UNIVERSITY OF CALIFORNIA,
IRVINE

Multi-scale modeling for cell fate specification during regeneration and development

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mathematics

by

Yangyang Wang

Dissertation Committee:
Professor Qing Nie, Chair
Professor German Enciso
Professor Maksim Plikus

2020

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to express my gratitude to all those who helped me out all the time. It is impossible to complete my doctor of philosophy degree without their help and supports.

I would like to extend my sincere gratitude to my advisor and my committee chair, Dr. Qing Nie, for his constant guidance and encouragement. He always gives me visionary suggestions on my research, and a lot of these suggestions help me to have a breakthrough of my research. His encouragement also supports me to keep my confidence and let me be proud of my research. I would be able to finish this thesis without his support.

I would like to express my appreciation to Plikus's Lab: Dr. Maksim Plikus and Christian Guerrero-Juarez. We published a great paper for skin wound healing. I also learned a lot about the biological knowledge and the idea to connect biology and mathematics from them.

I am thankful to Cho's Lab: Dr. Ken Cho. We collaborated on the embryo development project and I learned a lot about the biology insight of the embryo. The valuable discussions are critical to the success of our work.

I would like to express my gratitude to my committe members, Dr Maksim Plikus and Dr. German Enciso, for their insightful comments and encouragement. I would like to thank the Department of Mathematics, Center for Complex Biological System and Center for Multiscale Cell Fate Research for their help and cares for my lives and studies. Especially, many thanks to Karen Martin, Naomi Carreon, Donna McConnell, Aubrey Rudd, Clare Cheng and Kate Haubert.

I would like to thank all my former and current members of Nie lab:, Tian Hong, Chunhe Li, Adam Maclean, Lina Meinecke, Christian Guerrero-Juarez, Axel Almet, Honglei Ren, Emmanuel Dollinger, Floyd Maseda, Matt Karikomi, Kevin Johnston, Daniel Bergman, Jiajun Zhang, Chris Rackauckas, Seth Figueroa, Catherine Ta. I appreciate the suggestions and support from many people for helping me go through the tough time: Zixuan Cang, Peijie Zhou, Yuchi Qu, Shuxiong Wang, Suoqin Jin, Lihua Zhang, Yutong Sha, Tao Peng, Weitao Chen, Qixuan Wang and Hujing Du.

VITA

Yangyang Wang

EDUCATION

Doctor of Philosophy in Computational Mathematics **2014-now**
University of California, Irvine *Irvine, CA*

Bachelor of Science in Mathematics **2009-2013**
University of Science and Technology of China *Hefei, Anhui, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2015–now**
University of California, Irvine *Irvine, CA*

TEACHING EXPERIENCE

Teaching Assistant **2015–2020**
University of California, Irvine *Irvine, CA*

PUBLICATION: published

H. Du, Y. Wang, D. Haensel, B. Lee, X. Dai, and Q. Nie. Multiscale modeling of layer formation in epidermis. *PLoS computational biology*, 14(2):e1006006, 2018.

Y. Wang, C. F. Guerrero-Juarez, Y. Qiu, H. Du, W. Chen, S. Figueroa, M. V. Plikus, and Q. Nie. A multiscale hybrid mathematical model of epidermal-dermal interactions during skin wound healing. *Experimental dermatology*, 28(4):493-502, 2019.

PUBLICATION: under review

Z. Cang*, Y. Wang*, Q. Wang, K. Cho, W. Holmes and Q. Nie. A multiscale model via single-cell transcriptomics reveals robust patterning mechanisms during early mammalian embryo development. *PLoS computational biology*.

*: equal contribution

ABSTRACT OF THE DISSERTATION

Multi-scale modeling for cell fate specification during regeneration and development

By

Yangyang Wang

Doctor of Philosophy in Mathematics

University of California, Irvine, 2020

Professor Qing Nie, Chair

Tissue regeneration and development involves activities in multiple scales: gene, growth factor, cell, and tissue. As an example for regeneration, in wound healing process, the tissue recovery is achieved by cell hyper-proliferation and cell movement through regulations of growth factors. Early embryo development, as an example for development, the overall embryo spatial organization depends on cell movement regulated by cell-cell adhesion genes while cell differentiation is controlled by cell fate genes. In order to study interplay among those scales, multi-scale hybrid models are developed to incorporate gene and cell information, and their emergent dynamics in tissue development and regeneration.

In chapter 2, I established a model to study skin wound healing, which focuses on mechanisms to reduce scar after wound, especially the epidermis-dermis interaction. Each individual cell in epidermis is modeled discretely using Subcellular Element Method (SEM) to study the heterogeneous cell activities, whereas the cells and extracellular matrix (ECM) in dermis are modeled by partial differential equations (PDE) in continuum. To systematically study the role of signaling factors produced by cells, the model incorporate the signaling factors in continuum as well. In further, to study the interface between discrete epidermis and continuum dermis and their interaction, the interface is modeled using Level Set Method (LSM). The model makes several predictions: First, the signaling factors in both epidermis and dermis

are essential to maintain dermal stability; Second, wound-triggered increase production of signaling factors in epidermis and fast re-epithelialization kinetics reduce wound size; Third, high density fibrin clot leads to a raised, hypertrophic scar phenotype, whereas low density fibrin clot leads to a hypertrophic phenotype. Fourth, shallow wounds, compared to deep wounds, result in overall reduced scarring.

In Chapter 3, to study pattern formation of early embryo development, I created a data-informed multi-scale model to reveal the time evolution of gene expression and spatial arrangement in single cell level, allowing us to study both the mechanisms and the effective times of the mechanisms. The cells are modeled in SEM to study the heterogeneity, and the gene expressions of each cell are modeled by stochastic differential equations (SDE). Analysis on single cell RNA sequencing data both validates modeled gene expression and calibrates coefficients for physical cellular interactions in the model. The model discovered that an Epha4/Ephrinb2 gene driven cell adhesion between epiblast (EPI) and primitive endoderm (PE) ensures spatial embryo organization; a good time window for gene regulation involving fibroblast growth factor (FGF) is essential for cell to change cell fate into EPI and PE successfully.

Chapter 1

Introduction

The tissue regeneration and development processes contain activities of cells, growth factors and genes (or transcription factors of genes). As an example of generation, in wound healing process, there're two major activities involving cells: (1) re-epithelialization that epithelial cells around the wound margin begin to migrate into the wound area [103, 35]; (2) dermal tissue restoration, especially the extracellular matrix (ECM) deposition produced by fibroblast cells [35]. These two separate activities are related through growth factors: both epithelial cells and fibroblast cells secrete TGF transforming growth factors-β (TGF-β), which affect the collagen production from fibroblast [88, 23, 73]. In development, for example, the early embryo development consists of the formation of different type of cells: trophectoderm cells (TE), inner cell mass (ICM), primitive endoderm cells (PE) and epiblast cells (Epi), whereas TE/ICM cell fate decision are controlled by Cdx2 and Oct4 genes, and PE/Epi cell fate decision are controlled by Nanog/Gata6 genes [105]. Those biological facts indicate that tissue development and generation are controlled by cell activities, and cell activities are controlled by growth factors and genes. To establish a mathematical model to study tissue development precisely, a mutiscale model containing both cell activity controlled by growth factors and genes is necessary. The question is: how to model cell activities and how to model growth

factors and gene regulation on cells?

There are three essential parts in cell activities: cell fate decision, cell division and cell movement. Cell fate could be controlled by some specific genes (For example, PE/Epi cells are controlled by Nanog/Gata6 genes). Cell fate could also change after division, for example, when stem cells perform division, the daughter cells may have different cell types compare to the mother cell [84]. The second case of cell fate decision and cell division can be modeled together in continuum using ordinary differential equations (ODE) [55]. In this case, the cell populations of each cell type are modeled in continuum using ODE and controlled by cell division with two parameters: cell cycle the probability that the cell type of the daughter cell is same as the mother cell. In further, cell population with spatial information can be modeled using partial differential equations (PDE) with a convection and diffusion term to model cell movement [76]. Continuum model could transfer cell fate decision, cell division and cell movement into terms in ODE and PDE, it could capture cell population but lack some single cell activities like the direction of division plane when cell divides. To capture single cell information, discrete model is necessary. One useful discrete model is subcellular element method (SEM) [69]. SEM model each cell in several elements, which automatically describes the shape of the cell, and such elements could also model the polarity of the cell by giving different property on different elements in one cell [30]. SEM could model cell division by choosing a division plane and separate all elements into two parts, and it is useful when modeling division with specific division plane, for example, A. Gord, et al. utilized SEM to study division plane of the asymmetric division of epidermal stem cells [30]. SEM could also model cell adhesion and repulsion by setting potential between intercelluar elements [69]. In further, selective cell adhesion can be modeled in SEM by setting different potentials between different cell types to achieve cell sorting between different cell types [17]. In summary, the discrete model such as SEM could model some properties in single cell level, like division plane of the cell, cell adhesion and repulsion and selective adhesions. However, the SEM model needs to model every cell in the tissue, which could cause high time complexity. One

solution to reduce the time complexity of SEM is to use the parallel computing with GPU [12].

