Dynamic Landscape Analysis of Cell Fate Decisions: Predictive Models of Neural Development From Single-Cell Data

Marine Fontaine[1,5], M. Joaquina Delas[2,5], Meritxell Saez[3], Rory J. Maizels[2], Elizabeth Finnie[2], James Briscoe[2,**], and David A. Rand[1,4,*]

[1]Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
[2]The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK
[3]IQS, Universitat Ramon Llull, Via Augusta 390, 08017 Barcelona, Spain
[4]Zeeman Institute for Systems Biology and Infectious Epidemiology Research, University of Warwick, Coventry CV4 7AL, UK
[5]These authors contributed equally
[8]Lead contact
[*]Correspondence: D.A.Rand@warwick.ac.uk
[**]Correspondence: James.Briscoe@crick.ac.uk

SUMMARY

Building a mechanistic understanding of cell fate decisions remains a fundamental goal of developmental biology, with implications for stem cell therapies, regenerative medicine and understanding disease mechanisms. Single-cell transcriptomics provides a detailed picture of the cellular states observed during these decisions, but building quantitative and predictive models from these data remains a challenge. Here, we present *dynamic landscape analysis* (DLA), a framework that applies dynamical systems theory to identify stable cell states, map transition pathways, and generate a predictive cell fate decision landscape from single-cell data. Applying this framework to vertebrate neural tube development, we reveal that progenitor specification by Sonic Hedgehog (Shh) can be captured in a landscape with an unexpected circular topology, where initially divergent lineages converge through multiple routes. This landscape accurately predicted cellular responses and cell fate allocation for unseen dynamic signalling regimes. By modelling the dynamic responses that drive cell fate decisions, the DLA framework provides a quantitative and generative framework for extracting mechanistic insights from high-dimensional single-cell data.

KEYWORDS

Dynamical systems, single-cell genomics, machine learning, unstable manifolds, catastrophe theory, developmental biology, cell fate decisions, stem cell biology, neural tube, sonic hedgehog

INTRODUCTION

During embryonic development and tissue homeostasis, cells follow characteristic differentiation pathways to produce the diverse sets of specialised cell types that form functional tissues. This process is driven by gene regulatory networks (GRNs), which interpret external signals to induce cell type-specific gene expression programmes and, consequently, cell fate decisions[1]. Understanding how signalling inputs and GRNs coordinate these fate decisions remains a fundamental aim of biology.

Recent work has approached this problem by using dynamical systems theory to formalise Waddington's pioneering metaphor[2] of cellular differentiation as balls rolling down a landscape of branching valleys[3–23]. In this view, cell states correspond to attractors in a multidimensional landscape and transitions between these states follow paths, referred to as *unstable manifolds*, defined by saddle points[17,19]. Fate decisions emerge from signal-induced changes to the stability of these attractors and to the saddle-point connections in the landscape topology. This provides a precise mathematical characterisation of the states and transition routes that links observed cellular behaviour with the underlying stochastic process defined by the GRN. While this approach offers powerful insights, constructing quantitative landscape models from experimental data remains challenging. In particular, there is a critical need for methods that bridge the gap between high-dimensional single-cell data and quantitative models of cell fate dynamics.

Here, we present a systematic framework - *dynamic landscape analysis* (DLA) - for constructing quantitative models of developmental landscapes from single-cell data. We apply this to an in vitro system in which mouse embryonic stem cells are directed to differentiate into distinct ventral neural tube progenitors in response to defined Sonic Hedgehog (Shh) signalling regimes[24–26] (Fig. 1A). Building on our previously developed methods for constructing quantitative landscape models from flow cytometry data[18], we developed an approach to identify cell states and transition pathways in high-dimensional scRNA-seq data. This approach controls data dimensionality without relying on non-differentiable dimensionality reduction techniques. DLA enabled identification of cell states, explicit handling of temporal information, determination of bifurcations driving branching decisions, and investigation of responses to varying Shh levels. Our analysis uncovered the global decision making network and revealed an unexpected circular topology in the decision-making landscape, with initially divergent lineages ultimately converging. Branch points were driven by global bifurcations, such as flip bifurcations[19], enabling cells to split between two distinct commitment paths. Using this information, we constructed and parameterised mathematical models that reproduce observed behaviours and predicted unseen scenarios. Select predictions were experimentally validated.

This perspective views the underlying GRN as a generative process for cell fate decisions, and models the dynamics of this process rather than the details of molecular interactions within the network. In doing so, the framework provides a mathematical link from gene regulation to cellular behaviour by characterising the generative process of cell fate decision-making. On a practical level, by constructing a landscape model and identifying the network topology and bifurcations underlying the decision branching, the generative models produced should greatly assist the construction of relevant GRN models.

RESULTS

Overview of Dynamical Landscape Analysis

We established a framework to construct quantitative landscape models of cell fate decisions from single cell transcriptomic data. This systematic, multistep approach first identifies *attractor clusters* (ACs) representing stable cell states using a bespoke supervised clustering algorithm, which begins with known marker genes and expands to identify genes associated with both stable states and transitions between states. We then characterise transitions between ACs by tracking gene expression changes along approximated unstable manifolds. This geometric analysis validates the bifurcations underlying cell fate decisions and determines the branching network topology. Using this topology, we then construct a parameterised stochastic mathematical model of the global developmental landscape. Fitting the model parameters to experimental data accurately captured cell fate decisions across various experimental protocols. We use the model to make predictions about cell fate decisions under different signalling conditions.

To prototype this approach, we took advantage of an in vitro system in which pluripotent mouse embryonic stem cells were differentiated into distinct neural progen-

itor subtypes in response to Sonic Hedgehog signalling (Shh)[26–28]. The system uses SAG, a Shh signalling agonist, to induce ventral neural identities (Fig. 1B). We analysed a single-cell RNA sequencing dataset (scRNA-seq), comprising approximately 40,000 transcriptional profiles across ten time points, generated in response to 500nM SAG treatment[28]. Analysis using an iterative semi-supervised clustering approach, working outward from established marker genes, confirmed the presence of ACs representing the expected neural progenitor subtypes. We additionally identified some previously uncharacterised or unappreciated cell states and the transitions between them. We then characterised the branching decisions in terms of well-understood dynamical systems bifurcations. We then extended our analysis using flow cytometry data from cells exposed to varying SAG concentrations[26] to determine the ways in which the landscape changed as a function of morphogen level. Finally, to provide a synthesis of these results and to facilitate a predictive approach, we constructed and fitted a predictive generative landscape model.

Additionally, to evaluate the broader applicability of the DLA approach, we analysed neural progenitor scRNA-seq data from human embryonic stem cells organoids[29]. The analysis revealed conserved decision-making architecture between mouse and human, with the dynamical landscape model accurately predicting cell state proportions despite species-specific differences in developmental timing. This cross-species validation demonstrates the fundamental nature of the identified landscape topology in neural progenitor specification.

Identification and Validation of Attractor Clusters in scRNA-seq data

The first step in DLA is the identification and characterisation of ACs. These represent the distinct cell states of the system. Based on prior knowledge of ventral neural tube development[26–28,30,31], we hypothesised that the ACs in our dataset would consist of: neuromesodermal progenitors (NMPs)[24]; the PreNeural state that links NMPs to neural fates[32]; the progenitor subtypes found in the ventral neural tube: p0/p1, pMN, p3, and floor plate (FP)[31]; and mesoderm cells also generated by NMPs[24]. We reasoned that cells in the p0/p1 and p2 states would be rare in the dataset as these cell types are only generated in high numbers at lower concentrations of SAG[26,28,33].

We developed a semi-supervised clustering approach that we termed *outward clustering*. Starting with a small set of marker genes characteristic of known cell states, we iteratively identified genes that defined and discriminated ACs or transition routes (SI Sect. S2.1). We illustrate this method on the states and transitions found between days D5 to D7 (Fig. 1C). We began with known markers Olig2 (pMN)[34], Nkx2.2 (p3)[35], and Shh (FP)[36] and identified samples of cells expressing only one of these markers. Genes that were differentially expressed between each pair of these one-marker samples were identified, producing gene modules of size around $d = 100$, defined by setting a significance threshold (Fig. 1D). These genes were used to to find and characterise the ACs and transitions routes (see below). The significance level was set at a level that ensured the same ACs were found when the level was relaxed (SI Sect. S2.1).

We refer to the corresponding d-dimensional space of gene expression as the *effective gene space*. We then used PCA to project the d-dimensional expression data for these genes into a lower-dimensional space (the *clustering space*, with dimension chosen using a SVD elbow plot (typically 5-7, SI Sect. S2). We applied Leiden clustering[37] at high resolution (SI Sect. S2.1) to this, afterwards merging clusters with similar gene expression profiles (Fig. 1C and SI Sect. S2).

The use of well-defined gene modules enabled the robust identification of ACs despite the inherent noise in single-cell RNA sequencing data. We validated the robustness of the clustering approach by substituting alternative genes correlated with the markers and checked that this yielded equivalent cluster structures (SI Sect. S2.2). Additionally, to ensure that no clusters were missed, since every cell belongs to a cluster, we examined any clusters with low or no expression of the chosen markers at the various time points. These cells underwent standard quality control assessment, followed by differential gene expression analysis between clusters. When we identified relevant differentially expressed genes, such as transcription factors, we incorporated them into our marker set and evaluated their impact on clustering outcomes. In this example, the gene modules obtained using the three markers were sufficient to identify all the ACs, and adding more markers to generate the gene samples slightly improved resolution without uncovering any new cell states.

Using this method we identified ACs and clusters containing cells transitioning between them (summarised in Fig. 1E). We determined ACs and transition-

ing clusters for early time points, D3 to D4 (Fig. 1F-H) using this methodology with the markers TBXT/BRA, Foxc2, Foxa2 and Olig2 that is expressed early due to high SAG concentration (SI Sect. S2.2).

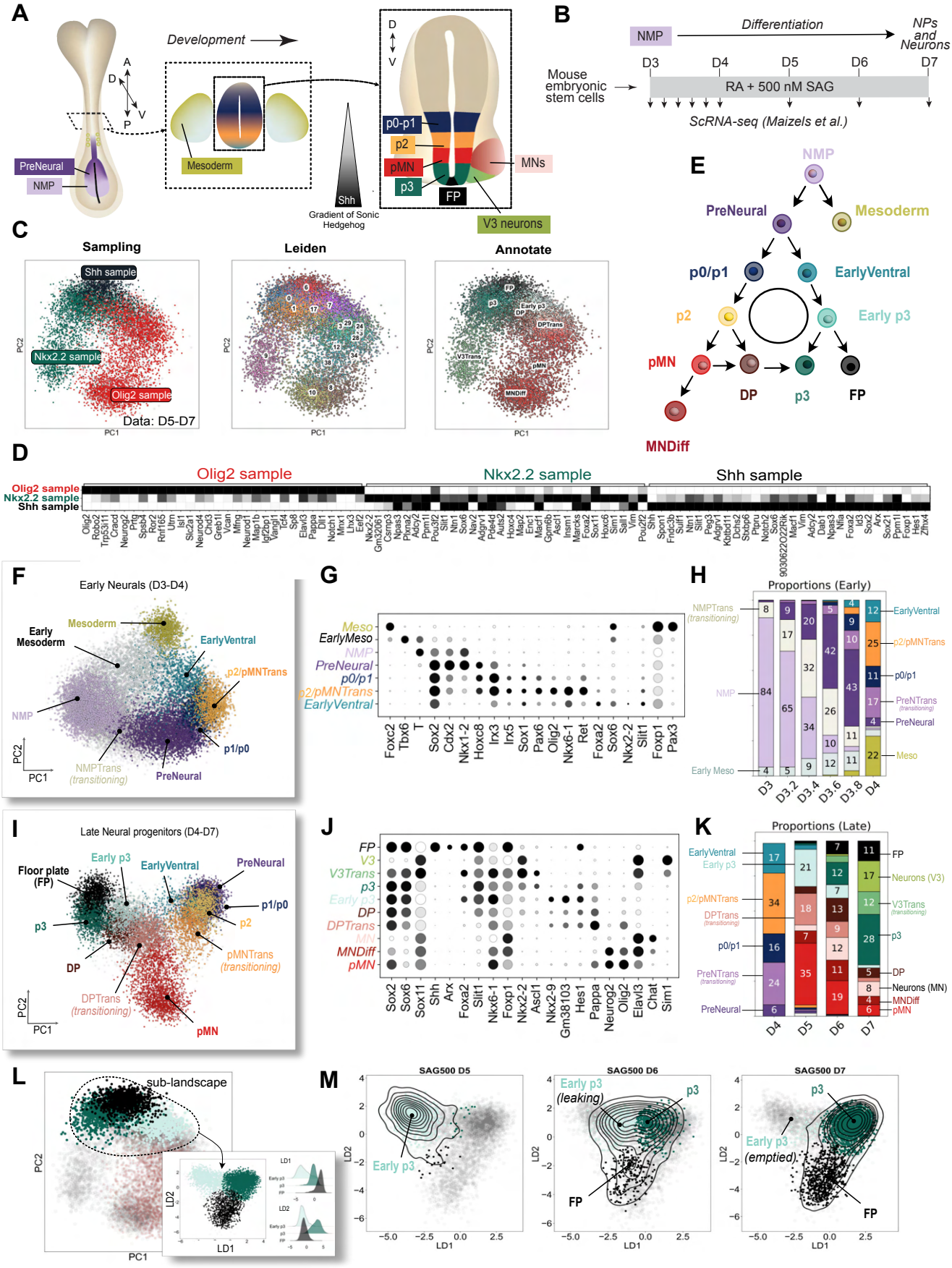
The identified ACs included: NMPs (characterised by TBXT/Bra expression[38]); mesodermal cells (marked by Foxc1 and Foxc2 expression[39]); and PreNeural cells (defined by co-expression of Nkx1.2, Msx1 and Sox2[40]). We found a transitioning cell population (NMPTrans) and Early Mesoderm cells expressing Tbx6[41,42]. As differentiation progressed (Fig. 1I-K), we observed the emergence of ventral neural progenitor states that include pMN cells, characterised by Olig2 expression; p3 cells, marked by Nkx2.2 expression; and floor plate (FP) cells, identified by co-expression of Foxa2, Arx and Shh. ACs representing p0/p1 (Pax6, Irx3) and p2 (Pax6, Irx3, Nkx6.1) were also recognisable but were only present in low numbers at D3.8 and D4. This is consistent with the high level of Shh signalling used, which promotes the production of more ventral cell types (pMN, p3, and FP) over more dorsal identities (p0-p2)[26,28].

The analysis also revealed, within the Olig2-expressing population, a group of cells (labeled "MN-Diff") coexpressing Olig2 and Neurog2[43] (Fig. 1J), likely representing pMN progenitors undergoing terminal differentiation to motor neurons (MN)[33] (SI Fig. S3F). Additionally, we identified a distinct population coexpressing Olig2, Nkx2.9, and Hes1 (labelled "DP" for double positive), which may represent a state between p3 and pMN identities[44]. Moreover, in D3.8 and D4 samples, a population of cells (labelled EarlyVentral) expresses Nkx6.1 and Foxa2, likely representing the common precursors of FP and p3 progenitors, consistent with recent findings[26]. At D5, we observed an AC (labelled Early p3) characterised by high expression of Nkx2.9, distinct from the conventional p3 state (Nkx2.2, Ascl1) observed in the D6 and D7 samples. We also identified V3 neurons, co-expressing Sim1 and Nkx2.2, representing the differentiated progeny of p3 progenitors[35] (SI Fig. S3F). Some intermediate populations were classified as transitioning (labelled using -Trans) as they appeared only transiently and did not form well-defined distributions that persist across time points (SI Sect. S2.3).

For visualisation and further analysis we used linear discriminant analysis (LDA)[45]. This provides a linear projection of $k+1$ clusters into k dimensions that optimises cluster separation by minimising within-cluster variance while maximising between-cluster variance (SI Sect. S2.4). Importantly, this projection preserves the geometry and fine-scale structure of the local landscape. It also allows a rigorous determination of the separation of the ACs and other clusters in the full effective gene space because if clusers are separated in the LDA projection in terms of Kullback–Leibler divergence then, by the data processing inequality, the same is true in the effective gene space.

We validated the ACs using multiple criteria. First, to check the ACs had different gene expression profiles, we performed a Wilcoxon rank-sum test to assess whether genes characteristic of known progenitor populations could distinguish between these ACs. Second, we checked that the ACs were well-separated in the projection and that, for each linear discriminant component, the distribution across cells in the AC of the corresponding LDA score was unimodal, indicating well-defined gene expression profiles for each AC (Fig. 1L) consistent with the idea that dynamics in the vicinity of attractors are approximately linear[46,47]. Third, we studied the temporal progression of the AC and in particular its loss of stability and the consequent transition of cells to adjacent attractors consistent with previous studies. For example, the way the Early p3 state loses its stability with cells transitioning to either p3 or FP is consistent with published data[48] (Fig. 1M). To do this we extracted groups of neighbouring ACs and transitioning clusters (that together form a structure we termed a *sub-landscape*) and projected them using LDA into either two or three dimensional space to analyse the geometry of the ACs and the temporal changes in the occupation of ACs. Finally, we examined the gene-gene correlations within each AC and assessed differences between clusters. Different correlations between clusters provided evidence for distinct regulatory interactions in each AC (SI Fig. S3D).

