-band segmentation and short germ-band segmentation, respectively. In the case of short germ-band segmentation, elongation was argued to be driven by a "posterior growth zone": an unsegmented region which, as it grows, provides the tissue required for new segments to form (Liu and Kaufman, 2005). This system of segmentation is, in general terms, similar to somitogenesis in vertebrates: sequential segmentation coinciding with/dependent on elongation. This similarity has led to authors using the term "posterior growth zone" when referring to the vertebrate tailbud/posterior PSM (Martin and Kimelman, 2009; Bénazéraf et al., 2010), and the assumption that axis elongation is driven by rapidly proliferating cells in this region in all vertebrates.

However, while this may be the case in some species, it is not the case in all species. Steventon et al., (2016) photo-labelled the posterior tip of the zebrafish embryo at the 15 somite-stage and found that the label, while elongating, did not increase in volume over the formation of the next 10 somites. They also counted the number of cells undergoing mitosis, using a cell-cycle reporter line, and found that the tailbud showed a *lower* proportion of proliferating cells than the rest of the axis, rather than a higher proportion. Similarly, it has been shown that zebrafish *emi1* mutants, which cease proliferation after the beginning of gastrulation, still undergo axis elongation and form the correct number of somites (Zhang et al., 2008; Riley et al., 2010). In fact, it seems that downregulation of proliferation is *required* for normal paraxial mesoderm specification and elongation (Bouldin et al., 2014). These results show that in zebrafish

at least, axis elongation is not driven by posterior growth. Steventon et al. (2016) also measured the volume of the unsegmented region at different stages in fixed embryos of lamprey, catsharks, zebrafish, and mice. They showed that this region simply decreases in volume (as it is segmented) over the course of somitogenesis in lamprey and zebrafish embryos. However, in catshark and mice embryos, the unsegmented region shows an initial net increase in volume (despite segmentation) before a later depletion phase. This suggests that axis elongation in catshark and mice embryos may be driven by growth, but not in lamprey and zebrafish. The authors hypothesised that posterior growth is related to embryo energy supply; embryos with a large yolk (or that are nourished by a placenta) undergo posterior growth, whereas embryos with a small yolk do not. Regardless of whether or not this is the case, it is clear that growth-driven axis elongation is not a conserved feature of vertebrate somitogenesis. Interestingly, recent work has also shown that many short germ-band invertebrates which were previously described as having a posterior growth zone do not show more growth in this region than in the rest of the embryo (Mayer et al., 2010; Auman et al., 2017). Regardless, there still exists a widespread assumption that in all vertebrates and short germ-band invertebrates, axis elongation is driven by posterior growth (Martin and Kimelman, 2009).

Neuromesodermal progenitors and axis elongation

The idea that vertebrate axis elongation is driven by posterior growth has been buoyed by the recent discovery of neuromesodermal progenitors (NMps). These are cells that, during somitogenesis, are located in the tailbud and give rise to the spinal cord, notochord and paraxial mesoderm (Steventon and Martinez Arias, 2017). Individual NMps co-express the early neural marker gene *sox2* and the early mesodermal marker gene *brachyury*, and have the potential to self-renew (Henrique et al., 2015; Cambray and Wilson, 2002). NMps have therefore been described as stem cells which drive axis elongation, by acting (at the population level) as a posterior growth zone (Martin and

Kimelman, 2012; Kimelman, 2016). However, there are problems with defining NMps, particularly when comparing between stages of development, and among different species. It is therefore important to clearly define NMps with respect to species and stage, to understand their contribution to axis elongation in different vertebrates.

In chicken and mouse embryos, gastrulation involves the ingression of epiblast cells through the anterior part of the primitive streak (just posterior to the node) to form the three primary germ layers: endoderm, mesoderm, and ectoderm. The primitive streak recedes posteriorly, and with it, the site of ingression – meaning that germ layers (and their tissue derivatives) form progressively, from anterior to posterior. When the streak has fully regressed, the cells in and around the node become incorporated into the newly formed tailbud. Primitive streak regression ends roughly halfway through somitogenesis in amniotes (Steventon and Martinez Arias, 2017), and so there is a significant overlap between the processes of gastrulation and somitogenesis. The NMps are those cells, initially in and around the node, which later become incorporated into the tailbud. NMps have been shown to have the capacity to self-renew and contribute to both spinal cord and paraxial mesoderm, along most of the axis, in chicken and mouse embryos (McGrew et al., 2008; Cambray and Wilson, 2002).

However, in fish and frog embryos, gastrulation does not involve gradual ingression along a primitive streak. Instead, cells involute at the blastopore, which rapidly decreases in size until blastopore closure. Throughout this, and for a period after this, all cells converge to the midline, which allows the embryonic tissues and structures (including the tailbud) to form. Tailbud NMps in fish and frog embryos have been shown to have the capacity to differentiate into both spinal cord and paraxial mesoderm (Martin and Kimelman, 2012; Davis and Kirschner, 2000), which has led to the assumption that these cells also self-renew and contribute to most of the axis in these species (Martin and Kimelman, 2012; Mongera et al., 2018). However, this is not the case. Experiments in which tailbud NMps were photolabelled show that these cells only contribute to the

posterior tip of the tail in frog and fish embryos (Davis and Kirschner, 2000; Attardi et al., 2018).

Importantly, it is not clear if *tailbud* NMps in chick or mouse embryos self-renew. While they have the potential to do so, as shown by their large axial contribution when grafted into the node region of younger embryos (McGrew et al., 2008; Cambray and Wilson, 2002), they appear to act as a depleting progenitor pool in the tailbud (Steventon and Martinez Arias, 2017).

In amniotes, because streak regression is a slow, gradual process that coincides with embryo growth, NMps at pre-tailbud stages contribute to most of the axis. However, in anamniotes, blastopore closure is rapid, and most of the axis is formed by convergent extension rather than by cell division and growth. So although there are regions of cells around the closing blastopore in which some cells will become neural and some will become mesodermal, an individual cell will undergo very few, if any, divisions over the course of blastopore closure (Attardi et al., 2018). This means very few individual cells will contribute to both neural and mesodermal tissues, and in terms of contributing to axis elongation, these cells only do so in that, as a group, they converge and extend – not because of any stem cell properties. In other words, while the population may contribute to a large portion of the axis, each cell will only contribute to a very small portion. It is therefore difficult to compare pre-tailbud NMps between anamniotes and amniotes, since it is such a brief, transient population in anamniotes. Comparing tailbud NMps between these groups is much more straightforward, as they seem to behave similarly – as a depleting population of bipotent progenitors. Even during streak regression in amniotes, it is not clear that NMps drive axis elongation by acting as a posterior growth zone. As it has not been shown that these posterior cells proliferate more than cells in other parts of the embryo, it could be that growth is uniform across the AP axis, and so NMps simply contribute to already growing tissues. Therefore, the

notion that tailbud NMps drive axis elongation may be the result of assuming that the behaviour of tailbud NMps is equivalent to that of pre-tailbud amniote NMps.

Possible mechanisms of paraxial mesoderm elongation

As axis elongation in species like zebrafish, frog, and lamprey is not driven by posterior growth, this raises the question: what is driving axis elongation in these species? To reiterate, axis elongation specifically refers to elongation of the AP axis during patterning and segmentation. In the case of the paraxial mesoderm, any elongation taking place in the somites will not contribute to this, as the somites are already segmented and separate from the unsegmented paraxial mesoderm (the PSM). Somite elongation will have no effect on wavefront movement and therefore will not affect somite number, and so it is only PSM elongation that is important for understanding the regulation and evolution of somite number. Several mechanisms (some of which involve posterior growth) have been proposed to explain PSM elongation in different vertebrates. As zebrafish represent a good model system to study axis elongation in the absence of growth, these proposed mechanisms will now be briefly discussed with their potential relevance to zebrafish PSM elongation.

Directional migration

Manning and Kimelman (2015) argued that PSM elongation in zebrafish is driven by anterior migration of posterior PSM cells. They transplanted labelled cells into host embryos (to achieve mosaic labelling), then explanted the tailbud of host embryos to live-image cell movements at high resolution. They showed that, in these explants, labelled PSM cells move anteriorly much more than they do posteriorly. They also showed that, when wild-type (WT) cells were transplanted into host embryos that had been injected with a morpholino (MO) against *tbx16* and *msgn1* (genes required for EMT of PSM progenitors), WT cells moved even further anteriorly than they did in WT host

explants. They therefore proposed a model in which all PSM cells are competing to migrate anteriorly. To test this idea, they analysed cell protrusions to see if cells preferentially form anterior protrusions. The authors found that, instead, protrusions were equally distributed in all directions. However, they found that anterior protrusions were more often followed by cell movement in the same direction than posterior protrusions were, so argued that this indicates a preferential anterior migration of cells – despite no preferential anterior protrusion formation.

There are several problems with this model. Firstly, defining movement as anterior or posterior requires a point of reference – something these authors do not identify. In an elongating tissue, cells will automatically displace along the axis. This presumably means that, relative to the posterior tip, all cells will be moving anteriorly, while relative to the anterior tip, all cells will be moving posteriorly. It is also difficult to rule out that, due to the tailbud being explanted, cells will move towards the cut (anteriorly) because of greatly reduced resistance from this direction. Additionally, their measurements appear to show that cells move medially and laterally as much as anteriorly. These issues make it difficult to argue that cells are preferentially migrating anteriorly. It is also not clear what substrate cells are migrating on. Other examples of migration during development, like neural crest migration and lateral line migration, involve small groups of cells migrating over non-migratory cells. There is no clear mechanism by which all cells within a tissue could migrate over one another – and how this could drive elongation of the whole tissue.

Random motility gradient and regulated tissue fluidity

In contrast to a model of anterior migration, other authors have proposed a mechanism of random motility driving axis elongation. Bénazéraf et al. (2010) first proposed this in chick embryos. Realising the need for a reference point, they measured both cell movement and ECM (fibronectin fibres) movement in the PSM, reasoning that

the ECM movement represents the global tissue elongation. They then subtracted the ECM movement from the PSM cell tracks to eliminate global tissue elongation from individual cell tracks, and found that cells appear to move randomly, showing no bias in directionality. They also found that there is a gradient of motility from posterior to anterior. In other words, all cells move randomly, but posterior cells move more than anterior cells. The authors argued that this random motility gradient could drive elongation when coupled with the addition of cells to the posterior PSM from the node. High motility in the posterior reduces local density, which provides space for node cells/NMps to move into, and this causes expansion of the posterior PSM, which drives axis elongation.

The requirement for a high level of cell addition makes this model not directly applicable to zebrafish, however a similar model has been proposed in zebrafish. Lawton et al. (2013) proposed that observed random movements of cells in the posterior PSM are important for tissue gene expression homogenization and bilateral symmetry, which are both necessary for proper elongation. They tracked cells and then subtracted the average movement of neighbouring cells from individual tracks, in order to remove global tissue elongation/movement from the individual cell tracks. They found that, while cell movement from the dorsal tailbud into the posterior PSM is highly ordered and directional, posterior PSM cells moved more randomly, and anterior PSM cells moved even more randomly (in contrast to the random motility gradient observed in chicks). These authors emphasised the importance of the ordered movement of dorsal tailbud cells into the posterior PSM, which they attributed to NMp addition. However, they carried out their analysis on embryos at the 10 somite-stage. Since it has been shown that tailbud NMps in zebrafish do not contribute to the PSM until much later stages (and even then, only in small numbers), this movement of cells into the PSM cannot be from NMps (Attardi et al., 2018). Rather, since cells during early somitogenesis stages are moving across the yolk into the tailbud, it is much more likely that this addition is

from these converging cells (Steventon et al., 2016). As such, it is difficult to know how much this model of PSM elongation applies to the rest of somitogenesis – or if it only applies to these early stages. However, it is theoretically possible that random motility in the PSM could still drive elongation without cell addition, by decreasing the density of the PSM and driving expansion. A decrease in density over time has not been shown in the zebrafish PSM, and so this remains unknown.

A fluid-to-solid transition

Other models have focussed less on cell movements and more on tissue properties. Mongera et al. (2018) proposed a model to explain zebrafish PSM elongation in which differential stresses along the AP axis cause a "jamming transition", which constrains tissue expansion into the AP axis, causing elongation. They injected magnetic droplets into different points along the PSM and measured droplet deformations as a readout of local mechanical stresses. They found a gradient of mechanical stress along the AP axis, increasing from posterior to anterior. In other words, the posterior PSM is more fluid-like, whereas the anterior PSM is more solid-like. They found that this AP stress gradient was not present in N-cadherin mutants, and so suggested that this gradient is dependent on N-cadherin in some way – although they did not show (or argue for) an AP gradient of N-cadherin itself. In keeping with a more fluid-like posterior, cells in this region underwent far more mixing than those in the anterior. Importantly, they did not consider this cell mixing to be itself a form of intercalation that would drive elongation. They looked at the alignment of neighbour exchanges and found that these did not orient along the AP axis, which they expected would be the case if cell intercalations were driving convergent extension. The authors proposed that this high level of cell mixing was instead required to keep the posterior PSM in a fluid-like state. They then used 2D mathematical modelling to show that, a stress gradient was essential for ensuring that tissue expansion was unidirectional (along the AP axis) rather than isotropic (in all axes).

The fluid posterior allowed expansion while the solid anterior focussed this expansion along the AP axis. However, their model involved a high degree of cell addition from NMps. Without this cell addition, it is likely that there would be no tissue expansion, and therefore nothing for a stress gradient to act upon to drive elongation. Since we know that NMp addition is very minor in zebrafish, this model cannot provide a mechanistic explanation of paraxial mesoderm elongation. However, it does provide important insights about the physical properties of the tissue.

Convergent extension and multi-tissue interactions

Although Mongera et al. (2018) argue against convergent extension, it is possible that their methods of testing for this were not appropriate. They measured the angles of cell junctions before rearrangement relative to the AP axis. While this approach may be appropriate in the context of a 2D epithelial sheet, it is not clear how appropriate it is in the context of a 3D mesenchymal tissue. As such, it is difficult to know if the lack of a clear trend is due to the absence of convergent extension, or if it is due to the complexities of measuring cell junction angles in a 3D tissue. Keller et al. (2000) argue that convergent extension may be an active process driven by cell intercalation, or it may be a passive process driven by forces generated in other tissues/regions.

Xiong et al. (2020) showed in the chick embryo, an example of this kind of passive convergent extension, driven by multi-tissue interactions. Building on previous work, they suggested that the random motility in the posterior PSM could be causing compression of the notochord and neural tube. They confirmed that the posterior PSM causes compression of the notochord and neural tube: when they ablated this region, or replaced it with a gel, the posterior notochord and neural tube were wider. In a separate experiment, they replaced the posterior notochord with a gel, and measured the deformation of this gel over time. They found that the gel thinned and lengthened over time, confirming that the posterior notochord undergoes passive convergent extension

due to compression from the posterior PSM. The authors also found that the posterior movement of the node/progenitor domain was reduced in the first set of experiments, suggesting that convergent extension of the axial tissues pushes the node posteriorly. They therefore proposed a model involving a mechanical feedback loop, whereby the posterior PSM drives convergent extension of posterior axial tissues, which pushes the node posteriorly, which contributes to cell addition to the posterior PSM. This further drives the expansion of posterior PSM and compression of axial tissues, and the process continues in a positive feedback loop. Again, this model relies heavily on high cell addition of NMps, and so it does not apply directly to zebrafish. However it is possible that similar multi-tissue compressive interactions are also involved in zebrafish axis elongation. Mclaren and Steventon (2021) ablated the posterior notochord in zebrafish and found no effect on overall tail elongation, suggesting that the notochord does not play a role in paraxial mesoderm elongation. However, the role of the neural tube and other tissues has not been tested.

A long-term, multi-scale, three-dimensional approach is required

It is clear that, despite several proposed models, there is no clear, consistent mechanistic explanation for paraxial mesoderm elongation in zebrafish – and other animals where there is little to no posterior growth. Most existing models rely heavily on a high degree of cell addition. While these models may explain paraxial mesoderm elongation during *early* somitogenesis stages when cells are still moving over the yolk and into the tailbud, they cannot explain elongation during the rest of somitogenesis. This problem is intensified by the fact that most analyses only take place over the formation of a few somites. As such, it is not clear if observed trends persist for the entirety of somitogenesis. False assumptions about levels of, and sources of, cell addition

are likely due to analyses focussing on a specific scale of observation (either individual cell movements, or tissue physical properties). Several current explanations also abstract the paraxial mesoderm to a 2D tissue, either in the analysis of cell movements, or in computer simulations, or both. This makes it difficult to know how well the conclusions apply to the real tissue. Particularly: are observed random movements truly random, or is this an artefact of ignoring one of three axes of movement? Previous work in the chick has involved multi-scale 3D measurements to show how different tissues elongate (Bénazéraf et al., 2017), however this work was also restricted to a few somite stages, and so it is not clear if these patterns can be generalised across all stages of axis elongation.

To answer the question of how the paraxial mesoderm elongates in zebrafish, a long-term, multi-scale, three-dimensional approach is required. Although Steventon et al. (2016) carried out work of this nature in zebrafish, they took the tailbud/unsegmented region as a whole, and so did not take the specific measurements of the paraxial mesoderm required to answer this question.

In this thesis, I show – through long-term 3D morphometrics, cell-tracking, 3D modelling, and molecular perturbations – that the zebrafish paraxial mesoderm elongates through some form of convergent extension in the posterior PSM. This convergent extension, however, is not driven by local cell intercalation, but through large "convergent flows" which result in random intercalations. An increase in tissue density over time is also observed, which I suggest is driving convergent flows through a novel form of convergent extension: "compression-extension". The source of this compression is not clear, but my results suggest that it could be a combination of extrinsic compression by the neural tube, and density regulation by TGFβ signalling.

I also applied similar 3D morphometrics to other species, in order to compare paraxial mesoderm elongation mechanisms among vertebrates. The recent emergence of cichlids as potential model organisms in evo-devo allows an interesting test of the

hypothesis that posterior growth is associated with larger yolk sizes (Steventon et al., 2016). Therefore, I carried out morphometric analyses on two cichlid species with different yolk sizes. The results show that, in both species, the paraxial mesoderm does not undergo volumetric growth, and likely elongates by some form of convergent extension. I also applied a similar 3D morphometrics approach to chicken and catshark embryos: two species that do undergo volumetric growth of the paraxial mesoderm during somitogenesis. These results show some amount of tissue convergence in both species, suggesting that, rather than compression-/convergent extension and volumetric growth being two distinct "modes" of elongation, these two processes may combine to drive elongation in these species. Convergent extension, in its most general sense, may therefore be a conserved feature of paraxial mesoderm elongation.

METHODS

Animal husbandry

Zebrafish embryos (*Danio rerio*) were raised in standard E3 media at 28°C. Wild type lines used were Tüpfel Long Fin (TL), AB, and AB/TL. The following transgenic lines were also used: *tbx16*::GFP (Wells, Nornes and Lardelli, 2011), *h2a*::mCherry, and *h2b*::GFP. Embryos staged according to Kimmel et al. (1995).

Cichlid embryos (*Astatotilapia calliptera* and *Rhamphochromis chilingali*) were raised in tap water with methylene blue (1 in 1000) at 25°C. Chicken embryos (*Gallus gallus*) were raised at 38°C and initially staged according to Hamburger and Hamilton (1951), although were later staged more precisely by somite number. Catshark embryos (*Scyliorhinus canicula*) were raised in natural seawater (~ 16°C) and initially staged according to Ballard, Mellinger and Lechenault (1993), but were also later staged more precisely by somite number.

Pharmacological treatments

Pharmacological inhibitors were dissolved in dimethyl sulfoxide (DMSO) to produce a stock, which was kept at -20°C. Stocks were diluted in E3 media to appropriate working concentrations. Embryos were manually dechorionated and treated in 6-well plates at 28°C. For SB-505124 ("SB50") treatments, a stock concentration of 100 mM was produced, and diluted to produce a working concentration of 50 µM.

Microinjection

mRNA (KikGR, DEP+) was recovered from *E. coli* plasmid stocks using an Invitrogen mMESSAGE mMACHINE Kit for *in vitro* transcription and Lithium Chloride precipitation. mRNA was then diluted to 100-250 ng/µl in nuclease-free water, with phenol red added (0.05% weight/volume) to help with visibility during

microinjection. Embryos were microinjected (using pulled capillary needles) at the 1-cell stage.

in situ hybridization chain reaction (HCR)

Embryos were fixed in 4% PFA in PBS (without calcium and magnesium) overnight at 4°C, then dehydrated in methanol and stored at -20°C. HCR was then performed using standard zebrafish protocol (Choi et al., 2018) and nuclei were stained with DAPI (1 in 1000, in 5x SSCT). The dehydration step was omitted when also staining membranes with phalloidin. Catshark embryos were also cleared in Sca*l*eCUBIC-1 solution (Susaki et al., 2015) for 24 hours after nuclear staining, and then imaged in this solution.

Immunohistochemistry

Zebrafish embryos were manually dechorionated and fixed in 4% PFA in PBS (without calcium and magnesium) overnight at 4°C. They were then washed in MABT and blocked in blocking solution (5% serum and 2% Roche block in MABT) for 1 hour at RT. Primary antibodies (PH3) were diluted in block solution at 1 in 500 and embryos were kept in this solution overnight at 4°C. The next day, embryos were washed in MABT. Secondary antibodies were diluted in block solution at 1 in 1000 and embryos were kept in this solution overnight at 4°C. The next day, embryos were washed in MABT and stained with DAPI (1 in 1000 in MABT).

Imaging & Analysis

Mounting

Fixed embryos were either mounted whole in LMP agarose (1% in E3 media) which was then covered in PBS, or the tail was cut off and mounted in 80% glycerol.

Live zebrafish embryos were anaesthetised with tricaine and mounted in LMP agarose (1% in E3 media) which was then covered in E3 media. The posterior part of the embryo was freed from agarose by cutting away the agarose with a microinjection needle, as described by Hirsinger and Steventon, (2017), which allowed normal development of the trunk and tail.

Widefield

Live widefield imaging was performed on a Nikon Eclipse Ti, using a 10x (air) objective and a heated chamber (heated to 28°C).

Confocal

All confocal imaging was performed on a Zeiss LSM 700 (inverted) or a Leica SP8 (inverted).

Fixed imaging was performed using 10x (air), 20x (air), or 40x (oil) objectives depending on the resolution required for analysis and on embryo size/mounting method.

Live confocal imaging was performed using a 20x (air) objective and a heated chamber (heated to 28°C). Photolabelling was performed using localised, continuous UV laser scanning at 15% laser power for 60 seconds.

Two-photon

Live two-photon imaging was performed on a TriM Scope II Upright 2-photon scanning fluorescence microscope equipped with a tuneable near-infrared laser and Inspector Pro software. Imagine was performed using a 25x (water) objective and a heated chamber (heated to 28°C). Laser ablations were performed using localised laser scanning for the duration of one z-stack at approximately 80% laser power.

Image analysis

Image analysis was performed using Imaris and Fiji (ImageJ). See relevant Results chapters for specific analysis methods (e.g. 1.1 Morphometric methods and 2.1 Tracking methods). Statistical analysis was performed in R Studio.

RESULTS

1 "Compression-extension" drives zebrafish paraxial mesoderm elongation

1.1 Morphometric methods

Presomitic mesoderm marker gene expression over time

In order to understand the tissue dynamics involved in zebrafish paraxial mesoderm elongation, a long-term, three-dimensional approach is required. To obtain 3D measurements of this tissue over time, I used *in situ* hybridization chain reaction (HCR) to stain for PSM marker genes in embryos fixed at a range of somitogenesis stages. Embryos were stained for *msgn1* and *tbx6*, which are expressed in the posterior and anterior PSM, respectively, and nuclei were stained with DAPI. Figure 3 shows the expression patterns of these genes from the 10 somite-stage to the 32 somite-stage (approximately the end of somitogenesis). For all subsequent measurements/analyses, only embryos from the 16 somite-stage onwards were used. There were two reasons for this. Firstly, this was done to ensure that cell movement over the yolk and contribution to the paraxial mesoderm was no longer happening. This allowed me to treat the paraxial mesoderm as a "closed system" (aside from NMp contribution) and exclude this gastrulation process from the mechanisms of paraxial mesoderm elongation. Although the exact stage at which this process stops is not known, previous work suggests it is between the 12 to 17 somite-stages (Steventon et al., 2016). Secondly, as the tailbud is well-formed at these stages, it can be cut off and imaged separately. This allowed the paraxial mesoderm to be imaged on higher resolution objectives. which was necessary to acquire accurate cell number counts (see below).

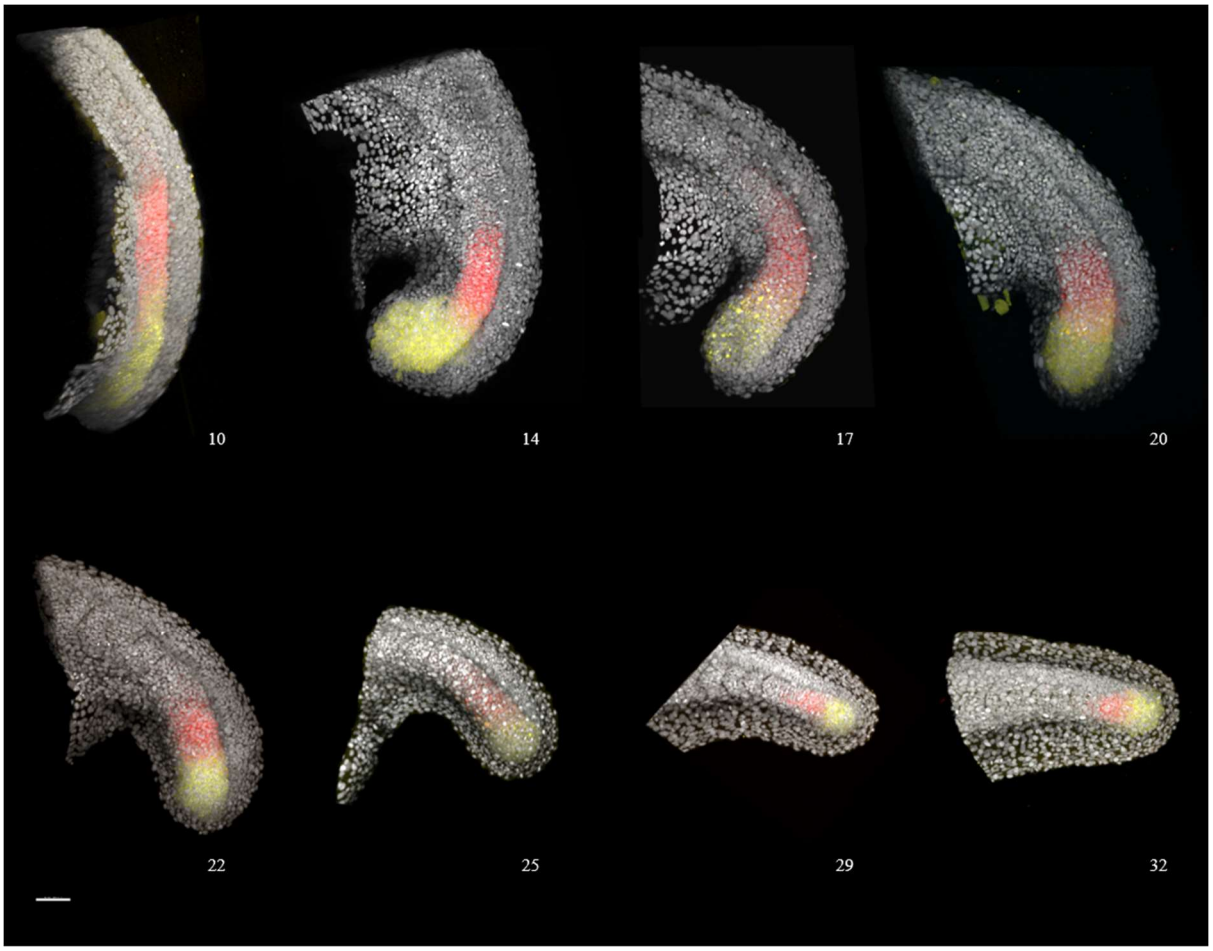


Figure 3: Zebrafish presomitic mesoderm (PSM) marker gene expression patterns. *in situ* hybridization chain reaction (HCR) was used to stain for *msgn1* (yellow) and *tbx6* (red), with nuclei stained with DAPI (grey). Each image is a maximum projection of a confocal stack. Embryos shown range in stage (white number, bottom right of each image) from the 10 somite-stage to the 32 somite-stage. The anteroposterior axis is uncurling during somitogenesis, so in these images it gradually changes from top-bottom to left-right. Scale bar: 50 µm.

3D tissue reconstructions: surfaces

Using Imaris image analysis software, I then created 3D reconstructions (or "surfaces") of the PSM and somites at each stage. This was done manually, by drawing 2D contours around the tissue, at regular z-intervals, for each image. For the PSM, I used *msgn1* expression to identify the posterior boundary of the tissue. For the anterior boundary, I used the nascent somite posterior boundary (visible in the DAPI channel) rather than *tbx6* expression, as cells stop expressing *tbx6* before they enter the nascent somite. For the somites, no marker genes were required as somite boundaries were visible in the DAPI channel. As somite boundary formation is gradual rather than instantaneous, I only considered a somite to be formed if the boundary was clearly visible across all z-slices. Otherwise, this almost-formed somite would be considered part of the PSM. As imaging all the way through the embryo was not always possible, and DAPI intensity decreased with depth of imaging, contours were only drawn for one lateral half of the paraxial mesoderm. In other words, in each image, surfaces were only made of the left (or right) half of the PSM, and the somites on the left (or right) side of the embryo. Whether this was the left or the right side depended on orientation of the embryo during imaging. The notochord was used to determine when the midline of the embryo had been reached, and therefore the point at which to draw the last contour. Once all 2D contours were drawn, a 3D surface reconstruction was generated automatically by Imaris. An automatic surface was also made of the *msgn1*-expressing region of the PSM to allow later length measurements (see below) to follow the curve of the tissue. Figure 4 shows the process of manual surface reconstruction of the PSM and a nascent somite, and an example of a *msgn1* automatic surface. Imaris surfaces provided volume measurements for paraxial mesoderm tissues, as well as allowing dimension measurements to be drawn (see below).

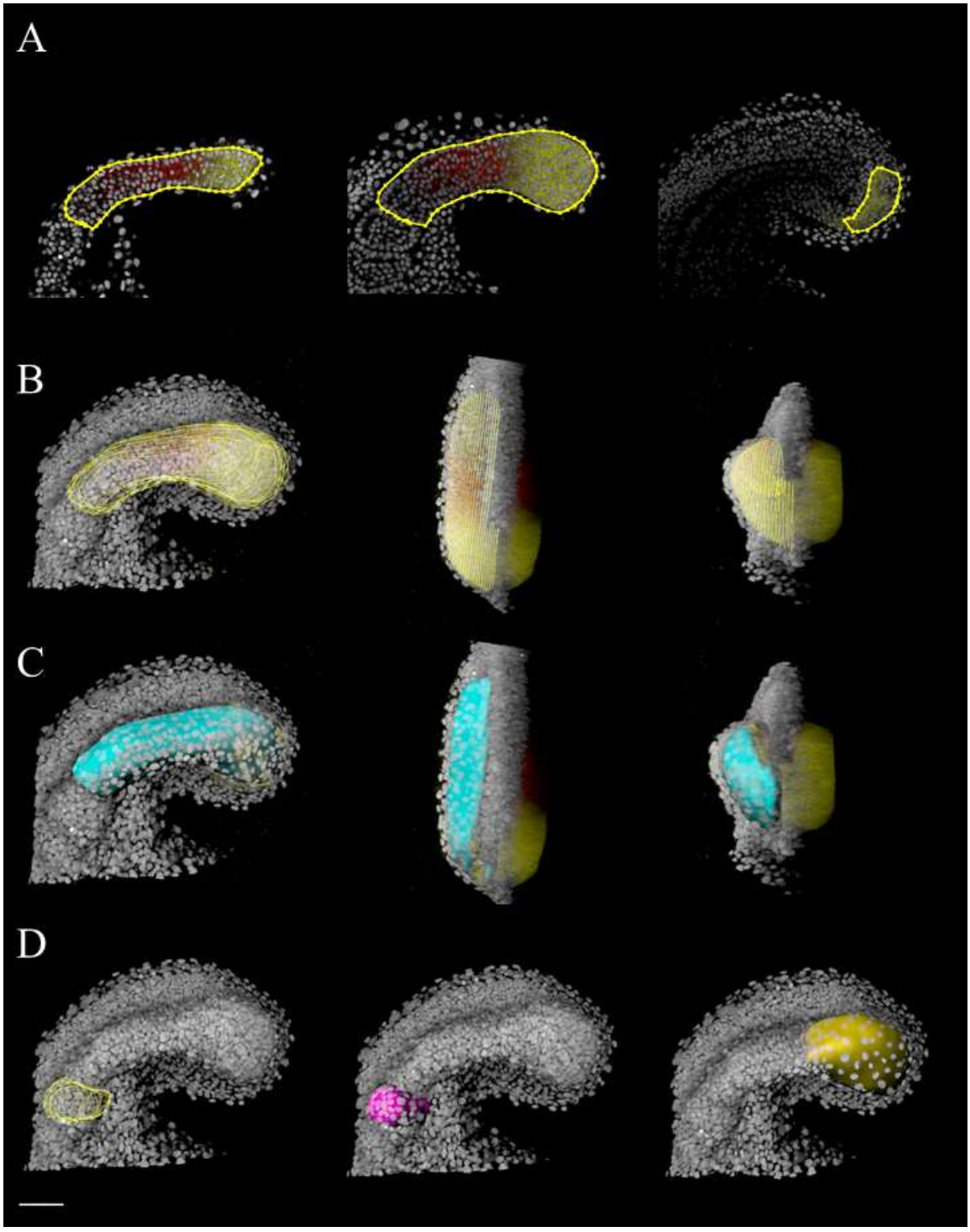


Figure 4: Tissue reconstructions of the PSM and a nascent somite. (A) 2D contours (yellow outlines) manually drawn around the PSM at regular z-slices, up to the midline. (B) Maximum projections showing all PSM contours, from lateral, dorsal, and posterior views (left to right). (C) Same images as in B but showing PSM surface (cyan). (D) Contours and surface (pink) of the nascent somite (left, middle), and automatic *msgn1* surface (gold, right). All images are of the same 23 somite-stage embryo. Scale bar: 50 µm.

3D cell positions: spots

To obtain information about cell numbers over time, I created "spots" in Imaris based on the DAPI channel. This was done automatically in Imaris, with a given estimated cell diameter and minimum fluorescence intensity threshold. The nuclei (DAPI signal) from each PSM/somite surface were isolated (Figure 5A) and spots were generated from these nuclei (Figure 5B). The estimated cell diameter used was 4 µm, and the minimum fluorescence intensity threshold was set to 0. These parameters gave the most accurate cell number estimates. This accuracy was validated by "masking" spots from the nuclei (to create a small black hole in each nucleus), then creating a new channel (given a different colour) with only DAPI signal inside the spots (to create a coloured dot in each nucleus). This method (Figure 5C) allowed me to check individual z-slices for the presence of one coloured dot in each nucleus. Estimated cell diameters below 4 µm led to multiple dots per nucleus (i.e. overestimation of cell number), while estimate cell diameters above 4 µm led to many nuclei with no dots (i.e. underestimation of cell number), as did setting any minimum fluorescence intensity threshold. Imaris spots provided cell numbers and 3D positions for paraxial mesoderm tissues over time.

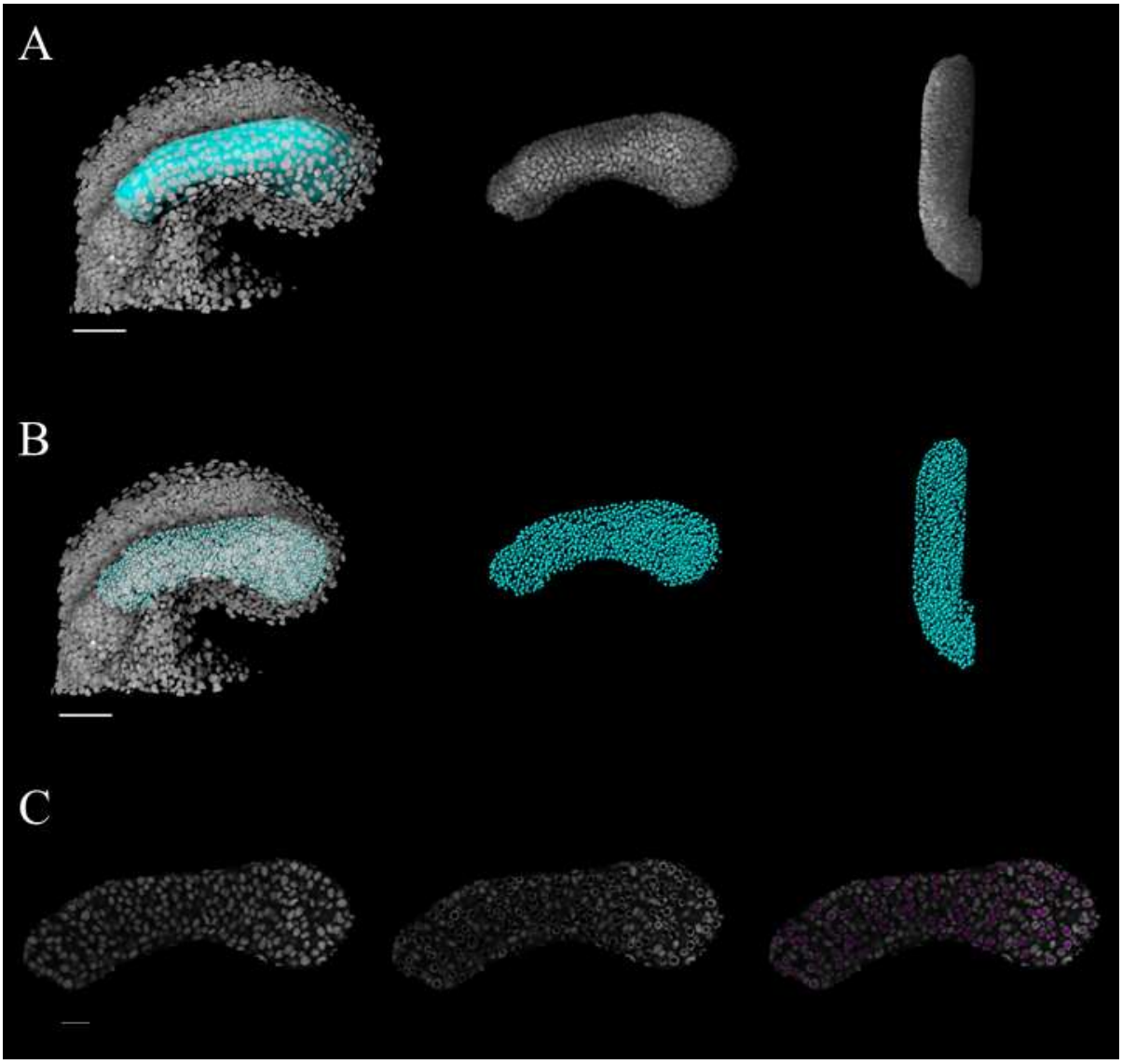


Figure 5: Cell positions of the presomitic mesoderm (PSM). (A) DAPI signal was isolated from the surface (left) to create channel of only nuclei in this tissue, shown in lateral and dorsal views (middle, right). (B) Imaris spots generated from this channel, providing cell positions and numbers. Images in A, B are maximum projections. Scale bar in A, B: 50 µm. (C) Single z-slice showing PSM nuclei (left) with "masked" spot fluorescence (middle) which was then used to create a new channel (right, pink). This enabled visualisation of spots in individual slices, which enabled validation. Scale bar in C: 20 µm. Spot creation and validation is only shown here for the PSM, but the same method was applied to the somites.

3D axes measurements

As well as volumes and cell numbers, specific dimensions of each axis were also measured. This was done manually in Imaris using "measurement points" and the surfaces previously created. To measure length (anteroposterior (AP) axis) of the PSM, both the manual PSM surface and the automatic *msgn1* surface were used. The measurement was taken through the middle of the tissue, following the tissue curve by taking two measurements: one from the posterior medial face of the *msgn1* surface to the anterior face, and one from here to the anterior face of the PSM surface (Figure 6). This length measurement thus considered the tissue curve, as well as the fact that the surface is only one lateral half of the true tissue. For both width (mediolateral (ML) axis) and height (dorsoventral (DV) axis) measurements, separate measurements were taken for the posterior and the anterior PSM, and were taken roughly at the halfway point (along the AP axis) of each region. While the ML axis keeps the same orientation between posterior and anterior PSM (Figure 6A), the DV axis does not (Figure 6B), in that the tail is curled ventrally. The two separate height measurements take this ventral curl into account. Somite measurements for each axis were also taken through the middle of the tissue. Together, these 3D axes measurements provided information about tissue shape change over time.

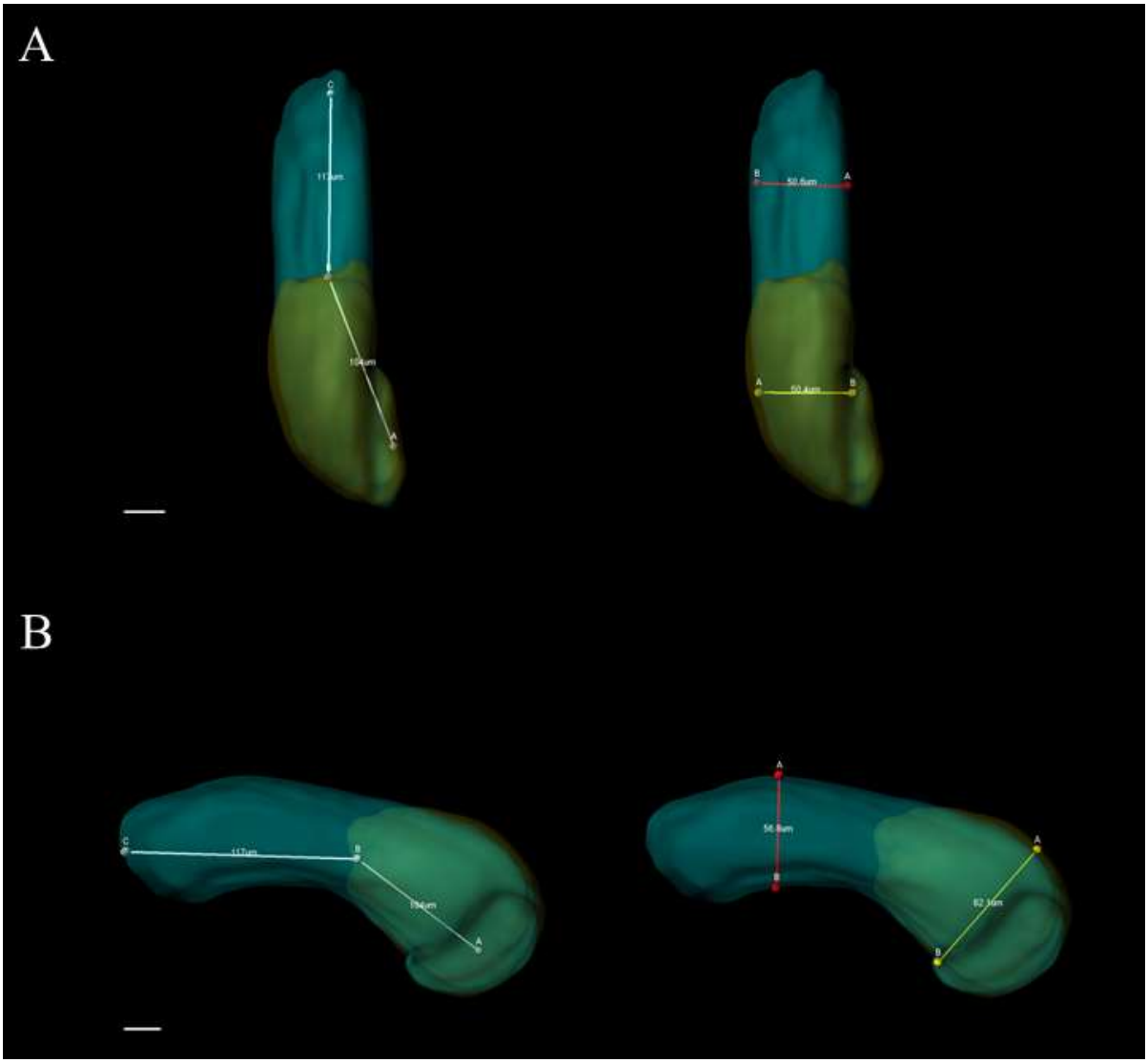


Figure 6: Dimension measurements of the presomitic mesoderm. (A) Dorsal view (anterior top) showing the length (AP axis) measurement of the PSM (left), and the width (ML axis) measurements (right) of the posterior and anterior PSM. (B) Lateral view (anterior left) showing the same length measurement as above (left), and the height (DV axis) measurements (right) of the posterior and anterior PSM. Scale bars: 20 µm. Measurements are only shown here for the PSM, but the same measurements were also taken of the somites.

