Multiscale Modelling and Analysis of Planar Cell Polarity in the Drosophila Wing

Qian Gao, David Gilbert, Monika Heiner, Fei Liu, Daniele Maccagnola, David Tree

Abstract—Modelling across multiple scales is a current challenge in Systems Biology, especially when applied to multicellular organisms. In this paper we present an approach to model at different spatial scales, using the new concept of hierarchically coloured Petri Nets (HCPN). We apply HCPN to model a tissue comprising multiple cells hexagonally packed in a honeycomb formation in order to describe the phenomenon of Planar Cell Polarity (PCP) signalling in Drosophila wing. We have constructed a family of related models, permitting different hypotheses to be explored regarding the mechanisms underlying PCP. In addition our models include the effect of well-studied genetic mutations. We have applied a set of analytical techniques including clustering and model checking over time series of primary and secondary data. Our models support the interpretation of biological observations reported in the literature.

Index Terms—Hierarchically coloured Petri nets, Continuous Petri nets, ordinary differential equations, multiscale modelling, continuous simulation, cluster analysis, model checking, planar cell polarity.

1 Introduction

With the rapid growth of data being generated in the biological field, it has become necessary to organise the data into coherent models that describe system behaviour, which are subsequently used for simulation, analysis or prediction. Modelling biological systems beyond one spatial scale introduces a series of challenges which should be addressed. These include:

1) *Repetition of components* – e.g. the need to describe multiple cells each of which has a similar definition.
2) *Variation of components* – sets of similar components with defined variations, e.g. mutants.
3) *Organisation of components* – e.g. how cells are organised into regular or irregular patterns over spatial networks in one, two or three dimensions.
4) *Hierarchical organisation* – enabling the description of (possibly repeated) components which contain repeated sub-components. e.g., cells containing several compartments. This feature enables the use of abstraction regarding the level of detail used to describe components.
5) *Communication between components* – in general communication is constrained to occur between immediate neighbours, but this may be further constrained according to the relationship between neighbours, and the position of a component within a spatial network.
6) *Mobility* – e.g. transport of components within a system, or actively motile cells.
7) *Replication of components* – e.g. cell division.
8) *Deletion of components* – e.g. cell death.

In this paper we have chosen the biological example of Planar Cell Polarity (PCP) signalling in the Drosophila wing, which illustrates several of these issues. The epithelial cells in this organ are hexagonally packed in a 2-dimensional honeycomb lattice. Signal transduction within each cell is coupled with inter-cellular communication through the formation of protein complexes, so that local (transmembrane) signalling produces a global effect over the entire organ. Our model of PCP includes the repetition of components in a two-level hierarchy of different geometries, permitting abstraction at the level of cells, with a static two-dimensional organisation which is different at each level. The higher inter-cellular level is that of cells in a rectangular honeycomb grid, representing the epithelium tissue, and the lower level being intra-cellular organisation represented by virtual compartments within one cell in a circular grid. Moreover we model variations among cells in the form of patches of mutant cells which lack a specific signalling protein.

A large variety of modelling approaches have already been applied to model a wide array of biological systems (see [16] for a review). Among them, Petri nets are particularly suitable for representing and modelling the concurrent and asynchronous behaviour of biological systems. However, standard Petri nets along with many other modelling approaches do not readily scale to meet the challenges

- *Q. Gao and D. Gilbert: School of Information Systems, Computing and Mathematics; D. Tree: School of Health Sciences and Social Care, Brunel University, UK.*
  E-mail: firstname.surname@brunel.ac.uk
- *M. Heiner and F. Liu: Computer Science Department, Brandenburg University of Technology Cottbus, Germany.*
  E-mail: firstname.surname@informatik.tu-cottbus.de
- *D. Maccagnola:Dipartimento di Informatica Sistemistica e Comunicazione, Università degli Studi di Milano-Bicocca, Italy.*

described above, and current attempts to simulate biological systems by standard Petri nets have been mainly restricted so far to relatively small models.

There are two orthogonal concepts which can be used to structure Petri net models. (i) Hierarchical Petri Nets reuse the well-established engineering principle of hierarchical decomposition to manage the design of large-scale systems. Sub-networks are hidden as building blocks within macro nodes. (ii) Coloured Petri Nets (CPN) overcome the constraints of standard Petri nets by allowing the modelling of large-scale systems in a compact, parameterised and scalable way.

The main contribution of our paper is the development of a general approach for modelling multiscale systems which combines the above two concepts, Hierarchically Coloured Petri nets (HCPN). In this study we illustrate the power of our approach for describing spatial multiscale problems in biological systems by a complex and challenging case study, which requires computational experiments over very large underlying models which are represented by systems of Ordinary Differential Equations (ODEs).

This paper is based on our earlier work published in CMSB 2011 [12], extended and improved by

- the use of a circular topology at the intracellular level, permitting the exploration of different intercellular communication schemes at the higher level in a more biologically faithful manner,
- the use of a variety of biasing mechanisms to induce polarity (feedback loops, external ligands, microtubule transport),
- the introduction of a variety of genetic mutations which can be placed in different clone shapes (rectangular, circular, elliptical),
- the development of a highly efficient computational technology exploiting constraint solving which permits the construction and exploration of models of much larger wing tissues than was previously possible,
- the development and application of automated clustering and visualisation techniques permitting improved analysis of the simulation results, complemented by PLTL model checking of primary and secondary data to explore the behaviour of the automatically identified clusters.

This paper is structured as follows: in Section 2 we introduce the biological background of planar cell polarity, followed by Section 3 on related work. Section 4 briefly describes coloured Petri Nets and in Section 5 we present our modelling approach by means of the PCP case study, followed by Section 6 on the analysis of our PCP model, and finally the conclusion.

2 Planar Cell Polarity

Planar cell polarity (PCP) refers to the orientation of cells within the plane of the epithelium, orthogonal

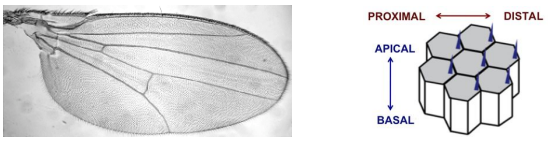


Fig. 1. Drosophila: whole wing (left); scheme of hexagonal cells with hairs (right).

to the apical-basal polarity of the cells. This polarisation is required for many developmental events in both vertebrates and non-vertebrates. Defects in PCP in vertebrates are responsible for developmental abnormalities in multiple tissues including the neural tube, the kidney and the inner ear (reviewed in [38]). The signalling mechanisms underlying PCP have been studied most extensively in the epithelia of the fruit fly *Drosophila melanogaster* including the wing, the abdomen, the eye, and the bristles of the thorax.

The adult Drosophila wing comprises about 30,000 hexagonal cells each of which contains a single hair pointing in an invariant distal direction, see Fig. 1. This hair comprises actin bundles and is extruded from the membrane at the distal edge of the cell during pupal development, at the conclusion of PCP signalling. Preceding this ultimate manifestation of PCP, signalling occurs such that the proteins adopt an asymmetric localisation within each cell. At the initiation of PCP signalling Fmi, Fz, Dsh, Vang and Pk are all present symmetrically at the cell membrane. At the conclusion of PCP signalling Fmi is found at both the proximal and distal cell membrane, Fz and Dsh are found exclusively at the distal cell membrane and Vang and Pk are found exclusively at the proximal cell membrane. These PCP proteins are thus thought to mediate the cell-cell communication that comprises PCP signalling and be involved in establishing the molecular asymmetry within and between cells which is subsequently transformed into the polarisation of the wing hairs (reviewed in [40]). The result is a polarisation of individual cells and local alignment of polarity between neighbouring cells. A proposed feedback loop mechanism was proposed by Tree et al. [41] which is believed to mediate a competition between proximal and distal proteins between adjacent surfaces of neighbouring cells. More details of the current views of the details of signalling can be found in [3]. Thus we will use Flamingo-Frizzled-Dishevelled complexes (FFD) as a proxy for hairs in our research.

Various studies postulated the existence of an unknown and as yet un-identified secreted morphogen signal to which the PCP proteins respond. Despite the lack of positive data identifying such a morphogen signal [32] such models remain popular in the literature. A second postulated upstream signal biasing PCP is associated with the protocaderhin molecules Fat and Dachsous (Ds) whose activities are modulated

by the golgi kinase Four-jointed (reviewed in [5], [46]). It has thus been proposed that the function of the Fat and Ds proteins is to facilitate the transport of the distal PCP proteins to the distal cell cortex where they are stabilised by interactions with proximal complexes.

In this paper, we use HCPN to model PCP signalling in a generic setting that encompasses a broad class of specific models, ranging from a single cell model to a multi-cellular model. We also provide alternative ways to build different neighbourhood relationships between adjacent cells. To this end, we have developed a set of models for the generation of PCP to investigate diverse mechanisms which have been proposed to underly PCP signalling by simulating and analysing the dynamic behaviour of the PCP proteins and complexes.

Initially we wish to recapitulate the phenotypes of all known mutant conditions, both loss and gain of function. We hope to be able to address questions about the mechanisms underlying the polarising signal, the dynamics of signalling by the individual components and the signals downstream of the PCP proteins which orchestrate the ultimate morphological manifestation of planar polarity. Ultimately we hope our model will make predictions about the mechanisms underlying the PCP signalling process which will be testable in a biological laboratory.