The complete characterisation of the identified ACs, including their molecular signatures and temporal progression, is presented in SI Sect. S2.3 and SI Fig. S2-S3. Here we focus on using these validated clusters as the foundation for constructing a quantitative model of the developmental landscape governing ventral neural tube patterning. Together, these findings highlight the power of single-cell transcriptomics to reveal the fine-grained cellular heterogeneity and developmental intermediates.



Attractor Clusters in Single-Cell RNA Sequencing Data Reveal Discrete Cell States During Neural Development.

Figure 1: (Caption next page.)

Figure 1: Attractor Clusters in Single-Cell RNA Sequencing Data Reveal Discrete Cell States During Neural Development.

A. Schematic diagram of neural tube patterning regulated by the Shh gradient.

B. Schematic of in vitro differentiation of mESCs and timing of the scRNA-seq data collection (D, Day).

C. Example of the outward clustering methodology illustrated by the analysis of the late time points (D5-D7). (Left) 2D PCA projection of cells that express one of the marker genes Olig2, Nkx2.2 and Shh. (Middle) PCA projection from d-dimensional effective gene space using the d most significantly differentially expressed genes (see SI Sect. S2.1) between the 3 samples of cells (here $d = 96$). Cells are clustered using Leiden clustering with fine resolution and coloured accordingly. (Right) The fine Leiden clusters are merged based on their molecular identities.

D. The most significantly differentially expressed genes distinguishing the 3 samples of cells shown in C (left panel).

E. Schematic diagram summarising all the ACs and possible transitions between them identified in the scRNA-seq dataset at 500nM SAG across all time points (D3-D7).

F.-H. Identification of early ACs (D3-D4) and transitioning cells between them labelled with -Trans.

F. PCA projection of cells from D3-D4 samples showing the early ACs.

G. Dot plot indicating the expression of markers used to identify the early ACs. The size of the circle represents the percentage of cells in the AC and the intensity of shading indicates the mean expression of the gene.

H. The proportion of cells (in percent) in each AC and transitioning clusters at the early time points (D3-D4).

I.-K. Identification of late ACs (D4-D7) and transitioning cells between them, labelled as -Trans. Same figures as in F.-H. using cells from the D4-D7 samples. The D8 sample is not displayed, it predominantly contains MNs, V3 neurons, MNDiff cells, and p3 cells.

L. Extraction of the group of adjacent ACs, Early p3, p3 and FP, that comprise the ventral sub-landscape, visualised in 2D LDA space (see also SI Fig. S3A-B). Histograms for each linear discriminant score (i.e the dot product of the data vector with the ith LDA eigenvector) show that ACs (such as FP) exhibit a well-defined approximately Gaussian structure in LDA space.

M. Temporal progression of the ACs within a sub-landscape comprising Early p3, p3, and FP identifies Early p3 as the initial ("head") attractor from which cells transition to either p3 or FP. Grey dots represent all cells in the group (D5-D7), coloured dots indicate cells present at the indicated time points with the colour indicating their cluster assignment. Contour lines indicate the density of cells at each time point.

Analysis of transitions using scRNA-seq

To map transitions between ACs, we developed *unstable manifold transition analysis*, a simple geometric computational method that provides a framework for understanding how cells navigate fate decisions during development (SI Sect. S3). This analysis identified three distinct transition types: direct transitions, binary flip decisions and binary choice decisions (SI Sect. S1.3). *Direct transitions* commit cells to a single new fate through fold bifurcations (aka saddle-node bifurcations) in which an attractor is eliminated by collision with a nearby saddle (Fig. 2A).

Binary flip decisions occur at branching points where cells choose between two distinct fates through flip bifurcations (Fig. 2B-C, SI Sect. S1.3). They involve a head AC from which cells transition to one of two downstream ACs along distinct routes after escaping through either a fold bifurcation of the head attractor or stochastic fluctuations over a nearby saddle. The route taken by a cell is determined by which a downstream attractor is connected to the head attractor via the unstable manifold of the saddle located close to the head attractor. Transitioning cells travel close to this unstable manifold. A change in signals or the cell's internal state can cause this connection to flip to the other downstream AC, changing the cell's fate. This is known as a flip bifurcation[18].

Binary choice decisions[18] also occur at branching points but when an attractor has two saddles close by and it can be destabilised by either saddle colliding with it. The cells then transition away from it along a route that depends upon which saddle was involved. (SI Sect. S1.4).

In both binary flips and binary choices, individual cells follow one of the two routes but, at the population level, cells may take both routes. This population level behaviour can arise for a number of reasons: (i) intrinsic cellular heterogeneity; (ii) heterogeneity in the signals that cells receive, resulting in the landscape differing between cells in a population; (iii) the escape route passing close to the saddle separating the two target attractors facilitating the distribution of transitioning cells into the two attractors via stochastic fluctuations (SI Fig. S1C). For such a binary flip decision, the landscape made up of the three attractors involved, the saddles and the stable and unstable manifolds of the saddles is called a *binary flip landscape*.

The transition or escape route from the destabilised AC is the unstable manifold of the saddle involved in the destabilisation. This is a 1-dimensional path between the attractor being destabilised and a destination attractor that will capture the state of the transitioning cell.

The decision from NMP to PreNeural or Early Mesoderm fates[24,25] exemplifies the binary flip (see also the PreNeural decision at D3.8-D4 (SI Fig. S5A)). Linear discriminant analysis revealed a clear separation between NMP, PreNeural, Early Mesoderm and Mesoderm, with transitioning cells forming bridges between these populations. At D3.2, the bifurcation manifests as a binary flip with two distinct trajectories emerging from NMP: one pathway leads to PreNeural identity (neural progenitor precursors), while the other directs cells towards Early Mesoderm (D3.4). From the Early Mesoderm they directly transition to Mesoderm (D4) (Fig. 2D).

To validate the bifurcation structures, we first approximated the unstable manifolds in the effective gene space using spline fitting to the data between ACs (Fig. 2E and SI Sect. S3.1). The approximated unstable manifolds (one per route) in the effective gene space (SI Sect. S2.1), are then projected down onto smooth curves in the LDA space (Fig. 2F-G). There is then a one-to-one correspondence between the points on the curves in different LDA projections and we can consistently map the expression of a given gene along any of these projected curves. We thus refer to these projected curves as the unstable manifolds of the flip when we use these curves in visualisations.

We then compute the mean expression level of genes along tubular neighbourhoods of the unstable manifolds to track gene expression changes along each route (SI Sect. S3.2-S3.3). Since we can monitor the expression of any gene along the tubular neighbourhood, our transition analysis is capable of detecting genes that, although not differentially expressed between clusters, vary in a notable way during the transitions between ACs and therefore might be components of the relevant GRN. Therefore, in addition to the genes that define the effective gene space, we also tracked a curated list of about 150 relevant transcription factors.

This approach revealed distinct expression trajectories that quantitatively demonstrated the exit patterns from initial attractors and the routes emerging from bifurcated attractors (Fig. 2H and SI Sect. S3). In each case where a binary flip landscape was identified there

were transitioning genes that had significantly different behaviour along the two exiting routes, providing strong support for the binary flip landscape.

The temporal progression of cells transitioning along the two unstable manifolds of this flip landscape is illustrated in Fig. 2I, showing how the distributions of all cells within the tubular neighbourhood at different time points progress along the unstable manifold. A broader variance in the distributions (e.g. at D3.4) supports the hypothesis that cells may experience delays and do not transition simultaneously. This aligns with the idea of "stickiness" of attractors, where some cells are retained by shallow attractors while others progress more quickly along the unstable manifold.

A similar analytical strategy was used to analyse all bifurcations associated with branching decisions. The results demonstrate robustness across different projections and datasets, including for flow cytometry data (see below), and effectively captures the progression of gene expression changes during transitions.

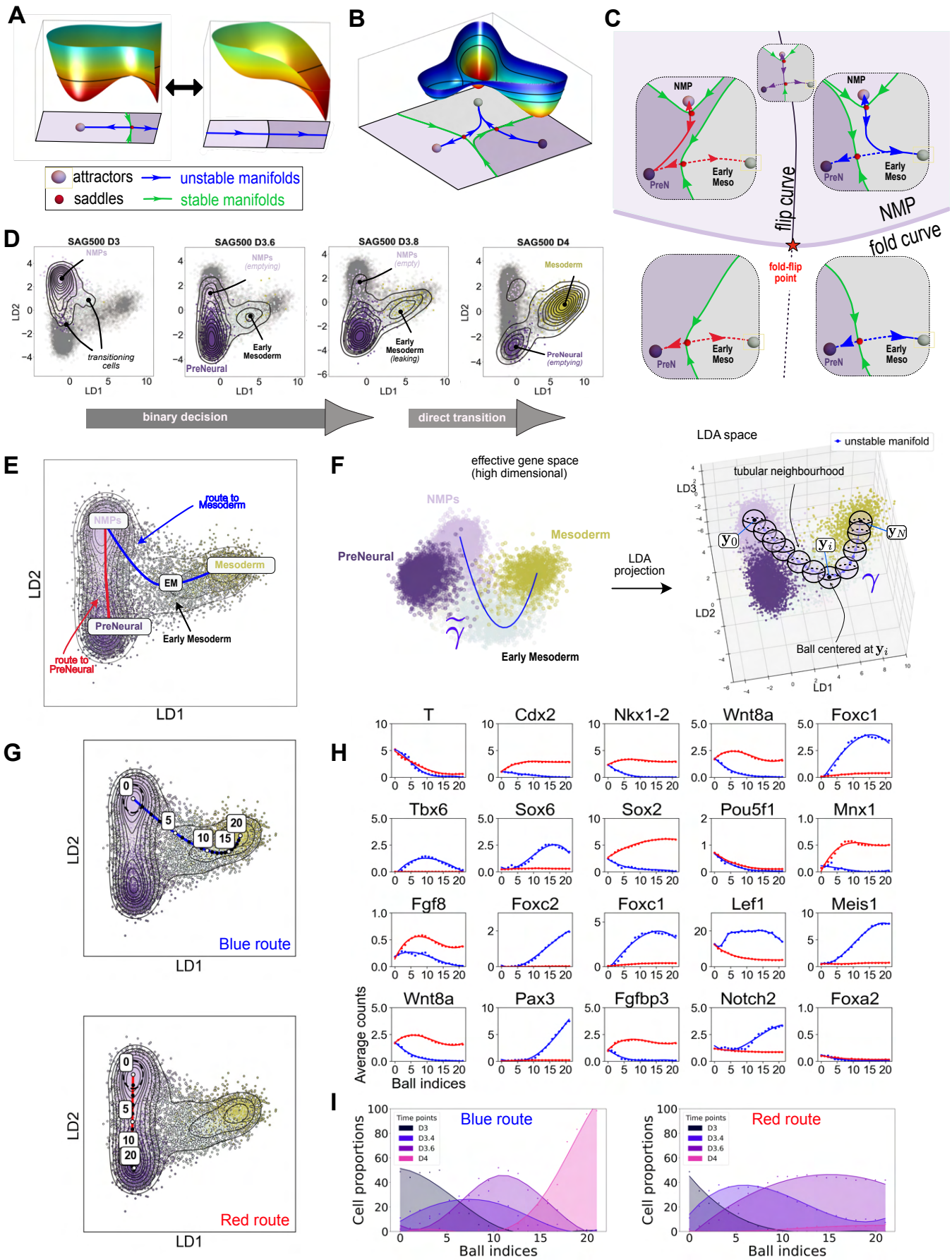
From the PreNeural state two distinct developmental trajectories emerge through a binary flip bifurcation (SI Fig. S5). One path leads cells to up-regulate Pax6, while the other generates the *ventral sub-landscape*, marked by the expression of Foxa2 and Nkx6.1 and the absence of Pax6. In the ventral sub-landscape (comprising the ACs EarlyVentral, Early p3, p3 and FP), cells progress through a defined sequence of states (SI Fig. S4). The EarlyVentral AC (Foxa2, Nkx6.1) serves as a common precursor for both p3 and FP states, consistent with p3 progenitors following a distinct regulatory programme from other ventral progenitors[26]. Cells first transition through EarlyVentral to an Early p3 state, which functions as a head state from which cells transition to either p3 or FP (Fig. 1M and SI Fig. S4B) again through a binary flip bifurcation. Transition analysis along these trajectories supports this bifurcation structure (SI Fig. S4F).

Once differentiated, PreNeural cells reach the p0/p1 AC they can directly transition to p2 and enter the *intermediate sub-landscape* (SI Fig. S6). At this SAG concentration (500nM), cells progress rapidly through p0/p1 and p2 states towards the pMN and DP states. At this junction, the landscape exhibits a rich structure involving another type of binary decision where pMN cells can transition to two distinct ACs (MNDiff and DP) via two distinct saddles. The choice of model for this decision is discussed below.

Further evidence for these bifurcations using the scRNA-seq data is given in SI Sect. S3.5. After p2 exit, those cells that did not proceed to MNDiff, showed coordinated upregulation of Olig2, Hes1, Rfx4 and Nkx2.9, indicating commitment to DP fate (SI Fig. S6C-E). The other population expressed Olig2 and Neurog2, indicating pMN specification. These pMN-specified cells progressed through a Neurog2-high MNDiff intermediate state before terminal differentiation into MNs[33]. Finally, we observed the progressive loss of DP cells by D7 converging towards p3 supporting a direct DP-to-p3 transition (SI Fig. S21B). Lineage studies support this connection: while FP cells rarely derive from Olig2-expressing precursors, a subset of p3 progenitors descend from cells with an Olig2 expression history[49]. This observation is compatible with either a flip bifurcation in which p2 is the head AC and pMN and DP are the target ACs or a binary choice decision in which pMN is the central AC and MNDiff and DP are the target attractors. The choice between these is difficult but the fitted mathematical model discussed below clearly chooses the binary choice (SI Sect. S3.5, SI Fig. S6C).

Examining gene expression changes along distinct transition routes revealed that the exit from a head AC at the origin of a decision is marked by the down-regulation of genes that define this AC and the simultaneous up-regulation of markers associated with the target ACs. As cells commit to a final fate, one group of markers prevails over the other, directing the transition towards a specific fate while closely following the approximated unstable manifold.

Together, the method captured the gradual changes in gene expression during transitions and identified route-specific gene programs. This provided evidence for well-defined transition routes and critical fate decision points which are crucial information for constructing a quantitative model of the developmental landscape.



Analysis of Cell State Transitions Using Unstable Manifolds Between Attractors.

Figure 2: (Caption next page)

Analysis of Cell State Transitions Using Unstable Manifolds Between Attractor Clusters.

A. Illustration of a fold bifurcation with a potential. (Left) Cells remain in the attractor which sits at the bottom of the potential well. There is a saddle close, over which some cells might escape. The blue curve is the approximated unstable manifold of the saddle. Escaping cells will, subject to noise, travel along this. (Right) A change in parameters leads to the collision with the saddle and destruction of the attractor and saddle so that all cells escape. The blue curve is the unstable manifold along which cells escape.

B. An illustration of a binary flip landscape with a potential. Cells start in the left hand attractor and, if they escape, they transition along the blue approximate unstable manifold to the grey attractor at the top of the illustration. A change in the parameters can result in the unstable manifold flipping so that cells transition to the bottom right attractor (dark purple) instead of the top attractor (grey).

C. Bifurcation set of a binary flip landscape. When a change in parameters causes them to cross the fold curve, the NMP attractor undergoes a fold bifurcation (i.e. it collides with a nearby saddle and disappears). When a change in parameters causes them to cross the flip curve, the system undergoes a flip bifurcation in which the unstable manifold of the saddle close to the NMP attractor changes its connection, flipping from one attractor to the other. If the parameters are to the left of the fold-flip point (asterisk), then the unstable manifold (red) connects to the PreNeural attractor. If the system is to the right of the fold-flip point then the unstable manifold (blue) connects to the Early Mesoderm attractor. The stable manifolds (green) of each saddle partition the phase space into basins of attraction, each containing a single attractor and coloured accordingly.

