Genomes and Developmental Control

Somitogenesis in the anole lizard and alligator reveals evolutionary convergence and divergence in the amniote segmentation clock

Walter L. Eckalbar a, Eris Lasku a, Carlos R. Infante b, Ruth M. Elsey c, Glenn J. Markov a, April N. Allen d, Jason J. Corneveaux e, Jonathan B. Losos b, Dale F. DeNardo a, Matthew J. Huentelman d, Jeanne Wilson-Rawls a, Alan Rawls a, Kenro Kusumi ab

a School of Life Sciences, Arizona State University, PO Box 874501, Tempe, AZ 85287, USA
b Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
c Louisiana Department of Wildlife and Fisheries, Rockefeller Wildlife Refuge, Grand Chenier, LA 70643, USA
d Neurogenomics Division, Translational Genomics Research Institute, 445 N. 5th St., Phoenix, AZ 85004, USA

ARTICLE INFO

Article history:
Received for publication 22 September 2011
Revised 22 November 2011
Accepted 29 November 2011
Available online 8 December 2011

ABSTRACT

The axial skeleton is a defining feature of vertebrates and is patterned during somitogenesis. Cyclically expressed members of the notch and other signaling pathways, described as the ‘segmentation clock’, regulate the formation of somite boundaries. Comparisons among vertebrate model systems have revealed fundamental shifts in the regulation of expression among critical genes in the notch pathway. However, insights into the evolution of these expression differences have been limited by the lack of information from non-avian reptiles. We analyzed the segmentation clock of the first Lepidosaurian reptile sequenced, the green anole lizard, Anolis carolinensis, for comparison with avian and mammalian models. Using genomic sequence, RNA-Seq transcriptomic data, and in situ hybridization analysis of somite-stage embryos, we carried out comparative analyses of key genes and found that the anole segmentation clock displays features common to both amniote and amniote vertebrates. Shared features with amniotes, represented by Xenopus laevis and Danio rerio, include an absence of lunatic fringe (lfng) expression within the presomitic mesoderm (PSM), a hes6a gradient in the PSM not observed in the chicken or mouse, and EGF repeat structure of the divergent notch ligand, dll3. The anole and mouse share cyclic expression of dll1 ligand in the PSM. To gain insight from an Archosaurian reptile, we analysed lfng and dll1 expressions in the American alligator. Lfng expression was absent in the alligator PSM, like the anole but unlike the chicken. In contrast, dll1 expression does not cycle in the PSM of the alligator, similar to the chicken but unlike the anole. Thus, our analysis yields novel insights into features of the segmentation clock that are evolutionarily basal to amniotes versus those that are specific to mammals, Lepidosaurian reptiles, or Archosaurian reptiles.

© 2011 Elsevier Inc. All rights reserved.

Introduction

One of the defining features of vertebrates is the segmented spine, which is patterned during the developmental process of somitogenesis (Dequéant and Pourquié, 2008; Gibb et al., 2010). Molecular analysis of somitogenesis has primarily focused on analysis of four vertebrate model systems: the zebrafish, Danio rerio (a teleost); the African clawed frog, Xenopus laevis (an anuran amphibian); the chicken, Gallus gallus (an avian reptile); and the mouse, Mus musculus (a eutherian mammal). From studies in these species, a model of genetic regulation of somitogenesis, described as the ‘segmentation clock’ has been developed, based on the theoretical framework of the ‘clock and wavefront’ model (Cooke and Zeeman, 1976). In the contemporary revision of this model, information from genes expressed in a gradient along the rostral-caudal axis of the unsegmented presomitic mesoderm (PSM) is integrated with expression of genes that are cyclically transcribed to shift cells between permissive or nonpermissive states for segment formation at the determination front. A new segmentation boundary is determined based on the periodic interaction of the cycling genes and these gradients. Comparative analysis has revealed conserved features of the segmentation clock in these four model systems, including: gradients of FGF8 and WNT3a proteins (Aulehla et al., 2003; Beck and Slack, 1998; Chapman et al., 2011; Crossley and Martin, 1995; Draper et al., 2003; Dubrulle et al., 2001; Pera et al., 2002; Reifers et al., 1998; Takada et al., 1994; Thorpe et al., 2005), cyclical expression of genes in the notch signaling pathway such as the hairy/enhancer of split (hes and her) genes (reviewed in Dequéant and Pourquié, 2008; Holley, 2007; Sparrow, 2008), and expression of mesp2 orthologues at the determination front (Buchberger et al., 1998; Hitachi et al., 2009; Saga et al., 1997; Sawada et al., 2000).
Comparative studies have also revealed divergence in the segmentation clock among the vertebrate models. Not all PSM gradient genes are conserved. As an example, the notch effector hes orthologues are present in a gradient in the amniotic models, X. laevis, and zebrafish, but not in the amniote models, chicken or mouse (Cossins et al., 2002; Fior and Henrique, 2005; Pissarra et al., 2000; Sieger et al., 2006). Furthermore, the orthologous genes do not necessarily share cycling expression patterns in the PSM, dividing the vertebrates studied so far into two major groups. In zebrafish and X. laevis, notch is cyclically activated by the delta ligand (deltaC/Delta-2) orthologous to mouse Dll3 (reviewed in Holley, 2007; Jen et al., 1997, 1999). In mouse and chicken, which are amniotes, notch is cyclically inhibited by lunatic fringe, a glycosyltransferase expressed in the Golgi complex that modifies the ability of the notch receptor to bind to the delta ligand. In addition, dynamic expression of delta-like 1 (Dll1) has been reported in the mouse, but not in other vertebrates to date.