1.2 Direct vs indirect measurements

Direct paraxial mesoderm measurements

The most intuitive way to measure the paraxial mesoderm over time is to do so directly, by taking measurements (as described above) of the PSM and the somites at multiple stages over time (Figure 7). As mentioned above, these measurements were taken from the 16th somite onwards. The results show that the length of the paraxial mesoderm increases substantially over time, while the number of cells decreases overall (despite a slight initial increase), and the volume of the total tissue decreases steadily over time. This decrease in cell number and volume is likely due to somite cells migrating out of the somites to contribute to other tissues (Lee et al., 2013). As previously discussed, processes occurring in the already segmented tissues (somites) are not relevant to understanding axis elongation. Therefore, from these measurements, it is impossible to determine how much growth/proliferation may be happening in the unsegmented paraxial mesoderm (PSM), and to what degree this contributes to elongation. However, simply measuring the PSM at different stages will also be confounded by the process of somitogenesis – as elongation, growth, and proliferation will be outweighed by the loss of length, volume and cells as the tissue is depleted to form somites. A different method is therefore required: one which takes measurements of the PSM and the somites it gives rise to; but does not include post-formation processes in those somites.

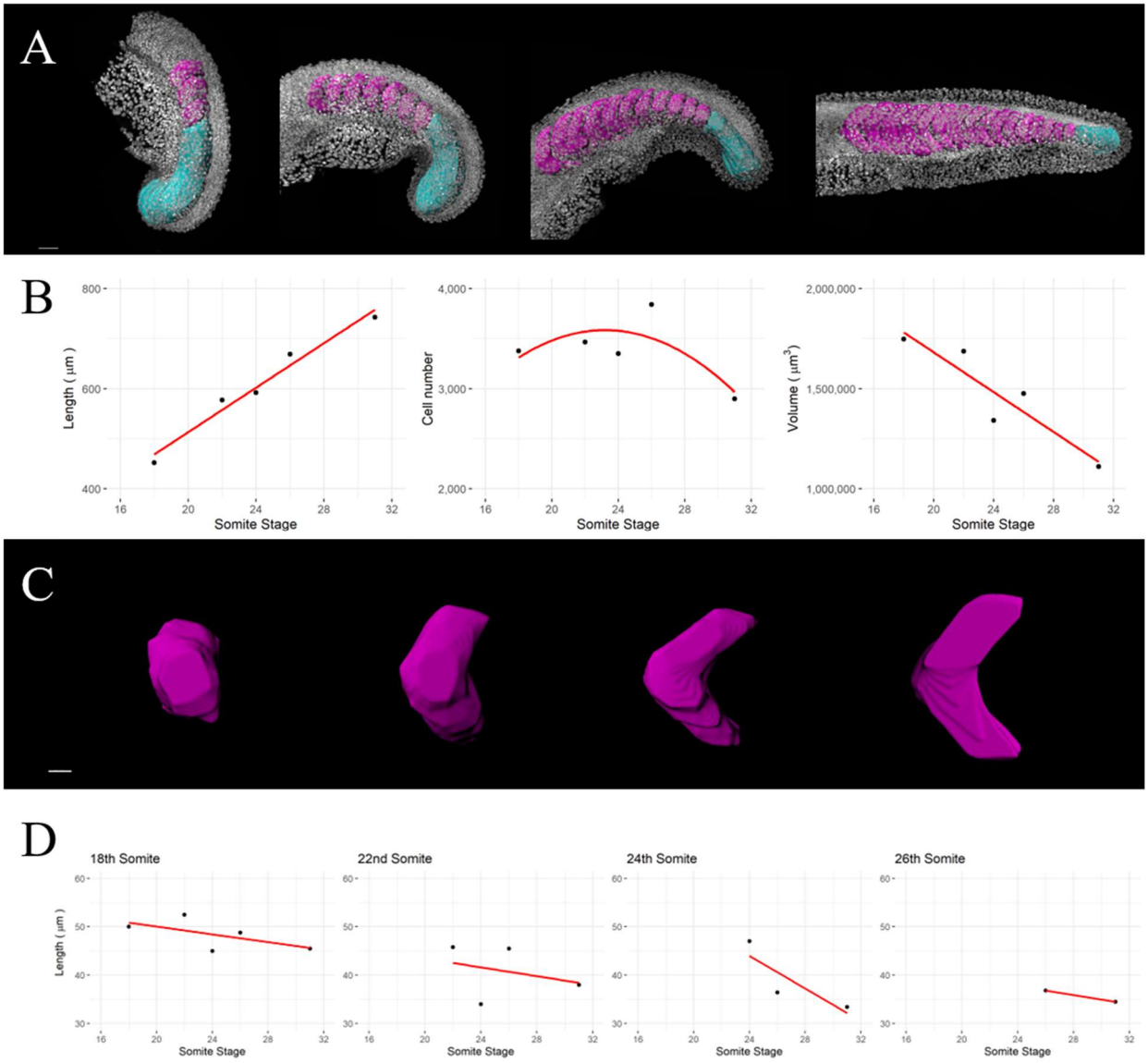


Figure 7: Direct paraxial mesoderm measurements. (A) The paraxial mesoderm from the 16th somite onwards at different stages of development (18, 22, 26, and 31 somite-stage), with somite surfaces (pink), PSM surfaces (cyan), and DAPI (grey) shown. Scale bar: 50 µm. (B) The length, cell number, and volume of the paraxial mesoderm (16th somite onwards) over time. The length increases, but cell number and volume decrease, most likely due to somite cells migrating away to contribute to other tissues. (C) An individual somite (18th somite) at different stages of development (18, 22, 26, and 31 somite-stage). Scale bar: 20 µm. (D) The lengths of different somites over time. All somites show a length decrease, indicating that paraxial mesoderm elongation is solely driven by PSM elongation, in spite of somite shortening. n = 5 embryos.

Indirect paraxial mesoderm calculations

To measure only *relevant* axis elongation (i.e. excluding the confounding processes of somite morphogenesis and post-formation cell loss) I only included measurements of the PSM and the nascent somite at each stage. In other words, to look at changes in a particular measurement (e.g. length) from the 16 somite-stage onwards, rather than simply taking the direct measurement of all paraxial mesoderm from the 16th somite onwards at each stage, I took the length of the PSM at each stage, and added on the length of each nascent somite immediately after its formation to calculate an "indirect" paraxial mesoderm measurement. This is shown as a schematic in Figure 8. In this way, the somites that the PSM gives rise to are included, but the post-formation processes of shortening and cell loss are excluded. The example given here is for length, but the same approach was taken for cell number and tissue volume.

Rather than using specific measurements, trendline equations were calculated from the data – in order to obtain average values for each stage that had more than one data point, and to provide values for any stages that had no data point. Indirect paraxial mesoderm values were then calculated from the PSM equation and the nascent somite equation. The exact equations are given in the next section, but in general terms, the indirect paraxial mesoderm function $f(x)$ can be expressed in terms of somite stage x, the PSM trendline $p(x)$, and the nascent somite trendline $s(n)$, which sums from $n = 17$ to the given somite stage x. The equation is as follows:

$$\begin{cases} x = 16, & f(x) = p(x) \\ x > 16, & f(x) = p(x) + \sum_{n=17}^{x} s(n) \end{cases}$$

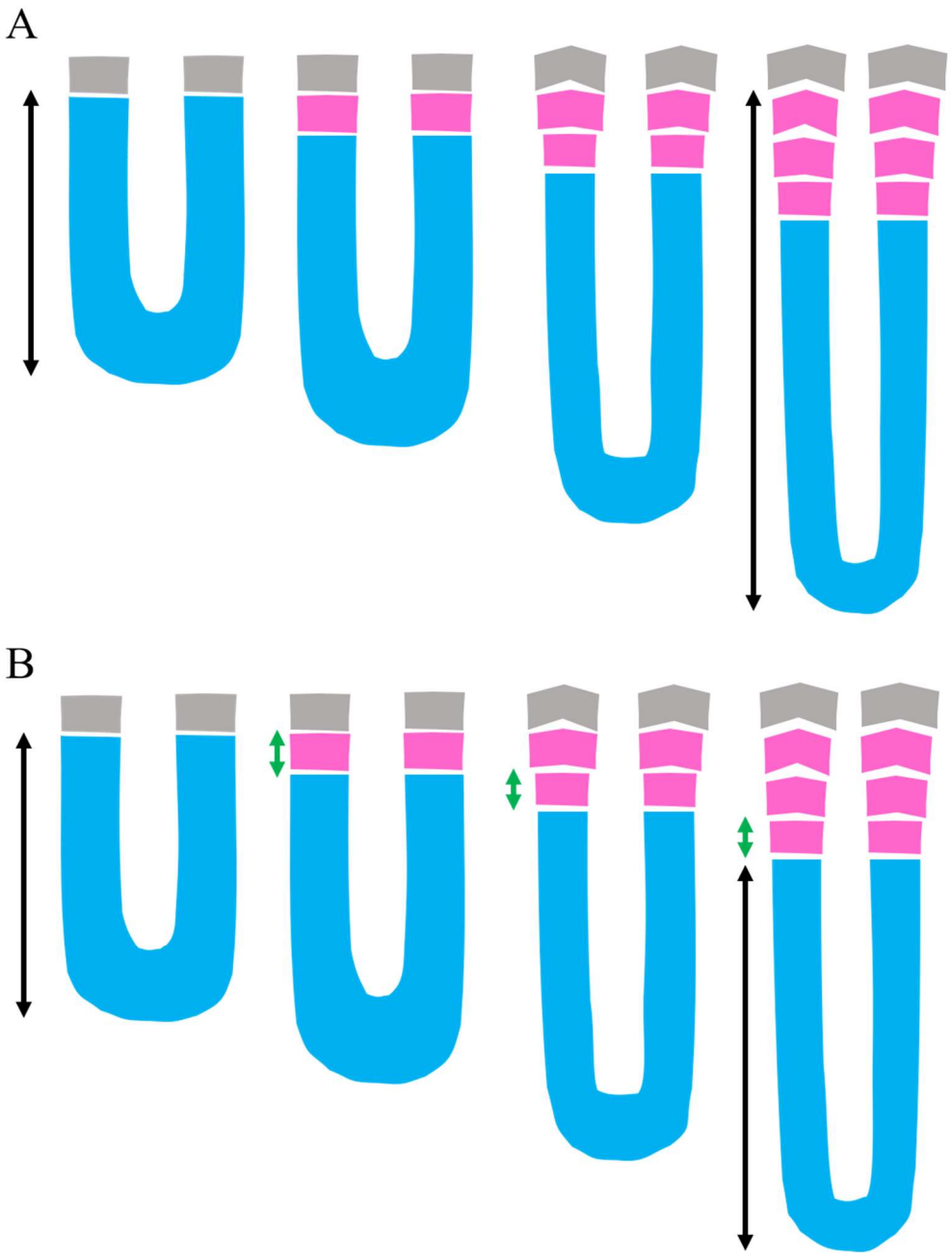


Figure 8: Direct vs indirect paraxial mesoderm measurements. (A) Direct paraxial mesoderm length measurements. To calculate the paraxial mesoderm length change between an initial stage (left) and a final stage (right), the length from a given somite (grey) (e.g. 16th somite) onwards is directly measured at both stages (black arrows). However, this length change will be confounded by post-formation processes occurring in the somites, which are not relevant to axis elongation mechanisms. (B) Indirect paraxial mesoderm length measurements. The lengths of the PSM are measured (black arrows) at the initial stage (left) and final stage (right). The length at the final stage is added to the lengths of each nascent somite at its time of formation (green arrows) to give an indirect paraxial mesoderm measurement. This measurement method excludes post-formation processes occurring in the somites, to provide a more informative measure of paraxial mesoderm elongation. The example shown here is for length measurements, but the same method was used for tissue volume and cell number measurements.

1.3 Elongation without growth

Measurements of the zebrafish PSM over time (Figure 9, left column) showed that this tissue is depleted over time, supporting the length measurement results of Gomez et al. (2008). My results show that the length decreases linearly, while the cell number and tissue volume decrease exponentially. In other words, while the length reduction is constant, the cell number and tissue volume reductions slow over time. The sizes of the nascent somites also follow these patterns (Figure 9, middle column). Each somite, at the time of its formation, is smaller than the previous somite (at the time of its formation). However, while this length trend is linear, the cell number and tissue volume trends are exponential. Possible trendline equations for each set of measurements were calculated in R, and AIC was used to determine the best statistical model (linear, exponential, quadratic, etc.). Figure 9 shows the trendline and equation for each set of measurements. These equations were then used to calculate the indirect paraxial mesoderm values for length, cell number, and tissue volume, as explained in the previous section. These values, shown in Figure 9 (right column, violet lines), therefore represent the unsegmented paraxial mesoderm at the 16 somite-stage onwards, including the somites it gives rise to (somites 17 to 33), but excluding post-formation processes in the somites.

As expected from the previous direct measurements, the length of the paraxial mesoderm increases considerably ($\sim$ 100%) throughout these stages of somitogenesis (Figure 9A). Interestingly this lengthening slows over time, which may be related to the downregulation of wavefront genes. Taken together with the PSM-only measurements, this result shows that the PSM is elongating, but that the length of each nascent somite specified is greater than the amount of elongation between somite stages, and so PSM length shows a gradual net decrease.

As an interesting aside, the correlation between somite stage and PSM length is so strong ($r^2 = 0.97$, $p < 0.001$) that PSM length can be used to estimate the stage of a

wild-type embryo with very high accuracy. This was particularly useful in other analyses/experiments, when embryo stage was difficult to determine (e.g. due to imaging constraints), or when validation was required.

The cell number results (Figure 9B) prove that this elongation is not driven by proliferation. In fact, paraxial mesoderm cell number initially stays constant (when elongation is fastest) before showing a very slight increase (~ 10%) at later stages. This late, slight increase in cell number is likely due, in part, to NMp addition – given that NMps only contribute to the paraxial mesoderm at these same late stages (Attardi et al., 2018). Again, taken with the PSM measurements, this shows that the PSM does show an increase in cell number, but that the number of cells lost to the somites is greater than this increase between somite stages, and so PSM cell number shows a net decrease.

Interestingly, paraxial mesoderm tissue volumes (Figure 9C) do not show the same trends as those of cell numbers. While both PSM and nascent somite volumes decrease exponentially (like cell numbers), this decrease is steeper. As a result, the paraxial mesoderm volume does not remain constant with a late increase, but instead *decreases* exponentially.

Taken together, these results prove that the paraxial mesoderm elongates considerably over time. However, this elongation is not the result of growth. While the number of cells does increase slightly, this does not cause tissue expansion – as the tissue *decreases* in volume over time. Therefore, the paraxial mesoderm elongates not only in the absence of growth, but despite tissue shrinkage.

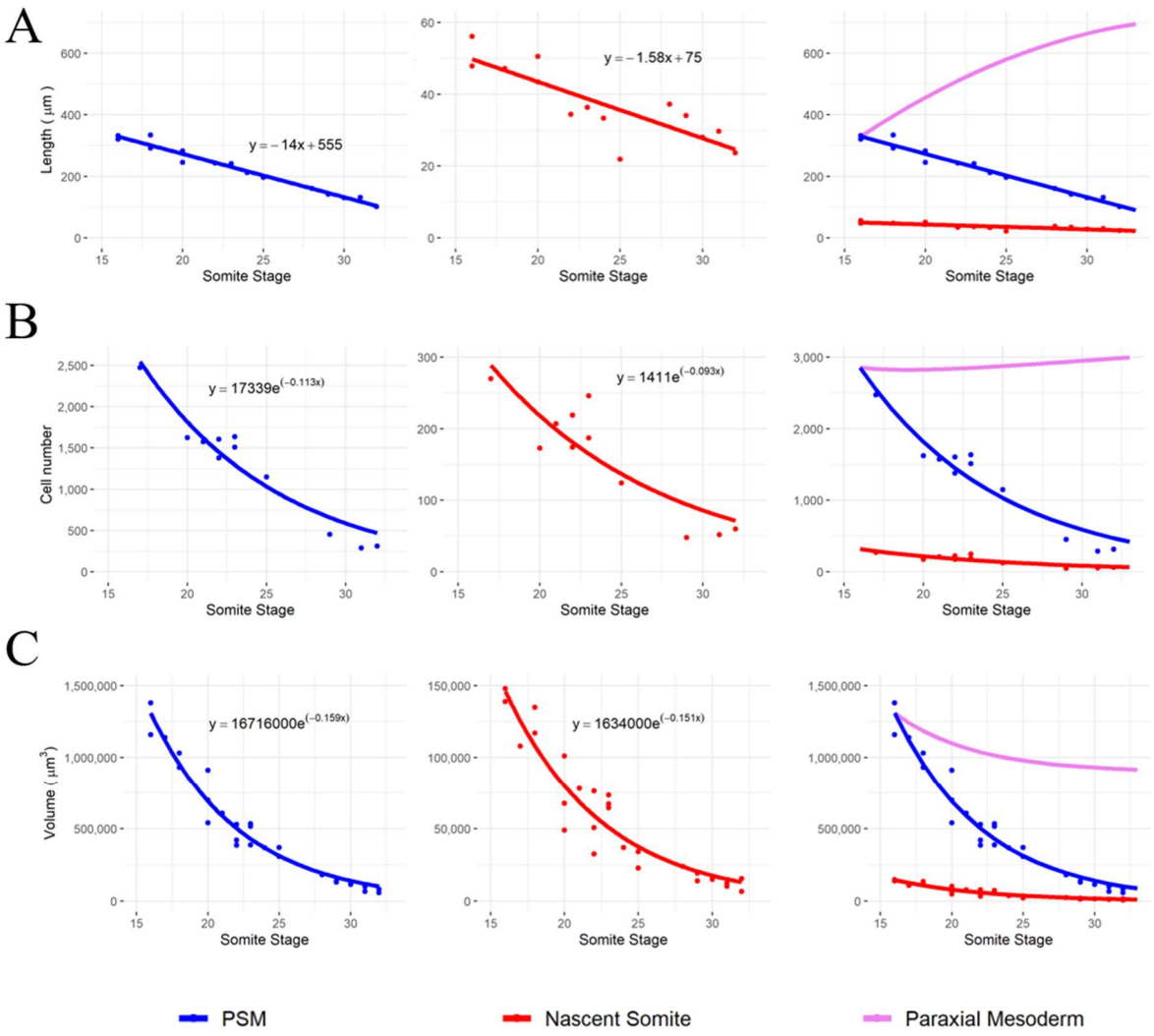


Figure 9: Length, cell number, and tissue volume measurements of the paraxial mesoderm over time (16 somite-stage onwards). (A) Length measurements of the PSM (left) and nascent somites at the stage of formation (middle) show that PSM length decreases, and each nascent somite is shorter than the one before. The linear trendline equations are combined as described previously to calculate the paraxial mesoderm length (right, violet) over time, which shows a logarithmic increase. n = 15 embryos. (B) Cell number counts of the PSM (left) and nascent somites at the stage of formation (middle) show that PSM cell number decreases exponentially over time, and each nascent somite is smaller than the previous one. The paraxial mesoderm cell number (right, violet) shows initial stasis with a slight increase at later stages. n = 11 embryos. (C) Tissue volume measurements of the PSM (left) and nascent somites at the stage of formation (middle) show that PSM volume decreases exponentially, and each nascent somite is smaller than the one before. The paraxial mesoderm volume (right, violet) shows exponential decrease (not growth), despite cell number increase. n = 26 embryos. Note that for (A-C), the y-axis scale is different between left and middle columns, but the same between left and right columns. Different embryo numbers are due to different imaging requirements for cell counts vs length and volume measurements.

1.4 Tissue convergence in both height and width

Given that elongation is not driven by growth, the most likely mechanism is some form of convergent extension. Measurements of both the posterior and the anterior PSM showed that both the height (DV axis) (Figure 10A) and width (ML axis) (Figure 10B) decrease over time by ~ 50% (Figure 10C). Accordingly, the height and width of each nascent somite is also less than that of the previous nascent somite. This strongly suggests that the paraxial mesoderm is elongating via convergent extension. However, from these measurements, the underlying cell behaviours are unknown – and so it is not clear whether this is active convergent extension (driven by directional intercalation), or another form of convergent extension (e.g. passive convergent extension).

It is also important to note that, from these measurements, it is unclear whether the anterior PSM thinning is a true process, or just the result of being progressively generated from a thinning posterior PSM. Given that nascent somite dimensions are highly similar to those of the anterior PSM (Figure 10C), the latter seems more likely, but live imaging is required to confirm this.

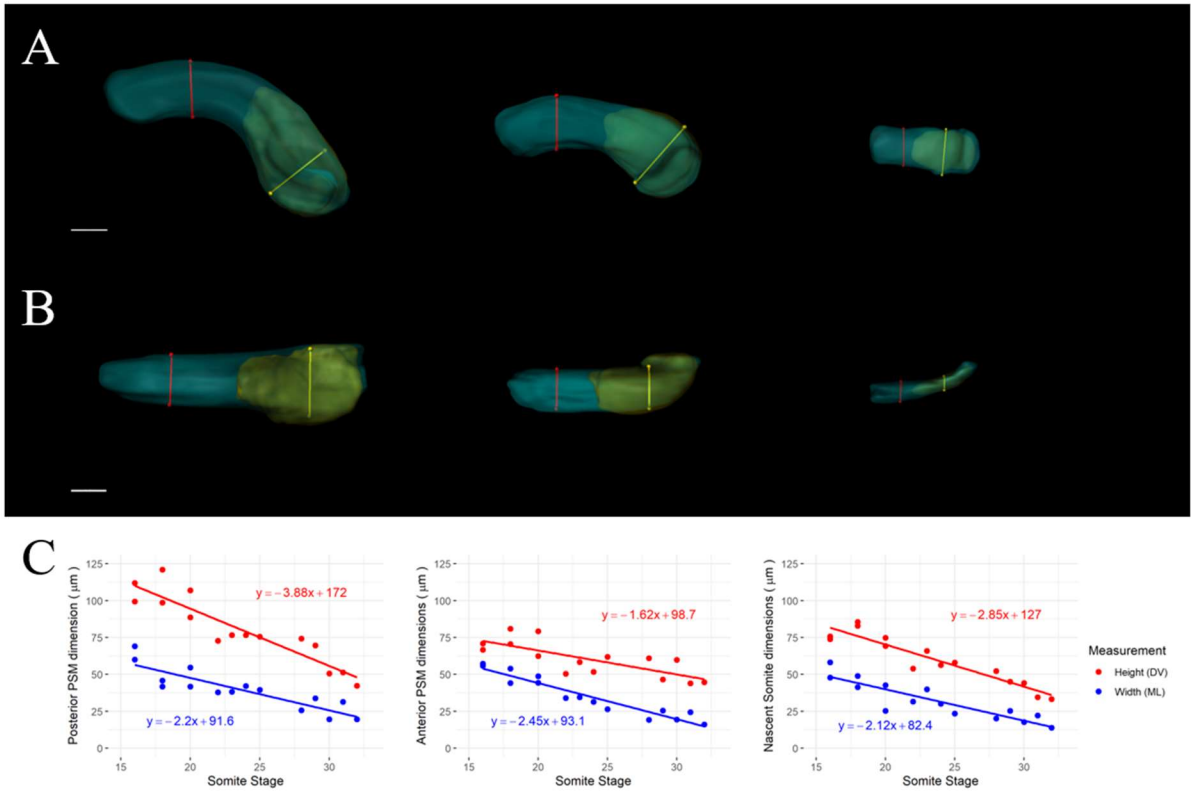


Figure 10: The PSM height and width decrease over time. (A) PSM surfaces, including full PSM manual surface (cyan) and automatic *msgn1* surface (yellow) are shown for three stages: 17, 23, and 32 somite-stages (left to right). Measurements of height (dorsoventral axis) are shown for the posterior (yellow line) and anterior (red line) PSM. Lateral view: anterior is left, dorsal is top. (B) Same images as above but in different orientation. Measurements of width (mediolateral axis) are shown for the posterior (yellow line) and anterior (red line) PSM. Dorsal view: anterior is left, medial is top. Scale bars in A and B: 40 µm. (C) Measurements of both dimensions shown for posterior PSM (left) and anterior PSM (middle), as well as for the nascent somites at the time of formation (right). Both height and width decrease over time in the posterior and anterior PSM, and each nascent somite is smaller (in height and width) than the previous one. n = 15 embryos.

1.5 A posterior-to-anterior gradient of convergence

To determine whether anterior PSM thinning is a true process or simply the result of being progressively generated from a thinning posterior PSM, I used photolabelling and live imaging. Embryos were injected at the one-cell-stage with mRNA for KikGR – a fluorescent protein that switches from green fluorescence to red fluorescence upon photoconversion with UV light (Habuchi et al., 2008). In one experiment (Figure 11A) regions were photolabelled in dorsoventral spot pairs at three different points along the PSM (posterior, middle, anterior), as a qualitative test of how much mixing is occurring at each point. The posterior label pair showed a large amount of mixing over 3 hours; the labels mixed together, and cells spread out along the PSM. The middle and anterior label pairs, however, remained as distinct spots and showed very little mixing and spreading. Increased mixing was observed, however, after labels entered a nascent somite, likely due to somite morphogenesis. In order to then quantitively measure tissue deformation along the axis, I performed a different type of photolabelling experiment in which dorsoventral stripes (across the full height of the PSM) were labelled at different positions along the PSM (Figure 11B). The length (AP axis) and height (DV axis) of these labels were measured immediately after labelling, and again after 2 hours. The results (Figure 11C) show that all labels show some length increase, but with a strong posterior to anterior gradient (i.e. posterior labels elongate the most). Posterior labels also show a reduction in height, which anterior labels do not (some even show an increase in height, although this is again likely related to somite morphogenesis). Taken together, these results show that tissue convergence displays a strong posterior to anterior gradient. The length-height ratio triples in posterior-most labels but stays constant in anterior-most labels (Figure 11C, right). Given that anterior PSM labels show very little deformation, this suggests that anterior PSM thinning is mostly the result of this tissue being progressively generated by a thinning posterior PSM.

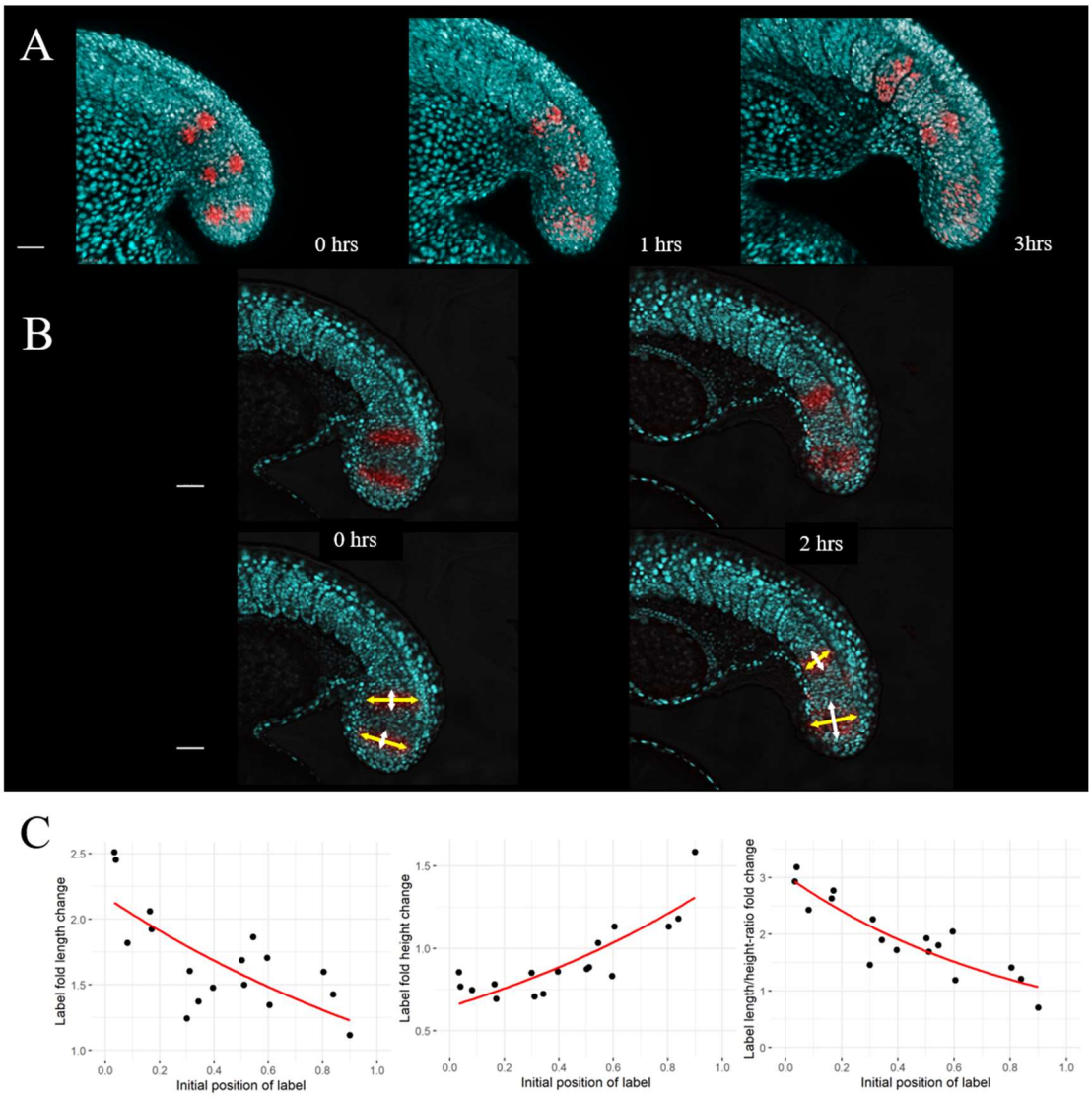


Figure 11: Photolabels show that tissue deformation is much greater in the posterior PSM. Embryos were injected at the one-cell-stage with mRNA for KikGR - a fluorescent protein which photoconverts from green (shown as cyan here) to red upon UV activation, allowing photolabelling of groups of cells. (A) Dorsoventral spot pairs were labelled at the posterior, middle, and anterior PSM of a 17 somite-stage embryo. Over 3 hours, posterior labels showed substantial mixing and spreading, whereas middle and anterior labels remained as distinct spots. Images are maximum projections. (B) For a more quantitative analysis, dorsoventral stripes were labelled at different points along the PSMs of 18-20 somite-stage embryos and imaged again after 2 hours (top images). The length (white arrows) and height (yellow arrows) were measured at 0 and 2 hours after labelling (bottom images: same as top images but showing measurements). Images are individual z-slices at the level of the PSM, with transmitted light shown in grey, Scale bars in A and B: 50 µm. (C) Graphs showing fold change of length, height, and length-height ratio over two hours. These are plotted against the initial AP position along the PSM, normalised from 0 (posterior tip) to 1 (anterior boundary). Note the different y-axis scales between graphs. n = 7 embryos (2-3 labels per embryo, 17 total labels).

1.6 Density increase via cell shrinkage

The surprising result that the paraxial mesoderm shrinks in volume over time suggests that PSM density increases over time. This was found to be the case, by dividing the number of PSM cells by the PSM tissue volume to obtain a density measurement in terms of cells per μm^3 (Figure 12A). This density increase is substantial ($\sim 80\%$), which suggests it could be biologically important. Given that there is very little cell addition/proliferation, this large increase in density cannot be caused by cleavage-like divisions (in which cells divide but do not grow). The alternative cause is that the cells that make up the tissue are themselves shrinking in volume. To test this, I fixed two batches of embryos (one at the 18 somite-stage, and one at the 26-somite stage, $n = 5$ embryos per batch) and stained the cell membranes with phalloidin (Figure 12B). Using Imaris, I made manual surfaces of randomly selected PSM cells (Figure 12C). For each embryo, I randomly selected five posterior cells and five anterior cells (Figure 12D), in order to determine if there were regional differences in cell size, as well as temporal differences. The results (Figure 12E) show that within each stage, there was no difference between anterior cell size and posterior cell size ($t = -0.36$, $p = 0.72$). However, there was a significant difference between stages ($t = 6.39$, $p < 0.001$) – PSM cells of 26 somite-stage embryos (mean = 220 μm^3) were smaller than those of 18 somite-stage embryos (mean = 367 μm^3). This shows that the decrease in PSM density is driven by a decrease in PSM cell sizes over time.

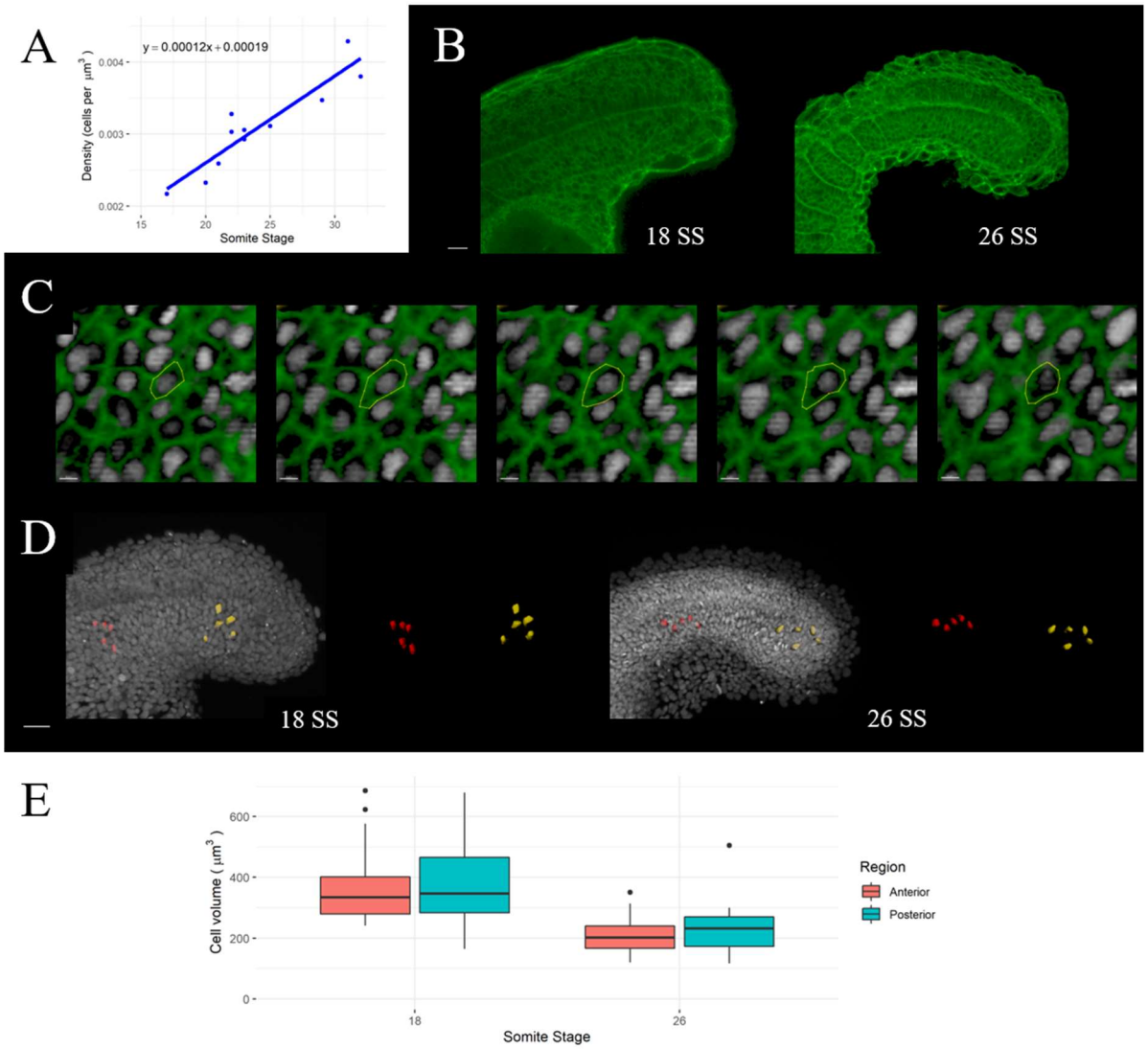


Figure 12: PSM density increases over time due to cells shrinking. (A) PSM density, as measured by number of cells divided by tissue volume, increases substantially over time. n = 11 embryos. (B) Phalloidin-stained (green) tails of 18 somite-stage and 26 somite-stage embryos were imaged and used for the subsequent cell size measurements. Images are optical sections (5 µm) at the level of the PSM. Scale bar: 20 µm (C) Manual surfaces creation. Each image shows the same cell at different z-slices, with the drawn contour (yellow) and DAPI (grey) shown. Scale bars: 3 µm (D) For each embryo, 5 posterior cells (yellow) were randomly selected and 5 anterior cells (red) were randomly selected for manual surface creation. Images show these surfaces with (left) and without (right) DAPI maximum projection, for both stages. Scales bar: 30 µm (E) Boxplot showing results of cell volumes, comparing between stages and between regions. PSM cells do not show regional size differences but do show temporal size differences. Later-stage PSM cells are smaller than earlier-stage PSM cells. n = 10 embryos (10 cells per embryo).

1.7 Tissue convergence accounts for the degree of elongation

The zebrafish PSM undergoes tissue convergence, but is the degree of convergence able to account for the substantial amount of elongation? By approximating the PSM as a simple geometric shape, a 3D elliptic cylinder, it is possible to apply convergence to a given set of dimensions (and a fixed volume) to determine how much elongation happens as a direct result. The volume of an elliptic cylinder is equal to the length multiplied by the cross-sectional area (CSA), which can be calculated based on the semi-height a and the semi-width b with the following equation:

$$CSA = \pi ab$$

As the previous axes measurements of the PSM involved two differing sets of height and width (one set for the posterior PSM, one set for the anterior), an approximation was required. The height and width trendlines (Figure 10C) were used to provide values for each stage, for both the anterior and the posterior PSM. Then, for each stage, both height values were averaged and halved to give an approximate semi-height a, and both width values were averaged and halved to give an approximate semi-width b. These were then used to calculate a CSA (using the formula above) for each stage. Approximating the PSM as an elliptic cylinder in this way provides a highly accurate approximation, as multiplying these CSA values by the PSM length trendline (Figure 9A) gives a volume trendline almost identical to that which was calculated from surface volume measurements (Figure 13A).

Given that the PSM can be accurately approximated as an elliptic cylinder, this can then be used to simulate the effect of convergent extension (using the most general definition). This cannot be done as a simple continuous process, due to the regular depletion of the tissue to form new somites. Instead, it must be done separately for each set of two consecutive somite stages. Assuming that the volume of the tissue remains

constant between two consecutive stages (ignoring the loss of volume when the nascent somite is formed, and treating this as still part of the tissue), the reduction in CSA between these two stages can be used to calculate the elongation length between these two stages. In other words, the CSA at the start of the second stage can be treated as the CSA at the end of the first stage (immediately before the loss of the somite tissue), to give a reduction in CSA over the period of the first stage. This reduction can then be used to calculate a length elongation value for that stage. When this process is repeated for all pairs of subsequent stages, the elongation values can be successively added to a starting PSM length to give the total paraxial mesoderm elongation. In other words, this modelling approach is not one continuous model, but the same model type applied to the PSM at each stage.

For example, take the 16 and 17 somite-stages. Let the volume of the cylinder PSM at the 16 somite-stage be 200 units^3, with a CSA of 20 units^2 and a length of 10 units. Then let the CSA decrease to 16 units^2 by the 17 somite-stage. Assuming the volume has stayed constant (200 units^3), the new length must be 12.5 units, since the length is equal to the volume divided by the CSA (200 / 16). This gives an elongation value of 2.5 units (12.5 – 10) for the 16 somite-stage. Then to calculate the elongation value for the 17 somite-stage, the true volume and length values for this stage must be taken. Let this be 144 units^3 with a length of 9 units (meaning the nascent somite volume must have been 56 units^3 (200 - 144) in volume and 3.5 units (12.5 - 9) in length). Then let the CSA decrease to 14 units^2 by the 18 somite-stage. The new length will be 10 (144 / 14), which gives an elongation value of 1 unit (10 - 9) for the 17 somite-stage. The two elongation values for each stage can be added together, to give a total elongation value from the 16 to the 18 somite-stage of 3.5 units. This example (visualised in Figure 13B) uses arbitrary units and only two sets of consecutive stages for the sake of simplicity, but the method is the same, which was applied to each set of two consecutive stages

from the 16 somite-stage to the 33 somite-stage. The calculation of the elongation value z for a give somite stage n can be expressed as follows:

$$z_n = \frac{V_n}{CSA_{n+1}} - l_n$$

Each elongation value z can then be progressively added to the PSM length to give the elongation over time l_n. Using the starting length from the PSM length trendline P_n, the equation is as follows:

$$\begin{cases} n = 16, & l_n = P_n \\ n > 16, & l_n = l_{n-1} + z_n \end{cases}$$

The results of this convergent extension (CE) model are shown in Figure 13C. Interestingly, the model very closely resembles the true elongation of the real paraxial mesoderm (indirect calculation from Figure 9A), with a total elongation of 410 µm (versus the true elongation of 366 µm). The difference in the total elongation amount and elongation dynamic (the trendline shows a lesser degree of curvature than the true elongation) is likely due to the fact that this is not a perfect model of the convergent extension process that is occurring in the PSM. As shown in the previous section, tissue convergence shows a posterior-anterior gradient, whereas this model took an average measure of convergence and applied it uniformly across the tissue. Additionally, the model treats the tissue as a simple elliptic cylinder, which, while being an accurate approximation, is still an approximation. Nevertheless, the differences between the model result and the true elongation are minor, and the overall result is very similar.

An important assumption in this modelling approach is that the volume of the cylinder remains constant during each stage, so that elongation can be calculated based on CSA reduction. However, my previous measurements of PSM density show that volume is not constant. The model does largely take this into account, by taking the

true volume at each stage (which will be affected by this density increase) rather than only taking the starting volume and applying convergent extension to this. However, it is possible that the minor density change within each stage could affect the elongation values. To test this, I applied compression to each stage. Taking the previously calculated PSM density trendline equation (Figure 12A), I calculated a volume compaction factor vc for each stage n, by dividing the density d of the next stage by the density of the current stage, expressed as follows:

$$vc_n = d_{n+1}/d_n$$

Assuming isometric compression, the length compaction factor c for each stage equals the cubic root of the volume compaction factor, i.e.

$$c_n = \sqrt[3]{vc_n}$$

This length compaction factor for each stage was then applied to the elongation value for each stage in the model by dividing the elongation value by the compaction factor to give an adjusted elongation value for each stage, which was then added to the PSM length at each stage. This was then used to calculate overall elongation of the paraxial mesoderm, as before.

As expected, this minor amount of unaccounted-for density decrease did not significantly affect the elongation result (404 µm versus 410 µm) (Figure 13C, green line vs blue line), although this minor reduction does give a result which is slightly closer to the true elongation value of 366 µm.

Importantly, this 3D modelling approach confirms that tissue convergence is sufficient to account for the observed elongation of the paraxial mesoderm.

