Physical Biology

PAPER

Positional information from oscillatory phase shifts: insights from in silico evolution

M Beaupeux and P François

Ernest Rutherford Physics Building, McGill University, H3A2T8 Montreal QC, Canada
1 Author to whom any correspondence should be addressed.
E-mail: paulf@physics.mcgill.ca

Keywords: gene networks, gene oscillators, vertebrate segmentation, evolution, phase encoding

Supplementary material for this article is available online

Abstract

Complex cellular decisions are based on temporal dynamics of pathways, including genetic oscillators. In development, recent works on vertebrae formation have suggested that relative phase of genetic oscillators encode positional information, including differentiation front defining vertebrae positions. Precise mechanisms for this are still unknown. Here, we use computational evolution to find gene network topologies that can compute the phase difference between oscillators and convert it into a decoder morphogen concentration. Two types of networks are discovered, based on symmetry properties of the decoder gene. So called asymmetric networks are studied, and two submodules are identified converting phase information into an amplitude variable. Those networks naturally display a 'shock' for a well defined phase difference, that can be used to define a wavefront of differentiation.

We show how implementation of these ideas reproduce experimental features of vertebrate segmentation.

Introduction

Signalling information can be efficiently encoded in temporal dynamics of regulatory pathways [1, 2]. In development as well, positional information appears now much more dynamical than first proposed in the classical 'French Flag model' paradigm [3]. Dynamical aspects of Hox genes patterning have been known for a long time [4, 5] and are intertwined with complex regulations of cell movements [6]. Other recent examples include ‘time-integration’ of Shh signalling for neural progenitor [7, 8], or Turing-like patterning in growing tissues [9]. Dynamical positional information is most strikingly observed during vertebrate segmentation (or somitogenesis), where multiple signalling pathways oscillate both temporally and spatially [10, 11] to define position, size and polarity of the vertebrate precursors, called somites [12–14].

An important role for oscillations had first been predicted by the classical clock and wavefront (CW) model by Cooke and Zeeman [15]. A front is proposed to sweep across an embryo considered as a linear array of oscillators, locally synchronized. As the front passes over cells, ticking information of the clock is used to define somite boundaries. Heuristically, this turns a temporal oscillation into a spatial one defining metameric structures (somites). Historically, the CW model was proposed to ensure vertebrae scaling with respect to total embryonic size. Dimensional analysis combining a clock (characterised by period $T$) to the regression of a hypothetical wavefront (speed $v$) naturally predicts a length scale $vt$ (corresponding to somite size), independent from any local property such as diffusion constant or number of cells (see e.g. [16] for an explicit computation of this length scale).

Different variants of the original CW model have been proposed e.g. the original model by Julian Lewis in appendix of [12], and subsequent works have suggested that patterning is mediated with a diverging clock period [17]. More explicit models account for different molecular controls [18], while other models assume bistability as a crucial effector of somite definition [19] or later polarity [20]. Existence of segmentation oscillators has now been observed in multiple vertebrates, and is most probably homologous from snakes to mammals [21]. Recently, oscillations controlling insect segmentation were also discovered [22–
suggesting that the ancestral mode of segmentation of chordates might be based on similar principles. Recent experiments in mouse have challenged the classical CW view [25], and especially the existence of a discrete, well-defined ‘wavefront’ independent of the clock. In particular, it was shown that scaling of somites with respect to pre-somitic mesoderm size (PSM) occurs in culture even without growth, which cannot be easily accounted by a CW model. Such scaling would require some unknown long range mechanism for PSM size sensing in the CW framework. However the scaling property can also be observed on the phase gradient of kinematic waves of Notch oscillations in culture, without a full-size PSM, and way before the appearance of any morphological somite boundary [25]. This strongly suggests that PSM-size is not actively ‘measured’ by the system, and excludes any signal propagating from the differentiated anterior.

A parsimonious way to explain these data is to assume that individual cells act as autonomous clocks, slowing down the phase of their Notch oscillation with respect to some reference phase, corresponding experimentally to the phase in the tail bud [25]. The differentiation front indeed appears to be well-defined by a phase difference of Notch oscillations with the tail bud, which suggests that positional information on somite boundary is defined directly by the oscillations. This is reminiscent of a 2-oscillator model proposed by Goodwin and Cohen [26], with an added feedback to account for the slowing down of one of the somitic oscillators. An important difference is that the Goodwin Cohen model necessitates very fast oscillations compared to timing of patterning, while in somitogenesis, the period of oscillation is roughly the period of pattern formation so that phase computation has to be done in real time and can not be averaged out over many oscillations.

These observations prompted us to propose a simple phenomenological ‘2-oscillator’ model driving somitogenesis [25] for single cells in PSM after their tail bud exit. The first oscillator (phase $\phi_0$) is synchronized in the embryo with fixed frequency while the second one (assumed to be Notch, phase $\phi_1$) slows down with respect to $\phi_0$ following

$$\dot{\phi}_0 = 1,$$  \hspace{1cm} (1)

$$\dot{\phi}_1 = 1 + \alpha (\phi_1 - \phi_0),$$  \hspace{1cm} (2)

where time has been rescaled so that the intrinsic period of segment formation is $2\pi$. To avoid confusion, we will refer to $\Delta \phi (x, t) = \phi_1 (x, t) - \phi_0 (x, t)$ as the local phase difference (between oscillator 1 and 0), and to $\Delta \phi_1 = \phi_1 (x, t) - \phi_1 (0, t)$ as the phase shift between local and tail bud Notch phase (which is the quantity described in [23]). These phenomenological equations strikingly reproduce many features of somitogenesis scaling [25]. Biophysical interpretation of equations (1) and (2) is straightforward: if $\alpha < 0$ Notch oscillator is slaved to the reference oscillator, while if $\alpha > 0$, Notch oscillator phase is ‘repelled’ by reference oscillator.

Biologically, dephasing of Notch oscillations occurs as cells move more and more anteriorly in PSM after their tail bud exit. Phenomenologically, segmentation occurs for some well-defined phase shift $\Delta \phi_1^*$ between local and tail bud phase for Notch [25], and thus is done in continuity with the autonomous slowing down of the clock described by equations (1) and (2). This suggests that the relative phase of the two oscillators is the crucial biological variable accounting for positional information, contrary to the classical the clock and the wavefront model where CW are essentially two independent entities until they interact.