There are major gaps in the representation of phylogenetic groups in the study of the evolution of the segmentation clock. In particular, reptiles are a diverse class and current classifications include the birds, together with crocodilian reptiles such as the alligator, within the infraclass Archosauromorpha. Analysis of the segmentation clock in the Lepidosauromorpha, the other main infraclass within the diapsids that includes the lizards and snakes, would be instructive given that avian reptiles display many convergent developmental features with mammals, such as a four-chambered heart and endothorax. However, without full genomic resources, molecular analysis of the segmentation clocks of reptiles has been limited.

Recently, the first genome sequence of a non-avian reptile was reported for the green anole lizard, Anolis carolinensis (Alfoldi et al., 2011). Using this genomic data, combined with transcriptome sequence generated by our group, we sought to analyze coding sequences and dynamic gene expression patterns, which may reflect non-coding regulatory changes. Our aim was to test whether the segmentation clock in the green anole would share greater commonality with the amniote models mouse and chicken, e.g., display lunatic fringe cycling expression in the PSM, or have unexpected similarities with the amniote representatives, X. laevis and zebrafish. This analysis would give insights into regulatory changes in the segmentation clock associated with vertebrate evolution. Our results show that the segmentation clock in the green anole displays transitional features of both amniotes and amniotiques, pointing to a major shift in the segmentation clock associated with the divergence of amniotes.

Materials and methods

Anole and alligator embryos

A. carolinensis lizards were housed at 70% humidity (14 h at 28 °C daylight and 10 h at 22 °C night) conditions. All animals were maintained according to Institutional Animal Care and Use Committee guidelines. Eggs were typically laid in the soil of a potted plant at a 25–30 somite pair stage. Embryos were dissected from eggs in PBS, fixed in 4% paraformaldehyde, dehydrated and stored in methanol at −30 °C.

Eggs from the American alligator, Alligator mississippiensis, were collected from a wild nest by staff of the Louisiana Department of Wildlife and Fisheries from the Rockefeller Wildlife Refuge. Eggs were collected approximately 2–3 days after laying and incubated until 10 days post-laying. At that point, embryos were dissected from these eggs stored in RNA Later (Qiagen) for transit and later fixed in 4% paraformaldehyde, dehydrated and stored in methanol at −30 °C.

RNA-Seq transcriptome analysis of anole embryos

To carry out our analysis of the segmentation clock in the lizard, we needed to improve annotation for somitogenesis genes from the A. carolinensis draft genome assembly (Anocar2.0) and cDNA sequences (Alfoldi et al., 2011). Embryos at 28 and 38 somite pair stages were collected for extraction using the total RNA protocol of the miR Vana kit (Ambion). Total RNA samples were prepared using the Ovation RNA-Seq kit (NuGEN) to generate double stranded cDNA, and Illumina reagents were used to generate paired end sequencing libraries following manufacturer protocols. Sequencing was carried out on a HiSeq 2000 (Illumina) using paired end chemistry with read lengths of 104 base pairs. Reads were mapped to the A. carolinensis genome using Bowtie and TopHat as described previously (Langmead et al., 2009). Based on the number of reads aligned to each transcript, Cufflinks was used to generate an estimation of transcript abundance as Fragments per Kilobase of exon per Million fragments mapped (FPKM). The FPKM estimation is generated by determining the likelihood for the abundances of the set of transcripts, based on the mapped fragments and reporting the abundances with the maximum likelihood (Roberts et al., 2011). RNA-Seq analysis of A. carolinensis embryos allowed us to identify coding sequences for orthologues of key somitogenesis genes, based on synteny conservation comparisons and sequence alignments as outlined by the Anolis Gene Nomenclature Committee (Kusumi et al., 2011; Table 1 and Supplemental Table 1; Supplemental Figs. 1 & 2). Together with the A. carolinensis second genome draft (Anocar2.0; Alfoldi et al., 2011), gene annotations were established and primers were selected for RT–PCR derived RNA probe synthesis.