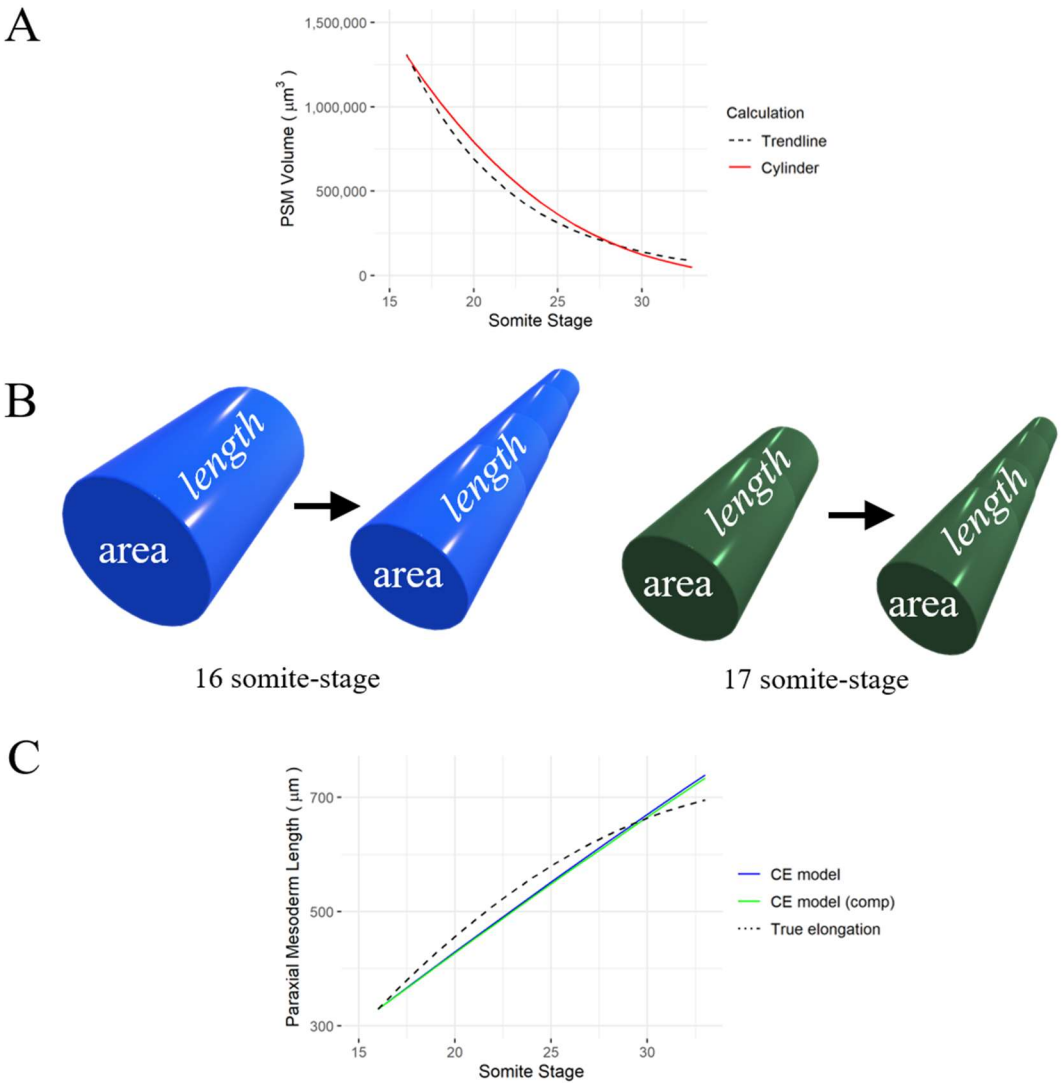


Figure 13: Tissue convergence is sufficient to explain paraxial mesoderm elongation. The PSM was approximated as an elliptic cylinder, in order to simulate the effects of convergent extension on the PSM. Height and width values were combined to give a cross-sectional area (CSA) (see text for details). (A) The elliptic cylinder approximation (red line) was highly accurate, showing an almost identical volume trend (by multiplying CSA by PSM length) to the surface volume trendline (black dotted lines). (B) Convergence was applied in a stepwise, rather than continuous, manner – to account for loss of somite tissue from the PSM. The area (CSA) reduction between two stages was used to calculate an elongation during the first stage (blue), assuming constant volume (see text for details). This process was then repeated for the next set of subsequent stages (green) using the true PSM length rather than elongation length (to account for somite formation). (C) Tissue convergence applied to the PSM (blue line) produced a very similar amount of elongation to the calculation of indirect paraxial mesoderm length (black dotted line), showing that convergent extension is sufficient to explain paraxial mesoderm elongation. Applying compression (comp) (green line) to this model reduced the elongation negligibly, as most compression was already accounted for by applying the model in a stepwise manner separately to each stage.

1.8 Compression-extension: a novel form of convergent extension

The finding that the PSM increases in density due to a reduction in cell volumes strongly suggests that the tissue is being externally compressed over time. This compression could be causally linked with the tissue convergence described in the previous sections. Therefore, I propose a novel form of convergent extension to explain zebrafish paraxial mesoderm elongation: "compression-extension". In this proposed mechanism, tissue compression partly results in convergence and extension of the tissue, and partly results in isometric tissue compaction. While it seems counterintuitive that a process causing tissue shrinkage could simultaneously drive tissue elongation, presumably the balance between these two outcomes could be regulated by the compressibility of the cells and of the tissue. It is possible, of course, that tissue compression and tissue convergence are simply co-occurring but not functionally related. This will be discussed more in later chapters, in which I describe the underlying cell behaviours. Importantly, however, the proposal that these processes are linked suggests a biological function of the substantial increase in density, which otherwise appears to have none. Compression-extension, therefore, provides a reasonable explanation for the observed tissue changes, and a reasonable mechanism of tissue elongation.

1.9 The leftover PSM

The mechanism by which somitogenesis and paraxial mesoderm elongation cease in zebrafish is unknown. Interestingly, I observed that in embryos at the end of somitogenesis (~ 33 somite-stage), there was still some unsegmented paraxial mesoderm in the tip of the tail, approximately 3-4 somite-lengths long, which was still expressing *msgn1* and *tbx6*. To my knowledge, this has not been reported in zebrafish. I shall refer to this as the "leftover PSM".

The presence of a leftover PSM suggests that somitogenesis does not simply end because the PSM is fully depleted. However, it is possible that the leftover PSM is simply segmented into somites later in development. To test this, I fixed embryos at regular intervals from the end of somitogenesis (25-28 hours post-fertilisation (hpf)) to 32 hours hpf, and performed HCR to stain for the following genes: *msgn1, tbx6*, *myod1* (somite marker gene), and *her1* (clock gene, homologous to *hes1* in chick). The results, shown in Figure 14, show that the leftover PSM does not segment into somites. Another possibility is that the whole leftover PSM differentiates into one long somite. However, the leftover PSM does not show *myod1* expression, suggesting that this is not the case (Figure 14A,B). Expression of both *msgn1* and *tbx6* ceases by 32 hpf (Figure 14A,B), as does expression of *her1* (Figure 14C). This suggests that somitogenesis ceases because clock gene expression ceases, perhaps due to decreasing Wnt/FGF activity over time. Importantly, DAPI staining shows no pycnotic nuclei, suggesting that there is little to no cell death in the leftover PSM.

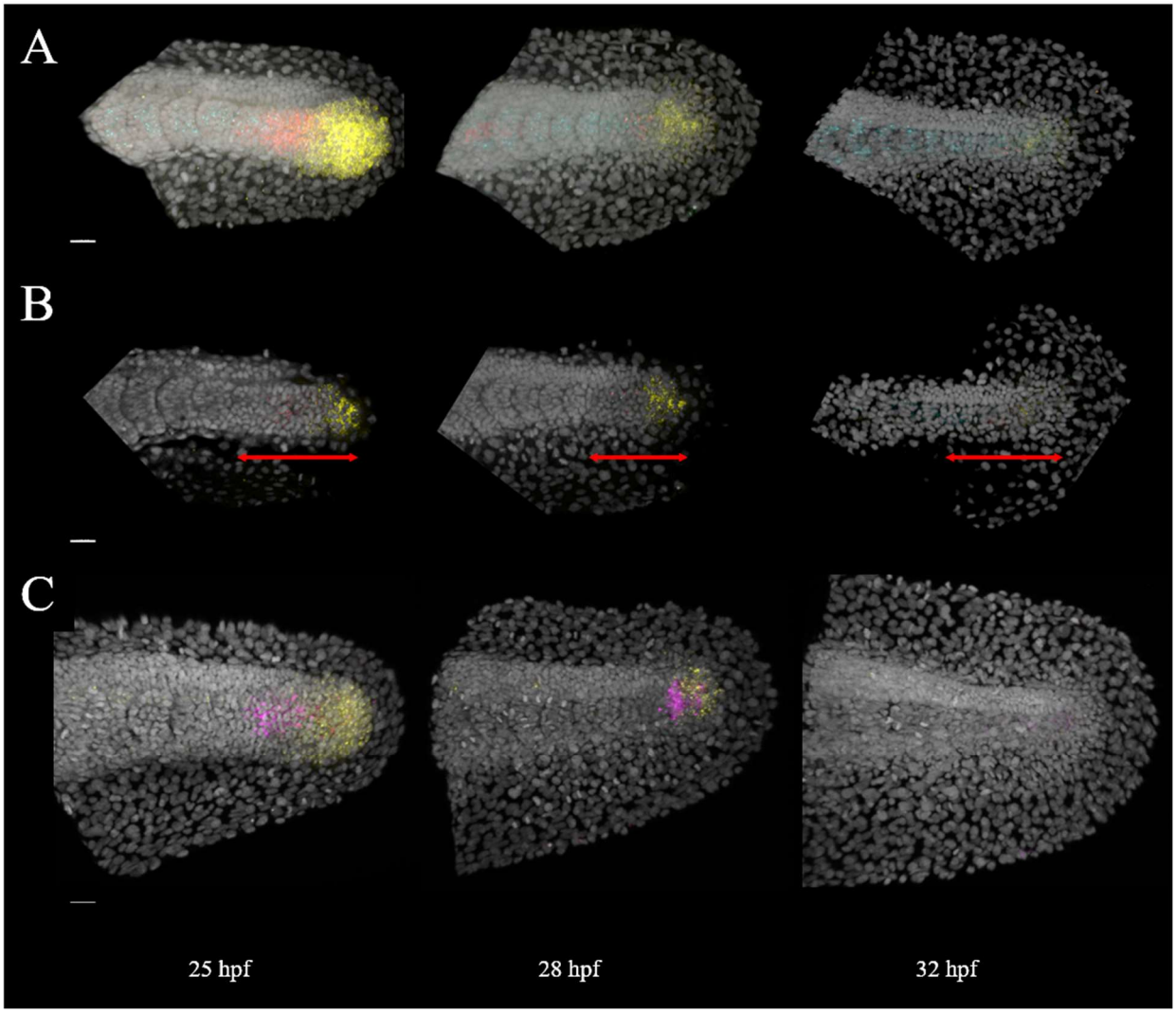


Figure 14: Somitogenesis stops before the PSM has been completely depleted. Embryos were fixed from the end of somitogenesis (25 hpf – 28 hpf) to 32 hpf and stained, using HCR, for various PSM and somite genes, as well as DAPI (grey). (A) HCR for *msgn1* (yellow), *tbx6* (red), and *myod1* (cyan). Expression of *msgn1* and *tbx6* decrease and cease in the leftover PSM, and *myod1* expression remains restricted to the somites. Images are maximum projections. (B) Same images as above, showing optical slices of 10 µm at the level of the PSM. Red arrow lines indicate the region of unsegmented paraxial mesoderm (leftover PSM). By 32 hpf, identifying the PSM/somite boundary is difficult as cells from both tissues spread into the fin mesenchyme (see Figure 15). (C) HCR for *msgn1* (yellow) and *her1* (magenta). Expression of *her1* decreases and then ceases by 32 hpf. Images are maximum projections. Scale bars: 20 µm.

It has been reported that cells from the somites migrate out and contribute to the dorsal and ventral fin mesenchyme (Lee et al., 2013). This raises the possibility that the leftover PSM also migrates out to contribute to the fin mesenchyme. To test this, I live-imaged *h2a::*mCherry x *tbx16::*GFP embryos from the end of somitogenesis for 6 hours. Because GFP is highly stable, it is still expressed in the anterior PSM and several somites, despite these tissues no longer actively expressing *tbx16*. This allowed me to use *tbx16::*GFP to follow paraxial mesoderm cells for several hours after the end of somitogenesis. The results show that the leftover PSM does also contribute to the dorsal and ventral fin mesenchyme (Figure 15), as cells move out of this region into the fin. Given that the leftover PSM cells do not go through the same differentiation process as somite cells (e.g. they do not express *myod1*), this raises the interesting question of how similar fin mesenchyme cells derived from the leftover PSM are to those derived from the somites.

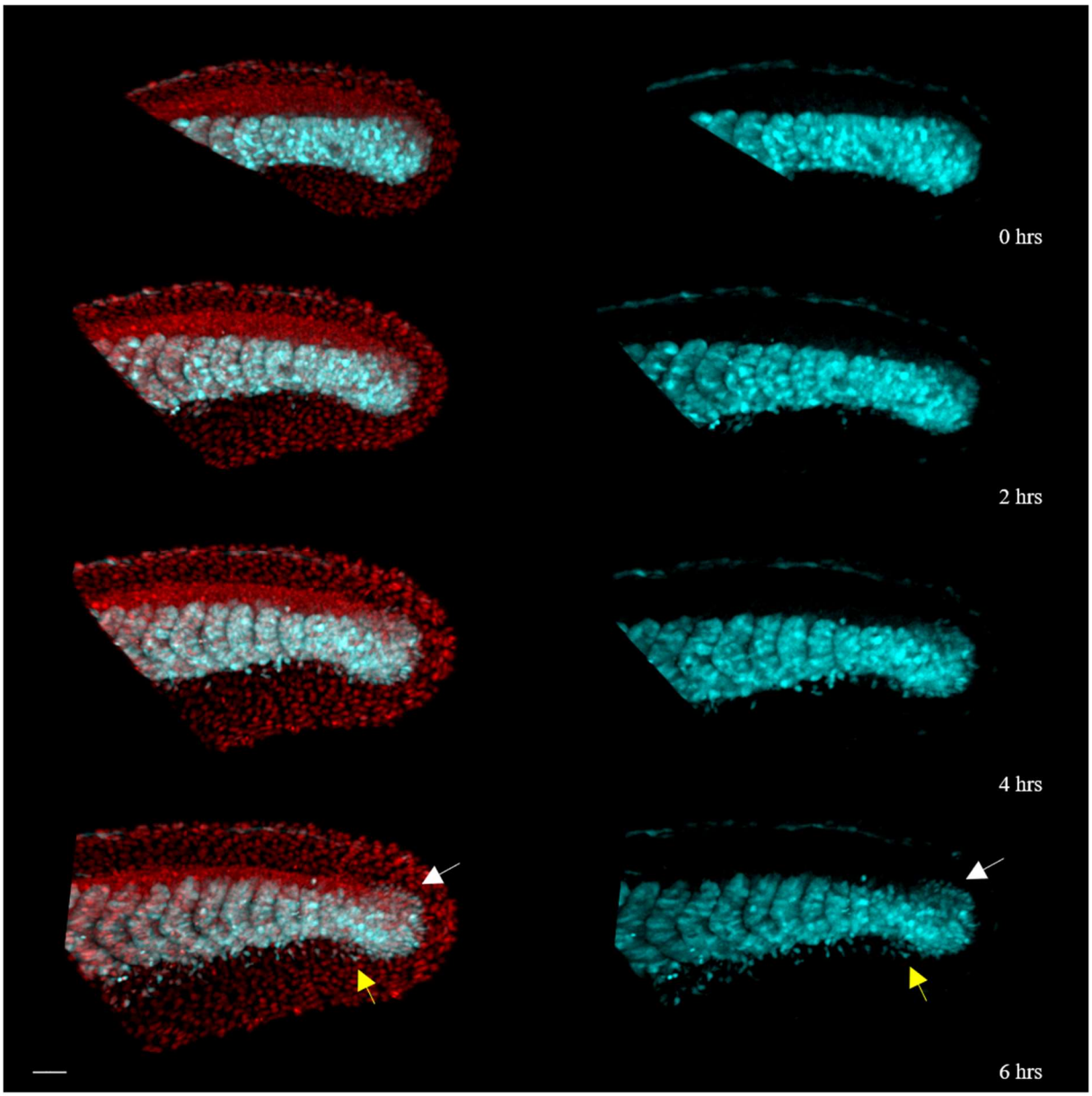


Figure 15: The leftover PSM contributes to fin mesenchyme. *h2a::*mCherry x *tbx16::*GFP were live-imaged from the end of somitogenesis (0-2hrs) for 6 hours (top to bottom). Left images show both nuclei (red) and *tbx16::*GFP (cyan), while right images show just *tbx16::GFP*. Cells can be seen moving out of the leftover PSM into the dorsal (white arrow) and ventral (yellow arrow) fin mesenchyme after 6 hours of imaging, as can cells from the somites. All images are maximum projections. Scale bar: 40 µm.

1.10 Conclusions

The results of the 3D morphometric measurements show that the zebrafish PSM elongates, not through growth/proliferation, but through tissue convergence. Tissue convergence displays a strong posterior-anterior gradient and is sufficient to drive the observed elongation of the tissue – as confirmed by a simple geometric modelling approach. The results also show that PSM density, while constant along the AP axis within stages, increases considerably over time, as cells reduce in volume. Therefore, I propose that tissue convergence and the increase in tissue density are linked by a novel form of convergent extension: "compression-extension".

Additionally, my observations of a leftover PSM suggest that somitogenesis does not end due to depletion of the PSM. Instead, it seems likely that the cessation of clock gene expression is responsible. The leftover PSM does not differentiate into a single somite, nor does it appear to undergo cell death, but instead contributes to the fin mesenchyme, as somite cells do.

There are several unanswered questions from these results. Most importantly, what are the cell behaviours underlying compression-extension? This question will be addressed in the next chapter.

RESULTS

2 Tissue convergence without directional intercalation

2.1 Tracking methods

Two-photon time-lapse imaging

To determine the cell behaviours and movements underlying the compression-extension of the zebrafish PSM, individual cell 3D tracks are required. To obtain cell tracks, I live-imaged *h2b::*GFP zebrafish embryos of two different stages on a two-photon microscope for 2-3 hours. I also used an existing two-photon movie*, taken by Ben Steventon, of an *h2a::*mCherry zebrafish embryo of an earlier stage. The different movies used are summarised in Table 1, and stills from these movies are shown in Figure 16. Each movie from now on will be referred to by its movie (M) number (given in Table 1).

Table 1: Summary of two-photon movies used for cell tracking analyses.

Movie	**Somite-stages**	**Frame interval (s)**	**Nuclear label**
M1*	14 - 20	70	*h2a::*mCherry
M2	18 - 22	180	*h2b::*GFP
M3	22 - 26	150	*h2b::*GFP

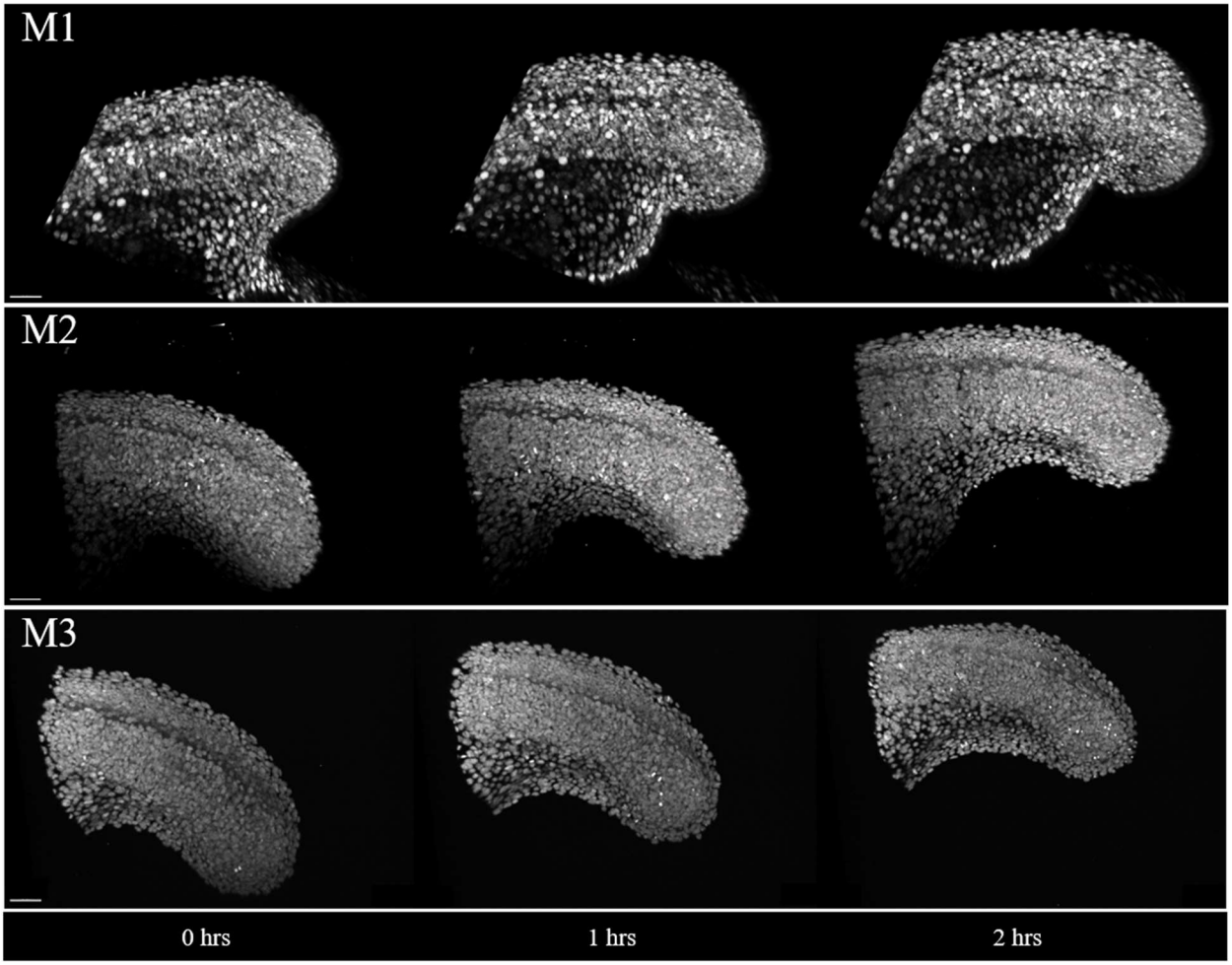


Figure 16: Two-photon time-lapse imaging of the zebrafish tailbud. Three embryos of different stages were live-imaged for 2-3 hours each. Nuclei are labelled with either *h2a::*mCherry or *h2b::*GFP (both shown as grey). See Table 1 for summary of each movie. Images are shown of the tailbud at 0, 1, and 2 hours after the start of imaging. All images are maximum projections. Scale bars: 40 µm.

Reference frames: normalising for global tissue movement

To draw meaningful conclusions from cell tracks, a frame of reference is required: what are cells moving relative to? The zebrafish tailbud is elongating and uncurling throughout most of somitogenesis, and so tracking PSM cells through absolute space will simply reflect the global tissue movement. Using Imaris imaging analysis software, a "reference frame" was placed at the end of the tailbud, at DV and ML midline (Figure 17A). The reference frame includes axes, which I orientated to match the three biological axes (x = AP, y = DV, z = ML). I placed and orientated the reference frame every 5 frames for the full movie duration, for each movie. Imaris then automatically calculated the movement and rotation of the reference frame, using linear interpolation, for in-between frames, and placed the reference frame accordingly for these frames. Tracking cells relative to this reference frame allowed cell movements to be normalised for the global uncurling of the AP axis. I then performed automatic "resampling" in Imaris to adjust each image so that, rather than the reference frame moving over time, the reference frame (i.e. the tailbud tip) would be a fixed point in space. This allowed visualisation (of movies/tracks) to reflect the normalisation for global movement (Figure 17B).

Other reference frames were also placed at different biologically relevant positions along the AP axis. This was done for analyses focusing on AP movement of cells (Section 2.3). In addition to the previously described tailbud tip reference frame, separate reference frames were placed at the posterior end of the notochord proper (DV & ML midline as before), and at the posterior boundary of the start-of-movie nascent somite (DV midline, medial side of the somite) (Figure 18). Reference frames were placed at regular frame intervals, as before.

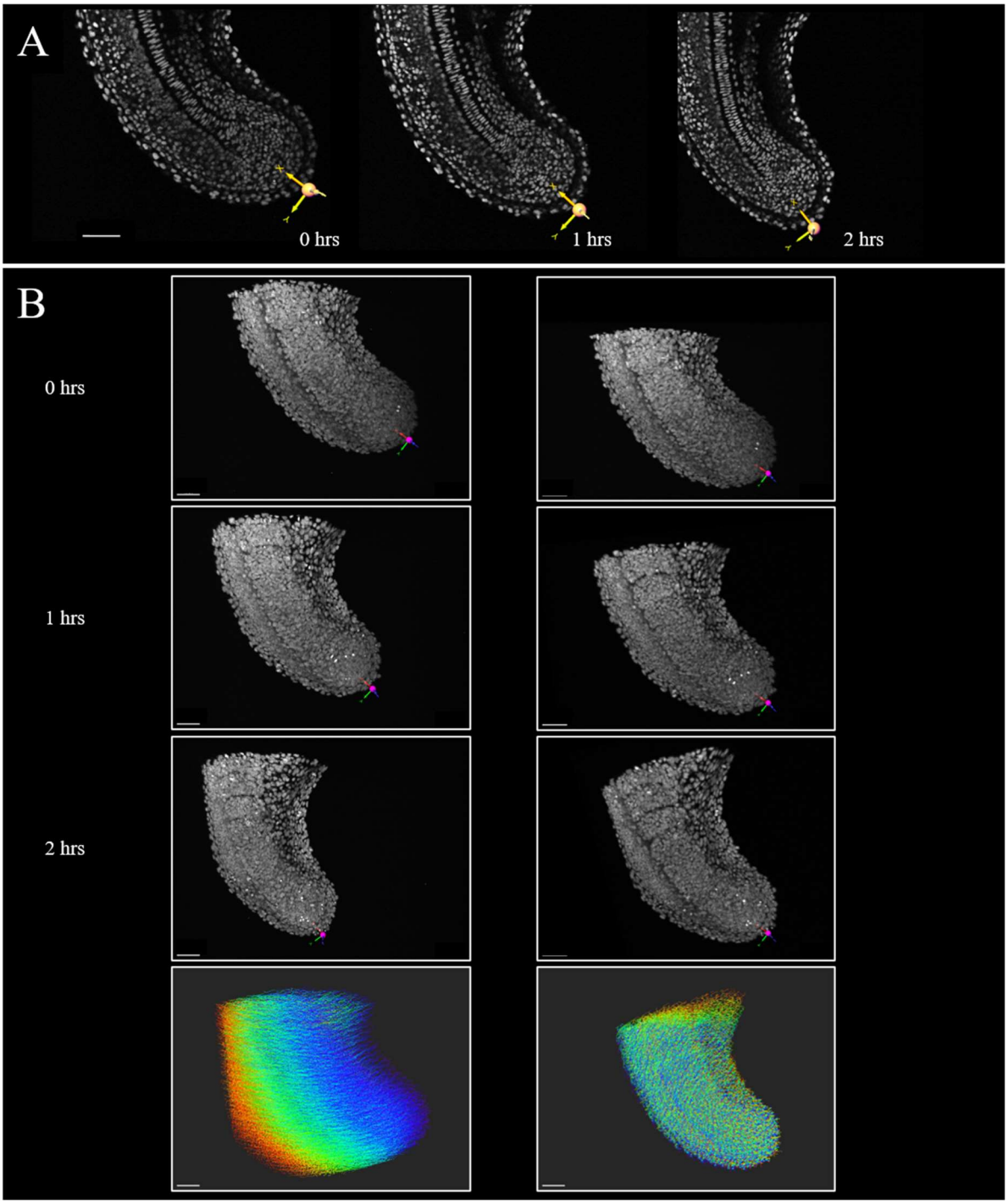


Figure 17: Normalising cell movements for global tail uncurling. A "reference frame" was placed at the DV/ML midline of the tailbud tip (with axes lined up to match biological axes) at regular intervals. (A) Individual z-slice showing reference frame placement over 2 hrs. Scale bar: 50 µm. (B) Images and cell tracks before (left) and after (right) "resampling" images to fix the reference point (tailbud tip) in space. Images and tracks (all cells) show a large dorsal movement before resampling, but none after resampling. Tracking colour code (blue to red) indicates time. Scale bars: 40 µm. All images from M3. Anterior is top, dorsal is left.

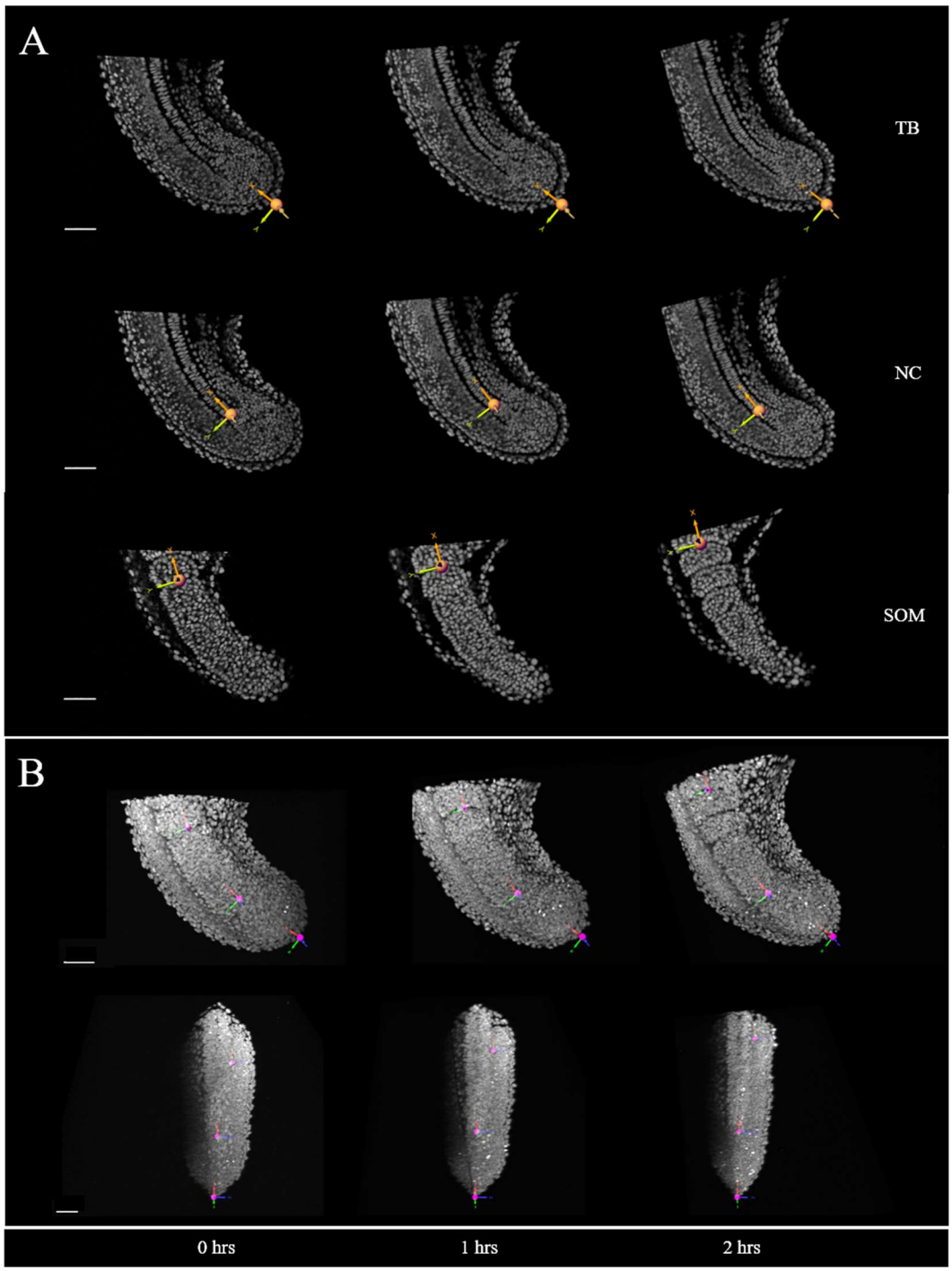


Figure 18: Multiple biological reference frames (RFs) along the AP axis. (A) Individual z-slice showing tailbud (TB) RF as previously described (Figure 17); notochord (NC) RF placed at the posterior end of the notochord proper (where the notochord "funnels out" into the notochord progenitors); and somite (SOM) RF placed at the posterior boundary of the start-of-movie nascent somite. (B) All three RFs shown together on maximum projection images from lateral (top) and dorsal (bottom) views. Scale bars: 40 µm. All images from M3. Anterior is top, dorsal is left (except bottom row in B, in which medial is left).

Automatic track creation and manual validation

Automatic nuclear tracking in Imaris involves two steps: spot reconstruction and spot tracking.

Spot reconstruction is the same process as previously described (section 1.1), requiring two input parameters: cell diameter and minimum fluorescence intensity threshold. Cell diameter was set to 4.0 µm (as before), while minimum fluorescence intensity threshold was adjusted manually, separately for each movie – due to differences between movies in background/autofluorescence. As before, spot parameters were validated by separately "masking" fluorescence inside and outside spots to create separate channels, showing dots inside each identified nucleus. Imaris runs spot reconstruction independently for each frame, and then, during the spot tracking step, attempts to link each spot in one frame with the correct spot in the next time frame, from the start to the end of the movie.

Spot tracking requires an algorithm input, for which there are five options in Imaris: Connected Components; Brownian Motion; Autoregression Motion; Autoregressive Motion Expert; and Lineage. Each algorithm has a different "expectation" of how cells should move. These will be briefly explained now, as subsequent tracking parameters, and what they mean, depend on the algorithm being used.

"Connected Components" (CC) calculates the overlap in position between frames, and only link spots that overlap. In other words, the algorithm expects cells to move by less than their diameter. If one cell overlaps with two cells in the next frame, this is treated as a cell division event, with both cells in the second frame treated as daughter cells. No manual input parameters are required.

"Brownian Motion" (BM) expects cells to move randomly, with frequent changes in direction and no discernible pattern. The first manual input parameter required is "Max Distance", which is the maximum distance (µm) the algorithm expects spots to move between frames. The other manual input parameter is "Gap Size", which is the

number of consecutive frames a cell can be lost for (i.e. number of gaps) before the track is considered finished (and a new track begins). As an example, if Max Distance is set to 5, if gap size is 1, then a spot can be linked with one two frames later (i.e. after a gap of 1), as long as it is within 10 µm, even if the nucleus is not identified as a spot in the intermediate frame.

"Autoregressive Motion" (AM) expects cells to move somewhat continuously. It assumes that a cell will move in the same direction, and by the same distance, between the current frame and the next, as it did between the previous frame and the current one. Here, the Max Distance parameter is the maximum distance the algorithm allows a cell to *differ* from the predicted trajectory (rather than the maximum distance a cell can move, as in Brownian Motion). Gap Size works similarly in this algorithm, although it will predict where the cell should be in the gap and calculate the Max Distance from this predicted point to spots in the next frame.

"Autoregressive Motion Expert" (AME) is the same as Autoregressive Motion, but with an added manual input parameter: "Intensity Weight". This algorithm takes fluorescence intensity into consideration by assuming that this should remain constant within a spot. The Intensity Weight parameter determines how much weight should be given to this, versus the weight given to the predicted position.

"Lineage" (LI) is the only algorithm other than Connected Components to include cell divisions within tracks. This algorithm uses Autoregressive Motion to predict positions, but also tries to identify the splitting of individuals spots into two - and considers these daughter cells (and therefore both part of the same initial track).

To determine which was the best tracking algorithm, I tested all five on the first hour of a two-photon movie (M3). For a first round of validation, tracks of the whole tailbud were made using the same spot reconstruction parameters, but different spot tracking algorithms. For each algorithm (except CC which does not include the parameter), two different Max Distance parameters were tested: 5 µm and 10 µm. Gap

Size was set to 3 for all three AM-based algorithms. Various Intensity Weight parameters (ranging from 0 to 10,000) were tried for AME. In this first round of testing, the main result being assessed was track duration – any algorithm which could not track cells for long durations (> 30 mins) were to be excluded. The results of track duration are shown in Figure 19A. Regardless of parameters, CC and AME failed to track cells for long durations, and so these were excluded.

To determine which was the best tracking algorithm out of the remaining three (AM, BM, and LI), I manually validated the accuracy of PSM tracks. First, I filtered tracks by duration to only include those which lasted the full hour. Then I manually selected 20 PSM tracks from each set of tailbud tracks. Manual validation involved using the same spot validation method as previously described, to follow the coloured spot over time and check if it stayed in the same nucleus. The initial LI track set (parameters: MD = 5, GS = 3) reported 8 cell divisions; however, manual validation confirmed these were all false positives. Given this level of inaccuracy, LI was excluded from further testing.

To choose between the two remaining algorithms (AM and BM), three different track sets were made for each algorithm, one for each MD parameter: 5, 7.5, and 10. Gap Size was kept constant (GS = 3) for all track sets. The results of validation are shown in Figure 19B. The best tracking algorithm and parameter set for this movie (M3) was found to be Autoregressive Motion (MD = 5), with 16/20 tracks showing no errors. Higher MD values led to more errors for both algorithms, and more gaps for AM. Track gaps were assumed to be a possible source of error, so it was noted whether errors occurred over gaps. However, for all track sets, gaps showed mostly accurate tracking, and for AM (MD = 5), all three gaps were error-free.

After choosing Autoregressive Motion as the best algorithm, this was tested on the other movies (M1, M2), using different MD parameters, to check if MD needed adjusting for differences in frame interval. For both M1 and M2 two sets of AM tracks were made,

one with MD = 5, and one with an "adjusted" MD (MD = 3 for M1, MD = 6 for M2). For each set of tracks, track durations were measured, and PSM tracks were manually validated as before (Figure 19C).

The results show that M1 (AM, MD = 5) tracks were highly accurate (20/20 accurate PSM tracks) but, on average, did not last very long (mean = 25 min). Adjusted MD tracks were not validated for M1, as they showed even shorter average duration (mean = 10 min) - and would not provide improved accuracy (given that 20/20 tracks were accurate). M2 showed highly inaccurate tracks for both MD values, but less so for MD = 6 (10/20 accurate PSM tracks). The difference in accuracy between M1 and M2 is likely due to the difference in frame interval (70s for M1, 180s for M2). The short track durations for M1 is likely due to slight twitching of the embryo during imaging, causing track breaking.

Based on all these results, Autoregressive Motion (AM) was used as the tracking algorithm for all three movies, with Gap Size set to 3 for all movies. For M1 and M3, a Max Distance (MD) of 5 was used, but for M2 an MD of 6 was used. Because M3 showed both long track durations (mean = 50 min) and high tracking accuracy, this movie was used as the main movie in tracking analyses, with the other two movies used as additional movies to check the generality of results.

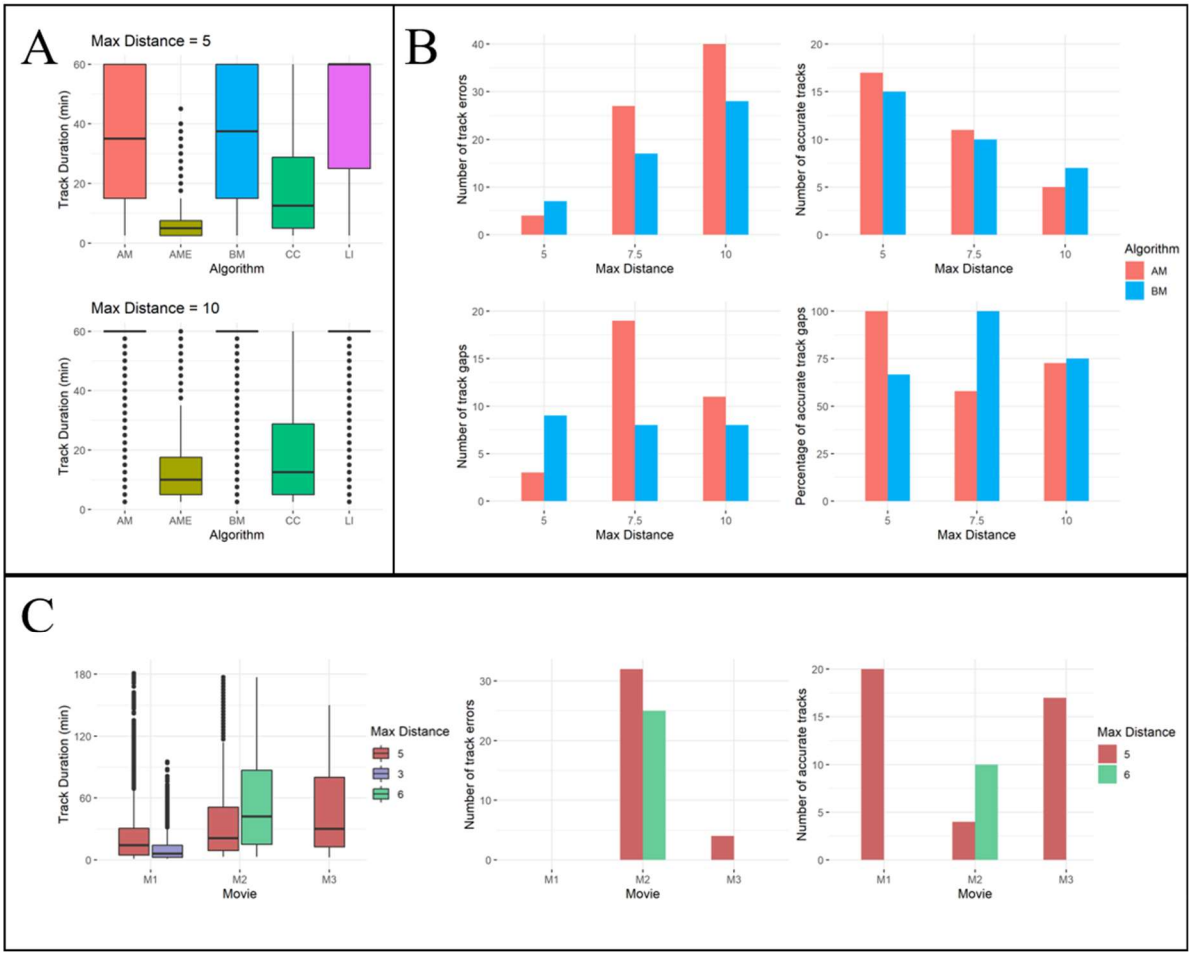


Figure 19: Validation of automatic tracks. (A) Tracks were generated of the whole tailbud (first hour of movie M3) using 5 different algorithms: Autoregressive Motion (AM); Autoregressive Motion Expert (AME); Brownian Motion (BM); Connected Components (CC); and Lineage (LI). For each algorithm, two sets of tracks were made for two Max Distance (MD) parameters (5, 10). Boxplots show track duration for each set of tracks. AME and CC were excluded from further analysis due to short track durations. (B) Manual validation of track accuracy (LI was excluded at the beginning of analysis due to high false positive reporting of cell divisions). Tracks were generated for AM and BM, for three MD parameters each (5, 7.5, 10), and 20 PSM tracks (with duration $>$ 60 min) were randomly selected for validation over 60 min. Number of errors (top left) and number of accurate tracks (error-free tracks) (top right) were recorded. Track gaps were also recorded (bottom left), and the percentage of accurate (error-free) gaps were noted (bottom right). (C) Having identified AM as the best algorithm, and MD = 5 as the best value for movie M3, tracks were made for all three full movies. For M1 and M2, two sets of tracks were made with an adjusted (for difference in frame interval) and non-adjusted MD value. Boxplot shows track duration for each set of tracks. For each track set, 20 PSM cells were randomly selected and validated as before. Bar charts show number of errors (left) and number of accurate tracks (right). (For M1, the adjusted MD tracks were not validated, given the 100% accuracy and longer duration of non-adjusted MD tracks.)

Isolation of paraxial mesoderm tracks

Having chosen the best tracking algorithm, and the best parameters for each movie, I created tracks for all three movies with these settings. I then isolated only the paraxial mesoderm tracks from these total tailbud tracks, partly using position-based automatic filters in Imaris, and partly by manually removing each non-paraxial mesoderm track. As in surface reconstructions (Section 1.1), only one lateral half of the paraxial mesoderm was included – cells past the midline of the embryo were excluded. Because the images only contained nuclear signal, and no gene expression information (unlike the HCR images used for surface/spots reconstruction), it is likely that some notochord and neural progenitors are included in the tracks. However, these will constitute a negligible minority of cells, especially as any tracks that moved into the notochord or neural tube were excluded. Figure 20 shows isolation of paraxial mesoderm tracks, alongside previous morphometric spots reconstruction for comparison. Tracks of the full visible paraxial mesoderm were included - somite tracks were not excluded at this stage, although they were excluded in some analyses based on position relative to the somite reference frame.