Phase equations similar to (1) and (2) are widely used to model coupling of similar oscillators at different positions [27]. Here we consider what happens if two different oscillators coexist and are coupled within the same cell, assuming phase difference between the two clocks is used to define temporal and positional information. Thus the natural question: what kind of signalling network can actively ‘compute’ a phase difference between two oscillators within the same cell, to eventually define a phase shift map in the embryo? In this manuscript, we turn to computational evolution to answer this question [28]. We use a mutual information based fitness to select for networks actively sensing the phase difference between two oscillating pathways. Characteristic network topologies evolve, and their behaviour is studied. Finally, improvements of these networks and implications for somitogenesis are discussed.

Method

Our goal is to evolve in silico networks that are capable of measuring the phase difference between two genetic oscillators through time, using computational evolution [28]. This method consists in simulating evolution of a population of gene networks, through a sequence of growth, selection and random mutation (see other examples of similar methods in [29–31]). The algorithm used here is similar to the one reviewed in [32]. We summarise in the following the most important features of these simulations.

We limit ourselves to purely transcriptional networks, where variables are transcription factor (TF) concentrations. To each network corresponds a set of ODEs encoding for network dynamics. Activations and repressions are modelled with Hill functions, in a similar way to [20], and degradation is assumed to be linear. As an illustration of the formalism used, consider a protein $P$, with degradation $\delta$, transcriptionally activated by two proteins $A$ and $B$ and repressed by $R$. In that case, given the two activating hill functions $\rho_A = \frac{A^n}{A^n + A_0^n}$ and $\rho_B = \frac{B^n}{B^n + B_0^n}$ ($n$, $m$ Hill coefficients, $A_0$, $B_0$ Hill thresholds), we assume a ‘OR’ like combinatorics, in the sense that the actual production
rate is the maximum of these two Hill activities
\( \rho_{\text{max}} = \rho_P \max (\rho_A, \rho_B) \) with \( \rho_p \) a parameter defining
maximum possible transcriptional rate. Repressors are
assumed to act multiplicatively, i.e. if \( R \) represses
the same gene, the production rate of this network is
with same notations, \( \rho = \rho_{\text{max}} \frac{R^k}{R^k + R^m} \), so that in the
end, the evolution equation for \( P \) is

\[
\frac{dP}{dt} = \rho_P \max \left( \frac{A^n}{A^n + A_0^n}, \frac{B^m}{B^m + B_0^m} \right) \times \frac{R_0^l}{R^l + R_0^l} - \delta P.
\]

For representation purpose, in the following, genes are indexed by integers and are called \( S_i \), al-
together with the protein they encode.

The algorithm acts on a population of networks
(typically 40). All networks in the population are ini-
tialised at generation 0 with two input genes (corre-
spending to our two oscillators, proteins are
subsequently called \( S_0 \) and \( S_1 \)), and one output gene
(corresponding protein is subsequently called \( D \); the
nature of \( D \) can be randomly changed by the algorithm
and is thus under influence of selection). We present
here results for networks evolved over 550 genera-
tions. At each generation, differential equations
corresponding to these networks are integrated. Then
a fitness function (see below) quantifies the per-
formance of these networks. Selection is made based on
this fitness: the worst half of the networks is discarded,
while the best half is kept, duplicated and then muta-
ted. Practically we also add a very small random num-
ber to the fitness (of order 0.001), to ensure that the
best network at generation \( N \) is not necessarily the
direct copy of the best network at generation \( N - 1 \) in
the absence of mutations.

There are three kinds of possible mutations:

- Random changes of kinetics. In this case, we
randomly draw new values of the kinetic parameter.

- Addition/removal of transcriptional activations.
When adding interactions, we randomly choose one
TF and one target, and randomly draw the kinetic
parameters associated to the interaction.

- Addition/removal of a TF. New TFs are assumed to
be initially produced with random constant rate,
and are degraded linearly.

When creating or changing an interaction, all rates
are randomly drawn in a predefined range so that all
concentrations stay of order 1 in arbitrary units. (see
supplement for more details on the range of parameters).

Mutations are drawn randomly with predefined
rates, for the simulations presented in the paper, the
probability to change parameters, add/remove a gene
or an interaction are identical. Mutation rates are
adjusted dynamically to have on average one mutation
per network per generation: this ensures incremental
evolution and implements a computational equivalent
of the Drake rule [33]. Like in our previous work, we
tested some variations in these parameters, and the
results do not depend much on those mutation rates in
the sense that the dynamics of 'working' networks are
qualitatively the same even when mutation rates are
changed (but success rate or speed of convergence
change).

The goal of the algorithm is to select for networks
computing phase difference between two oscillators.
\textit{A priori}, a gene network downstream of two oscillators
could encode much information about each of the two
oscillators, but the question of how this information is
later decoded is too generic as is. We therefore impose
some constraints: we assume there is a single output
gene \( D \) (for decoder) containing all 'information'
about the phase difference \( \Delta \phi \) between \( S_1 \) and \( S_0 \). To
compute its information content, we use a heuristic
approach similar to the ones in \([34, 35]\): we construct
pseudo-probability density functions on both \( D \) and
\( \Delta \phi \), then compute mutual information on these
pseudo PDF. More precisely, at each generation, dif-
ferential equations corresponding to each network are
integrated, with initial conditions 0 for each gene.
Dynamics of \( S_0 \) and \( S_1 \) are imposed to be sinusoidal. \( S_0 \)
is oscillating at constant predefined frequency \( S_0(t) \propto 1 + \cos(\omega_0 t) \) while \( S_1 \) is oscillating with a
slightly different random frequency and different
initial phase \( S_1(t) \propto 1 + \cos(\omega_1 t + \phi_0) \) imposing a
time-evolving phase shift of the two oscillators (see
details of sampling for \( \omega_1 \) and \( \phi_0 \) in supplement).
Numerical integration is reproduced many times with
randomized values of \( \phi_0 \) and \( \omega_1 \) to ensure representa-
tive sampling of \( \Delta \phi \) and of its dynamics.

For each time point, we store the couple of values
\( (D(t), \Delta \phi(t)) \). These values are then binned to
reconstruct an effective probability distribution
\( P(D, \Delta \phi) \) (we considered 64 bins). Such binning has
a natural biological interpretation: realistically there will
be some intrinsic noise preventing a perfect measure,
and finer binning means less intrinsic noise.