D. Temporal progression of the NMP, PreNeural, Early Mesoderm and Mesoderm ACs in LDA coordinates. Grey dots represent all cells in the group, coloured dots indicate cells present at the indicated time points with the colours indicating their cluster assignment. Contour lines indicate the density of cells at each time point. This exemplifies the binary flip decision from NMP to PreNeural or to Early Mesoderm, followed by a direct transition from Early Mesoderm to Mesoderm. At D3 almost all cells (dots) are in the NMP AC. Cells in both PreNeural and Early Mesoderm (Tbx6$^+$) are evident at D3.6, and increase at D3.8. There is then a direct transition from Early Mesoderm to Mesoderm at D3.8 and PreNeural empties as cells transition to other states (see SI Fig. S2G).

E. The two unstable manifolds (red and blue curves) for the landscapes of cells projected into LDA space. The red approximate unstable manifold outlines the route to PreNeural and the blue unstable manifold the route to Early Mesoderm and Mesoderm.

F. (Left) The unstable manifolds are estimated in the higher dimensional effective gene space. This avoids artificial ordering of the points along the curve that might occur if, for example, the LDA components were strongly determined by a single gene. This curve is then projected into the LDA space of interest (2D or 3D). (Right) Balls are placed along the projected curve in LDA space. Together they form a tubular neighbourhood of the unstable manifold. Raw counts of individual genes are averaged over the cells inside each ball to determine how gene expression changes along the unstable manifold (SI Sect. S3). This construction ensures that the change in expression levels does not depend upon the particular LDA projection used.

G. LDA plots of the binary flip decision from NMP to either PreNeural or Early Mesoderm using D3-D4 data showing the projected unstable manifolds. The positions of the balls is shown by the index of their centers.

H. The variation of key genes along the estimated unstable manifolds from G. Genes such as Cdx2, Sox2, Nkx1.2 and Tbx6 show diverging behaviour along the two routes consistent with the expected behaviour in a binary flip landscape.

I. Temporal progression of cells along the projected unstable manifolds in G. Proportions of cells (dots) at stages D3, D3.4, D3.6, and D4 within each ball along the unstable manifolds connecting NMP to Mesoderm (left) and NMP to PreNeural (right). We used a high smoothing factor to obtain smooth densities. In the right panel, there are very few cells at D4 (pink density) because cells at this stage have exited the PreNeural AC (SI Fig. S2G).

The effect of changing Shh levels on the landscape

We next set out to understand how the Shh signalling level shapes the decision landscape, specifically how signal perturbations affect the ACs and the transitions between them. To this end, we identified five markers - Sox2, Pax6, Olig2, Nkx6.1, and Nkx2.2 - sufficient to classify most ventral progenitor populations[26]. Using these markers, we generated time resolved flow cytometry data using four SAG concentrations (0, 10, 100, 500nM) at days D4, D5 and D6 of differentiation (Fig. 3A). We conducted three independent experimental series to ensure reproducibility. We selected the neural progenitors (Sox2$^+$) and used Gaussian Mixture Models (GMM) to identify clusters in the 5-dimensional marker gene space[18] (SI Sect. S4). Using statistical validation criteria analogous to those applied to the ACs from the scRNA-seq data (SI Sect. S4.3 and SI Fig. S8), we identified the major progenitor populations (Fig. 3A-B), analysed the local transitions between them in LDA coordinates (SI Sect. S5), and tracked their proportions over time across SAG concentrations (Fig. 3D). The proportions and the molecular composition for the transitioning populations for each experimental replicate are presented in SI Fig. S8B-C. As explained in SI Sect. S4.2, unlike the scRNA-seq dataset, MN and V3 neurons were not identified in the flow cytometry because only Sox2$^+$ neural progenitors were selected.

Changes in SAG concentration resulted in different proportions of cell types and affected the topology of the connections in the landscape (Fig. 3E-F). The branch from PreNeural towards the ventral sub-landscape, which was observed in the RNAseq dataset at 500nM SAG (SI Fig. S4A-B), is dependent on high SAG concentrations as no p3 cells were formed at 0-10nM SAG (Fig. 3G). In the flow cytometry data, the EarlyVentral AC emerges at 10nM with cells remaining in this state rather than transitioning further (see SI Fig. S11D Panel SAG10). We thus hypothesised that it was an AC rather than a transient state as it also appeared in the scRNA-seq data. This transition, and its behaviour under different SAG concentrations, is detailed in SI Sect. S5.2 and SI Fig. S11D-F for the flow cytometry data (see also SI Fig. S19C-D where a FP-like cluster was hypothesised). For the scRNA-seq data it is detailed in SI Sect. S3.1 and SI Fig. S3A-B.

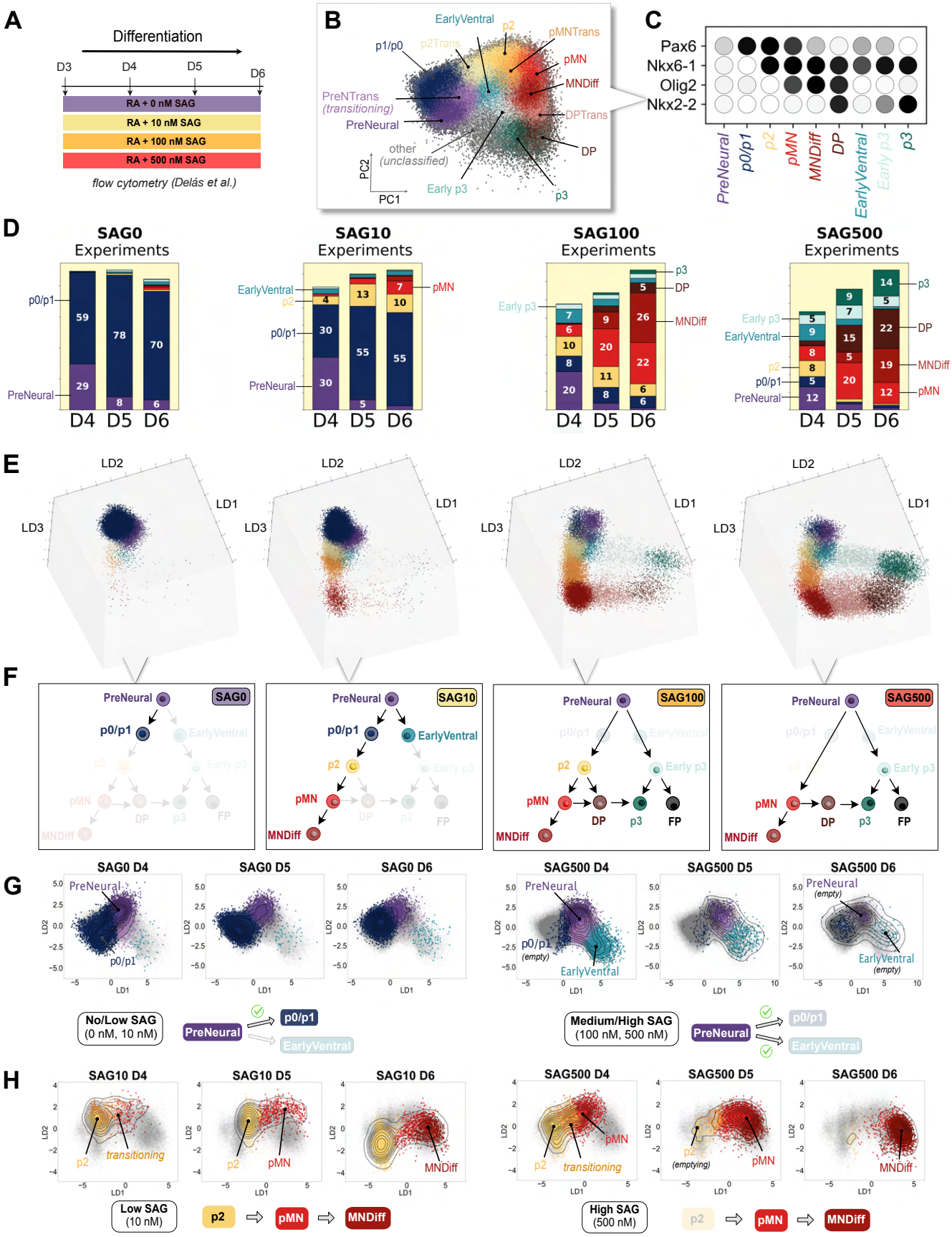
The AC containing exclusively Pax6$^+$ cells, labelled p0/p1, captures most cells at 0-10nM SAG but not at higher concentrations as Nkx6.1 and Olig2 levels rapidly rise within the Pax6$^+$ cells, leading to the emergence of p2 and pMN (Fig. 3G).

While the p2 AC is present and stable for all days at 10nM SAG, it appears to lose stability at higher SAG concentrations, with these cells instead adopting a pMN identity through up-regulation of Olig2 (Fig. 3H) or a DP identity through up-regulation of Olig2 and Nkx2.2 (SI Fig. S12F). The pMN AC is populated by a few cells at 10nM SAG but increased occupation of pMN and MNDiff is apparent at 100-500nM SAG. By contrast, the Nkx2.2-expressing ACs DP and p3 required a concentration above 100nM SAG to appear (Fig. 3D-F).

Although we can distinguish the Early p3 AC from the p3 AC, using the limited markers available we cannot distinguish FP, which is expected to co-express Sox2 and Nkx6.1 and to be present at D6. Nevertheless the scRNA-seq data indicate it should be present at 500nM SAG. Indeed, when the models are fitted to the scFACS data, they predict the presence of FP cells with an appropriate population size comparable to that of the scRNA-seq dataset at 500nM SAG. We therefore included it as part of the schematics (Fig. 3F).

The analysis of the transitions in each individual sub-landscape using LDA projection, as previously applied can be similarly performed for the flow cytometry data and is detailed for all the identified sub-landscapes and transition routes in SI Sect. S5 and SI Fig. S11-S12. In this context, the approach is particularly informative, as varying morphogen concentrations reveals critical bifurcations as exemplified in Fig. 3G-H. These include the presence or absence of attractors, suggesting that a change in signal level may have triggered a fold bifurcation. The emergence of alternative transition pathways, such as the ventral branch (EarlyVentral, Early p3, p3) emerging from PreNeural, indicates the occurence of a flip bifurcation.

Together, the analysis of time-resolved flow cytometry data confirmed the transition routes identified in the scRNA-seq data and examination of cell types generated at different SAG concentrations revealed the presence of specific SAG-dependent bifurcations. By combining transcriptional profiling with protein-level measurements and statistical validation, we established a comprehensive map of stable cell states and their connecting transition paths as a function of SAG concentration (Fig. 4A).



Flow Cytometry Analysis of Cell States Reveals SAG Concentration-Dependent Neural Progenitor Specification.

Figure 3: (Caption next page)

Flow Cytometry Analysis of Cell States Reveals SAG Concentration-Dependent Neural Progenitor Specification.
A. Schematic of the signalling conditions used for flow cytometry.
B. 2D PCA projection of the ACs found using flow cytometry data. Data for all SAG concentrations and time points are combined.
C. Dot plot indicating the expression of markers used to characterise the ACs in B (for the other clusters see SI Fig. S7B). The Floor Plate cluster (FP) was not identifiable in this dataset (see discussion in SI Sect. S5.2).
D. Percentage occupancy of the ACs as a function of Day (D) and SAG concentration (0, 10, 100, 500nM SAG). This shows the average proportions over 3 experimental replicates for 10nM and 100nM and 5 replicates for 0nM and 500nM (SI Sect. S4). The blank spaces in the bar plots correspond to the proportions of transitioning and unclassified cells and these are shown in SI Fig. S7C for all replicates.
E. 3D representation in LDA coordinates of the clusters at all time points (unclassified cells excluded). At 0nM SAG cells remain confined to the p0/p1 AC but start transitioning to p2 and pMN at 10nM SAG. The ventral cell branch (green/cyan dots) that includes EarlyVentral, Early p3, and p3 only appears at high SAG concentrations (100nM, 500nM).
F. Schematic diagrams of attractors present at different SAG levels and the connections between them. This represents the consensus state because a landscape is associated with an individual cell and these are heterogeneous in state and signal.
G-H. Temporal evolution of selected ACs (D4-D6) in LDA coordinates as a function of time and SAG concentration. Grey dots represent all cells in the group, coloured dots indicate cells present at the indicated time points with the colours indicating their cluster assignment. Contour lines indicate the density of cells at each time point.
G. Binary flip decision from PreNeural to either p0/p1 or EarlyVentral. The figure shows the concentrations 0nM, 500nM SAG (see SI Fig. S11A for other concentrations). Grey dots represent all cells in the group, coloured dots indicate cells present at the indicated time points with the colour indicating their cluster assignment. Contour lines indicate the density of cells at each time point. (Left) The p0/p1 AC is stable when SAG is absent (0nM) and there is no transition to EarlyVentral. (Right) The transition to EarlyVentral only appears at high SAG concentrations (100nM, 500nM). PreNeural, p0/p1, and EarlyVentral empty over three days at the highest SAG level (500nM). The attractors have bifurcated, but the unstable manifolds from PreNeural to pMN and Early p3 still pass through the regions in gene space formerly occupied by the bifurcated attractors.
H. Direct transitions from p2 to pMN and from pMN to MNDiff. The figure shows the concentrations 10nM, 500nM SAG (see SI Fig. S12A for other concentrations). (Left) At 10nM SAG, the p2 AC is relatively stable but shallow as cells can escape and transition to pMN over time. (Right) The same transition at 500nM SAG where p2 has already bifurcated showing more rapid occupation of the pMN AC.

Constructing a Global Landscape

The combined analysis of scRNA-seq and flow cytometry data enabled the precise characterisation of both stable populations and cells undergoing fate transitions, along with the effect of Shh signalling on the transitions. This revealed the complete branching structure of ventral neural progenitor specification (Fig. 4A).

We next set out to develop a mathematical model of this multi-decisions landscape. While the model should be based on an appropriately parameterised deterministic differential equation we required that it should realistically represent the stochasticity inherent in such molecular systems as well as the intrinsic heterogeneity in relatively uniform cellular populations.

The key step is the construction of the global deterministic model. For this we partitioned the system into overlapping local sub-landscapes, each comprising a small number adjacent ACs together with the transitioning cells between them (Fig. 4A-C). The use of such local sub-landscapes made it possible to use catastrophe and bifurcation theory to find appropriate normal form dynamical systems. This provides a highly flexible but minimally parameterised model that can accurately capture the behaviour in each local sub-landscape (e.g. Fig. 4C). These are then combined to produce a global model as described below. We assume that the cells follow trajectories within each sub-landscape governed by a dynamical system in normal form, which may vary between individual cells[10,50,51]. These variations can be encapsulated by differences in the parameter values within a parameterised family of dynamical systems with parameters regulated by biological signals[17] (SI Sect. S6 Definition S6.1). These parameters enable the model to capture the key bifurcations observed in the experimental data, including folds and flip bifurcations. A *landscape model* for an individual cell corresponds to a representative within this parameterised family of dynamical systems, defined by assigning specific values to the parameters. The bifurcation set together with stable and unstable manifold configurations for the parameterised family of the sub-landscape containing pMN is shown schematically in Fig. 4C. The other parameterised families used for the various sub-landscapes used in this study can be found in SI Fig. S13C-F and SI Fig. S14D-E.

These models are intentionally low-dimensional, despite the inherently high-dimensional nature of the underlying gene expression space. For this particular system, given the nature of the observed bifurcations and the topology of the transition routes, two dimensions are sufficient to encapsulate the core local dynamics governing each decision. In total, we identified five key sub-landscapes. Collectively, these describe the topology of the connections, attractor states, and their bifurcations that are capable of reproducing the developmental trajectories of this system.

The initial decision from NMP to either PreNeural or Early Mesoderm/Mesoderm can be modelled using a binary flip sub-landscape[18] after combining Early Mesoderm and Mesoderm into a single state (SI Sect. S6.1-S6.2 and SI Fig.S13). This model positions NMP as a shallow attractor close to undergoing a fold bifurcation, consistent with our observation that, by D3.6, cells have exited this state (Fig. 2D) and also consistent with previous findings[18].

The second decision, from PreNeural to either p0/p1 or EarlyVentral is another example of a binary flip sub-landscape (SI Fig. S13D). The data suggest a direct transition from p0/p1 to p2, characterised by a simple fold bifurcation (SI Fig. S8B).

Third, from the p2 state, cells transition towards either pMN or DP states, continuing along distinct developmental pathways (SI Fig. S6C and SI Fig. S12F-H). To accommodate both direct DP emergence from p2 cells and later emergence from pMN cells, we propose a parameterised family derived from the A_7 elementary catastrophe[52] that contains a binary choice bifurcation (Fig. 4C). This captures the key transitions: a direct shift from p2 to pMN as cells exit the p2 state, a transition from DP to p3, and direct transitions from pMN to MNDiff and DP, facilitated by the presence of a dual cusp in the bifurcation set. This cusp is a key component of the binary choice bifurcation. If the parameters are close to the cusp point the two saddles either side of the pMN attractor are close to it and at the cusp point they collide with it. This is further discussed below and illustrated in Fig. 5E-F.