3 Related Work

Modelling PCP. Several mathematical and computational approaches have been applied to study PCP signalling [1], [7], [23], [24], [33], [35]. However, these models lack an approachable way to generate the cell geometry or a grid of hexagonal cells and are hard to reproduce by other researchers. This motivated us to develop an approach which permits the systematic construction of large scale mathematically tractable models in which cell geometry is clearly formalised. Compared with ODEs and PDEs, hierarchically coloured Petri nets are more intuitive for users with limited knowledge about modelling. In addition, we include new aspects of signalling, such as biased transport of molecules [3], [37], which were not considered in previous models but have been suggested as mechanisms which might influence the regulation of PCP signalling [35]. Furthermore, we have developed advanced analysis methods and visualisation to compare simulation results with experimental data (images).

Coloured Petri nets in Systems Biology. There are a few applications that take advantage of the power and ease of modelling offered by CPN, e.g. [4], [14], [25], [45], for more details see [12], [26]. These existing studies are rather small and usually resort to `Design/CPN` [8] or its successor `CPN Tools` [20], which, however, were not specifically designed with the requirements of Systems Biology in mind, and do not support stochastic or continuous modelling and simulation. In our work, we use colour not only to express repetition, but also to encode (spatial) locality. We will show that the standard colouring concept can be further enhanced by hierarchically organised colours, which are able to directly reflect the hierarchical organisation of the objects modelled. We will demonstrate our approach by a case study which yields a design pattern for similar modelling problems.

4 Modelling Language

4.1 Coloured Petri nets

Coloured Petri nets (CPNs) [18], [19] are an established discrete event modelling formalism combining the strengths of Petri nets with the expressive power of programming languages. Petri nets provide a sound graphical notation for modelling systems with concurrency, communication and synchronisation. Programming languages offer the constructions of data types, which permit the creation of compact and parameterisable Petri net models. This is the most important advantage of CPN which we are going to exploit in this paper.

Syntax. CPNs comprise – as do standard Petri nets – places, transitions, and arcs. In systems biology, places typically represent species (chemical compounds), while transitions represent any kind of chemical reactions or transport steps [17]. In this paper places represent PCP proteins whereas transitions represent physical interaction and/or signalling events between proteins and polarised transport of proteins. Additionally, a CPN model is characterised by a set of colour sets (i.e. discrete data types). Each place gets assigned one colour set and may contain distinguishable tokens coloured with a colour of this colour set. As there can be several tokens of the same colour at a given place, the tokens at a place define a multiset over the place's colour set. A specific distribution of coloured tokens at all places constitutes a marking of a CPN. Each arc is assigned an expression; the result type of this expression is a multiset over the colour set of the connected place. Each transition has a guard, which is a Boolean expression, typically over variables occurring in the expressions of adjacent arcs. The guard must be evaluated to true for the enabling of the transition. The trivial guard 'true' is usually not explicitly given.

Behaviour. The variables associated with a transition consist of the variables in the guard of the transition and in the expressions of adjacent arcs. Before the expressions are evaluated, the variables must be assigned values with suitable data types, which is called binding [19]. A binding of a transition corresponds to a transition instance in the underlying unfolded net.

Enabling and firing of a transition instance are based on the evaluation of its guard and arc expressions. If the guard is evaluated to true and the preplaces have sufficient and appropriately coloured tokens, then the transition instance is enabled and may fire. When a transition instance fires, it removes coloured tokens from its preplaces and adds coloured tokens to its postplaces, i.e. it changes the current marking to a new reachable one. The colours of the tokens that are removed from preplaces and added to postplaces are decided by the arc expressions. The set of markings reachable from the initial marking constitutes the state space of a given net. These reachable markings and transition instances causing the marking change constitute the possibly infinite reachability graph (state transition system) of the coloured net.

Folding and unfolding. Coloured Petri nets with finite colour sets can be automatically unfolded into uncoloured Petri nets, which then allows the application of all the powerful standard Petri net analysis techniques. Vice versa, uncoloured Petri nets can be folded into coloured Petri nets, if partitions of the place and transition sets are given. These partitions of the uncoloured Petri net define the colour sets of the coloured Petri net. However, the algorithmic identification of suitable partitions is an open research issue. The conversion between uncoloured and coloured Petri nets changes the style of representation, but does not change the structure of the underlying reaction network, see Fig. 2 for some introductory examples and Section 5 for more details.

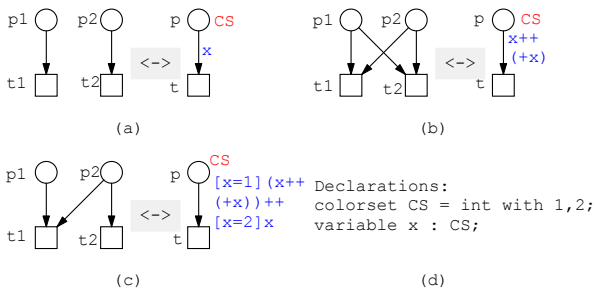


Fig. 2. Three typical examples for folding-unfolding; notation: ++ multiset addition, (+x) successor, [x=2] guard.

The key challenge when unfolding coloured Petri nets is to compute all transition instances. Their computation can be considered as a combinatorial problem, which suffers from combinatorial explosion. However, when the number of transition instances is only decided by guards, which are in fact logical expressions, a constraint satisfaction approach [6], [42] can be employed. We have used the efficient search strategies of Gecode [13] to greatly improve the unfolding efficiency of coloured Petri nets, reported in more detail in [26]; see Table 2 for results.

Extensions. We allow special arcs, e.g. read arcs or inhibitor arcs in our coloured Petri nets. If transitions are additionally associated with stochastic (deterministic) firing rates, as discussed in [15], [17], we will obtain coloured stochastic (continuous) Petri nets. The rate functions defining the usually state-dependent rates can be specified for coloured transitions, or individually for each transition instance; for more details see [27], [28].

4.2 Hierarchically coloured Petri nets

CPNs enable the modelling of (biological) systems comprising repeated components, each of which is associated with a particular colour.

Moreover, colour can be used to encode spatial locality. For example, to model cells in a 2-dimensional lattice, we can represent one cell as a CPN with colour sets denoting the number of copies (cells), and use functions to describe the connectivity between the cells. A colour is a 2-tuple which can then be read as an address identifying locality in space. This can be easily extended to higher dimensions. Moreover, the model is adjustable to different lattice sizes by just changing some constants.

Typically colour sets are not read as unordered sets, but are assumed to be totally ordered, with an implicit successor relation. We generalise this idea, by permitting a partial order relation over a set. One instance of this would result in a hierarchy, i.e. a tree-like structure. Combining the base colour set with the partial order relation, we obtain a hierarchical colour set, and represent each node in the tree by the path describing the branch from the root to the node, which enables direct referencing of the position of the node in the hierarchy, and thus permitting the formulation of operations over hierarchy levels.

We can define operations over the hierarchically defined colours which correspond to navigation over the hierarchy tree. Operations which move up/down one node are a generalisation of the predecessor/successor function over a totally ordered set. However, we can also define operations which move over more than one node at a time or jump to nodes in different branches in the tree.

The concept of hierarchy is orthogonal to, and can be freely combined with other colour set constructors such as product or union.

Specifically, we adopt the notation that colour sets are described by one, two or three-tuples, corresponding to the number of spatial dimensions being modelled, and note that the underlying colour set is given by the Cartesian product expansion of the colour set tuple. Thus, for example, when modelling cells in, for instance, a rectangular $M \times N$ grid, each cell is associated with a colour (x, y) where $x \in \{1..M\}, y \in \{1..N\}$. A colour set can be associated with a set of constraints which effectively describe the

topology of the one, two or three-dimensional grid used to model the layout of the components. Thus we may embed a component in a honeycomb (hexagonal) grid by imposing the requisite constraints over an underlying rectangular grid. Guards over transitions permit the description of the patterns of connection allowed between cells.

Furthermore if we want to describe regular organisation within a cell, we can extend this concept by having a grid at the intra-cellular level and another set of colours to indicate the position inside the inner grid. A separate set of arc functions at the intra-cellular level indicates the inter-component communication at this level. Consequently, we get a sequence of tuples, each tuple referring to the address within a certain level. We call this concept hierarchically structured colours, and the corresponding net class Hierarchically Coloured Petri Nets (HCPN).

In order to support the description of hierarchically organised systems, we extend the notation for colour sets as follows. The colour set of the highest level of an L hierarchy is denoted by a tuple T_L, that of the next level by a tuple T_{L-1} and the lowest level by a tuple T_1. When referring to the colour set of a level, we will give its position in the hierarchical tree of colour sets by prefixing the colour sets above it. Thus for example, the colour set for the $L-1$ level is given as $T_L \cdot T_{L-1}$, and the colour set for the lowest level as $T_L \cdot T_{L-1} \cdot \ldots \cdot T_1$. The number of colours in the underlying colour set is given by the product of the number of underlying colours in the colour set tuple from each level.

In order to further facilitate modelling, we can denote each level by a descriptive label, thus the levels in a three level hierarchy could be denoted by *wing, cell, subcellular-location* and the entire colour set by $wing_{cs} \cdot cell_{cs} \cdot subcellular\text{-}location_{cs}$. We further expand our notation for the guards so that they are associated with the level at which they operate. See Section 5 for more details on how a HCPN model can be constructed.

In summary, hierarchically structured colours are useful to express repetition and (spatial) locality as we will demonstrate in our case study. Unlike [43], we do not use a concept of nets-within-nets to describe hierarchy, but colours-within-colours. We use two two-tuples (nested two dimensions) to address hierarchical systems.

