Arthropod segmentation

Erik Clark\textsuperscript{1,2,*}, Andrew D. Peel\textsuperscript{3}, and Michael Akam\textsuperscript{2}

\textsuperscript{1} Department of Systems Biology, Harvard Medical School
\textsuperscript{2} Department of Zoology, University of Cambridge
\textsuperscript{3} School of Biology, Faculty of Biological Sciences, University of Leeds

* correspondence: erik_clark@hms.harvard.edu

Summary

There is now compelling evidence that many arthropods pattern their segments using a clock-and-wavefront mechanism, analogous to that operating during vertebrate somitogenesis. In this review, we discuss how the arthropod segmentation clock generates a repeating sequence of pair-rule gene expression, and how this is converted into a segment-polarity pattern by “timing factor” wavefronts associated with axial extension. We argue that the gene regulatory network that patterns segments may be relatively conserved, although the timing of segmentation varies widely, and double-segment periodicity appears to have evolved at least twice. Finally, we describe how the repeated evolution of a simultaneous (\textit{Drosophila}-like) mode of segmentation within holometabolan insects can be explained by heterochronic shifts in timing factor expression plus extensive pre-patterning of the pair-rule genes.

Introduction

Arthropods are an ecdysozoan phylum defined by their segmented bodies and jointed limbs. True arthropods (euarthropods) comprise three living clades: Chelicerata (spiders, scorpions and mites), Myriapoda (centipedes and millipedes), and Pancrustacea (crustaceans and insects). The closest relatives of arthropods are onychophorans (velvet worms) and tardigrades (water bears); together these phyla form the segmented superphylum Panarthropoda (Fig. 1A).

The great diversity of arthropod species is testament to the evolutionary potential of a segmented body plan: a modular organisation of fundamentally similar units arrayed serially along the anteroposterior (AP) axis (Hannibal and Patel, 2013). Arthropod segments, and their associated appendages, have diversified remarkably through adaptation to specific functions, such as feeding, locomotion, or reproduction. In addition, segment number can vary enormously, from fewer than twenty in insects and malacostracan crustaceans, to over a hundred in certain centipedes and
millipedes, resulting in a wide spectrum of organismal forms (Brusca et al., 2016). With over a million
named species, arthropods have colonised and exploited almost every environment on Earth, thanks
in no small part to the evolution of segmentation.

Our understanding of how segments are patterned in arthropod embryos has traditionally been
heavily influenced by study of the fruit fly Drosophila melanogaster. Over the past two decades,
research into sequentially-segmenting species has complemented the well-established Drosophila
model, resulting in the discovery of an arthropod “segmentation clock”, and an outline of conserved
and divergent aspects of arthropod segment patterning networks. In the light of these findings,
recent studies have re-examined segmentation in Drosophila, uncovering new subtleties and
interpreting their evolutionary significance.

In the sections that follow, we provide a general overview of arthropod segmentation and review
our current understanding of three key issues: (1) the nature of the arthropod segmentation clock;
(2) how the “pair-rule” genes pattern segments; and (3) the evolution of Drosophila-style
simultaneous segmentation from a sequentially-segmenting ancestral state. We also reflect on the
origins of arthropod segmentation (Box 1) and the control of segment number (Box 2). As we have
chosen to focus on the time window when segments are actively being patterned, we do not discuss
earlier AP patterning processes, such as axis specification, or later ones, such as segment
morphogenesis.

Overview of arthropod segmentation

Segments and parasegments

In arthropods, morphological segmentation is built upon a more fundamental developmental unit,
the “parasegment” (Martinez-Arias and Lawrence, 1985). Parasegment boundaries are established
during embryogenesis by “segment-polarity” genes such as engrailed and wingless, which are
expressed in a series of persistent stripes along the AP axis. Interestingly, parasegments are offset
slightly from morphological segments: parasegment boundaries fall at the anterior edge of each
engrailed domain and line up with the middle of each appendage, while segment boundaries fall at
the posterior edge of each engrailed domain and lie in between the appendages (Fig. 1B). Analogous
to vertebrate “resegmentation” (each vertebra being formed from portions of two different somite
pairs), this developmental phase shift makes sense if the role of the parasegments is chiefly to
organise the nervous system and associated appendicular structures, while the role of morphological
segmentation is to protect these centres and form exoskeletal articulations between them (Deutsch, 2004).

Each segment-polarity gene is expressed at a particular position within a segmental unit, and the overall arrangement is remarkably conserved across Panarthropoda (Damen, 2002; Janssen and Budd, 2013). A central goal of segmentation research is to understand how upstream regulatory processes establish this important pattern within the embryo.

**Sequential segmentation and the segment addition zone**

Most arthropods pattern their segments sequentially, from head to tail, coupling the segmentation process to progressive axial extension (Sander, 1976). They usually specify some number of anterior segments in the blastoderm, but the majority of the segments emerge rhythmically from a posterior “segment addition zone” (SAZ) after the blastoderm to germband transition. The SAZ retracts posteriorly as new segments are added to the trunk, generally shrinking in size, until the embryo reaches full germband extension (Fig. 1C).

“SAZ” is now preferred over the traditional term “growth zone”, because it makes no assumption of localised and continuous cell proliferation in the posterior of the embryo (Janssen et al., 2010). The material for new segments is generally provided by a combination of cell division and convergent extension, but – as in vertebrates – the relative contributions of these cell behaviours to axial elongation vary widely across species (Auman et al., 2017; Benton, 2018; Benton et al., 2016; Mito et al., 2011; Nakamoto et al., 2015; Steventon et al., 2016). Accordingly, while cell division may in some species be coordinated with segment addition, segment patterning processes do not appear to be mechanistically dependent on the cell cycle (Cepeda et al., 2017), aside from in special cases such as malacostracan crustaceans. This group exhibits a highly derived mode of segmentation in which patterning occurs through regimented asymmetrical divisions of rows of posterior cells (Scholtz, 1992).

While the shape, size, and proportions of the SAZ vary considerably across species, certain features are conserved. Segment-polarity stripes emerge at the anterior of the SAZ, and Wnt is expressed at its posterior (Williams and Nagy, 2017). Between these limits, we define the “anterior SAZ” as the portion of the SAZ that contains segments in the process of being patterned, and the “posterior SAZ” as the portion that contains cells not yet assigned to any particular prospective segment. These functionally-defined regions correlate with the differential expression of key developmental transcription factors; for example, Caudal (the arthropod homolog of the vertebrate Cdx proteins)
appears to be specifically associated with the posterior SAZ (Auman et al., 2017; Clark and Peel, 2018).

Importantly, SAZ identity is transient and dynamic for any given cell. With the generation of each new segment, newly-patterned tissue “leaves” the anterior SAZ, which is simultaneously “replenished” by cells from the posterior SAZ. (Whether cells flow anteriorly out of the SAZ or the SAZ retracts posteriorly along the embryo depends on one’s choice of reference frame.) Thus, a cell which starts out within the posterior SAZ, expressing one set of genes, will at some point end up within the anterior SAZ, expressing a different set of genes, and finally within the segmented germband, expressing yet another (Fig. 1C). This provides a mechanistic explanation for the tight coupling between axial elongation and the segmentation process, because the changing expression levels of SAZ-associated factors such as Caudal are likely to trigger coordinated expression changes in segment patterning genes as the SAZ retracts (Clark and Peel, 2018; El-Sherif et al., 2014).

**Segment patterning by a clock-and-wavefront mechanism**

Arthropod segmentation is frequently compared to vertebrate somitogenesis (reviewed in Hubaud and Pourquié, 2014; Oates et al., 2012). While segments and somites are not homologous morphological structures, it is now becoming clear that both arthropods and vertebrates have converged on a “clock-and-wavefront” strategy (Cooke and Zeeman, 1976) to pattern their AP axis. Temporal periodicity is generated by an oscillator (the “clock”), and progressively translated into spatial periodicity by a second signal (the “wavefront”), which travels along an axis and freezes (or reads out) the phase of the clock.