Cell activities are also controlled by growth factors and genes. Growth factors and transcription factors of genes could be modeled in continuum using PDE by describing the concentration. In this case, cells can be modeled in continuum, and the cell movement (convention term) or cell proliferation can be controlled by the density of the growth factor or transcription factor. As an example, S. N. Menon, et al. established a wound healing model that describes epidermal cells, fibroblast cells and TGF-β in continuum, and TGF-β controls epidermal cell proliferation and fibroblast movement [62]; Cells can also be modeled discretely and single cell property can be controlled by the density of the growth factor/transcription factor. H. Du, et al. created a mathematical model controlling single cell division cycle and regeneration probability based on the density of Ovol transcription factors [17]. Genes and growth factors can also be modeled for each single cell using ODE, and it is possible to model ligand-receptor based cell-cell interaction using ligand level of the neighbor cells. S. Bessonnard, et al. studied the cell-cell interaction on Nanog/Gata6 level via a model that describes Nanog/Gata6 level of each single cell using ODE, and the Nanog/Gata6 level are controlled by Fgf4/Fgfr2 ligand-receptor pairs of the neighbor cells [4].

Here I developed two models combining cells activities and growth factor/gene regulation on cells: (1) In Chapter 2, I established a wound healing multiscale model on interaction of epidermal and dermal cells [108], where the novelty of this work is that I modeled epidermis in discrete and dermis in continuum with a moving boundary between them, and describes the interaction between them via growth factors (modeled in continuum); (2) In Chapter 3, I cretaed a multiscale model via early embryo development (this paper is under review), and combining single-cell transcriptomics data with SEM models is the novel part of the model.

Chapter 2

A multiscale hybrid mathematical model of epidermal-dermal interactions during skin wound healing

Following injury, skin activates a complex wound healing programme. While cellular and signalling mechanisms of wound repair have been extensively studied, the principles of epidermal-dermal interactions and their effects on wound healing outcomes are only partially understood. To gain new insight into the effects of epidermal-dermal interactions, we developed a multiscale, hybrid mathematical model of skin wound healing. The model takes into consideration interactions between epidermis and dermis across the basement membrane via diffusible signals, defined as activator and inhibitor. Simulations revealed that epidermal-dermal interactions are critical for proper extracellular matrix deposition in the dermis, suggesting these signals may influence how wound scars form. Our model makes several theoretical predictions. First, basal levels of epidermal activator and inhibitor help

to maintain dermis in a steady state, whereas their absence results in a raised, scar-like dermal phenotype. Second, wound-triggered increase in activator and inhibitor production by basal epidermal cells, coupled with fast re-epithelialization kinetics, reduces dermal scar size. Third, high-density fibrin clot leads to a raised, hypertrophic scar phenotype, whereas low-density fibrin clot leads to a hypotrophic phenotype. Fourth, shallow wounds, compared to deep wounds, result in overall reduced scarring. Taken together, our model predicts the important role of signalling across dermal-epidermal interface and the effect of fibrin clot density and wound geometry on scar formation. This hybrid modelling approach may be also applicable to other complex tissue systems, enabling the simulation of dynamic processes, otherwise computationally prohibitive with fully discrete models due to a large number of variables.

2.1 Introduction

Skin functions as a vital interface between organism and its environment. When injured, skin rapidly heals via a wound healing programme characterized by four distinct yet overlapping phases: hemostasis, inflammation, proliferation and remodelling [19, 100]. The primary goal of this repair programme is to re-establish barrier function by reforming stratified epidermis and restoring dermal tissue integrity [103, 35]. During the remodelling phase, regeneration of new hair follicles [45, 67, 28, 52] and adipose tissue [80, 32] can also take place. Cellular or molecular defects during any of these phases can lead to pathological wound healing outcomes.

Molecular composition and high-order structure of collagen bundles laid by wound fibroblasts determine the architecture and the "quality" of the dermal scar, and collagen deposition is tightly regulated by paracrine, autocrine and mechanical signals [101]. While collagen architecture of normal skin is intricate and weave-like, scars typically have "less desirable"

highly parallel collagen [112, 113]. Epidermal keratinocytes can signal to dermal fibroblasts to decrease collagen production [27], while direct contact between two cell types stimulates keratinocyte proliferation and migration during wound re-epithelialization [109]. Epidermal-dermal crosstalk occurs via signalling growth factors [111], and, for example, both wound keratinocytes and fibroblasts secrete transforming growth factor beta (TGFβ) ligands, the key wound healing mediatorsc [88, 23, 73]. TGFβ1 is also abundant in the platelet-rich fibrin clot and has been linked to excessive extracellular matrix (ECM) deposition[21]. This suggests that together with keratinocytes, fibroblasts and immune cells, fibrin clot serves as an important signalling centre during wound repair. While multiple lines of evidence point towards the importance of signalling crosstalk between key cellular and molecular components of the wound, the overall logic of these interactions during scar formation remains incompletely understood.

Mathematical modelling offers a useful approach to study principles of wound healing. In the past, reaction-diffusion model has been implemented to study biochemical regulation of cell cycle in the epidermis during wound re-epithelialization [94]. Another study implemented hybrid model to study how collagen fibres organize during wound healing and their role in scarring [13, 61]. The Cellular Potts model also has been used to investigate sprouting and branching during angiogenesis [14] and, more recently, to examine proliferation and migration of skin wound fibroblasts [90].

While the crosstalk between keratinocytes and fibroblasts via soluble signalling factors has been previously analysed using a continuum model [63], epidermal and dermal compartments were treated as one spatially homogeneous region, without taking into consideration their interactions and other important anatomical aspects of the skin. A finite element method also has been used to study interactions between keratinocytes, fibroblasts and endothelial cells [104]. Although important, this method did not account for the interface between epidermis and dermis and, instead, modelled both of these distinct skin compartments as a

continuum. Unlike dermis, epidermis consists only of a few layers of densely packed cells, and its discrete nature likely introduces important biological effects that cannot be captured by a continuum model.

To simulate interactions between individual cells within spatially distinct epidermis and dermis and their effects on scarring, we developed an optimized two-dimensional, multiscale hybrid model of wound healing. This model incorporates dynamically continuous epidermal-dermal interface, and allows studying individual proliferating and migrating keratinocytes using discrete single-cell model, and dermal fibroblast functions using continuum model. Our simulations predict that both dermal and epidermal signalling factors are necessary to sustain steady-state ECM levels in normal skin, and to regulate new ECM deposition after wounding. Our model also predicts that high-density fibrin clot, serving as the source for signalling factors, can induce formation of raised hypertrophic scars, whereas low-density fibrin clot has the opposite effect. Finally, the model predicts causal relationships between initial wound width and depth and scarring outcomes.

2.2 Methods

The two-dimensional, multiscale hybrid mathematical model consists of two submodels allocated on two separate homogeneous regions: the dermal (D) and epidermal (E) regions. Each region is distinct and was modelled with a dynamic interface between them, termed Ω. Epidermal cells in E were modelled individually to evaluate their signal production and proliferation and migration abilities. In contrast, dermal fibroblasts, immune cells, ECM (including collagen bundles), and signals produced by cell types not explicitly accounted for (such as vascular cells) were modelled in a continuum.

2.2.1 Discrete model on epidermal cell dynamics using subcellular element method

Epidermal keratinocytes were modelled individually through a linear cell lineage composed of basal, spinous and granular layer keratinocytes. Following our previous work [17], a selective differential cell adhesion mechanism was implemented to allow for proper layer stratification. Individual cells and their divisions were modelled by a subcellular element method [69] that has been adapted to specifically study epidermis [30] (see Section 2.5).