Comparative analysis of coding sequences

Phylogenetic analysis was carried out using MEGAS (Tabura et al., 2011). Amino acid sequences were aligned using MUSCLE (Edgar, 2004). Evolutionary history was inferred using the Neighbor-joining method (Saitou and Nei, 1987). Bootstrap tests with 500 replicates were used to estimate the confidence in each branch point, which is displayed as the percentage of replicates in which the associated taxa cluster together (Felsenstein, 1985). Evolutionary distances were computed using the Poisson correction method and are in the units of the number

Table 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome/scaffold</th>
<th>5′ end start bp</th>
<th>3′ end finish bp</th>
<th>28 Somite (FPKM)</th>
<th>38 Somite (FPKM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>axin2</td>
<td>GL343260.1</td>
<td>882504</td>
<td>926506</td>
<td>11.5</td>
<td>12.1</td>
</tr>
<tr>
<td>dll1</td>
<td>GL343261.1</td>
<td>222341965</td>
<td>223355331</td>
<td>43.9</td>
<td>36.0</td>
</tr>
<tr>
<td>dll3</td>
<td>GL343263.1</td>
<td>268700</td>
<td>251452</td>
<td>1.0</td>
<td>0.6</td>
</tr>
<tr>
<td>dll4</td>
<td>GL343264.1</td>
<td>67781</td>
<td>53723</td>
<td>2.3</td>
<td>1.6</td>
</tr>
<tr>
<td>loss6</td>
<td>GL343239.1</td>
<td>1277301</td>
<td>1254864</td>
<td>3.9</td>
<td>13.1</td>
</tr>
<tr>
<td>hes4</td>
<td>GL343334.1</td>
<td>1044217</td>
<td>1038209</td>
<td>9.6</td>
<td>7.4</td>
</tr>
<tr>
<td>hes6a</td>
<td>GL343400.1</td>
<td>28075541</td>
<td>28070611</td>
<td>1.1</td>
<td>2.9</td>
</tr>
<tr>
<td>hes7a</td>
<td>GL343410.1</td>
<td>571697</td>
<td>574062</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>hey1</td>
<td>GL343415.1</td>
<td>25153734</td>
<td>25159192</td>
<td>4.3</td>
<td>2.5</td>
</tr>
<tr>
<td>hey2</td>
<td>GL343415.1</td>
<td>180536752</td>
<td>180585826</td>
<td>4.3</td>
<td>5.3</td>
</tr>
<tr>
<td>hey3</td>
<td>GL343415.1</td>
<td>134087655</td>
<td>134141003</td>
<td>33.1</td>
<td>19.5</td>
</tr>
<tr>
<td>hey4</td>
<td>GL343415.1</td>
<td>209497</td>
<td>250732</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td>mesp2</td>
<td>GL343415.1</td>
<td>13650</td>
<td>5835</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>notch1</td>
<td>AAW202036232</td>
<td>2523</td>
<td>29071</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>spry2</td>
<td>GL343279.1</td>
<td>96689658</td>
<td>96690599</td>
<td>2.5</td>
<td>3.2</td>
</tr>
<tr>
<td>tbx6</td>
<td>GL343279.1</td>
<td>3650052</td>
<td>356893</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>wnt3a</td>
<td>6</td>
<td>3407098</td>
<td>3441632</td>
<td>0.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>
A. carolinensis hes6b, hes7a and hes7b predicted sequences were determined through blast analysis of orthologues in mouse, identification of open reading frames and possible splice junction sequences. Sequences for Anolis hes6b and hes7a were further validated with RNA-sequencing data and RT-PCR for ISH probe DNA template generation. Annotated sequences used for MEGA analysis were as follows:

```plaintext
> A. carolinensis hes6b
MG551188.1
Cloned from 5'UTR

> A. carolinensis hes7a
MG551189.1
Cloned from 5'UTR

> A. carolinensis hes7b
MG551190.1
Cloned from 5'UTR
```

**In situ hybridization expression analysis**

For generation of antisense probes for in situ hybridization, the T7 conserved sequence and clamp sequences (5'-GCGTAATCCGGCATCGT-3') were added to the 5' end of each reverse primer sequence. Primers were designated as forward when in the sense strand, and reverse when corresponding to antisense sequence. Whole mount in situ hybridization was carried out as described previously (Sewell et al., 2009). Forward and reverse primers for each gene are listed below:

**Anolis carolinensis**

- *axin2* (5'-GCTACCAACCTGTCATC-3', 5'-ACATTTGCTGCGTCTCTTTG-3')
- *dll1* (5'-TGTTCACAATGGTGAGA-3', 5'-GTTCCACCCCTTTTGA-3')
- *dll3* (5'-GGTCCCTCCATTTCAAGT-3', 5'-CAGTGACGAGGAGGAC-3')
- *dll4* (5'-GAAGACAGGGTGACCAAACA-3', 5'-AACCCCTTGGGGGAC-3')
- *disp6* (5'-GAGGGGAAATGTGGGATTTCT-3')