5 Modelling Approach

Overview. We start off by modelling each cell as a (standard) Petri net, initially at a highly abstract level, neglecting details of the intra-cellular pathways, and divide the cell into several virtual compartments to facilitate the description of PCP. Because Drosophila wing cells form a regular honeycomb lattice, we impose this organisation at the top level of the model as a hexagonal grid of cells, see Fig. 3. We get the HCPN model by step-wise colouring this spatial information. In the next step we create a refined model by adding details of the intra-cellular signalling mechanisms, and reusing the colouring template of the abstract model.

We first construct an unbiased model (Fig. 6), which does not include any proposed biased factors, to recapitulate the symmetrical localisation of proteins before the polarisation of PCP signalling. We then build a set of models based on this unbiased model by introducing various proposed biased factors in order to discover the underlying mechanism of PCP. These models are built up by inserting (i) feedback loops, (ii) feedback loops plus morphogen factor X, (iii) feedback loops plus biased transport, and (iv) feedback loops plus factor X and biased transport.

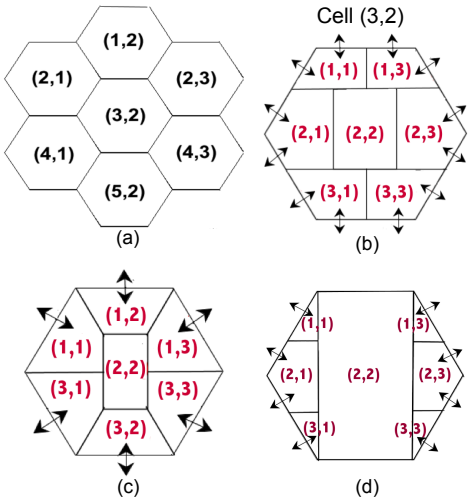


Fig. 3. Drosophila wing epithelial cells. (a) Fragment of wing tissue; coordinates represent honeycomb grid position; (b) Cell with seven virtual compartments in **one-to-two** relationship, arrows denote inter-cellular communication with adjacent neighbouring cells. (c) Likewise, for **one-to-one** relationship. (d) Cell asymmetrically divided into seven virtual compartments, see [12].

Abstract Petri net model for a single cell. Our initial model of the wing epithelial cell, illustrated in Fig. 4, is a high-level representation of the PCP network in order to establish the colour sets required.

We first sub-divide each cell into four spatial regions (Fig. 4, from left to right): (i) the extracellular space (labelled as communication), where the intercellular complexes form, (ii) the proximal cell margin (left-hand side of each cell) in order to process intercellular signals between two neighbouring cells, (iii) production (read arcs cause an infinite supply of proteins) and intracellular transport of core species, and (iv) the distal cell margin (right-hand side of each cell).

TABLE 1
Declarations for the abstract and refined HCPN models. Abbrev used: "const" – "constant", "cs" – "color set", "var" – "variable", and "fun" – "function".

Category	Declaration
const	$int : M = 15, N = 15, R = 3, C = 3$
cs	Row = int with $1 - M$
cs	$Column$ = int with $1 - N$
cs	$CS1$ = product with $Row \times Column$
cs	CS_Cell = $CS1$ with $x\%2 = 1\&y\%2 = 0 \vert x\%2 = 0\&y\%2 = 1$
cs	$ComR$ = int with $1 - R$
cs	$ComC$ = int with $1 - C$
cs	CS_ComP = product with $ComR \times ComC$
cs	$CS2$ = product with $CS\_Cell \times CS\_ComP$
cs	$CSmembrane$ = $CS2$ with $b = 1 \vert b = 3$
cs	$CSmiddle$ = $CS2$ with $a = 2\&b = 2$
cs	CSI = int with $1 - 2$
cs	$CS3$ = product with $CMembrane \times CSI$
var	$x : Row, y : Column,$ $a : ComR, b : ComC, r : CSI$
fun	NN $(Row\ x, Column\ y, ComR\ a, ComC\ b, CSI\ r)$ $CSMembrane$: $\{[(!(x = 1 \vert x = 2))\&(r = 2\&a = 1\&b = 1 \vert r = 1\&a = 1\&b = 3)], (((x - 2, y), (a + 2, b)), +r)\}$
fun	NW $(Row\ x, Column\ y, ComR\ a, ComC\ b, CSI\ r)$ $CMembrane$: $\{[(!(x = 1 \vert y = 1))\&(r = 1\&a = 1\&b = 1 \vert r = 2\& a = 2\&b = 1)], (((x - 1, y - 1), (a + 1, b + 2)), +r)\}$
fun	SW $(Row\ x, Column\ y, ComR\ a, ComC\ b, CSI\ r)$ $CMembrane$: $\{[(!(x = M \vert y = 1))\&(r = 1\&a = 2\&b = 1 \vert r = 2\& a = 3\&b = 1)], (((x + 1, y - 1), (a - 1, b + 2)), +r)\}$
fun	$Gd1$ $(Row\ x, Column\ y)$ $bool$: $\{x\%2 = 1\&y\%2 = 0 \vert x\%2 = 0\&y\%2 = 1\}$
fun	$Gd2$ $(Row\ x, Column\ y, ComR\ a, ComC\ b)$ $bool$: $\{(!(x = 1 \vert x = 2))\&(a = 1)\&(r = 1\&b = 3 \vert r = 2\& b = 1) \vert (!(x = 1 \vert y = 1))\&(b = 1)\&(r = 1\&a = 1 \vert r = 2\&a = 2) \vert (!(x = M \vert y = 1))\&(b = 1)\& (r = 2\&a = 3 \vert r = 1\&a = 2)\}$
fun	$RectangleReg$ $(Row\ x, Column\ y)$ $bool$: $\{x >= 15\&x <= 25\&y >= 20\&y <= 25\}$
fun	$DiskReg$ $(Row\ x, Column\ y, Row\ x0, Column\ y0, int$ radius$)$ $bool$: $\{(abs(y - y0) <= radius)\& (abs(x - x0) + abs(y - y0) <= 2 * radius)\}$
fun	$OvalReg$ $(Row\ x, Column\ y, Row\ x0, Column\ y0, int$ rx$, int$ ry$)$ $bool$: $\{(rx < ry)\&(abs(x - x0) <= rx)\& (abs(x - x0) + abs(y - y0) <= ry) \vert (rx >= ry)\&(abs(y - y0) <= ry)\& (abs(x - x0) + abs(y - y0) <= rx)\}$

In order to facilitate the detection of PCP asymmetry, we then partition each cell into seven *virtual compartments*, three each for the proximal and distal membrane compartments arranged in a circular manner, and one middle compartment for the cytosol, see Fig. 3, right. This circular structure is imposed on an underlying 3×3 rectangular grid, which explains the numbering of the compartments. We also create different neighbourhood relationships between adjacent

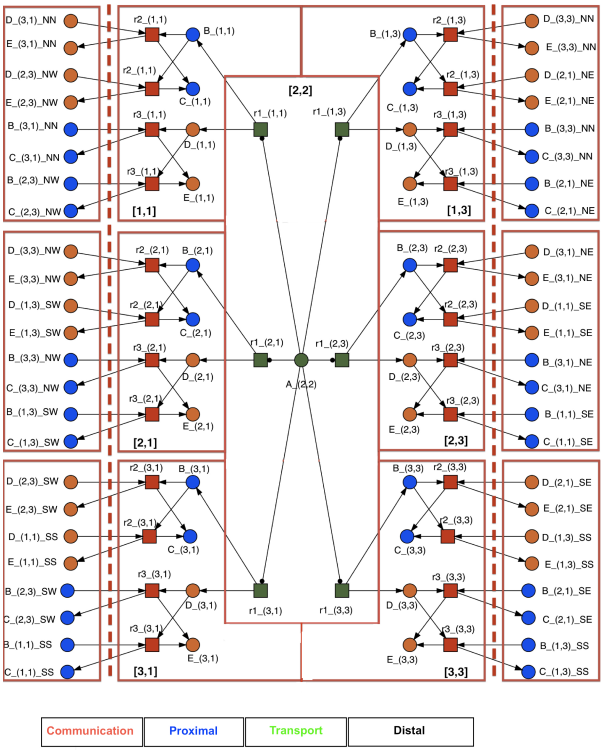


Fig. 4. Abstract PN for a single cell, one-to-two neighbourhood relationship, four spatial regions: communication, proximal, transport and distal, and the seven virtual compartments. Dashed red lines indicate cell boundaries.

cells in order to investigate their influence on the overall behaviour. The two neighbourhood relationships are (i) *one-to-two*, where one individual compartment in a cell communicates with two other compartments in the neighbouring cell(s), see Fig. $3(b)$; (ii) *one-to-one*, where each compartment in a cell communicates with only one compartment in the neighbouring cell, see Fig. $3(c)$.

Both of these relationships require the definition of six neighbourhood functions (north, north-east, south-east, south, south-west and north-west, abbreviated as NN, NE, SE, SS, SW, NW respectively), see Fig. 4. Because a honeycomb is a tessellated structure, we can ignore pairwise complementary symmetries between neighbouring cells, thus reducing the need for the 6 rotational axes of symmetry to 3 axes (NN, SW, NW), see Table 1.

In our previous work [12] we designed an asymmetric model, see Fig. $3(d)$, where the cell was divided into virtual compartments, three each at the proximal and distal sides, and one large central block. This configuration introduced an inherent proximal-distal topological bias across the wing tissue. In addition our previous model included three specific intracellular biasing mechanisms in terms of feedback loops, morphogen factor X and cytosolic transport. Our current work eliminates the tissue-level bias by using the

symmetrical configurations of the virtual compartments in Fig. 3$(c)\&(d)$, and then focuses on controlled addition of the intracellular biasing mechanisms (see "Refined models" below).