The fourth sub-landscape is a binary flip from Early p3 to either p3 or FP (Fig. 1M and SI Fig. S14F).

Finally, our analysis indicates that cells populate the p3 AC via two distinct routes: a route to p3 through the ventral sub-landscape, passing through the EarlyVentral state, marked by markers such as Foxa2 and Slit1, and another route to p3 through the intermediate sub-landscape where cells transition through pMN and DP before finally reaching p3 at later time points (D6-D7)

(SI Fig. S21B-C). To represent this dual pathway, we combined the ventral and intermediate routes to p3 as part of a circular sub-landscape, the parameters of which can control the fold bifurcations of p0/p1, p2, EarlyVentral and Early p3 (SI Fig. S14D). This topology necessitates the presence of a central repeller and is further tested and validated below.

We then used techniques from differential topology [53] to merge the phase spaces of these parameterised local models along their decision regions to allow smooth transitions (Sect. S6.3), combining the modular "LEGO block" approach [15,54] with sigmoid techniques that seamlessly connect them through their shared attractors [18,19]. This resulted in a global parameterised family of dynamical systems which models the global landscape, $\dot{\mathbf{x}} = \mathcal{L}(\mathbf{x}, \boldsymbol{\theta}^{\mu}, \mathbf{vel}^{\mu})$, where $\mathbf{x}$ is the state variable representing the state of a cell μ. The vector of parameters $\boldsymbol{\theta}^{\mu}$ contains all the sub-landscape parameters that control the configuration of the connections and the bifurcations of the attractors. The landscape depends on extra cell-dependent velocity scaling parameters used to adjust the duration of cell transitions between states (Sect. S6.4).

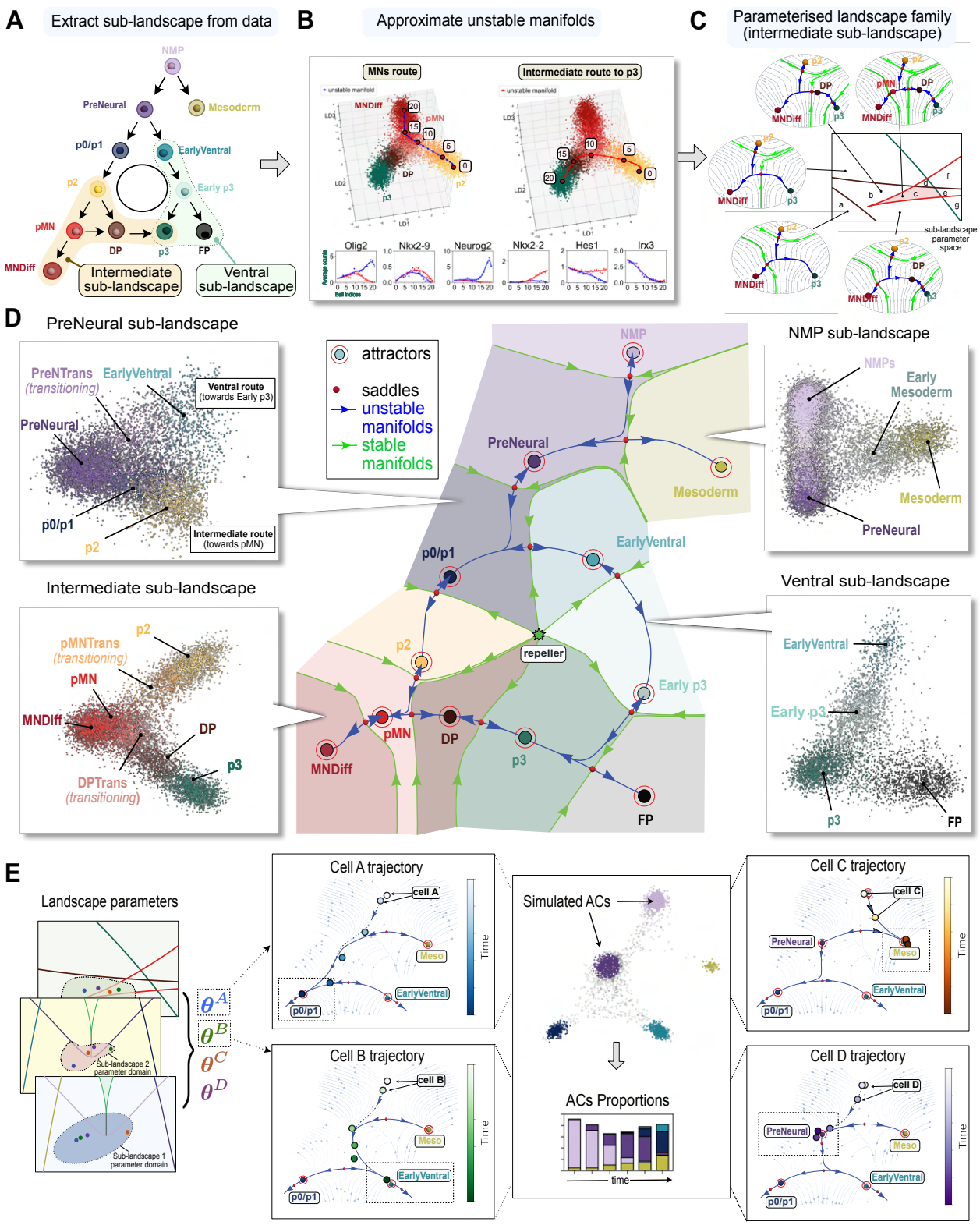
One particular global landscape instance for this global parameterised model that a cell might have, obtained by fixing the parameters, is shown in Fig. 4D. This model incorporates all attractors but only a subset of the transitions in Fig. 4A. Changing parameters produces all possible attractors and connection configurations for the cells. Cell specific landscapes can be envisioned by reversing selected flips and removing certain attractors through bifurcations, while maintaining cellular progression along the blue curve through these bifurcated states. We propose that this representation is the modern equivalent of Waddington's epigenetic landscape.

Stochastic simulations of the landscape

The trajectory of a single cell μ is governed by a stochastic differential equation of the form

$$d\mathbf{x}(t) = \mathcal{L}(\mathbf{x}(t), \boldsymbol{\theta}^{\mu}, \mathbf{vel}^{\mu}, t)dt + \eta(\mathbf{x}(t), \boldsymbol{\sigma}^{\mu}, t)d\mathbf{w}(t)$$

where the deterministic part $\mathcal{L}(\mathbf{x}(t), \boldsymbol{\theta}^{\mu}, \mathbf{vel}^{\mu}, t)$ is the global landscape supplemented by a state-dependent diffusion with amplitude $\eta(\mathbf{x}(t), \boldsymbol{\sigma}^{\mu}, t)$ in order to incorporate the stochastic nature of cellular dynamics (SI Sect. S7). While our model does not rely on known or inferred molecular interactions, when fitted to the data, it reproduces the observed behaviour as a function of the signalling state. For this reason we call it a *generative landscape model*. As we show below, it predicts the result of changes to signalling. Moreover, it will likely be helpful, perhaps crucial, for constructing a molecular model because a molecular model will have to reproduce the dynamics of cell fate choices and this will greatly reduce the arbitrariness of the mechanistic model and guide parameter fitting.



Global Landscape Model.

Figure 4: (Caption next page)

Global Landscape Model.
A. A schematic diagram of all the identified transitions between cell states. A group of neighbouring ACs and transitioning cells is extracted from the data and then projected in LDA space. We highlighted the ACs belonging to the intermediate (yellow shaded) and the ventral sub-landscape (green shaded).
B. The projection in 3D LDA space of the intermediate sub-landscape highlighted in A. Unstable manifolds are approximated using splines in the higher-dimensional effective gene space and projected onto the 3D LDA space. Gene expression variations along these pathways reveal consistent patterns of lineage splitting.
C. We model the intermediate sub-landscape in A and B, by a parameterised family of local vectorfields (SI Sect. S6 Definition S6.1). The configuration of the sub-landscape's attractors and unstable/stable manifolds depends upon the parameters and varies as parameters change in the different regions (marked by a-g) inside the rectangle. For the parameters controlling the transitions among the linear arrangement of the 4 attractors MNDiff, pMN, DP and p3, the bifurcation set consists of fold curves decorated with dual cusps in parameter space (one dual cusp in the rectangle).
D. Phase portrait of the global landscape for a fixed set of parameters when all attractors are present. The blue and green curves are respectively the unstable and stable manifolds of the various saddles. The green stable manifolds of the saddles separate the basins of attraction of the various attractors which are coloured accordingly. If the state of a reprogrammed cell falls in one of these coloured domains, it will converge towards the corresponding attractor and adopt that specific fate. In this figure we have arbitrarily chosen one direction for each flip bifurcation. 2D LDA plots of the ACs are shown for 4 of the 5 sub-landscapes.
E. Schematic of the simulation procedure, illustrated for the first two binary decisions (further transitions are not shown). For each cell μ, parameters for the individual sub-landscapes are selected within their respective parameter domains (two components in each 2D bifurcation panel represented by a dot coloured differently for each cell). Together, they form the components of a parameter vector $\boldsymbol{\theta}^{\mu}$ of the global landscape (illustrated here for 4 cells labelled A, B, C and D with parameters $\boldsymbol{\theta}^{A}$, $\boldsymbol{\theta}^{B}$, $\boldsymbol{\theta}^{C}$, $\boldsymbol{\theta}^{D}$). These parameters define the configuration of attractors, saddles and unstable manifolds of the cell's deterministic dynamics. Adding stochastic fluctuations to these dynamics, we simulate a stochastic trajectory for each cell, recording their positions in state space at discrete time points. Each of the panels on the left shows the deterministic (local) landscape for each cell determined by their respective parameters. In each panel, the evolving cell is coloured according to its time step. Different set of parameters produce different landscape models resulting in distinct possible fates for each cell (dashed rectangles in each panel). If an attractor is present in a cell's landscape (e.g PreNeural for cell D), the cell may remain trapped in this attractor as we see for the purple dots, indicating the position of the cell over time, staying near the PreNeural attractor. Expanding the parameter sampling allows the simulation of a large number of cells (central panel). Clustering the simulated cells identifies ACs (coloured according to the cell state they correspond to), and the proportion of cells in each AC at each time points were compared to the experimental data.

Fitting the model to experimental data

We used Approximate Bayesian Computation (ABC) particle filtering[55–57] to estimate the model parameters[15,18,54] by fitting to the experimental proportions averaged across the different experimental series for each fixed SAG concentration. This involves simulating the system, quantifying each simulated AC at different time points and comparing to the experimental proportions (Fig. 4E). For details see SI Sect. S8-S9. To fit the model to the scRNA-seq dataset (SI Sect. S11), we standardised populations by grouping neuronal states with their progenitors (MN with MNDiff, V3 and V3Trans with p3). By focusing the analysis on the neural progenitors, this enabled a direct comparison between experimental modalities.

In broad terms, the particle filtering algorithm works by finding a sequence of distributions of parameters that produce increasingly good agreement between the simulated and experimental proportions of cells in each AC at each time point (SI Sect. S8.3). Upon convergence, this gives the posterior parameter distribution (PPD) of parameter values that produced simulated ACs proportions that closely match the experimental proportions (Fig. 5A). This approach immediately produces predictions. For example, the flow cytometry dataset could not distinguish FP proportions but the fitted models to these data provided estimates of FP proportions that can be treated as model predictions, as described below.

A significant advantage of our approach is that, because the landscape is constructed using normal forms of the bifurcations, we are able to examine the bifurcation sets for each decision. This allows better understanding of how outcome depends upon the parameters and the relationship between the level of heterogeneity in the cell population to the level of heterogeneity in the outcome. These bifurcation sets are made up of the curves in parameter space which, when crossed, cause either a fold or a flip bifurcation. Their local structure is highly constrained and predictable from dynamical system theory[17]. The analysis identifies the critical points in the landscape that organise cell fate decisions. Flip and fold curves divide the parameter space into regions with the same attractors and same connection topology. These (and therefore the decision-making) change when such a curve is crossed. This reveals how specific signalling conditions control cell fate choices. From each parameter associated with a cell, sampled from the PPD, we can extract the informative sub-landscape parameters and project them onto the corresponding bifurcation sets. Comparisons of these projected PPDs for the flow cytometry samples exposed to different SAG concentrations revealed distinct patterns that successfully validated all the hypothesised bifurcations (SI Sect. S9.1).

To illustrate this, we show the PPDs that control the fold bifurcation of the p2 AC (Fig. 5B). The corresponding sub-landscape bifurcation set contains a curve of fold bifurcations (aka saddle-nodes). A parameter crossing the bifurcation curve results in the p2 attractor bifurcating, forcing the cell to leave p2 and adopt a pMN or a DP fate. At SAG 10nM, the parameters from the PPD fall on both sides of the fold curve. This indicates that the p2 attractor is present for some of the cells, but is close to bifurcation, and has bifurcated for the others. For the former, the attractors are shallow and stochasticity can result in cells escaping and transitioning to pMN. This leakage explains the low proportion of pMN cells observed at 10nM SAG (Fig. 3H). At higher SAG concentrations, the PPDs lie below the p2 fold curve, indicating that the p2 attractor has bifurcated for most cells, explaining the large increase in pMN cells and reduction in p2 cells in these conditions. While the 10nM SAG example demonstrates a heterogeneous set of cells producing heterogeneous behaviour, the higher SAG examples show a heterogeneous set of cells producing homogeneous behaviour. The PPDs in Fig. 5B are expected to vary smoothly as the morphogen changes and this defines a direction in parameter space. Generically, a second morphogen will similarly define a different direction. Therefore, such an approach allows one to understand the combined effects and dynamical systems theory constrains the structures that can occur.

A more elaborate example is the binary flip sub-landscape from the PreNeural state in which cells choose between the p0/p1 and EarlyVentral fates. The bifurcation diagram shows a flip bifurcation curve and a fold curve that intersect at a fold-flip point (SI Fig. S10). PPDs mapped onto this bifurcation set demonstrate distinct clustering patterns that explain the observed cell state transitions (Fig. 5C and SI Fig. S11A). At low SAG concentrations (0-10 nM), the PPDs localise away from and to the left of the flip bifurcation curve, driving transitioning PreNeural cells preferentially towards p0/p1. By contrast, at high SAG concentrations (100-500 nM), the PPDs concentrate near the fold-flip point, creating uncertainty in the PreNeural cell fate decision between p0/p1 and EarlyVentral (SI Sect. S1.3). This

parameter distribution explains the observed splitting at high SAG concentrations of the PreNeural transitioning cell population between p0/p1 and EarlyVentral.

All the sub-landscape bifurcation sets with the PPDs for each SAG concentration are discussed in SI Sect. S9.1 and presented in SI Fig. S18.

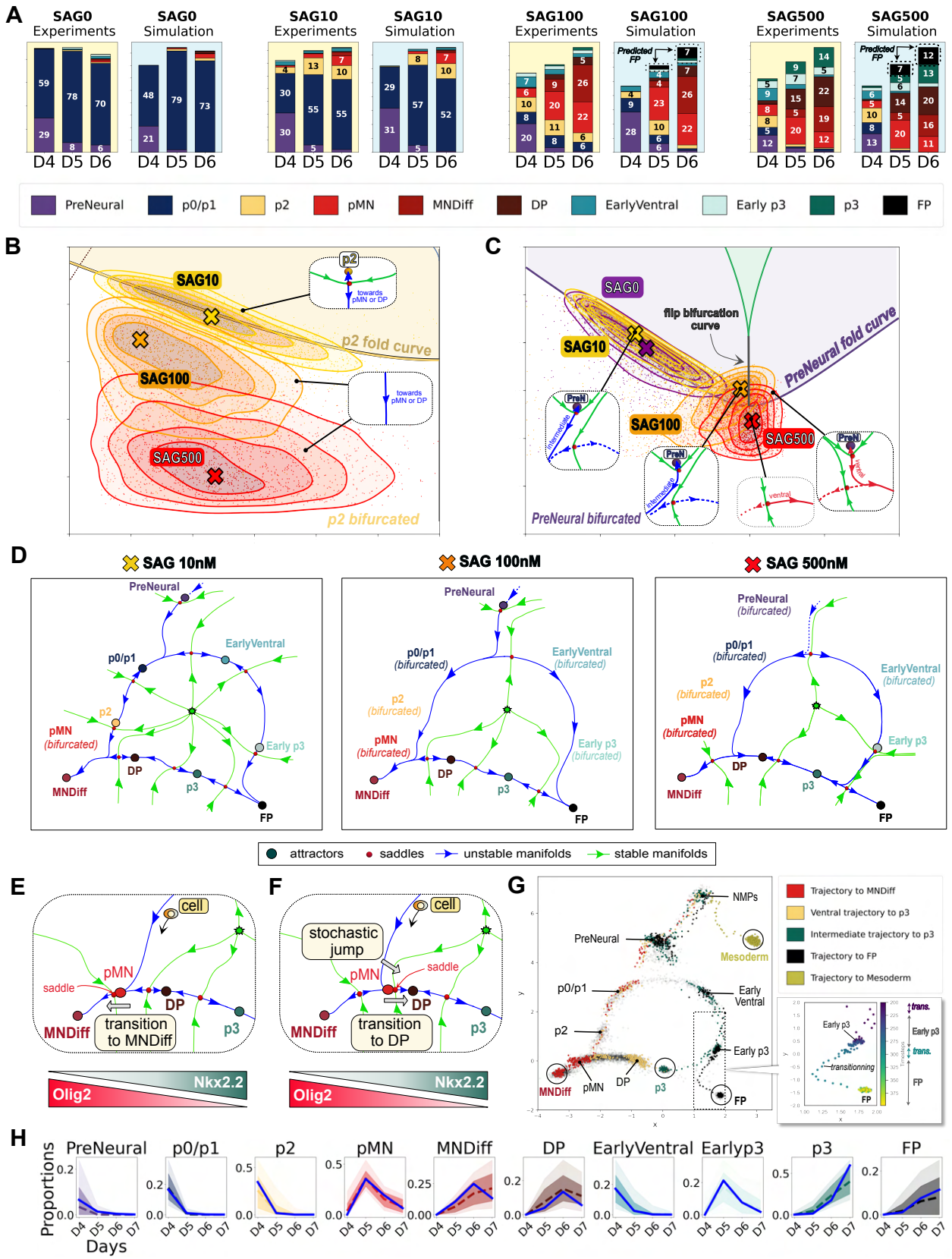
Integrating the full set of fitted parameters at each SAG concentration, including those for each sub-landscape, results in a global landscape model for each cell. Selected examples of the landscape for three cells are shown in (Fig. 5D). In these examples, the pMN attractor is bifurcated but for other cells with different parameter values it may be present (SI Fig. S18A Panel 5).