In vertebrates, the clock consists of cycles of gene expression in the presomitic mesoderm (PSM), while in arthropods it consists of cycles of gene expression in the posterior ectoderm. In both the vertebrate anterior PSM and the arthropod anterior SAZ, the oscillations are slowed by the retraction of posterior signals associated with axial extension, converting them into a series of stripes. These stripes then pattern other genes, which determine the AP polarity of somites (in vertebrates) or segments (in arthropods).

Curiously, the periodicity of the segmentation clock is not fixed across arthropods. Most groups pattern a single new segment for each cycle of the clock (as do vertebrates), but some species pattern two segments in each cycle, meaning that their clock has a double-segment (or “pair-rule”) periodicity (Chipman et al., 2004; Sarrazin et al., 2012).
Other modes of segmentation

The sequential mode of segmentation is widespread and almost certainly ancestral within arthropods. However, across species, the timing of segmentation can vary dramatically relative to other developmental events.

For example, arthropod embryos differ widely in the number of segments they pattern at the blastoderm stage, versus afterwards during germband extension. In insects, this variation is roughly correlated with a spectrum of “germ types” defined in the pre-molecular era (Davis and Patel, 2002; Krause, 1939), but for simplicity and generality, we have chosen to eschew such terminology in this review. Instead, we will refer to sequential segmentation (usually occurring in a germband, under the control of a segmentation clock) versus simultaneous segmentation (usually occurring in a blastoderm, downstream of non-periodic spatial cues). The mechanisms underlying simultaneous segmentation are discussed in more detail below.

Outside of the insects, many arthropod groups undergo post-embryonic segmentation, i.e. delay the development of a portion of the AP axis until after hatching. In crustaceans with naupliar larvae, for example, only the head segments are patterned in the embryo, and trunk segments develop sequentially from a SAZ-like region after the larva has begun feeding (Anderson, 1973). Other, less extreme, examples are found within myriapods: these pattern the head and the first trunk segments in the embryo, but may add one or more trunk segments after each moult (Blower, 1985).

Our focus here is on the segmentation of the trunk (i.e. the axial patterning of the gnathal, thoracic, and abdominal segments), but note that there are other parts of the arthropod body that are segmented by different mechanisms, such as the anterior head (Posnien et al., 2010) or the jointed appendages (Angelini and Kaufman, 2005a). Within the trunk itself, the mechanisms we describe specifically control ectodermal segmentation; mesodermal segmentation occurs later, apparently directed by inductive signals from the segmented ectoderm (Azpiazu et al., 1996; Green and Akam, 2013; Hannibal et al., 2012). Finally, there is evidence that dorsal segmentation in millipedes is decoupled from ventral segmentation, which later leads to segment fusions (Janssen, 2011; Janssen et al., 2004).

Segment patterning genes

Most of the arthropod segmentation genes we know about were originally identified from a genetic screen in Drosophila (Nüsslein-Volhard and Wieschaus, 1980). Drosophila represents an extreme example of simultaneous segmentation, patterning all but its most terminal segments in the
blastoderm. It has taught us a lot about how segmentation genes regulate one another’s expression (Akam, 1987; Nasiadka et al., 2002), but studies in other arthropods were (and are) necessary to reveal how these networks relate to more ancestral modes of segmentation (Peel et al., 2005).

In Drosophila, as in other arthropods, the segment-polarity genes are patterned by the “pair-rule” genes, which code for various transcription factors. In Drosophila, the pair-rule genes are expressed in stripes in the blastoderm, but in sequentially-segmenting species they are also expressed in the SAZ (Patel et al., 1994). In general, the pair-rule genes that turn on earliest in Drosophila (“primary” pair-rule genes) are expressed in the posterior SAZ in sequentially-segmenting species, and may oscillate, while those that turn on later (“secondary” pair-rule genes) are expressed in the anterior SAZ. The periodicity of pair-rule gene expression can be segmental or double-segmental depending on the species (in Drosophila it is double-segmental, hence the term “pair-rule”), but the genes are always referred to as the “pair-rule genes” regardless. There has been some confusion over the years as to which Drosophila pair-rule genes should be classed as primary and which as secondary or even tertiary. However, the most recent analysis (Schroeder et al., 2011), which classifies only paired (prd) and sloppy-paired (slp) as secondary, and all of hairy, even-skipped (eve), runt, odd-skipped (odd) and fushi tarazu (ftz) as primary, meshes well with the comparative evidence.

In Drosophila, the primary pair-rule genes are patterned by the “gap” genes, which code for another set of transcription factors. In Drosophila, these genes are expressed in broad, partially-overlapping domains along the length of the blastoderm, but in sequentially-segmenting species some portion of this pattern is generated over time, in the SAZ (Box 2). Gap genes in sequentially-segmenting species do not seem to be important for directing pair-rule gene expression. They do, however, appear to play a relatively conserved role in patterning the Hox genes, which regulate segment identity (Hughes and Kaufman, 2002a; Marques-Souza et al., 2008; Martin et al., 2016).

| BOX 1: The evolutionary origins of arthropod segmentation |

The major segmented phyla – arthropods, annelids, and chordates – are evolutionarily distant and separated by many unsegmented groups. While losses of segmentation are possible in evolution (e.g. from spoon worms and peanut worms within annelids), we are sceptical about the existence of a segmented urbilaterian ancestor that could have given rise to all three phyla (Couso, 2009). Instead, segmentation appears to have evolved repeatedly during animal evolution, involving various developmental mechanisms (Graham et al 2014).
Some of the developmental commonalities between different segmented phyla may reflect bilaterian homologies that predate segmentation itself, such as elongation of the body from a posterior zone (Jacobs et al 2005, Martin and Kimelman 2009). Other similarities may reflect the convergent adoption of generic patterning strategies, such as molecular oscillators (Richmond and Oates 2012). Finally, certain similarities may reflect the parallel redeployment of ancient molecular mechanisms (Chipman, 2010), and therefore require both homology and convergence to fully explain. For example, segment boundary formation in some, but not all, annelids shows striking similarities to parasegment boundary formation in arthropods (Dray et al., 2010; Prud’homme et al., 2003; Seaver et al., 2001; Seaver and Kaneshige, 2006). Probably, this boundary specification mechanism evolved before trunk segmentation, possibly in the context of patterning the head and anterior nervous system (Vellutini and Hejnol, 2016).

The evolutionary success of segmented phyla emphasizes the adaptive value of diversified metameric structures, but it does not explain why segmentation evolved in the first place. One long-standing hypothesis stresses the advantages of a segmented body for generating coordinated waves of muscular activity to drive locomotion (Clark, 1964). Given that most of the earliest arising segmented lineages have many similar segments, this seems a likely explanation for the initial origins of serial repetition along the body axis, which was likely the forerunner for metameric segmentation. Under this scenario, repetition would be expected first in the nervous system and body wall musculature. Interestingly, onychophorans have distinct mesodermal somites, and show clear parasegmental boundaries in the limbs and nervous system (Eriksson et al., 2009), but show no obvious segmentation of the body wall ectoderm.

**BOX 2: Regulation of segment number**

In arthropods, segment number is determined by the total number of pair-rule stripes (and the periodicity with which they regulate segment-polarity genes). In simultaneously-segmenting insects such as *Drosophila*, individual pair-rule stripes are positioned by gap factors at specific locations along the AP axis, hardcoding segment number. In sequentially-segmenting species, segment number instead depends on the temporal duration of segmentation, divided by the period of the segmentation clock.