2.2.2 Continuum model on signalling and dermal components

Extracellular matrix deposition and diffusive signalling molecules in both dermis and epidermis were modelled by reaction-diffusion differential equations. Dermal fibroblasts and immune cells were modelled using the Keller-Segel model[37], which includes reaction-diffusion-advection equations controlling their diffusion, chemotactic movement, self-renewal and decay.

2.2.3 Coupling of discrete and continuum models with dynamic interface

We modelled epidermal basement membrane as a dynamic interface (Ω) to separate the dermis and epidermis and used the movement of such interface to mimic dermal scar formation dynamics during wound healing. The dynamic nature of the interface is implicitly modelled by the level set method [75] (see Section 2.5).

2.3 Results

2.3.1 A multiscale model with a dynamic interface and epidermal-dermal interactions

We constructed a two-dimensional, multiscale mathematical model composed of a hybrid epidermal-dermal interface to explore the mechanisms that regulate aspects of skin physiology during homeostasis and after wounding. Epidermis (E) was modelled individually with discrete keratinocytes, whereas dermis (D) was modelled in a continuum and considers diffusive signals produced by keratinocytes, dermal fibroblasts, immune cells and other skin cells not explicitly accounted for (see Methods). Both compartments were modelled independently and separated by a dynamic interface, termed Ω . During wound healing, ECM production and fibroblast proliferation are known to be regulated by multiple signalling factors [24, 42, 2, 54, 95]. Key signalling factors involved in wound healing are members of the TGF pathway, although their effects are complex. While TGFβ1 promotes collagen synthesis [24, 22], TGFβ3 shows downregulating effects [42, 66]. Both fibroblasts and keratinocytes secrete TGFβ1 [2], and fibroblasts respond to TGFβ1 by proliferating [54]. This suggests that TGFβ1 serves an activator for fibroblast proliferation and ECM production, whereas TGFβ3 works as an inhibitor of ECM production. Signalling factors belonging to canonical WNT [106, 110] and PDGF pathways [95, 46, 39] also regulate fibroblast activity and can serve the role of activators and inhibitors in the model.

To systematically study the role of putative regulatory signals during scar formation, we modelled diffusive signals as activator (A) and inhibitor (I). A and I assume specific roles in our model—A promotes fibroblast proliferation and ECM production, whereas I inhibits ECM production. The model obeys three basic conditions (Figure 2.1). First, both basal epidermal keratinocytes and dermal fibroblasts produce A and I. Immune cells can also

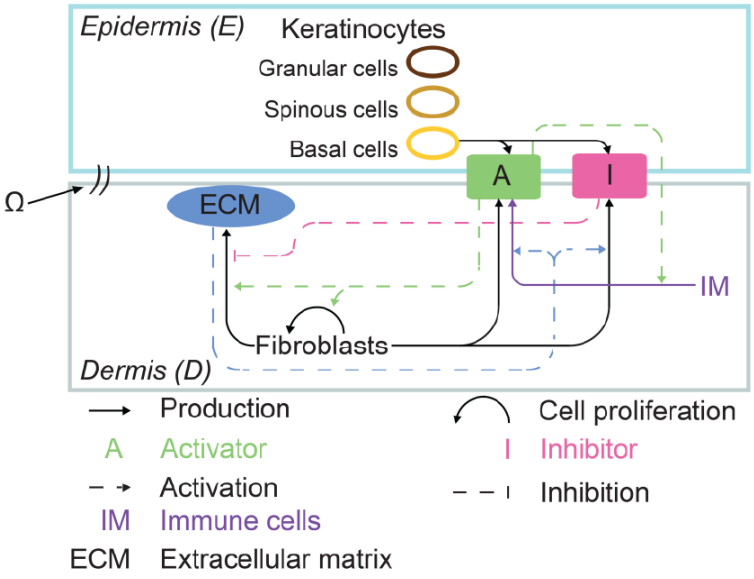

Figure 2.1: Schematic of two-dimensional multiscale hybrid model of wound healing. The modelling domain is separated into epidermis (E) (light blue) and dermis (D) (grey) by a dynamic interface (Ω) to mimic the basement membrane. In E, keratinocytes are modelled discretely as basal (yellow), spinous (light brown) and granular (dark brown) cells. Basal cells are set to produce activators A (shaded green box) and inhibitors I (shaded dark pink box) (rule (1)). In D, fibroblasts, ECM (shaded blue oval) and immune cells (IM, purple) are modelled in a continuum, where fibroblasts produce A and I, immune cells produce A, and these processes are directly activated by ECM (rule (2)). Additionally, fibroblast proliferation is activated by A, production of A by immune cells is activated by A, and ECM production is activated by A and inhibited by I (rule (3)). A and I exist in both D and E and can diffuse across Ω

produce A. Second, A and I diffuse through the dermis, and their production rates are directly regulated by fibroblast density and ECM production. Production of A by immune cells is regulated by their activation state, which they enter when local A signalling levels are high and above a set threshold. Third, A positively regulates ECM production, while I does the opposite. These processes are modelled in a continuum and assume the following sets of partial differential equations:

$$\frac{\partial F}{\partial t} = \nabla(D_F \nabla F - v_F F \nabla p_2) + f_F(F, p_2) - d_F F \tag{2.1}$$

$$\frac{\partial C}{\partial t} = D_C \Delta C + f_C(F, p_1, p_2) - d_C C \tag{2.2}$$

$$\frac{\partial p_1}{\partial t} = D_{p_1} \Delta p_1 + f_{p_1}(F, C, p_1) + E_{p_1} - d_{p_1} p_1 \tag{2.3}$$

$$\frac{\partial p_2}{\partial t} = D_{p_2} \Delta p_2 + f_{p_2}(F, C, p_2) + E_{p_2} + C_{Mp_2} M + f_{IM}(IM, p_2) - d_{p_2} p_2 \tag{2.4}$$

where F and C represent fibroblast proliferation/density and overall ECM concentration (measured as a direct readout of collagen bundle production/deposition), respectively; p_1 and p_2 represent the concentrations of I and A, respectively. The term E_{p_i} represents the production rate of p_i (where i = 1,2 in epidermal basal keratinocytes), and IM represents immune cells, which also serves as a source of A (see Section 2.5). M represents fibrin clot, which forms during wound healing (see Section 2.5).

The dynamic interface between E and D was implicitly modelled by the level set method and implemented by the phase function Φ, which was determined by the velocity field ν. The change in phase function Φ was defined by:

$$\frac{\partial \Phi}{\partial t} = -v \cdot \nabla \Phi \tag{2.5}$$

where Φ is the phase function representing the epidermal region (Φ ¿ 0) and dermal region

(Φ ¡ 0). The initial condition for Φ is the signed distance between the grid point and the interface (see Section 2.5). The velocity field v is encoded by:

$$v = -K\nabla(C - C_0) \tag{2.6}$$

where $C_0(x, y, t)$ is zero when (x, y) is in D, and $C_0(x, y, t)$ is $C_{\text{stable}} \neq 0$ when (x, y) is in E. C_{stable} is ECM density in the dermis under homeostatic conditions. The interface will rise up if $C > C_{\text{stable}}$ is satisfied near the interface and decrease $C < C_{\text{stable}}$ is satisfied.

2.3.2 Coupled signalling between basal keratinocytes and fibroblasts maintains dermal homeostasis in the model

To examine the effects of epidermally derived A and I on dermis stability, we varied the parameters regulating signal production by basal keratinocytes. Initially, A and I were modelled such that they are found at near-constant levels throughout D and in basal layer of E. (Figure 2.2 A", A"'). Simulations suggest that if both basal keratinocytes and fibroblasts produce A and I at relatively high levels and immune cells do not produce extra A because of the combined keratinocyte-and fibroblast-derived A levels are below the threshold required for immune cell activation (see Section 2.5), ECM and fibroblast densities (which serve as proxies for dermal stability) are uniformly distributed across D (Figure 2.2 A, A'), but form an upward gradient in E (Figure 2.5A–A"'). The average ECM and fibroblast densities, as well as A and I levels, remained stable and did not fluctuate over a simulated timescale of 6 days (Figure 2.2 E, F, black line). We used these simulation parameters as a baseline to represent the homeostatic condition in the following simulations, where A and I production rates and sources were perturbed.