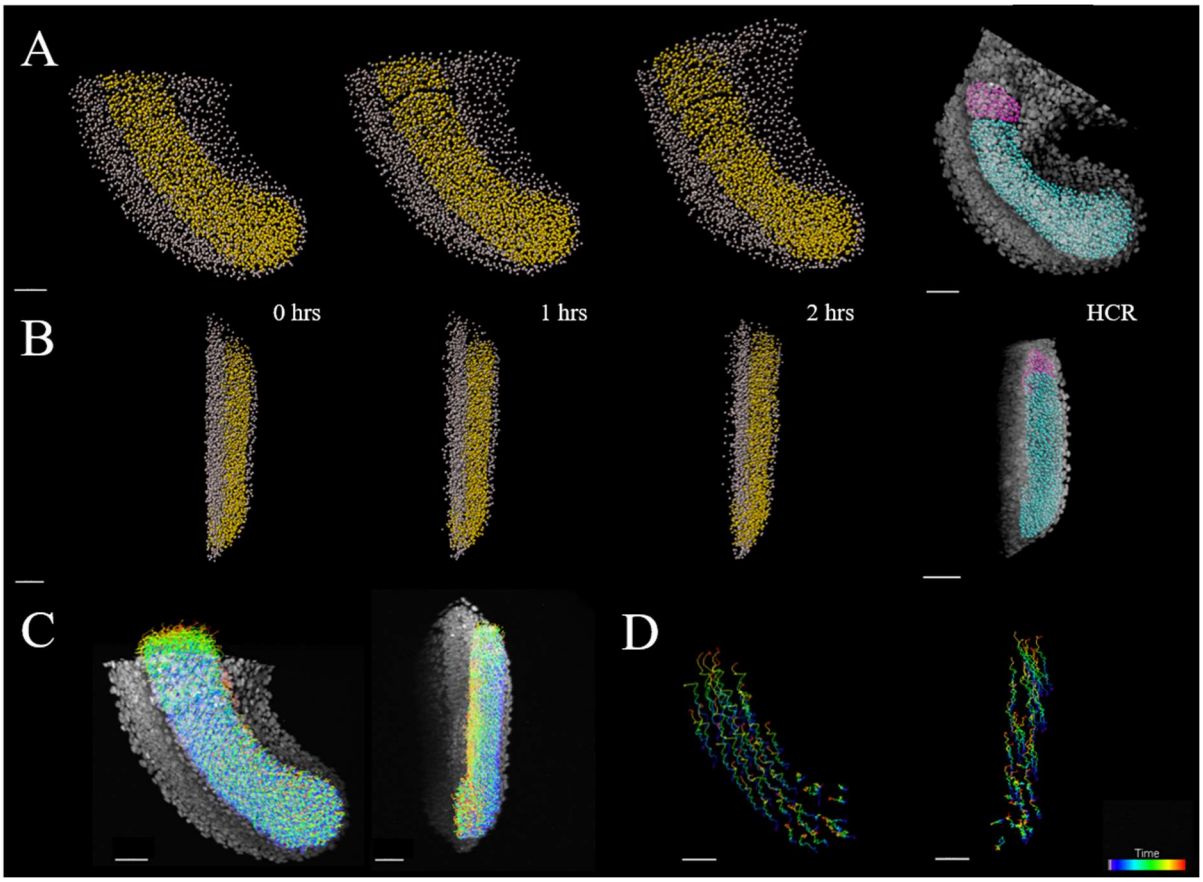


Figure 20: Paraxial mesoderm tracks selection. Tracks were generated of the whole tailbud and all non-paraxial mesoderm tracks were then manually removed. (A) Paraxial mesoderm (PM) tracks (yellow spots) and non-paraxial mesoderm tracks (grey spots) shown at 0, 1, and 2 hours after imaging. PM spots reconstructions (cyan & pink) from a similar stage HCR image is shown for comparison/validation of selection accuracy. All images are lateral view: anterior is top, dorsal is left. (B) Same images as above, but dorsal view: anterior is top, medial is left. (C) PM tracks of full movie (colour coded by time) superimposed over first frame image, shown for lateral and dorsal views (left, right). (D) A random selection (n = 50) of long-duration (>2 hr) PM tracks are shown, for lateral and dorsal views (left, right). Images are all from movie M3, but this process was performed on all three movies. Lateral view scale bars: 40 µm, dorsal view scale bars: 50 µm.

2.2 Cell tracking statistics match photolabelling results

The manual validation of selected PSM tracks in the previous section gave an idea of individual track accuracy for each movie. However, I also wanted to test tracking accuracy in a more general way for the whole tissue. Since manually validating all ~ 2,000 cells for each movie was not feasible, I instead used whole-tissue tracking patterns that could be compared qualitatively with photolabelling results (Section 1.5), to test if cell tracks gave similar results.

Given that posterior labels showed much more mixing than anterior labels (Figure 11A), I looked at track mixing along the AP axis using a MATLAB script for "neighbourhood analysis" (written by and provided by Leila Muresan). Neighbourhood analysis involves, for every frame, taking a given number k of nearest cells (neighbours) for each cell to form a neighbourhood for that cell at every frame. Combining this with track ID numbers, the script calculates the number of new cells that enter this neighbourhood over a given time t. A schematic of this analysis is shown in Figure 21A.

The results of this analysis (Figure 21B) show that posterior cells exchange neighbours more than anterior cells – in other words, posterior cells mix more than anterior cells. This difference appears as a steep gradient from the posterior tip to the (approximate) middle of the PSM, at which point there is a slight increase in mixing from this point to the anterior PSM/somites. This slight increase is consistent with that of photolabelled regions once they enter the somites (Figure 11A). These results were consistent between small ($k = 10$) and large ($k = 50$) neighbourhoods, and between movies. However in M1 and M2, the initial gradient was less steep. This is likely due to track breaking in M1, and track error in M2. M1 also showed much higher numbers of new neighbours, but this is likely due to track breaking – as a cell which appears as multiple, separate tracks over time will be counted as multiple new neighbours.

Given this link between track duration and cells being counted as "new" cells in the neighbourhood, it was important to check that the observed trends were not simply

an artefact of a possible correlation between track duration over track start position (e.g. if average track duration was much lower in the posterior). However, there was no clear correlation between track duration and track start position (Figure 21C), and certainly no trend matching that of cell mixing along the axis.

The consistent similarities between the cell tracks and the photolabels confirms the reliability of the cell tracks. Even in the movies with short track durations (M1) and lower track accuracy (M2), while the trend was less pronounced, it was still present.

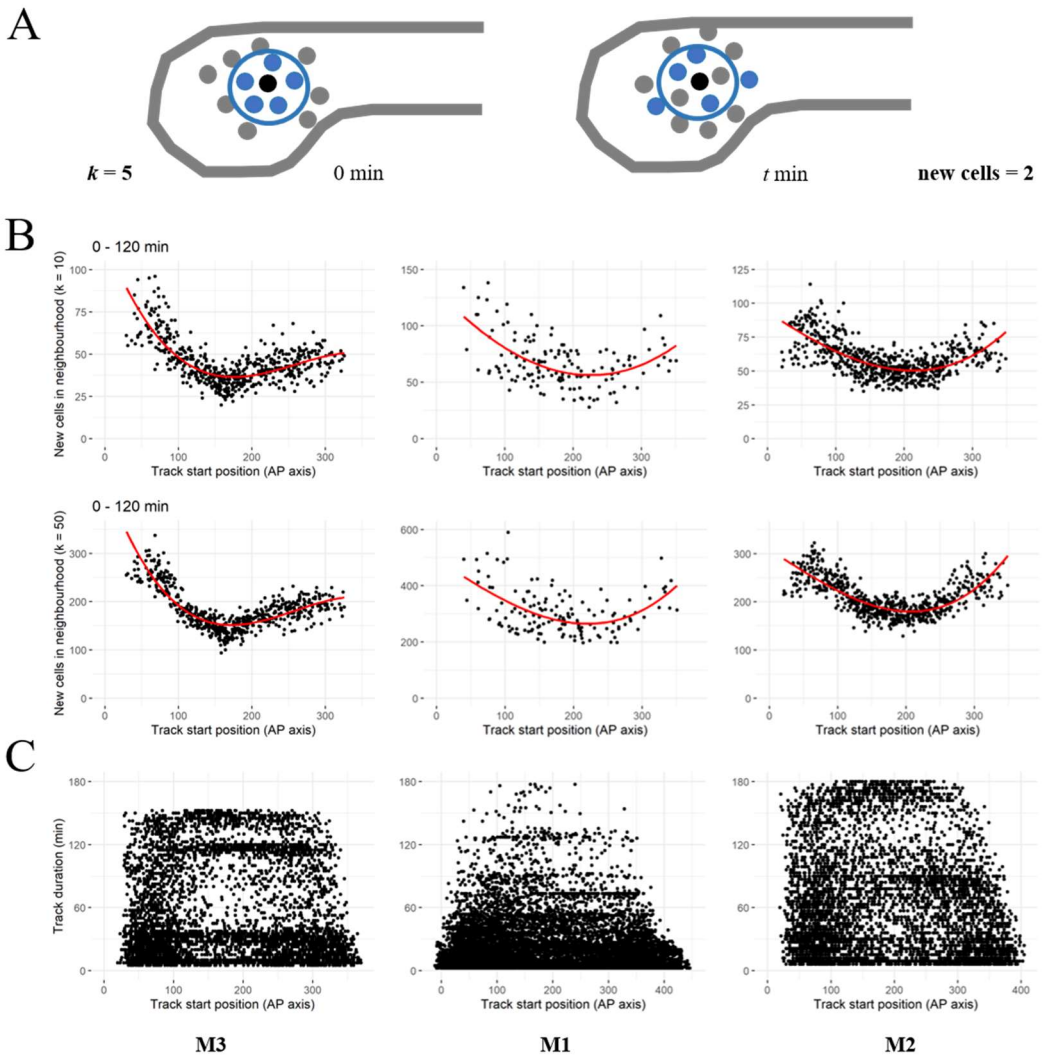


Figure 21: Neighbourhood analysis. Analysis was performed using MATLAB scripts written and provided by Leila Muresan. (A) Schematic of cells in the PSM (grey outline), showing the analysis approach. Taking track positions, a neighbourhood (blue circle) was defined for each cell (black), as the k nearest cells (blue) ($k = 5$ in this example). This was done for every time frame. Over a specified time period t, the number of initially non-neighbour cells (grey) that became neighbours was counted. In this example, over t min, 2 new cells entered the neighbourhood of the main cell. (B) Neighbourhood analysis results for all three movies (left to right), over a time period of 120 min, for small ($k = 10$, top row) and large ($k = 50$, bottom row) neighbourhoods (note different y-axis scales). Each point represents the neighbourhood of one cell. The number of new cells which entered that cell's neighbourhood is plotted against the starting position (AP) of that cell (0 = posterior). The results show a strong gradient of cell mixing, from the posterior (high mixing) to the middle (low mixing) of the paraxial mesoderm. From the middle to the anterior, there is a slight increase, likely related to somite morphogenesis. These results are highly similar to those of photolabelling experiments, confirming the reliability of cell tracks. (C) Track duration plotted against track start position, for each movie, to confirm that these patterns are not an artefact of track durations.

2.3 Anterior movement is an artefact of elongation

To test the idea that cells are actively migrating/moving anteriorly, as suggested by Manning and Kimelman (2015), I measured the AP displacement of tracks over time. I used three reference frames along the paraxial mesoderm: the tailbud tip, the posterior end of the notochord proper, and the posterior boundary of the start-of-movie nascent somite.

The results do indeed show that, relative to the tailbud tip, almost all cells displace anteriorly. However, relative to the nascent somite, almost all cells displace posteriorly. These trends are consistent across all three movies. The results relative to the notochord end are the most variable between movies, likely due to this reference point being the least exact (and therefore the most likely to differ between movies). In all movies, all cells anterior to the notochord reference frame displace anteriorly, but in M3 and M1, cells posterior to the notochord show roughly half displacing posteriorly, and roughly half displacing anteriorly. In M2, almost all cells displace anteriorly. The results are shown in Figure 22 for all three movies.

These results highlight the difficulty of defining movement as anterior or posterior in an elongating tissue. While it is certainly the case that cells disperse along the AP axis, defining this dispersal as anterior vs posterior is purely subjective, based on which end of the embryo displacement is measured relative to. Given that the anterior end of the tissue/embryo is equally valid as a reference point as the posterior end is, there is no reason to define the AP dispersal as anterior movement.

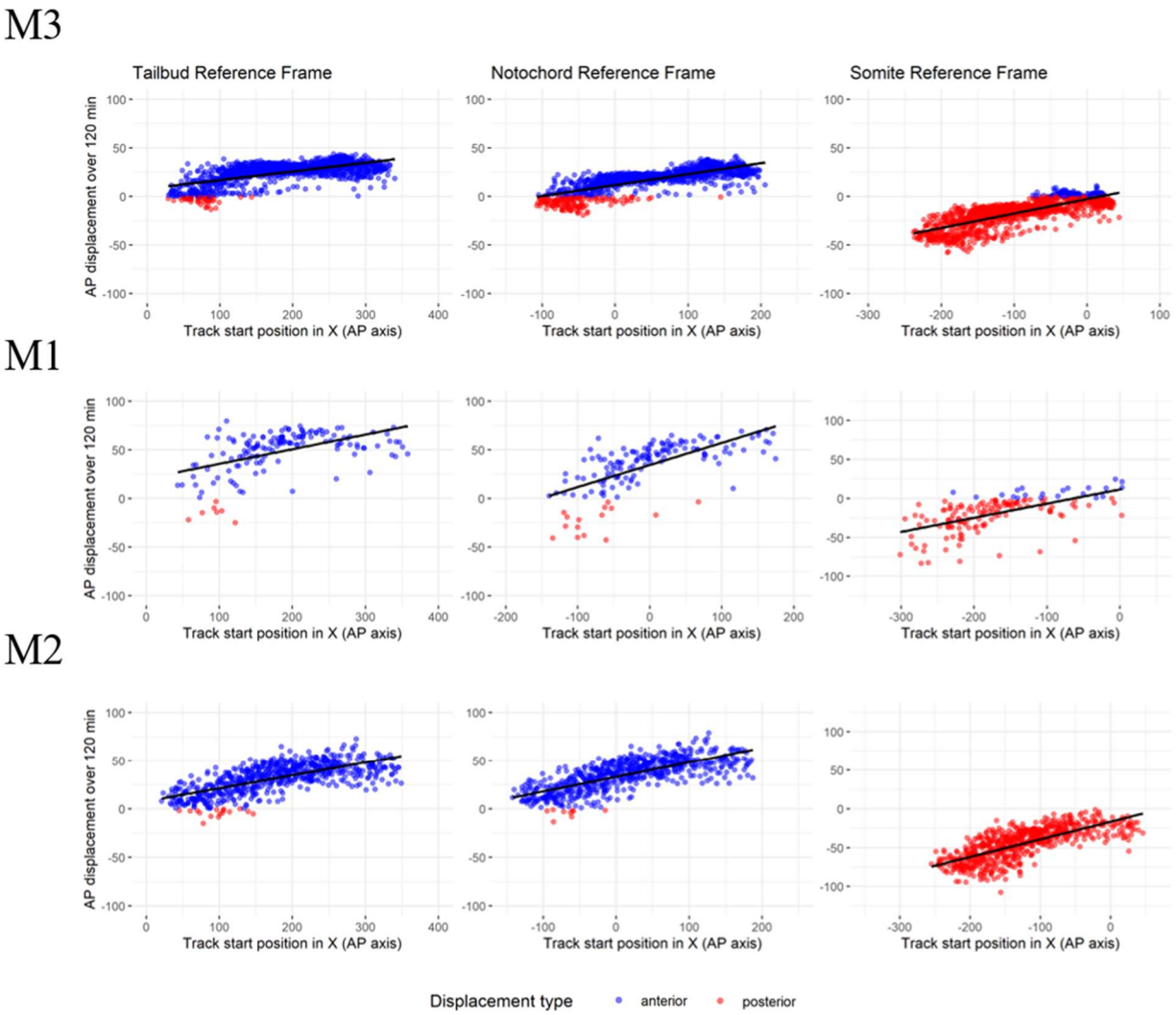


Figure 22: Anteroposterior displacement is relative to the reference frame. Measurements of AP displacement of each track over 120 min are shown for all three movies (top to bottom), relative to three different reference frames: tailbud tip; notochord proper end; and the start-of-movie nascent somite (left to right). AP displacement is shown on the y-axis, with positive values indicating anterior displacement, and negative values indicating posterior displacement (points are also colour coded for displacement type to highlight this). This is plotted against track start position, where 0 is the reference frame position (negative values are posterior to this, positive values are anterior). The data shows that almost all cells move anteriorly relative to the tailbud tip, but also that almost all cells move posteriorly relative to the nascent somite, with displacement relative to the notochord showing similar results to that of the tailbud tip, but with slightly more posterior displacement. All trendlines (black) show a positive correlation i.e. anterior cells move more anteriorly while posterior cells move more posteriorly.

2.4 Compression-extension involves convergent flows in two axes

As my morphometric data showed convergence at the tissue level, and photolabels confirmed this (at least for the DV axis), I measured track displacements in the DV and ML axes. For this analysis, I focussed on the posterior PSM, as this region showed much greater convergence than the anterior PSM (Figure 11C). All tracks which began in the first frame and were, at this timepoint, located between the tailbud tip and the notochord proper, were selected for analysis. This selection is shown from both lateral view (Figure 23, top) and dorsal view (Figure 24, top). For both DV and ML analyses, the tailbud tip reference frame was used, as this was the best reference frame to give a constant, correct midline for this region.

DV displacement was measured over different timescales: 30, 60, and 120 min (Figure 23). The results show that posterior PSM cells converge in the dorsoventral axis, with ventral cells displacing dorsally and dorsal cells displacing ventrally. However, this trend is only clear over longer timescales. At short timescales (30 min), while the correlation between track start position (DV) and track displacement (DV) is negative, it is a very weak correlation. By 120 min, this negative correlation is much stronger.

ML displacement was also measured over these timescales: 30, 60, and 120 mins (Figure 24). The results show a weaker convergence than those of DV results; but do show a slight convergence in the ML axis, with medial cells displacing laterally and lateral cells displacing medially. Again, this trend becomes clearer over time, although even after 120 min, the negative correlation between track start position (ML) and track displacement (ML) is still weak. Overall, more cells displace medially, suggesting that cells are displacing to the midline of the embryo, rather than midline of the lateral half of the PSM.

These results show that compression-extension involves convergent flows of cells towards the midline (in both DV and ML). The difference in the degree of convergence between the DV and ML axes is also consistent with the morphometric results Figure 10, which showed a greater decrease in tissue height than in width over time.

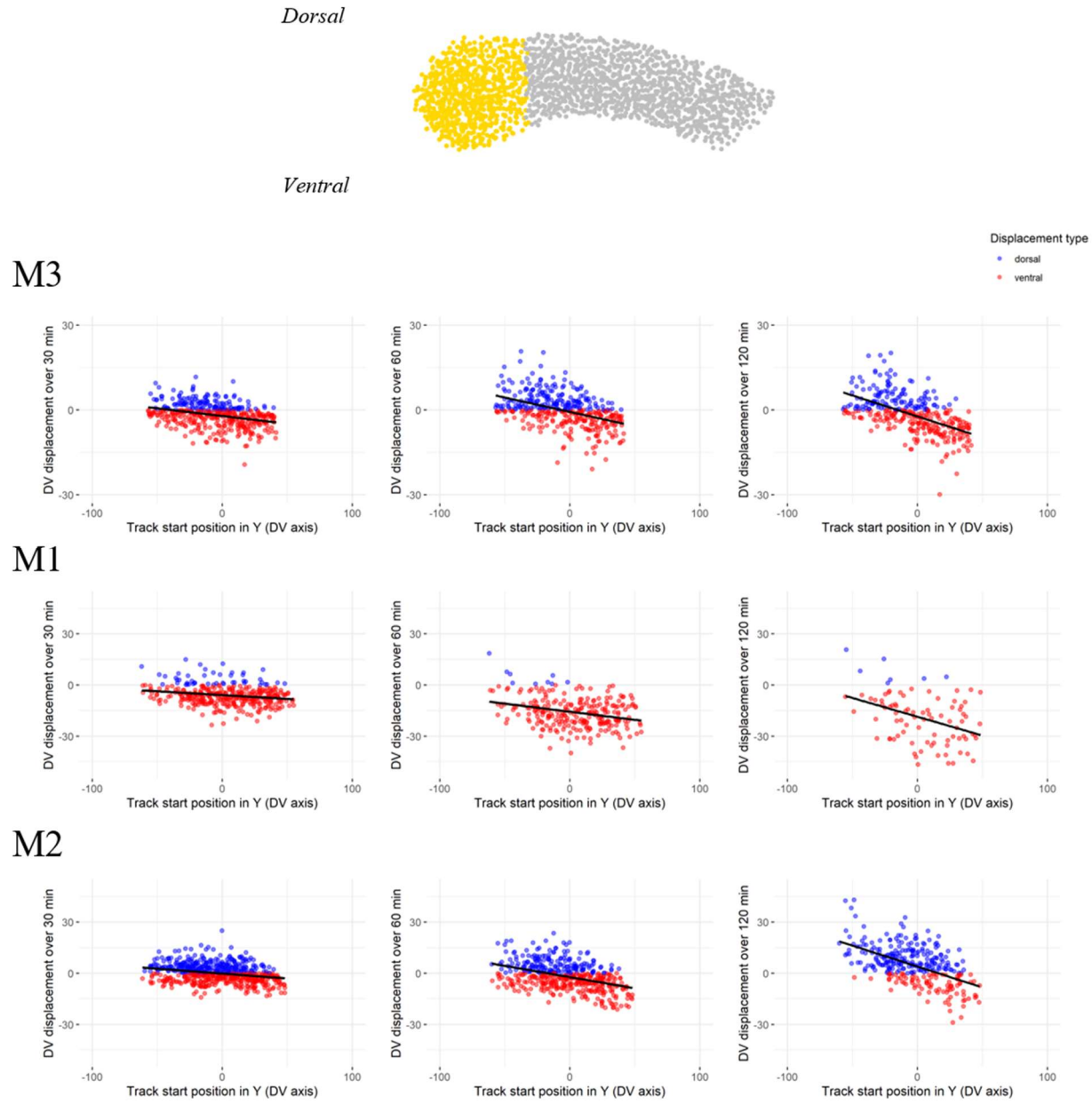


Figure 23: Posterior PSM cells show dorsoventral convergence over time. Top reconstruction shows posterior PSM cell selection (yellow), with the DV axis labelled. Measurements of DV displacement of each track over 30, 60, and 120 min (left to right) are shown for all three movies (top to bottom). DV displacement is shown on the y-axis, with positive values indicating dorsal displacement, and negative values indicating ventral displacement (points are also colour coded for displacement type to highlight this). This is plotted against track start position, where 0 is the tailbud midline (negative values are ventral to this, positive values are dorsal). All trendlines (black) show a negative correlation i.e. dorsal cells move ventrally while ventral cells move dorsally. This correlation increases over time.

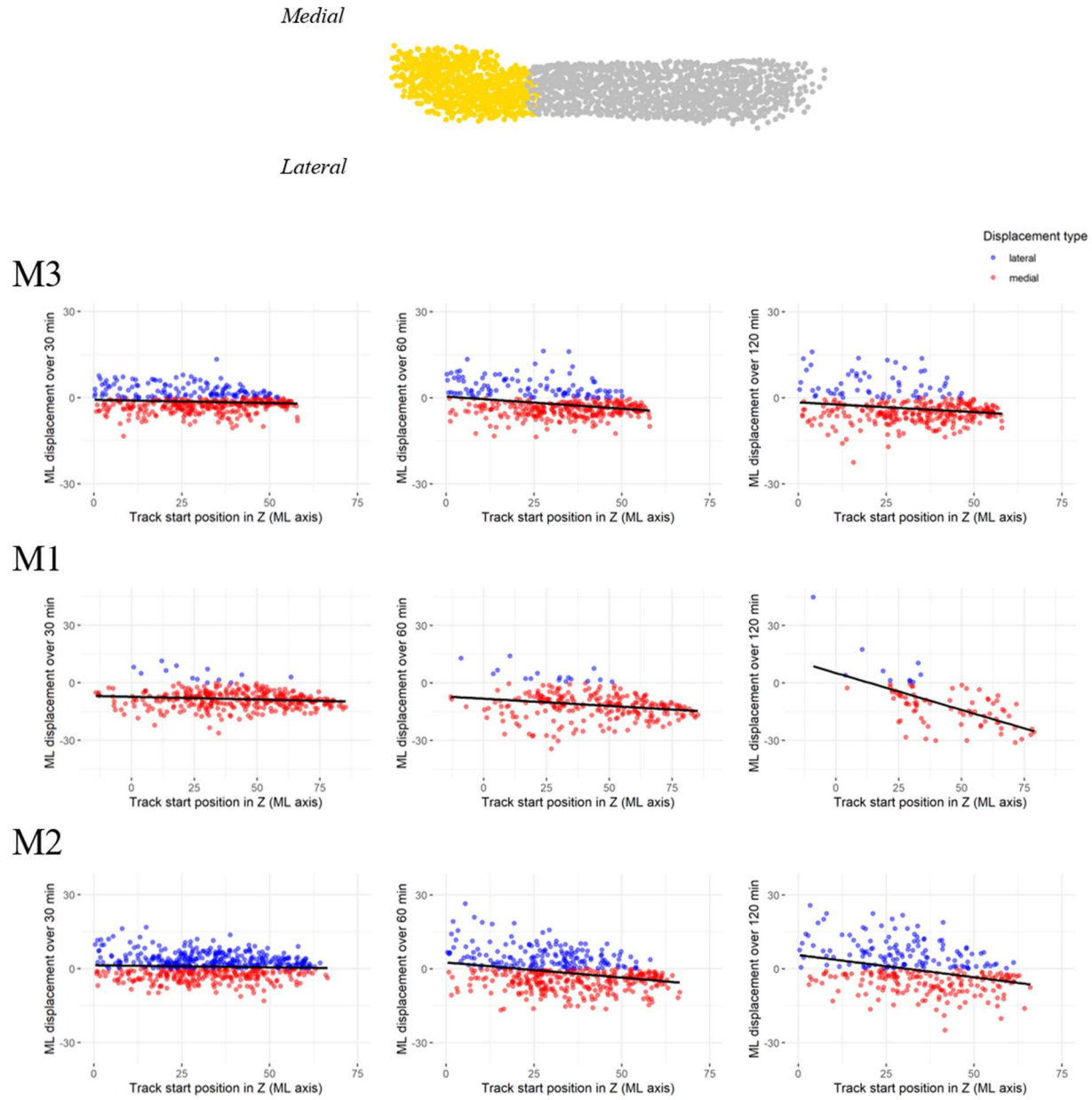


Figure 24: Posterior PSM cells show weak mediolateral convergence over time. Top reconstruction shows posterior PSM cell selection (yellow), with the ML axis labelled. Measurements of ML displacement of each track over 30, 60, and 120 min (left to right) are shown for all three movies (top to bottom). ML displacement is shown on the y-axis, with positive values indicating lateral displacement, and negative values indicating medial displacement (points are also colour coded for displacement type to highlight this). This is plotted against track start position, where 0 is the tailbud midline (positive values are lateral to this). All trendlines (black) show a weak negative correlation, as most cells move medially but lateral cells are more likely to move medially, and medial cells are more likely to move laterally. This negative correlation increases over time.

2.5 Compression-extension does not involve directional intercalation

Classical convergent extension (i.e. active convergent extension) involves directional intercalation between neighbouring cells. In other words, cells rearrange with their neighbours to line up along the AP axis, causing the tissue to decrease in width and increase in length (Keller et al., 2000). If the tissue convergence happening in the zebrafish paraxial mesoderm is due to tissue compression, one would expect that directional intercalation is not involved. To test this, I measured angle changes between neighbours, relative to the AP axis, over time. This was done using a MATLAB script written by and provided by Leila Muresan. "Angles analysis" involves taking the nearest neighbour for each cell at the start of the movie, then taking the vector between those two cells, and measuring the angle between this vector and the AP axis. Then, after a specified time-period *t*, the same two cells are taken (regardless of whether they are still neighbours), and their vector angle relative to the AP axis is measured again. This analysis is illustrated in Figure 25A. In this way, the orientation of cell pairs along the AP axis can be measured over time, to test if neighbours are aligning with each other along the axis to drive elongation.

The results of this analysis are shown in Figure 25B for the posterior PSM of movie M3 (as this movie has high track accuracy and high track duration - both of which are essential for this analysis). Interestingly, while there is substantial rearrangement happening between neighbour pairs, these rearrangements show no evidence of directional intercalation, confirming previous findings by Mongera et al. (2018). Neighbour pairs are not orientating along the AP axis, even when this is measured over long time periods.

In case directional intercalation was happening at the level of groups of cells, rather than at the level of neighbouring cells, I extended the analysis to include the *nth* nearest

neighbour. In other words, rather than taking the angle with the nearest neighbour, the analysis would take the angles with the 10^{th}, 50^{th}, and 100^{th} nearest neighbours. However, rearrangements between the nth nearest neighbours also showed no evidence of directional intercalation. In fact, the angles between further away cells changed less, not more (Figure 25C). This is presumably due to the fact further away cells will have to move a greater distance to change their angle than closer cells will.

These results confirm that cells are not intercalating with their neighbours to drive tissue convergence, and that neighbour rearrangements appear to be randomly orientated. Instead, tissue convergence involves cell displacements towards the midline and non-directional intercalation. This supports the hypothesis that compression is the cause of convergence, as external compression would likely cause midline displacements and random mixing between neighbours. Therefore, these results support the proposed mechanism of compression-extension - and show clear mechanistic differences between this process and that of classical/active convergent extension.

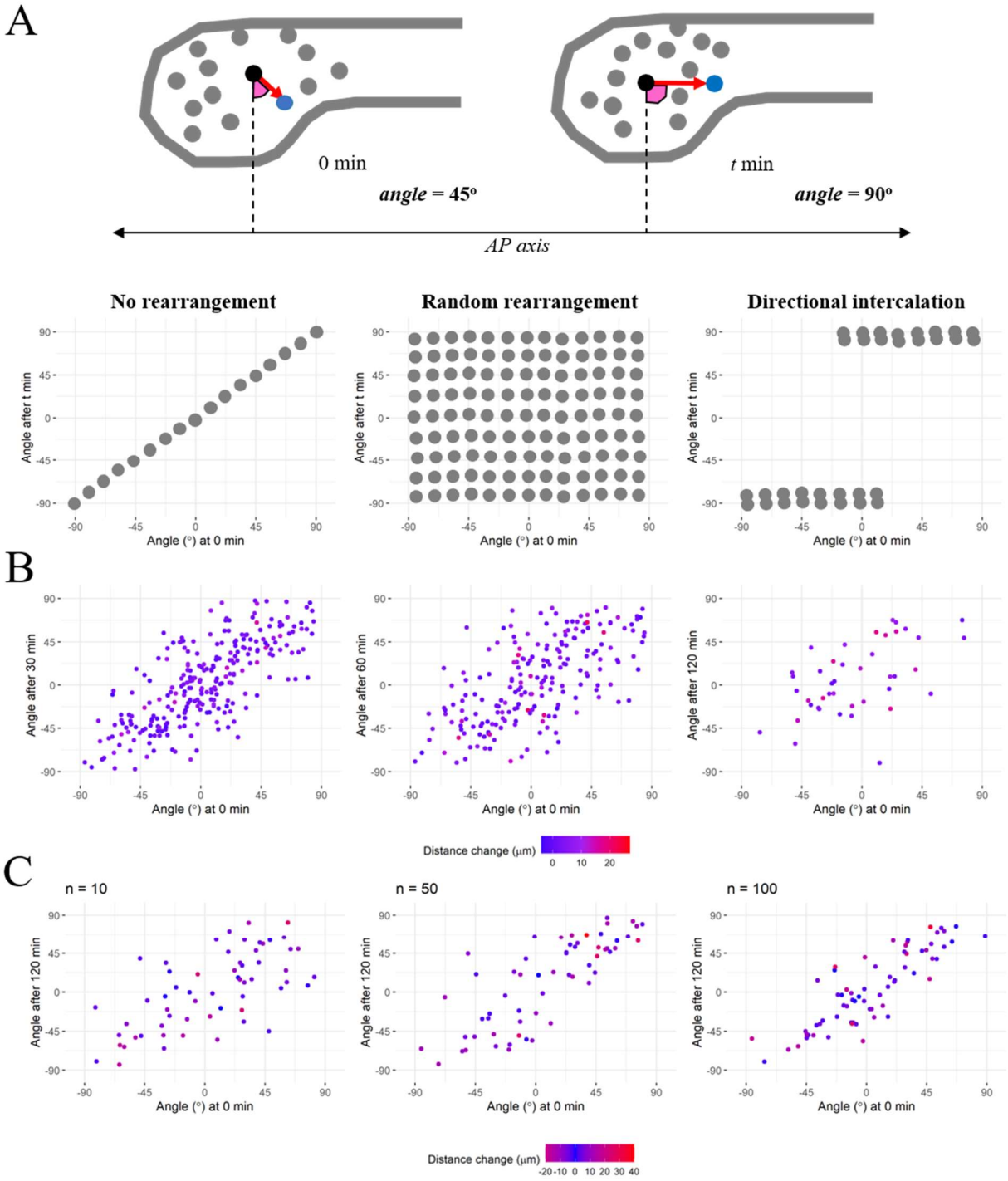

A
0 min
angle = 45°
t min
angle = 90°
AP axis
No rearrangement
Random rearrangement
Directional intercalation
Angle after t min
Angle (°) at 0 min
B
Angle after 30 min
Angle after 60 min
Angle after 120 min
Distance change (μm)
0 10 20
C
n = 10
n = 50
n = 100
Angle after 120 min
Distance change (μm)
-20 -10 0 10 20 30 40

Figure 25 (previous page): Neighbour angles analysis. Analysis was performed using MATLAB scripts written and provided by Leila Muresan. (A) Schematic (top) showing the analysis approach. The angle (pink) between the vector (red) from a cell (black) to its nearest neighbour (blue) and the AP axis is taken at the start of the movie. After a specified amount of time t, this angle is measured again for the same cells. In the example shown, the angle changes from 45° to 90°. These angle changes can be plotted on graphs with the initial angle as the x coordinate, and the later angle as the y coordinate. So, in this example, the coordinate of the black cell on a graph would be (45,90). Three example graphs (bottom) are shown for different tissue behaviours: no rearrangement (angle after t min is the same as the angle at 0 min, i.e. y = x); random rearrangement (no bias towards any angles, and no correlation between initial and later angle); and directional intercalation (cells rearrange such that they lie directly in front or behind their neighbour, i.e. y = 90 & -90, with those already slightly behind tending to move behind, and those already slightly in front tending to move in front). (B) Analysis results shown for the posterior PSM cells (movie M3), at 30, 60, and 120 min (left to right). Neighbour rearrangement increases over time, but rearrangements do not show any bias in orientation, suggesting that they are random. Points are colour coded (bottom) by the change in distance between the cell and its nearest neighbour. Later timepoint graphs show fewer data points due to the limitation of track duration. (C) The analysis was extended the 10^{th}, 50^{th}, and 100^{th} nearest neighbours (left to right), to see if looking at rearrangements between cells over greater distances would show clearer patterns. However, further apart cell pairs rearrange less than close cell pairs.

2.6 A 4D agent-based model of compression-extension

To test the assumption that compression-driven tissue convergence would lead to non-directional intercalation, we designed a 4D agent-based model of compression-extension. The model was built in Python by Berta Verd. An initial tissue was specified as a cube of 2744 spherical cells (14 x 14 x 14). The x-axis corresponds to the "AP" axis of the tissue. At every timepoint t, each cell calculates its shortest vector to the midline (in y and z) along the AP axis of the tissue and moves along this vector (up to a maximum specified distance between stages). This simulates cells being pushed towards the midline by tissue compression. To maintain tissue cohesion, two zones were specified for each cell: a zone of repulsion and a zone of attraction (similar to models of flocking behaviour (Huth and Wissel, 1990)). If, after one timepoint, a cell has moved within the zone of repulsion its neighbour, both cells are displaced equally along their vector until they are no longer within each other's zone of repulsion. Conversely, if a cell has moved too far apart from its nearest neighbour (into the zone of attraction), both cells are pulled back towards each other along their vector until they are no longer within each other's zones of attraction. These two parameters provide an approximation of cell-cell adhesion, by ensuring that cells remain within sensible distances of their neighbours. Importantly, they also allow cells to push other cells out of the way as each cell moves towards the midline. This model is iterative, in that movement vectors are calculated for each timepoint, and at the end of each timepoint, cells adjust their distances with each other based on the zone parameters. Then the process repeats for the next timepoint, until the end.

A simulation of this model with 500 timepoints is shown in Figure 26A. The tissue quickly collapses from a cube to a cylindrical shape. Dimensions of the tissue were measured for each timepoint (taken as the maximum value for a coordinate minus the minimum value of that coordinate, for each timepoint). The length increases throughout the simulation, with an initial rapid increase, followed by a slower, constant increase.

Both height and width show the same pattern as each other: an initial increase (likely due to a loss in density as the tissue initialises), before rapidly returning to the starting dimension, followed a constant decrease. These measurements are shown in Figure 26B.

To confirm that individual cells in the model were moving as expected, I measured track displacement for these cells in the same way as was done for the posterior PSM tracks in Figure 22 - Figure 24. Measurements were taken, not from t0, but from t70 – to exclude the initialisation period when the tissue shape changes drastically from a cube to a cylinder. The results, shown in Figure 26C, confirm that cells converge to the midline (dorsal cells move ventrally and vice versa, cells on the right move to the left and vice versa), and cells disperse along the AP axis (anterior cells move anteriorly, posterior cells move posteriorly).

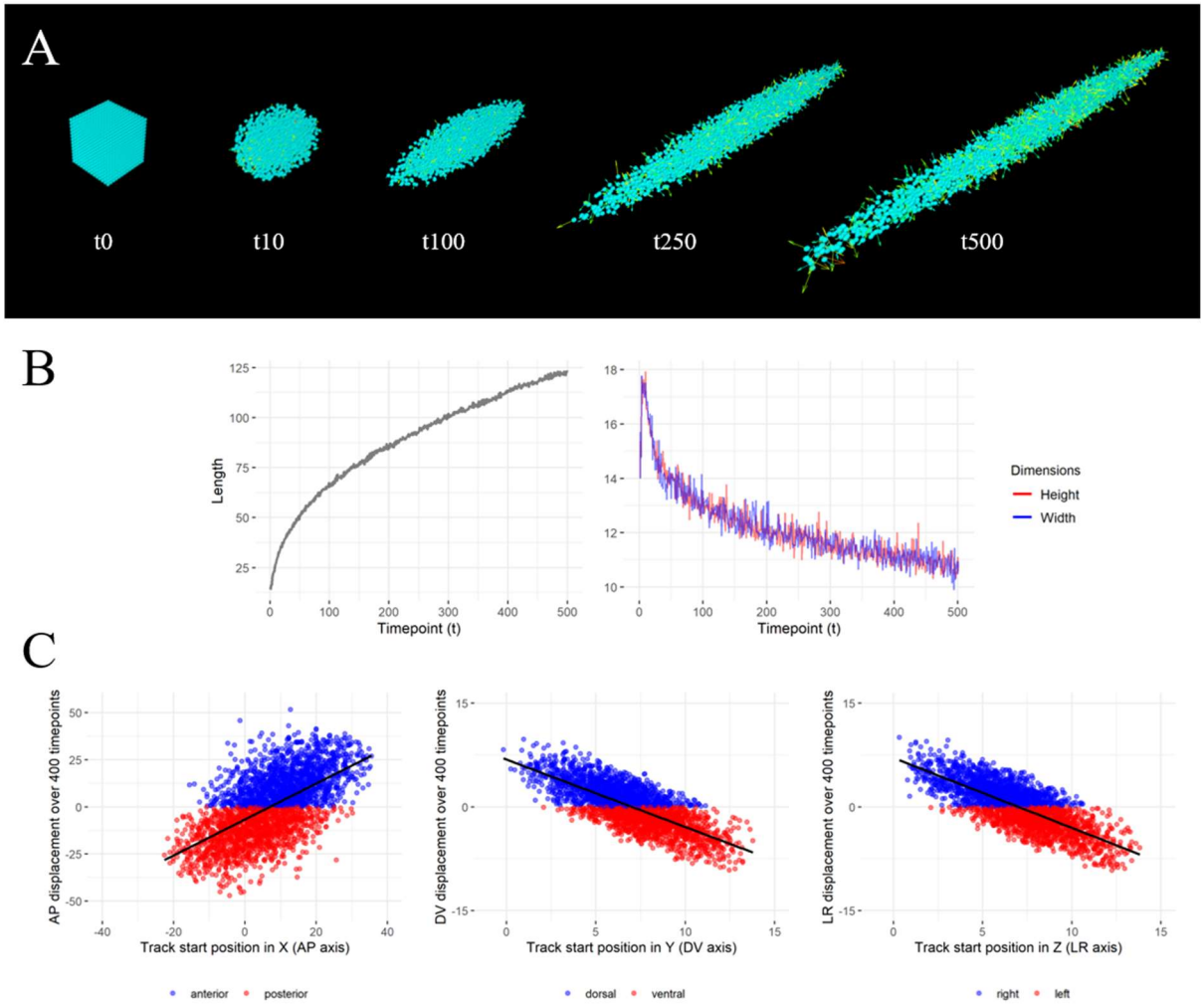


Figure 26: A 4D model of compression-extension. The model was built in Python by Berta Verd and is composed of 2744 individual "cells" that together constitute a tissue. Each timepoint (0-500), every cell calculates its shortest distance to the midline (y, z) and moves towards that point, while maintaining semi-fixed distances to other cells, meaning cells can displace each other. (A) Snapshots of one simulation of the model at timepoints 0, 10, 100, 250, and 500. (B) Measurements of "tissue" dimensions. The dimension for each timepoint is taken as the highest coordinate value minus the lowest coordinate value (e.g. the length of the tissue at any timepoint is the highest x value at that timepoint minus the lowest x value at that timepoint). Length increases over time while height and width decrease over time (aside from a brief increase in the first few timepoints, which are likely due to model initialisation). (C) Displacements of individual "tracks" from t70 to t470 are shown for AP, DV, and LR axes (left to right graphs). Displacements are colour-coded for displacement type (see below each graph). AP displacement shows a positive correlation (black trendline) with initial position, meaning anterior cells displace anteriorly, while posterior cells displace posteriorly. DV and LR displacements show negative correlations with initial position, meaning dorsal cells move ventrally (and vice versa), and cells on the right move to the left (and vice versa).

Having confirmed that the model behaved as expected (after a brief initialisation phase), I then performed the same nearest neighbour angles analysis on the model tracks as I previously performed on posterior PSM tracks (Figure 25). As with measuring model track displacements, this was done from t70 onwards to exclude the initialisation period (Figure 27A).

Neighbour angle changes over 50 and 100 timepoints (representing 20% and 40% elongation of the tissue) appeared fully random. Neighbour pairs were not orientating along the AP axis, and there was no correlation between initial angle and later angle. Over much longer timescales of 200 and 400 timepoints (representing 65% and 110% elongation of the tissue), a pattern of AP orientation did appear (Figure 27B). After 400 timepoints, most neighbour pair angles were -90^{o} or 90^{o} (in other words, the two cells were lined up along the AP axis). However, it is important to note that by this stage these cells are no longer neighbours – those pairs of cells that line up along the AP axis are those that move the furthest away from each other (as shown by the colour code in Figure 27B). This shows that neighbour cells are not rearranging with each other in a "productive" manner. Rather, they are all moving towards the midline; displacing and pulling each other as they do so, leading to highly random/chaotic neighbour rearrangements. Because the tissue is getting thinner, cell pairs that move far away from each other will inevitably line up along the AP axis, but this is due to changes in tissue dimensions, rather than any directional intercalation between neighbours. This is further highlighted by the fact that cell pairs that already lie along the AP axis (e.g. an angle of 90^{o}) appear to be just as likely to stay at this angle (90^{o}) as they are to swap with one another (i.e. to give a later angle of -90^{o}). In other words, there is no correlation between initial angle and later angle, even for those neighbour pairs already at the x = -90 or x = 90 lines.

As with the posterior PSM, extending the analysis to the nth nearest neighbour did not cause the AP orientation trend to appear at earlier timepoints (Figure 27C).

Although extending the analysis to the 200th neighbour did appear to show a concentration of angles at -90^{o}/90^{o}, this was the case for both the initial angle (x-axis) and the later angle (y-axis). Presumably because, in a tissue which is longer than it is wide, a cell that is far away enough to be the 200th neighbour is more likely to be far along the AP axis (and therefore the initial angle more likely to be -90^{o}/90^{o}).

It is important to note that this model is not a perfect model of compression-extension, particularly in that cell volumes remain constant. However, it does show that convergence of cells towards a midline does not require directional intercalation to drive tissue elongation. This supports the idea that external compression of the paraxial mesoderm - provided this caused cells to displace towards the midline - could drive tissue elongation without any requirement for directional intercalation. Taken together with the results of previous sections, this suggests that while active convergent extension is driven by cell-level mechanisms of directional intercalation, compression-extension is driven by tissue-level mechanisms of compaction - which causes convergence of cells towards the midline and chaotic/random rearrangements between neighbours.