Gap genes appear to play some role in controlling the duration of segment addition (Cerny et al., 2005; Nakao, 2016). Over time, gap genes are expressed sequentially within the SAZ, their turnover driven by cross-regulatory interactions (Boos et al., 2018; Verd et al., 2018). This process, effectively
a developmental “timer”, shows intriguing similarities to the “neuroblast clock” (Isshiki et al., 2001; Peel et al., 2005). It evidently exerts some control over the body plan, since perturbing hunchback expression can both decrease (Liu and Kaufman, 2004a; Marques-Souza et al., 2008; Mito et al., 2005) and increase (Boos et al., 2018; Nakao, 2016) segment number in sequentially-segmenting insects. These phenotypes are not well understood, but might result from gap genes directly or indirectly regulating cell behaviour within the SAZ. Such effects are unlikely to be mediated via the Hox genes, since significant perturbations of Hox gene expression in insects and crustaceans have not been found to affect segment number (Angelini et al., 2005; Martin et al., 2016; Stuart et al., 1991).

Despite varying widely among arthropods, segment number is usually fixed within a species. However, there are certain groups, such as geophilomorph centipedes, where naturally occurring variation might provide clues as to how this number evolves (Kettle and Arthur, 2000; Vedel et al., 2008; Vedel et al., 2010). Another interesting question is how species which undergo post-embryonic segmentation coordinate segment patterning with the moult cycle. Ecdysone-related genes play segmentation roles in some embryos (Erezyilmaz et al., 2009; Heffer et al., 2013), suggesting these two processes might be deeply related.

Nature of the arthropod segmentation clock

Oscillating gene expression in the SAZ

Some segmentation genes exhibit extremely variable expression patterns in the posterior SAZs of fixed embryos, suggesting that they continually turn on and off over time. In the beetle Tribolium, split-embryo experiments have confirmed that this variability results from a temporally dynamic “segmentation clock” within individuals rather than spatially variable expression between individuals (Sarrazin et al., 2012). Expression dynamicity has also been demonstrated in Tribolium by comparing the average patterns of finely-staged cohorts of embryos, visualising discrepancies between the transcript and protein domains of a given gene, and gaining an understanding of cell dynamics within the SAZ via live imaging (Benton, 2018; El-Sherif et al., 2012; Sarrazin et al., 2012). In other species, gene expression dynamics within the SAZ have rarely been studied in detail. However, convincing “pseudo time-series” assembled from carefully-staged Strigamia (centipede) and Parasteatoda (spider) embryos imply that oscillatory dynamics are widespread (Brena and Akam, 2013; Schönauer et al., 2016).
Candidate gene approaches in species including Tribolium, Strigamia, the millipede Glomeris, and a second spider, Cupiennius, indicate that oscillating SAZ genes include the primary pair-rule genes hairy, eve, runt and odd (Choe et al., 2006; Damen et al., 2005; Green and Akam, 2013; Janssen et al., 2011). (The segmentation role of ftz is less widely conserved (Pick, 2016).) In addition, Notch signalling components appear to oscillate in many clades (see below), as do prd and hedgehog in spiders (Davis et al., 2005; Schoppmeier and Damen, 2005a; Schwager, 2008). However, since there has not yet been an exhaustive screen for cyclic expression, we don’t know how many other genes may have been missed.

Measurements from Tribolium (El-Sherif et al., 2012; Nakamoto et al., 2015; Sarrazin et al., 2012) and Strigamia (Brena and Akam, 2012) suggest an oscillation period in these species of ~3 hours at 18-20°C (or equivalently ~6 hrs at 13 °C or ~1.5 hours at 30°C, as segmentation speed scales with developmental rate). Adjusted for temperature, these numbers are comparable to the fastest segmenting vertebrates, such as zebrafish or snakes (Gomez et al., 2008). Interestingly, the rate of segment addition is not constant throughout development (Brena and Akam, 2013; Nakamoto et al., 2015). This implies that there is stage-specific variation in the oscillation period, the axial elongation rate, and/or the dynamics of tissue maturation in the SAZ (Schröter et al., 2012; Soroldoni et al., 2014).

At present, the mechanistic basis for the oscillations is not well understood. Nonetheless, it is useful to think about contributing regulatory processes using a three-tier framework (Oates et al., 2012): (1) gene expression dynamics within cells; (2) signalling interactions between cells; and (3) the changing regulatory context along the SAZ.

**Gene expression dynamics within cells**

In vertebrates such as zebrafish, (auto)repressive interactions between Her/Hes transcription factors (homologs of the Drosophila pair-rule gene hairy) are thought to form the core of the segmentation clock, driving oscillations by time-delayed negative feedback (Lewis, 2003; Schröter et al., 2012). Analogously, it is possible that the arthropod segmentation clock is driven by an intracellular negative feedback loop formed by some or all of the oscillating pair-rule genes.

The main evidence for this is that knocking down primary pair-rule genes can block segmentation and truncate the body axis, as has been found in Tribolium (Choe et al., 2006), the silkmoth Bombyx (Nakao, 2015), a second beetle species Derestes (Xiang et al., 2017), and the hemipteran bug Oncopeltus (Auman and Chipman, 2018; Liu and Kaufman, 2005). It can also cause the expression of
other primary pair-rule genes to become aperiodic (Choe et al., 2006; Nakao, 2015), suggesting that at least some of the oscillations are mutually interdependent. This observation distinguishes these knockdowns from those of downstream patterning genes, which may also yield asegmental phenotypes but do not perturb expression dynamics in the SAZ (Choe and Brown, 2007; Farzana and Brown, 2008).

The topology for a pair-rule gene segmentation clock is not clear. An early RNAi study in Tribolium found that eve, runt, or odd knockdown resulted in truncation, while hairy knockdown resulted only in head defects (Choe et al., 2006). This led to the hypothesis that eve, runt, and odd are linked into a three-gene ring circuit, and that even though hairy oscillates in the SAZ, it is not required for segmentation. Specifically, it was proposed that Eve activates runt, Runt activates odd, and Odd in turn represses eve, returning the sequence to the beginning (Fig. 2A). However, more recent evidence has raised issues with this proposal.

First, whether hairy is involved in the Tribolium segmentation clock or not remains unclear. A later study found that hairy knockdown gave a pair-rule phenotype for gnathal and thoracic segments (Aranda et al., 2008), and the iBeetle screen (Dönitz et al., 2015) additionally recovered posterior truncations. hairy also has a paralog, deadpan, expressed with similar dynamics in the SAZ (Aranda et al., 2008), and so its role might be masked by functional redundancy. Finally, hairy knockdown was recently found to produce truncations in Dermestes (Xiang et al., 2017), and hairy is also known to regulate segment patterning in the cockroach Periplaneta (Pueyo et al., 2008), the parasitic wasp Nasonia (Rosenberg et al., 2014), and of course Drosophila, indicating that a role in segmentation is widely conserved.

Second, whether eve and odd are part of the primary oscillator is also not certain. eve expression may be necessary for establishing and/or maintaining the SAZ (Cruz et al., 2010; Liu and Kaufman, 2005; Mito et al., 2007; Xiang et al., 2017), and therefore its severe truncation phenotype may be independent of its potential role in the segmentation clock. odd, on the other hand, has been found to cause pair-rule defects rather than truncations in Dermestes (Xiang et al., 2017) and Oncopeltus (Auman and Chipman, 2018), although the interpretation of these phenotypes is complicated by the existence of odd paralogs, such as sob. Notably, neither eve nor odd shows dynamic expression in the posterior SAZ of Oncopeltus (Auman and Chipman, 2018; Liu and Kaufman, 2005), indicating that periodicity is likely to be generated by other genes in this species.

Finally, the specific regulatory interactions proposed for the circuit seem unlikely. In holometabolous insects (and also Strigamia), eve, runt, and odd are expressed sequentially within each pattern repeat (Choe et al., 2006; Clark, 2017; Green and Akam, 2013; Nakao, 2015; Rosenberg et al., 2014).
In both Tribolium and Bombyx, Eve is necessary for runt expression, and Runt is necessary for odd expression (Choe et al., 2006; Nakao, 2015). However, it is probably not the case that Eve directly activates runt and Runt directly activates odd, as was proposed for Tribolium. Instead, genetic evidence from Bombyx and Drosophila (and wild-type expression dynamics from Tribolium) suggest something closer to a “repressilator” scenario (Elowitz and Leibler, 2000), where each gene in the sequence represses the one before it (Fig. 2A).