We can define mutual information between \( D \) and
\( \Delta \phi \) as

\[
F = I(\Delta \phi : D) = S(\Delta \phi) - S(\Delta \phi | D),
\]

where \( S \) is the standard Shannon entropy
\( S(X) = -\sum P(X = x) \log_2 P(X = x) \). We use \( F \) as
the fitness to optimise for each network. Figure 1
summarises how the fitness is computed. Note that
since we are computing the mutual information for
the entire time-course starting at 0 initial condition,
there could be some influence of a transient behaviour,
but with many values of \( \phi_0 \) and \( \omega_1 \), rapid convergence
of \( D \) towards a value faithfully tracking the phase shift
should be ensured. Indeed we typically observe very
fast convergence so that the transient behaviour is
irrelevant, see e.g. simulated time courses on figure 3.
We can thus consider that $D$ reaches some stationary behaviour very quickly.

Finally, it should be noted how the relation between $D$ and $f_D$ on the example of figure 1 looks qualitatively similar to actual biological input/output relationship in developmental context (e.g. the Bcd/Hb relationship in fly from [36]). This is not a coincidence: as detailed by Gregor and co-worker, such probabilistic input/output relationship gives quantification of the positional information transmitted by gap genes network [37]. The problem we consider here is very similar, except information is encoded in a dynamical way (phase difference $\Delta \phi$) in contrast to a static morphogen such as Bcd. Furthermore, the dynamical nature of this problem imposes more specific constrains: for instance, given the periodicity of the input signals, the behaviour of $D$ is expected to be periodic in $\Delta \phi$. This is indeed what we observe, and this is the reason why we display $\Delta \phi$ only over a $2\pi$ interval on all our figures. $D$ should also be continuous, which means, because of periodicity, any value of $D$ should correspond to at least two different values of $\Delta \phi$ (except for extrema), a problem computational evolution circumvented as described below.

**Results**

A sample of networks evolved are displayed on figure 2.

Networks found by the procedure can be essentially cast into two groups:

- A first group of networks has the property to relate $D$ to two values of the phase difference, $\Delta \phi$ and its
Figure 2. Sample of networks evolved. Upper row shows the topologies of the networks whose corresponding PDF are displayed on the lower row. Oscillatory input genes $S_0$ and $S_1$ are represented with inverted triangle, output ‘decoder’ genes are represented with a triangle, i.e. respective Outputs for columns (A)–(C) are genes $S_2$, $S_4$ and $S_5$. Other genes are represented by circles, green arrows are activation while red arrows correspond to repressions. $S$ symbol for genes is omitted for simplicity in these graphs and only gene index is displayed. Respective fitnesses for networks (A)–(C) are $F = 1.93, 2.62, 2.66$ bits. Equations for these networks are given in supplement.

Figure 3. (A) Topology of an evolved symmetric network (fitness $F = 1.93$ bits). Input genes are genes $S_0$ and $S_1$, output ‘decoder’ gene is gene $S_2$. (B) Typical dynamics of the symmetric network, when the phase difference between genes $S_1$ and $S_0$ is slowly changing ($\Delta \phi$ linearly spans the $[0, 4\pi]$ interval). Note how species $S_2$ is considerably varying with phase difference. (C) Distribution of values of species $S_4$ as a function of $\Delta \phi$, neglecting transient entrainment. Note the shock for low values of $\Delta \phi$ ensuring the $2\pi$-periodicity in $\Delta \phi$. (D) Topology of an evolved asymmetric network (fitness $F = 2.31$ bits). Input genes are genes $S_0$ and $S_1$, output ‘decoder’ gene is gene $S_4$. (E) Typical dynamics of the network, when the phase difference between genes $S_1$ and $S_0$ is slowly changing ($\Delta \phi$ linearly spans the $[0, 4\pi]$ interval). (F) Distribution of values of species $S_4$ as a function of $\Delta \phi$, neglecting transient entrainment. Note the shock for low values of $\Delta \phi$ ensuring the $2\pi$-periodicity in $\Delta \phi$. 

Phys. Biol. 13 (2016) 036009 M Beaupeux and P François
symmetric with respect to $\pi$, $2\pi - \Delta \phi$. Thus those networks are not able to discriminate which oscillator is ahead of the other. We call them symmetric. An example of such network can be found on figure 2(C) $(F = 2.66$ bits) and figure 3(A) $(F = 1.93$ bits). Note that both inputs $(S_0$ and $S_1$) are equivalent topologically in the graphical representation of network on figure 3(A).

- A second group is able to discriminate between a phase difference $\Delta \phi$ and $2\pi - \Delta \phi$, therefore is able to discriminate which oscillator is ahead of the other. We call them asymmetric. Examples of such networks can be found on figures 2(A), (B) and 3(D). (Respective fitnesses $F = 1.93, 2.62$ and 2.31 bits.)

Figure 3 displays the simplest networks found (in terms of gene numbers) by our algorithm for each category.

Our discrimination between asymmetric and symmetric is most visible on the reconstructed distribution $P(D, \Delta \phi)$, panels (C) and (F) on figure 3. As said above, because of the periodic behaviour, there are always two values of $\Delta \phi$ corresponding to the same value of $D$. However, the binning of both $\Delta \phi$ and $D$ ‘compresses’ many values of $D$ within a small region of $\Delta \phi$ for asymmetric network, resulting in an apparent discontinuity or ‘shock’ in the relationship between $D$ and $\Delta \phi$. A limit case of asymmetric network would be a perfect shock, where $D$ suddenly changes its value for many values of $\Delta \phi$. This value of $\Delta \phi$ would essentially have very small measure given the binning we impose, so that asymmetric networks essentially realise one-to-one mapping between $D$ and $\Delta \phi$.

In principle, we would expect asymmetric networks to have more information content (up to 1 extra bit) and to systematically dominate symmetric networks; practically $D$ still oscillates a bit for a given value of $\Delta \phi$, which effectively decreases the fitness. This explains why, depending on parameters, some asymmetric networks like on figure 2(A) have lower fitness than some symmetric networks with very damped oscillations like the one on figure 2(C). Nevertheless, the simplest asymmetric network we found contains 0.4 extra bit of information compared to the simplest symmetric network of figure 3. Our simulations tend to show that asymmetric networks are more difficult to evolve than symmetric networks (less than 30% of networks that we obtained can be qualified as asymmetric), so that the evolution algorithm often leads to symmetric networks and optimise them by reducing their residual oscillations.