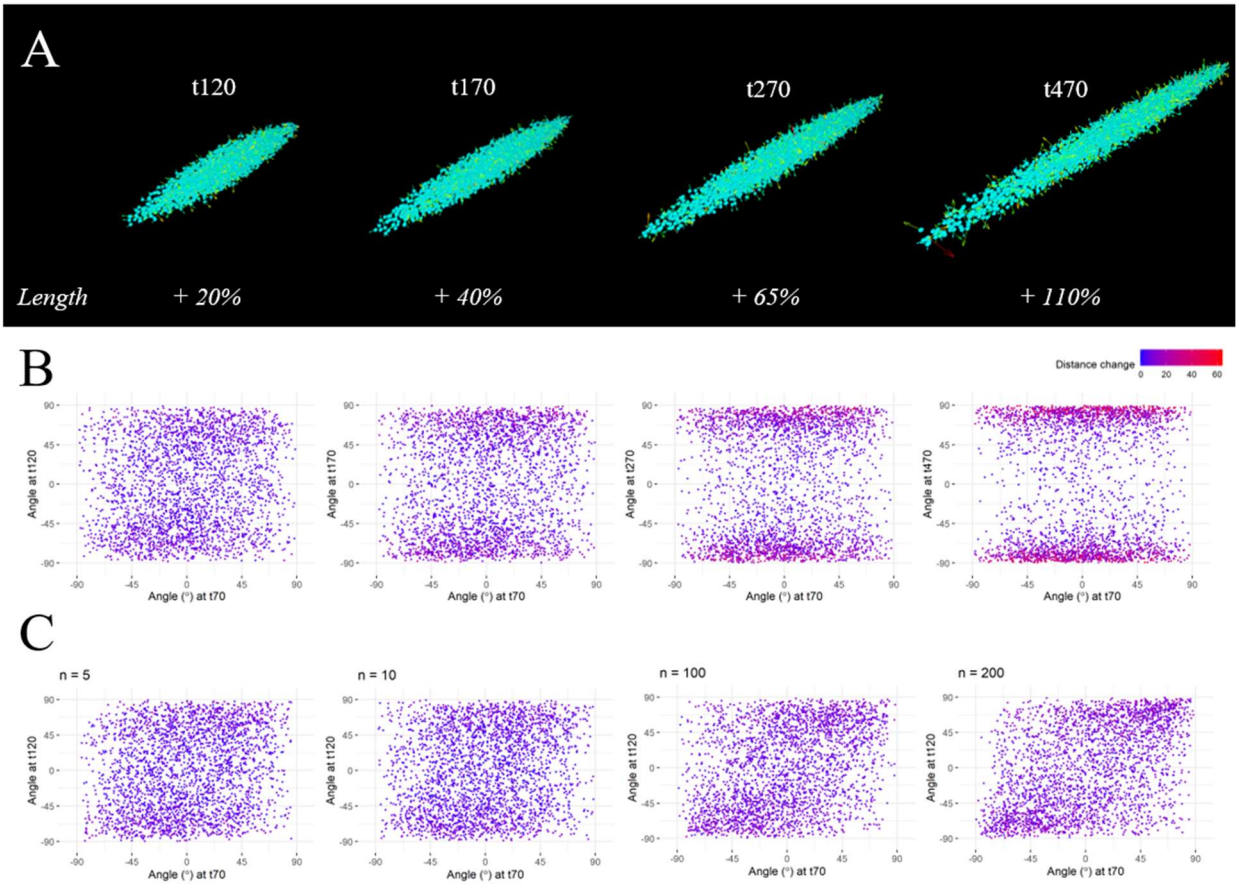


Figure 27: Compression extension involves non-directional intercalation. (A) Taking t70 of the model simulation as the start timepoint, the timepoints chosen for angles analysis are shown. The length increase between t70 and each timepoint is also shown (bottom), as a percentage of initial t70 length. Angles analysis was performed as previously described in Figure 25. (B) Neighbour pair angle changes over different time periods. Despite significant length increases, t120 and t170 graphs show random rearrangement of cell pairs. Only at much later timepoints (t270 and t470) are directional rearrangements observed. These rearrangements are not neighbour rearrangements, as these cells are no longer neighbours (as shown by the colour code, top). (C) Extending the analysis to the 5th, 10th, 100th, and 200th nearest neighbours does not cause the trend to appear sooner (all graphs show angles at t120), although initial angles are more likely to be -90°/90° for far away cells, given that the tissue is substantially longer than it is wide at t70.

2.7 Conclusions

The results of tracking cells in the PSM confirm that cells disperse along the AP axis, but that defining this as "posterior" or "anterior" movement is purely based on the chosen reference point, as it is an elongating tissue. The tracks also confirm the findings of the previous chapter; that posterior PSM cells displace towards the midline in both the DV and the ML axes, and that DV convergence is greater than ML convergence. These patterns are slight, and therefore are only observable over longer timescales than are often considered. However, even over long timescales, neighbour rearrangements show no evidence of directional intercalation – instead, neighbour pairs appear to orientate randomly with respect to the AP axis. An agent-based model of compression-extension accurately recapitulates these patterns: convergent displacements towards the midline without any directional intercalation. This supports the idea that external compression of the tissue drives elongation through the proposed mechanism of compression-extension.

RESULTS

3 Investigating compression-extension mechanisms

3.1 Tissue convergence is not PCP-dependent

The results of previous chapters have shown that the paraxial mesoderm elongates through a form of convergent extension that does not involve directional intercalation. I propose that this is driven by external compression of the tissue, which also results in increased tissue density, through a process of compression-extension. During zebrafish gastrulation, cells converge across the yolk and extend to form the tailbud-stage embryo. This form of convergent extension is dependent on the Planar Cell Polarity (PCP) pathway to facilitate directional intercalation (Roszko, Sawada and Solnica-Krezel, 2009). The hypothesis that paraxial mesoderm convergence during somitogenesis is driven by tissue-level compression, rather than by cell-level directional intercalation, would suggest that the PCP pathway is not involved. To test this, I injected mRNA for a dominant-negative form of *Xenopus* Dishevelled (DEP+) into one-cell-stage zebrafish embryos. This form of Dishevelled can still function normally in canonical Wnt signalling, but not in the PCP pathway, due to deletion of the DEP domain (Tada and Smith, 2000). Injections resulted in post-gastrulation deformities (Figure 28A). However, despite the paraxial mesoderm being shorter between injected and uninjected embryos, the elongation rate was highly similar (Figure 28B). This suggests that, while convergent movements involved in paraxial mesoderm *formation* (during gastrulation) were perturbed, those driving paraxial mesoderm *elongation* were not affected. This is supported by the observation that, in injected embryos, anterior somites (i.e. those formed during gastrulation) showed abnormal shape, whereas posterior somites (i.e. those formed during axis elongation) showed normal shape (Figure 28A).

These results show that tissue convergence during paraxial mesoderm elongation is not simply a continuation of earlier gastrulation movements, and instead represents a different form of convergent extension.

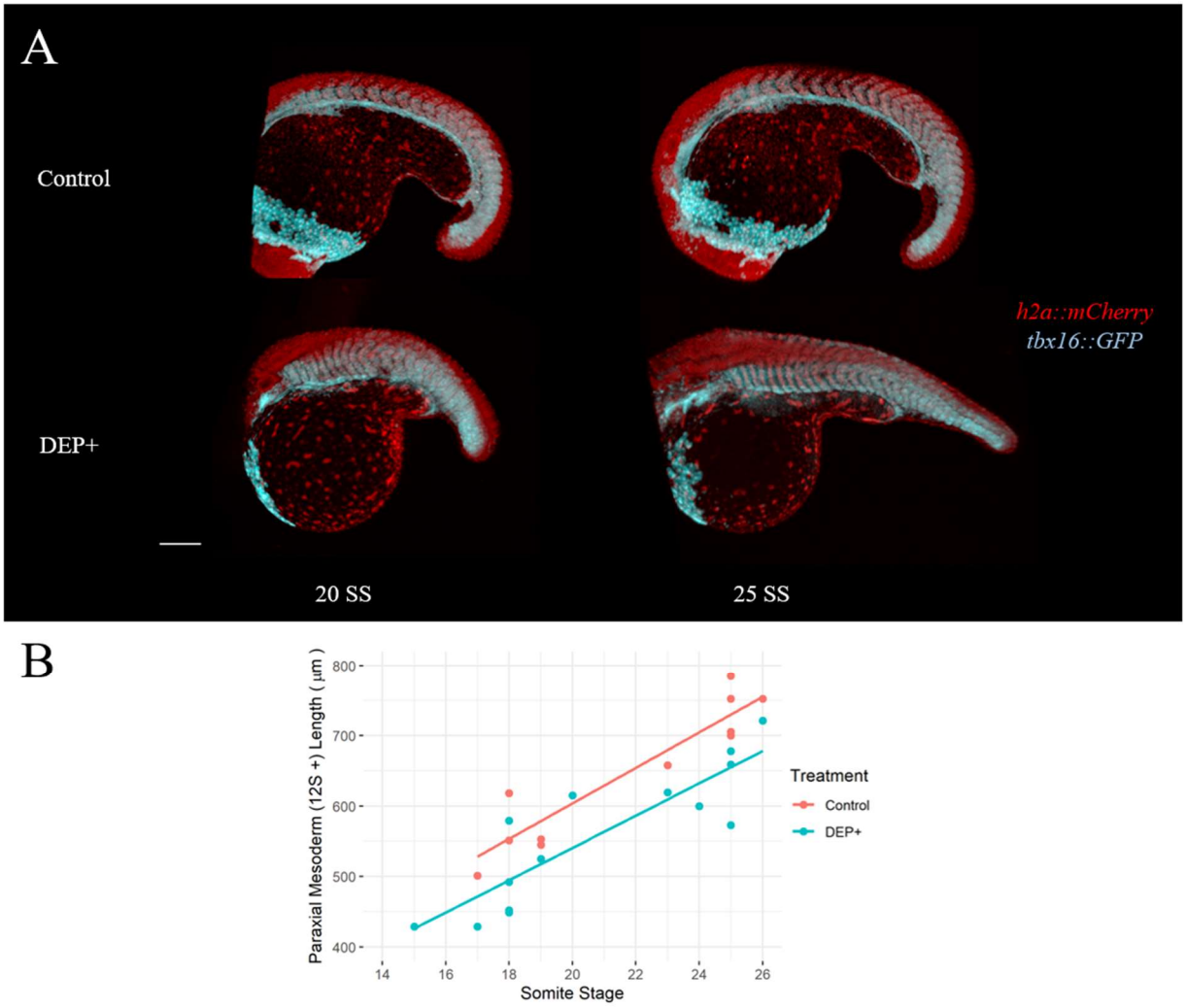


Figure 28: DEP+ injection does not affect paraxial mesoderm elongation rate. (A) *h2a::*mCherry x *tbx16::*GFP embryos were either injected with mRNA for a dominant-negative form of Dishevelled (DEP+, n = 7) which inhibits the PCP pathway, or were left uninjected (Control, n = 6). Embryos were then live imaged for 3.5 hours. Anterior is left in all images. Scale bar: 150 µm. (B) Paraxial mesoderm length, from the 12th somite to the end of the PSM, was measured for each embryo at the start and at the end of imaging. The results show that DEP+ embryos have a shorter paraxial mesoderm, but that the elongation rate of this tissue is unaffected.

3.2 Possible extrinsic compression by the neural tube

The proposal of compression-extension raises the obvious question: what is causing this compression? Particularly, is this compression intrinsic to the tissue itself, or is it extrinsic – caused by other tissues? There are a few candidate tissues that could be causing extrinsic compression of the PSM. One possibility is the notochord, which, at least in chicks, pushes into the posterior PSM (Xiong et al., 2020). However, recent notochord ablation experiments have shown that ablating large regions of the zebrafish notochord has no effect on paraxial mesoderm elongation (Mclaren and Steventon, 2021). The other main tissue that is in contact with the PSM is the neural tube. As this tissue sits dorsal to the PSM, it could be exerting a pushing force down onto the PSM, driving compression and convergence (at least in the DV axis). To test this, I used two approaches to try to perturb potential pushing by the neural tube on the PSM.

Neural tube laser ablations

The first approach I used was to ablate the posterior neural tube. Ablations were performed by Dillan Saunders on a two-photon microscope at the 12 somite-stage. A region of cells in the posterior neural tube approximately 150 µm in length (AP) and 30 µm in both height and width (DV, ML) was manually selected and then ablated with high laser power (see Figure 29A). Ablated embryos (and stage-matched non-ablated control embryos) were then left to develop for a further 5 hours before being fixed. After fixing, I performed HCR for *tbx16* as a marker for the posterior PSM, in order to make manual PSM surfaces in Imaris (Figure 29B). I used these surfaces to mask nuclear signal and make spots as before, to calculate the density of the PSM (cell number/tissue volume). I also measured the height of the posterior PSM.

The results (Figure 29C) show non-significant differences between control and ablated embryos in both PSM density ($t = 0.76$, $p = 0.47$) and PSM progenitor height ($t = 1.45$, $p = 0.18$). This could suggest that the neural tube is not causing compression

or height reduction of the PSM. However, measurements of unsegmented region neural tube volumes showed no effect of ablation on tissue volume ($t = -0.46$, $p = 0.66$) (Figure 29C, right graph). Therefore, while laser ablation caused high cell death in the neural tube (as shown by the loss of nuclear fluorescence and high prevalence of pycnotic nuclei in the tailbud (Figure 29A,B)), it is still possible that the tissue is pushing on the PSM - if the growing anterior neural tube is pushing the dead cells of the posterior neural tube into the PSM. Therefore, in the next section, I used another approach to perturb possible pushing on the PSM by the neural tube.

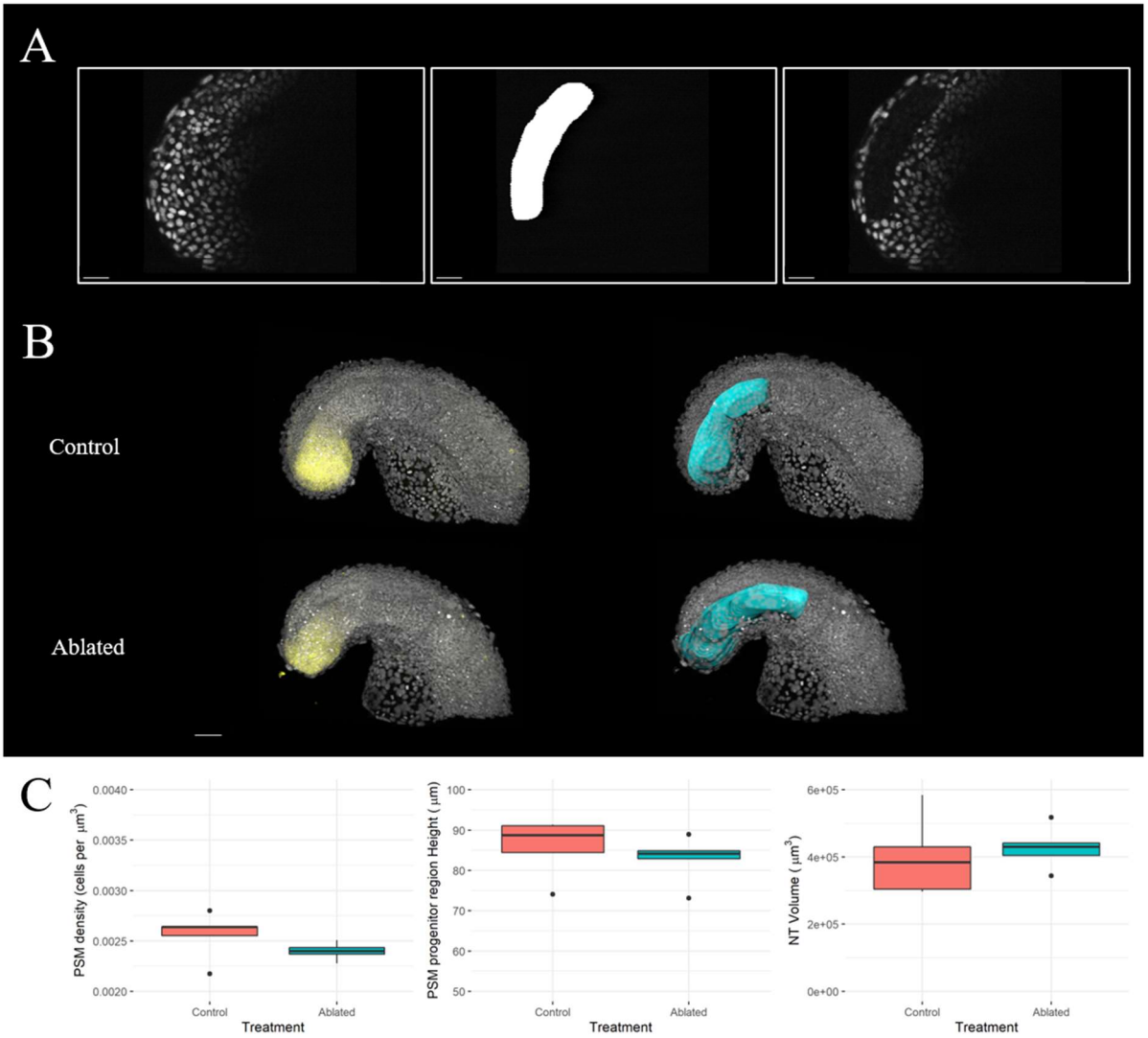


Figure 29: Neural tube laser ablations do not significantly affect PSM density or convergence. (A) Individual z-slices before, during, and after and ablation of the posterior neural tube (left to right). Ablations were performed on *h2b::GFP* (grey) embryos by Dillan Saunders, who manually selected the laser ablation region (white, middle image). After ablation, fluorescence is absent from this region. Anterior is top, dorsal is left. Scale bars: 30 µm. (B) Control (n = 5) and ablated (n = 5) embryos were stained using HCR (left) for *tbx16* (yellow), with nuclei stained using DAPI (grey). High cell death is apparent in ablated embryos, as shown by pycnotic nuclei in the DAPI channel. Manual surfaces of the PSM (cyan) were then reconstructed as before (right). Scale bar: 50 µm (C) Neither PSM density (left graph) nor PSM progenitor height (middle graph) are significantly different between control and ablated embryos. However, unsegmented region neural tube (NT) volume (right graph) was also not affected by neural tube ablation, and so the hypothesis that the neural tube exerts a compressive force on the PSM cannot be ruled out from these results. See text for t-test results.

Cdx4 mutants

cdx4 is expressed along the posterior neural tube and generally throughout the tailbud (Shimizu et al., 2005). This gene plays a role in axis elongation – as shown by the fact that zebrafish mutants have a truncated AP axis (Davidson et al., 2003; Davidson and Zon, 2006). This suggests that this gene is involved in neural tube growth/elongation. *Cdx4* mutants, therefore, could be used to test the idea the neural tube is compressing PSM during axis elongation. As these mutants show reduced elongation, one would expect PSM density to stay constant, or at least decrease by less, if the neural tube is driving PSM compression-extension.

To test this, I fixed wild-type (WT) and *cdx4*- embryos at two stages of somitogenesis (20 and 30 somite-stages). Embryos were stained with DAPI and phalloidin and manual PSM surfaces were made in Imaris (Figure 30A). PSM density was measured as before, as was the paraxial mesoderm length (from the 16th somite to the posterior end of the PSM).

The results (Figure 30B) show a reduction in paraxial mesoderm length at both stages in mutants compared with wild-types, suggesting that *cdx4* is, in some way, important for paraxial mesoderm elongation, possibly through neural tube elongation. On the other hand, given that the reduction at both stages is similar, it could be argued that elongation rate is unaffected. However, as these measurements were taken from fixed embryos (which displayed a wide range of phenotypes), this is not clear. Paraxial mesoderm length would need to be measured in live embryos to better test this, but fixing was required to accurately measure PSM density. These PSM density measurements, however, also gave inconclusive results. Density is not significantly different between groups at the 20 somite-stage ($t = -1.34$, $p = 0.22$) or at the 30 somite-stage ($t = 0.92$, $p = 0.39$), suggesting that PSM density is not affected by the loss of *cdx4*. However, the trends over time/between stages are different between wild-types and mutants. WT embryos showed a significant difference in density between the 20 and

30 somite-stages (t = -4.26, p = 0.003), whereas *cdx4* mutants did not (t = -0.76, p = 0.47). This could suggest that *cdx4* loss prevents normal PSM density increase. However, given that at each stage there is no clear significant difference in density, this conclusion is only tentative. It is possible that multiple effects of the loss of *cdx4*, as well as variation in phenotype strength, is confounding the results. A higher sample size, combined with live imaging experiments, may help clarify the effects of *cdx4* loss on paraxial mesoderm compression-extension.

These results, therefore, are somewhat inconclusive. They suggest that normal neural tube development/elongation is required for normal paraxial mesoderm elongation and normal density increase – tentatively supporting the hypothesis that the neural tube contributes to paraxial mesoderm compression-extension. However, given that *cdx4* is expressed throughout the tailbud (including in the posterior PSM), a possible intrinsic role of *cdx4* in paraxial mesoderm elongation cannot be ruled out.

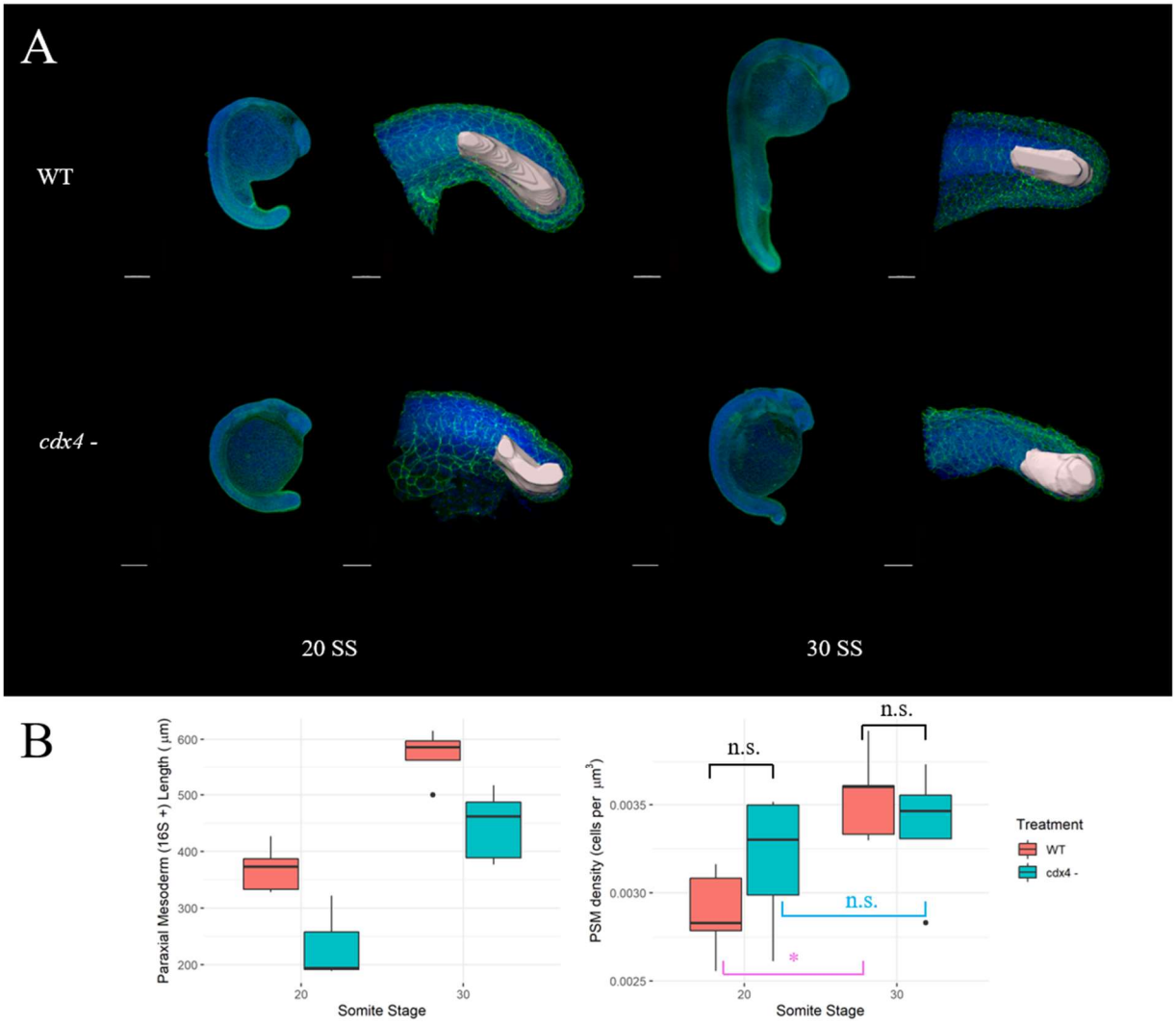


Figure 30: *cdx4* mutants show reduced paraxial mesoderm length and no density differences between stages. (A) Wild type (WT) and *cdx4* mutant (cdx4 -) embryos were fixed at the 20 and 30 somite-stages (n = 5 per group per stage). Embryos were stained with DAPI (blue) and phalloidin (green), and manual PSM surfaces (white) were made in Imaris. For each group and stage, left images show whole embryo imaged at 10x (scale bars: 200 µm) and right images show the tailbud imaged at 40x (scale bars: 50 µm) with the PSM surface shown. (B) Paraxial mesoderm length (measured from the 16th somite to the end of the PSM) is shorter in *cdx4* mutants at both stages (left graph). PSM density (right graph) is not significantly different between groups at each stage (black brackets), but WT embryos show a significant increase in density (pink bracket) which mutants do not show (cyan bracket) (see text for t-test results).

3.3 TGFβ signalling regulates tailbud density

During an initially unrelated project, some interesting results were observed that may relate to the questions asked in this chapter. When treating zebrafish embryos with SB505124 ("SB50") during early somitogenesis, I observed that some embryos developed tumour-like swellings in the tailbud by the end of somitogenesis (Figure 31A,B). This phenotype was highly variable: most embryos showed no visible phenotype, and the appearance of the phenotype varied in rate between experimental repeats (ranging from 1 in 20 to 1 in 5). These swellings were not specific to one tissue, sometimes appearing in the paraxial mesoderm and sometimes appearing in the neural tube (Figure 31C).

SB50 is often described as a Nodal signalling inhibitor (Hagos and Dougan, 2007). However, it competitively binds to ALK-4, ALK-5, and ALK-7 membrane receptors, which are involved in TGFβ signalling as well as Nodal signalling (Casari et al., 2014). Given that TGFβ ligands are expressed in the zebrafish tailbud during somitogenesis and Nodal ligands are not (Monteiro et al., 2016), the observed phenotype is more likely the result of TGFβ signalling inhibition, rather than Nodal signalling inhibition.

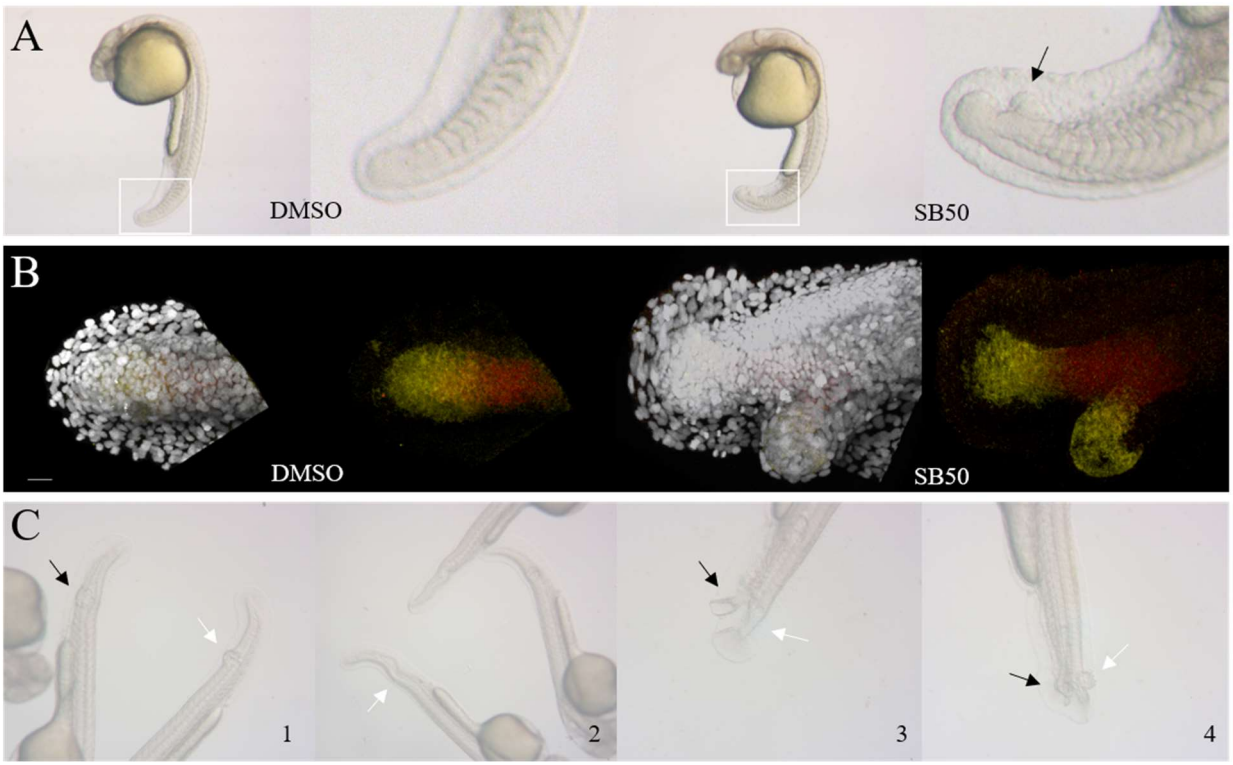


Figure 31: Tissue swellings form after inhibition of TGFβ signalling with SB50. Embryos were treated at early somitogenesis stages (~ 8–16 hpf) with either DMSO (control) or SB50 (an inhibitor of TGFβ signalling). (A) Control embryo (left images), and SB50 embryo (right images) at 24hpf (30 somite-stage). White boxes show the magnified region (shown to the right of each embryo). Black arrow indicates tissue swelling in the PSM. Scale bar: 20 µm. (B) HCR stains for *msgn1* (yellow) and *tbx6* (red) on embryos fixed at the 30 somite-stage. For both control (left images) and SB50 (right images), the image is shown with (left) and without (right) DAPI (grey). In this SB50-treated embryo, the PSM swelling shows the same AP patterning as the true PSM. (C) Four images showing different SB50 phenotypes. All embryos are imaged at 48 hpf. Swellings appeared in the neural tube (white arrows) as well as the paraxial mesoderm (black arrows). Images 3 and 4 show more extreme phenotypes. Several embryos also show a dorsal bend to the tail which is not seen in control embryos.

TGFβ signalling has been shown to regulate cell division, as inhibition leads to over-proliferation in some contexts (Casari et al., 2014). Therefore, to test whether these swellings were extreme cases of increased proliferation, I treated embryos with SB50 and fixed them at regular intervals. I then performed antibody stains for phosphorylated histone H3 (PH3), a marker of cells undergoing mitosis (Figure 32A,B). I used Imaris to automatically detect PH3 "spots", and then manually removed all spots anterior to the nascent somite – to only include PH3+ cells in the unsegmented region/tailbud. I also removed any spots past the midline, to ensure differences between samples in terms of imaging depth did not affect the results. The results show that after 1 2, 4, and 6 hours of treatment, there is no significant difference in the number of PH3+ tailbud cells between treated and control embryos (Figure 32C) (1 hour: $t = 0.31$, $p = 0.76$; 2 hours: $t = -1.45$, $p = 0.17$; 4 hours: $t = 1.50$, $p = 0.15$; 6 hours: $t = 0.97$, $p = 0.34$). This suggests that tailbud swellings are not the result of over-proliferation. At 1 and 2 hours after treatment, there does seem to be more variation in treated embryos than in control embryos, but this difference in variation was not statistically significant (1 hour: $F = 0.28$, $p = 0.13$; 2 hours: $F = 0.23$, $p = 0.10$).

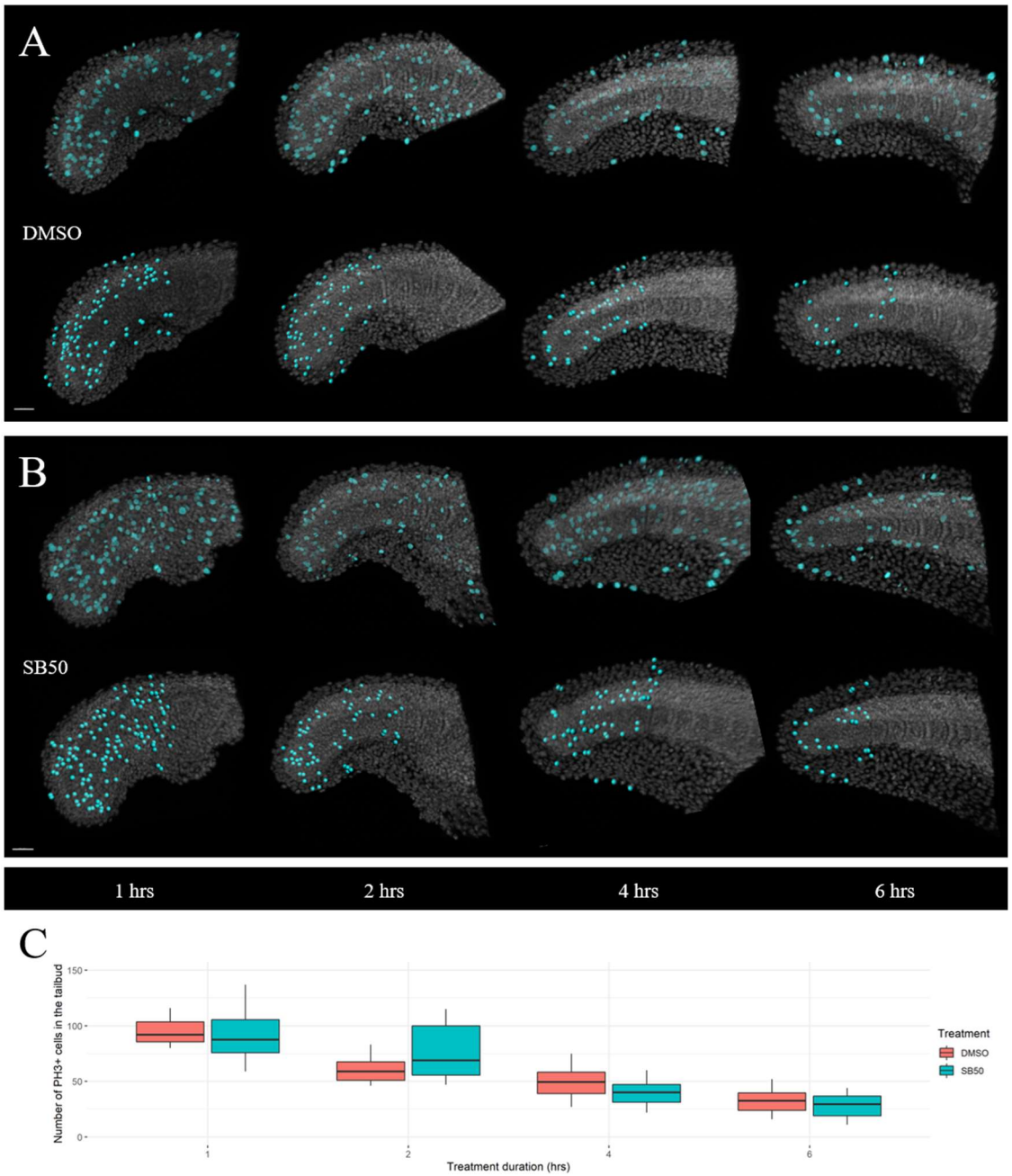


Figure 32: Tissue swellings are not the result of increased proliferation. Embryos were treated with DMSO (control) (A) or SB50 (B) at the 16 somite-stage and embryos were fixed after 1, 2, 4, and 6 hours of treatments (left to right). Top row of images shows DAPI (grey) and PH3 (cyan), and bottom row of images shows DAPI with tailbud "spots" (cyan) – representing all PH3+ cells posterior to the nascent somite, and on one lateral half of the embryo. Scale bars: 30 µm. (C) Boxplot showing the number of PH3+ tailbud cells at each timepoint for both treated (n = 10 embryos per timepoint) and control embryos (n = 10 embryos per timepoint). No significant differences were detected between treated and control embryos at each stage (see text for details).

In a separate experiment, I performed HCR stains for *msgn1* on SB50-treated embryos, to look more directly at tissue volume and cell number in the PSM progenitor region (Figure 33A). In keeping with previous results of proliferation, the number of cells in the PSM progenitor region was not significantly different between treated and control embryos ($t = -1.77$, $p = 0.08$). However, the volume was significantly higher in treated embryos ($t = -2.42$, $p = 0.02$). In other words, treated embryos had a slightly larger progenitor region, not because of increased cell number, but because of decreased tissue density ($t = 3.11$, $p = 0.004$) (Figure 33B). It is possible, then, that tissue swellings are an extreme phenotype of reduced density. The reason for the high variability of this phenotype is unknown, but the simplest explanation would be that perhaps SB50 does not diffuse well, and so some embryos are exposed to higher levels than others. The reason for swellings sometimes appearing in the paraxial mesoderm and other times appearing the neural tube is less clear - and remains unknown.

Interestingly, the effect of SB50 treatment on PSM density suggests that TGFβ signalling could play a role in increasing PSM density over time. However, the density difference observed between treated and control embryos is only very slight, and so this may indicate that TGFβ signalling is not the only factor responsible for density increase. It is also not clear *how* TGFβ signalling increases PSM density – whether this is through intrinsic or extrinsic mechanisms.

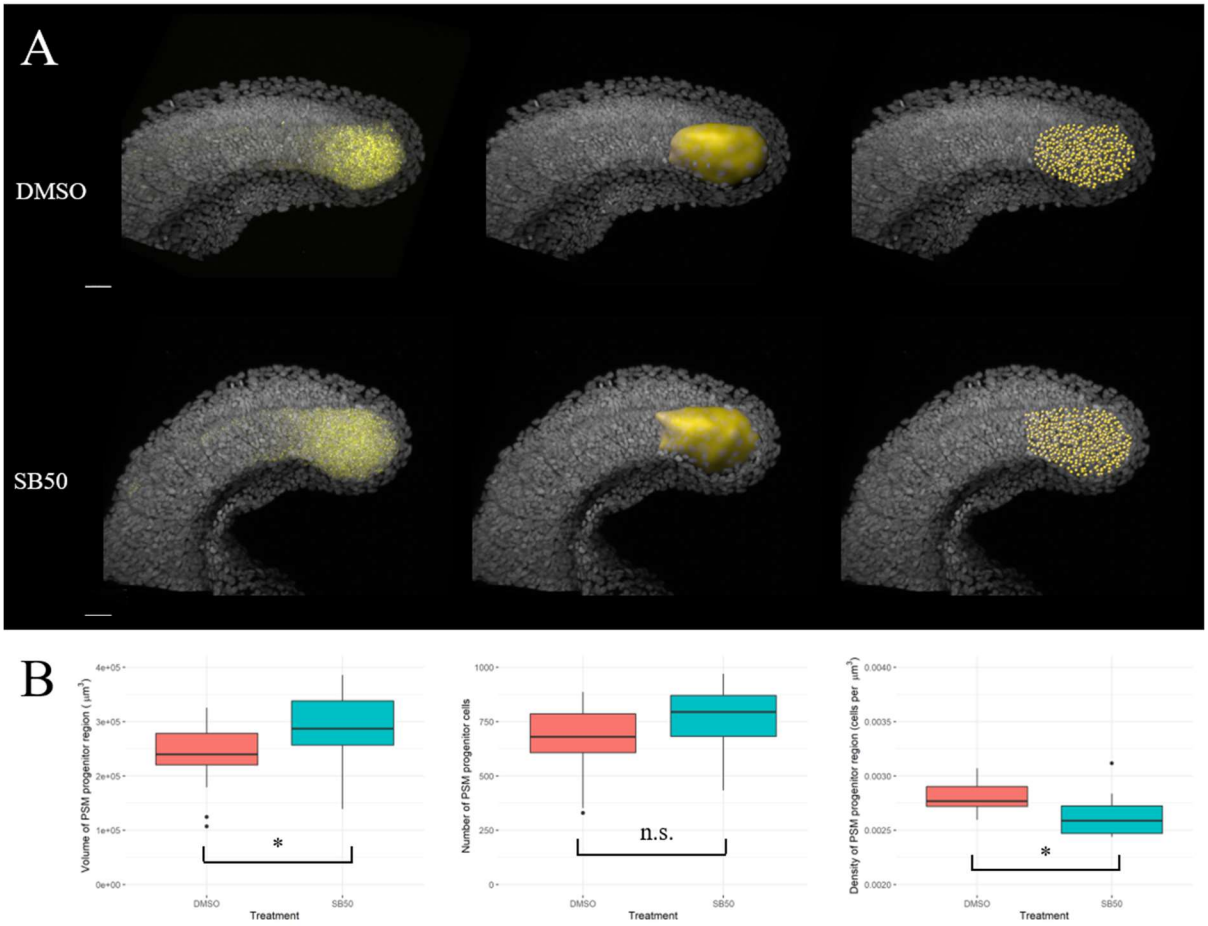


Figure 33: SB50 treatment causes a minor but significant reduction in PSM progenitor density. Embryos were treated with DMSO (control) or SB50 at the onset of somitogenesis and fixed at the 28 somite-stage. HCR stains were then performed for *msgn1* (yellow). DAPI (grey) is shown in all images with *msgn1* HCR signal (left), PSM progenitor automatic surface (middle), and PSM progenitor automatic spots (right). Scale bars: 30 µm. (B) Boxplots showing the volume, number of cells, and density of the PSM progenitor region (left to right). Significant differences were detected for volume and density, but not for cell number, between control embryos (n = 17) and treated embryos (n = 17).

3.4 Conclusions

The mechanisms controlling paraxial mesoderm compression-extension remain unclear. However, the results of PCP pathway inhibition show that this pathway is not involved. This confirms that tissue convergence in the paraxial mesoderm during somitogenesis is not simply a continuation of convergent extension movements during gastrulation. Neural tube ablations appeared to have no effect on paraxial mesoderm compression or convergence. However, neural tube volume was also unaffected in these ablations, making it difficult to draw conclusions from this regarding a possible role of neural tube-driven compression of the PSM. *Cdx4* mutants also showed inconclusive results, but did show reduced paraxial mesoderm length, and no difference between stages in PSM density. This supports the idea that the neural tube is contributing to PSM compression-extension - but is not conclusive proof. TGFβ signalling appears to play a role in regulating PSM density, however the mechanism by which it regulates this is unknown – this could be related to neural tube elongation, or it could be unrelated. These findings, together with those of previous chapters, confirm that the zebrafish paraxial mesoderm elongates not through classical convergent extension, but through compression-extension. The source of this compression remains unclear, but may involve extrinsic forces from the neural tube and TGFβ signalling.

RESULTS

4 Evolutionary comparisons of paraxial mesoderm elongation

4.1 Introduction

In the previous three chapters, I have focussed on zebrafish as a model system for studying paraxial mesoderm elongation in the absence of posterior growth. However, the use of any species as a "model system" always requires justification: how well does this model species represent others? And how much can findings from this species be generalised? In this chapter, I extend my analyses to other vertebrates: firstly, to two other teleost species (from the cichlid family), then to another bony vertebrate (chick), and then to a chondrichthyan (catshark). This comparative approach will allow me to test how general the findings of the previous chapters are, and form hypotheses about the evolution of paraxial mesoderm elongation mechanisms.

4.2 Somitogenesis and paraxial mesoderm elongation in two cichlid species

The cichlid family is one of the most species-rich vertebrate families, comprising over 3,000 species spread across Central and South America, Africa, and India (Kocher, 2004). Most (~ 2,000) of these species arose in the last 10 million years in three African Great Lakes: Lake Victoria, Lake Tanganyika, and Lake Malawi (Seehausen, 2006), making this group also one of the most rapidly-evolving vertebrate families. These huge adaptive radiations involved diversification of habitats, feeding niches, and pigmentation patterns. Because of this, cichlids have long been used in the fields of behavioural ecology and evolutionary biology. However, it is only recently that their embryonic development has been studied (Streelman and Albertson, 2006; Fraser et al., 2009). Because of this, very little is known about cichlid embryonic development, and almost nothing is known about somitogenesis and axis elongation - and the degree to which these processes differ among species. It has been reported, however, that vertebral number differs between species within a genus of cichlids (Mattsson, 2018) – suggesting that somitogenesis and paraxial mesoderm elongation have diversified among cichlids.