HCPN model for honeycomb lattice of cells. We begin by demonstrating how to construct a HCPN model for a one-to-two neighbourhood relationship and later illustrate how simple it is to adapt this approach to generate a HCPN model for a one-to-one relationship. For the following see Table 1.

We define two constants M, N and a two-dimensional colour set ($CS1$) representing a rectangular $M \times N$ grid, and select the subset denoting the coordinates of the hexagonally packed cells (CS_Cell), (Fig. 3). At this level of hierarchy (wing tissue comprising folded cells) we obtain a HCPN model, which has a similar structure to that of Fig. 4, but each place has been assigned the colour set CS_Cell.

Next we assign a colour to each of the seven virtual compartments of a cell. We do this by using a 3×3 grid (CS_ComP) and ignoring colours (2,1) and (2,3) so that the proximal compartments are (1,1), (2,1) and (3,1), the middle compartment is (2,2), and the distal compartments are (1,3), (2,3) and (3,3). We combine information about cell and compartment locality by defining $CS2$ as product of the colour set to address cells (CS_Cell) and the colour set to address the virtual compartments (CS_ComP). We introduce two subsets $CSmembrane$, and $CSmiddle$ of $CS2$ to facilitate addressing the components of a cell in a specific region, (i.e. distal or proximal, and middle). The colour sets we define are hierarchical, so we can locate each place in terms of the coordinates $((x,y),(a,b))$, where (x,y) denotes the position of a cell in the honeycomb grid, and (a,b) denotes the position of a virtual compartment within that cell.

We continue folding using these colours to obtain a more compact HCPN model – a tissue of cells comprising virtual compartments. This is achieved by folding the six membrane compartments into one, by assigning the colour set $CSmembrane$ to its places. The central compartment $(2,2)$ is denoted by the colour set $CSmiddle$.

Finally, we fold the two similar communication components (the transitions given in red in Fig. 4) in each membrane compartment into one. For this we define a colour set CSI of two colors, and a product colour set $CS3$ based on $CSmembrane$ and CSI.

In the following, we describe the necessary steps to construct the compact HCPN model. Having defined the colour sets, we create variables that are used in transition guards and arc expressions. All the transitions in the six membrane compartments are assigned guard $Gd1$ that selects the coordinates of the hexagonally packed cells, while the communication transitions are assigned guard $Gd2$ which mediates the communication between two neighbouring cells. Each arc in the six membrane compartments for the proximal and distal cell edges is assigned an expression $(((x,y),(a,b)),r)$, where the coordinate tuples $((x,y),(a,b))$ describe the arcs linking the associated place to a particular transition in compartment (a,b) in cell (x,y), and r takes a value of the colorset CSI. In the middle compartment, the arc expression is $((x,y),(2,2))$. In the communication region, each arc is assigned an expression $NN(x,y,a,b,r)$++$NW(x,y,a,b,r)$++$SW(x,y,a,b,r)$ which defines how neighbouring cells communicate with each other. Finally we obtain the HCPN of the abstract PCP model illustrated in Fig. 5. Unfolding this model gives the plain Petri net model in Fig. 4.

This is a generic model able to generate honeycomb tissues of arbitrary size by adjusting the two constants M, N. The colour sets define a pattern which can easily be reused to model similar scenarios of spatial locality. For example, the colour sets for a one-to-one HCPN model retain most of what we have established for the above HCPN model. Since this neighbourhood relationship only enables each compartment of the cell to communicate with one compartment in an adjacent cell, we can simply remove CSI, $CS3$, and the parameter CSI r in the neighbourhood functions. Due to space considerations, in this paper we concentrate on the one-to-two model.

In summary, the procedure to construct a HCPN model for a multi-cellular tissue with regularly structured compartments inside each cell can be divided into two steps.

- Fold cells: build up the structure of multi-cells representing the locality of each cell.
- Fold compartments: create the required localisation of compartments within a cell.

This can be trivially extended to model systems with more than two levels of hierarchy.

Refined models of PCP. We develop step-wise a set of more detailed PCP models by refining the model in Fig. 5. The refined models of PCP signalling consider a set of five core proteins (Fz, Dsh, Vang, Pk and Fmi, see Section 2).

(i) We start off with an unbiased model of PCP (Fig. 6) which creates an equal distribution of FFD in all membrane compartments in those cells having six neighbours (see Fig. 6). Remark: boundary cells or cells communicating with less than six neighbours will not have a symmetric localisation of FFD due to the lack of intercellular communication.

The following models are available as supplementary material.

(ii) In order to build up a basic biased model with which to test different hypotheses, we include the intracellular inhibitory loops which mediate a competition between proximal and distal proteins displayed on adjacent surfaces of neighbouring cells [41].

(iii) Subsequently we create a set of HCPN models based on this basic biased model, by introducing either factor X or biased transport alone, or both

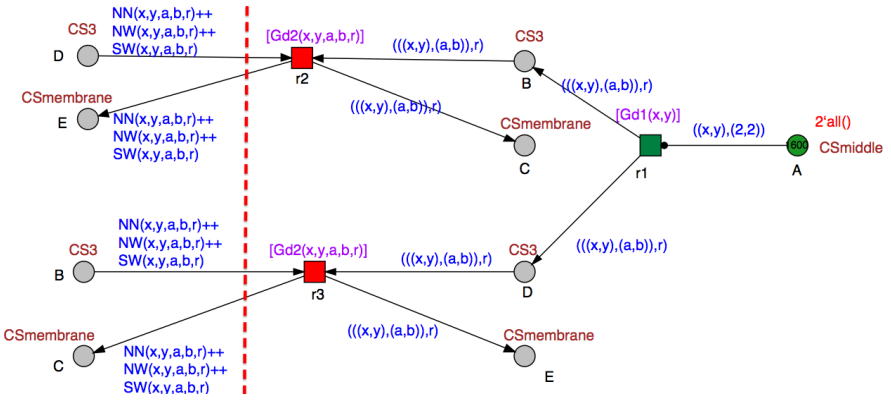


Fig. 5. Abstract HCPN model, folded version of Fig. 4. Places B, C, D, and E are logical nodes which are in the distal compartments of each cells. See Table 1 for declarations.

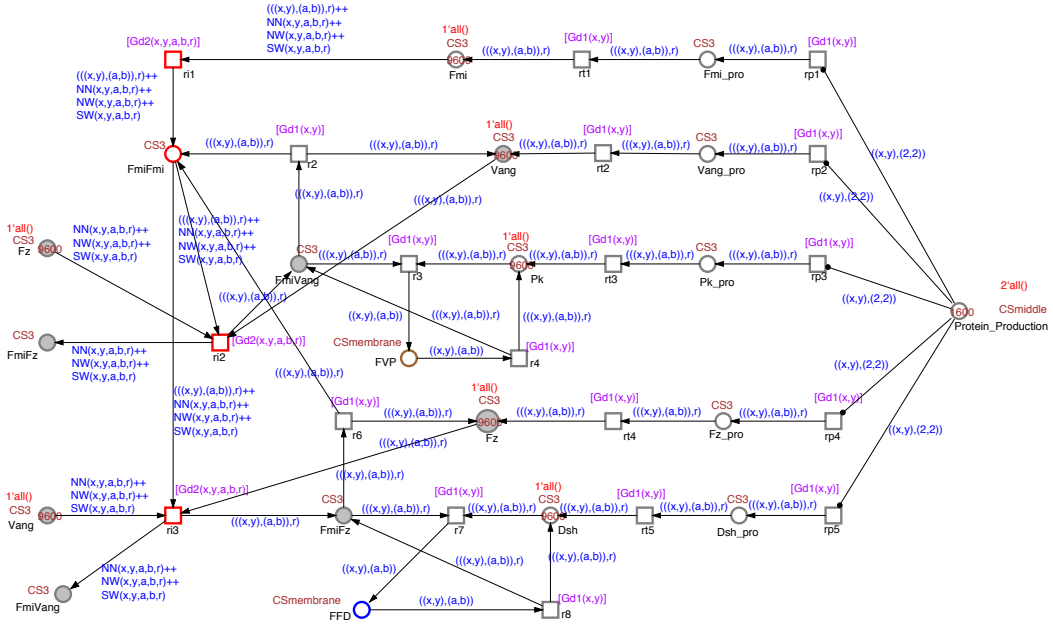


Fig. 6. HCPN model of unbiased PCP. It refines the abstract HCPN model given in Fig. 5.

combined. These variants enable us to test various hypotheses underlying PCP signalling.

In order to simplify the identification of these models, we distinguish between the biased and unbiased versions. If the model is biased (B) then it always includes feedback loops (F), and optionally either or both of biased transport (T) or factor X ligand (X). Finally a biased model can be wild-type (Wt) or contain one of two mutant clones, Fz- or Vang-. We can describe this naming convention by the following simple BNF: **B F [X] [T] [Wt|Fz|Vang]**.

To construct these refined models we use the same declarations as those used for Fig. 5 (see Table 1); thus, we do not need to start from scratch. We group transitions and places into different spatial regions and virtual compartments, and then assign the same colour sets to each region or compartment as done for Fig. 5, likewise for arc expressions. Additionally, we define three functions, $RectangleReg$, $DiskReg$ and $OvalReg$ to generate a rectangle, circle-like (hexagonal) or oval region within the whole tissue, respectively (see Table 1) which will be used to place mutant cells on the tissue to model biological mutant clones. Thus, with our models we can not only perform simulations of PCP signalling in normal, wild-type cells but also on patches of mutant cells in a wild-type background.