Thus computational evolution is able to identify working networks, however their dynamics is not trivial, nor the correspondence between topology and behaviour. Clearly, from the examples displayed here on figure 2, there is a relatively broad family of networks able to compute phase difference. One advantage of computational evolution is that it can generate simple efficient networks with coarse-grained interactions, which can help extracting simple working principles (see e.g. [38] for a discussion in developmental context). To get some more general principles for phase computation, that could potentially be recognised in actual networks, we focus on the simplest networks obtained, and since asymmetric networks are closer to an ideal perfect mapping, on network of figure 3(D). Interestingly, the principles and topologies identified in the following can be identified on the more complex networks and thus appear to be potentially generic.

Networks compute phase difference, irrespective of oscillators drift and periods

Our selective pressure imposes values of $D$ as a function of $\Delta \phi$ for randomized input profiles with slowly drifting phases. However, it is important to check that the networks are still able to compute phase differences for different conditions. For instance, it is known that somitogenesis period itself slows down during development, or can be modified by changes of temperature. So we checked that the evolved networks perform the same computation irrespective of the periods of the input oscillators (faster and slower). This is illustrated on figure 4, for period 20% faster and slower: indeed, the PDF of $D$ as a function of $\Delta \phi$ hardly varies for different periods. The only slight difference is observed close to the maximum value of $D$: for shorter period, the remaining oscillation on $D$ has a smaller amplitude and thus $D$ is defined more precisely. This corresponds to an effective better time-averaging: if the time-scale of oscillations is smaller, keeping the same kinetic parameters, we expect the higher frequency to be more efficiently filtered out (because longer time-scale can not ‘follow’ shorter one). It is nevertheless quite remarkable that the phase shift information and shock persist despite this, and indeed fitness is maximum for shorter period.

Methods for phase computation: dissection of networks

As an illustration of the general principles underlying phase computation, we proceed in a very similar way to ‘classical’ genetics by successively removing genes and interactions on network of figure 3(D), and then checking the consequence on the network dynamics.

Positive feedback creates asymmetry

Positive auto regulatory feedback loops are systematically observed in the upstream part of asymmetric networks, on gene $S_2$ and $S_1$ in network of figure 3(D), but also on 2 genes on network of figure 2(A) and 4 genes on network of figure 2(B). On the contrary, in
Asymmetry due to positive feedback can be well understood theoretically. If the combinatorics of activation is akin to an ‘OR’ gate (as we assume here), a gene can activate another self-activating gene, inducing complex history dependency in the dynamics of this gene (e.g. transition from oscillation to bistability in [20]) and, here, asymmetry. This is illustrated on figure 5 using a toy model for the subnetwork made of symmetric networks, there is no such positive feedback loop (figures 2(C) and 3(A)).
genes $S_0$, $S_1$ and $S_2$. In this model, which is fully analytic, Hill functions for $S_0$ and $S_1$ are replaced by step functions (i.e. taking the limit of infinite Hill coefficients). In that case the value of the production rate is then completely determined by the relative concentrations of all activators and repressors with respect to their threshold.

Without this feedback loop (figures 5(B) and (C)), there is no strong difference in the behaviour of $S_1$ for $\Delta \phi$ or $2\pi - \Delta \phi$. $S_1$ activates $S_2$ above threshold for activation (which represents a periodic window $W_3$) and $S_0$ does not repress $S_1$ below threshold for repression (which represents another periodic window $W_0$), so that $S_3$ is periodically active only over the intersection $I = W_0 \cap W_1$. Given the value of thresholds evolved, $W_0$ and $W_1$ are approximately of the same size, so that $S_3$ is fully activated only when $W_0 = W_1$, i.e. for $\Delta \phi \sim \pi$. In this simplified model, where activations and repressions are step functions, the average value of $S_3$ is a simple function of the size of its activation window $I$. If the phase difference $\Delta \phi$ is changing from $\pi$, the size of $I$ decreases, and so does the value of $S_3$. But the size of $I$ is independent whether $S_1$ or $S_0$ are advanced or delayed with respect to each other, so that amplitude of $S_3$ is always symmetrical with respect to $\Delta \phi \sim \pi$.

The situation changes when the positive feedback loop is present (figures 5(D) and (E)). Crucially here, $S_3$ can be on if either $S_1$ or $S_0$ is active. If $I$ occurs at the end of window $W_0$, there is no difference with the previous situation, in the sense that $S_3$ is only activated over the overlap window $I$. In particular, at the end of window $W_0$, $S_3$ is fully repressed by $S_0$ (see figure 5(D)). On the contrary, if $I$ occurs at the end of $W_1$ and at the beginning of $W_0$, once $W_1$ is over, auto activation by $S_3$ can sustain $S_3$ production for a longer time, resulting in a window of activation longer than $I$, until $W_0$ ends (red rectangle on figure 5(E)). As a consequence, there is a higher overall amplitude for $S_3$ and an asymmetry between $\Delta \phi$ and $2\pi - \Delta \phi$.

The importance of this auto activation is also visible if the activator combinatorics is changed from an ‘OR’ logic to an ‘AND’ logic where activators act multiplicatively, just as repressors here. If ‘AND’ logic is used, there can not be any boost of production due to direct self-activation, and asymmetry might only evolve through more complex combinatorics. To check this we ran evolutionary simulations with ‘AND’ logic for activations, and as expected they led to a significantly lower number of asymmetric networks (around 5%). An example of such asymmetric network evolved with the ‘AND’ logic is presented in supplement (figure S1). In this example, positive feedback is effectively replaced by purely feedforward interactions, via one extra gene, leading to similar history dependency and asymmetry in the amplitude of a downstream gene. Functional correspondence between feedback and feedforward interactions has been observed in other contexts [35, 39].

**Layers of feedforward interactions reinforce asymmetry**

Many evolved networks present several feedforward interactions starting with the input genes. The role of specific feedforward interactions is not clear a priori, but in general feedforward loops are used to sense non trivial temporal dynamics (from persistence detection [40] to active sensing of kinetic parameters [35]).