Our initial analysis of the data suggested that, in any model, the connections between MDiff, pMN, DP were on a single smooth curve and that therefore there should be saddles between them. Moreover, it suggested that the two saddles on either side of pMN could each be potentially close to pMN and therefore could both collide with pMN. We therefore modelled the binary decision from pMN to either MNDiff or DP by a sub-landscape featuring a dual cusp in its bifurcation set (SI Fig. S14E). This allows the pMN attractor to bifurcate through one of two distinct saddles, and the saddle selected determines the resulting cell fate (Fig. 5E-F).

The simulation procedure captures timing variations in transitions between cells, as illustrated in Fig. 5G, which shows one stochastic trajectory ending at each of the three ACs MNDiff, FP and Mesoderm, and the intermediate and ventral trajectories to p3. This is particularly informative as it shows how cells can remain trapped in shallow attractors for extended periods before escaping due to stochastic fluctuations, resulting in noisy transitions along unstable manifolds.

Fitting the model to the scRNA-seq data, discussed in SI Sect. S11, also gave good agreement with the experimental proportions of the neural progenitors (Fig. 5H). The resulting FP proportions at D6 were around 7%, consistent with the prediction (12%) obtained from flow cytometry at the same SAG concentration (Fig. 5A and SI Fig. S20A-B). Further validation using flow cytometry is given in SI Sect. S10 and SI Fig. S19E. We then compared the resulting PPDs for the progenitor landscape in the scRNA-seq with those of the flow cytometry at SAG 500 nM (SI Fig. S20F). As discussed in SI Sect. S11, the PPDs showed striking similarities across most sub-landscapes, despite cell states being identified based on protein concentrations in flow cytometry and mRNA counts in scRNA-seq, and minor differences in timing and experimental protocols. A noticeable difference however appeared at D7, where the scRNA-seq PPD crosses the DP fold curve (SI Fig. S20F Panel 5), forcing DP cells to rapidly transition to p3. This effect is likely due to the increased number of p3 cells in the scRNA-seq data, as these were merged with V3Trans and V3 neurons. The significant increase of p3 cells at late time points cannot be reached by simulation by a direct transition from PreNeural to the ventral sub-landscape. Instead, the PPD crossing the DP fold curve suggests that this increase results from cells transitioning to p3 via an route involving a direct transition from DP to p3. The ability of simulated cells to reach p3 through two distinct developmental pathways is facilitated by the circular structure of the model which we validate further below.

Overall, the model accurately captured the experimental dynamics (SI Video S1-S3), and successfully reproduced the observed bifurcation behaviours. This provides a quantitative and rigorous explanation for Shh signalling dependent patterning of ventral neural tissue and makes testable predictions.



Parameter Estimation and Bifurcation Analysis of the Dynamic Landscape Model.

Figure 5: (Caption next page)

Parameter Estimation and Bifurcation Analysis of the Dynamic Landscape Model.

A. Experimental proportions (yellow background) compared to the simulated proportions (blue background) after parameter fitting for each SAG concentration using the flow cytometry data. The predicted FP proportions (12%) obtained from simulations at 500 nM SAG are consistent with the FP proportions observed in the scRNA-seq at the same SAG concentration in H below (7%). The experimental proportions were obtained by averaging the proportions for 3 individual experimental series for 10nM, 100nM and for 5 experimental series for 0nM and 500nM (SI Sect. S4).

B.-C. Dots represents parameters accepted by the ABC particle filter. Contour lines are used to estimate the PPDs projected on the corresponding sub-landscape parameter space (purple: 0nM, yellow: 10nM SAG, orange: 100nM, red: 500nM). The other curves show the bifurcation set where fold and flip bifurcations occur for the different sub-landscapes. PPDs projected on the other sub-landscapes parameter space can be found in SI Fig. S18A.

B. Bifurcation set of the sub-landscape controlling the p2 transition. The yellow fold curve represents the set of sub-landscape parameter values where the p2 attractor is eliminated in a fold bifucation. The PPD at 10nM SAG lies on both sides of the fold curve, suggesting that the p2 attractor is present, but shallow, for half of the cells and bifurcated for the other half in which case cells transition to a downstream attractor. This demonstrates that the p2 AC remains stable over time, and the leakage to pMN observed in Fig. 3H occurs either because a cell escapes from the p2 attractor or because its parameter fall outside of the shaded yellow region of the bifurcation set where the p2 attractor is bifurcated. At higher SAG concentrations (100nM, 500nM) the PPDs lie outside the shaded region suggesting that the p2 attractor is bifurcated for all cells. The 0nM PPD (purple) is not shown, as the parameter did not converge due to the absence of p2 cells at this SAG concentration (the parameters are distributed across the entire region).

C. Bifurcation set of the sub-landscape controlling the binary decision from PreNeural to a downstream attractor modelled by a binary flip. Parameters on opposite sides of the gray flip bifurcation curve give rise to distinct sub-landscape representatives: one with an unstable manifold (blue) connecting PreNeural to the p0/p1 attractor or, if p0/p1 is bifurcated, to an attractor part of the intermediate sublandscape and the other (red) connecting PreNeural to the next attractor, part of the ventral sub-landscape (Fig. 4A), that is present on the way down. When the purple fold curve is crossed, the PreNeural attractor bifurcates. Upon bifurcation, the cell's fate is determined by the basin of attraction it was in, leading it to the downstream attractor belonging to this basin. At low SAG concentrations (0nM, 10nM) the PPDs are to the left of the flip bifurcation curve implying that most PreNeural cells transition towards p0/p1 or an intermediate state. At higher concentrations (100nM, 500nM) the PPDs cluster near the fold-flip point where the fold and the flip curves intersect. In this case, cells can transition to both intermediate and ventral states.

D. Examples of a cell's landscape for three specific set of landscape parameters. The sub-landscape parameter components are indicated in B and C by the yellow, orange and red crosses which are in a high density region of one of the PPDs. The pMN attractor appears bifurcated away for this specific set of parameters, but it is present for other parameter values and remains accessible to other cells (see other parameters in SI Fig. S18A Panel 5).

E-F. Examples of dynamical configurations suggested by the fitting for the p2 transition from p2 to pMN/MNDiff or DP obtained by fixing specific landscape parameters values. A cell (yellow) that has just left the p2 attractor transitions along the blue unstable manifold to reach the pMN attractor.

E. The unstable manifold connects p2 to pMN with the latter well-separated from the saddle between it and DP. The saddle between pMN and MNDiff is close to pMN and therefore the cell can transition from pMN to MNDiff.

F. In this case, the pMN attractor is near bifurcation with its right saddle, thus when a cell reaches pMN, it is likely to jump over this saddle and move towards a DP fate.

G. Five stochastic single cell trajectories followed for 400 simulation timesteps and coloured according to the eventual fate. Gray dots indicate other cell trajectories, included to help visualise the shape of the landscape. Notice that while pseudotime makes sense for the transitioning segments of a trajectory, it does not make sense for a full trajectory because cells spend random long periods near attractors.

H. Distribution of the simulated proportions (coloured by cell state) retained by the ABC particle filter compared with the experimental proportions for the scRNA-seq dataset (percent values in decimal form). For each cell state, the thick blue curve represents the experimental proportions and the dashed curve represents the mean of the simulated proportions. Shaded region indicates 95% confidence interval for the simulated proportions.

Testing Model Predictions Using Signalling Perturbations

To test the model and gain deeper insight into signal-dependent fate decisions we focused on the principal decision points: the transition from PreNeural to either the intermediate or ventral sub-landscapes; the transition from p0/p1 to p2 that then leads to pMN; and the connection from DP to p3 that is responsible for the circular topology of the landscape.

To examine the first two decisions, we recast the model into a simplified version that focused on these transitions. This reduced model aggregated several related cell states (combining NMP/PreNeural, EarlyVentral/Early p3/p3, and pMN/MN/DP) while preserving the essential topology of a binary flip landscape followed by a linear transition pathway (SI Sect. S12). In this framework, the ventral sub-landscape consists of a single state p3*, and the intermediate sub-landscape comprises two states: p2 and the combined pMN/MNDiff/DP state, relabelled pMN* (Fig. 6A). The model accurately reproduced the proportions of cell states observed in various SAG concentrations (Fig. 6C).

We then used the model to simulate the effect of timed pulses of Shh signalling on neural progenitor proportions. We first simulated the effect of 24h with no SAG followed by 48h with 500nM SAG, assuming the effect of changing signal level was instantaneous on the landscape. This predicted that at D4, after 24h with 0nM SAG, cells would adopt a p0/p1 state, at which point an abrupt increase to 500nM SAG would destabilise this state, driving the cells to transition further into the intermediate sub-landscape. Since the p2 state is absent at 500nM SAG, cells would rapidly adopt a pMN* identity. Importantly, because most cells are predicted to have left the PreNeural attractor and committed to the p0/p1 state by D4, few cells would be available to transition to the ventral sub-landscape. Consequently the model predicts a smaller fraction of cells adopting p3* fate at D5-D6, compared to the outcome under constant 500 nM SAG exposure.

We tested this prediction experimentally by culturing cells in the absence of SAG for 24h, from D3 to D4, and then adding 500nM SAG for the subsequent 48h (D4-D6). We analysed the proportions of cell types at D4, D5 and D6 using flow cytometry. The experimental data were in good agreement with the simulations, with the majority of cells adopting a pMN* identity at D6 (Fig. 6D and SI Fig. S22D).

We next simulated the opposite pulsing regime - 500nM SAG for 24h followed by 0nM SAG for 48h - to explore the impact on the landscape of an abrupt removal of SAG at D4. This scenario would address the state of the p2 and pMN* attractors at 0nM SAG because, at constant 0nM SAG, cells remain confined to p0/p1 and never transition to p2 or pMN*. Simulations predicted that 500nM SAG for 24h would be sufficient for some cells to have passed the saddle that separates p0/p1 from p2. Consequently, if the p2 and pMN* attractors are stable at 0nM SAG, then a proportion of cells would remain in these states at D5 and D6. The removal of the signal would also result in the reestablishment of the p0/p1 state, as the parameters cross back over the p0/p1 fold line (SI Fig. S22E). Hence cells that had not passed the p0/p1-p2 saddle by D4 were predicted to remain trapped in p0/p1. Consequently, approximately equal proportions of p0/p1 and pMN* cells were predicted at D5 and D6, with minimal presence of other cell types.

Experimental results showed a markedly different outcome from the prediction: pMN* cells dominated at D5 and D6 and p0/p1 cells were largely absent (Fig. 6 E). While the presence of pMN* cells after SAG removal was consistent with the stability of the pMN* attractor at 0nM SAG, the almost complete absence of p0/p1 indicated an irreversible commitment to pMN* fates. Thus, once the p0/p1 attractor has bifurcated, it appears not to be reestablished after signal withdrawal. This is consistent with previous experimental observations that documented bistability in the factors specifying neural identities and hysteresis in the allocation of ventral neural progenitor identities[58].

To capture this effect in the model, we re-fitted to determine the distributions of parameters 'post-500nM-SAG' i.e. to the experimental results for D4-D6. The fitted parameters for this experimental condition closely matched those observed under continuous 500nM SAG exposure for D4-D6 and were far from those for 0nM SAG, reflecting a sustained signalling response even after SAG withdrawal (Fig. 6B and (SI Fig. S22E). With this new set of parameters the model accurately reproduced the experimental cell state distributions for this condition (Fig. 6E).

To test whether the irreversibility of the p0/p1 attractor depends on the concentration of SAG, we tested the effect of exposure to 100nM SAG for 24h followed by a transition to 0nM SAG. This resulted in the predom-

inant allocation to pMN* fate and some p2 cells, suggesting that both pMN* and p2 attractors are present at 0nM SAG. Similar to the results with the 500nM SAG pulse, there were few, if any, p0/p1 cells indicating that 100nM SAG is also sufficient for the irreversible bifurcation of the p0/p1 attractor.

In the absence of Shh signalling, the p0/p1 attractor in a cell's landscape is far from its bifurcation point and maintains stability. Cells exiting the PreNeural state are captured by it. When Shh signalling exceeds a critical threshold, model parameters shift to regions where the p0/p1 attractor destabilises, enabling cells to transition towards p2 and pMN* fates. The experimental data indicate that the bifurcation of p0/p1 is irreversible, as cells do not accumulate in the p0/p1 attractor if they have been exposed to a pulse of SAG. As p2 is always a shallow attractor under SAG exposure, cells can only reach the pMN* state if they can leave the p0/p1 attractor. The persistence of pMN* identity after signal removal demonstrates how transient signalling can induce lasting changes in cell fate through the attractor stability dynamics.

Validating the circular topology

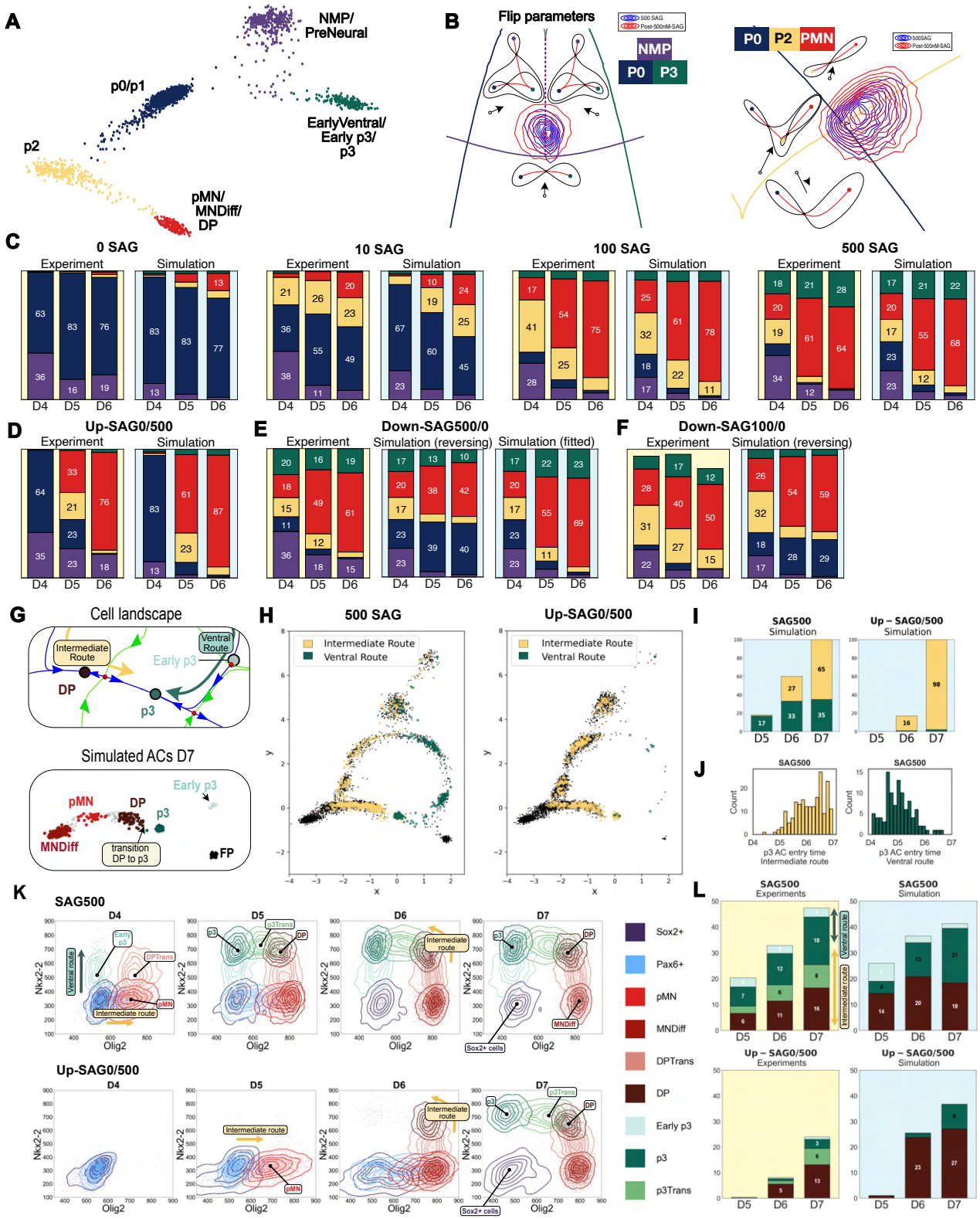
Finally, we set out to test the circular topology of the landscape that predicted two routes to access p3 (Fig. 6G) - a direct path through the ventral sub-landscape under constant high Shh signaling (*ventral route*), and an indirect path through the intermediate sub-landscape via the DP attractor (*intermediate route*). Consistent with this, analysis of the scRNAseq data showed a concurrent decrease in DP cells and increase in p3 cells between D6 and D7 (SI Fig. S21B).