Quantitative models. We quantify our PCP models by assigning to each transition a rate function following mass-action kinetics. The kinetic parameters have been optimised by using simulated annealing with wet-lab time series target data, taken from [2].

These quantitative models can be equally read as stochastic or continuous models, with appropriate scaling of the kinetic constants. Particularly, a continuous HCPN model uniquely defines a set of ODEs. To simulate wild-type and different mutant conditions,

we consider individual marking sets, function sets, and parameter sets (which are maintained within one model file).

All the models we have discussed are available at http://people.brunel.ac.uk/~cspgqqg.

6 ANALYSIS

HCPNs enjoy a large variety of analysis techniques, ranging from informal animation to formal static or dynamic analysis techniques. In the following we confine ourselves to simulative methods to analyse dynamic properties of our models in the continuous setting, because the infinite state space caused by the infinite supply of proteins precludes the application of dynamic exhaustive analysis techniques. Specifically we apply automated clustering and visualisation techniques to identify sets of cells with similar behaviours, complemented by PLTL model checking and visual analysis to both primary and secondary time-series data from representative cells of automatically identified clusters.

6.1 Analysis techniques

Cluster analysis. The strategy to analyse the models is based on the behaviour of each cell, represented by the FFD complex concentrations for each of the six membrane compartments because these are associated with the PCP response [39]. In the first step, we take the time points of the FFD time series from the six membrane compartments and use them as 'features' for each cell; they comprise the characteristics which define the behaviour. With e.g. 200 time points for each of the six time series we obtain 1,200 features; too many for a clustering algorithm, which would be slowed down and its result diluted by the background noise.

We then reduce the number of features by applying a feature selection technique: Principal Component Analysis (PCA) [21]. This way, the original features are converted into new features, which are a linear combination of the previous ones, see Fig. 7. PCA allows us to compute the variance explained by each new feature in order to find how many new features we need to describe the space. We then check how many features are needed to explain at least 95% of the variance.

Once we have obtained the new dataset through PCA, we apply a clustering algorithm to find groups of cells with similar behaviour. There are many clustering algorithms available in literature, such as the commonly used K-means [30]. However we cannot rely on K-means because we do not know a priori the number of clusters we are looking for, and most importantly the clusters are not necessarily spherical. For this reason we employ a density-based clustering algorithm called DBScan [11], which permits finding

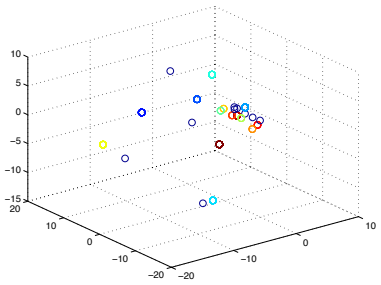


Fig. 7. Representation of the most important three features derived by PCA for model BFXFz.

clusters of any shape. DBScan identifies dense areas of space by looking at the neighbourhood of each point.

An area of a given radius is considered to be dense if it contains at least a given number of points. The algorithm then merges connected dense areas to form a cluster. All the points left out of the clusters are labelled as outliers. We can find the best number of clusters to fit the scenario by inspecting the three dimensional data space and adjusting the parameters (radius and size of the neighbourhood). The results are then plotted on a hexagonal grid, which represents the layout of the cells in the model, see experiments below. The cells in each cluster are allocated a unique colour and are labelled by the cluster number generated by our technique. All outliers are allocated -1 as their cluster number.

Model checking. In order to explore in more detail the behaviour of a model, we perform model checking on representative cells from the clusters identified by the cluster analysis. Model checking is a technique which is used to check the validity of properties of a dynamic model expressed in a temporal logic; this can be performed either analytically (requiring the entire state space to be explored), or over finite time-series traces produced by the simulation of the model. In our work we use simulative model checking because of the infinite state space of our models, and specifically employ the MC2 model checker [10] which operates over Probabilistic Linear Temporal Logic with constraints (PLTLc). We check both the primary data output from a model (i.e. time series of concentrations) as well as secondary data – in this case time series of accumulated concentrations (known as cumulative rewards in CSL model checking [36]).

We check this secondary data because the localisation of PCP signalling components at any given time point is the result of the cumulative effect of the sum over the signalling events until that point. We do this by computing the accumulation of the concentration of FFD at each point in the time series trace, and call this a *cumulative time series*. The cumulative time series data "smooths away" small variations which are apparent in the primary data.

6.2 Models

For reasons of space all experiments reported in this section refer to one-to-two neighbourhood relationship models, namely the unbiased model and variants of the biased model. We focus on the BFX family of models in order to further conserve space.

A crucial point is how many cells we can simulate in terms of current computational capabilities, i.e. what tissue size can we actually analyse. In this respect, we have to address three technical key problems: unfolding, ODEs construction, and simulation. The coloured continuous Petri nets are automatically unfolded which can be considered as a kind of compilation. Afterwards, the corresponding ODE needs to be constructed for each place; the runtime for this step is negligible. Finally, the model has to be simulated using an appropriate continuous simulation algorithm.

We performed a simple test by increasing the number of cells in the tissue. We report the runtime for unfolding and simulation for increasing size of the PCP model in Table 2. From the ratio of unfolding or simulation runtime to number of cells, respectively, we can see that both the unfolding and simulation runtime increase approximately linearly.

We begin the experiments by using our unbiased and biased HCPN models to generate an in-silico wildtype tissue based on a honeycomb grid comprising 800 cells imposed on an underlying 40×40 rectangular grid. Next, we have modelled the effect of a patch of mutated cells (clone) lacking the key signalling molecules Frizzled (Fz) or Van Gogh (Vang), denoted by Fz- and Vang- respectively, in an otherwise wild-type field of cells by zeroing the concentration and switching off transport in the corresponding Petri net places. Using the $DiskReg$ function (see Table 1), we can produce a mutant clone of Fz- or Vang- in a circle-like shape with 80 mutant cells, i.e. 10% of the in-silico tissue.

We conduct our analysis as a proof of principle that our models indeed have the ability to capture the given biological phenomena and make sensible predictions. All simulations were run over 1,000 time units reported at 200 time points. The time span represents 33 hours of PCP signalling in-vivo.

6.3 Experiments

In order to validate the ability of our models to recapitulate biological phenomena observed in the wet lab, we start off by simulating and analysing the wild-type models, before considering the effect on neighbouring wild-type cells of a patch of mutant cells lacking Frizzled or Van Gogh proteins.

Experiment 1: Unbiased model

Current biological models of PCP in the literature [39] show that at initialisation of signalling PCP proteins are symmetrically distributed along the cell membrane. In order to mimic this we use our unbiased model which reflects this symmetry.

Clustering. We expect that FFD complexes will be equally distributed over the six membrane compartments in those cells which communicate with six neighbouring cells (i.e. except boundary cells). The continuous simulation result confirms the expected behaviour, as shown in Fig. 9; all cells with six neighbours belong to cluster 3 (labelled in green in Fig. 9). Cells at each boundary are separately assigned to other clusters; due to the different number of communication partners resulting from the honeycomb grid. In the following analyses we ignore tissue border cells, since these are an artifact, and there are no real biological data or observations for these.

Model checking, primary data. The unbiased model has been constructed to reproduce the lack of polarisation, causing proteins to be symmetrically distributed along the cell membrane. This is confirmed by our results, see Fig. 12 which shows identical time-series plots for all six membrane compartments.

For this, we form the following query, where P_1, P_2, P_3, D_1, D_2 and D_3 denote the concentration of FFD in the proximal and distal compartments.

$$P_{=?} \quad [G(P_1 = P_2 \pm \delta \wedge P_1 = P_3 \pm \delta \wedge P_1 = D_1 \pm \delta \wedge P_1 = D_2 \pm \delta \wedge P_1 = D_3 \pm \delta)]$$

Note that $P_1 = P_2 \pm \delta$ is a short-hand for $P_2 - \delta < P_1 < P_2 + \delta$. We use a small value $\delta = 0.003$ because we are dealing with real numbers, and there can be slight variations between traces due to, for example, effects in the solver. This query holds for all non-boundary cells.

Model checking, secondary data. Because the wild type cells (except boundary cells) exhibit the same time-series behaviour in all compartments, the accumulation of FFD is identical in all membrane compartments (Fig. 18).

We use the same query as above, but set $\delta = 0.04$ due to the cumulative nature of variations; again this query holds for all non-boundary cells.

Experiment 2: Biased model, wild-type (BFXWt)

Clustering. The feedback loops mediate the competition between proximal and distal proteins between adjacent surfaces of neighbouring cells and amplify the asymmetric localisation of these proteins. The results obtained by clustering analysis for this model exhibit similar characteristics to the wild-type unbiased model (Experiment 1, Fig. 9).

Model checking, primary data. The time-series behaviour for a representative cell is shown in Fig. 13. We choose to compare the middle proximal and distal compartments because they exhibit the largest differences in proximal-distal intra-cellular behaviour. At the initial time point all concentrations are zero; after this the concentration of FFD in the middle

TABLE 2
Unbiased PCP model size and runtime[a] for unfolding and continuous simulation over 1000 time units, illustrating the efficiency of our computational technology.