**Genomic sequence of delta-like 1 and lunatic fringe orthologues**

**Allecat mississippiensis**

**Analysis for conserved regulatory elements**

Genomic sequence of delta-like 1 and lunatic fringe orthologues from *A. carolinensis*, *X. tropicalis*, zebra finch, chicken, and mouse, including the coding region plus 10 kb of both 5' and 3' flanking sequence, were analyzed by MUS2A (Kunz et al., 2008) for potential N- and E-box regulatory elements. Elements were mapped to the genomic sequence, were analyzed by MUS2A (Kunz et al., 2008).
**Accession numbers**

Genomic and cDNA data described in this paper have been deposited into GenBank under accession numbers 28S RNA-seq data, GSM848765; 38S RNA-seq data, GSM848766; for *A. carolinensis* (axin2, JQ303083; dll1, JQ303084; dll3, JQ303085; dusp6, JQ303086; hes6a, JQ303087; hes7a, JQ303088; hey1, JQ303089; hey2, JQ303090; ifng, JQ303091; notch1, JQ303092; spry2, JQ303093; tbx6, JQ303094; wnt3a, JQ303095) and *A. mississippiensis* (DLL1, JQ303096; LFNG, JQ303097). RNA-Seq data are available from the NIH Gene Expression Omnibus (GEO) record series GSE34415.

**Results**

*Transcriptomic analysis of anole embryos identifies segmentation clock genes*

To carry out analysis of the segmentation clock in the green anole lizard, an essential first step was to identify orthologous genes. Public databases have generated predicted gene models based on the *A. carolinensis* Anocar2.0 genome assembly (Ensembl, 2011); however, these predictions were incomplete and data from available cDNA sequences were not incorporated. Furthermore, segmentation clock genes are not well represented in the current cDNA libraries sequences, which derive mostly from adult tissues, so we carried out RNA-Seq transcriptome analysis focused on somite-stage embryos. We collected two *A. carolinensis* embryos from newly laid eggs, at 28 and 38 somite pair stages (Figs. 1B,C). RNA-Seq analysis determined that 53% of the total predicted genes were expressed at this developmental stage, i.e., 11,170 and 11,314 transcripts from the 28 and 38 somite-stage embryo samples, respectively, displayed a range of non-zero Fragments per Kilobase transcript per Million reads (FPKM) values (Figs. 1D,E; Table 1; Supplemental Table 1). The FPKM value provides an estimate of mRNA quantity, based on a probability distribution of transcript abundance derived from the number of aligned sequencing reads. Transcripts with zero FPKM represented genes that were expressed at extremely low levels in the whole embryo. These two embryos were both collected at mid-somitogenesis stages, and as might be expected, the FPKM levels for all transcripts had a very high Pearson’s correlation coefficient of 0.99534. We were able to refine annotations for key segmentation clock genes in the anole, based on transcriptomic and genomic data (Table 1). Assignments of orthology were confirmed by analysis of syntenic conservation (Supplemental Fig. 1).

Using these refined gene annotations for *A. carolinensis*, we carried out comparative sequence analysis of segmentation clock genes (Fig. 1F; Supplemental Fig. 2). Compared with housekeeping genes such as gapdh and actn1, the segmentation clock genes varied considerably in ClustalW scores (Fig. 1F; Larkin et al., 2007) and multiple sequence alignment based analysis by MEGA (Supplemental Fig. 2). Predicted protein sequence similarity scores of the notch effector genes hes6a and hes7a and the notch ligand dll3 were particularly divergent (Fig. 1F; Supplemental Fig. 2) compared with other segmentation clock genes such as notch1, ifng, and axin2.

**hes6a is expressed in a gradient in the PSM of the anole**

Somite boundaries are determined through gradients of gene expression within the PSM interacting with genes in the determination front. Anole orthologues of three genes, wnt3a, fgf8, and tbx6, displayed expression in the posterior PSM (Figs. 2A–C), similar to reports in other vertebrate models (Aulehla et al., 2003; Beck and Slack, 1998; Chapman et al., 2011; Crossley and Martin, 1995; Draper et al., 2003; Dubrulle et al., 2001; Pera et al., 2002; Reiflers et al., 1998; Takada et al., 1994; Thorpe et al., 2005). Similarly, the anole orthologue of mesp2, a key gene in the determination front
that regulates the prepatterning of somite boundaries (Morimoto et al., 2005; Morimoto et al., 2007; Takahashi et al., 2007), was expressed in the −1 region of the PSM (Fig. 2G). In contrast, the notch pathway hairy-enhancer of split (hes) gene 6 is expressed in the posterior PSM in Xenopus and zebrafish but not in mouse and chicken (Fior and Henrique, 2005; Holley, 2007; Koyano-Nakagawa et al., 2000; Pissarra et al., 2000). The hes6a orthologue was expressed in the posterior PSM (Figs. 2D,E) and in a band in the rostral PSM of the anole, unlike the chicken and mouse. Surprisingly, hes6a is a PSM gradient gene for anoles, similar to Xenopus and zebrafish, suggesting that squamate reptiles have transitional features in common with amniotes.