In summary, while it is likely that cross-regulation plays a considerable role in shaping dynamic pair-rule gene expression, it is not yet clear whether the oscillating genes are linked into a single circuit, whether this circuit is sufficient to generate oscillations, what the topology of this circuit is likely to be, nor indeed the extent to which it may have diverged in different lineages (Krol et al., 2011).

**Signalling interactions between cells**

Regardless of whether the pair-rule gene network is capable of producing intracellular oscillations autonomously, the segmentation clock must also involve intercellular communication to keep oscillations synchronised across the SAZ. Notch signalling, known to synchronise oscillations during vertebrate somitogenesis (Liao and Oates, 2017), is the key candidate for this role. Indeed, Notch signalling components appear to oscillate along with the pair-rule genes in chelicerates (Schoppmeier and Damen, 2005b; Stollewerk et al., 2003), myriapods (Chipman and Akam, 2008; Kadner and Stollewerk, 2004), crustaceans (Eriksson et al., 2013), and some insects (Pueyo et al., 2008), suggesting that arthropod segmentation ancestrally involved Notch.

Experiments in Cupiennius, Periplaneta, and the branchiopod crustacean Daphnia have found that segment boundaries and the expression of segmentation genes become disorganised when Notch signalling is perturbed (Eriksson et al., 2013; Pueyo et al., 2008; Schoppmeier and Damen, 2005b; Stollewerk et al., 2003). Inhibiting Notch signalling also blocks segmentation (but not axial elongation) in anostracan crustaceans (Williams et al., 2012). These findings indicate that Notch may play an explicit role in generating and/or coordinating pair-rule gene oscillations, perhaps via regulation of hairy (Fig. 2B).

However, the pleiotropy of the Notch pathway means that characterising this potential segmentation function may be difficult. During development, Notch signalling also regulates cell proliferation (Go et al., 1998), SAZ establishment (Chesebro et al., 2013; Oda et al., 2007; Schönauer et al., 2016), and fertility (Xu and Gridley, 2012). Accordingly, strong Notch perturbations in
sequentially-segmenting arthropods often result in uninterpretable axial truncations, or simply a failure to lay many eggs (Kux et al., 2013; Mito et al., 2011; Stahi and Chipman, 2016).

Surprisingly, in the insects Gryllus, Oncopeltus, and Tribolium, the Notch ligand Delta is not expressed in the posterior SAZ (Aranda et al., 2008; Auman et al., 2017; Kainz et al., 2011). Either Notch signalling acts through a different ligand in these species, or it does not directly regulate the clock. Delta also seems not to play a segmentation role in the honeybee Apis (a simultaneously-segmenting species), even though it is expressed in stripes at an appropriate time (Wilson et al., 2010).

If a role for Notch signalling in sequential segmentation has indeed been lost in some insect lineages, it is not clear what mechanism(s) might synchronise cells instead. One possibility is the Toll genes, which are thought to influence intercellular affinity and are expressed dynamically in the SAZ across arthropods (Benton et al., 2016; Paré et al., 2014). However, they seem only to affect morphogenetic processes downstream of segment establishment, rather than segment patterning. Another possibility that has been raised is intercellular communication via Tenascin major (Ten-m) (Hunding and Baumgartner, 2017), a transmembrane protein that was erroneously identified as a Drosophila pair-rule factor owing to an opa mutation present on the balancer chromosome of its stock (Zheng et al., 2011). However, mutation/knockdown of Ten-m does not affect segmentation in either Drosophila or Tribolium (Choe et al., 2006; Zheng et al., 2011), and Ten-m is expressed periodically only after segment-polarity stripes have formed (Baumgartner et al., 1994; Jin et al., 2019).

The changing regulatory context along the SAZ

The segmentation clock oscillates in the posterior SAZ and its phase is read out in the anterior SAZ. Therefore, the “wavefront” can be loosely identified with the boundary between these regions, which retracts posteriorly across the embryo over time. The posterior SAZ and the anterior SAZ are apparently defined by the differential expression of specific regulatory factors (“timing factors” in our terminology), which are expressed dynamically over the course of axial elongation, determining where and when segment patterning takes place (Clark and Peel, 2018). Understanding the mechanistic basis for the wavefront therefore entails characterising (1) the identities of these factors, (2) how they regulate segmentation gene expression, and (3) how they themselves are regulated in the embryo.
Many genes are specifically expressed in a subregion of the SAZ (Oberhofer et al., 2014). However, most studies to date have focused on Wnt and caudal, supplemented recently by Dichaete/Sox21b and odd-paired (opa)/zic. The expression patterns of these genes are relatively consistent across species (Fig. 2C). Wnt is expressed in a small zone around the proctodaeum (Janssen et al., 2010). (We note that this population of cells appears to be distinct from the SAZ proper, and may not contribute to segmental tissue). In Tribolium, two of its receptors are expressed ubiquitously in the embryo, and one is expressed in the anterior SAZ and in segmental stripes (Beermann et al., 2011).

caudal is expressed in the posterior SAZ (Copf et al., 2004; Schulz et al., 1998), and Dichaete is expressed in a similar zone to caudal, but does not overlap with posterior Wnt (Clark and Peel, 2018; Janssen et al., 2018; Paese et al., 2018). In contrast, opa is expressed in the anterior SAZ, i.e. anterior to or slightly overlapping caudal and Dichaete, and also in segmental stripes (Clark and Peel, 2018; Green and Akam, 2013; Janssen et al., 2011). Across arthropods, Wnt, caudal and Dichaete are required to establish and maintain the SAZ (Angelini and Kaufman, 2005b; Bolognesi et al., 2008; Chesebro et al., 2013; Copf et al., 2004; McGregor et al., 2008; Miyawaki et al., 2004; Nakao, 2018; Paese et al., 2018; Schönauer et al., 2016; Shinmyo et al., 2005). In Tribolium, opa is required for segmentation, following earlier roles in blastoderm formation and head specification (Clark and Peel, 2018).

Caudal and Dichaete are strong candidates for activating the segmentation clock, since their expression domains roughly correlate with the extent of its oscillations, and they positively regulate pair-rule gene expression in Drosophila. Caudal has also been shown to be necessary for eve and runt expression in Parasteatoda (Schönauer et al., 2016). Opa, on the other hand, may be important for reading out the phase of the clock, since it activates segment polarity genes and regulates late pair-rule gene expression in Drosophila (Clark and Akam, 2016). Given that all three are transcription factors, they might regulate segmentation by activating or repressing specific genes, modulating the regulatory effects of other transcription factors, or switching expression control between different enhancers. However, the severity of their knockdown phenotypes in sequentially-segmenting species means that uncovering the details may require precisely targeted functional perturbations, and probably transgenic reporters.

In sequentially-segmenting species, the relative expression patterns of different timing factors remain consistent across development, suggesting that they regulate each other’s expression. Wnt is thought to act as a posterior organiser (Chesebro et al., 2013; Oberhofer et al., 2014), and we have hypothesised that regulatory interactions between caudal, Dichaete and opa drive their sequential expression over time (Clark and Peel, 2018). In addition, caudal has been found to be activated by Wnt in diverse arthropods (Beermann et al., 2011; Chesebro et al., 2013; McGregor et al., 2008;
Miyawaki et al., 2004), while Opa, as a Zic factor, might physically bind the Wnt effector TCF and modulate its effects on downstream genes (Murgan et al., 2015; Pourerebrahim et al., 2011).

Therefore, while details are currently sketchy, it seems probable that the timing factors are integrated into a regulatory network that ensures the maintenance of the SAZ over time, and also governs its gradual posterior retraction. Given the numerous parallels between posterior development in arthropods and posterior development in other bilaterian phyla, a similar network might have ancestrally coordinated cell differentiation during axial extension, and only later been exploited to regulate segmentation.