Size				Unfolding runtime (seconds)		Simulation runtime (seconds)
Grid($M \times N$)	Cells	Places	Transitions	Before optimisation	After optimisation	
5×5	12	2,028	2,802	3.195	1.154	3.145
10×10	50	8,450	11,826	9.714	2.613	14.618
15×15	112	18,928	26,622	22.771	4.495	42.586
20×20	200	33,800	47,646	44.818	9.231	88.886
40×40	800	135,200	191,286	280.598	83.162	371.647
40×40^b	800	164,000	229,686	329.384	120.186	7,399.544

[a] performed on a Mac Quad-core Intel Xeon, CPU 2× 2.26GHz, memory (DDR 3) 8 GB; [b] for the biased model BFXWt.

distal compartment (D_2) is always the highest of all the compartments, and that in the middle proximal compartment (P_2) is always the lowest. The upper and lower distal compartments are equal, lower than the middle distal and higher than the corresponding upper and lower proximal compartments.

We first of all construct the following query to reflect the observation above, where $\delta = 0.002$:

$$P_{=?} \quad [time > 0 \rightarrow G(D_2 > D_1 \wedge D_1 = D_3 \pm \delta \wedge D_1 > P_1 \wedge P_1 = P_3 \pm \delta \wedge P_1 > P_2)]$$

We also observe that all traces exhibit one peak followed by a trough, and formulate a query for compartment D_2 as an example, using the differential function:

$$P_{=?}[F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$$

We can extend our queries in the obvious manner to show that the proximal compartments have two peaks whilst the distal compartments have one peak in the traces over the time period shown here.

Model checking, secondary data. We observe that the middle distal compartment always has a higher cumulative time series than the other compartments, whereas the middle proximal compartment is consistently the lowest. The upper and lower distal compartments are very similar, less than the middle distal and higher than the corresponding upper and lower proximal compartments (Fig. 19).

We denote the cumulative variables for the membrane compartments $P_1, \ldots, D_3$ by $CP_1, \ldots, CD_3$. The equivalent query to that which we formulated for the time series traces holds for the cumulative signal, again with $\delta = 0.002$:

$$P_{=?} \quad [time > 0 \rightarrow G(CD_2 > CD_1 \wedge CD_1 = CD_3 \pm \delta \wedge CD_1 > CP_1 \wedge CP_1 = CP_3 \wedge CP_1 > CP_2)]$$

We can use the free variables in MC2 [10] to compute the range over the difference in the maximum and minimum cumulative signals:

$$P_{=?}\ [\$\gamma = max(CD_2) - max(CP_2)]$$

In this case $\gamma = 194$, indicating the clear signal in the Wt for the formation of the hairs in the central part of the distal edge of the cell.

Experiment 3: Biased models with mutated clones (BFXFz and BFXVang)

We present the results of both Fz- and Vang- mutant clones. As reported in [44], cells in a Fz- clone have incorrect polarity and occasional multiple hairs. Wild-type cells distal, but not proximal to the clone have incorrect polarity, pointing more proximally towards the clone [22], see Fig. 8. Regarding the capability of our current model, we expect that: (i) cells in the clone have incorrect polarity (FFD does not form); (ii) wild-type cells distal to the clone have FFD accumulated at the proximal rather than distal edge of cell.

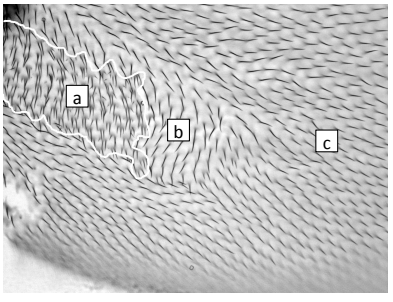


Fig. 8. Image of Drosophila wing. (a) Fz- clone, (b) wild-type cells distally adjacent to clone, (c) wild-type cells far from clone.

Cells in a Vang- clone have proximal domineering nonautonomy characteristic, the wild-type neighbouring cells proximally next to the clone point away from the clone rather than towards the clone [29].

Clustering: Fz- mutant clone. The result demonstrates the impact of a clone of Fz- mutant cells on the neighbouring wild-type cells surrounding it, see Fig. 10 where all of the mutant cells are in one cluster, and we detect disruption in all the wild-type cells directly adjacent to the clone. In more detail, we detect different clusters, and hence different responses in the wild-type cells to each of the six sides of the clone. Note that each of the six cells at the vertices of the clone are detected as individual outliers by our technique. Because Fz is knocked out in all mutant cells, the supply of Fz to form FFD is completely cut off, leading to zero FFD in the mutant cells. In addition, FFD can be affected by adjacent cells due

to intercellular communication. These factors have an influence on the clusters that are detected.

Clustering Vang- mutant clone. The result shows the impact of a clone of Vang- mutant cells on their neighbouring cells around it (see Fig. 11). In this case the central cells in the clone are in one cluster, while the mutant cells at the edges of the clone are in other clusters – one cluster at each of the six edges. We detect disruption in all the wild-type cells directly adjacent to the clone, resulting in several clusters, one at each edge of the clone. Only cells at four of the vertices of the clone are detected as individual outliers by our technique. The lack of Vang in the central mutant cells indirectly affects the formation of FFD in adjacent cells due to intercellular communication. Thus while the central Vang- clone cells have zero FFD due to the lack of Vang in their surrounding cells, mutant cells at the boundary of the clone have some FFD because they maintain some communication with adjacent wild type cells.

Model checking, primary data, Fz- clone. In the following, we only consider the middle distal and middle proximal compartments, D_2 and P_2, because they exhibit the highest PCP signal.

Unlike in the wild-type cells, for the cells distally neighbouring to the clone the concentration of FFD in the middle distal compartment is always lower than that of the middle proximal compartment, see Fig. 14:

$P_{=?}\ [time > 0 \rightarrow G(D_2 < P_2)]$

Moreover, the trace of D_2 exhibits a peak followed by a trough, which is not true for P_2:

$P_{=?}[F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$

However the cells immediately proximal to the clone exhibit behaviour similar to wild-type cells, i.e. the middle distal compartment is always higher than the middle proximal compartment (Fig. 15). The concentration of FFD in all the clone cells is zero at all time points.

Model checking, primary data, Vang- clone. All the non-boundary cells within the clone have zero FFD at all time points, however the clone cells directly adjacent to the clone boundaries do exhibit some accumulation of FFD in those compartments directly abutting wild-type cells; this is to be expected because of the formation of FFD can take place with participation of Vang from the adjacent wild-type cells.

As in the wild-type cells, the concentration of FFD in the middle distal compartment of the cells distally neighbouring to the clone is always higher than that of the middle proximal compartment, but this latter is always zero. The trace for the distal compartment exhibits a peak and a trough, see Fig. 16.

$P_{=?}[(time > 0 \rightarrow G(D_2 > P_2)) \wedge G(P_2 = 0)]$

$P_{=?}\ [F(d(D_2) > 0 \wedge F(d(D_2) < 0 \wedge F(d(D_2) > 0)))]$

The behaviour of the cells immediately proximal to the clone is similar to wild-type cells, in that middle proximal compartment is always higher than the middle distal compartment, while the latter is always zero (Fig. 17).

$P_{=?}[(time > 0 \rightarrow G(P_2 > D_2)) \wedge G(D_2 = 0)]$

$P_{=?}\ [F(d(P_2) > 0 \wedge F(d(P_2) < 0 \wedge F(d(P_2) > 0)))]$

Model checking, secondary data, Fz- clone. All biased models with the Fz- clone (BFXFz, BFTFz and BFXTFz) always exhibit a relatively higher cumulative signal in the middle proximal compartment compared to the middle distal compartment in those cells distally directly next to the Fz- clone (see Fig. 20 for model BFXFz):

$P_{=?}\ [time > 0 \rightarrow G(CD_2 > CP_2)]$

In this case the range $\gamma = 72$, indicating a severe loss of polarising signal.

The wild type cells in the tissue (i.e. away from the clone area) display behaviour similar to the wild type cells in the BFXWt model (Fig. 19).

$P_{=?}[time > \epsilon \rightarrow G(CD_2 > CD_1 \wedge CD_1 = CD_3 \pm \delta \wedge CD_1 > CP_1 \wedge CP_1 = CP_3 \wedge CP_1 > CP_2)]$

where $\epsilon = 50$ and $\delta = 0.2$

The clustering technique is highly sensitive, and over-identifies clusters. However, the cumulative behaviour of the wild-type cells proximally adjacent to the Fz- clone (Fig. 21) is very similar to that of the wild-type cells in the BFXWt model (Fig. 19). All clone cells within the Fz- clone have zero accumulation of FFD in all compartments because there is no supply of Fz to form the FFD complex.

Model checking, secondary data, Vang- clone. In contrast to the Fz- clone, for all Vang- clone models (BFXVang, BFTVang and BFXTVang) the cumulative signal in the proximal cells next to the clone is always higher in the middle proximal compartment compared to the middle distal compartment, and the latter is always zero (see Fig. 23 for the BFXVang model):

$P_{=?}[(time > 0 \rightarrow G(CP_2 > CD_2)) \wedge G(CD_2 = 0)]$

In this case the range over the difference in the maximum and minimum cumulative signals between the compartments is given by $\gamma = 421$ over compartments CP_1 and CP_3 as the maxima and CD_2 as the minima indicating strong polarisation in the opposite direction to the Wt cells.

The behaviour of the distal cells next to the clone (Fig. 22) exhibits a similar trend to that of the wild-type cells; however the cumulative value for FFD in the middle proximal compartment is always zero, and the values in the upper and lower proximal compartments are lower than in the wild-type cells. This is because the knockout of Vang in the adjacent cells in the clone causes the lack of FFD in the distally neighbouring cells.

$P_{=?}[(time > 0 \rightarrow G(CD_2 > CP_2)) \wedge G(CP_2 = 0)]$

In this case $\gamma = 472$, indicating very strong polarisation in the normal (Wt) direction.