Using the full model, we compared simulations up to D7 of 500nM SAG and simulations of 24h with no SAG followed by 72h with 500nM SAG signalling (Fig. 6H). We tracked the route cells followed to reach the p3 AC. Comparing the simulated proportions of cells in the p3 AC with respect to the total population of p3 cells at D7 for the two signalling conditions we observed that, in the second signal regime, the vast majority of p3 cells took the indirect DP-to-p3 intermediate route, rather than the ventral route via EarlyVentral and Early p3. (Fig. 6I). A consequence of our model capturing cellular heterogeneity is that the timing of arrival at the p3 attractor varies widely across cells. Histograms of first entry time points for the simulated trajectories reveal that a larger proportion of cells reach p3 earlier via the ventral route, whereas those taking the alternative intermediate route tend to arrive later, as the shallow intermediate states delay transitions (Fig. 6J).

To test experimentally the simulation predictions of the DP-to-p3 transition, we performed experiments extending the time course to D7 and analysed the effect of adding 500nM SAG at D4 after 24h without SAG using flow cytometry. Under these conditions, the model predicts that PreNeural cells are forced to take the intermediate route through the p0/p1 attractor, as without SAG almost all cells transition to p0/p1. Subsequently elevating SAG levels to 500nM SAG destabilises the p0/p1 and p2 attractors, making p3 accessible via the DP saddle (Fig. 6G). Flow cytometry data at D7 revealed that cells had transitioned from DP to p3, characterised by decreased Olig2 expression while maintaining that of Nkx2.2 (Fig. 6K). This is particularly clear at D7 when SAG exposure is delayed by 24h: p3 cells emerge despite the near absence of Early p3 at earlier time points, ruling out the possibility of a ventral route.

We compared the simulated proportions in the DP, Early p3 and p3 ACs for both experimental conditions, using the PPDs previously obtained for 0nM and 500nM, noting that this experiment was not included in the fitting of the model (Fig. 6L and SI Fig. S21G). The model captured the increased proportion of DP cells and the leakage from DP to p3 via the intermediate route. In the SAG delay experiment, the simulated proportions of p3 cells at D7 via the ventral route are nearly absent (less than 1%) while p3 cells from the intermediate route (9%) are lower than the proportion observed under constant SAG 500nM exposure (SI Video S3-S4).

Together, these results validate the model and demonstrate the complex relationship between signal dynamics and cell fate allocation. The analysis indicates that transient high-level Shh signalling induces irreversible commitment to ventral fates, while also providing support for the dual pathways to p3 identity. These findings establish how cells integrate and retain signalling information to make robust fate decisions, with the mathematical model providing a framework for understanding the underlying dynamical principles.



Experimental Validation of Model Predictions Through Signalling Perturbations.

Figure 6: (Caption next page)

Experimental Validation of Model Predictions Through Signalling Perturbations.

A. Simulated data showing the attractors and transition routes. Dots represent simulated cells retained at 3 time points (D4, D5, D6). These are aggregated into simulated ACs coloured by the cell state to which they correspond.

B. PPDs from model fitting for the 500nM SAG condition (blue) and the post-high-SAG condition (red) are close to identical, revealing that the two conditions are approximately equivalent. (Left) Projection of the PPDs on the bifurcation set of the binary flip sub-landscape. (Right) Projection of the PPDs on the bifurcation set of the linear sub-landscape.

C. Comparison of average simulation results (blue background) for the constant SAG conditions with the corresponding experimental results (yellow background). Results are average over 10000 simulated proportions using parameters drawn from the fitted PPDs (SI Sect. S12).

D. As C but for the Up-SAG0/500 condition. The prediction aligns with the experimental data. It shows that delaying SAG application by 24h commits the cells to differentiate into pMN* rather than p3*.

E. As C but for the Down-SAG500/0 condition but modified as follows: (middle) uses the fitted PPDs (500nM and 0nM SAG) obtained from H, (right) uses the refitted PPD. The latter shows better agreement with the experimental results.

F. As for the first two panels of E for the Down-SAG100/0 condition but where the fitted PPDs (100nM and 0nM SAG) obtained from C are used. We observe the same phenomenon as in E (middle panel), despite lower levels of the signal.

G. (Top) Local view of a cell's landscape near the DP and p3 attractors where the DP-to-p3 transition occurs. The p3 attractor can be reached via two distinct pathways: an intermediate route (yellow), following the unstable manifold which passes through the DP attractor and connects to p3 from the left, and a ventral route (green), following the unstable manifold which passes through the Early p3 attractor and connects to p3 from the right. Each route is characterised by different gene expression profiles. Cells transitioning through the intermediate route have a history of Olig2 expression, whereas cells following the ventral route express Foxa2 early at D4. (Bottom) Local view of simulated ACs (500nM SAG condition) at D7, capturing the DP-to-p3 transition. Each dot represents a single cell. The cells that reached DP (brown) had their dynamics locally governed by a dynamical system as shown above where there is a well defined intermediate route. A few cells remain in the pMN AC but the majority of pMN cell have transition into motor neurons or DP.

H. (Left) Simulated cell trajectories using the PPD from model fitting (500nM SAG condition). Dots represent simulated cells retained at 4 time points (D4-D7). Cells reaching the p3 AC by D7 are coloured based on their transition history: yellow for cells that follow the intermediate route and green for those that follow the ventral route. Black dots represent cells that adopted alternative end fates (such as MNDiff or FP) or have not yet reached p3 by D7. (Right) Same setup, but with pulsing applied in the simulation by switching the parameters from 0 nM to 500 nM after D4. Significantly fewer cells opt for the ventral route as most cells had already taken the intermediate route by D4 when the signal changed, as in D.

I. Proportions of simulated cells in the p3 AC for each day (D5 to D7) computed with respect to the total number of p3 cells at D7 and coloured by the route they originate from as in H. (Left) 500nM condition, (Right) Up-SAG0/500nM condition. The majority of p3 cells under this latter condition originate from the intermediate route.

J. Histograms showing cell counts (y-axis) by the time point (x-axis) at which simulated cells first reach the p3 AC, with colours indicating their route of origin: intermediate (yellow) or ventral route (green). The distributions reflect timing heterogeneity: cells that reach p3 via the intermediate route tend to arrive at late time points (D6-D7). By contrast, cells following the ventral route tend to reach p3 early (D4-D5) while few new cells arrive via this route after D6.

K. 2D Gene plots (Olig2-Nkx2.2 shown) of the clusters identified in the flow cytometry dataset at each time point. The dots represent individual cells, and contour lines represent the density of cells (plotted only when the cluster contains more than 150 cells). Without the Nkx6.1 marker p0/p1 cannot be distinguished from p2 cells and both are combined into Pax6$^+$ cells, and PreNeural from EarlyVentral and FP are combined into Sox2$^+$ cells. (Top) 500nM SAG condition: as early as D4, a group of cells take the ventral route by up-regulating Nkx2.2 and reach p3 by D5. (Bottom) Up-SAG0/500nM condition: with no SAG for 24h, cells are forced to take the intermediate route and most of them had up-regulated Pax6 by D4. When SAG is applied, they can only reach p3 by transiting through pMN (D5), then DP (D6), followed by a direct DP-to-p3 transition (D7).

L. (Top left) 500nM SAG condition: experimental proportions of the DP AC and p3Trans (intermediate route) and the Early p3 and p3 ACs (ventral route) at late time points (D5-D7).

(Top right) 500nM SAG condition: simulated proportions of the DP, Early p3 and p3 ACs at late time points (D5-D7). The p3 proportions can be further split into two type of cells depending on which route they come from as in I.

(Bottom) Same as above for the Up-SAG0/500nM condition. In the simulation (bottom right), most p3 cells at D7 originate from the intermediate route as shown in I (right) which is consistent with the experiment (K bottom).

Testing the model on in vitro human data

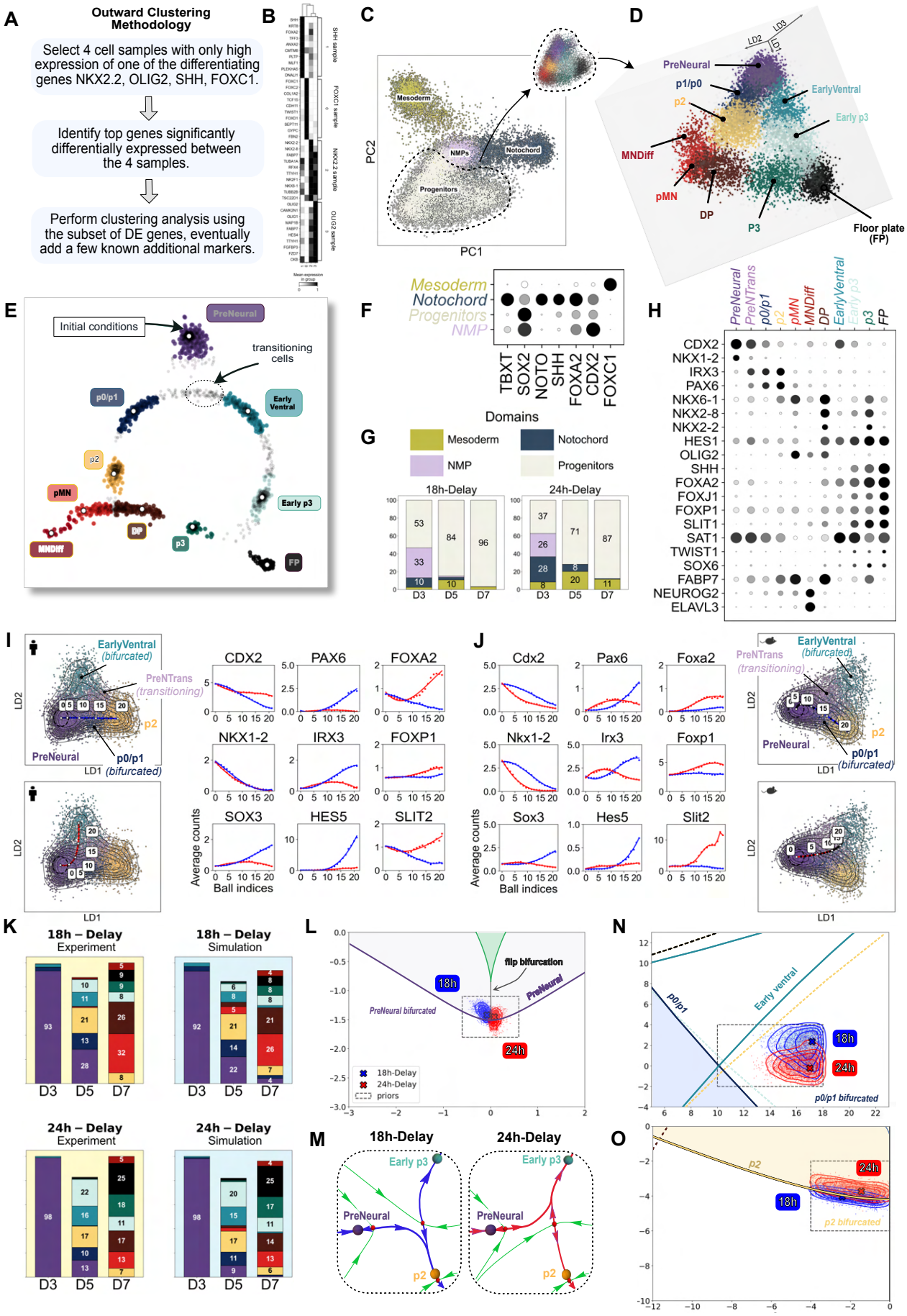
To evaluate the generality of the landscape model, we investigated its applicability to an in vitro model of human trunk development[29]. This system generates multiple embryonic tissues, including neural progenitors, mesodermal cells, and notochord cells, which are a source of endogenous Shh signalling to the other tissues. The presence of the Shh producing notochord in these organoids results in the generation of ventral neural tube progenitors (p2, pMN, p3) and floor plate (FP). We analysed two experimental conditions - an *18h-delay* and *24h-delay* in TGFβ addition - which produced different proportions of notochord cells and consequent ly varying levels of Shh signalling[29]. The data comprised scRNAseq collected at 3 time points D3, D5, D7.

Using the clustering method described above (Fig. 7A-B), we identified the expected cell populations: NMPs, mesoderm, ventral neural progenitors, and notochord (Fig. 7C-G). When examining neural progenitors (SOX2$^+$, TBXT$^-$), we discovered remarkably similar ACs to the mouse dataset (SI Fig. S23). Moreover, an LDA projection revealed the ACs were arranged in a similar pattern to the previous data: the PreNeural state leading to intermediate and ventral sub-landscapes and comparable transitions between ACs including a circular topology in which the two routes converge to the p3 AC (Fig. 7D-E).

Despite differences in developmental timing, gene expression patterns characterising the ACs (Fig. 7H) and individual analysis of each sub-landscape showed striking conservation between mouse and human systems (Fig. 7I-J for the PreNeural sub-landscape and SI Fig. S24-S25 for the other sub-landscapes). We therefore fitted our landscape model to the two datasets – *18h-delay* and *24h-delay* – focusing on the neural progenitor landscape. We anticipated that, since the *18h-delay* condition contains less notochord and therefore less Shh signalling than the *24h-delay* condition, the *18h-delay* dataset would behave as if it had been exposed to lower SAG concentrations than the *24h-delay* dataset. Cells in the D3 PreNeural AC were used as initial conditions and comparisons were made to the proportions of cells in each simulated AC with the experimental cell states proportions (SI Sect. S13.3). The model successfully reproduced the experimental data (Fig. 7K).

Analysis of the bifurcation sets revealed key differences between the *18h-delay* and *24h-delay* conditions (SI Sect. S13.4). The 24h-delay generated more notochordal cells (Fig. 7G), resulting in higher proportions of floor plate and p3 progenitors and fewer pMN cells by D7 than the 18h-delay. Importantly, the bifurcation analysis showed that these differences primarily affected the PreNeural flip bifurcation (Fig. 7L-M), while other bifurcations remained unchanged between conditions (SI Fig. S26F). In both experimental conditions, the p0/p1 attractors were unstable across all cell landscapes, making the p0/p1 AC transient - mirroring the behaviour observed at high SAG concentrations in the mouse experiments (Fig. 7N). Most p2 attractors remained shallow, with the corresponding PPDs clustering just within the p2 bifurcation fold curve (Fig. 7O). This suggests that the increased number of notochordal cells in the 24-delay condition, and consequently the higher level of Shh signalling, primarily influenced the first neural progenitor cell fate decision, directing more cells towards the ventral sub-landscape, providing quantitative insight into how notochord-derived Shh signalling affects ventral patterning.

These findings demonstrate that the landscape model successfully captures neural patterning dynamics across species while revealing features of human development. The analysis identified how cell state transitions are affected by notochord-derived Shh: while the PreNeural flip bifurcation showed high sensitivity, other transitions displayed qualitative robustness to signalling variations. This raises important questions about how dorsal fates are established in the presence of strong ventralising signals from the notochord, suggesting the necessity of additional regulatory mechanisms. Future studies will need to explore how combinations of signals shape the early neural tube landscape.