In the basic clock-and-wavefront model, the clock stops abruptly when it is hit by the wavefront. However, in both arthropod segmentation and vertebrate somitogenesis, segmentation clock oscillations may resolve into narrowing travelling waves before they stabilise, indicating that the clock winds down relatively gradually. The way in which the oscillation period varies along the SAZ is described phenomenologically by a “frequency profile” (Morelli et al., 2009), and this can vary over developmental time, as well as between species. While the shape of the frequency profile is not predicted to affect segmentation rate or segment size, models suggest that a graded profile might make patterning more robust (El-Sherif et al., 2014; Vroomans et al., 2018).

Wnt signalling perturbations distort the size and proportions of the SAZ (as judged by the expression of caudal), and cause equivalent distortions to the frequency profile (as judged by the expression of eve) (El-Sherif et al., 2014). This indicates that Wnt signalling affects the dynamics of the segmentation clock, and that its effects might be mediated by SAZ timing factors. However, the mechanism for modulating the oscillation period is not clear. One hypothesis proposes that the clock is quantitatively regulated by a morphogen gradient of Caudal (El-Sherif et al., 2014; Zhu et al., 2017), but the effects of specific timing factors are yet to be disentangled and assessed. Currently, it is unknown whether the period of the clock is indeed explicitly determined by the concentrations of particular timing factors (i.e. given control of these levels one could produce sustained oscillations of arbitrary period), or whether the slowing of the segmentation clock is an inherently transient phenomenon inseparable from its temporal transition from an oscillating to a non-oscillating state (Verd et al., 2014).

Segment patterning by the pair-rule network

Reading out the pattern
In the anterior SAZ, each segmentation clock cycle resolves into an anterior-to-posterior array of partially overlapping stripes of pair-rule gene expression. Because the pair-rule genes are expressed in a strict sequence across a clock repeat (e.g. first eve, then runt, then odd), they convey unambiguous phase information to the cells they are expressed in, which provides significant patterning benefits over a single-gene oscillator (Fig. 3A). The internal organisation of a parasegment consists of at minimum three distinct segment-polarity states (Jaynes and Fujioka, 2004; Meinhardt, 1982). Therefore, each pair-rule gene expression repeat must specify at least three output domains in species with single-segment periodicity, and at least six output domains in species with double-segment periodicity (Fig. 3B).

In Drosophila, the relative expression patterns of pair-rule genes and segment-polarity genes have been characterised in a variety of genetic backgrounds, allowing us to infer the regulatory interactions involved in specifying and resolving the segment pattern (reviewed in Clark and Akam, 2016; Jaynes and Fujioka, 2004). Equivalent data is generally lacking from other arthropod species. However, as far as we can tell from what does exist (mainly single or double stains in wild-type embryos) the overall process appears to be fairly conserved, at least in its broad outline (Auman and Chipman, 2018; Damen et al., 2005; Green and Akam, 2013; Xiang et al., 2017).

First, the primary pair-rule genes pattern the secondary pair-rule genes. Across arthropods, prd and slp are expressed in a conserved, partially overlapping arrangement, which aligns with prospective parasegment boundaries (Choe and Brown, 2007; Green and Akam, 2013). In both Drosophila and other arthropods, prd turns on earlier than slp, at a time when upstream pair-rule gene expression is still dynamic. In Drosophila, both genes are patterned by Eve, and we have proposed that the dynamic nature of the Eve stripes (see below) helps differentially position the two domains (Clark, 2017) (Fig. 3C).

Next, the segment polarity-genes are activated. Each segment-polarity gene is activated or repressed by particular pair-rule factors, which combinatorially define where it is expressed within the pattern repeat (Bouchard et al., 2000; Choe and Brown, 2009; DiNardo and O’Farrell, 1987). In species with double-segment periodicity, odd-numbered and even-numbered segment-polarity stripes may be driven by different regulatory logic (Fig. 3D).

At the same time, some of the pair-rule genes also start being expressed in segment-polarity patterns. In pair-rule species, this involves the splitting of existing stripes or the intercalation of new ones. The new patterns are explained by a new network of regulatory interactions between the pair-rule genes (Clark and Akam, 2016). In contrast to the earlier network, which drives dynamic expression, this later one behaves like a multistable switch, “locking in” specific segment-polarity
Interestingly, different primary pair-rule genes undergo frequency doubling in each of *Drosophila*, *Bombyx*, *Tribolium*, and *Nasonia* (Choe et al., 2006; Clark and Akam, 2016; Nakao, 2015; Rosenberg et al., 2014), contrasting with the conserved expression of the segment-polarity and secondary pair-rule genes.

The resulting segmental patterns go on to regulate morphological segmentation. Note that the pair-rule genes are therefore pleiotropic: they are involved in generating the segment pattern, but some additionally play roles in maintaining segment-polarity, and they also regulate the development of other structures, such as the nervous system. In some cases, these functions have become distributed between multiple paralogs, e.g. *prd/gooseberry/pox-neuro* in *Drosophila* (He and Noll, 2013), or the three copies of *eve* in *Strigamia* (Green and Akam, 2013). Across species, there can be considerable variation in both the number of paralogs present in the genome and the degree of subfunctionalization between them, complicating the interpretation of genetic perturbations.

**The evolution of pair-rule patterning**

In several insect species, and also the centipede *Strigamia* (Chipman et al., 2004), segmentation gene expression undergoes a striking transition from double-segment periodicity to single-segment periodicity as the segment pattern is resolved. However, there is no indication of an initial double-segment periodicity during sequential segmentation in the spiders *Cupiennius* (Davis et al., 2005; Schoppmeier and Damen, 2005a) and *Parasteatoda* (Schwager, 2008), the millipede *Glomeris* (Janssen et al., 2011), or the crustacean *Daphnia* (Eriksson et al., 2013) ([Fig. 1A](#)). This suggests that the ancestral arthropod segmentation clock had a single-segment periodicity, and that pair-rule patterning in insects and centipedes originated independently.

Beyond this, it is not clear exactly when or how many times pair-rule patterning evolved in either of the centipede or insect lineages. *eve* is expressed segmentally rather than in pair-rule stripes in a different centipede species, *Lithobius* (Hughes and Kaufman, 2002b), which could indicate that pair-rule patterning evolved relatively recently within the centipede clade, possibly correlating with the origin of longer bodied forms. However, the dynamics of the *Lithobius* segmentation clock will need be investigated to rule out a transient or cryptic double-segment periodicity.

In insects, most of the available data come from holometabolon or orthopteran species, as well as the cockroach *Periplaneta* and hemipteran bug *Oncopeltus* ([Fig. 1A](#)). Holometabolans (Binner and Sander, 1997; Nakao, 2010; Patel et al., 1994; Rosenberg et al., 2014) and orthopterans (Davis et al., 2001; Mito et al., 2007) both show obvious transitions from double-segment to single-segment
periodicity, but the mapping between the pair-rule pattern and the segmental pattern is different in the two groups, suggesting that their respective pair-rule mechanisms might have evolved independently. Consistent with this possibility, gene expression in *Periplaneta* (more closely related to orthopterans than to holometabolanst) appears to be single-segmental (Pueyo et al., 2008), although, as with *Lithobius*, the dynamics of its segmentation clock have not been explicitly investigated. Finally, *Oncopeltus* is a rather strange case: based on the expression and function of *eve*, it appears to lack pair-rule patterning, but pair-rule expression and/or function of certain other genes hints at an underlying double-segment periodicity (Auman and Chipman, 2018; Benton et al., 2016; Erezyilmaz et al., 2009; Liu and Kaufman, 2005).