Next, we tested the effect of disabled epidermal A and I on dermal homeostasis (Figure 2.2

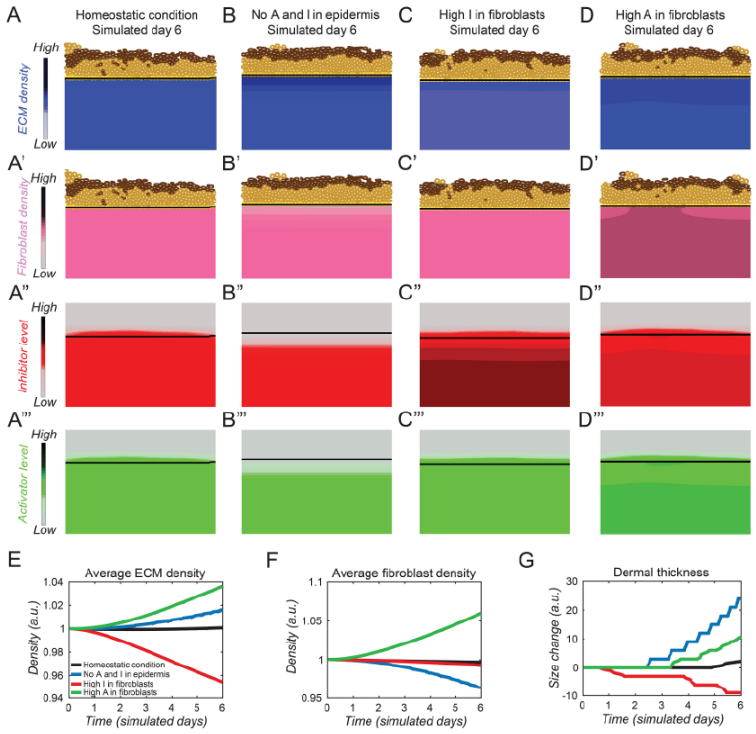

Figure 2.2: Dermal homeostasis in normal skin depends on epidermal and dermal signalling. (A-A"') Simulated ECM and fibroblast densities in unwounded skin with A and I produced both by keratinocytes and by fibroblasts. Black line denotes the position of basement membrane relative to the simulated skin surface. Fibroblast density, and I and A levels are shown in (A'), (A") and (A"'), respectively. (B) Simulated ECM and fibroblast densities in the model where only fibroblasts produce A and I. Significantly decreased fibroblast density, and inhibitor and activator levels are shown in (B'), (B") and (B"'), respectively. (C) Simulated ECM and fibroblast densities in the model where fibroblasts produce 10% more I relative to the homeostatic condition (high I). Corresponding fibroblast density, and inhibitor and activator levels are shown in (C'), (C") and (C"'), respectively. (D) Simulated ECM and fibroblast densities in the model where fibroblasts produce 10% more A relative to the homeostatic condition (high A). Corresponding fibroblast density, and inhibitor and activator levels are shown in (D'), (D") and (D"'), respectively. (E) Temporal change in average ECM density across modelling time. (F) Temporal change in average fibroblast density across modelling time. (G) Temporal change in dermal thickness, as measured by the position of basement membrane relative to the skin surface, across modelling time. Colour definitions for each line on E-G are provided on the figure. Values along X-axis are in simulated days, and values along Y-axis are in arbitrary units (a.u.)

B", B"'). This change led to near-constant levels of ECM and fibroblast densities in D, except near the epidermal-dermal interface (Figure 2.2B, B'), where A and I gradients shifted downwards into D (Figure 2.5B–B"') and over time fibroblast density decreased (Figure 2.2F, blue line), while ECM density increased (Figure 2.2E, blue line). This suggests that epidermal activator signals are primarily involved in promoting fibroblast self-renewal, while epidermal inhibitor signals prevent excessive ECM synthesis within the immediate basement membrane microenvironment. Increase in ECM output under these signalling perturbations led to dermal thickening despite modest decrease in fibroblast density (Figure 2.2 G, blue line). This is because in the model ECM contributes to dermal volume changes substantially more as compared to fibroblasts. This assumption is in line with the observations that ECM occupies larger proportion of a given dermal volume in adult mouse skin as compared to fibroblasts [90]. These results also suggest that epidermal signals (primarily I) can supplement dermal signals and contribute to maintaining dermal skin compartment in homeostasis.

Next, we modelled the effects of perturbed dermal signalling. We varied the production rate of A or I in fibroblasts to two opposite extremes and simulated changes in ECM, fibroblast density and dermal thickness as measures of dermal stability. Simulation results suggest that when the production rate of I in fibroblasts is increased by more than 10% in the upper D compartment (Figure 2.2 C"), fibroblast density modestly decreased and ECM production significantly decreased (Figure 2.2 C, C'), and continued to decrease over the entire simulation period (Figure 2.2E, F, red lines). As expected, these changes resulted in continuously reducing dermal thickness (Figure 2.2 G, red line). On the other hand, when the production rate of A in fibroblasts is increased by more than 10% (Figure 2.2 D"'), densities of ECM and fibroblasts as well as dermal thickness increased (Figure 2.2 D, 2D'), and these changes continued over the simulated time (Figure 2.2E–G,green lines). Together, these modelling results suggest that both epidermal and dermal sources for A and I signals and their balance are likely necessary for dermal homeostasis.

2.3.3 Increased signalling and fast re-epithelialization kinetics are essential for scar-free healing in simulated epidermal abrasion wounds

Next, we asked how A and I signals may regulate dermal repair during wound healing. First, we modelled healing of epidermal abrasion wounds. Epidermal abrasions mainly heal by re-epithelialization, and no dermal scar typically forms [85]. We carried out several simulations in which basal and suprabasal keratinocytes are stripped, while dermal fibroblasts and ECM remain intact (Figure 2.6A, B; see Section 2.5 for modelling details). We modelled two parameter kinetics that allow for "fast" and "slow" re-epithelialization in order to simulate kinetics of normal and delayed wound healing, respectively, and evaluated their effects on dermal homeostasis during and after re-epithelialization (Figure 2.3). Both re-epithelialization kinetics were modelled under "basal" (ie unchanged) and "high" (ie increased) epidermal A/I production rates (Figure 2.7; see Section 2.5). We also assumed that immediately after epidermal abrasion, A/I levels in the wound area drop below normal levels due to loss of keratinocytes, which function as the source for signals (Figure 2.6C, D).

First, we simulated the effects of unchanged vs increased A/I epidermal signalling under fast re-epithelialization kinetics (Figure 2.3A–D"'). Under basal signalling conditions, A/I levels were normal at the wound edges but dropped directly beneath the epidermal wound (Figure 2.3C", C"'). Under these simulated conditions, wound re-epithelialization was accompanied by an increase in dermal thickness (Figure 2.3K, blue line), an outcome that is unusual for epidermal abrasions. This is underlined by an increased ECM deposition (Figure 2.3D and 2.3I, blue line), which is the consequence of sustained lower I levels. Density of fibroblasts experienced a modest reduction (Figure 2.3J, blue line), opposite to ECM changes. This is also an unusual outcome, and it is driven by the disproportionally higher sensitivity of ECM production to I vs A. In contrast, when epidermal A/I levels were allowed to increase after