Comparative Analysis of Human Neural Progenitor Development Validates the Circular Topology of the Decision Landscape

Figure 7: (Caption next page)

Comparative Analysis of Human Neural Progenitor Development Validates the Circular Topology of the Decision Landscape

A. Summary of the outward clustering methodology.

B. The most significantly differentially expressed genes found using the 4 samples defined in A.

C. The four cell type domains including neural progenitors were extracted, annotated and separately projected using LDA from the effective gene space defined by the gene modules in B.

D. 3D LDA projection of the Neural Progenitor ACs across all time points and all experimental conditions. The topological structure reveals the circular topology of the transition pathways.

E. Simulated Neural Progenitors ACs produced by the landscape model of this system after fitting in a similar fashion to that described above. Dots represents simulated cells retained at 3 time points corresponding to D3, D5, D7 using the parameters from the fitted PPDs for the 18h-Delay condition. Gray dots represent transitioning cells that were not assigned to any simulated AC.

F. Dot plot indicating the expression of markers used to identify the tissue domains. The intensity of shading indicates the mean expression of each gene in each domain.

G. Proportions of cells in each domain of C and F by time points under the two experimental conditions.

H. Dot plot indicating the expression of markers used to identify the ACs and transitioning clusters of the Neural Progenitors domain in C.

I.-J. PreNeural sub-landscape extracted from the human dataset (I) and the mouse datasets (J) and transition analysis for marker genes that characterise the binary decision as cells exit PreNeural. Unstable manifolds defining routes between pairs of ACs where computed in effective gene space defined by the genes in B and projected onto 2D LDA space. Marker genes were analysed along both routes. Exit from PreNeural is marked by down-regulation of Cdx2 and Nkx1.2. A branching event occurs (around index 10), followed by up-regulation of markers characteristic of ventral states along the red route (leading to EarlyVentral), and markers such as Pax6 characteristic of the intermediate states p0/p1, and p2 along the blue route.

K. (Top) Experimental compared to simulated proportions after parameter fitting for the 18h-Delay condition. (Bottom) Same as above for the 24h-Delay condition. It shows an increase of ventral states such as p3 and FP.

L. Bifurcation set of the PreNeural binary flip shows the qualitative difference between the 18h- and the 24h-Delay conditions. The 18h-Delay PPD (blue) cluster on the left of the flip curve and above the PreNeural fold curve. Under this condition, the PreNeural attractor is shallow for most cells and is connected to p0/p1. By contrast, the 24h-Delay PPD (red) lies to the right of the flip curve, favouring cell trajectories from PreNeural to EarlyVentral and promoting the generation of more p3 and FP cells.

M. Typical configurations of attractors, saddles and unstable manifolds for the cell landscape models for the PreNeural sub-landscape for parameters in each of the PPDs in L. In the 18h-Delay, cells are most likely to transition via the intermediate route towards p2 and to Early p3 via the ventral route in the 24h-Delay. As indicated by the PPDs in N, the attractors p0/p1 and EarlyVentral are bifurcated for all cells in both experimental conditions. Hence the connections from PreNeural lead directly to p2 (blue unstable manifold) or to Early p3 (red unstable manifold).

N. Bifurcation set of the p2 transition shows that, under both conditions, the p2 AC is shallow as most parameters cluster near (and above) the p2 fold curve.

O. Bifurcation set of the circular sub-landscape controlling the p0/p1 and EarlyVentral fold bifurcations. Under both conditions, the PPDs are well-separated from the blue shaded region where the p0/p1 and the EarlyVentral attractors are bifurcated for all cells. It shows in particular that the p0/p1 and EarlyVentral ACs are not stable.

DISCUSSION

In this study, we established Dynamic Landscape Analysis (DLA), a systematic framework for constructing quantitative and predictive landscape models from single-cell data. By applying this approach to a model of neural tube development, we uncovered an extensive landscape with a rich bifurcation structure and an unexpected circular topology, where initially divergent lineages converge through multiple routes.

A key component of this approach was the identification and characterisation of attractor clusters (ACs). This follows a semi-supervised strategy, starting with known marker genes and iterative expansion to identify genes associated with the ACs and the transitions between them. This approach identified established neural progenitor states and less well-understood states, such as a double-positive population expressing both Olig2 and Nkx2-9, along with Hes1[44]. The method is characterised by greater control of dimensionality, motivated by the need to avoid noisy or irrelevant dimensions in the data and the extensive experimental work suggesting that cell fate decisions are controlled by a relatively small number of genes[1]. While current techniques for analysing scRNA-seq data often start with highly variable genes (typically numbering in the thousands), the ratio of relevant genes to all features in such samples is very small. As the fraction of relevant features decreases, existing clustering and dimensionality reduction techniques often fail to discover the identity of relevant features because correlations between samples become dominated by noise and fluctuating irrelevant genes[59].

A second important feature of DLA is the use of dimension reduction methods that preserve the geometry of the local landscape and the transitions. Our analysis of the bifurcation structures depends upon the dimension reduction projection being differentiable. Non-continuous or non-differentiable methods would eliminate critical geometric features, particularly unstable manifold smoothness. While neural networks can be used to modify t-SNE[60] and UMAP[61] to create smooth embeddings[62,63], LDA provides superior results in our context of analysing related attractors because it maximises separation through a controlled, understandable linear embedding.

The availability of a fitted stochastic model that reproduces experimental data allows a better understanding of analytical methods for single cell data (e.g. Fig. 5F). For example, there has been much interest in trajectory inference (TI) methods which aim to infer a 1-dimensional graph-like structure underlying the dynamic process from which the cells are sampled[64,65]. By projecting the cell states to this inferred structure, their properties can be compared over pseudotime, an inferred unit of progress along the graph-like structure. In this approach, cell state branches correspond to nodes of the graph with one inward and two outward edges. Flip bifurcations have a similar branching topology, but cell states traverse it with radically different dynamics to that given by pseudo-time.

Our theory predicts that at the branching point there is a head attractor that cells pass through. Experimental results and models validate this. Cells can either remain in a strong attractor until a signal triggers a bifurcation, or linger near a shallow attractor for an extended period before escaping through stochastic fluctuations (Fig. 5G). Consequently, a global pseudotime cannot accurately represent a cell's passge through such an attractor. Simulations demonstrate this limitation, showing cells arriving at p3 via the intermediate route are approximately equally distributed across Days 5, 6, 7, despite starting from the NMP state simultaneously (Fig. 6J). By contrast, pseudotime aligns more closely with actual simulation time during transitions between attractors (Fig. 5G). The "stickiness" of attractors explains the difficulty in capturing transitioning cells in experimental data with limited temporal resolution.

Our analysis of the transitions between ACs revealed three transition types: fold bifurcations that drive a direct transition from one state to another (for example in the transition from p2 to pMN); binary flip and binary choice bifurcations that govern the branching decisions between alternative fates. The bifurcation types provide insight into biological features of cell fate decisions. Fold bifurcations explain how an incremental change in signal strength can trigger a switch in cell identity once a threshold value is crossed. On the other hand, a binary flip bifurcation describes the concurrent production of two downstream states from a progenitor population. In this case, differences in cell state or signals received by a cell determine which of the two possible escape routes a cell chooses. Thus, at a population level, it will be common, but not mandatory, for cells to go both ways. An important implication of the flip bifurcation is that a cell first leaves the head attractor before determining which downstream fate it adopts. In other words, in

these transitions, a cell commits to differentiate before specifying which fate it will choose. This could explain the phenomenon of multi-lineage priming, where cells simultaneously express low levels of genes associated with multiple potential fates before committing to a specific lineage[66]. Finally, the binary choice differs from the flip in that there are two saddles determining the escape of the cells rather than one. This allows for two independent signals to move the two saddles independently and so to move all the cells to the same attractor[18]. We speculate that the reason we see it in defining the p2-pMN-DP region of the landscape is because if a flip was used to choose between pMN and DP, the head attractor controlling the bifurcation would be p2, which is not present at high Shh levels. Therefore, the more stable pMN attractor is used to control the bifurcation, and since this sits on the unstable manifold linking MNDiff and DP, the binary choice is used.

A striking feature of our analysis is the circular topology of the neural progenitor landscape, where two developmental trajectories converge on the p3 identity. Experimental perturbations confirmed that p3 cells can be generated through both paths, with the route followed depending on the timing and level of Shh signalling. Applying our landscape model to human neural organoid development revealed consistency between human and mouse systems, suggesting that this circular topology represents a fundamental feature of vertebrate neural development. This finding reconciles previously contradictory findings of the p3 domain being generated through either a Foxa2$^+$/Nkx6.1$^+$/Pax6$^-$ route[26] or through a Pax6$^+$/Olig2$^+$ route[49]. It also suggests greater complexity and redundancy in developmental programmes than can be explained by the conventional view of a strictly hierarchical 'tree' of cell type diversification.

While GRN models have been successfully constructed for some systems, there has been a notable lack of progress in others, and current single-cell computational approaches have been disappointing[67]. GRN construction has been feasible in systems such as oscillators where there is an obvious dynamical phenotype that constrains the possible dynamics of the GRN with easy and informative readouts such as period, amplitude and phase that can be used to explore the effects of perturbations. The generative modelling approach we have developed could play an analogous role for other systems by providing similar tools such as a clear phenotype (the route and timing of a cell passing through the ACs and the proportion of cells doing this) and a precise prescription of the dynamical phenomena that the GRN model must display. Moreover, the simplest gene circuits that produce the local normal forms will likely be very highly constrained and this will provide a limited set of gene interaction modules that can be searched for in the data[17,19].

Our approach to modelling allows us to readily identify the bifurcation set in the model parameter space. Combining this with the posterior parameter distribution (PPD) found by parameter optimisation provides a potentially powerful tool. It links the PPDs corresponding to different morphogen levels to dynamical behaviour. The relationship between the estimated PPD and the bifurcation curves predicts the behaviour of a heterogeneous group of cells exposed to a particular level of signal. For example, in cell states with stable identities, such as p0/p1 in low SAG concentrations, we find parameters clustering away from critical bifurcation curves so that even a heterogeneous cell population displays homogeneous behaviour. In other cases the PPD overlaps bifurcation curves in a way that suggests a cell population makes a mixture of alternative decisions (e.g. to transition, remain in place, or choose between options in a binary flip landscape). For different morphogen levels all are seen in the transition from the PreNeural state (Fig. 5G). More varied cellular outcomes can occur when the PPDs for different morphogen levels overlap more complex structures in the bifurcation curves, as is the case for the transition from p2 to pMN or DP (SI Fig. S18A Panel 5). A key insight is that distinct fate choices for a group of initially near-identical or "micro-heterogeneous" cells[68] can emerge from the deterministic landscape dynamics (bifurcations) rather than solely from noise induced stochastic fluctuations affecting the trajectory of a cell[69].

The bifurcations that we observe are essentially irreversible, providing an explanation for previously identified hysteresis in the system[58]. By modelling, we see that the saddle controlling the fold bifurcation or stochastic escape from each attractor is always on the ventral side of the attractor implying that once a cell has escaped from the attractor it is very difficult for signals to reverse this. Consequently, we observe that even moderate SAG exposure (100nM SAG) results in the permanent loss of the otherwise stable p0/p1 identity, while the pMN identity can persist even after SAG removal.

By systematically characterising the decision points and the parameters that control cell fate decisions, we can design differentiation protocols that precisely manipulate bifurcation parameters towards desired cell type proportions. This systematic approach provides a deeper understanding of cellular decision making providing a precise and quantitative approach for designing *in vitro* protocols which, in time, could be applied to organoid models, stem cell therapies and regenerative medicine.

The simplest hypothesis concerning the linear neural progenitor patterning in the neural tube (namely that there is a monotonically ordered set of thresholds determining the boundaries) is contradicted by our results. Instead the analysis suggests that cells progress through a series of binary decisions, first choosing between the intermediate or ventral route and then patterning according to monotonic thresholds or further flips within the separate routes and sometimes adopting a fate that could have been reached by the other route. We conjecture that is an intrinsically more robust way to pattern compared to using a set of sequential monotonic thresholds and that the circular structure increases the dynamical robustness of the landscape.

Intriguingly, similar hierarchical decision-making dynamics have been observed in the learning process of generalized Hopfield networks, where memories progress through well-defined saddles before splitting into progressively specialised states[70]. In both neural development and machine learning, these decision landscapes appear to follow reproducible, low-dimensional trajectories despite the high-dimensional nature of the underlying systems. This similarity between biological cell decision making and machine learning suggests that hierarchical binary decisions may represent a fundamental organising principle for robust decision-making across different complex systems.

A challenge for future work is to extend this DLA framework to describe how combinations of different signals shape developmental decision-making by including a model of the way such multidimensional signals change the parameters of the model, as was done for a simpler system[18]. For these more sophisticated applications, integration of epigenetic data with transcriptional profiles may be required.

Taken together, DLA provides a framework for generating predictive mathematical models of cell fate decisions from single-cell data. With this framework, we have gained insight into how cells navigate fate decisions in neural development. By bridging the gap between high-dimensional single-cell data and mathematical models of cellular behaviour, DLA can advance our understanding of the principles governing cellular differentiation.

Acknowledgments

We gratefully acknowledge many very useful discussions with Eric Siggia and Dillon Cislo. We thank the Flow Cytometry Science Technology Platform at the Francis Crick Institute for their assistance in carrying out experiments.

Funding

This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (CC001051), the UK Medical Research Council (CC001051) and the Wellcome Trust (CC001051); by the Wellcome Trust (220379/D/20/Z); and by the UK Engineering and Physical Science Research Council (EPSRC) (grants EP/P019811/1 and EP/T031573/1). M.J.D. was supported by the Wellcome Trust Career Development Award (227326/Z/23/Z). This research was supported in part by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP).

Author contributions

J.B., M.J.D. and D.A.R. conceived the project. M.J.D. and E.F. designed and performed flow cytometry experiments. R.M. generated and processed scRNA-seq data. M.F. and D.A.R. developed methodology. M.F. developed python software and analysed scRNA-seq data. M.F., M.S. and M.J.D. analysed FACS data. Visualisation M.F. All authors interpreted data and designed experiments. M.F. and D.A.R. developed the mathematical models with contributions from M.S. M.F. constructed the global landscape model and performed related analysis. M.S. performed the pulses analysis and constructed the related model. M.F., J.B. and D.A.R. wrote the manuscript with additional material from M.S. All authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing or financial interests.

Materials and Methods

Cell lines

Experiments were performed with the mouse embryonic stem cell line HM1[71] maintained at 37°C with 5% CO_2.

ES cell culture and differentiation

Mouse ES cells were maintained on a feeder layer of mitotically inactivated mouse embryonic fibroblasts (MEFs, derived and expanded in-house) in ES cell medium (Dulbecco's Modified Eagle Medium (DMEM) Knock Out (Gibco; 10829-018) supplemented with 1% Foetal Bovine Serum (Pan Biotech; P30-2602), Penicillin/Streptomycin (Gibco; 15140122), 2 mM L-Glutamine (Gibco; 25030024), 2 mM Non-essential amino acids (Gibco, Cat No. 11140-035), and 0.1 mM 2-mercaptoethanol (Gibco; 21985-023)) with 1000 U/ml LIF (Chemicon, Int ESG1107)). Media was changed every day and cells were passaged every other day at a density of 500,000 cells per 60mm dish.

For differentiations, cells were washed once with PBS and dissociated using 0.05% Trypsin-EDTA (Gibco; 25300054) for 4 min at 37°C and resuspended in 10 ml ES media. Cells were plated in 10 cm plates coated with 0.1% gelatin to remove feeder cells ('panning'). Cells were incubated for 15 min at 37°C to allow feeder cells to attach. Without disturbing any attached cells, the cells in suspension were transferred to a second gelatinised 10 cm for another 15 min. The process was repeated a third time.

Differentiations were carried out as previously described[26], by plating 'panned' mES cells resuspended in N2B27 media (Advanced DMEM - F12 (Gibco, Cat. No. 21331-020) and Neurobasal medium (Gibco, Cat. No. A35829-01) (1:1), supplemented with 1xN2 (Gibco Cat no. 17502001), 1xB27 (Gibco Cat no. 17504001), 2 mM L-glutamine (Gibco, Cat No. 25030024), 40 μg/ml BSA (Sigma-Aldrich, Cat No. A7979-50ML), and 0.1 mM 2-mercaptoethanol) at a density of 20,000 cells in 1.5 ml of media onto 6-well plates (Corning, Cat. No. 353046) precoated in Matrigel (Corning, Cat. No. 356231) diluted 1/100 in Advanced DMEM - F12. The media was supplemented on the different days as follows: Day 0 to Day 2, 10ng/ml bFGF (R&D, Cat. No. 100-18B) and 5 μM LGK (Cayman Chemical Company, Cat. No. 1.800.364.9897); Day 2 to Day 3 for 20 h, 10 ng/ml bFGF, 5 μM CHIR99021 (Axon Medchem, Cat. No. 1386), 10 μM SB-431542 (Tocris, Cat. No. S0400), and 2μM DMH1 (Adooq Bioscience, Cat. No. A12820); from Day 3 onwards, 100 nM RA (Sigma, Cat. No. R2625) and the indicated concentrations and timings of Smoothened Agonist SAG (Merck, Cat. No. 566660-5mg).