Thus, while the evidence from some of these species is ambiguous, the current picture suggests that pair-rule patterning may have evolved within crown-group insects, possibly multiple times. This is puzzling, because the specialised and relatively invariant body plan of insects presents a morphological constraint that is hard to reconcile with a saltational doubling of segmentation rate. (Instead, it is much easier to imagine pair-rule patterning evolving in remipedes, which are thought to be the sister group of hexapods (Schwentner et al 2017), and have homonomous, centipede-like bodies.) How was the evolution of double-segment periodicity coordinated with compensatory changes to Hox dynamics and the duration of axial extension, so as to keep segment number (Box 2) and segment identity constant? Given that *Strigamia* seems to switch to a single segment periodicity when adding its most posterior segments (Brena and Akam, 2013), and that pair-rule patterns are seen during the anterior patterning of otherwise segmental species (Dearden et al., 2002; Janssen et al., 2012), one possibility is that pair-rule patterning was introduced gradually along the AP axis, allowing other developmental parameters the chance to adapt.

Since pair-rule patterning requires half the number of clock cycles to generate a given number of segment-polarity stripes, its evolution may have been driven by selection for faster development (in holometabolanst) or a longer body (in centipedes). However, it is currently not obvious how the ancestral segment patterning mechanism was modified to become pair-rule. Segmental frequency could have been doubled by changing the “readout” of a conserved clock, i.e. by evolving new enhancers to drive additional segment-polarity stripes in between the originals, or altering the control logic of existing enhancers to drive a pair of stripes instead of just one. Alternatively, the clock itself could have been modified, e.g. by recruiting new genes into the original cyclic repeat and thereby expanding its patterning potential. To reconstruct the specific regulatory changes that occurred, it will be informative to find out how the gene expression and enhancer logic of pair-rule species compares to their closest segmental relatives.
The evolution of simultaneous segmentation

Reconciling sequential and simultaneous segmentation

A segmentation clock is one strategy for generating periodicity, but another is simply to regulate each stripe individually, exploiting whatever positional information is locally available (François et al., 2007; Salazar-Ciudad et al., 2001; Vroomans et al., 2016). This latter method is used in the *Drosophila* blastoderm, where over 20 “stripe-specific elements” (SSEs) regulate the expression of the five primary pair-rule genes (Schroeder et al., 2011). These elements receive spatial information from gap factors, and each drives expression at a different AP position along the blastoderm, contributing just one or two stripes to a gene’s overall 7-stripe pattern. Sepsid flies (which diverged from drosophilids about 100 million years ago) are also known to use this kind of element (Hare et al., 2008), and it is likely that similarly ad hoc regulatory mechanisms are used wherever periodicity emerges simultaneously, e.g. in the blastoderms of *Nasonia* (Rosenberg et al., 2014) and *Oncopeltus* (Stahi and Chipman, 2016), or in the chelicerate prosoma (Pechmann et al., 2011; Schwager et al., 2009). While less “elegant” than using temporal oscillations, this explicitly spatial mode of segmentation can—in principle—occur much faster, since a number of different pattern repeats can be initialised at once.

Simultaneous segmentation, typified by *Drosophila*, is traditionally thought of as mechanistically distinct from sequential segmentation, typified by e.g. *Tribolium* or *Gryllus*. The textbook model of the hierarchical “subdivision” of a syncitial blastoderm by morphogen gradients seems a world away from waves of gene expression within a cellularised, elongating germband. However, the *Drosophila* blastoderm is now known to be more dynamic than was previously imagined, and the basic structure of its segment patterning network seems remarkably similar to that of other arthropods ([Fig. 4A](#)).

As the *Drosophila* blastoderm stage is so short, the effects of dynamic gene expression are subtle, and for years were overlooked. However, quantitative expression atlases suggest that expression domains in the posterior half of the blastoderm travel anteriorly across cells over time (Jaeger et al., 2004; Keränen et al., 2006; Surkova et al., 2008), and this has recently been demonstrated through live imaging (El-Sherif and Levine, 2016; Lim et al., 2018). The shifts reflect sequential patterns of transcriptional states within cells, and trace back to asymmetric repressive interactions in the gap gene network (Jaeger, 2011; Verd et al., 2018) ([Fig. 4B](#)) – perhaps similar to the ones driving their temporal expression in the SAZs of sequentially-segmenting species.
In the *Drosophila* blastoderm, the expression dynamics of the gap genes are directly transferred to pair-rule genes via their SSEs (Fig. 4B). In addition, the pair-rule genes cross-regulate each other through “zebra elements”: enhancers that drive expression in all of the trunk stripes simultaneously (Schroeder et al., 2011). (Some primary pair-rule genes, and both secondary pair-rule genes, possess zebra elements.) These regulatory interactions are also dynamic, and they combine with the stripe shifts driven by the gap genes to generate a staggered sequence of pair-rule gene expression within each double-segment repeat (Clark, 2017) (Fig. 4B3). This spatiotemporal sequence is the same as that driven by the segmentation clock in sequentially-segmenting species such as *Tribolium* and *Strigamia* (Choe et al., 2006; Green and Akam, 2013), suggesting that zebra enhancers and “clock” enhancers may be homologous.

Once primary pair-rule gene expression is properly phased within each double-segment repeat, *Drosophila* segment patterning proceeds just as it would in the anterior SAZ of a sequentially-segmenting species, beginning with the activation of *prd* and *slp*, and moving on to segment-polarity gene expression and stripe doubling. This conserved process of pattern resolution is apparently regulated by a conserved sequence of timing factor expression: posterior SAZ factors Caudal and Dichaete are expressed throughout the trunk during the early, dynamic stages of pair-rule gene expression in *Drosophila*, and are replaced by anterior SAZ factor Opa as the segment-polarity pattern is being resolved (Clark and Peel, 2018).

The *Drosophila* blastoderm therefore seems effectively equivalent to a SAZ, except that rather than maturing gradually from anterior to posterior, it does so all at once (Fig. 4C). We suspect that much of the ancestral segmentation machinery remains intact. However, since spatial information is no longer conveyed by the delayed maturation of posterior tissue, gap genes and SSEs preload it into the system instead (Fig. 4A). Importantly, while genetic perturbations tend to result in different phenotypes in the two modes of segmentation (e.g. primary pair-rule genes cause pair-rule phenotypes in *Drosophila* rather than truncations), this might often be explained by the divergent deployment of the genes in the embryo, rather than divergent function.

**The evolution of stripe-specific elements**

Simultaneous segmentation differs from sequential segmentation in two key respects: its temporal regulation (determined by the expression profiles of the timing factors), and the spatial pre-patterning of the pair-rule genes by gap genes (Fig. 4C). Simultaneous segmentation is also associated with an anterior shift of the blastoderm fate map and an increase in the number of...
segments patterned prior to gastrulation. (Note, however, that although segment patterning in the blastoderm is often simultaneous and regulated by gap genes, this need not be the case: Tribolium patterns its blastoderm segments sequentially, using retracting timing factors and a clock (El-Sherif et al., 2014, 2012).)

The evolution of simultaneous segmentation appears to be constrained by early embryogenesis (French, 1988). Some insects, such as orthopterans, have “panoistic” ovaries, in which all germline cells become oocytes, and the eggs contain little but yolk (Bünning, 1994). These species pattern their segments sequentially. Other insects, such as hemipterans and holometabolans, have “meroistic” ovaries, in which germline-derived “nurse” cells load oocytes with maternal mRNA. These species frequently have a biphasic mode of segmentation, in which anterior segments are patterned simultaneously. Meroistic ovaries (which facilitate pre-patterning of the egg), may therefore be a pre-adaptation for simultaneous segmentation.

Extreme examples of simultaneous segmentation (e.g. Drosophila) have evolved independently within each of the major holometabolan orders (Davis and Patel, 2002). (Intriguingly, there has also been at least one reversion to sequential segmentation, within braconid wasps (Sucena et al., 2014)). A Drosophila-like mode of segmentation likely requires far-reaching changes to early embryogenesis, such as a novel anterior patterning centre to help spatially pattern gap genes along the entire AP axis of the egg (Lynch et al., 2006) (Fig. 4A). Here, we focus on understanding how SSEs and gap genes are together able to take over stripe patterning from the clock. It seems likely that this transition to intricate spatial regulation involves a series of selectively favourable regulatory changes, which incrementally increase the speed or robustness of segmentation, while strictly preserving its output (Fig. 5).