**hes7a is a cycling gene in the PSM of the anole**

The hes7 orthologue is a primary molecular oscillator driving cyclical expression of other notch pathway cycling genes in vertebrates (Bessho et al., 2001; Niwa et al., 2007). We identified two tandem duplications of hes7 in the anole, and the hes7a orthologue displayed cyclical expression (Figs. 2J–L). Other hairy/enhancer of split genes such as hey1 and hey2, which display cycling expression in the mouse (Leimeister et al., 1999), were not evidently dynamic in anole embryos (Figs. 2H,I). The anole hes4 gene is an orthologue of the mouse cycling gene Hes1 (Jouve et al., 2000) but was not expressed within the PSM (Supplemental Fig. 3C). The segmentation clock regulates the activation of the notch receptor, and consistent with this role, notch1 is expressed in the determination front and the newly formed somites (Fig. 2F). The anole jag1 orthologue is expressed in a similar pattern to the mouse notch ligand Jag1, which is seen in a static band in the anterior PSM and in the tailbud (Xue et al., 1999; Supplemental Fig. 3D). Wnt and FGF pathway genes have been identified with cyclical expression in the mouse and chicken (Aulehla et al., 2003; Gibb et al., 2009; Krol et al., 2011), with no evidence for oscillatory expression of these genes in the corn snake, Pantherophis guttatus (Gomez et al., 2008). The expression of the anole orthologues of the wnt pathway gene axin2 and the FGF gene dusp6 were not dynamic in expression (Supplemental Fig. 3A,B). Thus, A. carolinensis shares components of the hes-driven segmentation clock, such as hes7a, but the cyclical expression of other genes is not conserved with the mouse and chicken.

**Lunatic fringe is not expressed in the PSM of anole embryos and lacks enhancer regions found in mouse and chicken**

Given the cyclical expression of the modulator lunatic fringe in both chick and mouse, this gene was thought to be a key dynamic notch regulator in amniotes (Barrantes et al., 1999; Evrard et al., 1998; Forsberg et al., 1998; Serth et al., 2003). However, we observed that anole lfng was not expressed in the PSM and instead localized to the rostral compartment of the first new somites (Fig. 3A). As in the mouse and the chicken, anole lfng is also expressed in the developing neural tube (Figs. 3B,C). Thus, lfng is not a cycling gene in the anole, suggesting either that PSM-specific expression was lost in ancestors of the anoles, or that there is potential convergence between birds and mammals. Amino acid sequences of A. carolinensis fringe genes were compared between the other vertebrate orthologues using MEGA 5 (Tamura et al., 2011). As expected, A. carolinensis fringe...
genes shared the most sequence similarity with chicken fringe genes (Fig. 3D), so coding sequence divergence is not a likely mechanism for the differences observed in the anole.

Cyclical expression in the PSM is driven by regulatory feedback loops involving binding of hes proteins such as HES7 to non-coding regulatory elements (Chen et al., 2005; Cole et al., 2002; Morales et al., 2002). These hes bHLH transcription factors bind to regulatory sequences, termed N and E boxes, which are conserved among metazoans (reviewed in Davis and Turner, 2001). To test whether changes in the number or distribution of N and E-box elements could account for the convergence of lunatic fringe expression in the mouse and the chicken, we compared genomic regions up to 10 kb upstream and downstream of these genes. Changes in sequences associated with lunatic fringe cycling expression were observed. Specifically, a large cluster of N and E-box elements was identified in −3 to −4 kb upstream region in the chicken LFNG,
comparable to a −1 to −2 kb upstream cluster in mouse Lfng that has been demonstrated to be required for cycling expression (Cole et al., 2002; Morales et al., 2002). No comparable enhancer cluster was identified in anole Lfng (Fig. 3E). Thus, the difference between the −3 to −4 kb upstream N and E-box cluster in the chicken vs. the −1 to −2 kb cluster in the mouse and lack of a major cluster in the anole may reflect convergent molecular evolution of non-coding elements required for lunatic fringe cycling expression.

Delta-like 1 is a cycling gene in the anole

Dynamic expression of delta-like 1 has been reported in the mouse PSM (Maruhashi et al., 2005); however, given expression of Dll1 throughout the mouse PSM, it is difficult to detect clearly sweeping bands of expression within this region. In the anole, in situ hybridization analysis showed that dll1 was limited in expression within the PSM and displayed clear phases of cycling expression (Figs. 4A–E). The dll1 expression pattern can be categorized into three distinct phases of expression, with expression shifting from the caudal to the rostral PSM (Figs. 4C–E). Comparative analysis of delta-like protein sequence did not reveal any unexpected similarity between the mouse and anole orthologues (Fig. 4F).