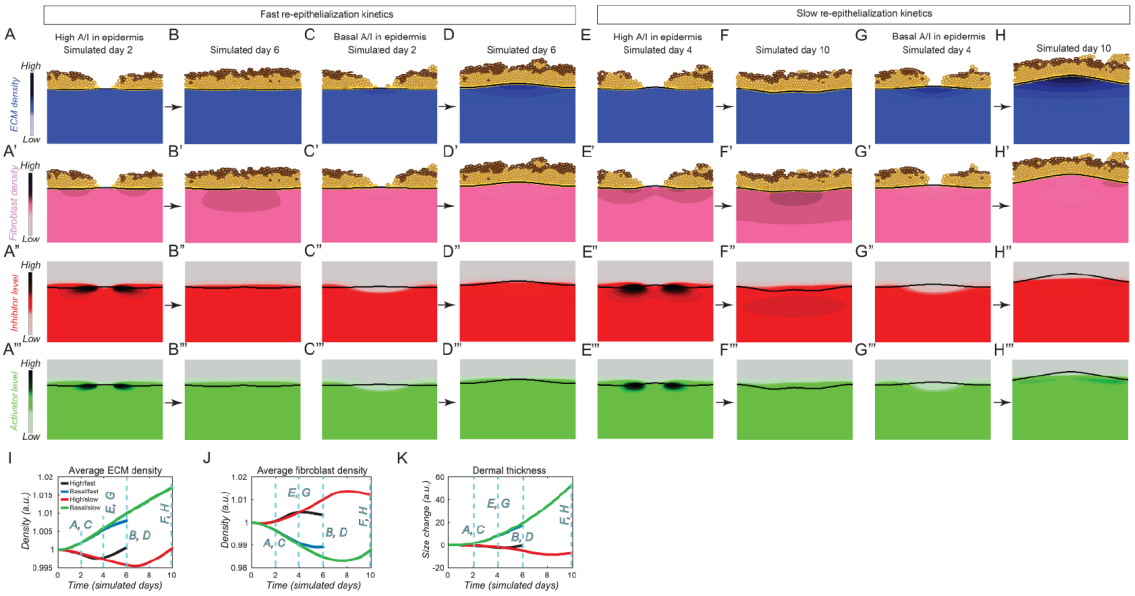

Figure 2.3: Healing outcomes of abrasion wounds depend on re-epithelialization kinetics and epidermal signal production rate.
Abrasion wound healing simulations were performed using fast (A-D"') and slow re-epithelialization kinetics (E-H"'). For each set of re-epithelialization kinetics, simulations were ran using "basal" (C-D"', G-H"') and "high" epidermal A/I production rates (A-B"', E-F"'). For each simulated condition, ECM density (blue), fibroblast density (pink), inhibitor level (red) and activator level (green) are shown. For each simulation, two consecutive snapshots are shown: day 2 and 6 for fast re-epithelialization simulations and day 4 and 10 for slow re-epithelialization simulations. Black line on each image marks the position of basement membrane. (I) Temporal changes in average ECM density across modelling time. (J) Temporal changes in average fibroblast density across modelling time. (K) Temporal changes in dermal scar thickness across modelling time. Colour definitions for each line on I-K are provided on the figure. Values along X-axis are in simulated days, and values along Y-axis are in arbitrary units (a.u.)

abrasion injury (Figure 2.3A", A"'), dermal thickness of re-epithelialized wounds did not significantly change compared to homeostatic pre-wounding state (Figure 2.3K, black line)—a more biologically realistic outcome. Underlying this dermal compartment behaviour in the model were near-normal fibroblast density (Figure 2.3J, black line) and ECM density (Figure 2.3I, black line). These modelling results suggest that rapid increase in A/I production by the wound edge keratinocytes is likely necessary to compensate for the loss of normal A/I levels due to epidermal injury and for eventual scar-free healing of epidermal abrasions.

Next, we modelled healing of epidermal abrasion wounds under the conditions of slow re-epithelialization kinetics, designed to simulate chronic wound healing (see Section 2.5). When modelling combined slow re-epithelialization with basal A/I production levels (Figure 2.3G– H"'), post-injury ECM deposition (Figure 2.3I, green line) and dermal scarring became exacerbated (Figure 2.3K, green line). Interestingly, when modelling combined slow re-epithelialization with increased rate of A/I production (Figure 2.3E-F"'), delayed re-epithelialization kinetics triggered basal keratinocytes into a state of sustained elevated production of A (Figure 2.3E"') and I (Figure 2.3E"). In turn, such chronically high A/I levels resulted in dermis in and around the wound area to overproduce fibroblasts and decrease ECM (Figure 2.3I and 2.3J, red lines), causing modest, albeit abnormal decrease in dermal thickness over simulated time (Figure 2.3K, red line). Dichotomous behaviour of fibroblasts and ECM in this simulation is driven by higher sensitivity to changes in A and I, respectively. Taken together, our simulations indicate that proper closure of epidermal abrasion wounds and return of injured skin to near-normal homeostasis require both rapid re-epithelialization and increase in epidermal signalling.

2.3.4 Simulating fibrin clot density and different wound geometries predicts their effects on the type and thickness of dermal scar

Previous studies showed that physiological inflammatory processes are necessary to achieve normal full-thickness wound repair, while abnormal inflammation levels may lead to pathological scarring [51]. Important components of the inflammation phase of wound healing are diverse immune cell types [18] and the formation of a fibrin clot [31]. Fibrin clot is initiated by the activation and aggregation of platelets,and platelets release α-granule content, containing mitogenic and chemotactic growth factors important for wound healing [72]. Among these platelet-derived growth factors is TGF$beta$1 [21], which functions as a positive regulator of ECM deposition52. Therefore, in our modelling, we incorporated the effect of immune cell–derived and fibrin clot–derived putative activator (modelled as A) on healing of full-thickness dermal wounds (Figure 2.8). We assumed that fibrin clot forms in the wound bed region during the transition between hemostasis and inflammation phases (the starting point of our simulations) and that it can have high density or be defective (ie low density) (Figure 2.9). We also assumed that fibrin clot serves as a "passive" source for A signal that diffuses through D and E compartments, acts on dermal components and degrades over time (Figure 2.4L, Figure 2.10). For the immune cells, we assumed that they serve as an "active" source for A, but that they produce A only when A levels are already above homeostatic. Naturally, this results in transient activation of immune cells near the wound edge and fibrin clot—the site of elevated A (Figures 2.11, 2.12).

We ran a series of simulations, starting with dermal wounds that are deep and wide (Figure 2.13A), imitating large excisional wounds, and that form high-density fibrin clot (Figure 2.4A-B"'). Under these conditions, A in the wound bed reached high level (Figure 2.4A"'), including from immune cells (Figure 2.11A), triggering fibroblast hyper-proliferation in and

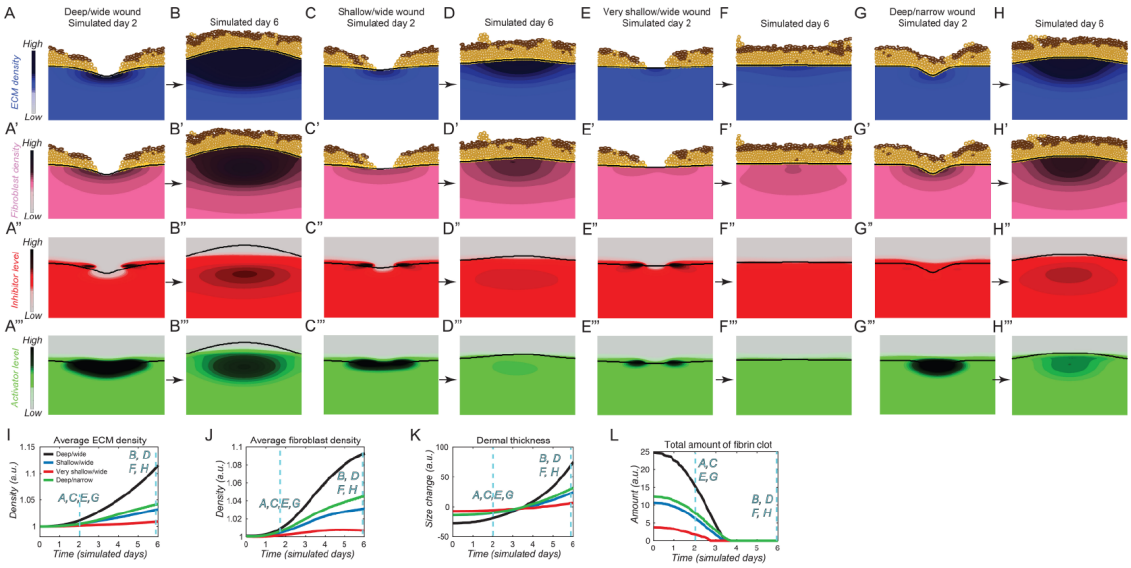

Figure 2.4: Healing outcomes of dermal wounds depend on wound geometry and fibrin clot density.