Intracellular flow cytometry

Sample collection: 1μl/ml of LIVE/DEAD Fixable Dead Cell Stain Near-IR fluorescent reactive dye (ThermoScientific, Cat. No. L34976) was added to cells in culture and incubated at 37°C for 30 mins. Cells were then washed twice with PBS (Gibco, Cat. No. 14190-094), and dissociated using 0.5ml Accutase (Gibco, Cat. No. 00-4555-56) per well of a 6 well plate incubated 5 min at 37ºC. Cell were collected, centrifuged at 400g for 4 min and resuspended in 100 μl of 4% paraformaldehyde (PFA) (ThermoScientific, Cat. No. 28908). PFA fixation was carried out for 10 min at room temperature. Cells were washed in PBS and resuspended in 500µl PBS + 0.5% BSA.

Staining: 1 million cells were stained for flow cytometry analysis. Pellets were resuspended in 0.1 ml of PBS supplemented with 0.5% BSA and 0.1% Triton-X100 (VWR Chemicals, Cat No. 28817.295) and the appropriate primary or directly-conjugated antibodies for 1.5 h protected from light at room temperature. Secondary antibodies were incubated under the same conditions for 40 min. Cells were washed in PBS supplemented with 0.5% BSA and 0.1% Triton-X100, and resuspended in 300μl of PBS with 0.5% BSA for analysis on a BD Fortessa analyzer (Becton Dickinson).

The antibody panels used were as follows: Primary antibodies were Sox2-V450 (1:100)(BD Biosciences Cat no 561610), Pax6-PerCPCy5.5 (1:100)(BD Biosciences Cat no 562388), Nkx6.1-AlexaFluor647 (1:100)(BD Biosciences Cat no 563338), Goat Olig2 unconjugated (1:400) (R&D Cat no AF2418) and Nkx2.2-PE (1:100) (BD Biosciences Cat no 564730) followed by secondaries donkey anti-goat AlexaFluor488 (1:1000) (Thermo Fisher Scientific cat no A11055).

Flow cytometry data pre-processing

The flow cytometry data were analyzed using FlowJo™ v10.8 Software (BD Life Sciences). Cells were selected for downstream analysis if they were negative for LIVE/DEAD Fixable Dead Cell Stain Near-IR (alive),

and with uniform FSC, SSC distribution. The cells were also required to be SOX2+ (neural progenitors).

References

1. Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. Nature *468*, 911–920.

2. Waddington, C. The strategy of the genes. allen (1957).

3. Huang, S., Eichler, G., Bar-Yam, Y., and Ingber, D. E. (2005). Cell fates as high-dimensional attractor states of a complex gene regulatory network. Physical Review Letters *94*, 128701.

4. Huang, S., Guo, Y. P., May, G., and Enver, T. (2007). Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. Developmental Biology *305*, 695–713. doi:10.1016/j.ydbio.2007.02.036.

5. Huang, S. (2012). The molecular and mathematical basis of waddington's epigenetic landscape: A framework for post-darwinian biology? Bioessays *34*, 149–157.

6. Corson, F., and Siggia, E. D. (2012). Geometry, epistasis, and developmental patterning. methods: model and numerical simulation. PNAS *0*, 1–23.

7. Ferrell Jr, J. E. (2012). Bistability, bifurcations, and waddington's epigenetic landscape. Current Biology *22*, R458–R466.

8. Marco, E., Karp, R. L., Guo, G., Robson, P., Hart, A. H., Trippa, L., and Yuan, G. C. (2014). Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. Proceedings of the National Academy of Sciences of the United States of America *111*, E5643–E5650. doi:10.1073/pnas.1408993111.

9. Bargaje, R., Trachana, K., Shelton, M. N., McGinnis, C. S., Zhou, J. X., Chadick, C., Cook, S., Cavanaugh, C., Huang, S., and Hood, L. (2017). Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. Proceedings of the National Academy of Sciences of the United States of America *114*, 2271–2276. doi:10.1073/pnas.1621412114.

10. Corson, F., and Siggia, E. D. (2017). Gene-free methodology for cell fate dynamics during development. eLife *6*, 1–25. doi:10.7554/eLife.30743.

11. Wang, J., Xu, L., Wang, E., and Huang, S. (2010). The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. Biophysical Journal *99*, 29–39. doi:10.1016/j.bpj.2010.03.058.

12. Huang, S., Li, F., Zhou, J. X., and Qian, H. (2017). Processes on the emergent landscapes of biochemical reaction networks and heterogeneous cell population dynamics: differentiation in living matters. Journal of the Royal Society Interface *14*, 20170097.

13. François, P., and Jutras-Dubé, L. (2018). Landscape, bifurcations, geometry for development. Current Opinion in Systems Biology *11*, 129–136.

14. Cao, J., Zhou, W., Steemers, F., Trapnell, C., and Shendure, J. (2020). Sci-fate characterizes the dynamics of gene expression in single cells. Nature biotechnology *38*, 980–988.

15. Camacho-Aguilar, E., Warmflash, A., and Rand, D. A. (2021). Quantifying cell transitions in c. elegans with data-fitted landscape models. PLoS Computational Biology *17*, 1–28. doi:10.1371/journal.pcbi.1009034.

16. Freedman, S. L., Xu, B., Goyal, S., and Mani, M. (2021). Revealing cell-fate bifurcations from transcriptomic trajectories of hematopoiesis.

17. Rand, D. A., Raju, A., Sáez, M., Corson, F., and Siggia, E. D. (2021). Geometry of gene regulatory dynamics. PNAS *118*.

18. Sáez, M., Blassberg, R., Camacho-Aguilar, E., Siggia, E. D., Rand, D. A., and Briscoe, J. (2021). Statistically derived geometrical landscapes capture principles of decision-making dynamics during cell fate transitions. Cell Systems.

19. Sáez, M., Briscoe, J., and Rand, D. A. (2022). Dynamical landscapes of cell fate decisions. Interface Focus *12*, 20220002.

20. Raju, A., and Siggia, E. D. (2023). A geometrical perspective on development. Development, Growth Differentiation *65*, 245–254. doi:10.1111/dgd.12855.

21. Raju, A., and Siggia, E. D. (2024). A geometrical model of cell fate specification in the mouse blastocyst. Development *151*, dev202467. doi:10.1242/dev.202467.

22. Howe, A. E. S., and Mani, M. (2024). Learning geometric models for developmental dynamics. bioRxiv. https://www.biorxiv.org/content/10.1101/2024.09.21.614191v2. doi:10.1101/2024.09.21.614191. Preprint.

23. D. J. Cislo, J. B., M. J. Delás, and Siggia, E. D. Reconstructing waddington's landscape from data (2025). arXiv:in preparation.

24. Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2014). In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. PLoS Biology *12*, e1001937.

25. Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2017). A gene regulatory network balances neural and mesoderm specification during vertebrate trunk development. Developmental Cell *41*, 243–261.

26. Delás, M. J., Kalaitzis, C. M., Fawzi, T., Demuth, M., Zhang, I., Stuart, H. T., Costantini, E., Ivanovitch, K., Tanaka, E. M., and Briscoe, J. (2023). Developmental cell fate choice in neural tube progenitors employs two distinct cis-regulatory strategies. Developmental Cell *58*, 3–17.e8.

27. Sagner, A., Zhang, I., Watson, T., Lazaro, J., Melchionda, M., and Briscoe, J. (2021). A shared transcriptional code orchestrates temporal patterning of the central nervous system. PLoS Biology *19*, e3001450. doi:10.1371/journal.pbio.3001450.

28. Maizels, R. J., Snell, D. M., and Briscoe, J. (2024). Reconstructing developmental trajectories using latent dynamical systems and time-resolved transcriptomics. Cell Systems *15*, 411–424.e9.

29. Rito, T., Libby, A. R. G., Demuth, M., Domart, M.-C., Cornwall-Scoones, J., and Briscoe, J. (2025). Timely tgfβ signalling inhibition induces notochord formation from human pluripotent stem cells. Nature *637*.

30. Jessell, J. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nature Reviews Genetics *1*, 20–29.

31. Frith, T. J., Briscoe, J., and Boezio, G. L. Chapter five - from signalling to form: the coordination of neural tube patterning. In: M., M., ed. *Vertebrate Pattern Formation* vol. 159 of *Current Topics in Developmental Biology* (168–231). Academic Press (2024):(168–231).

32. Diez del Corral, R., Breitkreuz, D. N., and Storey, K. G. (2002). Onset of neuronal differentiation is regulated by paraxial mesoderm and requires attenuation of FGF signalling. Development *129*, 1681–1691. doi:10.1242/dev.129.7.1681.

33. Sagner, A., Gaber, Z. B., Delile, J., Kong, J. H., Rousso, D. L., Pearson, C. A., Weicksel, S. E., Melchionda, M., Gharavy, S. N. M., Briscoe, J., and Novitch, B. G. (2018). Olig2 and hes regulatory dynamics during motor neuron differentiation revealed by single cell transcriptomics. PLoS Biology *16*, e2003127.

34. Novitch, B. G., Chen, A. I., and Jessell, T. M. (2001). Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. Neuron *31*, 773–789. doi:10.1016/s0896-6273(01)00407-x.

35. Briscoe, J., Pierani, A., Jessell, T. M., and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell *101*, 435–445.

36. Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A., and McMahon, A. P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. Cell *75*, 1417–1430. doi:10.1016/0092-8674(93)90627-3.

37. Traag, V. A., Waltman, L., and van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. Scientific Reports *9*, 5233.

38. Wilkinson, D. G., Bhatt, S., and Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. Nature *343*, 657–659. doi:10.1038/343657a0.

39. Hiemisch, H., Monaghan, A. P., Schütz, G., and Kaestner, K. H. (1998). Expression of the mouse Fkh1/Mf1 and Mfh1 genes in late gestation embryos is restricted to mesoderm derivatives. Mechanisms of Development *73*, 129–132. doi:10.1016/s0925-4773(98)00039-2.

40. Schubert, F. R., Fainsod, A., Gruenbaum, Y., and Gruss, P. (1995). Expression of the novel murine homeobox gene Sax-1 in the developing nervous system. Mechanisms of Development *51*, 99–114. doi:10.1016/0925-4773(95)00358-8.

41. Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou, V. E., and Kondoh, H. (2011). Tbx6-dependent sox2 regulation determines neural or mesodermal fate in axial stem cells. Nature *470*, 394–398.

42. Javali, A., Misra, A., Leonavicius, K., Acharyya, D., Vyas, B., and Sambasivan, R. (2017). Co-expression of tbx6 and sox2 identifies a novel transient neuromesoderm progenitor cell state. Development *144*, 4522–4529.

43. Mizuguchi, R., Sugimori, M., H, T., H, K., Nagao, M., Yoshida, S., Nabeshima, Y., Shimamura, K., and Nakafuku, M. (2001). Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. Neuron *31*, 757–771. doi:10.1016/S0896-6273(01)00413-5.

44. Jang, S., Gumnit, E., and Wichterle, H. (2024). A human-specific progenitor sub-domain extends neurogenesis and increases motor neuron production. Nature Neuroscience *27*, 1945–1953. doi:10.1038/s41593-024-01739-8.

45. Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence *19*, 711–720. doi:10.1109/34.598228.

46. Kurtz, T. G. Approximation of Population Processes. Society for Industrial and Applied Mathematics (1981). doi:10.1137/1.9781611970333.

47. Jacobsen, M. Homogeneous gaussian diffusions in finite dimensions. Preprint No. 3 Institute of Mathematical Statistics, University of Copenhagen (1991).

48. Schäfer, M., Kinzel, D., and Winkler, C. (2007). Discontinuous organization and specification of the lateral floor plate in zebrafish. Developmental Biology *301*, 117–129. doi:https://doi.org/10.1016/j.ydbio.2006.09.018.

49. Dessaud, E., Yang, L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitch, B. G., and Briscoe, J. (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. Nature *450*, 717–720. doi:10.1038/nature06347.

50. Zhang, S., and Stumpf, M. P. H. (2023). Learning cell-specific networks from dynamical single cell data. bioRxiv. https://www.biorxiv.org/content/10.1101/2023.01.08.523176v2. doi:10.1101/2023.01.08.523176. Preprint.

51. Stumpf, M. P. (2021). Inferring better gene regulation networks from single-cell data. Current Opinion in Systems Biology *27*, 100342. doi:10.1016/j.coisb.2021.05.003.

52. Arnold, V., Afrajmovich, V., Il'yashenko, Y. S., and Shil'nikov, L. Dynamical systems V: bifurcation theory and catastrophe theory vol. 5. Springer Science & Business Media (2013).

53. Hirsch, M. W. Differential Topology vol. 33. 'Springer Science & Business Media' (2012).

54. Camacho-Aguilar, E. New mathematical methods for the study of stem cell differentiation. Ph.D. thesis University of Warwick Warwick, UK (2018). https://wrap.warwick.ac.uk/id/eprint/108668/.

55. Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. (2009). Approximate bayesian computation scheme for parameter inference and model selection in dynamical systems. Journal of The Royal Society Interface *6*, 187–202.

56. Klinger, E., Rickert, D., and Hasenauer, J. (2018). pyabc: distributed, likelihood-free inference. Bioinformatics *34*, 3591–3593.

57. Schälte, Y., Klinger, E., Alamoudi, E., and Hasenauer, J. (2022). pyabc: Efficient and robust easy-to-use approximate bayesian computation. Journal of Open Source Software *7*, 4304.

58. Balaskas, N., Azevedo, G. R., Panov, J. K., Marti, E., Briscoe, J., and Anderson, K. R. M. (2012). Gene regulatory logic for reading the sonic hedgehog signaling gradient in the vertebrate neural tube. Cell *148*, 273–284. doi:10.1016/j.cell.2011.10.047.

59. Melton, S., and Ramanathan, S. (2021). Discovering a sparse set of pairwise discriminating features in high-dimensional data. Bioinformatics *37*, 202–212.

60. van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-sne. Journal of Machine Learning Research *9*, 2579–2605.

61. McInnes, L., Healy, J., Saul, N., and berger, L. G. (2018). Umap: Uniform manifold approximation and projection. Journal of Open Source Software *3*, 861. https://doi.org/10.21105/joss.00861. doi:10.21105/joss.00861.

62. Van Der Maaten, L. Learning a parametric embedding by preserving local structure. In: *Artificial intelligence and statistics*. PMLR (2009):(384–391).

63. Sainburg, T., McInnes, L., and Gentner, T. Q. (2021). Parametric umap embeddings for representation and semisupervised learning. Neural Computation *33*, 2881–2907.

64. Deconinck, L., Cannoodt, R., Saelens, W., Deplancke, B., and Saeys, Y. (2021). Recent advances in trajectory inference from single-cell omics data. Current Opinion in Systems Biology *27*, 100344.

65. Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. Nature biotechnology *37*, 547–554.

66. Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Development *11*, 774–785. doi:10.1101/gad.11.6.774.

67. Badia-i Mompel, P., Casals-Franch, S., Aguilar-Mogas, A., Piulachs, M.-D., Guigó, R., Pérez-Brocal, V., Moya, A., Montañez, R., Pérez-R., I., and Conesa, A. (2022). Comparison and evaluation of methods to infer gene regulatory networks from single-cell rna-seq data. bioRxiv. doi:10.1101/2022.12.20.521206.

68. Kamenev, D., Kameneva, P., and Adameyko, I. (2025). The role of microheterogeneity in cell fate decisions in neural progenitors and neural crest. Current Opinion in Neurobiology *65*, 101316. doi:10.1016/j.conb.2025.101316.

69. Balázsi, G., van Oudenaarden, A., and Collins, J. J. (2011). Cellular decision making and biological noise: From microbes to mammals. Cell *144*, 910–925. doi:10.1016/j.cell.2011.01.030.

70. Boukacem, N. E., Leary, A., Thériault, R., Gottlieb, F., Mani, M., and François, P. (2024). Waddington landscape for prototype learning in generalized hopfield networks. Physical Review Research *6*, 033098. doi:10.1103/PhysRevResearch.6.033098.

71. Magin, T. M., McWhir, J., and Melton, D. W. (). A new mouse embryonic stem cell line with good germ line contribution and gene targeting frequency. *20*, 3795–3796. doi:10.1093/nar/20.14.3795.