Delta-like 3 does not display cycling expression in the anole but conserves EGF repeat structure with amniotes orthologues

Amniotes and amphibians have three paralogues of delta-type notch ligands, and the second delta group, which includes frog X-Delta-2, mouse Dll3, and zebrafish deltaC, is divergent in domain structure and cyclical expression (Fig. 5D). Anole Dll3 is expressed in stationary, non-cycling bands within the rostral PSM and tailbud (Figs. 5A–C), but its localization to the caudal somite compartment (Fig. 3C) differs from mouse Dll3 localization to the rostral compartment (Dunwoodie et al., 1997; Kusumi et al., 1998). The zebrafish del-taC and the X. laevis X-Delta-2 are cycling genes within the PSM (Holley, 2007; Holley et al., 2000, 2002; Jen et al., 1997, 1999; Jiang et al., 2000). In contrast, the mouse orthologue Dll3 is expressed throughout the PSM (Dunwoodie et al., 1997; Kusumi et al., 1998), has lost two EGF repeat domains, and functions as an inhibitor of notch signaling in somitogenesis (Fig. 3D; Chapman et al., 2011; Geffers et al., 2007). A delta-like 3 orthologue has not been discovered in the three available bird genomes and may have been deleted during avian evolution (Supplemental Fig. 1). The sequence of the second delta ligand in the anole was most similar to X. laevis X-Delta-2 and X. tropicalis dll3 (Fig. 4F). Therefore, while the domain structure of anole Dll3 shares similarity with anamniotes, the lack of cycling in anole and mouse may reflect an ancestral change in non-coding regulatory function prior to amniote divergence.

In the alligator, lunatic fringe is not expressed in the PSM and Dll1 does not display cycling expression

To further investigate the divergence of lunatic fringe and delta-like 1 expression among amniotes, particularly in Archosaurian evolution, we examined the expression of LFNG in the American alligator, A. mississippiensis using in situ hybridization. Embryos collected at 10 days after laying in the alligator were comparable to the stages of newly laid anole eggs (Fig. 6B). Birds and crocodilian reptiles such as the alligator are both classified in the division Archosauria, and commonality of gene expression between chicken and alligator would suggest that the regulatory changes occurred prior to the divergence of this group. Intriguingly, alligator LFNG was expressed in the neural tube and somites (Fig. 6C), but not in the PSM of somite-stage embryos (Figs. 6C–D), like the anole but unlike the chicken. Furthermore, alligator Dll1 was expressed in a band within the rostral PSM similar to that observed in the green anole, but unlike the lizard, no evidence of cycling expression was observed in 15 embryos examined (Fig. 6E). Thus, the expression of LFNG in the alligator was more comparable to the anole than the chicken, but both the chicken and alligator do not demonstrate Dll1 cycling expression in somitogenesis.

Discussion

Our analysis of somitogenesis in the green anole lizard, A. carolinensis, and the American alligator, Alligator mississippiensis, provides novel insights into the evolution of vertebrate somitogenesis and the segmentation clock in amniotes. Shared features between anoles and the chicken, mouse, Xenopus, and zebrafish include the fgf8 and wnt3a PSM gradients, mesp2 expression at the determination front, and cyclical expression of the notch pathway hes genes. However, we have identified four divergent features from other vertebrate models, which suggest major shifts in the regulation of the segmentation clock associated with the evolution of the amniotes: 1) retention of a hes6a PSM gradient in the anole, suggesting this is a basal feature of the amniotes lost in the mouse and chicken; 2) loss of cyclical expression of dll3/deltaC orthologues in the PSM in anamniotes, including the anole; 3) cyclical or dynamic expression of dll1 orthologue in the anole and mouse but not the alligator, suggesting this oscillatory expression arose in the amniotes but was lost in the Archosaurian ancestor, and 4) cyclical expression of lunatic fringe in the mouse and chicken but not the alligator and alligator. These changes are associated with divergence in coding and non-coding sequences that has arisen during the evolution of vertebrate somitogenesis.