Cichlids, therefore, represent an exciting group to study these processes in. As teleosts, they can be compared easily with zebrafish, to determine how general zebrafish processes and mechanisms are to the wider teleost group. They also present a unique possibility of comparing development between very closely related vertebrate species – to look at the evolution of development over microevolutionary, rather than macroevolutionary, timescales. Additionally, as cichlid embryos show variation in size (particularly in yolk size), they also present the possibility to test the hypothesis of Steventon et al. (2016); that posterior growth is associated with large energy supply during embryonic development (e.g. through a large yolk or via a placenta).

I focussed on two cichlid species: *Astatotilapia calliptera* and *Rhamphochromis chilingali*. Both species are endemic to Lake Malawi and both are mouthbrooders, meaning the females store their embryos in their mouths until hatching. *R. chilingali* adults are much larger than *A. calliptera* adults, and their eggs are also substantially larger (Figure 34). *R. chilingali* is a midwater piscivore, whereas *A. calliptera* is a shallow-water generalist feeder (Malinsky et al., 2018), and these ecological differences are likely related to the differences in body and egg size. *R. chilingali* have ~ 35 vertebrae whereas *A. chilingali* have ~ 30 vertebrae, and these numbers appear to roughly match final somite number of embryos (Berta Verd, unpublished data). This makes for an interesting comparison of somitogenesis.

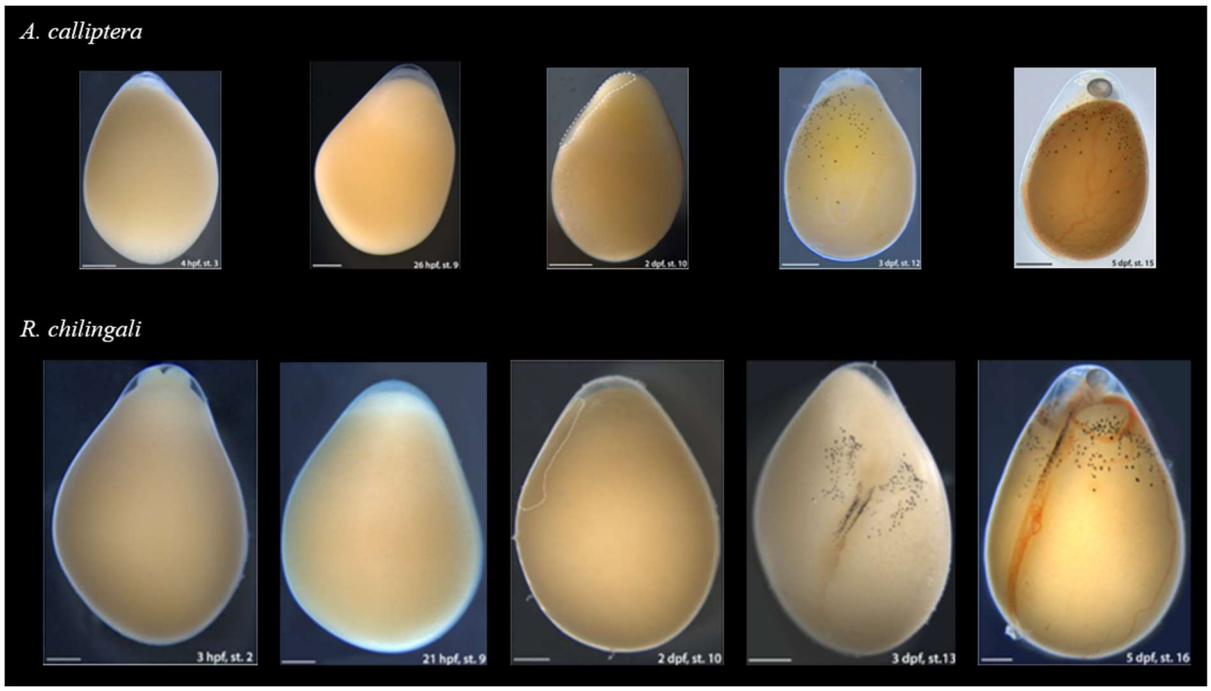


Figure 34: Two cichlid species with different sizes of eggs/embryos. *Astatotilapia calliptera* and *Rhamphochromis chilingali* embryos are shown at different developmental stages (left to right): cleavage, gastrula, early somitogenesis, late somitogenesis, and pharyngula. Images have been adjusted to keep the scale bar (1 mm) constant, to show the relative sizes of the two species. Embryos were staged and imaged by Aleksandra Marconi, who provided these images. The white dotted lines in the early somitogenesis images highlight embryo location.

Somitogenesis rate and axis elongation

The difference in somite/vertebral number between these two species suggests that somitogenesis has evolved differently in the two lineages. As discussed previously, differences in somite number may arise due to changes to the clock rate (i.e. increasing or decreasing somitogenesis rate), or due to changes in axis elongation – as somite number is determined by the relative rates of these two processes. I fixed embryos from one clutch of each species at different timepoints (Figure 35A) and counted the number of somites in every embryo. Because both species are mouthbrooders, the exact time of fertilisation is usually not known. Therefore, time was measured relative to the time I fixed the first batch, for each species. This allowed me to calculate somitogenesis rate for each species (Figure 35B), but not the time (relative to fertilisation) that each species begins and ends somitogenesis. I also measured total embryo length and total paraxial mesoderm length (from the anterior boundary of the 1st somite to the posterior tip of the PSM) for each embryo, to compare axis elongation rates between species (Figure 35C).

The results (Figure 35) show that somitogenesis rate is the same between the two species: roughly forming one somite pair every 2.5 hours. However, axis elongation rate was found to be much greater in *R. chilingali*. Although at earlier stages of somitogenesis, this species was smaller than *A. calliptera,* embryo length (and paraxial mesoderm length) increased much quicker - meaning that, by the end of somitogenesis, *R. chilingali* embryos were much longer. This suggests that differences in somite number are due to changes in axis elongation rate or wavefront movement rate rather than changes in clock cycle/somitogenesis rate.

However, from these simple measurements I cannot rule out the possibility that this elongation is only occurring in the anterior, segmented part of the embryo. As discussed previously, only elongation of the unsegmented region/PSM is relevant to axis elongation in terms of providing sufficient length for more segments to form. Therefore,

in the following sections I apply the same morphometric analysis to both cichlid species as I previously applied to zebrafish, to calculate indirect paraxial mesoderm measurements.

PSM marker gene expression over time

To enable 3D measurements of the PSM, I used HCR to stain for *tbx16* and *tbx6* (and DAPI to stain nuclei). *Tbx16*, like *msgn1*, marks the posterior PSM, while *tbx6* marks the anterior PSM. Both species showed similar expression patterns to each other (and to zebrafish) for the stages observed: the expression domains of both genes shrink over time as the PSM is depleted (Figure 36). Earlier stage embryos were imaged from the dorsal/ventral side, rather than the lateral side.

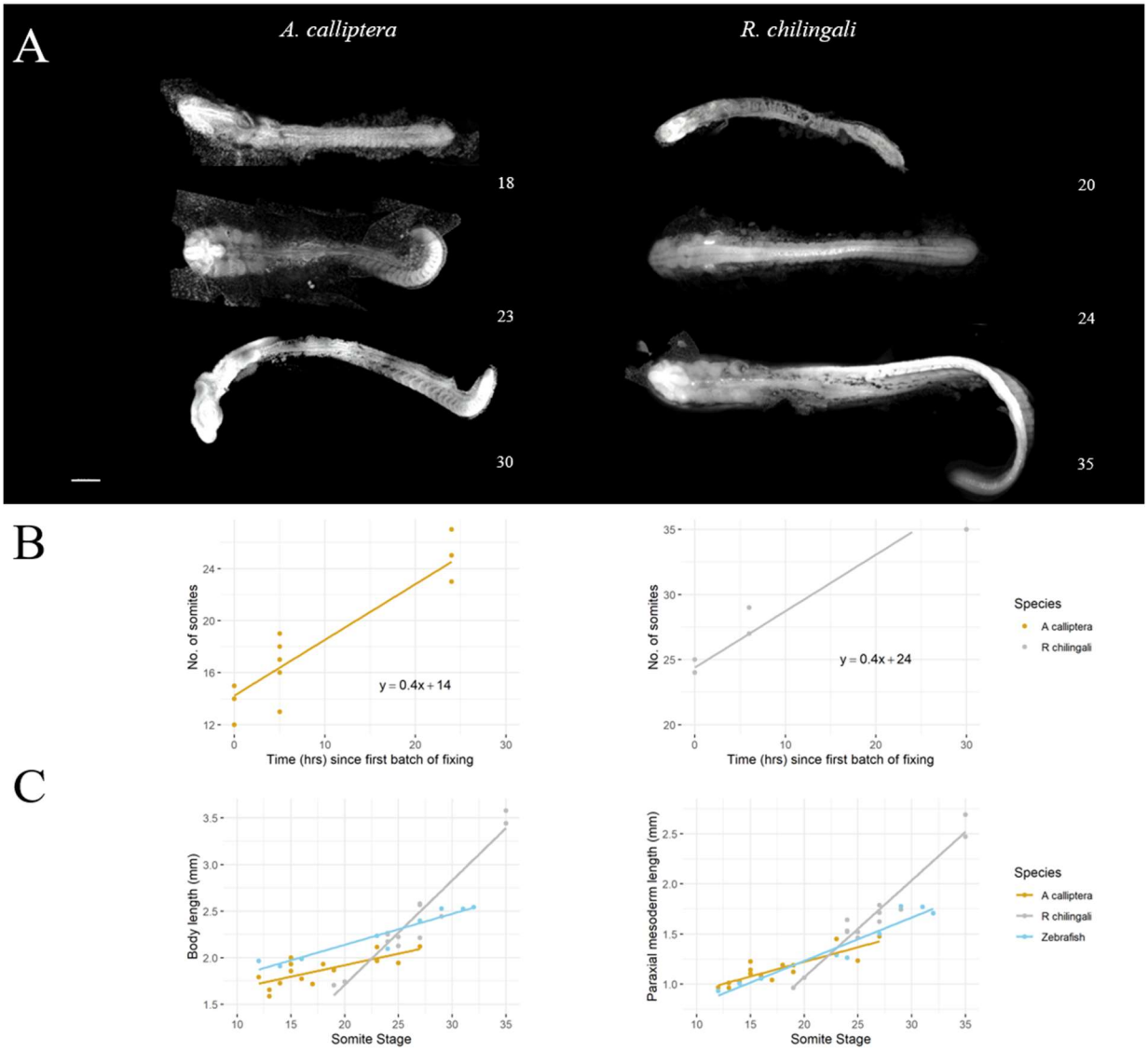


Figure 35: Cichlid somitogenesis and axis elongation rates. Embryos of *A. calliptera* (n = 16) and *R. chilingali* (n = 11) were fixed at different timepoints, with time measured relative to the first timepoint. (A) Embryos at different somite stages (bottom right number for each image), stained with DAPI. Scale bar: 200 µm. (B) Somitogenesis rates for both species. Number of somites are plotted over time (hours since first fixing timepoint), with a linear trendline shown[1]. The rate for both species is y = 0.4x, i.e. 0.4 somites per hour (1 somite every 2.5 hours). (C) Total embryo body length and total paraxial mesoderm length over time for both species (and zebrafish, for comparison). For body length and paraxial mesoderm length, *R.* chilingali shows a much faster rate of elongation (relative to somite stage) than *A. calliptera* and zebrafish.

[1] When calculating the trendlines *R. chilingali* embryos with 35 somites (n = 2) were not included, as these appeared to have finished somitogenesis, and, given the time gap since the previous timepoint, they may have reached this somite stage several hours earlier – meaning the trendline would give a rate that was slower than the true rate.

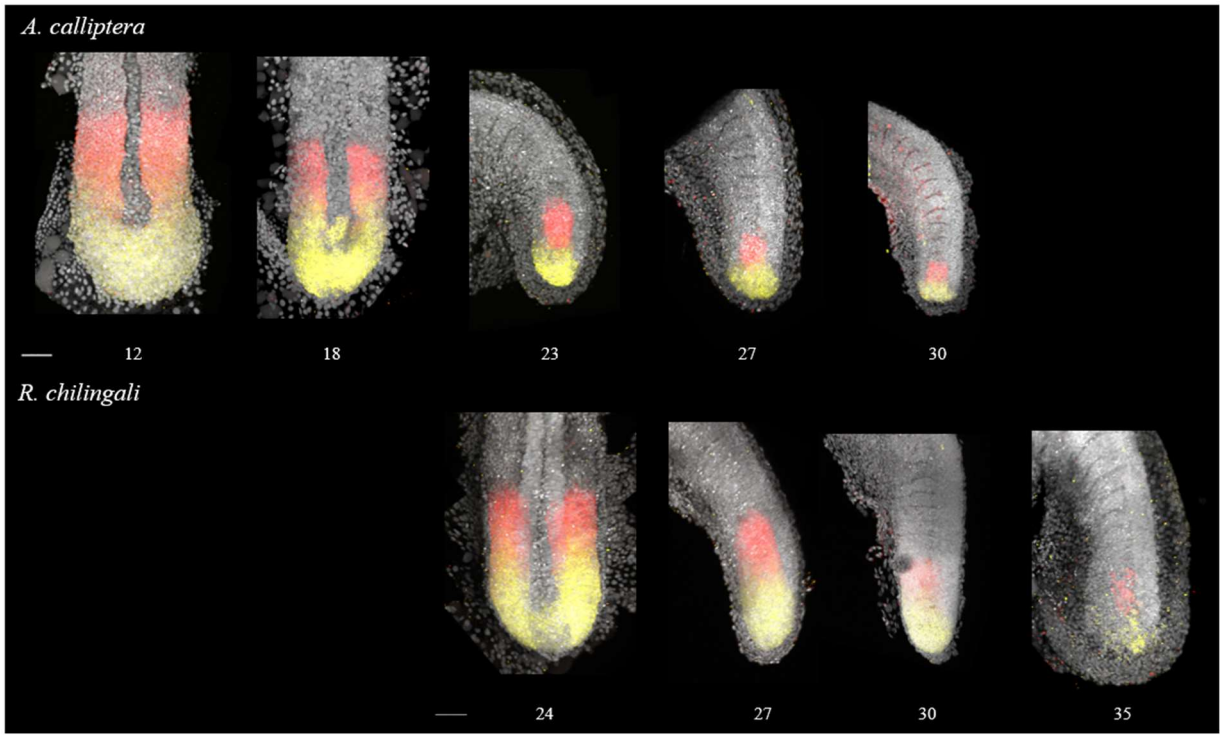


Figure 36: Cichlid PSM marker gene expression patterns. *in situ* hybridization chain reaction (HCR) was used to stain for *tbx16* (yellow) and *tbx6* (red), with nuclei stained with DAPI (grey). Each image is a maximum projection of a confocal stack. *A calliptera* embryos ranged from 12 to 30 somite-stage, and *R. chilingali* ranged from 24 to 35 somite-stage. The first two images for *A. calliptera* and the first image for *R. chilingali* are dorsal/ventral view. The final image for *R.* chilingali shows fading gene expression, as well as significant fin growth, suggesting that somitogenesis has finished by this stage. All other images are lateral view. Anterior is top in all images. Scale bars: 50 µm.

Morphometric methods

The image analysis methods used to generate morphometric measurements were largely the same as for zebrafish. 3D tissue reconstructions of the PSM and nascent somites for each stage were generated by creating manual "surfaces" in Imaris. The main difference was that, for earlier stage embryos which were imaged from a dorsal/ventral view, I created a surface of the whole PSM (rather than a lateral half) and both somites from the nascent somite pair (rather than one) (Figure 37A). Measurements from these images were halved (PSM) or averaged (nascent somites) to then be comparable with later-stage embryos which were imaged from a lateral view (Figure 37B). Cell numbers were obtained using automatic spots, as before, which were validated with the same "masking" method as before. The same parameters as used for zebrafish (cell diameter = 4 µm, no minimum fluorescence intensity threshold) were used for both cichlid species, as these provided accurate spot detection. Dimension measurements were also taken as before, with separate height (DV axis) and width (ML axis) measurements taken for the anterior and posterior PSM at each stage, and a PSM length measurement that followed the 3D curvature of the tissue.

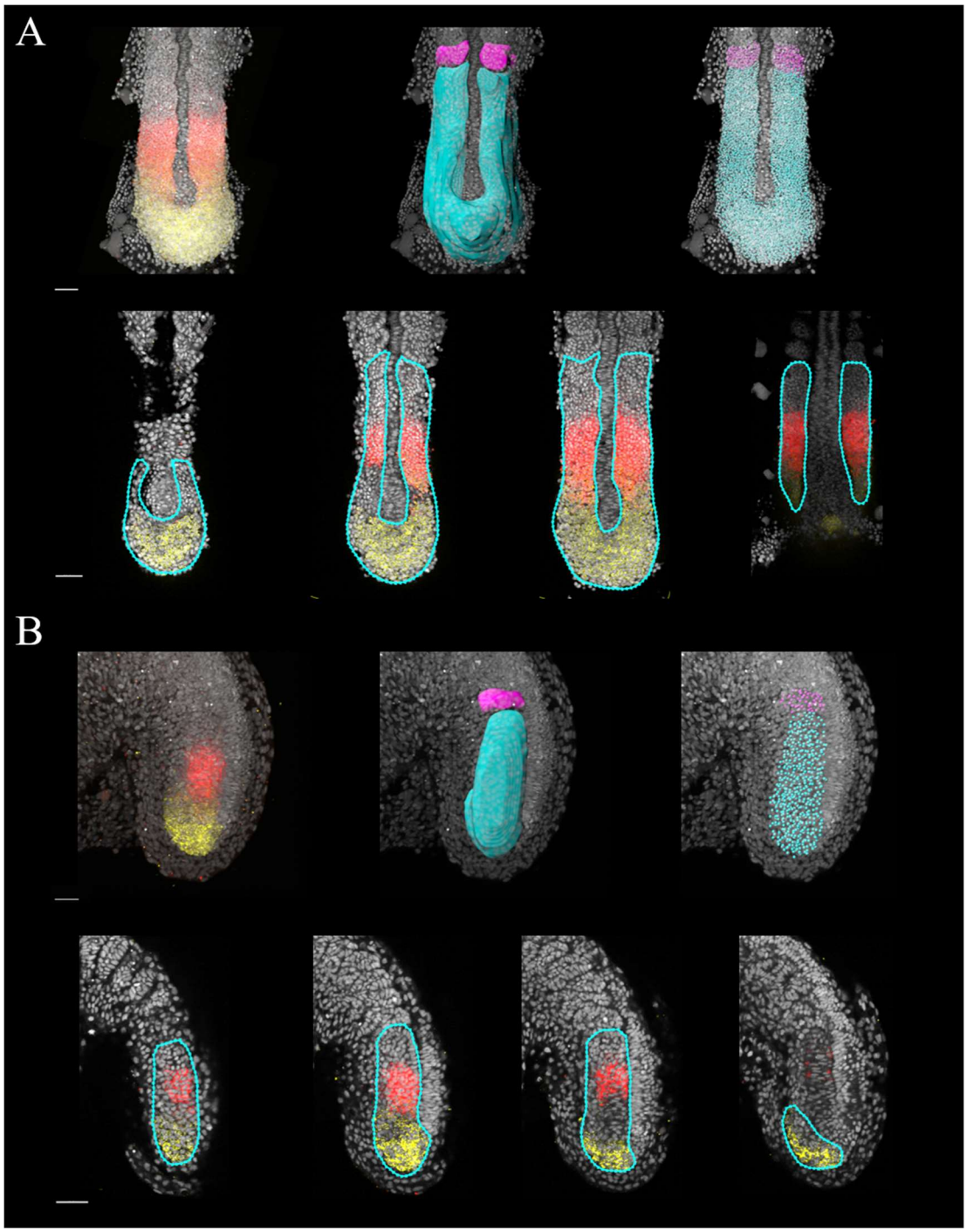


Figure 37: Cichlid morphometric methods. Manual surfaces and automatic spots were made in Imaris as before. (A) Process for dorsal/ventral-view images. Top row: HCR image, surfaces, and spots are shown (left to right) Scale bar: 50 µm. Bottom row: Individual z-slices are shown with the PSM surface contours for the whole PSM. Scale bar: 50 µm. (B) HCR image, surfaces, and spots are shown (left to right). Scale bar: 30 µm. Bottom row: Individual z-slices are shown with the PSM surface contours for one lateral half of the PSM (up to the midline). Scale bar: 40 µm.

A heterochronic shift in tailbud formation

Although imaging orientation may seem irrelevant to understanding development, the differences between stages did raise an interesting point. The reason that earlier stage embryos were imaged from a dorsal/ventral view rather than a lateral view was simply due to the way that samples naturally lay during mounting/imaging. I presumed that this was due to whether the tailbud had properly formed – as, in zebrafish at least, the unsegmented region during early stages is flat against the yolk (which means cells can converge across the yolk to join the PSM), and at later stages (i.e. tailbud stages) is raised up off the yolk. The observation that, between species, embryos of the same somite stage (23-24) naturally lay in different orientations therefore suggested that the stage of tailbud formation is different between species. To test this, I examined images from around these stages more carefully, to look at the connection between the embryo and the extra-embryonic tissue. The images (Figure 38A) show that in *A. calliptera*, tailbud formation occurs around the 21-23 somite-stages, whereas in *R. chilingali*, it occurs later - around the 25-27 somite-stages.

This difference is roughly the same as the number of "extra" somites that *R. chilingali* has compared to *A. calliptera*. In other words, both species form roughly the same number of somites after tailbud formation, but tailbud formation is "delayed" in *R. chilingali* such that it has formed more somites by this point. This suggests that as well as differences in axis elongation, a heterochronic shift in tailbud formation may have contributed to differences in somite/vertebral number. It is important to note that, as I am comparing two extant species to each other (rather than comparing an extant with an extinct, ancestral species), I cannot conclude in which direction this shift occurred, as we do not know the ancestral state. Nevertheless, a heterochronic shift appears to have occurred in one of (or both of) the two lineages.

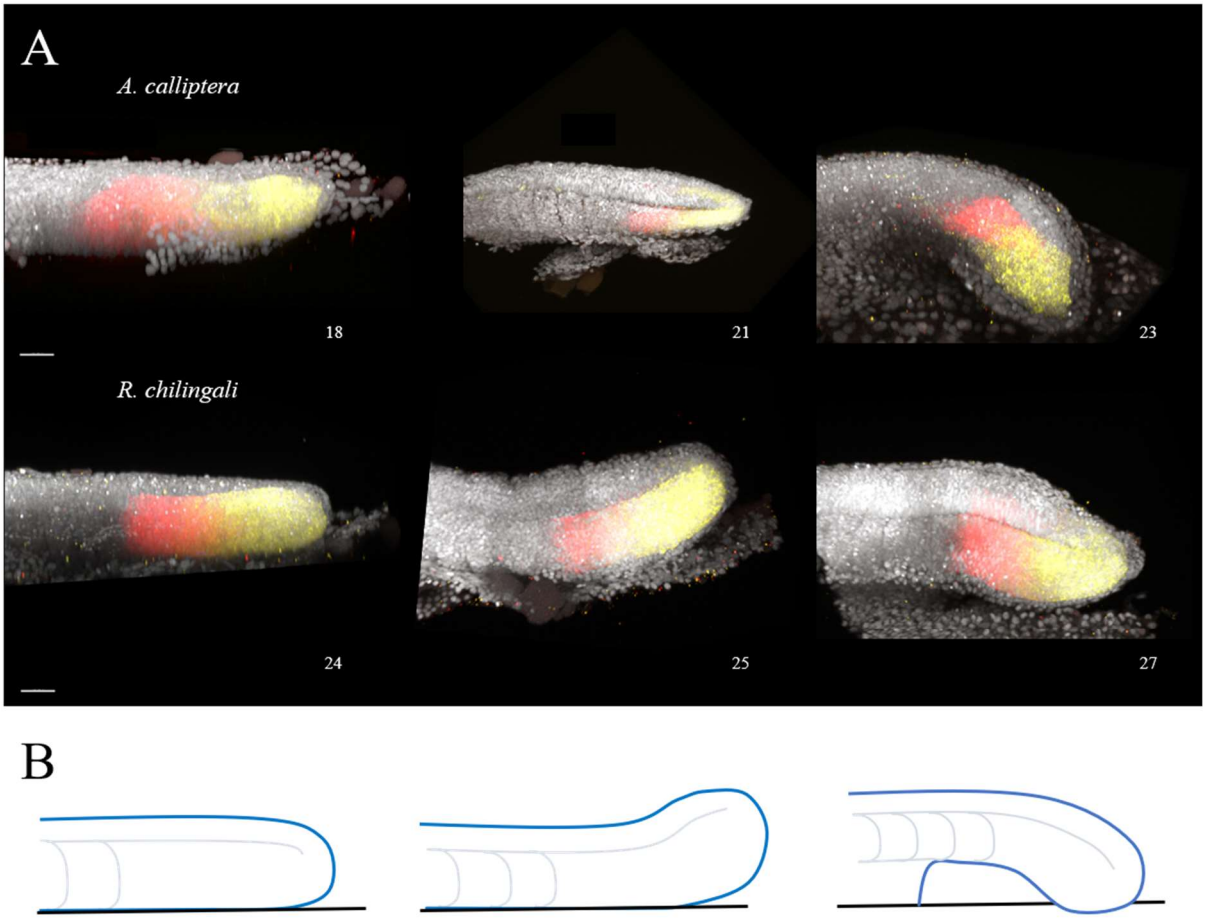


Figure 38: Tailbud formation occurs later in *R. chilingali* than in *A. calliptera*. (A) Images before, during, and after tailbud formation (left to right) are shown for both species, along with the somite stage of each image (bottom right numbers). All images are lateral view (those imaged from a dorsal/ventral view are rotated to lateral view). Tailbud formation occurs around the 21-23 somite-stages in *A. calliptera* and around the 25-27 somite-stages in *R. chilingali*. (B) Schematic of tailbud formation (left to right). Initially, the posterior of the embryo (blue) lies flat against the yolk (dark grey). Then the posterior tip begins to lift off from the yolk. Finally, the whole unsegmented region is fully raised up off the yolk, forming a distinct tailbud.

Different forms of convergent extension appear to drive paraxial mesoderm elongation between cichlids

Morphometric measurements of both cichlid species (Figure 39A,B) show similar trends to each other, and to zebrafish, in terms of length: the PSM decreases in length over the course of somitogenesis in a highly linear manner, as it is depleted to form somites. However, like in zebrafish, the PSM is elongating (shown by the indirect paraxial mesoderm measurements) – but the rate of depletion is faster, resulting in a net length decrease. Like zebrafish, both species also do not show any volumetric growth in the paraxial mesoderm (Figure 39A,B). Instead, the results indicate that the paraxial mesoderm *decreases* in *both* cell number and volume over time (whereas zebrafish cell number stayed constant before a slight increase late in somitogenesis, and volume decreased over time). The decrease in cell number is surprising, and could suggest that PSM cells are contributing to tissues other than the somites (as the number of cells lost from the PSM at each stage is greater than the number of cells in the nascent somites), or that they are undergoing cell death. However, the most likely explanation is that the apparent cell number decrease is due to slight measurement error (e.g. if somite cell numbers were underestimated) combined with the small sample size. Nevertheless, even allowing for substantial error, it is clear that paraxial mesoderm elongation is not driven by growth in either species.

A. calliptera showed similar results to zebrafish in terms of dimension measurements: both height (DV) and width (ML), of both the posterior and anterior PSM, decreased over time (Figure 39C). Although in *A. calliptera* it is width that decreases the most, whereas in zebrafish it is height that decreases the most. The PSM also shows a similar increase in density in *A. calliptera* as in zebrafish (Figure 39C). Importantly, this suggests that compression-extension is not unique to zebrafish; and is also occurring in *A. calliptera*. The results for *R. chilingali*, on the other hand, are quite different (Figure 39D). While the width of both the posterior and anterior PSM decreases

over time, the height does not (and instead shows a very slight increase over time). Additionally, the PSM does not increase in density, and instead shows a very slight decrease. Again, it is not clear if these differences are true developmental differences, or simply an artefact of cell number measurement error and small sample size. If these results are accurate, this suggests that, while convergent extension appears to be happening in both species, it is happening in different ways/through different forms. In *A. calliptera*, it involves width and height decrease, and density increase (like in zebrafish). In *R. chilingali*, it involves only width decrease, and no density increase.

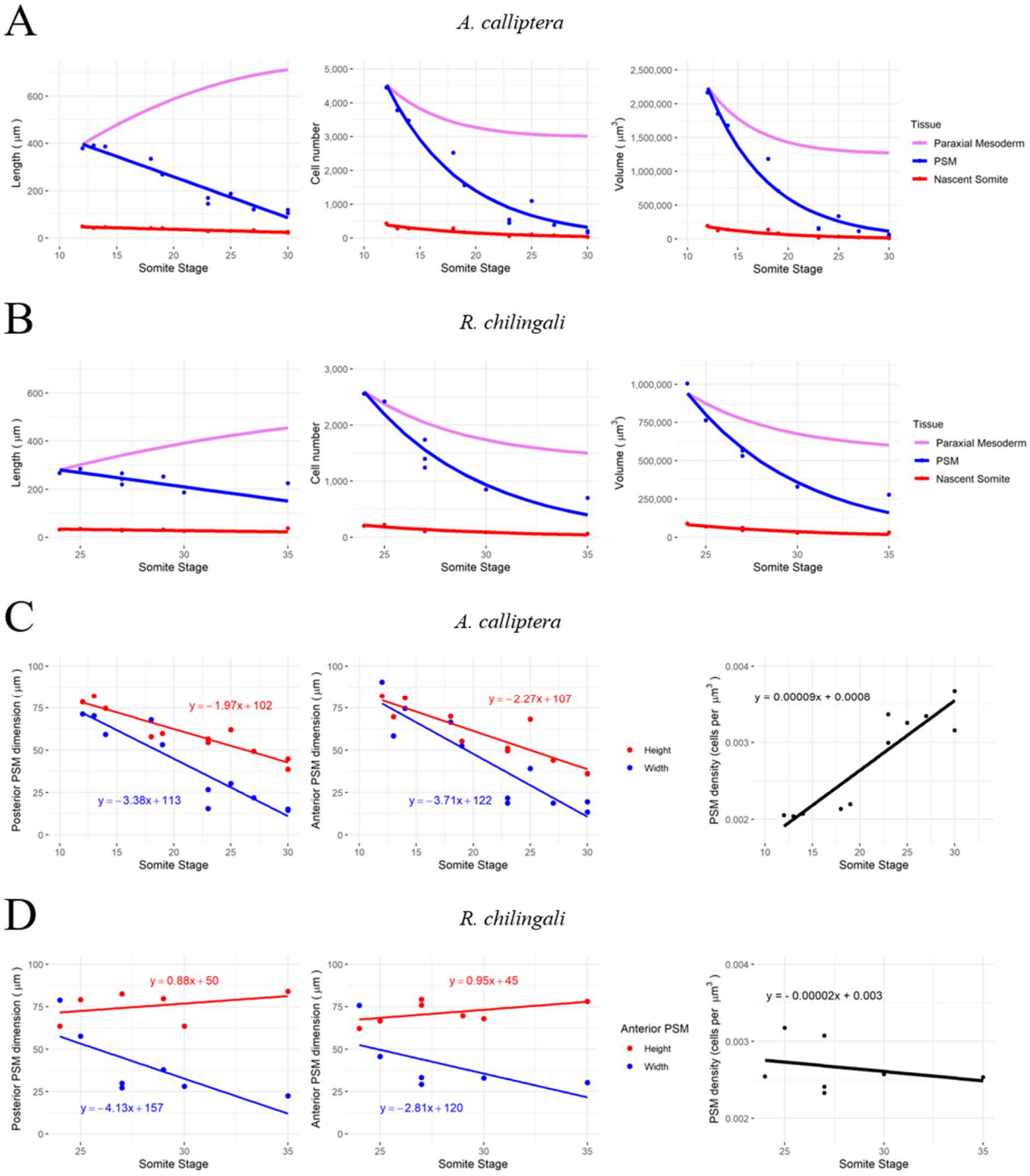


Figure 39: Morphometric results for *A. calliptera* (n = 11 embryos) and *R. chilingali* (n = 8 embryos). (A, B) Length, cell number, and tissue volume measurements (left to right) for the PSM (blue lines) and nascent somite (red lines) across somite stages. The trendline equations for these were then used to calculate the indirect paraxial mesoderm values (violet lines, see Section 1.2 for explanation). Results are shown for both *A. calliptera* (A) and *R. chilingali* (B), with highly similar results for both species. (C, D) Height (DV axis, red lines) and width (ML axis, blue lines) measurements of the posterior PSM (left graphs) and anterior PSM (right graphs) across somite stages, as well as PSM density measurements (right graphs). Results are shown for both *A. calliptera* (C) and *R. chilingali* (D), with differences in height and density trends. Note the different x-axis scales between species (A-D), due to available stages for measurement.

From this data, comparing the two species quantitatively to each other is difficult, as the range of stages sampled for the two species are different (*A. calliptera:* 12 to 30 somite-stage, *R. chilingali*: 24 to 35 somite-stage). However, I selected the region of overlap (24 to 30 somite-stage) to compare measurements between species more quantitatively – taking the trendline equations from the full datasets (shown in Figure 39) to give values for this specific range of stages. The results of this are shown in Figure 40, along with zebrafish values for comparison.

The indirect paraxial mesoderm length from the 24th somite onwards (i.e. the length PSM from 24 somite-stage onwards, including the somites it gives rise to) was calculated based on the full dataset trendlines. The results show that, for these stages, PSM elongation rate is greater in *R. chilingali* than in *A. calliptera*. This is surprising, given that convergence happens in both width and height in *A. calliptera* but only in width in *R. chilingali* – especially as the width decrease rate is similar for both species. However, it does fit with the measurements of total embryo body length and total paraxial mesoderm length, which showed faster elongation rates in *R. chilingali* (Figure 35). This supports the idea that increased somite number could be due to changes in axis elongation rate.

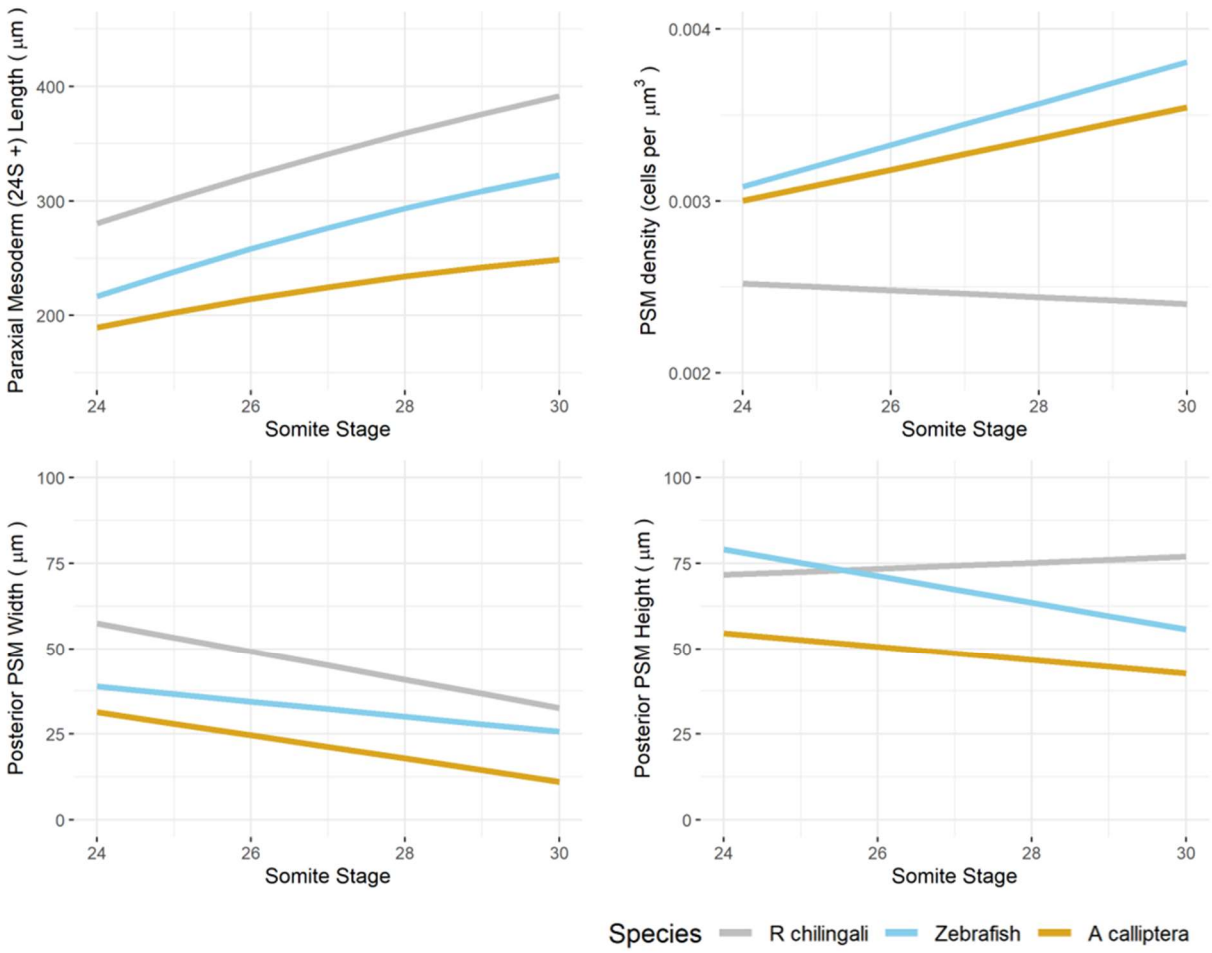


Figure 40: Direct species comparison of morphometric trends. Trendlines from Figure 39 were used to plot values over the same range of stages (24 to 30 somite-stage) for each species, to allow direct comparisons. Graphs are shown for indirect paraxial mesoderm length from the 24th somite onwards (top left), PSM density (top right), posterior PSM width (bottom left) and posterior PSM height (bottom right).

4.3 Chick paraxial mesoderm elongation

Morphometric methods

The previous results, showing that tissue convergence appears to be a conserved feature of paraxial mesoderm elongation in distantly related teleosts, raises the question: is this process also conserved across other vertebrate groups? Particularly, is it conserved even in species displaying volumetric growth of the PSM? In this section I test this in chicks – another bony vertebrate species, but one in which embryos undergo volumetric growth of the PSM (Bénazéraf et al., 2017).

Chick embryos were fixed across a range of HH stages (10-18) and stained using HCR for *tbx6* expression. Whereas in teleosts *tbx6* is a marker of the anterior PSM (with *tbx16* marking the posterior PSM), in other bony vertebrates (including chick) there is only *tbx6*, which marks the entire PSM. This difference is likely due to the extra whole genome duplication event in teleosts (Glasauer and Neuhauss, 2014), generating two copies of an ancestral gene - with each copy taking on a more specialised role/expression pattern. Chick *tbx6* HCR stains thus allowed me to create manual surfaces of the whole PSM, as before. Due to their large size/depth, embryos could not be imaged using a 40x objective, and instead were imaged using a 20x objective. Because of the reduced resolution that this meant (particularly in the z-axis), it was not possible to generate accurate cell "spots" in Imaris. Therefore, the morphometric results presented here only include tissue volumes and dimensions – cell number and tissue density could not be measured.

Figure 41 shows *tbx6* expression for a range of stages (12 to 28 somite-stage), alongside my PSM and nascent somite surfaces. As chick embryos form ~ 55 somites, this analysis does not represent the entirety of somitogenesis, but nevertheless covers a much longer period than is usually observed in chick axis elongation studies. Although all embryos were imaged dorsally (allowing surface reconstruction of the whole PSM and

both nascent somites), volumes of the PSM and nascent somites were halved and averaged, respectively, to maintain consistent methods with previous sections.

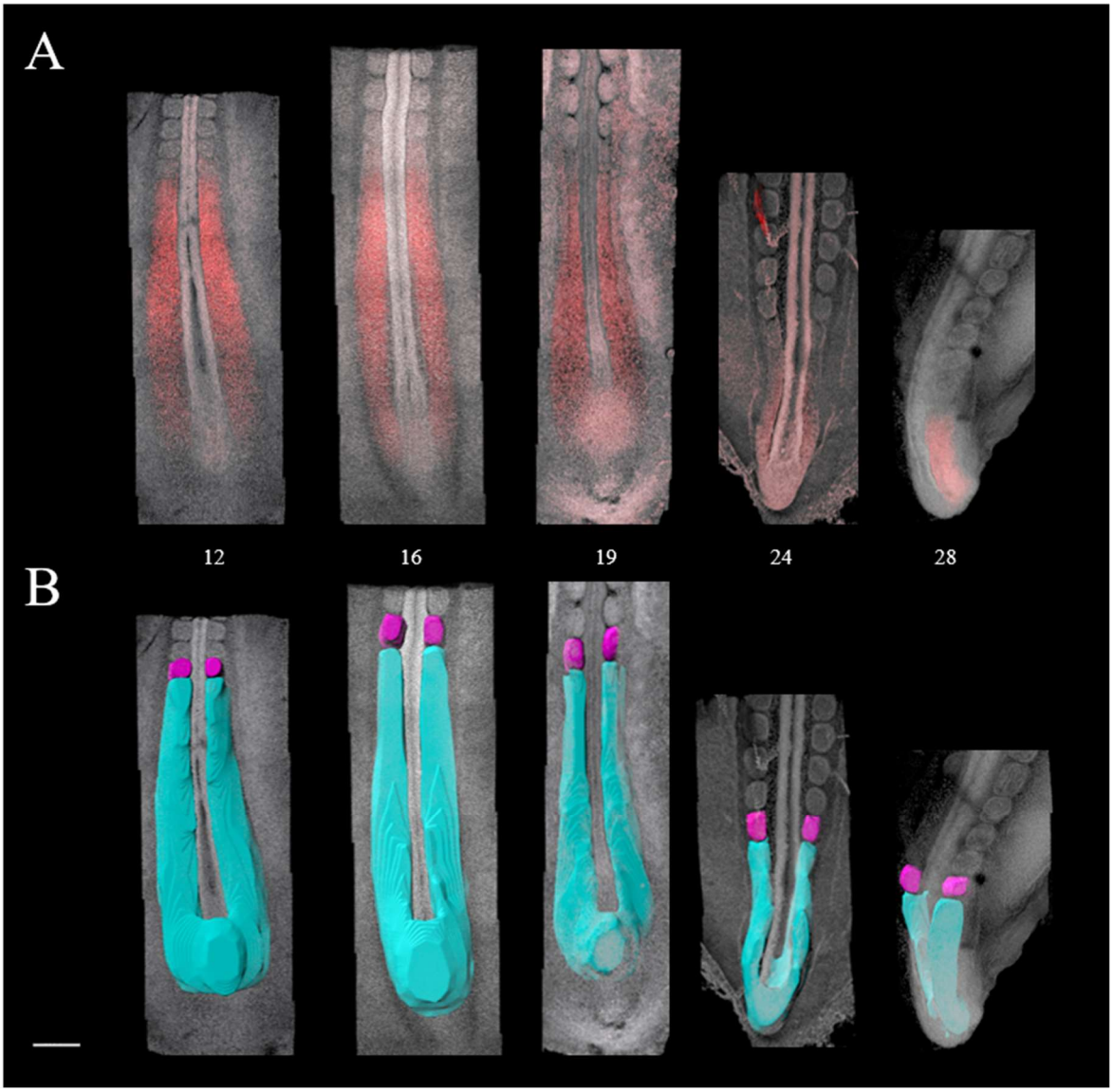


Figure 41: Chick PSM morphometric methods. (A) Embryos were fixed at different somite stages (bottom number for each image) and stained using HCR for *tbx6* (red), with nuclei stained by SYTOX-DR (grey). (B) Same images as above, showing manual surfaces of the PSM (cyan) and nascent somites at each stage (magenta). For the subsequent analysis, volume measurements for each stage were halved (PSM) or averaged (nascent somite pairs). All images are dorsal view, although the tailbud forms twists as it forms and so the 28 somite-stage image is between dorsal and lateral view. Measurement points were also taken of the width (ML) and height (DV) of the posterior and anterior PSM, as before (not shown here). Scale bar: 200 µm.