Remark. Our findings regarding the analysis of the secondary data are consistent with observations reported in the literature, for example [3], namely that PCP disturbance is reported in cells distal to Fz- clones, whereas the disturbance is on the proximal side of Vang- clones. The cause of this disturbance is indicated by the cumulative plots for these cells, where there is a lack of an orienting signal in the distal neighbours to the Fz- clone (Fig. 20), and the near-zero values in the distal compartments for the proximal neighbours to the Vang- clone (Fig. 23).

7 Reproducibility

All Petri net models in this paper were constructed with Snoopy [34], recently extended to support coloured Petri nets [27], [26], which can be obtained from http://www-dssz.informatik.tu-cottbus.de/DSSZ/Software/Snoopy.

Simulations were done with Snoopy's built-in continuous simulators, simulation traces have been written to csv files, which have then been further processed by Matlab®7.11.0 [31] and MC2 [10] for model checking. Clustering analysis techniques have been developed in Matlab. DBScan for Matlab was taken from [9]. The models in Snoopy format and high resolution diagrams, the simulation data, clustering routines and model checking queries can be found at http://people.brunel.ac.uk/~cspgqqg/.

Thus, all our results can be easily reproduced by the interested reader.

8 Conclusions

In this paper, we have presented our work applying Petri net techniques to construct a family of multiscale computational models in order to explore the mechanisms that drive Planar Cell Polarity in Drosophila wing tissue. We have shown that our family of models recapitulates signalling phenomena known to occur in wild-type and mutant clones, and in this paper we have focussed on the the effect of morphogen factor X on PCP. Our approach has involved developing sophisticated patterns of communication over hierarchically organised components. We have demonstrated the power of the new concept of hierarchically coloured Petri nets and associated analysis techniques (cluster analysis and simulative model checking of primary and secondary data) which we have developed. Our computational experiments have been over very large underlying models each comprising about 150,000 ODEs.

The behaviour of our models reflects the fact that the major accumulation of actin (from which the hairs are formed) occurs in the most distal part of wild-type cells, corresponding to the location of the prehair formation in wing cells of Drosophila. Moreover our models confirm that the introduction of mutant clones disrupts the pattern of actin accumulation and hence hair orientation in wild-type cells on the distal side of a Fz- clone and the proximal side of a Vang- clone.

Our HCPN models and the software required to simulate and analyse them are freely available, thus ensuring that our results are reproducible by the scientific community.

9 Acknowledgments

Qian Gao was supported by a bursary from Brunel University, Fei Liu by German Federal Ministry of Education and Research grant 0315449H, and Daniele Maccagnola by a scholarship from Universita di Milano-Bicocca. Monika Heiner and David Gilbert were supported by UK EPSRC grant EP/I036168/1.

References

[1] K. Amonlirdviman, N. A. Khare, D. R. P. Tree, W. S. Chen, J. D. Axelrod, and C. J. Tomlin. Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science*, 307, 2005.

[2] J. D. Axelrod. Unipolar membrane association of dishevelled mediates frizzled planar cell polarity signaling. *Genes & Development*, 15:1182–1187, 2001.

[3] J. D. Axelrod and C. J. Tomlin. Modeling the control of planar cell polarity. *WIREs Systems Biology and Medicine*, 2011.

[4] N. Bahi-Jaber and D. Pontier. Modeling transmission of directly transmitted infectious diseases using colored stochastic Petri nets. *Mathematical Biosciences*, 185:1–13, 2003.

[5] J. Barrow. Wnt/PCP signaling: A veritable polar star in establishing patterns of polarity in embryonic tissues. *Seminars in Cell and Developmental Biology*, 17(2):185–193, 2006.

[6] S. C. Brailsford, C. N. Potts, and B. M. Smith. Constraint satisfaction problems: Algorithms and applications. *European Journal of Operational Research*, 119(3):557–581, 1999.

[7] Y. Burak and B. I. Shraiman. Order and stochastic dynamic in drosophila planar cell polarity. *PLoS Computational Biology*, 5(12):e1000628, 2009.

[8] S. Christensen, J. B. Jørgensen, and L. M. Kristensen. Design/CPN - a computer tool for coloured Petri nets. In *Proc. TACAS*, LNCS 1217, pages 209–223. Springer, 1997.

[9] M. Daszykowski, B. Walczak, and D. Massart. Looking for natural patterns in data: Part 1. density-based approach. *Chemometrics and Intelligent Laboratory Systems*, 56(2):83 – 92, 2001.

[10] R. Donaldson and D. Gilbert. A model checking approach to the parameter estimation of biochemical pathways. *LNCS/LNBI*, 5307:269–287, 2008.

[11] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD*, pages 226–231, 1996.

[12] Q. Gao, F. Liu, D. Gilbert, M. Heiner, and D. Tree. A multiscale approach to modelling planar cell polarity in drosophila wing using hierarchically coloured Petri nets. In *Proc. 9th International Conference on Computational Methods in Systems Biology (CMSB 2011)*, pages 209–218. ACM digital library, September 2011.

[13] Gecode. Gecode: An open constraint solving library, 2011.

[14] H. Genrich, R. Küffner, and K. Voss. Executable Petri net models for the analysis of metabolic pathways. *International Journal on Software Tools for Technology Transfer*, 3(4):394–404, 2001.

[15] D. Gilbert, M. Heiner, and S. Lehrack. A unifying framework for modelling and analysing biochemical pathways using Petri nets. *LNCS/LNBI*, 4695, 2007.

[16] A. P. Heath and L. E. Kavraki. Computational challenges in systems biology. *Computer Science Review*, 3(1):1–17, 2009.

[17] M. Heiner, D. Gilbert, and R. Donaldson. Petri nets in systems and synthetic biology. In *Schools on Formal Methods (SFM)*, pages 215–264. LNCS 5016, Springer, 2008.

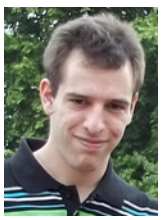
[18] K. Jensen. Coloured Petri nets and the invariant-method. *Theoretical Computer Science*, 14:317–336, 1981.
[19] K. Jensen and L. M. Kristensen. *Coloured Petri nets*. Springer, 2009.
[20] K. Jensen, L. M. Kristensen, and L. M. Wells. Coloured Petri nets and cpn tools for modelling and validation of concurrent systems. *International Journal on Software Tools for Technology Transfer*, 9(3/4):213–254, 2007.
[21] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, New York, 1986.
[22] P. A. Lawrence, G. Struhl, and J. Casal. Planar cell polarity: A bridge too far? *Current Biology*, 18(20):959–961, 2010.
[23] J. F. Le Garrec and M. Kerszberg. Modeling polarity buildup and cell fate decision in the fly eye: Insight into the connection between the pcp and notch pathways. *Development Genes & Evolution*, 218:413–426, 2008.
[24] J. F. Le Garrec, P. Lopez, and M. Kerszberg. Establishment and maintenance of planar epithelial cell polarity by asymmetric cadherin bridges: A computer model. *Development Dynamics*, 235:235–246, 2006.
[25] D. Lee, R. Zimmer, S. Lee, and S. Park. Colored Petri net modeling and simulation of signal transduction pathways. *Metabolic Engineering*, 8:112–122, 2006.
[26] F. Liu. *Colored Petri nets for Systems Biology*. PhD thesis, Brandenburg University of Technology Cottbus, Department of Computer Science, January 2012.
[27] F. Liu and M. Heiner. Colored Petri nets to model and simulate biological systems. In *International Workshop on Biological Processes and Petri nets (BioPPN), satellite event of Petri nets 2010*, June 2010.
[28] F. Liu and M. Heiner. Manual for Colored Petri Nets in Snoopy. Technical report, Brandenburg University of Technology Cottbus, Department of Computer Science, July 2011.
[29] D. Ma, C. H. Yang, H. McNeill, and M. A. Simon. Fideility in planar cell polarity in drosophila. *Nature*, 421:543–547, 2003.
[30] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proc. of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, Berkeley, CA, 1965. University of California Press.
[31] MATLAB. *version 7.11.0*. The MathWorks Inc., Natick, Massachusetts, 2011.
[32] M. Povelones and R. Nusse. The role of the cysteine-rich domain of frizzled in wingless-armadillo signaling. *EMBO Journal*, 24:3493–3503, 2005.
[33] R. L. Raffard, K. Amonlirdviman, J. D. Axelrod, and C. J. Tomlin. An adjoint-based parameter identification algorithm applied to planar cell polarity signaling. *IEEE Transaction on Automatic Control*, 53:109–121, 2008.
[34] C. Rohr, W. Marwan, and M. Heiner. Snoopy - a unifying Petri net framework to investigate biomolecular networks. *Bioinformatics*, 26(7):974–975, 2010.
[35] S. Schamberg, P. Houston, N. A. M. Monk, and M. R. Owen. Modelling and analysis of planar cell polarity. *Bulletin of Mathematical Biology*, 72:645–680, 2010.
[36] M. Schwarick, C. Rohr, and M. Heiner. Marcie - model checking and reachability analysis done efficiently. In *Proc. 8th International Conference on Quantitative Evaluation of SysTems (QEST 2011)*, pages 91–100. IEEE CS Press, September 2011.
[37] Y. Shimada, S. Yonemura, H. Ohkura, D. Strutt, and T. Uemura. Polarized transport of frizzled along the planar microtubule arrays in drosophila wing epithelium. *Developmental Cell*, 10:209–222, 2006.
[38] M. Simons and M. Mlodzik. Planar cell polarity signaling: From fly development to human disease. *Annual Review of Genetics*, 42:517–540, 2008.
[39] D. I. Strutt. Asymmetric localization of frizzled and the establishment of cell polarity in the drosophila wing. *Molecular Cell*, 7:367–375, 2001.
[40] D. I. Strutt. Asymmetric localization of frizzled and the establishment of cell polarity in the drosophila wing. *Molecular Cell*, 7:367–375, 2002.
[41] D. R. P. Tree, J. M. Shulman, R. Rousset, M. Scott, D. Gubb, and J. D. Axelrod. Prickle mediates feedback amplification to generate asymmetric planar cell polarity signaling. *Cell*, 109:371–381, 2002.
[42] E. P. K. Tsang. *Foundations of Constraint Satisfaction*. Academic Press, London and San Diego, 1993.
[43] R. Valk. Petri nets as token objects: An introduction to elementary object nets. In J. Desel and M. Silva, editors, *Application and Theory of Petri Nets 1998*, volume 1420 of *LNCS*, pages 1–25. Springer-Verlag, Berlin, 1998.
[44] C. R. Vinson and P. N. Adler. Directional non-cell autonomy and the transmission of polarity information by the frizzled gene of drosophila. *Nature*, 14(329):549–551, 1987.
[45] K. Voss, M. Heiner, and I. Koch. Steady state analysis of metabolic pathways using Petri nets. *In Silico Biology*, 3:0031, 2003.
[46] C. Yang, J. D. Axelrod, and M. A. Simon. Regulation of frizzled by fat-like cadherins during planar polarity signaling in the drosophila compound eye. *Cell*, 108:675–688, 2002.