Focusing on network of figure 3(D), asymmetry in $S_3$ gets reinforced downstream by a similar process, where $S_2$ itself auto activates and is in turn repressed by the asymmetric $S_3$. In addition, gene $S_2$ regulation by $S_0$ and $S_1$ is inverted compared to $S_3$; i.e. $S_0$ activates $S_2$ (while it represses $S_3$) while $S_1$ represses $S_2$ (while it activates $S_3$). This creates two interlocked coherent feedforward loops reinforcing the asymmetry effect, and allowing for better phase shift detection.

To see this, we plot on figure 6 the maximum values of $S_2$ and $S_3$ as a function of phase shift, for different simulated ‘mutants’. To focus on the feedforward part of the network, we further remove $S_4$ feedback on $S_2$. Figure 6(A) shows the maximum level of $S_3$ for the normal network: it clearly saturates in an interval around $\Delta \phi = 3\pi/2$, so there would be no good phase shift detection in this region if we were only using $S_3$. If $S_3$ is not repressed by $S_0$, it is sensitive to phase shift only in a limited region, symmetric to the region where $S_3$ detects well the phase shift (figure 6(B)). Gene $S_2$ encoding a good phase detection for the whole phase range only when repression from $S_0$ on $S_3$ is present as finally illustrated on figure 6(C).

**Dampening, amplification and shock refinement are due to negative feedback**

In many networks we observed a clear ‘amplification module’, based on negative feedbacks, where the output decoder gene represses an upstream activator. This is apparent on both networks 3(A) and (D), where respective decoder $S_2$ and $S_3$ are embedded into negative feedbacks with respectively $S_1$ and $S_2$ that appear to amplify them.

Focusing again on the network from figure 3(D), there is a dual role for this module. First, the step of activation of $S_3$ by $S_1$ both dampens oscillations and amplify $S_2$ with respect to $S_2$. This effect is mathematically detailed in the present context in supplement. Second, the negative feedback from $S_3$ on $S_2$ refines the behaviour of the network close to the shock. This is illustrated on figure 6(D), where we have removed the negative interaction from $S_1$ to $S_2$: $S_3$ then clearly displays a more symmetrical profile (black curve), and the shock is less pronounced compared to normal behaviour (red curve). The shock corresponds to a region where maximum production rate of $S_2$ changes rapidly, so that the faster convergence of the oscillating system towards its average value, the more refined the shock is. It is well known that negative feedback indeed helps fast convergence towards steady state in
transcriptional networks [41], and here this effect still exists for slowly varying dynamics of $S_2$ amplitude. More details on this feedback are included in supplement.

**Evolutionary pathway**

As an alternative to the *a posteriori* network study above, it is informative here to study the evolutionary pathway leading to the network of figure 3(D). As we will see, evolutionary transitions correspond almost perfectly to the apparition of functional modules described in the previous section, which validates their adaptive value.

Supplementary video S1 shows the time evolution of the best network (i.e. the one with the highest fitness) at each generation with its corresponding pseudo PDF and fitness (the opposite of the fitness is actually represented, so that a better network corresponds to a lower value on this graph).

- Very quickly in evolution, the output gene is connected to the input genes. Until generation 120, only symmetric networks are observed, with a typical topology that contains three to four genes, with many possible interactions which are essentially neutral (figure 7(A)).
- A transition occurs from generation 120 to generation 121, when the first asymmetric network appears as the best network, with the positive feedback on $S_3$ studied in previous section. Interestingly, as can be seen on supplementary video S1, the qualitative change of PDF at generation 121 is not continuous: there is no smooth change of PDF that leads from the first symmetric networks to the asymmetric network of generation 121 (figure 7(B)).
- The generation 121 sets the core topology of the network: $S_3$ and $S_2$ appear with their interactions that are then adapted. Subsequent parameter and interaction addition leads to better fitness. $S_2$ gets its auto activation at generation 145. $S_4$ is progressively fixed as the output gene, so that the adaptation starting at generation 121 continuously deforms the best network’s PDF to make it span this entire allowed output interval. Figure 7(C) shows the generation 338’s best network and PDF before the next qualitative jump in network and PDF structure occurs.
- The final qualitative jump appears at generation 339 and progressively establishes as the dominant topology. At this generation, a negative feedback appears from $S_4$ to gene $S_2$. First this mutation is neutral, but
quickly, positive selection changes parameters to render it functional and the mutation is then fixed. The initial effect of this negative feedback, shown on figure 7(D) and discussed above, is to reduce the amplitude of the PDF. This negative effect on the fitness is counterbalanced by the positive effect that is the reduction of the noise. Further adaptation contributes to extend the amplitude of the PDF while keeping a low noise level. The best network’s topology is slightly modified but the structure is essentially the same. Further slow evolution of the parameters will lead to the final network displayed on figure 3(D).

Summing up, the sequential evolution of the network itself recapitulates the functional roles described in the previous section, a feature already pointed out in several recent works more focused on dynamics of evolution itself, e.g. [32, 42].

Reconstructing somitogenesis phenomenology using evolved networks

Using computational evolution, we have found gene networks setting a particular gene expression according to the value of the phase difference between two oscillators. Our asymmetric networks are essentially realizing the operation \( D = G(\Delta \phi) \), where \( G \) is an invertible decreasing function on \([0, 2\pi]\) (if we neglect the weak oscillations of the output species \( D \) and the finite width of the shock). In this section, we check that such evolved networks can be indeed embedded into a more general model of somitogenesis, and in particular can help recover the phenomenological properties of mouse somitogenesis observed in [25].

As derived in appendix of [25], from equations (1) and (2), a phenomenological equation describing oscillations of individual cell is

\[
\Delta \phi_i(x, t) = \alpha \Delta \phi_1(x, t)
\]

with \( \Delta \phi_i(x, t) = \phi_i(x, t) - \phi_{i_0}(0, t) \). Figure 3(F) shows that \( D \) itself is a linear function of \( \Delta \phi = \phi_1 - \phi_0 \). As \( \Delta \phi_1 \) and \( \Delta \phi \) only differ by a constant (see supplement for details), it is natural within our framework to replace \( \Delta \phi \) by some linear function of \( D \), effectively implementing feedback of oscillators on themselves. Biologically, this could be done if \( D \) modifies the angular velocity of the Notch oscillator.