Volumetric growth combined with tissue convergence drives elongation

These results of this morphometric analysis are shown in Figure 42. As expected from the literature, both PSM length and volume initially increase, before decreasing at later stages (~ 20 somite-stage). The fact that this decrease begins so early is interesting, in that it suggests that for the majority of somitogenesis the PSM is in a net depletion phase – much like zebrafish. Nascent somite lengths and volumes, on the other hand, stay relatively constant (with an initial slight increase, and then a slight decrease) – contrary to what is observed in zebrafish (continuous decrease). As before, I used the trendline equations to calculate indirect paraxial mesoderm values for length and volume. Paraxial mesoderm length increases over these stages, although the elongation slows down and appears to plateau by the 28 somite-stage. Paraxial mesoderm volume also increases initially, but then begins to slowly decrease around the 20 somite-stage – roughly when the tailbud is starting to form. Whether this decrease is genuine is not clear – given the large deviation from the nascent somite volume trendline there is likely to be some error. However, even taking into account a large amount of error, the data suggests that, from this stage onwards, there is little to no volumetric growth (and there may even be volumetric shrinkage). At these stages, although paraxial mesoderm elongation is slowing, there is still some elongation. This suggests that volumetric growth is not the only process contributing to paraxial mesoderm elongation in chick embryos.

Measurements of the height and width of the posterior and anterior PSM showed that the width of the posterior PSM substantially decreases (~ 33%) over these stages (posterior height and anterior width/height all remain constant). This could suggest that there is some degree of convergence, at least in the ML axis, that contributes to elongation. Importantly, this decrease is linear throughout these stages. In other words, tissue convergence appears to be occurring even during stages in which volumetric growth is occurring (and then continues after volumetric growth ceases). This argues against any idea that volumetric growth and convergent extension represent different

"modes" of elongation, and instead suggests that both processes can co-occur and contribute to elongation together.

It is important to note that this analysis was performed on a small sample size, and so these conclusions are only preliminary. Further work would be required to conclusively show that tissue convergence is contributing to elongation, but these results do raise interesting questions, and at least suggest that some form of convergent extension could be a conserved feature of paraxial mesoderm elongation.

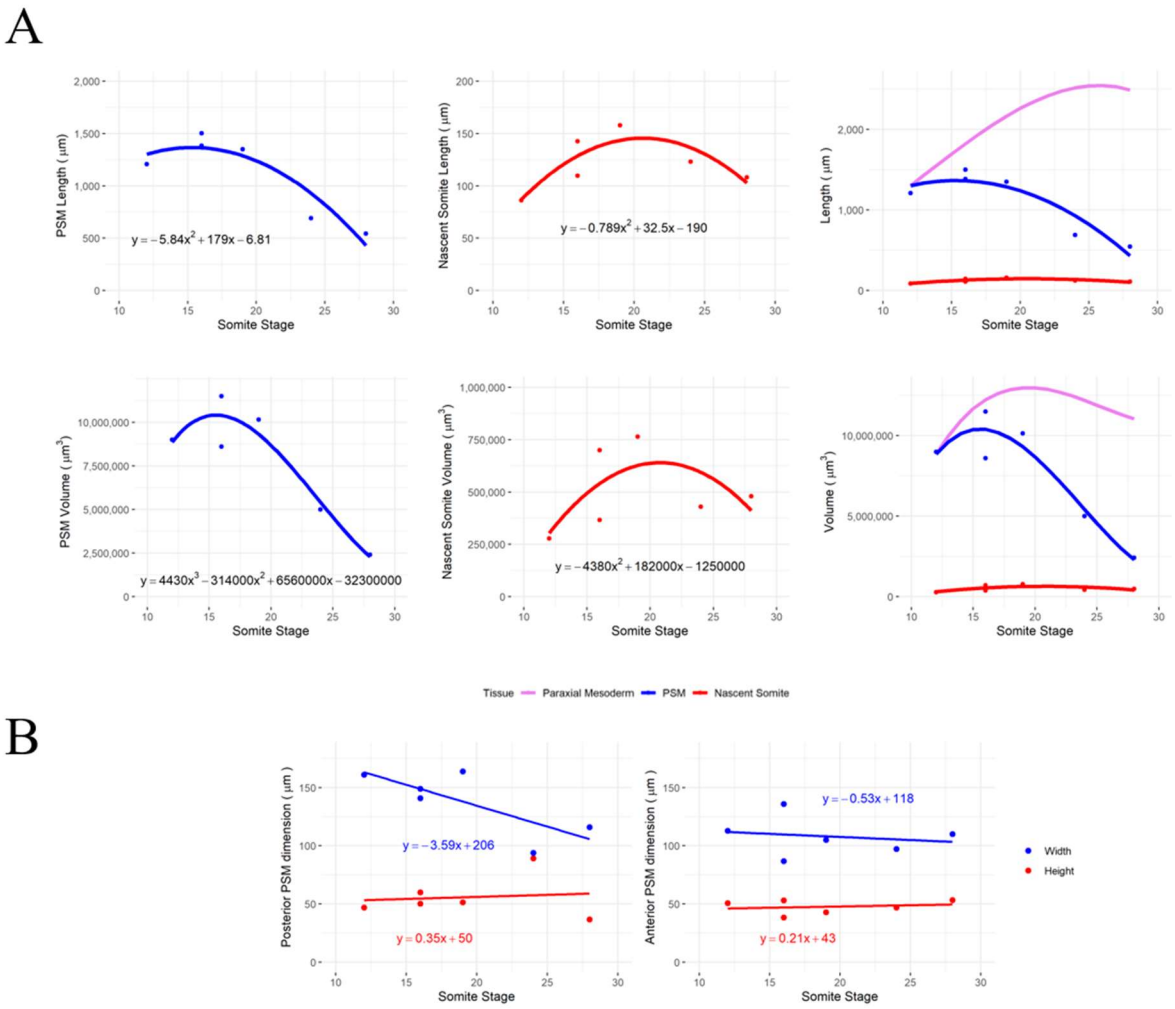


Figure 42: Chick morphometric results. (A) Length (top row) and volume (bottom row) measurements for the PSM (left) and nascent somites (middle). Trendline equations for both were then used to calculate indirect paraxial mesoderm values (right, violet lines) from the 12 to 28 somite-stage. PSM length is initially constant and then decreases. The paraxial mesoderm length increases and then plateaus. PSM and paraxial mesoderm volume increase and then decreases, although in the case of the paraxial mesoderm this slight decrease is likely within the degree of error. (B) Width (blue lines) and height (red lines) of the posterior (left) and anterior (right) PSM. The width of the posterior PSM decreases throughout these stages, whereas the height of the posterior PSM, and both height and width of the anterior PSM, stays constant. n = 6 embryos.

4.4 Catshark paraxial mesoderm elongation

Morphometric methods

Posterior growth-driven elongation is often described as an amniote-specific trait (Kimelman, 2016), however the work of Steventon et al. (2016) showed that catshark embryos (*Scyliorhinus canicula*), which are cartilaginous fishes, undergo posterior growth. Those authors suggested that posterior growth is less likely to be phylogenetically determined, and more likely related to embryo ecology. This species, therefore, represents a good system to test the relative roles of various processes involved in paraxial mesoderm elongation.

Unfortunately, no clear PSM marker candidate genes were found for catshark – no similar genes to tbx6 were identified in the draft genome. Therefore, my catshark morphometric analyses were performed using only YO-PRO (nuclear stain) (Figure 43), to create rough estimations of paraxial mesoderm tissue volumes and dimensions. However, this was informed by some gene expression information – separate embryos were stained for *brachyury* (a marker of axial & paraxial mesoderm progenitors and notochord cells), and this was used to help identify where mesoderm progenitors would be in other tailbuds (Figure 44).

In catshark embryos, the gut tube lies immediately ventral to the notochord and, at the posterior end, the lumen appears continuous with that of the neural tube (whereas in zebrafish the gut tube exits through the yolk stalk, and so is not continuous with tailbud tissues). *Brachyury* expression appears to wrap (laterally) around this lumen, suggesting that mesoderm progenitors reside in the very posterior tip of the tailbud, then move laterally around the neural/gut lumen to join either the notochord or paraxial mesoderm (Figure 44C,D). In my subsequent creation of manual PSM surfaces, I therefore included tissue in the posterior tip (posterior to the neural/gut lumen), and lateral tissue between this and the notochord. As such, a small amount of notochord progenitor tissue is likely included in the PSM surfaces.

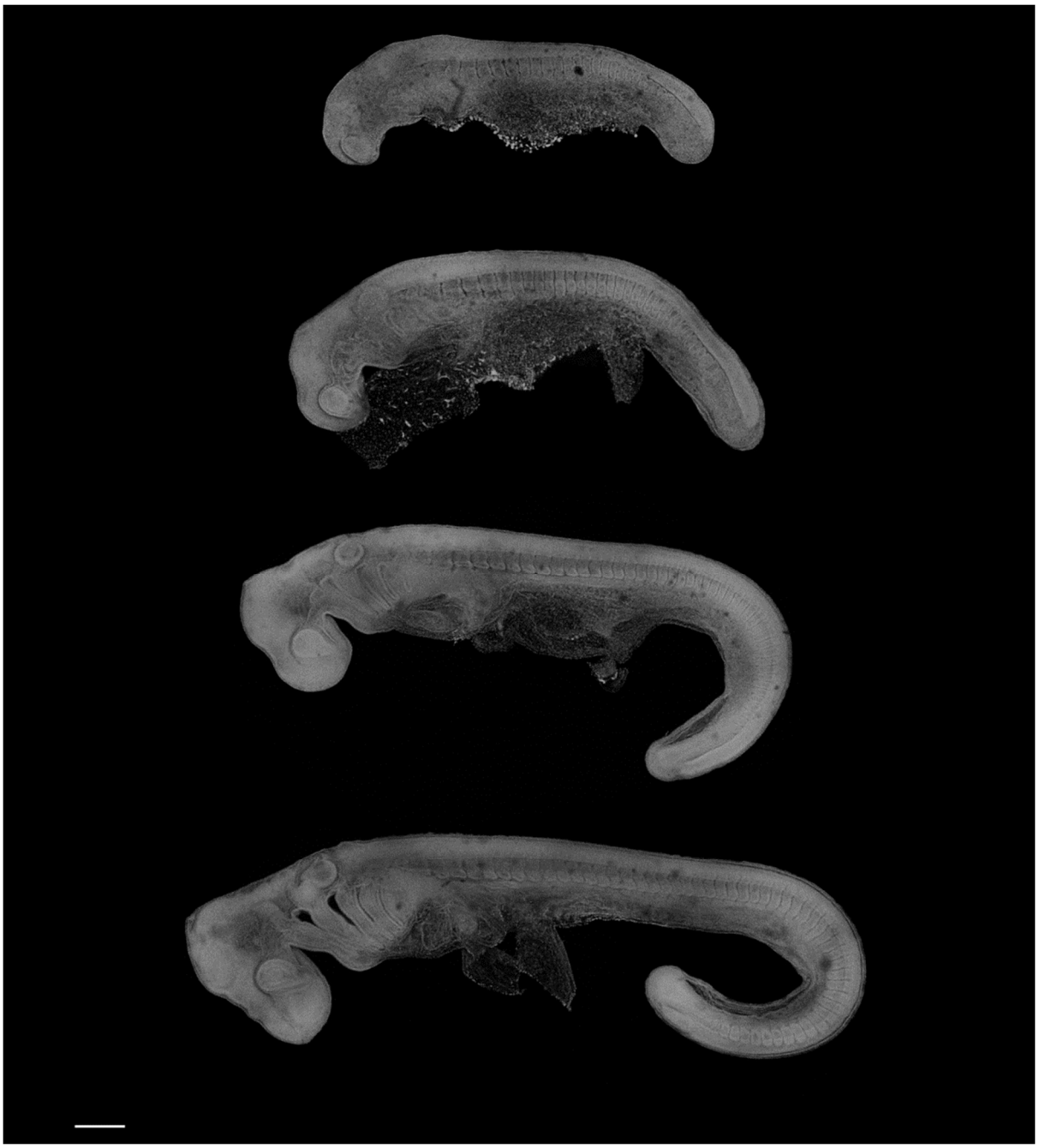


Figure 43: Catshark embryos at different stages of somitogenesis. Embryos were fixed and stained with YO-PRO1 (nuclear stain). Embryos span stages 18 – 22 of standard catshark staging (Ballard, Mellinger and Lechenault, 1993). Specific somite stages: 21, 35, 42, and 51 (from top to bottom). Scale bar: 200 µm.

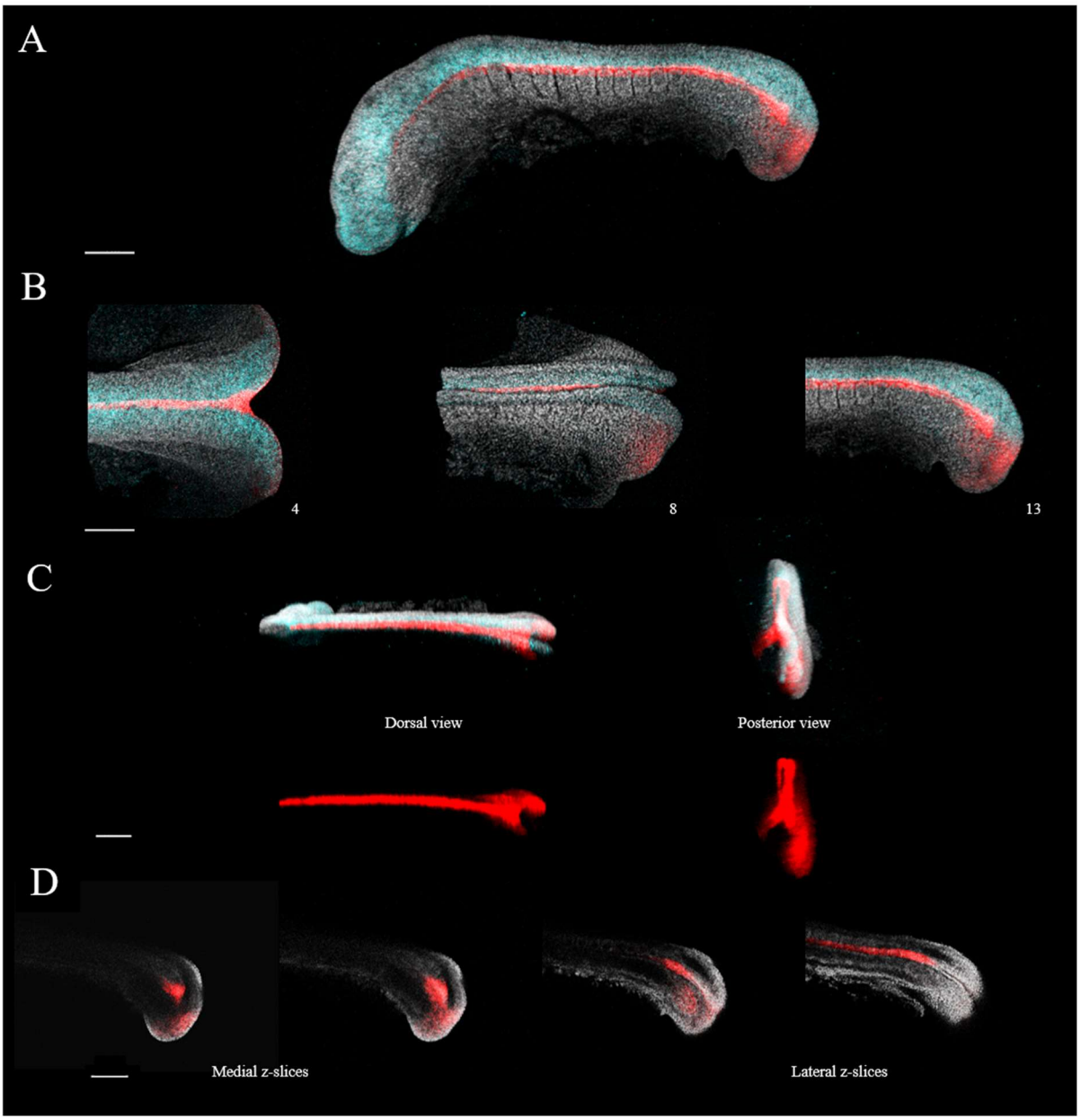


Figure 44: Catshark *brachyury* and *sox2* expression. Embryos were fixed and stained using HCR for *brachyury* (red) and *sox2* (cyan), with nuclei stained with YO-PRO1 (grey). (A) A whole catshark embryo (13 somite-stage) HCR, lateral view. (B) Tailbud formation from the 4 somite-stage (dorsal view) to the 13 somite-stage (lateral view). (C) Dorsal and posterior views show that, at the 13 somite-stage, *brachyury* expression is in the posterior tailbud tip and wraps laterally around the midline to join the notochord. (D) Individual z-slices, going from the midline (left) to more lateral slices (right). At the midline, *brachyury* is expressed in the posterior tip and the notochord progenitors, but not in between. More laterally, expression is continuous between the two, as well as with the posterior PSM. In all images, anterior is left, posterior is right (except posterior view in C). All scale bars: 300 µm.

Figure 45A shows, for one embryo (35 somite-stage), the PSM surface contours drawn at different z-slices, and Figure 45B shows the 3D surface generated, from dorsal and posterior views. It was observed that the paraxial mesoderm has a significant "overhang" which lies lateral to the gut tube (Figure 45C). Because of this, height of the tissue varied between medial and lateral regions (Figure 45D). In the subsequent measurements, therefore, I took separate height (and width) measurements for medial and lateral regions, for both the anterior and posterior PSM (see Figure 47). Figure 46 shows PSM and nascent somite surfaces for a range of stages (21 to 51 somite-stage). Catsharks form ~ 90 somites, and so this analysis does not cover the entirety of somitogenesis but does cover a substantial portion.

As with chick embryos, the large size of these embryos meant that high resolution imaging was not possible. Embryos were imaged using 10x objectives, and so the results presented here only include tissue volumes and dimensions – cell number and tissue density could not be measured. All embryos were imaged from a lateral view, and so only one lateral half of the PSM and one nascent somite was reconstructed for each embryo, as with zebrafish.

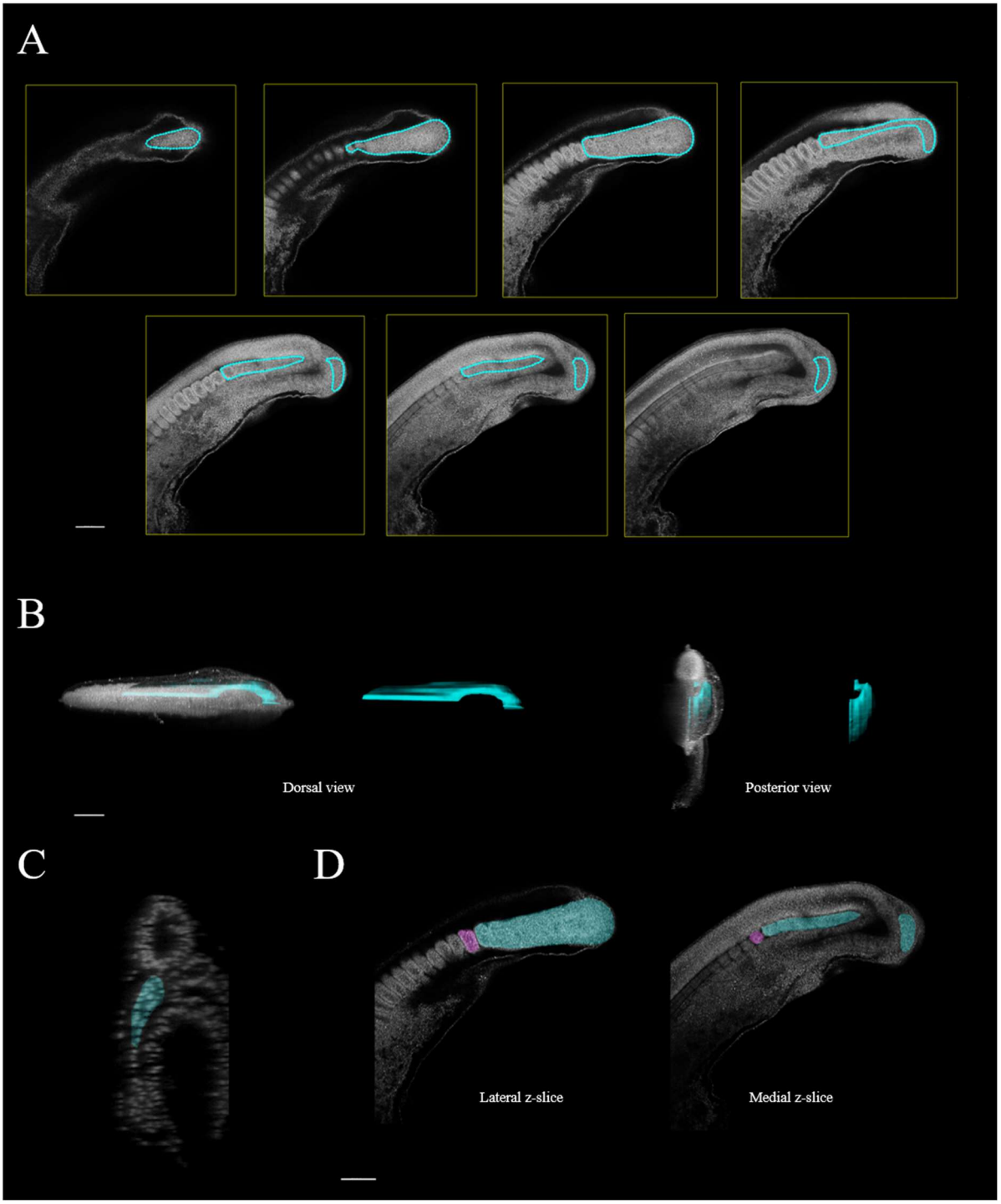


Figure 45: Catshark PSM surface reconstruction. (A) Manual PSM surface contours (cyan) were drawn at different z-slices from lateral (top left) to medial (bottom right). The shared lumen between the neural tube and gut tube is visible in the medial-most image. (B) 3D PSM surface shown from dorsal and posterior views. In dorsal view, lateral is top, anterior is left. In posterior view, dorsal is top, lateral is right. The surface wraps around the neural/gut lumen in the posterior, to resemble *brachyury* expression. (C) An embryo transverse cross-section, showing the "overhang" of the paraxial mesoderm. PSM tissue (cyan) hangs over the gut tube. (D) This overhang results in greater tissue height at lateral regions than medial regions, for both the PSM (cyan) and somites (magenta). All scale bars: 200 µm.

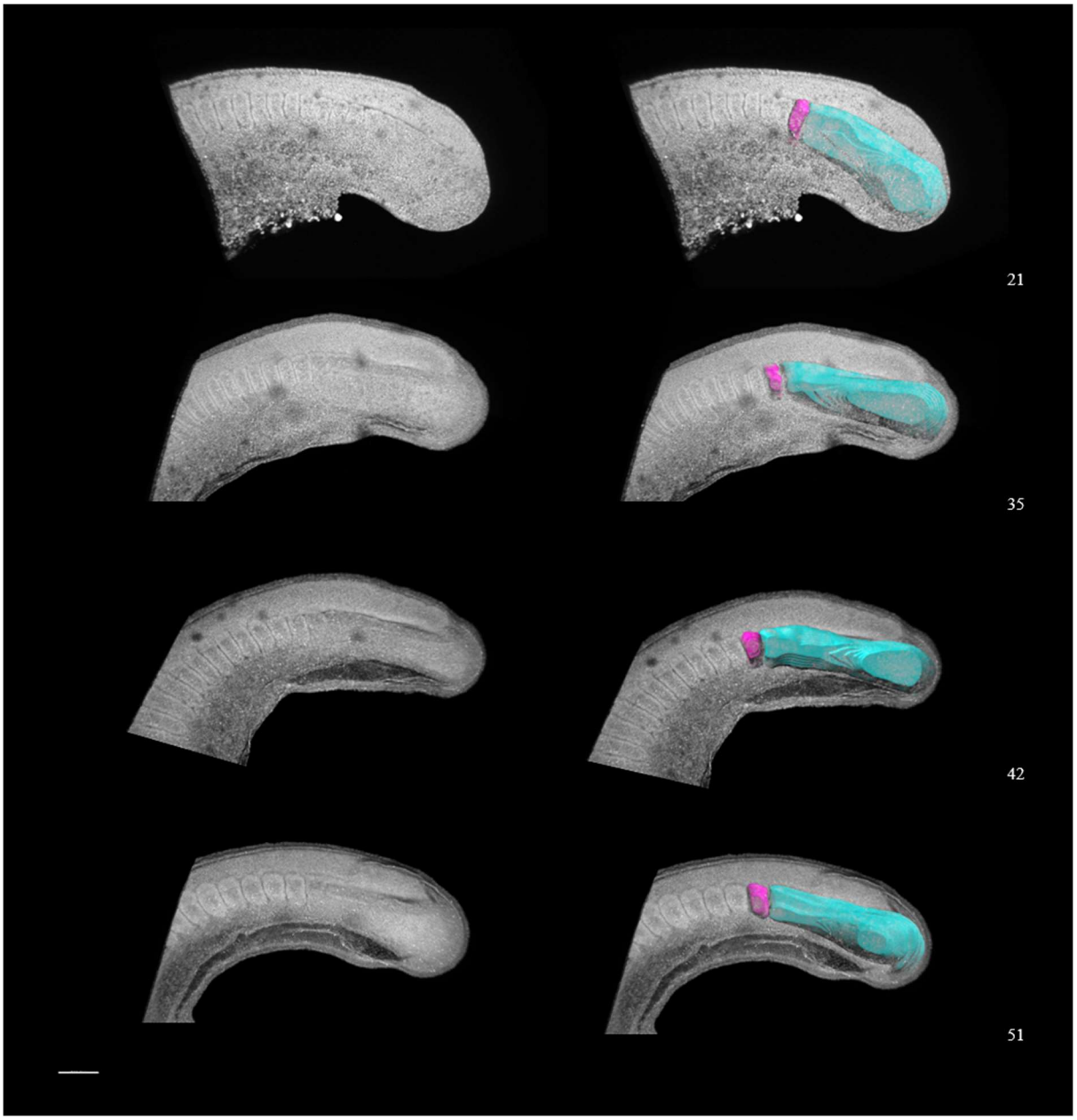


Figure 46: Catshark paraxial mesoderm surfaces across stages. Left images show only nuclei (YO-PRO1, grey). Right images show 3D surfaces of the PSM (cyan) and nascent somite (magenta) at each somite-stage (given at the bottom right of each row). Scale bar: 200 µm.

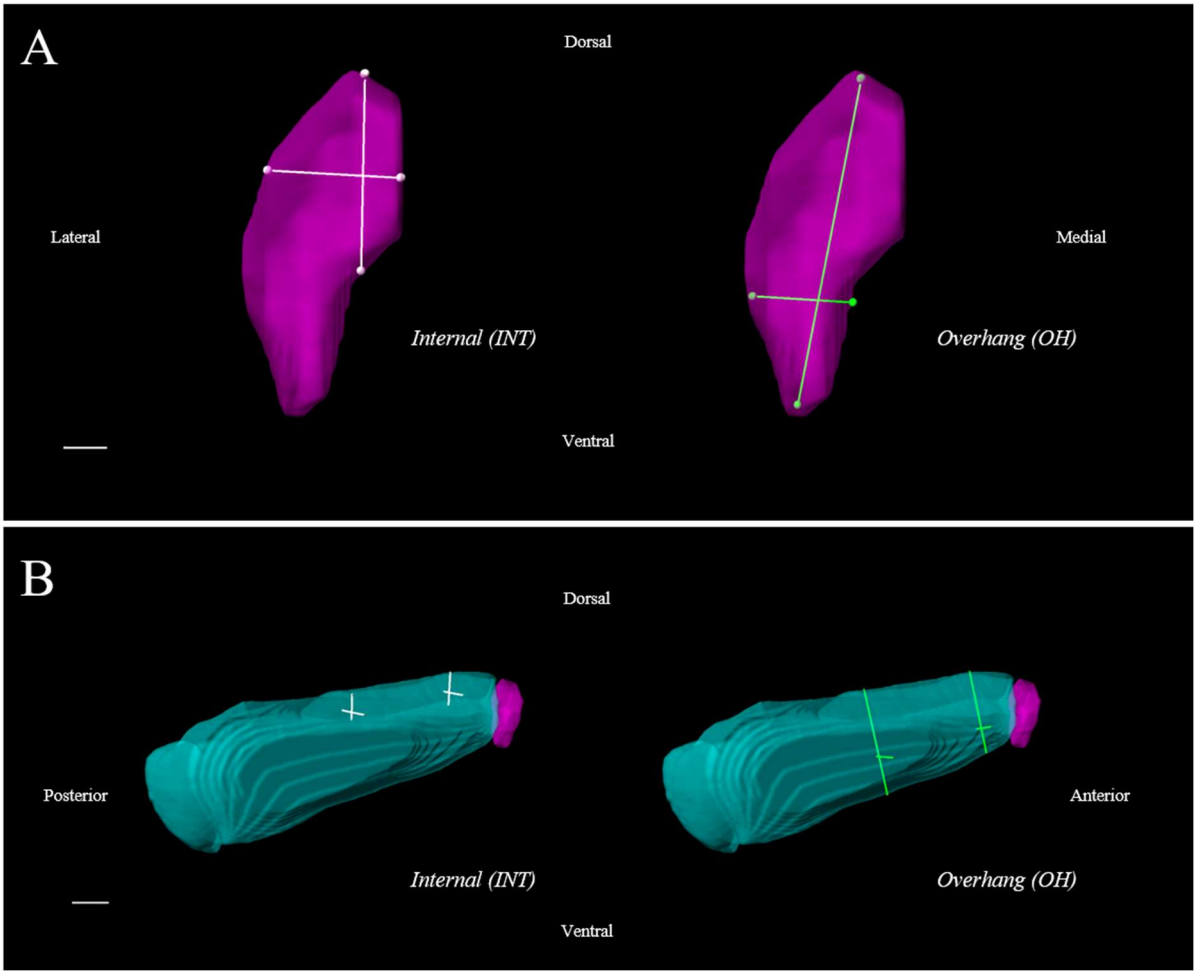


Figure 47: Height and width measurements. Due to the considerable amount of "overhang" of the paraxial mesoderm tissues, separate height and width measurements were taken of the overhang (OH) and the internal (INT) parts of the somite. (A) Internal (white) and overhang (green) height and width measurements shown for a nascent somite (magenta, anterior view). Scale bar: 50 µm. (B) Internal (white) and overhang (green) height and width measurements shown for the PSM (cyan, lateral view). Scale bar: 100 µm.

Substantial growth and a possible minor role of tissue convergence

The results of this morphometric analysis are shown in Figure 48. The length measurements (Figure 48A, top row) show that the PSM maintains a relatively constant length throughout this period of somitogenesis, and that the length of each nascent somite is also relatively constant. The paraxial mesoderm values calculated from these trendline equations show a substantial and continuous amount of elongation. Similarly, the volume of the PSM, and the volume of each nascent somite, remain relatively constant over time (Figure 48A, bottom row). As expected from these results, and from previous work that measured tailbud growth in catsharks (Steventon et al., 2016), the paraxial mesoderm values calculated from these trendlines showed a substantial and continuous amount of volumetric growth, which is likely the main driver of the observed elongation. All stages used in the analysis were post-tailbud formation, which refutes a possible interpretation of previous results, that volumetric growth is always limited to pre-tailbud stages.

Interestingly, some amount of tissue convergence was also observed. While both width measurements (internal and overhang) of the posterior and anterior PSM remained constant, the height (overhang) decreased by ~ 50% over these stages (Figure 48B). While this is not enough to show that convergent extension is happening (as the decrease in height could simply be DV compaction, and not translated into elongation), it does suggest that it could be happening. While volumetric growth is likely to be the main driver of elongation, it is possible that this tissue convergence also plays a minor role.

As with the chick analysis, these findings are taken from a small sample size, and so further work would be required to conclusively prove these findings.

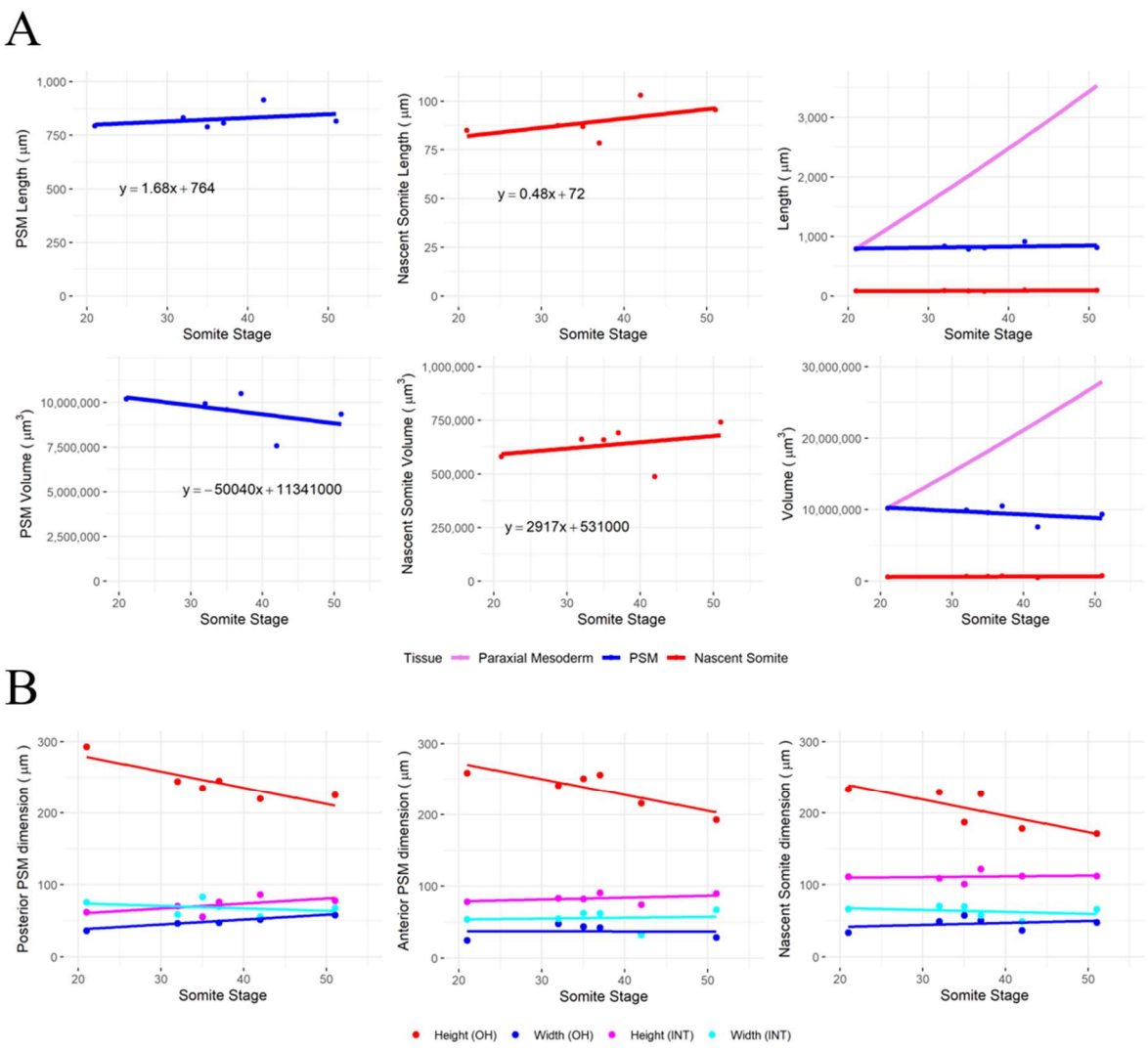


Figure 48: Catshark morphometric results. (A) Length (top row) and volume (bottom row) measurements for the PSM (blue lines) and nascent somites (red lines), along with indirect paraxial mesoderm values (violet lines) calculated from those trendlines. PSM length and volume stay relatively constant over time, despite depletion to form somites, due to continuous, substantial amounts of elongation and volumetric growth. (B) Height and width measurements for the posterior and anterior PSM (left, middle), as well as the nascent somites (right). Two sets of measurements were taken for each dimension, one internal set (INT) excluding the overhang, and one of the overhang (OH) (see Figure 45 & Figure 47, and text, for more details). While most height and width measurements remained constant, tissue height (including the overhang, red lines) decreased by ~ 50% for both the posterior and anterior PSM (also reflected in the somite heights). This suggests that DV convergence may also contribute to elongation. n = 6 embryos.

4.5 Conclusions

The results of this comparative analysis (summarised below in Figure 49) show that compression-extension of the paraxial mesoderm is not unique to zebrafish but is also occurring in cichlids. While there were observed differences between the two cichlid species (with *R. chilingali* showing no height decrease or density decrease), these may be the result of sample sizes/measurement error. Regardless, the conservation of convergent extension (in general terms) between zebrafish and cichlids is interesting. Both are teleosts, but this is a huge and ancient group of vertebrates; cichlids and zebrafish are separated by ~ 250 million years of evolution (Near et al., 2012). Morphometric measurements in chick embryos suggests that, even in species that undergo large volumetric growth, there may be some degree of convergent extension occurring (both during the growth phase and after it). This suggests that, rather than convergent extension and volumetric growth representing two opposite modes of elongation, both are conserved features in vertebrate development – although the relative contributions are different between species. This is further supported by the catshark morphometric measurements, which, while showing substantial volumetric growth, also show some tissue convergence. Together, these findings suggest that convergent extension – rather than being unique to zebrafish - is a conserved process of paraxial mesoderm elongation, but that its relative contribution (relative to other processes like volumetric growth) has evolved differently among lineages. Because cell number could not be accurately measured in chickens and catsharks, it is not clear whether tissue convergence in these species is due to compression-extension, or another form of convergent extension.

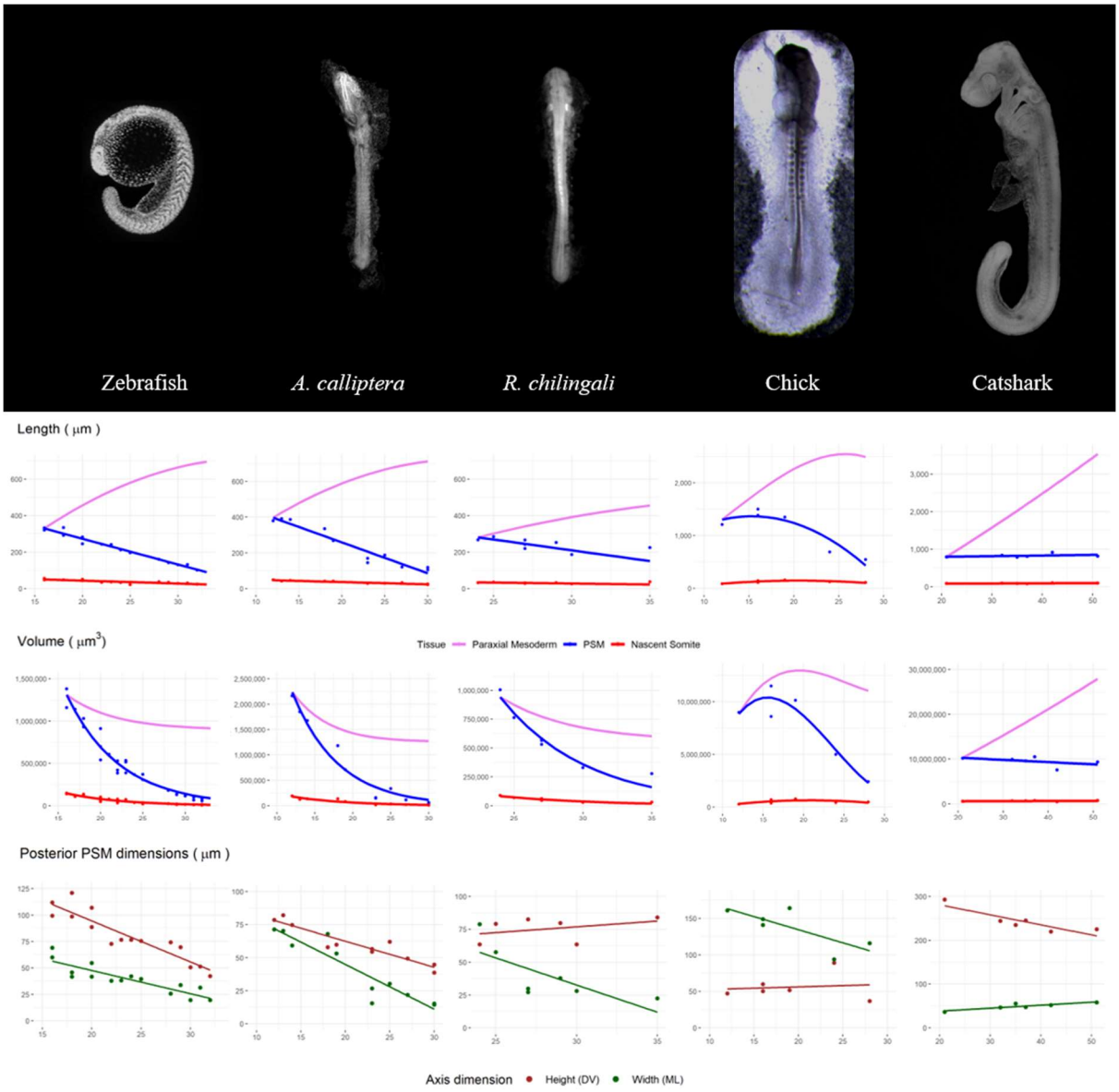


Figure 49: Summary of comparative morphometric results. Images of all five species are shown (not to scale), with three graphs per species shown below: paraxial mesoderm tissue lengths, paraxial mesoderm tissue volumes, and posterior PSM dimensions (height and width). In all graphs, the x-axis corresponds to somite-stage (note that the scale is different among species). For all species, paraxial mesoderm length (indirect measurement) increases over time. Paraxial mesoderm volume (indirect measurement) decreases over time in zebrafish and cichlids, increases then decreases slightly in chicks, and increases over time in catsharks. Posterior PSM height and width both decrease over time in zebrafish and *A. calliptera*, while only the width decreases in *R. chilingali* and chicks, and only the height decreases in catsharks (height and width measurements shown for catshark are of the lateral PSM i.e. the overhang).

DISCUSSION

Compression-extension and paraxial mesoderm elongation

The results I have presented in this thesis conclusively show that the zebrafish paraxial mesoderm does not elongate its axis through volumetric growth or cell addition. In fact, my results show that, despite a small overall increase in the number of cells, the tissue does not expand in volume at all – but instead decreases in volume and increases in density. This refutes existing assumptions that all vertebrates elongate their axis by the process of posterior growth (Martin and Kimelman, 2009). My results show that the posterior PSM (the region containing *mgsn1*-expressing progenitor cells) undergoes a substantial amount of height (DV axis) and width (ML axis) reduction, and my simple 3D geometric modelling approach shows that this convergence is sufficient to explain the observed elongation. Together, these results suggest that some form of convergent extension is driving elongation.

This convergence, which is largely limited to the posterior PSM, does not fit the classical definition of active convergent extension. Although photolabels show that regions of cells converge in the DV axis and extend in the AP axis, tracking analysis shows that neighbouring cells are not locally rearranging with each other through junctional rearrangement/directional intercalation. Instead, cells displace towards the tissue midline in "convergent flows", which involve non-directional rearrangements between neighbours. Additionally, perturbation experiments with DEP+ show that paraxial mesoderm elongation does not require functional Planar Cell Polarity (PCP) signalling during post-gastrulation stages. This shows that the tissue convergence during these stages is not simply a continuation of gastrulation movements over the yolk – which do involve PCP signalling and directional intercalation between neighbours (Roszko, Sawada and Solnica-Krezel, 2009). Rather, it appears that, once the PSM is incorporated into a fully formed tailbud (between the 12-16 somite-stages), its

mechanism of elongation changes from PCP-dependent directional intercalation to PCP-independent convergent flows.

Together, these results support a hypothesis that tissue convergence during axis elongation is driven by external compression of the tissue – which also results in tissue and cell compaction and an increase in tissue density. I therefore propose a novel form of convergent extension: compression-extension. Whereas classical convergent extension is driven by cell-level mechanisms (PCP-dependent directional intercalation), compression-extension is driven by a tissue-level mechanism of compaction, which results in convergent flows and non-directional intercalation.

The 4D agent-based model of compression-extension (presented in Section 2.6) confirms that, when cells displace towards the midline of a tissue, this causes tissue convergence and elongation, but does not involve directional intercalation between neighbours. Cell tracks from this model displayed highly similar behaviours to cell tracks in the embryo: displacement towards the midline with random rearrangements between neighbours. Cell pairs only showed evidence of orientating along the AP axis when tracked over extremely long time periods, by which point cells were no longer neighbours, but had dispersed along the AP axis. As such, this AP orientation of cell pairs is simply a consequence of tissue morphogenesis: all cell pairs would eventually orientate along the AP axis, as the tissue continues to converge in other axes and extend in this axis.