All dermal wound healing simulations shown on this figure were performed using high-density fibrin clot condition. See Figure 2.14 for low-density fibrin clot condition. The following wound geometries were compared: (A-B"') deep and wide wounds, (C-D"') shallow and wide wounds, (E-F"') very shallow and wide wounds, and (G-H"') narrow and deep wounds. For each simulated condition, ECM density (blue), fibroblast density (pink), inhibitor level (red) and activator level (green) are shown. For each simulation, two consecutive snapshots are shown: day 2 and 6. Black line on each image marks the position of basement membrane. (I) Temporal changes in average ECM density across modelling time. (J) Temporal changes in average fibroblast density across modelling time. (K) Temporal changes in dermal scar thickness across modelling time. (L) Degradation dynamics of the fibrin clot across modelling time. Colour definitions for each line on I-L are provided on the figure. Values along X-axis are in simulated days, and values along Y-axis are in arbitrary units (a.u.)

around the wound (Fig 2.4A', B', J, black line). Fibroblast overproduction was accompanied by ECM overproduction (Fig 2.4A, B, I, black line), and the resulting scar became prominently raised and hypertrophic-like (Figure 2.4K, black line). Importantly, although fibrin clot decayed to zero only after three simulated days (Figure 2.4L, black line), signalling effects of clot-derived A on ECM deposition and fibroblast density persisted beyond that period. This is because expanding fibroblasts overtook as the dominant source of excessive A production (Figure 2.4B"') and immune cells continued to produce A beyond day 3 (Figure 2.11I, black line). Confirming these findings about fibrin clot was simulation of deep and wide wounds with defective, low-density fibrin clot (Figure 2.14A, B"'). Under low-density clot conditions, signalling levels for A and I in and around wound bed remained relatively low (Figure 2.14A", A"') and fibroblast density decreased (Figure 2.14J, black line), while that of ECM remained almost unchanged (Figure 2.14I, black line). Consequently, dermal thickness recovery dynamics were slow, such that raised scar did not develop and, instead, healed wounds acquired depressed, hypotrophic-like appearance.

Next, we modelled the effects of wound geometries on wound healing, comparing deep and wide wounds with: (a) shallow and wide, (b) very shallow and wide, and (c) deep and narrow wounds (Figure 2.13). Under high fibrin clot conditions, wounds of all sizes healed with raised scars of varying degree (Figure 2.4K) that had increased ECM density (Figure 2.4I) and increased fibroblast density (Figure 2.4J). This was underlined by increased A/I levels (Figure 2.4), including increased A from immune cells (Figure 2.11), and depended on initial high A levels originating from high-density clot. At the same time, prominent size-dependent differences were observed. Simulations suggest that the degree of dermal scarring strongly correlates with the wound's depth rather than its width. Indeed, deep and narrow wounds (Figure 2.4G–H"') resulted in higher degree of scarring as compared to very shallow but wide wounds (Figure 2.4E–F"', K, green vs red line).

Interestingly, the above correlations between wound size and scarring outcome were altered

under low-density fibrin clot conditions. Wounds of all sizes acquired depressed, hypotrophic-like appearance (Figure 2.14K), with scars resulting from deep wounds showing decreased fibroblast density (Figure 2.14J). At the signalling level, compared to high-density clot wounds, simulated low-density clot wounds quickly restored A/I levels to pre-wounding state (Figure 2.14).

Lastly, as proof of principle of the model's scalability, we simulated healing of deep and wide wounds with dense fibrin clot with the version of the model that contains two activators: A_1 that only stimulates ECM production and A_2 that only stimulates fibroblast proliferation (Figure 2.15). We tested the dependency of ECM and fibroblasts on fibrin clot–derived activator by assigning the clot to contain only A_1 (Figure 2.16, yellow lines) or A_2 (Figure 2.16, green lines). Simulations showed that fibroblast density highly depended on its activator A2 (Figure 2.16B), but that lack of A_1 can be largely compensated by A_2 for ECM production (Figure 2.16A). This result lays the framework for modelling more complex wound healing scenarios in the future.

2.4 Discussion

In this work, we describe new hybrid mathematical model designed to study the effects of epidermal-dermal interactions on the molecular and cellular dynamics, and outcomes of skin wound healing. Our simulations suggest that putative activator and inhibitor signalling factors produced both in the epidermis and in the dermis are important for proper wound repair. If either of these signalling sources were perturbed in the model, dermal homeostasis and repair became altered, underlined by the defects in maintenance and restoration of ECM and fibroblast densities. For example, our model predicts how different signalling regimes affect the degree of skin scarring in the scenario when wound closure kinetics are perturbed. In the context of epidermal abrasion injuries, the ability of basal keratinocytes to rapidly

increase the production of A and I over the baseline homeostatic levels was predicted to be critical for preventing skin scarring in fast re-epithelializing wounds and to minimize scarring in wounds with defective re-epithelialization, that simulates chronic epidermal wounds.

Our simulations also make several predictions regarding the effects of geometry and fibrin clot in dermal wounds. First, modelling results suggest that dermal wounds of all geometries heal by forming scar tissue that differs in its ECM and fibroblast compositions compared to unwounded skin. This is consistent with the available experimental and clinical data that adult partial-thickness and full-thickness wounds repair by scarring [77]. Second, wide but shallow simulated wounds—an equivalent of superficial partial-thickness dermal wounds—repaired with smaller scars as compared to deep wounds of different width, an equivalent of deep partial-thickness and full-thickness dermal wounds. Indeed, this modelling prediction is broadly consistent with the available experimental and clinical data. Superficial dermabrasion wounds, even when relatively wide, typically repair with no noticeable scar [60], while deep, full-thickness wounds even when relatively narrow, such as full-thickness incisional surgical wounds, repair with visible scars [38]. Third, simulations suggest strong effects of fibrin clot on the trajectories and healing outcomes of dermal wounds. Clot density had the strongest effect on the simulated repair of deep wounds, with high-density clots causing distinctly raised, hypertrophic-like scars, while deficient, low-density clots causing somewhat depressed, hypotrophic-like scars. The initial clot density had progressively decreasing influence on the healing outcomes of shallow wounds. In our model, the effect of clot is mediated by the amount of activator that it releases, with dense clots releasing larger activator quantities. Indeed, fibrin clots contain platelet-derived growth factors and recent proteomic studies started to define their composition [98, 1]. Our modelling predictions warrant new study on the effect of fibrin clot in the animal model for wound healing, where clot formation can be regulated.

From the mathematical perspective, our model provides several advantages. Discrete nature

of the epidermal compartment enables to model behaviour of individual keratinocytes, including cell proliferation, migration and signal production. Within its current framework, the model can be easily adjusted to incorporate additional epithelial cell types, such as hair follicle and/or sweat gland cell types. The model accounts for dermal compartment dynamics using continuum description, which eliminates the need for a substantially more complex and computation cost-heavy discrete description. Additional components, as shown with immune cells and two activator species, can be relatively easily added into the current model of the dermal compartment without having to fundamentally change it. Our model also implicitly considers the epidermal (E) and dermal compartments (D) and the Ω boundary between them via phase function, so that a uniform mesh can be used for both E and D. This eliminates the necessity of generating two separate meshes for E and D in order to achieve sharp compartment boundary. This approach can be applied to modelling additional sharp boundaries within the skin, such as hair follicle/dermal boundary.

In terms of its limitations, our model does not consider potential heterogeneity in skin fibroblast populations, and the possibility that different fibroblast subtypes can respond differently to signalling cues and exert distinct effects on wound healing. Indeed, several recent studies have identified distinct populations of mouse skin fibroblasts with distinct roles in ECM deposition during development and wound healing [32, 95, 16, 47, 87, 78]. Rinkevich et al. identified two populations of mouse skin fibroblasts: En1 (Engrailed homeobox 1)-negative and En1-positive cells. The former are abundant during skin development, and the latter increase in numbers late in adulthood and predominantly deposit collagen and remodel ECM during wound healing [87]. Developmental change in En1-positive vs En1-negative fibroblast abundance affects scarring outcomes in skin wounding experiments [47]. Driskell et al. showed that during mouse skin development, two distinct fibroblast types differentially contribute to the formation of the upper papillary and lower reticular dermal layers [16]. Moreover, during wound healing, reticular fibroblasts populate the wound first, before papillary fibroblasts, and they deposit early ECM. Shook et al. showed that unwounded

mouse skin contains three fibroblast populations, that one of them shares high similarity with En1-positive cells, and that their contributions to wound healing are distinct [95]. Using single-cell RNA-sequencing, our group showed that early wound scars in mice contain as many as twelve fibroblast clusters and that they form at least three distinct fibroblast differentiation trajectories [32]. Drawing on the above experimental evidence for fibroblast heterogeneity, it will be of interest to incorporate it into wound healing models.