Computational evolution also predicts the existence of a shock in \( D \) when the two oscillators are in phase. So it is tantalizing to assume that our predicted shock might be detected by cells to trigger genetic expression giving rise to differentiation into somites. This would biologically correspond, for instance, to the expression of gene \( Mesp2 \) [43, 44].

Such process can be easily modelled: the shock observed for \( \Delta \phi^* = 0 \) in our simulation can be used as an input of an adaptive network [45], the output being \( Mesp2 \). The rationale is that an adaptive network would not detect the slow increase of \( D \)’s average value, but would be very sensitive to the shock itself when oscillators are in phase. Then we finally assume than once \( Mesp2 \) has been activated, some process turns on after some short delay to switch off oscillators. Details and a MATLAB implementation of this mechanism are given in the supplement.
Figure 8 recapitulates this model, including feedback of $D$ on oscillator $\phi_f$, and adaptive reading of the shock, as well as its behaviour. Figure 8(D) presents a simulation of the experiments in figure 4 from [25]. The two oscillators are initialised close to phase opposition, with a spatially linear phase gradient (and corresponding gradient of $D$ measuring the phase difference). Those initial conditions mimic a process where the cells are initially similar to tail bud but slightly phase shifted, close to their stationary limit cycle, and, most importantly recapitulate the observations from [25]. At time $t = 0$ in this simulation, we turn on the feedback by $D$ and then cells start to slow down, except for the posterior bottom cells, which are assumed to oscillate forever at same frequency for the two oscillators. This simulates the tail bud-like cells similarly to simulations/assumptions from [25], and more recently described in [46].

Quite strikingly, even though the relation between $D$ and $\Delta \phi$ is not perfectly linear, and despite the small remaining oscillations of $D$ itself, this model reproduces extremely well typical kymographs of the scaling experiments from [25]. In particular, we check on figures 8(B) and (C) that the speed of propagating wave of oscillation of $S_1$ is an exponential function of time, as observed experimentally. Finally, there is a clear band of simulated $Mesp2$ gene (blue region on figure 8(D)). Two aspects are interesting: first, the width of the simulated $Mesp2$ stripes scale with the field of oscillations, which is consistent with the overall scaling of the experimental system (including somite) with PSM size. Second, the simulated band of $Mesp2$ can be visually separated in almost discrete blocks, suggestive of discrete somites. This block structure indicates that the moment at which $Mesp2$ is expressed in this model does not purely depend on the phase difference $\Delta \phi$.

This is further illustrated on figure 9. Starting with different initial phases at $t = 0$ for $S_0$, we integrate the network equations, imposing as before that $S_1$, initially out of phase ($\phi_f - \phi_p = \pi$), slows down. After some time, the phase difference between the two oscillators reaches the critical value $\Delta \phi^* = 0$, they are in phase (corresponding phase of $S_0$ is called $\phi_p$). In a realistic situation, the moment when the decoder gene drops, thus triggering $Mesp2$ expression and defining the front, occurs shortly after this. In figure 9(A), we define $\Delta t$ as the corresponding time difference between front definition and $\phi_p$, and $\phi_f$ as the corresponding value of the phase of reference oscillator at the front. We possibly have different values for $\Delta t$ and $\phi_f$ depending on initial phase of $S_0$. Two extreme cases could arise, illustrated on figures 9(B) and (C):

1. The front is a pure function of the phase of the reference oscillator, i.e. $\phi_f$ coincides with the time when phase $\phi_0$ reaches a specific value. In that case $\Delta t$ decreases linearly with $\phi_p$. This situation is represented on figure 9(B), and the corresponding kymograph would exhibit discrete segments corresponding to phases $\phi_0$ modulo $2\pi$.

2. The front is a pure function of the phase difference, i.e. appears for a constant $\Delta t$. In that case $\phi_f$ is linear in $\phi_p$. This is represented on figure 9(C). In this case, the kymograph would exhibit a smooth expression pattern for $Mesp2$, with no visible segments.

As illustrated on figure 9(D), our network is closer to situation of figure 9(C), which is not a surprise since
decoder gene has been selected to be a function of $\Delta \phi$. Nevertheless, there is a non negligible oscillation of the phase difference on figure 9(D) that modulates the front signal as a function of $\phi_0$. Because of this dependency on the absolute phase of oscillators, the system displays more discrete blocks of expression like in figure 8(D). Thus a decoder variable sensing phase difference of two oscillators can naturally implement two contradictory aspects: a continuously regressing shock combined with a mechanism for discrete somite formations. Further rostro-caudal patterning of somites would necessitate another mechanism downstream of the clock, e.g. the slowing down of the clock can provide somite polarity [47].

Discussion

Recent experiments on somitogenesis have challenged the classical CW paradigm [25]. The behaviour of Lfng waves in mouse explants [25] suggests a phenomenological model where two genetic oscillators might be crucial for positional information leading to vertebrae formation. In this manuscript, we have reformulated this idea using the concept of mutual information, and used in silico evolution to generate nonlinear gene networks encoding phase difference with a decoder morphogen. Our evolutionary computation gives clear reproducible network structures, which raises new experimental questions.

The first natural question is the nature of the oscillators implicated. There are many candidates given the plethora of oscillating genes [10]. The oscillatory models described here should be interpreted as coarse-grained descriptions, with one phase variable possibly describing oscillations of complete pathways. From [25], the natural candidate for phase $\phi_0$ is the Notch pathway, with Lfng as a natural representative. A prediction from the 2-oscillator model from [25] is that reference oscillator $\phi_0$ should be synchronized in the whole PSM, so that there would be no clear wave of genetic expression like what is observed for Lfng. Rather the whole PSM would simply turn ‘on/off’ in response to this oscillator. Dynamical reporters would be necessary to identify such first oscillator. Microarrays have clustered two groups of oscillating genes: Notch/FGF pathways and Wnt pathways [10]. So Wnt might be a priori a good candidate for $\phi_0$ [13], but only detailed spatial and temporal resolution of its oscillations would answer this question. One can not exclude for instance that other components of Notch or FGF, while in phase in tail-bud, present different patterns than Lfng as PSM cells move more anteriorly.

Our study identifies the important combinatorics of positive and negative feedbacks necessary to compute
phase difference. An upstream positive feedback creates asymmetry in the decoder profile, while downstream negative feedback plays a role in both dampening of the oscillations and amplifying the final signal. However, as can be seen from the variety of networks evolved in figure 2, those feedbacks can be potentially implemented in many different ways. Nevertheless, these properties are rather generic, and therefore constitute experimental predictions on coarse-grained network organisation.