This agent-based model, together with the real PSM tracks, shows that directional intercalation between neighbours is not necessary for all forms of convergent extension. Importantly, this also shows that measuring directional intercalation alone is not enough to refute the occurrence of convergent extension – only the occurrence of classical/active convergent extension. Rather than measuring cell-level details, tissue-level dimension measurements must also be taken, as convergent extension is a tissue-level process – which, depending on the form of convergent extension, may or may not involve cell-level

processes like junctional rearrangement. This idea challenges a prevailing reductionist view that tissue-level changes are simply a consequence of cell-level processes.

My tracking analysis, through the use of separate biological frames of reference, also highlights significant flaws in the notion that PSM cells are migrating anteriorly (Manning and Kimelman, 2015). While all cells do displace anteriorly relative to the posterior tip of the tail, these same cells all displace posteriorly relatively to a given somite. This emphasises the importance of the reference point that cells are tracked relative to; and highlights the need to recognise this importance when drawing conclusions from cell tracks. Defining AP movement as anterior vs posterior in an elongating tissue is arbitrary; and depends solely on which position along the tissue cells are tracked relative to. Instead, it seems that cell displacement along the AP axis is a consequence of convergence, as cells move towards the midline and displace each other along the AP axis, rather than an active process of anterior migration.

Balancing compaction with elongation

The idea that a process (compression) which causes tissue shrinkage could simultaneously cause tissue elongation is, admittedly, counterintuitive. However, it is possible that the physical properties of the cells determine how much they are compressed in volume by tissue compaction, and how much they are displaced by it. These physical properties (e.g. how rigid the membrane is) would therefore be important in this balance between compaction and elongation. If cells were not rigid enough, tissue compression would simply result in all cells decreasing in size, and no displacement (therefore no elongation). On the other hand, if cells were highly rigid, tissue compaction would only result in displacements, and cells would stay the same size – or in extreme cases, tissue compaction would not even possible, and there would be no displacements. Therefore, the physical properties of these progenitor cells would be extremely important in determining the amount of elongation. Co-expression of *tbx16* and *msgn1* in the

posterior PSM causes these cells to become more mesenchymal than they otherwise would be (Manning and Kimelman, 2015), and mutants of these genes display large swollen tailbuds – hence the original tbx16 mutant name of *spadetail* (Griffin et al., 1998). Perhaps then, these genes help control the compressibility of progenitor cells, in order that tailbud compaction can drive the correct amount of elongation. Taking this hypothesis, this would mean that mutants for these genes have highly incompressible cells – which means that tailbud compaction cannot occur, and therefore no cell displacements/tissue-level convergent extension can occur. This hypothesis is not only consistent within itself, but also with the work of Mongera et al. (2018), which characterised the difference in stiffness/rigidity between posterior and anterior PSM. Those authors argued that the fluidity of the posterior PSM is essential for allowing high levels of cell addition to be translated into unidirectional elongation. Compression-extension, on the other hand, does not require high levels of cell addition. Instead, in this process, the controlled fluidity of the posterior PSM is essential for allowing tailbud compaction to drive convergence and extension.

It is of course possible that tissue compaction is the result of cell shrinkage, rather than the cause. However, if compression was being driven by cells contracting, it is unlikely that this would lead to cell displacements and elongation. On the other hand, it is much more feasible that tissue-level compaction could lead to a combination of cell shrinkage, cell displacements, and tissue elongation. Therefore, understanding the mechanism of PSM compression – and identifying the level that this is happening at – could help answer the question of whether this compression is driving elongation (as I propose), or if it is simply a co-occurring process.

The source(s) of compression

Posterior neural tube ablations appeared to have no effect on paraxial mesoderm elongation, and previous work has shown similar results with posterior notochord

ablations (Mclaren and Steventon, 2021). However, given that – at least in the case of posterior neural tube ablations – tissue volume is not affected by the ablation, it is possible that dead, ablated cells could still exert a pushing force on the PSM because of more anterior cells pushing them posteriorly. Therefore, from these laser ablations alone, I cannot rule out a possible role of other tissues in driving PSM compression.

Cdx4 mutants, in which neural tube development was significantly perturbed, did show a reduction in paraxial mesoderm length, as well as no significant difference in density between developmental stages. Together, these two results could suggest that the neural tube contributes to PSM compression-extension. However, due to small sample sizes and a highly variable phenotype, this is not conclusive. Further experiments would be required involving live imaging to measure elongation of individual embryos (to remove confounding effects of phenotype variability). Measuring density changes in live embryos would be more difficult, due to the necessity of using high resolution objectives to obtain accurate cell counts, but could theoretically be done.

TGFβ signalling appears to play a role in regulating/increasing PSM density – as shown by reduced PSM density in SB50-treated embryos. This would suggest that tailbud swellings, which were infrequently observed in SB50 treatments, could be extreme cases of reduced density. Given that proliferation was not affected in treated embryos, this is the most likely explanation. However, the mechanism by which TGFβ regulates/increases PSM density is not clear. It may relate to the above role of extrinsic neural tube compression, or it may relate to a separate intrinsic process within the PSM itself.

There are, of course, other possible sources of compression. The PSM is surrounded by a sheath of extracellular matrix (ECM), largely composed of fibronectin (Koshida et al., 2005), and, without PSM-ECM interaction, body elongation is greatly reduced (Dray et al., 2013). If compression is indeed causally linked to elongation, this could suggest that the ECM surrounding the PSM is driving compression-extension by constricting

the posterior PSM. Again, however, it is not clear what the mechanism of this would be, in terms of what could cause the ECM to actively constrict. One possibility is that the continuous expression of fibronectin genes in the PSM causes accumulation of fibronectin, and a gradual increase in the amount/thickness of ECM surrounding the tissue, and this then causes compression. Although it has been shown that accumulation of fibronectin at the somite boundaries is required for the maintenance of these boundaries (Koshida et al., 2005), the amount of fibronectin around the PSM (and whether or not this increases throughout somitogenesis) has not been quantified. Quantifying this, and disrupting fibronectin production, would help determine the importance of the ECM in PSM compression. Given how fundamental ECM is to the development of the entire embryo, perturbation would ideally be specific to the PSM/paraxial mesoderm, and adjustable in terms of degree (e.g. using a heat-shock dominant negative construct and varying the duration of the heat shock, rather than using mutants). It is also possible that the SB50 phenotype is due to disruption in ECM formation, and so quantifying the amount of fibronectin around the PSM of SB50-treated embryos could elucidate the mechanism by which TGFβ signalling regulates PSM density.

Compression-extension vs passive convergent extension

So far, I have compared compression-extension with active convergent extension. However, there is another (often overlooked) form of convergent extension: passive convergent extension. This is defined as "a passive response to force generated elsewhere… [causing the tissue] to be stretched and narrowed by the active tissue" (Keller et al., 2000). A clear example of this is seen in the chick notochord, which is compressed laterally by the PSM to undergo passive convergent extension (Xiong et al., 2020). Is compression-extension the same process, therefore, as passive convergent extension? Given that the source of force generation in zebrafish paraxial mesoderm compression-extension is not clear, this remains an open question. If the PSM is completely passive

and only compresses due to forces from other tissues (e.g. the neural tube), then this would fit the definition of passive convergent extension. However, in the work of Xiong et al. (2020), compaction (in terms of volume reduction) of the notochord tissue/cells was not shown. If this is not occurring, then the compression-extension observed in zebrafish represents a different process from the passive convergent extension observed in chickens – or at least is a specific type of passive convergent extension. Measuring cell volumes and tissue density of the chick notochord over long periods of somitogenesis would help answer these questions of definitions. Bénazéraf et al. (2017) measured the density of the chick notochord (and other tissues) and observed no clear density increase. However, the analysis focussed on a period of four somite-stages. As the density increase in zebrafish is a slow, gradual process, a longer period of measurement would be required in chick embryos to provide certainty either way.

Random motility vs chaotic motility

The result that directional displacements (towards the midline) involves non-directional rearrangements between neighbours raises the question: how do we define "random"? Is it accurate to describe cell rearrangements as "random" because we do not see a clear, simple pattern; or is it better to say that cell rearrangements are highly complex and chaotic due to convergence in multiple axes of thousands of cells? This question could also be extended from cell rearrangements to cell movement. Is the reported "random motility" of cells in chick embryos (Bénazéraf et al., 2010; Xiong et al., 2020) truly random, or is it just too complex to show clear, simple patterns? While this may seem like a purely philosophical distinction, it is important because, in order to understand development, we need to be able to distinguish between mechanisms and observations. Are cells being instructed in some way (either by external signals, or intrinsic gene expression, or both) to move in any/all directions? Or, are cells being instructed to move (or forced to move by physical forces) in a specific direction, but

cannot do so in a straightforward, linear manner because of the movements of cells around them?

As an analogy, we can imagine a large crowd of people all trying to move through a narrow gate. This will involve some degree of "convergent extension": those on the edge of the crowd will need to move towards the centre (converge) in order to move through (extend). However, despite this being a relatively simple process, if we were to "track" the movement of an individual woman in this crowd, it would likely appear highly chaotic. As she moves towards the centre, others around her are doing the same, and so she may be pushed to the sides or even back the way she came, as individuals on the other side of the crowd push towards her. So even though she has a highly directional intended path, her movement could appear random. This is only a 2D analogy – a crowd of people can only move in the x and y axes – whereas PSM cells are doing this in 3D. Therefore, there is likely to be even more chaos in individual tracks, as each individual cell is not only displaced by cells to the front, back, and sides of them, but also by cells above and below them.

In the above analogy, the timescale of analysis would be very important. If movement was only observed over a few seconds, it would not be at all clear what the intended path of each individual was. However, after a few minutes, it would be clear – despite some deviation from this due to the individual being pushed around by others. In the case of cells, analysis timescale is also very important. Cell displacements in the DV and ML axes (Section 2.4) showed clearer convergence trends over longer timescales. After 2 hours of development, these trends were clear (although the ML axis trend was still only very slight), but over short timescales (30 min) they were not. This supports the idea that reports of "random" motility could simply be *chaotic* motility, observed over timescales that are not long enough to see the true "intended" paths emerge. In other words, perhaps – like gene expression – cell movements are "noisy", with fluctuations and deviations from set trajectories.

This brings us back to the importance of distinguishing mechanisms from observations. Although cell movements may appear random, this does not mean that the mechanism is randomness. Random motility may be a transient property of highly directional processes like tissue convergence. This is especially important when thinking about the evolution of these processes. It is not clear how a completely random developmental process could be altered in evolution to provide variation. On the other hand, the compressibility of cells and the degree of tissue convergence are developmental "parameters" that could easily change (through selection or drift) to provide variation in axis length and segment number. Given the necessity for developmental processes to be both robust and evolvable, a truly random developmental mechanism seems unlikely.

Tissue convergence is conserved across vertebrates

Importantly, the cichlid results presented here show that paraxial mesoderm compression-extension is not unique to zebrafish – but is also occurring in at least one cichlid species. *A. calliptera* showed highly similar results to zebrafish, with tissue convergence in height and width, a decrease in tissue volume and an increase in tissue density. If compression is driving elongation in zebrafish, this would suggest that it is also driving it in *A. calliptera*. *R. chilingali*, on the other hand, also showed some degree of tissue convergence – but only in width, not in height. Additionally, while tissue volume did decrease, this was not associated with a decrease in tissue density. Given that both sample size and stage range were smaller in *R. chilingali,* it is not clear if these are genuine differences, or simply the result of slight measurement error. If these do represent genuine differences between *A. calliptera* and *R. chilingali*, this raises an interesting point about evolutionary conservation and convergence. These species are very closely related (separated by less than 1 million years of evolution (Malinsky et al., 2018)), while they are very distantly related to zebrafish (separated by ~ 250 million years of evolution (Near et al., 2012)). The possibility, then, that a developmental process has been

conserved between zebrafish and *A. calliptera* for 250 million years, but then changed over a few hundred thousand years along the *R. chilingali* lineage, seems very unlikely. Rather, it seems more likely that these developmental processes are highly evolvable - and may have been altered multiple times along each lineage. This would suggest that close similarities between zebrafish and *A. calliptera* are the result of convergent evolution. This is not to say that the general process of convergent extension is not a conserved developmental process – but, rather, that the particular details could be highly evolvable, and just so happen to be very similar, at this point in evolutionary time, between these two species.

However, given the uncertainty of the *R. chilingali* results, it is important not to over-interpret the differences between *R. chilingali* and *A. calliptera.* It is possible that these species are more similar than the results suggest, and that these developmental processes are deeply conserved. Regardless, the general conservation of tissue convergence and growth-less elongation is clear and important: these processes are not unique to zebrafish.

The stark difference between the primary mechanism of paraxial mesoderm elongation in zebrafish and cichlids (convergent/compression- extension) and that of mice (volumetric growth) could suggest that two distinct "modes" of elongation have evolved in different lineages (anamniotes vs amniotes). However, a more parsimonious explanation is that both convergent/compression- extension and growth are conserved processes in paraxial mesoderm elongation, and that natural selection has only altered the relative contributions of these processes. This would suggest that, rather than zebrafish/cichlids and mice representing two different "modes", they are simply on opposite ends of a spectrum from no growth/high convergence to high growth/low convergence. This finding is supported by the results in chick and catshark embryos.

My results show that, in chickens, the paraxial mesoderm is not constantly growing throughout somitogenesis, but instead undergoes growth at early stages and then volume

remains constant/decreases slightly at later stages. However, elongation is still occurring at these later stages (although it is slower). This suggests that processes other than growth must be involved in, and sufficient for, elongation. The observation that the posterior PSM continuously decreases in width points to a possible role of convergent extension. Importantly, this width convergence does not begin at the point that growth stops; but is happening throughout somitogenesis. This suggests that, rather than there being an early growth phase and a later convergent extension phase, convergent extension could be occurring throughout – even when growth is also occurring.

Additionally, my catshark results also show that, despite a constant, substantial volume increase, there appears to be some level of tissue convergence – with the lateral overhang region of the PSM decreasing in height. This suggests that, while volumetric growth is undoubtedly the main mechanism of paraxial mesoderm elongation, some form of convergent extension may provide a minor contribution.

Importantly, if convergent extension is occurring in chicken and catshark embryos, there is no reason to expect that this will meet the classical definition of active convergent extension. It is far more likely that if it is occurring, it will resemble the process in zebrafish, which is the result of large convergent flows with non-directional intercalation. As imaging restrictions made it impossible to accurately estimate cell numbers in these species, it was not possible to measure tissue density across stages. As such, it is unclear whether tissue convergence could be driven by – or at least correlated with – compression of the tissue. While it may seem counterintuitive, it is feasible that compression and volumetric growth could co-occur – as compression would only be occurring in one or two axes, and not the AP axis. Therefore, cells could be dividing and growing, while being compressed in the ML and/or DV axes. Indeed, given that there is such a large amount of growth in the catshark PSM, tissue compression could even help ensure that growth is translated into unidirectional elongation, rather than isotropic expansion.

As mentioned in the previous section, evolvability/adaptability is an important quality for developmental processes to have. If paraxial mesoderm elongation had a clear either-or mechanism (growth vs convergence), this would not allow for much adaptability. On the other hand, if, as a conserved feature of elongation, both mechanisms were involved, this would allow for much greater adaptability. If, as suggested by Steventon et al. (2016), volumetric growth is associated with high maternal investment, then the relative contribution of tissue convergence could increase in a population evolving from high to low maternal investment (or vice versa). In this way, paraxial mesoderm elongation could remain robust to ecological changes, due to adaptability of underlying processes.

Robustness and evolvability: the leftover PSM

The observation that not all of the PSM in zebrafish (and, it seems, in cichlids) segments into somites raises several questions. What is the mechanism by which somitogenesis ends, given that it is not simply PSM depletion? And what is the purpose of this leftover PSM?

In chick embryos, the end of somitogenesis is associated with (and therefore could be brought about by) apoptosis in the PSM (Olivera-Martinez et al., 2012). However, this has not been observed in any other species – suggesting that it may be a unique feature of bird development, due to the evolutionary reduction in tail vertebral number. In zebrafish, chickens, and mice, expression of wavefront genes (*fgf8* and/or *wnt3a*) decreases and ceases towards the end of somitogenesis (Zhang et al., 2018; Olivera-Martinez et al., 2012; Cambray and Wilson, 2007). Given the positive interactions between these genes and clock genes, this expression decline is likely sufficient to explain how somitogenesis ends – although it is not clear what causes this expression decline.

The observation that the zebrafish leftover PSM contributes to fin mesenchyme in much the same way as somites do (Lee et al., 2013) suggests that there is nothing unique

about the contribution of leftover PSM cells. In other words, there is no clear reason for some paraxial mesoderm to remain unsegmented, given that it does not appear to perform a unique function from segmented paraxial mesoderm. Why, then, does somitogenesis finish "early" (before the PSM has been fully segmented)? Perhaps this leftover PSM acts as a buffer of developmental defects. If embryo elongation is perturbed somehow, this "extra" amount of tissue might ensure that the correct number of somites (of the correct sizes) still form. In other words, perhaps the leftover PSM *is* segmented into somites in the case of elongation defects. This would allow robustness of somite/vertebral number.

Additionally, this "extra" tissue could facilitate evolvability of somite number. In zebrafish, somite number could increase by ~ 3-4 somites by maintaining wavefront/clock gene expression for longer, without the need for more tissue. Similarly, the fact that the end of somitogenesis is not dependent on all of the PSM being segmented, means that somite number could decrease without the need to decrease the amount of tissue. This system is much more evolvable than one in which all of the PSM is segmented and the end of somitogenesis is determined by this point.

Vertebral number in vertebrates

To return to the wider evolutionary questions: how important is axis elongation in the specification of, and evolution of, vertebral number? Vertebral number is largely determined by somite number, which is thought to be determined by a combination of somitogenesis rate and axis elongation rate (Gomez et al., 2008; Gomez and Pourquie, 2009). However, it has been shown that, at least in frog and zebrafish embryos, somite size scales with embryo size/PSM length (Cooke, 1975; Ishimatsu et al., 2018). Embryos that have been surgically reduced in size still form the correct number of somites. Therefore, it could be argued that, since axis length does not affect somite number, axis elongation is not important in determining somite number. However, Ishimatsu et al.

(2018) measured both somitogenesis rate and axis elongation rate in these size-reduced zebrafish embryos, and found that neither were different between control and size-reduced embryos. It is likely, therefore, that while embryo size/axis length can change during evolution without any effect on somite number, the *rate* of axis elongation (relative to somitogenesis rate) does affect somite number – and changes to this rate in evolution will change somite number, and thus, vertebral number. Therefore, understanding the mechanisms of axis elongation, and how these mechanisms evolve, is an important part of understanding how vertebral number evolves.

Final conclusions

In this thesis, I have shown that the zebrafish paraxial mesoderm does not elongate through volumetric growth. Instead, my results show that elongation is driven by a form of convergent extension that does not involve PCP-dependent directional intercalation, but instead involves convergent flows of cells to the midline and non-directional intercalation. As my measurements show a co-occurring decrease in cell and tissue volumes (with a consequent increase in tissue density), I propose that this form of convergent extension is driven by tissue compression, hence the proposed term "compression-extension". The source of compression is not clear, although my results suggest that the neural tube may be involved. TGFβ signalling also appears to play a role in the increase in tissue density, but the mechanism by which it does this is unknown. Other sources, including the ECM, may also contribute to compression-extension of the paraxial mesoderm.

The comparative work I have presented here also shows that at least some of the above findings are not unique to zebrafish. Two species of cichlid fishes also undergo paraxial mesoderm elongation that appears to be driven by tissue convergence, and at least one of these species displays an increase in tissue density – suggesting a common process of compression-extension. Chick embryos do display volumetric growth, but not

throughout axis elongation. Tissue convergence is happening throughout, suggesting that convergence and growth could be combining to drive elongation at early stages, with convergence solely driving elongation at later stages. Similarly, catshark embryos undergo a substantial amount of growth, which is likely the main driver of elongation, but they also undergo some degree of tissue convergence, which could be contributing a minor amount to elongation. Together, these findings suggest that both posterior growth and tissue convergence could be conserved mechanisms of paraxial mesoderm elongation, and that the relative contributions of these two mechanisms have evolved differently across vertebrate lineages, resulting in a spectrum of elongation strategies.

Axis elongation is an important part of embryo development, but one which is not well understood. There exists a widespread assumption that this process, in all vertebrates, is primarily driven by posterior growth. However, it is clear that the mechanisms are more complex (and more interesting) than this. Understanding this process is important, not only because it is a major aspect of development, but because of how many other developmental processes are linked to this. Fate specification and differentiation, somitogenesis, and axial patterning all take place in this morphologically dynamic context, and so, understanding axis elongation is essential to better understand these processes. It is also important to understand how axis elongation evolves, in order to help explain the vast diversity in vertebral number across vertebrates. Segmentation is undoubtedly one of many features that has allowed the vertebrate body to diversify so greatly. Therefore, the more we understand the developmental processes regulating vertebral number, the more we can understand – and appreciate – the astonishing diversity of vertebrates.

References

Abu-Abed, S., Dollé, P., Metzger, D., Beckett, B., Chambon, P. and Petkovich, M., 2001. The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. *Genes and Development*, 15(2), pp.226–240.

Aoyama, H. and Asamoto, K., 2000. The developmental fate of the rostral/caudal half of a somite for vertebra and rib formation: Experimental confirmation of the resegmentation theory using chick-quail chimeras. *Mechanisms of Development*, 99(1–2), pp.71–82.

Asher, R.J., Lin, K.H., Kardjilov, N. and Hautier, L., 2011. Variability and constraint in the mammalian vertebral column. *Journal of Evolutionary Biology*, 24(5), pp.1080–1090.

Attardi, A., Fulton, T., Florescu, M., Shah, G., Muresan, L., Lenz, M.O., Lancaster, C., Huisken, J., van Oudenaarden, A. and Steventon, B., 2018. Neuromesodermal progenitors are a conserved source of spinal cord with divergent growth dynamics. *Development*, 145(21), p.dev166728.

Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M.B., Deng, C., Taketo, M., Lewandoski, M. and Pourquié, O., 2008. A β-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nature Cell Biology*, 10(2), pp.186–193.

Auman, T., Vreede, B.M.I., Weiss, A., Hester, S.D., Williams, T.A., Nagy, L.M. and Chipman, A.D., 2017. Dynamics of growth zone patterning in the milkweed bug Oncopeltus fasciatus. *Development*, 144(10), pp.1896–1905.

Ballard, W.W., Mellinger, J. and Lechenault, H., 1993. A series of normal stages for development of Scyliorhinus canicula, the lesser spotted dogfish (Chondrichthyes: Scyliorhinidae). *Journal of Experimental Zoology*, 267(3), pp.318–336.

Bar-On, Y.M., Phillips, R. and Milo, R., 2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences of the United States of America*, 115(25), pp.6506–6511.

Bénazéraf, B., Beaupeux, M., Tchernookov, M., Wallingford, A., Salisbury, T., Shirtz, A., Shirtz, A., Huss, D., Pourquié, O., François, P. and Lansford, R., 2017. Multi-scale quantification of tissue behavior during amniote embryo axis elongation. *Development*, 144(23), pp.4462–4472.

Bénazéraf, B., Francois, P., Baker, R.E., Denans, N., Little, C.D. and Pourquié, O., 2010. A random cell motility gradient downstream of FGF controls elongation of an amniote embryo. *Nature*, 466, pp.248–252.

Berenguer, M., Lancman, J.J., Cunningham, T.J., Dong, P.D.S. and Duester, G., 2018. Mouse but not zebrafish requires retinoic acid for control of neuromesodermal progenitors and body axis extension. *Developmental Biology*, 441(1), pp.127–131.

Berggren, K., Mccaffery, P., Drä, U. and Forehand, C.J., 1999. Differential Distribution of Retinoic Acid Synthesis in the Chicken Embryo as Determined by Immunolocalization of the Retinoic Acid Synthetic Enzyme, RALDH-2. *Developmental Biology*, 210(2), pp.288–304.

Bouldin, C.M., Snelson, C.D., Farr, G.H. and Kimelman, D., 2014. Restricted expression of cdc25a in the tailbud is essential for formation of the zebrafish posterior body. *Genes and Development*, 28(4), pp.384–395.

Brent, A.E. and Tabin, C.J., 2002. Developmental regulation of somite derivatives: Muscle, cartilage and tendon. *Current Opinion in Genetics and Development*, 12(5), pp.548–557.

Cambray, N. and Wilson, V., 2002. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development*, 129, pp.4855–4866.

Cambray, N. and Wilson, V., 2007. Two distinct sources for a population of maturing axial progenitors. *Development*, 134(15), pp.2829–2840.

Casari, A., Schiavone, M., Facchinello, N., Vettori, A., Meyer, D., Tiso, N., Moro, E. and Argenton, F., 2014. A Smad3 transgenic reporter reveals TGF-beta control of zebrafish spinal cord development. *Developmental Biology*, 396(1), pp.81–93.

Choi, H.M.T., Schwarzkopf, M., Fornace, M.E., Acharya, A., Artavanis, G., Stegmaier, J., Cunha, A. and Pierce, N.A., 2018. Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development*, 145(12), p.dev165753.

Cooke, J., 1975. Control of somite number during morphogenesis of a vertebrate, Xenopus laevis. *Nature*, 254, pp.196–199.

Cooke, J. and Zeeman, E.C., 1976. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *Journal of Theoretical Biology*, 58(2), pp.455–476.

Criswell, K.E. and Gillis, J.A., 2020. Resegmentation is an ancestral feature of the gnathostome vertebral skeleton. *eLife*, 9.

Dale, J.K., Malapert, P., Chal, J., Vilhais-Neto, G., Maroto, M., Johnson, T., Jayasinghe, S., Trainor, P., Herrmann, B. and Pourquié, O., 2006. Oscillations of the snail genes in the presomitic mesoderm coordinate segmental patterning and morphogenesis in vertebrate somitogenesis. *Developmental Cell*, 10(3), pp.355–366.

Davidson, A.J., Ernst, P., Wang, Y., Dekens, M.P.S., Kingsley, P.D., Palis, J., Korsmeyer, S.J., Daley, G.Q. and Zon, L.I., 2003. cdx4 mutants fail to specify blood progenitors and can be rescued by multiple hox genes. *Nature*, 425, pp.300–306.

Davidson, A.J. and Zon, L.I., 2006. The caudal-related homeobox genes cdx1a and cdx4 act redundantly to regulate hox gene expression and the formation of putative hematopoietic stem cells during zebrafish embryogenesis. *Developmental Biology*, 292(2), pp.506–518.

Davis, G.K. and Patel, N.H., 1999. The origin and evolution of segmentation. *Trends in Genetics*, 15(12), pp.68–72.

Davis, R.L. and Kirschner, M.W., 2000. The fate of cells in Xenopus tailbud. *Development*, 127, pp.255–267.

Dequéant, M.L., Glynn, E., Gaudenz, K., Wahl, M., Chen, J., Mushegian, A. and Pourquié, O., 2006. A complex oscillating network of signaling genes underlies the mouse segmentation clock. *Science*, 314(5805), pp.1595–1598.

Dequéant, M.L. and Pourquié, O., 2008. Segmental patterning of the vertebrate embryonic axis. *Nature Reviews Genetics*, 9(5), pp.370–382.

Diez del Corral, R., Olivera-Martinez, I., Goriely, A., Gale, E., Maden, M. and Storey, K., 2003. Opposing FGF and Retinoid Pathways Control Ventral Neural Pattern, Neuronal Differentiation, and Segmentation during Body Axis Extension. *Neuron*, 40(1), pp.65–79.

Dray, N., Lawton, A., Nandi, A., Jülich, D., Emonet, T. and Holley, S.A., 2013. Cell-fibronectin interactions propel vertebrate trunk elongation via tissue mechanics. *Current Biology*, 23(14), pp.1335–1341.

Dubrulle, J., McGrew, M.J. and Pourquié, O., 2001. FGF Signaling Controls Somite Boundary Position and Regulates Segmentation Clock Control of Spatiotemporal Hox Gene Activation. *Cell*, 106(2), pp.219–232.

Dubrulle, J. and Pourquié, O., 2004. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature*, 427, pp.419–422.

Forsberg, H., Crozet, F. and Brown, N.A., 1998. Waves of mouse Lunatic fringe expression, in four-hour cycles at two-hour intervals, precede somite boundary formation. *Current Biology*, 8(18), pp.1027–1030.

Fraser, G.J., Hulsey, C.D., Bloomquist, R.F., Uyesugi, K., Manley, N.R. and Streelman, J.T., 2009. An Ancient Gene Network Is Co-opted for Teeth on Old and New Jaws. *PLoS Biology*, 7(2), p.e1000031.

Glasauer, S.M.K. and Neuhauss, S.C.F., 2014. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Molecular Genetics and Genomics*, 289(6), pp.1045–1060.

Gomez, C., Özbudak, E.M., Wunderlich, J., Baumann, D., Lewis, J. and Pourquié, O., 2008. Control of segment number in vertebrate embryos. *Nature*, 454, pp.335–339.

Gomez, C. and Pourquie, O., 2009. Developmental control of segment numbers in vertebrates. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 312(6), pp.533–544.

Griffin, K.J., Amacher, S.L., Kimmel, C.B. and Kimelman, D., 1998. Molecular identification of spadetail: regulation of zebrafish trunk and tail mesoderm formation by T-box genes. *Development*, 125(17).

Habuchi, S., Tsutsui, H., Kochaniak, A.B., Miyawaki, A. and van Oijen, A.M., 2008. mKikGR, a Monomeric Photoswitchable Fluorescent Protein. *PLoS ONE*, 3(12), p.e3944.

Hagos, E.G. and Dougan, S.T., 2007. Time-dependent patterning of the mesoderm and endoderm by Nodal signals in zebrafish. *BMC Developmental Biology*, 7(1), pp.1–18.

Hamburger, V. and Hamilton, H.L., 1951. A series of normal stages in the development of the chick embryo. *Journal of Morphology*, 88(1), pp.49–92.

Henrique, D., Abranches, E., Verrier, L. and Storey, K.G., 2015. Neuromesodermal progenitors and the making of the spinal cord. *Development*, 142(17), pp.2864–2875.

Hirsinger, E. and Steventon, B., 2017. A Versatile Mounting Method for Long Term Imaging of Zebrafish Development. *Journal of Visualized Experiments*, (119), p.e55210.

Holland, L.Z., Albalat, R., Azumi, K., Benito-Gutiérrez, È., Blow, M.J., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L.J., Ferrier, D.E.K., Garcia-Fernàndez, J., Gibson-Brown, J.J., Gissi, C., Godzik, A., Hallböök, F., Hirose, D., Hosomichi, K., Ikuta, T., Inoko, H., Kasahara, M., Kasamatsu, J., Kawashima, T.,

Kimura, A., Kobayashi, M., Kozmik, Z., Kubokawa, K., Laudet, V., Litman, G.W., McHardy, A.C., Meulemans, D., Nonaka, M., Olinski, R.P., Pancer, Z., Pennacchio, L.A., Pestarino, M., Rast, J.P., Rigoutsos, I., Robinson-Rechavi, M., Roch, G., Saiga, H., Sasakura, Y., Satake, M., Satou, Y., Schubert, M., Sherwood, N., Shiina, T., Takatori, N., Tello, J., Vopalensky, P., Wada, S., Xu, A., Ye, Y., Yoshida, K., Yoshizaki, F., Yu, J.K., Zhang, Q., Zmasek, C.M., De Jong, P.J., Osoegawa, K., Putnam, N.H., Rokhsar, D.S., Satoh, N. and Holland, P.W.H., 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Research*, 18(7), pp.1100–1111.

Holley, S.A., Geisler, R. and Nüsslein-Volhard, C., 2000. Control of her1 expression during zebrafish somitogenesis by a Delta- dependent oscillator and an independent wave-front activity. *Genes and Development*, 14(13), pp.1678–1690.

Huth, A. and Wissel, C., 1990. The Movement of Fish Schools: A Simulation Model. In: W. Alt and G. Hoffmann, eds. *Biological Motion*, 1st ed. Berlin: Springer.pp.577–595.

Iimura, T., Yang, X., Weijer, C.J. and Pourquié, O., 2007. Dual mode of paraxial mesoderm formation during chick gastrulation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(8), pp.2744–9.

Ishimatsu, K., Hiscock, T.W., Collins, Z.M., Sari, D.W.K., Lischer, K., Richmond, D.L., Bessho, Y., Matsui, T. and Megason, S.G., 2018. Size-reduced embryos reveal a gradient scaling-based mechanism for zebrafish somite formation. *Development*, 145(11), p.dev161257.

IUCN, 2020. .

Keller, R., Davidson, L., Edlund, A., Elul, T., Ezin, M., Shook, D. and Skoglund, P., 2000. Mechanisms of convergence and extension by cell intercalation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1399), pp.897–922.

Kelly Kuan, C.Y., Tannahill, D., Cook, G.M.W. and Keynes, R.J., 2004. Somite polarity and segmental patterning of the peripheral nervous system. *Mechanisms of Development*, 121(9), pp.1055–1068.

Kimelman, D., 2016. Tales of Tails (and Trunks). Forming the Posterior Body in Vertebrate Embryos. In: P.M. Wassarman, ed. *Current Topics in Developmental Biology*. Academic Press Inc.pp.517–536.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. and Schilling, T.F., 1995. Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3), pp.253–310.

Knezevic, V., De Santo, R. and Mackem, S., 1998. Continuing organizer function during chick tail development. *Development*, 125(10), pp.1791–1801.

Kocher, T.D., 2004. Adaptive evolution and explosive speciation: The cichlid fish model. *Nature Reviews Genetics*, 5(4), pp.288–298.

Koshida, S., Kishimoto, Y., Ustumi, H., Shimizu, T., Furutani-Seiki, M., Kondoh, H. and Takada, S., 2005. Integrinα5-dependent fibronectin accumulation for maintenance of somite boundaries in zebrafish embryos. *Developmental Cell*, 8(4), pp.587–598.

Lawton, A.K., Nandi, A., Stulberg, M.J., Dray, N., Sneddon, M.W., Pontius, W., Emonet, T. and Holley, S.A., 2013. Regulated tissue fluidity steers zebrafish body elongation. *Development*, 140(3), pp.573–582.

Lee, R.T.H., Knapik, E.W., Thiery, J.P. and Carney, T.J., 2013. An exclusively mesodermal origin of fin mesenchyme demonstrates that zebrafish trunk neural crest does not generate ectomesenchyme. *Development*, 140(14), pp.2923–2932.

Li, Y., Fenger, U., Niehrs, C. and Pollet, N., 2003. Cyclic expression of esr9 gene in Xenopus presomitic mesoderm. *Differentiation*, 71(1), pp.83–89.

Linville, A., Gumusaneli, E., Chandraratna, R.A.S. and Schilling, T.F., 2004. Independent roles for retinoic acid in segmentation and neuronal differentiation in the zebrafish hindbrain. *Developmental Biology*, 270(1), pp.186–199.

Liu, P.Z. and Kaufman, T.C., 2005. Short and long germ segmentation: unanswered questions in the evolution of a developmental mode. *Evolution & Development*, 7(6), pp.629–646.

Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F. and Durbin, R., 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nature Ecology & Evolution*, 2, pp.1940–1955.

Manning, A.J. and Kimelman, D., 2015. Tbx16 and Msgn1 are required to establish directional cell migration of zebrafish mesodermal progenitors. *Developmental Biology*, 406(2), pp.172–185.

Mara, A. and Holley, S.A., 2007. Oscillators and the emergence of tissue organization during zebrafish somitogenesis. *Trends in Cell Biology*, 17(12), pp.593–599.

Martin, B.L. and Kimelman, D., 2009. Wnt Signaling and the Evolution of Embryonic Posterior Development. *Current Biology*, 19(5), pp.215–219.

Martin, B.L. and Kimelman, D., 2012. Canonical Wnt Signaling Dynamically Controls Multiple Stem Cell Fate Decisions during Vertebrate Body Formation. *Developmental Cell*, 22(1), pp.223–232.

Mattsson, C., 2018. A morphological and molecular analysis of the species diversity of the cichlid genus Petrochromis from Lake Tanganyika (Teleostei: Cichlidae). *bioRxiv*.

Mayer, G., Kato, C., Quast, B., Chisholm, R.H., Landman, K.A. and Quinn, L.M., 2010. Growth patterns in Onychophora (velvet worms): Lack of a localised posterior proliferation zone. *BMC Evolutionary Biology*, 10(1), p.e339.

McGrew, M.J., Sherman, A., Lillico, S.G., Ellard, F.M., Radcliffe, P.A., Gilhooley, H.J., Mitrophanous, K.A., Cambray, N., Wilson, V. and Sang, H., 2008. Localised axial

progenitor cell populations in the avian tail bud are not committed to a posterior Hox identity. *Development*, 135(13), pp.2289–99.

Mclaren, S.B.P. and Steventon, B., 2021. Anterior expansion and posterior addition to the notochord mechanically coordinate embryo axis elongation. *bioRxiv*.

Mongera, A., Rowghanian, P., Gustafson, H.J., Shelton, E., Kealhofer, D.A., Carn, E.K., Serwane, F., Lucio, A.A., Giammona, J. and Campàs, O., 2018. A fluid-to-solid jamming transition underlies vertebrate body axis elongation. *Nature*, 561, pp.401–405.

Monteiro, R., Pinheiro, P., Joseph, N., Peterkin, T., Koth, J., Repapi, E., Bonkhofer, F., Kirmizitas, A. and Patient, R., 2016. Transforming Growth Factor β Drives Hemogenic Endothelium Programming and the Transition to Hematopoietic Stem Cells. *Developmental Cell*, 38(4), pp.358–370.

Moreno, T.A. and Kintner, C., 2004. Regulation of segmental patterning by retinoic acid signaling during Xenopus somitogenesis. *Developmental Cell*, 6(2), pp.205–218.

Morin-Kensicki, E.M., Melancon, E. and Eisen, J.S., 2002. Segmental relationship between somites and vertebral column in zebrafish. *Development*, 129(16), pp.3851–3860.

Mueller, R.L., Huang, C. and Ho, R.K., 2010. Spatio-temporal regulation of Wnt and retinoic acid signaling by tbx16/spadetail during zebrafish mesoderm differentiation. *BMC Genomics*, 11(1), p.e492.

Near, T.J., Eytan, R.I., Dornburg, A., Kuhn, K.L., Moore, J.A., Davis, M.P., Wainwright, P.C., Friedman, M. and Smith, W.L., 2012. Resolution of ray-finned fish phylogeny and timing of diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 109(34), pp.13698–13703.

Niederreither, K., Abu-Abed, S., Schuhbaur, B., Petkovich, M., Chambon, P. and Dollé, P., 2002. Genetic evidence that oxidative derivatives of retinoic acid are not involved in retinoid signaling during mouse development. *Nature Genetics*, 31(1), pp.84–88.

Oates, A.C., Morelli, L.G. and Ares, S., 2012. Patterning embryos with oscillations: Structure, function and dynamics of the vertebrate segmentation clock. *Development*, 139(4), pp.625–639.

Olivera-Martinez, I., Harada, H., Halley, P.A. and Storey, K.G., 2012. Loss of FGF-Dependent Mesoderm Identity and Rise of Endogenous Retinoid Signalling Determine Cessation of Body Axis Elongation. *PLoS Biology*, 10(10), p.e1001415.

Palmeirim, I., Henrique, D., Ish-Horowicz, D. and Pourquié, O., 1997. Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis. *Cell*, 91(5), pp.639–648.

Riley, B.B., Sweet, E.M., Heck, R., Evans, A., McFarland, K.N., Warga, R.M. and Kane, D.A., 2010. Characterization of *harpy/Rca1/emi1* mutants: Patterning in the absence of cell division. *Developmental Dynamics*, 239(3), pp.828–843.

Roszko, I., Sawada, A. and Solnica-Krezel, L., 2009. Regulation of convergence and extension movements during vertebrate gastrulation by the Wnt/PCP pathway. *Seminars in Cell and Developmental Biology*, 20(8), pp.986–997.

Sachs, S., Kear, B.P. and Everhart, M.J., 2013. Revised Vertebral Count in the "Longest-Necked Vertebrate" Elasmosaurus platyurus Cope 1868, and Clarification of the Cervical-Dorsal Transition in Plesiosauria. *PLoS ONE*, 8(8), p.e70877.

Sawada, A., Shinya, M., Jiang, Y.-J., Kawakami, A., Kuroiwa, A. and Takeda, H., 2001. Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development*, 128(23), pp.4873–4880.

Seehausen, O., 2006. African cichlid fish: A model system in adaptive radiation research. *Proceedings of the Royal Society B: Biological Sciences*, 273(1597), pp.1987–1998.

Shimizu, T., Bae, Y.K., Muraoka, O. and Hibi, M., 2005. Interaction of Wnt and caudal-related genes in zebrafish posterior body formation. *Developmental Biology*, 279(1), pp.125–141.

Sirbu, I.O. and Duester, G., 2006. Retinoic-acid signalling in node ectoderm and posterior neural plate directs left–right patterning of somitic mesoderm. *Nature Cell Biology*, 8(3), pp.271–277.

Slack, J.M.W., Holland, P.W.H. and Graham, C.F., 1993. The zootype and the phylotypic stage. *Nature*, 361(6412), pp.490–492.

Steventon, B., Duarte, F., Lagadec, R., Mazan, S., Nicolas, J.-F. and Hirsinger, E., 2016. Species-specific contribution of volumetric growth and tissue convergence to posterior body elongation in vertebrates. *Development*, 143(10), pp.1732–1741.

Steventon, B. and Martinez Arias, A., 2017. Evo-engineering and the cellular and molecular origins of the vertebrate spinal cord. *Developmental Biology*, 432(1), pp.3–13.

Streelman, J.T. and Albertson, R.C., 2006. Evolution of novelty in the cichlid dentition. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 306B(3), pp.216–226.

Susaki, E.A., Tainaka, K., Perrin, D., Yukinaga, H., Kuno, A. and Ueda, H.R., 2015. Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nature Protocols*, 10(11), pp.1709–1727.

Tada, M. and Smith, J.C., 2000. Xwnt11 is a target of Xenopus Brachyury: regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. *Development*, 127(10).

Vermot, J., 2005. Retinoic Acid Controls the Bilateral Symmetry of Somite Formation in the Mouse Embryo. *Science*, 308(5721), pp.563–566.

Vermot, J. and Pourquié, O., 2005. Retinoic acid coordinates somitogenesis and left–right patterning in vertebrate embryos. *Nature*, 435, pp.215–220.

Wagner, G.P. and Altenberg, L., 1996. Perspective: Complex Adaptations and the Evolution of Evolvability. *Evolution*, 50(3), pp.967–976.

Wellik, D.M., 2007. Hox patterning of the vertebrate axial skeleton. *Developmental Dynamics*, 236(9), pp.2454–2463.

Wells, S., Nornes, S. and Lardelli, M., 2011. Transgenic Zebrafish Recapitulating tbx16 Gene Early Developmental Expression. *PLoS ONE*, 6(6), p.e21559.

Wilson, J., 2005. Overview of Sauropod Phylogeny and Evolution. In: K. Rogers and J. Wilson, eds. *The Sauropods: Evolution and Paleobiology*. Indiana University Press.pp.15–49.

Xiong, F., Ma, W., Bénazéraf, B., Mahadevan, L. and Pourquié, O., 2020. Mechanical Coupling Coordinates the Co-elongation of Axial and Paraxial Tissues in Avian Embryos. *Developmental Cell*, 55(3), pp.354–366.

Young, T., Rowland, J.E., van de Ven, C., Bialecka, M., Novoa, A., Carapuco, M., van Nes, J., de Graaff, W., Duluc, I., Freund, J.N., Beck, F., Mallo, M. and Deschamps, J., 2009. Cdx and Hox Genes Differentially Regulate Posterior Axial Growth in Mammalian Embryos. *Developmental Cell*, 17(4), pp.516–526.

Zhang, L., Kendrick, C., Jülich, D. and Holley, S.A., 2008. Cell cycle progression is required for zebrafish somite morphogenesis but not segmentation clock function. *Development*, 135(12), pp.2065–2070.

Zhang, W., Ducos, B., Delagrange, M., Vriz, S. and Bensimon, D., 2018. Quantitative study of the somitogenetic wavefront in zebrafish. *bioRxiv*.

Zhang, Z.Q., 2013. Animal biodiversity: An update of classification and diversity in 2013. *Zootaxa*, 3703(1), pp.5–11.