Qian Gao is a PhD student of Systems Biology at School of Information Systems, Computing and Mathematics, Brunel University, UK. Her interests centre around System Biology, Bioinformatics and their applications in biomedical and health sciences. Her current research focuses on the development and application of computational techniques to model and analyse multiscale systems. Her web page is http://people.brunel.ac.uk/~cspgqqg.

David Gilbert is a Professor of Computing and Co-Director of the Centre for Systems and Synthetic Biology at Brunel University, UK. He holds a PhD in computational logic from Imperial College, London. David's research interests lie in the development and application of computational techniques to model and analyse biological data, with a focus on multiscale systems, and approaches to support the design and development of synthetic biological systems. His web page is http://people.brunel.ac.uk/~csstdrg.

Monika Heiner received her Ph.D. at the University of Technology at Dresden. Since 1994 she holds a professorship for data structures and software dependability at the Brandenburg University of Technology in Cottbus. She has a long history of expertise in Petri nets, shifting her application focus from technical to natural systems in the late nineties. She applied various techniques from the standard body of Petri net theory to biochemical network validation, and also employed computational methods for stochastic and continuous network evaluation, including model checking. Methodology-oriented aspects of her work include a unifying framework of computational models and BioModel Engineering. Over the last 15 years her group at Cottbus has developed a sophisticated toolkit to support the methodology, comprising Snoopy, Charlie, and Marcie.

Fei Liu was a scientific staff member of the Data Structures and Software Dependability group, Brandenburg University of Technology. He is an associate professor of the Control and Simulation Center, Harbin Institute of Technology. He received a doctorate in control science and engineering from HIT and a second doctorate in computer science from BTU. His research interests are modelling and simulation, colored Petri nets, and systems biology.



Daniele Maccagnola received a bachelor degree in Computer Science and a master degree in Bioinformatics at the University of Milano-Bicocca. From March 2012, he is a research grant holder at the Department of Computer Science (DISCo), University of Milano-Bicocca. His main interests are in data mining, clustering and their application to systems biology. His email address is daniele.maccagnola@gmail.com.

David Tree is a lecturer in the Division of Biosciences, School of Health Sciences and Social Care, Brunel University, UK. He is a developmental geneticist who uses Drosophila genetics to study the genetic, molecular and cell biological mechanisms controlling patterning during embryonic development and the homeostasis of adult tissues and organs. Current work in his lab focuses on cellular reception and integration of extracellular signals to regulate cell polarity, tissue homeostasis and aging.

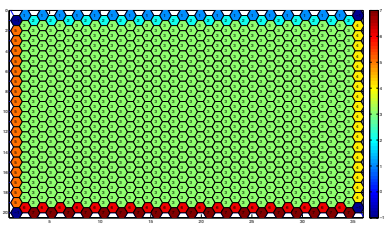


Fig. 9. Clustering for continuous simulation of unbiased and BFXWt models.

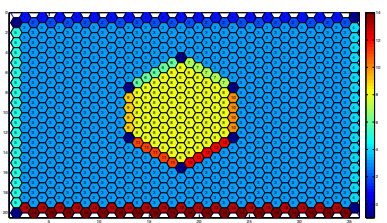


Fig. 10. Clustering for continuous simulation of model BFXFz.

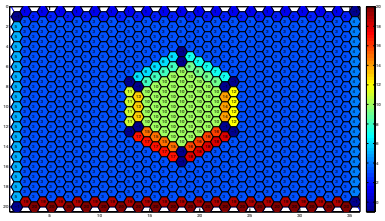


Fig. 11. Clustering for continuous simulation of model BFXVang.

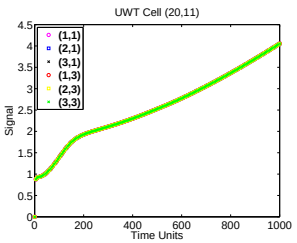


Fig. 12. Unbiased model, continuous simulation for a representative of cluster 3 (see Fig. 9).

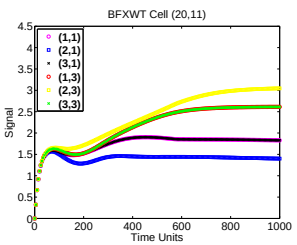


Fig. 13. BFXWt model, continuous simulation for a representative of cluster 3 (see Fig. 9).

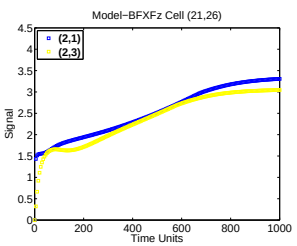


Fig. 14. BFXFz model, continuous simulation for a representative of cells distally adjacent to the Fz clone (cluster 10, see Fig. 10).

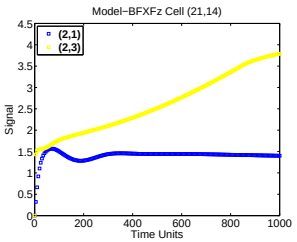


Fig. 15. BFXFz model, continuous simulation for a representative of cells proximally adjacent to the Fz clone (cluster 9, see Fig. 10).

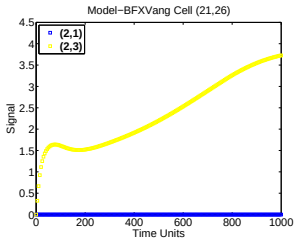


Fig. 16. BFXVang model, continuous simulation for a representative of cells distally adjacent to the Vang clone (cluster 12, see Fig. 11).

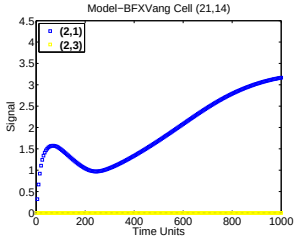


Fig. 17. BFXVang model, continuous simulation for a representative of cells proximally adjacent to the Vang clone (cluster 11, see Fig. 11).

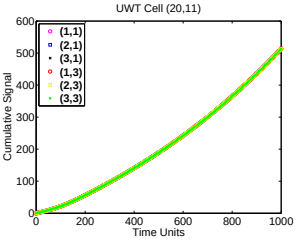


Fig. 18. Unbiased model, cumulative signal for a representative cell, relating to Fig. 12.

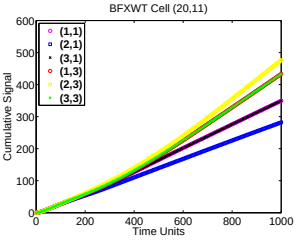


Fig. 19. BFXWt model, cumulative signal for a representative cell, relating to Fig. 13.

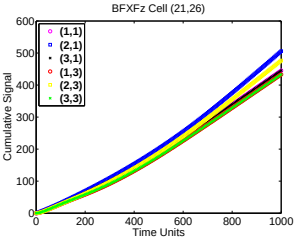


Fig. 20. BFXFz model, cumulative signal for a representative of cells distally adjacent to the Fz clone, relating to Fig. 14.

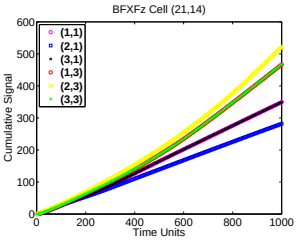


Fig. 21. BFXFz model, cumulative signal for a representative of cells proximally adjacent to the Fz clone, relating to Fig. 15.

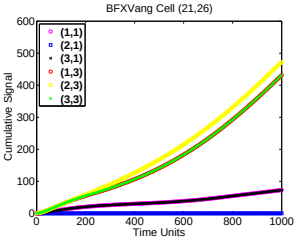


Fig. 22. BFXVang model, cumulative signal for a representative of cells distally adjacent to the Vang clone, relating to Fig. 16.

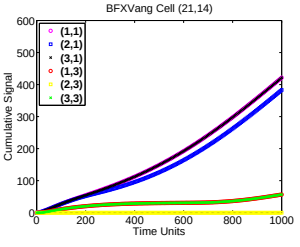


Fig. 23. BFXVang model, cumulative signal for a representative of cells proximally adjacent to the Vang clone, relating to Fig. 17.