The feedbacks inside the somitogenesis network appear themselves rather complex and convoluted (see e.g. the multiple delta feedbacks in zebrafish [48]), and it is clear that there are many redundancies in the system since one can knock out entire pathways and still observe oscillations [10]. Phase/amplitude information would be important to disentangle the coarse-grained feedbacks predicted here, but to date there are only few reporters allowing for such fine measurements: Notch and Wnt in mouse [13, 25], and Hes reporters in fish [14, 49]. Nevertheless, screening of oscillating pathways have identified oscillating genes common to Notch and Wnt such as Bcl9L and Nkd1 [10], that could thus be part of the feedforward pathways downstream of \( \phi_h \) and \( \phi_f \).

The decoder gene \( D \) in our model couples the two oscillators and helps defining the differentiation front. FGF pathway controls somite formation [50], and thus is expected to be related to \( D \) in our framework. Interestingly, many members of FGF pathways oscillate as well [10]. Functional roles of FGF pathway oscillations are not clear in a picture where FGF gradient sets a threshold of differentiation [51], but on the contrary, they are expected if oscillations control somite formation like in the two-oscillator framework assumed here. Interestingly, FGF indeed regulates Hes7 oscillations [32] and the whole FGF cluster appears in phase in Notch [10], thus suggesting a feedback between differentiation and oscillation itself.

It should also be pointed out that behaviours qualitatively similar to decoder \( D \) here are actually observed experimentally on reporters of core oscillating genes: there is a significant increase of both average level and oscillation amplitude, before a complete collapse in expression at the front (that would correspond to the shock). Examples include Lfng in mouse [25], or Hes in zebrafish [14]. So such genes, while belonging to the core Notch signalling pathway, might actually be part of the actual decoder, feeding back on the core oscillator. If so, their average level (or their signalling intensity) should directly control their period, which might be testable experimentally.

A last phenotypic prediction of our approach is the existence of a shock. This shock appears as a ‘phenotypic spandrel’ (i.e. a phenotype appearing as a consequence of another phenotype [53]) related to decoding of phase differences by a linear variable. We showed on a simple toy-model how such shock could actually define what is usually considered as the ‘wavefront’, upstream of genes such as Mesp2, that can nevertheless display more discrete patterns of expression corresponding to somites. While the idea of a shock emerging from phase equation has been recently explored in a model bearing resemblance to the Burger’s equation [54], the shock there is due to cell-to-cell coupling, while we explore here a different framework based on decoding the phase difference between two oscillators that is purely local and can be implemented in a single cell.

A major and surprising evolutionary insight from comparison of somitogenesis networks is that homolog genes might have different roles in different species [11]. The principles described here are general enough to be shared by many oscillatory pathways during somitogenesis: phase computation can be performed by combining small positive and negative feedback loops like in the asymmetric network from figure 3(D). The associated shock in genetic expression could be shared by many components and be used to define differentiation front, giving evolutionary plasticity. This suggests again that focusing on ‘phenotypic’ description of developmental networks might be key to functional understanding, as already suggested in [38].

### Acknowledgments

We thank Alexander Aulehla for useful comments, discussions, and an anonymous referee for suggesting to explore more explicitly the role of feedforward loops. This project has been partially funded by the Natural Science and Engineering Research Council of Canada (NSERC), Discovery Grant Program, McGill University Physics Department and a Simons Investigator Award for Mathematical Modelling of Biological Systems to PF.

### References

[1] Sorre B, Warmflash A, Brivanlou A H and Siggia E D 2014 Encoding of temporal signals by the TGF-β pathway and implications for embryonic patterning Dev. Cell 20 334–42

[2] Lin Y, Sohn C H, Dalal C K, Cai I and Elowitz M B 2015 Combinatorial gene regulation by modulation of relative pulse timing Nature 527 54–8

[3] Wolpert L 2006 Principles of Development (Oxford : Oxford University Press)

[4] Nelson C E, Morgan B A, Burke A C, Laufer E, DiMambro E, Murtaugh L G, Gonzales E, Tessarollo L, Parada L F and Tabin C J 1996 Analysis of Hox gene expression in the chick limb bud Development 122 1449–66

[5] Wacker S A, Jansen H J, McNulty C L, Houtzager E and Durston A J 2004 Timed interactions between the hox expressing non-organiser mesoderm and the spemann organiser generate positional information during vertebrate gastrulation Dev. Biol. 268 207–19

[6] Iimura T and Pourquié O 2006 Collinear activation of Hoxb genes during gastrulation is linked to mesoderm cell ingestion Nature 442 566–71

[7] Balaskas N, Ribeiro A, Panovska J, Dessaud E, Sasaki N, Page K M, Briscoe J and Ribes V 2012 Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube Cell 148 273–84

[8] Tufcea D E and François P 2015 Critical timing without a timer for embryonic development Biophys. J. 109 1724–34
[9] Raspopovic I, Marcon L, Russo L and Sharpe J I 2014 Modeling digits, digit patterning is controlled by a Bmp-Sox9-Wnt turing network modulated by morphogen gradients Science 345 566–70

[10] Dequante M-L, Glynn E, Gaudenzi K, Wahl M, Chen J, Mushegian A and Pourquié O 2006 A complex oscillating network of signaling genes underlies the mouse segmentation clock Science 314 1595–8

[11] Krol AJ, Roellig D, Dequante M-L, Tassy O, Glynn E, Hattem G, Mushegian A, Oates A C and Pourquie O 2011 Evolutionary plasticity of segmentation clock networks Development 138 2783–92

[12] Palmettlin I, Henrique D, Ish-Horowicz D and Pourquié O 1997 Avian hairy gene expression identifies a molecular clock linked to vertebrate segmentation and somitogenesis Cell 91 639–48

[13] Aulehla A, Wiegabrae W, Baubet V, Wahl M B, Deng C, Taketo M, Lewandoski M and Pourquié O 2008 A beta-catenin gradient links the clock and wavefront systems in mouse embryo segmentation Nat. Cell. Biol. 10 186–93

[14] Deltaue E A, François P, Shih N P and Amacher S L 2012 Collective modes of coupled phase oscillators with delayed coupling Phys. Rev. Lett. 104 208101