The evolution of coding versus non-coding sequences of developmental genes

The segmentation clock is driven by the expression of genes with tightly regulated spatial and temporal patterns of expression. Amino acid conservation of the dll1, wnt3a, fgf8, and Lfng matches expectations between the anole and other vertebrates (Fig. 1F). The hes genes, including hes6 and hes7 orthologues (Fig. 1F) have high divergence when comparing full amino acid sequence, but are much more conserved in the functional basic and helix–loop–helix domains. The standout exceptions include the delta-like 3 orthologue, which is highly divergent in the mouse due both to loss of two EGF repeat domains in this ligand and general sequence divergence (Fig. 5D). The divergence of mammalian Dll3 has also been associated with a change in cellular localization to the Golgi complex and shift in functional role from a trans-activatory ligand to primarily a cis-inhibitory factor (Chapman et al., 2011; Geffers et al., 2007; Ladi et al., 2005). In addition, there has been no avian orthologue of Dll3 identified (Supplemental Fig. 1C). Within the vertebrates, there is divergence in both the localization of the dll3 expression within the PSM and whether there is cycling expression. In the anole, Xenopus, and zebrafish, the dll3 orthologues are expressed in two regions, i.e., the tailbud and the rostral PSM; however, this pattern is static in anoles but cycling in X. laevis and zebrafish (Figs. 5A–C). Like the anole, Dll3 in the mouse is static in expression, but unlike the anole, the expression extends to the entire PSM (Dunwoodie et al., 1997; Kusumi et al., 1998). Thus, the similarity in the PSM expression of dll3 orthologues between the anole and X. laevis/zebrafish suggests an ancestral amniote pattern, with potential regulatory changes leading to a loss of cycling expression.

To display cycling expression, segmentation clock genes first must be expressed in the PSM, which requires transcriptional activation specific to these cells. Paraxial mesoderm-specific enhancers have been identified for mouse Dll1 (Beckers et al., 2000; White and Chapman, 2005; White et al., 2005). For lunatic fringe, there is no PSM expression in the anole and alligator but there is cycling expression in the chicken and mouse. PSM-specific enhancers have been identified in mouse Lfng
Insights from analysis of anole and alligator somitogenesis into the evolution of the segmentation clock

Since developmental and molecular tools are currently available for only a limited number of organisms, each class or infraclass has been typically represented by a single species at best. Based on our analysis in the anole and alligator, we can generate new hypotheses about the evolution of the segmentation clock. With the rapidly decreasing cost of next-generation sequencing and ability to test many additional vertebrates, these hypotheses can help us to select the most informative species for further analysis.

1. Predicted segmentation clock components of the ancestral vertebrate (Fig. 7): Positional information along the rostral–caudal axis of the PSM was likely established by gradients of soluble ligands WNT3A and FGF8 with gene expression gradients of tbx6 and hes6. Expression of the hes6 orthologue was then later lost in both the mammalian and avian radiations. The determination front is regulated by mesp orthologues, interacting with cyclical patterns of gene expression driven by hes genes (hes1 or hes7 orthologues) and their downstream targets. The cycling gene regulating notch activation was the delta-like 3/deltaC/X-Delta-2 orthologue. While it is possible that all three genes—dll3, dll1, and lfng—displayed cycling expression in the PSM of an ancestral vertebrate, cycling expression would have to have been lost for dll1 and lfng in both amphibian and teleost lineages.

2. Predicted segmentation clock components in the ancestral amniote (Fig. 7): For cycling genes in amniotes, analysis of additional reptile models will help to distinguish between two alternate models for both lunatic fringe and delta-like 1.

For lfng, the amniote ancestor could have displayed cycling expression, which was subsequently lost in the anole and alligator, or cycling expression of lfng in the PSM could have arisen independently in birds and mammals. For dll1, the amniote ancestor could have displayed cycling expression as seen in the anole, and mouse, with loss by deletion analysis (Cole et al., 2002; Morales et al., 2002). Potential N and E-box binding sites were not identified — 500 to —2,000 bp 5’ flanking sequence in the chicken, anole or X. laevis (Fig. 3E), but intriguingly, the chicken 5’ flanking region contains a more distant cluster of N- and E-boxes that are potential hes binding sites. Expression of lunatic fringe has been reported for the corn snake (Gomez et al., 2008). The PSM of the snake appears to have undergone an axial extension, with expansion of the region rostral to the determination front. This has resulted in a region rostral to the determination front with multiple lunatic fringe bands of expression, which are dynamic. However, lunatic fringe in the corn snake does not appear to be a cycling gene, i.e., a gene whose expression oscillates between the caudal-to-rostral PSM caudal to the determination front. For other segmentation clock genes such as hes6 or dll3, enhancers required for PSM expression have not yet been identified.

Oscillatory expression of the hes genes has been demonstrated to derive from auto-inhibitory negative feedback loops, with HES hetero- and homo-dimers binding to N- and E-box binding sites. The presence of these binding sites has also been identified in key enhancers required for downstream cycling genes, such as has been shown for lunatic fringe in the mouse (Cole et al., 2002; Morales et al., 2002). While the enhancers driving paraxial mesoderm-specific expression have been identified for delta-like 1 (Beckers et al., 2000; White and Chapman, 2005), the elements required for cyclical expression have yet to be defined. The delta-like 3 orthologues, deltaC and X-Delta-2, display cycling expression in the zebrafish and Xenopus, respectively, but analysis for regulatory sites directing this oscillation has also not been reported. The amniote orthologues of dll3 do not display cycling expression, and analysis of 10 kb 5’ flanking regions has been inconclusive.