Indeed, a recently reported mathematical model of wound healing that utilized Cellular Potts model accounted for two fibroblast subtypes—proliferative and collagen-producing fibroblasts [90]. In the model, the switch between two fibroblast types was positively regulated by ECM and was required to achieve dermal scar maturation. Our model herein accounts only for one type of dermal fibroblasts, and ECM production is controlled by putative activator and inhibitor factors with ECM feedback onto both, rather than by a fibroblast state switch. We posit that signalling regulations in combination with uncoupled ECM and fibroblast components within the context of a hybrid modelling approach may afford a more precise description of fibroblast and ECM dynamics and their roles in wound healing.

Lastly, our current model does not fully recapitulate all wound healing phases. Because the model lacks blood vessel and blood flow elements, it does not reproduce the hemostasis phase. Because it contains fibrin clot and immune cells, both of which are signalling sources, the model partially reproduces the inflammation phase. Our model is best suited for studying the proliferation phase but is not fully optimized for simulating long-term scar remodelling. Additional elements with negative feedback function in ECM remodelling will be required for the model to enter new stable steady state, which would recapitulate scar tissue maturation.

Overall, our multiscale hybrid model provides a flexible and efficient computational framework to investigate epidermal-dermal interactions and their effects on wound healing. By systematically adding various biological processes or elements that are critical to wound healing, one may use such modelling framework to delineate and predict novel mechanisms

that have not been previously explored using experimental approaches.

2.5 Modeling details

2.5.1 Subcellular element method for wound healing model

Subcellular element method for cellular dynamics. The subcellular element method describes each individual cell as a collection of subcellular "elements". Cell movement is driven by biomechanical forces, including: (1) intra-cellular forces among elements in the same cell; (2) inter-cellular forces between elements in different cells; and (3) external interactions between the basement membrane and cells. The movement of the i^{th}cell among all N cells (N may be different at different time intervals) is described by the interactions among a collection of elements, which we index by α_i . The position (r_{α_i}) of these elements, whose dynamics are described by the sum total of their interactions with all other elements of the system, evolve according to:

$$\frac{dr_{\alpha_i}}{dt} = -\nabla_{\alpha_i} \sum_{\alpha_i \neq \beta_i} V_{\text{intra}}(|r_{\alpha_i} - r_{\beta_i}|) - \nabla_{\alpha_i} \sum_{i \neq j} \sum_{\beta_j} V_{\text{inter}}(|r_{\alpha_i} - r_{\beta_j}|) + F_{\text{external}}(r_{\alpha_i}) \quad (2.7)$$

where V_{intra} is a pairwise potential energy between elements α_i and β_i within the same cell i , V_{inter} is a pairwise energy between element α_i of cell i and element β_j of cell j , F_{external} is any external force.

In our model each cell contains 12 elements. Within each cell there is only one element type, and both inter- and intra-cellular forces are determined by Morse type potentials:

$$V = U_0 \exp(-\frac{r_{ij}}{\zeta_1}) - V_0 \exp(-\frac{r_{ij}}{\zeta_2}) \quad (2.8)$$

where r_{ij} is the distance between element i and element j . $U_0 = U_{\text{intra}}$, $V_0 = V_{\text{intra}}$, $\zeta_1 = \zeta_{1\text{intra}}$, and $\zeta_2 = \zeta_{2\text{intra}}$ for intra-cellular force, and $U_0 = U_{\text{inter}}$, $V_0 = V_{\text{inter}}$, $\zeta_1 = \zeta_{1\text{inter}}$, and $\zeta_2 = \zeta_{2\text{inter}}$ for inter-cellular force. Without external forces, the intra-cellular forces will arrange the inner elements so that cells will always have a roughly round shape. There is an equilibrium bound distance (r_e) such that when $r_{ij} < r_e$ or $r_{ij} > r_e$, the force generated by V is repulsive or adhesive, respectively, between element i and element j.

The adhesion between a basal epidermal cell and the basement membrane is modeled by the adhesion between the cell's elements and the basement membrane as:

$$F_{\text{external}}(r_{\alpha_i}) = \text{di}_{\alpha_i} \cdot \frac{Z}{h_{\alpha_i} + \epsilon} \tag{2.9}$$

where di_{α_i} is the direction from the element to the basement membrane and h_{α_i} is the distance between the element and the basement membrane. In the case where element r_{α_i} is located in the dermis, the element position will reset by reflecting along the basement membrane.

2.5.2 Cell division and lineage specification

We modeled three types of epidermal cells: basal, spinous, and granular keratinocytes. During wound healing, basal cells near the wound center experience a high self-renewal probability and division frequency. For basal keratinocytes (c_0), the division is determined by cell cycle T_0 and self-renewal probability pro_0 . During wound re-epithelialization, basal cells increase their self-renewal probability and frequency. For every division, each basal cell has a probability pro_0 to divide symmetrically into two basal cells, where the division line is perpendicular to the basement membrane, and a probability $1 - \text{pro}_0$ to divide asymmetrically into one basal cell and one spinous cell, where the division line is parallel to the basement

membrane.

Spinous keratinocytes are represented as two generations: proliferative spinous keratinocytes (c_1) and mature spinous keratinocytes (c_1). For the first generation (c_1), each cell has cell cycle $T_1 = (1 + \xi_{T1})T_{1c}$, and self-renewal probability pro_1 . In this case, orientation of the division line is stochastic. For the second generation of spinous keratinocytes (c_2), cell cycle is described by $T_2 = (1 + \xi_{T2})T_{2c}$ and each spinous keratinocyte will differentiate into one granular keratinocyte (c_3).

After every division, each daughter cell is described by 6 elements. New elements are gradually added to the cell using the following rule: one new element is added to a uniformly random position inside the convex hull of the cell's current elements every time $T_{g_i} = (1+\xi_{g_i})Tg_{ic}$,$i = 0, 1, 2, 3$ for cell type c_i . Cell growth stops when cell contains 12 elements. When a granular keratinocyte grows to have 12 elements, this cell is removed from the system in the following way: one element is randomly selected and deleted progressively $T_d = (1 + \xi_d)Td_c$ until only two elements remain. Once only two elements remain, the cell is removed by deleting all its remaining elements.

2.5.3 Selective cell adhesion

We assume that cell sorting is driven by differential adhesion between cells, and that adhesion between cells of the same type is stronger than adhesion between cells of different types. In this model, first and second generation of spinous keratinocytes are treated as different cell types. Cell sorting is modeled by replacing $F_{\text{inter}}(\alpha_i, \beta_j)$ with $c_{\text{sorting}}F_{\text{inter}}(\alpha_i, \beta_j)$ in the Morse potential function between cells of different types. Here, c_{sorting} represents cell sorting strength. We assume that $c_{\text{sorting}} > 1$ for strong adhesion (repulsion) and $c_{\text{sorting}} < 1$ for weak adhesion (repulsion). Cell sorting is assumed to be driven by cell-cell contact. Therefore, we require that cell sorting strength becomes 1, meaning that cells resume normal interactions,

if the distance between two cells is larger than the distance required for cell-cell contact r_{contact}. We set r_{contact} as two times the average cell diameter. This is represented by the equation:

$$c_{\text{sorting}} = \begin{cases} 1, \text{if } |r_{\alpha_i} - r_{\beta_j}| > r_{\text{contact}} \\ c_{\text{repulsion}}, \text{if } |r_{\alpha_i} - r_{\beta_j}| > r_{\text{contact}} \text{ and } \arg(F_{\text{inter}}(\alpha_i, \beta_j), r_{\alpha_i} - r_{\beta_j}) = 0 \\ c_{\text{adhesion}}, \text{if } |r_{\alpha_i} - r_{\beta_j}| > r_{\text{contact}} \text{ and } \arg(F_{\text{inter}}(\alpha_i, \beta_j), r_{\alpha_i} - r_{\beta_j}) = \pi \end{cases} \tag{2.10}$$

where $c_{\text{repulsion}} = c_{\text{repulsion·diff}} > 1$ and $c_{\text{adhesion}} = c_{\text{adhesion·diff}} < 1$ represent strong cell repulsion and weak cell adhesion if cells i and j are different cell types; while $c_{\text{repulsion}} = c_{\text{repulsion·same}} < 1$ and $c_{\text{adhesion}} = c_{\text{adhesion·same}} > 1$ represent strong cell repulsion and weak cell adhesion if cells i and j are the same cell type.