First, new SSEs seem to be easy to evolve, because they tend to be short, with simple regulatory logic and high sequence turnover between closely related species (Hare et al., 2008; Ludwig et al., 1998). Some of them may have been selected simply to increase the robustness of segmentation clock expression; this might have occurred in either a blastoderm or a SAZ context. (There is one report from Tribolium suggesting the existence of SSEs that drive expression in the germband (Eckert et al., 2004)). Importantly, because gap gene expression is inherently dynamic (whether in the blastoderm or the SAZ), SSE-regulated stripes are predicted to “shadow” stripes driven by the clock, allowing them to take over downstream functions quite gradually (Verd et al., 2018) (Fig. 5A).

Second, only a single new SSE need evolve at one time. Simultaneous patterning seems likely to have evolved progressively, from anterior to posterior, with each new SSE-driven stripe reducing the number of cycles needed from the clock (Peel and Akam, 2003) (Fig. 5B). Furthermore, cross-
regulation between the pair-rule genes means that a SSE for one gene could in principle go on to
organise a whole pattern repeat, with the remaining genes evolving their own SSEs afterwards, to
make patterning faster or more robust (Clark, 2017) (Fig. 5C). This process might be highly
contingent: in Drosophila, eve and runt have full sets of SSEs and odd is patterned largely through
cross-regulation (Schroeder et al., 2011), but RNAi evidence from Bombyx suggests precisely the
opposite (Nakao, 2015).

Finally, SSEs can be reused. In Drosophila there are several SSEs that drive a pair of stripes, typically
arranged symmetrically around a particular gap domain (Schroeder et al., 2011). This suggests that
posterior gap gene expression evolved to duplicate the regulatory environments of anterior stripes,
initialising additional pair-rule gene stripes without the need to evolve additional SSEs (Fig. 5D).

Interestingly, Drosophila eve stripes 3 and 7, which are co-driven by a single SSE, are regulated by
the same gap genes as are eve stripes 3 and 6 in Anopheles (Goltsev et al., 2004), which has led to a
proposal that certain stripes have been lost or gained from these lineages over time (Rothschild et
al., 2016). This hypothesis is hard to reconcile with the gradualist scenario we favour, since the
transitional states would have severely compromised fitness. We think it more likely that the
posterior gap gene domains were recruited in a different order in the Drosophila and Anopheles
lineages, resulting in a homologous “stripe 3” element additionally driving non-homologous
posterior stripes. In support of this alternative, a midge species more closely related to Drosophila
than to Anopheles patterns only five eve stripes before gastrulation (Rohr et al., 1999), indicating
that the two lineages probably evolved fully simultaneous segmentation independently (Jaeger,
2011).

Conclusion

Our current understanding is that arthropod segment patterning is an inherently dynamic and a
significantly conserved process, ancestrally taking the form of a clock-and-wavefront system. Note,
however, that many of the conclusions in this review extrapolate from fragmentary data gathered
from a small number of model species, with functional data available from an even smaller number.
This is certainly not the last word on arthropod segmentation, but we hope to have provided a
coherent framework for further thought and experiment.

We anticipate that future investigation will centre on two contrasting but interrelated tasks. First,
better resolving the nature of the ancestral arthropod clock-and-wavefront system: the topology of
the gene regulatory networks comprising the clock, the production of timing factor wavefronts by a
retracting SAZ, and the mechanistic basis for the interactions between them. Second, reconstructing how arthropod segmentation networks have diversified over time, giving rise to such remarkable novelties as simultaneous patterning and double-segment periodicity. In addition, we believe that sequentially-segmenting arthropod models are well placed to complement and inform the study of vertebrate axial patterning, especially given their benefits of cost-efficiency, short generation times, experimental tractability, and relatively simple genomes.

The most pressing next step is to collect good-quality multiplexed expression data from a variety of arthropod species (Choi et al., 2018, 2016) and cross-reference this with information about tissue dynamics (Wolff et al., 2018), to better characterise how segmentation gene expression changes over space and time. Building on a solid descriptive foundation, there are numerous exciting directions to pursue: genome editing to generate mutants, misexpression constructs, and live reporters (Gilles et al., 2015; Lai et al., 2018); construction and analysis of data-informed dynamical models (Sharpe, 2017); single-cell sequencing of segmenting tissues (Griffiths et al., 2018); ex vivo culturing of SAZ cells (Lauschke et al., 2013). Over the past four decades, arthropod segmentation has contributed enormously to our understanding of developmental gene networks and their evolution. As we enter a new “golden age” of developmental biology, we see great promise for this legacy to continue.

**Acknowledgements**

We thank Olivia Tidswell, Matthew Benton, Lauren Bush, Mariana Wolfner and Roger Keynes for comments on the manuscript.

**Funding**

MEA and EC were supported by BBSRC research grant BB/P009336/1. EC was also supported by a Junior Research Fellowship from Trinity College, Cambridge, and an EMBO Long Term Fellowship. ADP was supported by BBSRC research grant BB/L020092/1.
References


Fig. 1. Overview of arthropod segmentation. (A) Phylogenetic tree of notable arthropod model species (based on Misof et al., 2014; Schwentner et al., 2017). Red text indicates species known to use pair-rule patterning; the status of Oncopeltus is currently unclear. Branch lengths not to scale. (B) Diagram showing the relationship between parasegments and segments. Pink=engrailed expression; ‘A’=anterior; ‘P’=posterior. (C) Schematic time series of an arthropod embryo undergoing sequential segmentation. engrailed stripes (pink) emerge sequentially from a retracting segment addition zone (SAZ, blue) as the germ band extends posteriorly. Green dots mark the progress of a specific individual cell that starts in the posterior SAZ (dark blue), passes through the anterior SAZ (light blue), and ends up in the segmented germ band.

Fig. 2. Within cell, between-cell, and tissue-level aspects of the arthropod segmentation clock. (A) Pair-rule gene oscillations may be driven by a cross-regulatory feedback loop within cells. The two hypothetical topologies shown (left) would be capable of driving similar, although not identical, cycles of eve, runt, and odd expression within cells (right). In Tribolium, the relative expression patterns of Eve protein, runt transcript and odd transcript resemble the predicted expression of model 2, rather than model 1 (see Supporting Information from Choe et al., 2006). Expression predictions assume Boolean regulatory logic and equal time delays for protein synthesis and protein decay (Clark, 2017). (B) Notch signalling might indirectly synchronise intracellular oscillations of eve, runt, and odd across cells, by acting through hairy. This figure shows a hypothetical regulatory network, which synthesises genetic interactions documented from various different arthropod species (Clark, 2017; Eriksson et al., 2013; Nakao, 2015; Pueyo et al., 2008; Stollewerk et al., 2003). The left half of the network (“oscillator 1”) would synchronise oscillations of hairy across neighbouring cells, by coupling hairy expression to Notch signalling. The oscillations of hairy would then influence the phase of the genetic ring oscillator that forms the right hand of the network (“oscillator 2”), by repressing some of its component genes. (C) Genes such as Wnt, caudal, Dichaete, and opa have distinct expression patterns within the SAZ, which correlate with different phases of segment patterning. ’A’=anterior; ‘P’=posterior. (Based on Tribolium data from Clark and Peel, 2018.) Note that Wnt and opa have segment-polarity patterns in the segmented germ band. caudal and/or Dichaete stripes (not shown) are seen in the anterior SAZ of some species, indicating that the clock feeds back on their expression (Chipman et al., 2004; Clark and Peel, 2018).