Fig. 4. Anolis carolinensis dll1 displays cycling expression in the PSM. A,B: Expression of dll1 is localized primarily in the lizard presomitic mesoderm (PSM), where dynamic shifts are observed. C–E: Dynamic expression of dll1 can be categorized into three phases characteristic of cycling genes (n=19). F: The evolutionary history comparing Anolis dll1—chicken DLL1—Xenopus dll1—mouse Dll1—zebrafish deltaA—zebrafish deltaD—Xenopus dll4—mouse Dll4—Anolis dll4—chicken DLL4—mouse Dll3—Anolis dll3—Xenopus dlc—zebrafish deltaB—zebrafish deltaC.

of cycling expression in Archosauria, or this cycling could have arisen independently in mammalian and squamate evolution.

In addition to the components detailed above, the ancestral amniote could have lost cyclical expression of $dll3$ in the PSM. If the $dll3$ orthologue was no longer an essential cycling activator of notch signaling, replaced by either or both lunatic fringe and delta-like 1 in the amniote ancestor, the loss or rapid divergence of the gene became possible. In mammals, the $Dll3$ gene could have undergone rapid divergence in coding sequence (Figs. 4F and 5), changes in intracellular localization (Chapman et al., 2011; Geffers et al., 2007; Ladi et al., 2005), and functional shift from trans-activation to cis-inhibition of notch signaling. In birds, the gene may have been completely deleted.

Expression of genes in the segmentation clock, particularly cycling genes, is a sensitive read-out of the changes in regulatory networks during vertebrate evolution. These findings highlight the challenge of integrating evolutionary studies of developmental networks with sequence-based phylogenetic analysis (Fig. 7). Further analysis of chelonian and additional crocodilian reptiles (Mansfield and Abzanov, 2010), urodele amphibians, paleognath birds, and monotreme mammals could yield further insights into the complex evolution of divergent and convergent processes in vertebrate development.

Supplementary materials related to this article can be found online at doi:10.1016/j.ydbio.2011.11.021.

Acknowledgments

We thank N. Emmert, A. Wooten, M. Ammar, I. Maayan, and E. Hutchins for technical assistance; N. Crawford and C. Schneider for assistance in bioinformatics analysis; S. Pratt and A. Sidow for comments on the manuscript. We thank the Broad Institute Genome Sequencing Platform and Genome Sequencing and Analysis Program, Jessica Alföldi, Federica Di Palma and Kerstin Lindblad-Toh for making the unpublished data for Anolis carolinensis available. This work was supported by Arizona State University (K.K.) and the David and Lucile Packard Foundation (C.R.I. and J.B.L.)
References


Fig. 6. In the Alligator mississippiensis embryonic PSM, LFNG expression was not detected and DLL1 does not display cycling expression. A, Based on consensus analysis of coding sequence variation, both the American alligator and the chicken are categorized as Archosaurian reptiles (Hedges et al., 2006). An adult alligator is shown on the right. B. Alligator mississippiensis embryo at approximately day 10 of development. Scale bar is 1 mm. C, Alligator mississippiensis LFNG is expressed in the neural tube and somites but expression is not observed in PSM (C,D; n = 5), which is similar to the pattern observed in the green anole. E, Alligator DLL1 expression is observed as band in the rostral PSM that is static and does not display cycling expression (n = 15).
Fig. 7. Diagram comparing how the phylogenetic relationships of the zebrafish, frog *Xenopus laevis*, lizard *Anolis carolinensis*, chicken and mouse (Hedges & Kumar, 2009) with conserved and divergent features of the segmentation clock. Hypothetical vertebrate ancestor (light blue VA) and amniote ancestors (purple AA) are indicated. Shared among all these vertebrate developmental models is the cycling expression of hairy/enhancer of split (hes) genes, *mesp2* expression at the determination front for prepatterning of somite boundaries, and *fgf8* and *wnt3a* gradients (light blue blocks; Aulehla et al., 2003; Holley, 2007; Sparrow; Dequeant and Pourquié, 2008; Gibb et al., 2010). Notch receptor activation in zebrafish and frog is regulated via cyclical expression of the ligand *deltaC/IX-Delta-2* (orange blocks; Jen et al.; Jen et al., 1999; reviewed in Holley, 2007). A gradient of *hes6* orthologue expression is observed in zebrafish, Xenopus and anole, but not in mouse and chicken (green blocks; Koyano-Nakagawa et al., 2000; Pissarra et al., 2000; Fior and Henrique, 2005; Holley, 2007). A key oscillatory gene in the anole is the *dll1* ligand, which has also been reported to display dynamic expression in mouse PSM (purple blocks; Maruhashi et al., 2005; Krol et al., 2011). Notch receptor activation in chicken and mouse is regulated by cyclical expression of the modulator *Ifng*, which is not expressed in the PSM of the anole or alligator (red blocks; McGrew et al., 1998; Cole et al., 2002; Morales et al., 2002).