[15] Giudicelli F, Ozbudak E M, Wright G J and Lewis J 2007 Setting the tempo in development: an investigation of the zebrafish somite clock mechanism PLoS Biol. 5 e150

[16] Baker R E, Schnell S and Maini P K 2006 A mathematical investigation of a clock and wavefront model for somitogenesis J. Math. Biol. 52 438–82

[17] Goldbeter A, Gonzalez D and Pourquié O 2007 Sharp developmental thresholds defined through bistability by antagonistic gradients of retinoic acid and FGF signaling Developmental Dyn.: Official Publ. Am. Assoc. Anatomists 236 1495–508

[18] Francois P, Hakim V and Sigga E D 2007 Deriving structure from evolution: metazoan segmentation Mol. Syst. Biol. 3 9

[19] Gomez C, Ozbudak E M, Wunderlich J, Baumann D, Lewis J and Pourquié O 2008 Control of segment number in vertebrate embryos Nature 454 335–9

[20] El-Sherrif E, Averof M and Brown S J 2012 A segmentation clock operating in blastoderm and germ band stages of the zebrafish somite clock mechanism PLoS Biol. 5 e1000310

[21] Sarrazin A F, Peal A D and Averof M 2012 A segmentation clock with two-segment periodicity in insects Science 336 338–41

[22] Rosenberg M I, Brent A E, Payre F, Desplan C and Pan D 2014 Dual mode of embryonic development is highlighted by expression and function of nasionia pair-rule genes eLife 3 1120

[23] Lauschke V M, Tsiairis C D, François P and Aulehla A 2013 Scaling of embryonic patterning based on phase-gradient encoding Nature 493 101–5

[24] Goodwin B C and Cohen M H 1969 A phase-shift model for the spatial and temporal organization of developing systems J. Theor. Biol. 25 49–107

[25] Kuramoto Y 1984 Chemical Oscillations, Waves, and Turbulence (Berlin: Springer)

[26] Francois P and Hakim V 2004 Design of genetic networks with specified functions by evolution in silico Proc. Natl. Acad. Sci. USA 101 580–5

[27] Fujimoto K, Ishihara S and Kaneko K 2008 Network evolution of body plans PLoS One 3 e2772

[28] ten Tusscher K H and Hogeweg P 2011 Evolution of networks for body plan patterning: interplay of modularity, robustness and evolvability PLoS Comput. Biol. 7 e1002208

[29] Feng S, Ollivier J F, Swain P S and Soyer O S 2015 BioLazz: in silico evolution of cellular networks with unbounded complexity using rule-based modeling Nucleic Acids Res. 43 e123–123

[30] Francois P 2014 Evolving phenotypic networks in silico Sem. Cell Developmental Biol. 35 90–7

[31] Lynch M 2007 The Origins of Genome Architecture (Sunderland, MA: Sinauer Associates)

[32] Francois P and Sigga E D 2010 Predicting embryonic patterning using mutual entropy fitness and in silico evolution Development 137 2385–95

[33] Lalanne J-B and Francois P 2013 Principles of adaptive sorting revealed by in silico evolution Phys. Rev. Lett. 110 218102

[34] Gregor T, Tank D W, Wieschaus E F and Bialek W 2007 Probing the limits to positional information Cell 130 153–64

[35] Dubuis J O, Tkacik G, Wieschaus E F, Gregor T and Bialek W 2013 Positional information, in bits Proc. Natl. Acad. Sci. 110 16301–8

[36] Francois P and Sigga E D 2012 Phenotypic models of evolution and development; geometry as destiny Curr. Opin. Genetics Dev. 22 627–33

[37] Ma W, Trusina A, El-Samad H, Lim W A and Tang C 2009 Defining network topologies that can achieve biochemical adaptation Cell 138 760–73

[38] Mangam S and Alon U 2003 Structure and function of the feed-forward loop network motif Proc. Natl. Acad. Sci. USA 100 11901–5

[39] Rosenfeld N, Elowitz M B and Alon U 2002 Negative autoregulation speeds the response times of transcription networks J. Mol. Biol. 323 785–93

[40] Kohosakabe T and Kaneko K 2016 Evolution–development congruence in pattern formation dynamics: bifurcations in gene expression and regulation of networks structures J. Exp. Zoology B 326 61–84

[41] Koski H, Saga Y, Takahashi Y, Koizumi K, Takagi A, Kitajima S and Inoue T 2000 Mesp2 initiates somite segmentation through the Notch signalling pathway Nat. Genet. 25 390–6

[42] Saga Y and Takeda H 2001 The making of the somite: molecular events in vertebrate segmentation Nat. Rev. Genet. 2 835–45

[43] Francois P and Sigga E D 2008 A case study of evolutionary computation of biochemical adaptation Phys. Biol. 5 26009

[44] Tsiairis C D and Aulehla A 2016 Self-organization of embryonic genetic oscillators into spatiotemporal wave patterns Cell 164 656–67

[45] Shih N P, Francois P, Deltaue E A and Amacher S L 2015 Dynamics of the slowing segmentation clock reveal alternating two-segment periodicity Development 142 1785–93

[46] Schwendinger-Schreck J, Kang Y and Holley S A 2014 Model of development of the zebrafish segmentation clock’s gene regulatory network constrained by expression data suggests evolutionary transitions between oscillating and nonoscillating transcription Genetics 197 725–38

[47] Webb A B, Lengyel I M, Jörg D J, Valentin G, Jülicher F F, Morelli L G, Oates A C and Whitfield T T 2016 Persistence, period and precision of autonomous cellular oscillators from the zebrafish segmentation clock eLife 5 e08438

[48] Dubrulle J, McGrew M J and Pourquié O 2001 FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation Cell 106 219–32

[49] Dubrulle J and Pourquié O 2004 fgf8 mrNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo Nature 427 419–22

[50] Niwa Y, Masamizu Y, Liu T, Nakayama R, Deng C-X and Kageyama R 2007 The initiation and propagation of Hes7 antagonistic gradients of retinoic acid and FGF signaling in the somite segmentation clock eLife 5 e08438

[51] François P, Johnson K A and Saunders L N 2015 Phenotypic spandrels: absolute discrimination and ligand antagonism arXiv.org

[52] Murray P J, Maini P K and Baker E 2011 The clock and wavefront model revisited J. Theor. Biol. 283 227–38