Fig 3. Resolving the segment pattern: from oscillations to stable stripes. (A) Comparison of patterning using a single-gene oscillator versus patterning using a three-gene oscillator. With a single-gene oscillator, different cell fates are determined by different expression levels of the oscillator. The output is sensitive to noise in the amplitude of, or measuring of, the signal, and must be palindromic, because the input signal is symmetrical. With a three-gene oscillator, different cell fates can be determined by different combinations of input factors. The output is more robust to noise, and has an inherent polarity. (B) Comparison of the segment-polarity fate readout for clocks with single-segment or double-segment periodicity. Parasegment boundaries (red lines) form wherever a cell with an anterior segment-polarity fate (‘A’; i.e. expressing engrailed) abuts a cell with a posterior segment-polarity fate (‘P’; i.e. expressing slp and wg). A third cell fate (light grey; e.g. odd in Drosophila) prevents ectopic boundaries. Note that species with double-segment periodicity have a different, more complex mapping between the input pattern (pair-rule gene expression) and the output pattern (segment-polarity gene expression). (C) Dynamic model for the patterning of prd and slp in Drosophila: the staggered expression boundaries of prd and slp are caused by the Eve stripes shifting anteriorly across the tissue over time. The posterior border of the prd stripe is patterned at timepoint t₁ (Eve expression shown by dotted line), while the posterior border of the slp stripe is patterned a short while later, at timepoint t₂ (Eve expression shown by solid line). (Based on Clark, 2017). (D) The staggered pattern of pair-rule gene expression comprises a positional code, which specifies narrow stripes of segment-polarity gene expression. The regulatory logic (top) and resulting expression pattern (bottom) of Drosophila engrailed (en) is shown as an example. Note that odd-numbered and even-numbered en stripes are regulated differently. (Based on Jaynes and Fujioka, 2004).
**Fig. 4. Reconciling sequential and simultaneous segmentation.** (A) Structural overview of arthropod segmentation gene networks. The core of the system (yellow box) is relatively conserved across species. In sequential segmentation, spatial information is provided by the timing factor network, which generates a wavefront. Gap genes do not play a major role in segment patterning, although late gap gene expression may be important for “shutting down” the SAZ, by repressing timing factors that maintain it (dashed blue arrow). In simultaneous segmentation, timing factors only provide temporal information. Spatial information is usually provided by a novel anterior patterning centre (i.e. a morphogen gradient such as Bicoid (Liu et al., 2018; McGregor, 2005)), which regulates gap gene expression. Gap genes pass this information to the primary pair-rule genes, through newly-evolved regulatory elements (SSEs). (B) Spatial patterning in *Drosophila* is inherently dynamic. (1) Regulatory interactions between gap genes cause gap domains to shift anteriorly across the blastoderm over time. (2) Stripes of pair-rule gene expression regulated by gap inputs also shift anteriorly. (3) Regulatory interactions between the pair-rule genes convert these shifts into a staggered pattern of expression overlaps across the pair-rule repeat. Note that each panel zooms in on a smaller region of the AP axis. (C) Schematic kymographs (i.e., plots of how gene expression along the AP axis changes over time) comparing the key spatiotemporal features of sequential and simultaneous segmentation. In sequential segmentation, timing factor expression (blue) matures from anterior to posterior across the tissue, producing a wavefront (diagonal line). Periodicity is generated by sustained oscillations (note how even-skipped turns on and off over time within the blue zone). The wavefront converts the oscillations into a stable segment-polarity pattern. In simultaneous segmentation, there is little spatial regulation of timing factor expression across the tissue, and pair-rule stripes are present from the start. Embryo diagrams depict the specific timepoints they line up with on the kymographs (eve expression is not shown). Patterning has double-segment periodicity. Note the different scales of the two time axes.

**Fig 5. The evolution of simultaneous segmentation involves a gradual replacement of the segmentation clock by SSEs.** (A) Clock enhancers (potentially homologous to zebra elements) and SSEs both drive stripes that shift anteriorly over time. SSEs can therefore gradually assume regulatory control over particular clock-driven stripes, without disrupting downstream patterning. (B) (1) Simultaneous patterning is likely to evolve stepwise along the AP axis, via the acquisition over evolutionary time of new SSEs that control expression in increasingly posterior stripes. Embryo diagrams assume a segmentation clock with double-segment periodicity. (2) Simultaneous patterning is likely to evolve stepwise within each pair-rule gene expression repeat, as more of the primary pair-rule genes evolve their own SSEs. Additional SSEs reduce the time required to organise pair-rule gene expression across the repeat. (D) Changes in gap gene expression can be sufficient to generate additional SSE-driven stripes, without accompanying changes in cis-regulatory logic. In *Drosophila* (right panel), SSEs such as eve 3+7 and eve 4+6 each drive a pair of stripes. The current situation likely evolved from a simpler scenario (left panel), in which the same enhancers drive expression in only one stripe each. Hb=Hunchback; Kr=Krüppel; Kni=Knirps; Gt=Giant. Note that eve 3+7 and eve 4+6 are both repressed by Kni and Hb, but with different relative strengths, represented by different arrow thicknesses (Samee et al., 2017). Diagrams are colour-coded such that transcription factor names (top) have the same colour as their corresponding expression domain(s) (below).
eve, runt, and odd might form a genetic ring oscillator

1) activation-based model (Choe et al 2006)

\[
\begin{align*}
\text{Eve} & \rightarrow \text{Odd} \\
\text{Odd} & \rightarrow \text{Runt}
\end{align*}
\]

expression within cell over time

2) repression-based model

\[
\begin{align*}
\text{Eve} & \leftarrow \text{Odd} \\
\text{Odd} & \leftarrow \text{Runt}
\end{align*}
\]

expression within cell over time

hypothesis: hairy links intercellular and intracellular oscillations

oscillator 1

Notch

Hairy

oscillator 2

Eve

Runt

Odd

1) Notch signalling synchronises Hairy oscillations across cells

2) cyclic Hairy expression within cells entrains oscillations of other pair-rule genes

3) pair-rule factors directly regulate segment patterning

functional evidence:

- Periplaneta
- Daphnia
- Cupiennius
- Drosophila
- Bombyx

"timing factors" coordinate segment patterning across the SAZ

expression

segmented germband: segment-polarity stripes maintained

anterior SAZ: the segment pattern is resolved

posterior SAZ: oscillations generate periodicity
A  single-gene oscillator:

expression:

interpretation:

fate:

palindromic repeat unit

three-gene oscillator:

expression:

interpretation:

fate:

polarised repeat unit

B  single-segment periodicity:  double-segment periodicity:

input:

output:

1 parasegment 1 parasegment

parasegment boundaries  parasegment boundaries

1 cycle 1 cycle

C  dynamic model for the patterning of prd and slp

Eve stripe shifts anteriorly over time

patterned later patterned earlier

anterior  posterior

D  segment-polarity genes are patterned by a "combinatorial code" of pair-rule gene expression

control logic: odd-numbered stripes even-numbered stripes

inputs:

outputs:

odd-numbered stripe  even-numbered stripe

odd-numbered stripe  even-numbered stripe
A  

**Sequential Segmentation**

- **Primary Pair-Rule Gene Network**
- **Secondary Pair-Rule Genes**
- **Segment-Polarity/Late Pair-Rule Network**

**Timing Factor Network**

- **Posterior**
- **Anterior**

**Fate Specification**

**Spatiotemporal Regulation**

**Duration/Termination**

B  

1) Cross-regulation causes gap expression shifts

2) Gap shifts cause SSE-driven stripes to shift

3) Shifts plus cross-repression organise pair-rule pattern

C  

**Sequential Segmentation**

- Even-skipped
- Engrailed

**Simultaneous Segmentation**

- Even-skipped
- Engrailed

**Time**

**A** (Anterior) **P** (Posterior)
A. Each SSE can take over from the clock gradually

1) Along the AP axis

- Segments patterned by the clock
- Segments patterned by new SSE

2) Within each clock repeat

- Size of shift needed to phase stripes driven by clock/zebra enhancers
- Larger size of shift needed to drive zebra stripes

B. SSEs are able to evolve one at a time:

1) Along the AP axis

- Evolution of new SSE
- A new SSE establishes stripe; clock enhancer then refines it

2) Within each clock repeat

- More stripes established by cross-regulation
- More stripes established by gap genes

C. Existing SSEs can be recruited to drive additional stripes

- Simpler gap gene pattern
- More complex gap gene pattern