2.5.4 Coupling discrete and continuum models of signal production in the epidermis

To distribute the signal produced by one discrete cell onto grids of the PDE mesh, we modeled that a cell will update the signal field where its center is located. Increase in concentration at four corner points of the grid is determined by bilinear interpolation of the signal secretion by a cell. Signal produced by keratinocytes must be located in one grid of the PDE mesh (grid is of the form $[x_{k_x}, x_{k_x+1}] \times [y_{k_y}, y_{k_y+1}]$).Production of A and I by one cell is E_{pj0} for $j = 1, 2$. Production of A and I is then mapped to four-corner points of the PDE grid, and

is encoded by:

$$E_{pj}(x,y,\alpha_i,t) = \begin{cases} E_{pj0}[1-(x_{\alpha ic}-x_{k_x})][1-(y_{\alpha ic}-y_{k_y})], (x,y) = (x_{k_x}, y_{k_y}) \\ E_{pj0}[1-(x_{k_x+1}-x_{\alpha ic})][1-(y_{\alpha ic}-y_{k_y})], (x,y) = (x_{k_x+1}, y_{k_y}) \\ E_{pj0}[1-(x_{\alpha ic}-x_{k_x})][1-(y_{k_y+1}-y_{\alpha ic})], (x,y) = (x_{k_x}, y_{k_y+1}) \\ E_{pj0}[1-(x_{k_x+1}-x_{\alpha ic})][1-(y_{k_y+1}-y_{\alpha ic})], (x,y) = (x_{k_x+1}, y_{k_y+1}) \\ 0, \text{otherwise} \end{cases} \tag{2.11}$$

for $j = 1, 2$, where $(x_{\alpha ic}, y_{\alpha ic})$ are the coordinates of the center of cell α_i . The total production is therefore encoded by:

$$E_{pj}(x,y,t) = \sum_{\alpha_i} E_{pj}(x,y,\alpha_i,t) \tag{2.12}$$

2.5.5 Description of equations (2.1)-(2.4)

In equations (2.1-2.2) (see Results in Chapter 2), first term on the right side of Eqn (2.1) represents random movement (diffusion) and active movement (chemotaxis) based on the Keller-Segel model [37]; second term represents self-renewal of F, promoted by p_2; and third term represents decay of F. First term on the right side of Eqn (2.2) represents diffusion of ECM; second term represents production of ECM by fibroblasts, which is inhibited by p_1 and activated by p_2, respectively. Third term represents decay of ECM. In equation (2.3) and (2.4), first term on the right side represents diffusion; second term represents production of p_1 and p_2 from fibroblasts and their promotion by ECM (see following section for details); third term represents production from keratinocytes; fourth term represents production by fibrin clot during wound healing; and last term represents decay of p_1 or p_2. In equation (2.4), fifth term represents production of p_2 by immune cells.

2.5.6 Activator production by immune cells

Immune cells are modeled as a source of activator (Figure 2.11). In the model, immune cells produce activator only when the surrounding activator level is high, which mimics activation of immune cells during inflammation phase of wound healing and lack (or low) activation in normal, unwounded skin. Density of immune cells (IM) is assumed to be constant in the dermis. Production rate of activator is $f_{IM}(IM, p_2) = IM \cdot \frac{(p_2 - p_{\lim})^{n_{\text{im}}}}{1+(p_2-p_{\lim})^{n_{\text{im}}}}$, when $p_2 > p_{\lim}$ ($p_{\lim}$ is modeled as the threshold) and $f_{IM}(IM, p_2) = 0$, when $p_2 < p_{\lim}$.

2.5.7 Production function in equaitons (2.1)-(2.3)

$$f_F(F, p_2) = \frac{c_F F}{1 + (\frac{F}{gm_F})^{n_F}} \cdot [b_F + \frac{a_F p_2^{n_{p2F}}}{1 + (\frac{p_2}{gm_{p2F}})^{n_p 2F}}] \tag{2.13}$$

$$f_C(F, p_1, p_2) = c_C F \cdot [b_{cp1} - \frac{a_{cp1} p_1^{n_{p1}}}{1 + (\frac{p_1}{gm_{p1C}})^{n_p 1C}}] \cdot [b_{cp2} + \frac{a_{cp2} p_2^{n_{p2}}}{1 + (\frac{p_2}{gm_{p2C}})^{n_p 2C}}] \tag{2.14}$$

$$f_{p1}(F, C, p_1) = c_{Fp1} F \cdot [b_{p1} + \frac{a_{p1} C^{n_{p1}}}{1 + (\frac{C}{gm_{p1}})^{n_p 1}}] \tag{2.15}$$

$$f_{p2}(F, C, p_2) = c_{Fp2} F / [b_{p2} + \frac{a_{p2} C^{n_{p2}}}{1 + (\frac{C}{gm_{p2}})^{n_p 2}}] \tag{2.16}$$

2.5.8 Fibrin clot description

Fibrin clot (M), which is only present during wound healing, is modeled by:

$$\frac{\partial M}{\partial t} = D_M \Delta M - d_M M \tag{2.17}$$

with initial condition:

$$M(x, y, 0) = \begin{cases} M_0, \text{if } \Phi(x, y) < 0 \text{ and } \Phi(x, y, 0) > 0 \\ \\ 0, \text{otherwise} \end{cases} \tag{2.18}$$

where Φ_0 is the phase function in stable, normal skin, and $\Phi(x, y, 0)$ represents the initial phase function of wounded skin. Equ (A.12) defines M .

2.5.9 Boundary condition for signal and dermal components

Combination of regions D and E forms a rectangular region described by $[0, L_x] \times [0, L_y]$. For $p1, p2, M$, whose domain is $[0, L_x] \times [0, L_y]$, the boundary condition in x-axis (when $x = 0$ and $x = L_x$) is periodic and no-flux in y-axis (when $y = 0$ and $y = L_y$).

For F, C, CB, whose domain is region D, the boundary condition in x-axis (when $x = 0 and x = L_x$) is periodic, and no-flux in y-axis (when $y = 0$ and along Ω).

2.5.10 Initial condition in normal skin

Normal skin is setup with the following conditions: (1) a flat epidermis-dermis interface where $\Phi(x, y, 0) = y - L_D$ is the initial phase function, with L_D denoting the height of the dermis; (2) F, C are constants F_{in}, C_{in} in region D, while p_1, p_2 are constants $p_{1\text{in}}$, $p_{2\text{in}}$ in region Ω; and (3) a stratified epidermis with basal, spinous, and granular cell layers.

2.5.11 Model of initial epidermal wound

Epidermal wound is modeled by damaging only epidermis and keeping dermis restricted to the following criteria: (1) interface between regions D and E is the same as in unwounded skin simulation; (2) ECM density and fibroblast density are the same as in unwounded skin simulation; (3) all keratinocytes are the same as in unwounded skin simulation except that cells are removed by deleting all cellular elements if the distance between x-direction of cell center and wound center is less than d_{wound} . x-axis of the wound center is defined to be the

$\frac{L_x}{2}$; (4) concentrations of activator and inhibitor are set to half of the concentration as in unwounded skin simulation in the region $[\frac{L_x}{2} - d_{\text{wound}}, \frac{L_x}{2} + d_{\text{wound}}] \times [0, L_y] \cap E$, and equal to unwounded skin simulation in other regions.

2.5.12 Model of initial full-thickness wound

In the full-thickness wound model, wounding occurs both in the epidermis and dermis. It is modeled by the following conditions: (1) epidermal-dermal interface is modeled by a phase function described by: $\Phi(x, y, 0) = y - L_D + \alpha_d \exp[1 - (\frac{x - 0.5L_x}{d_{\text{wound}}})^2]$ where L_D is stable height of dermis in simulations of unwounded skin; (2) ECM and fibroblast densities are equal to these in unwounded skin, when $\Phi(x, y, 0) < 0$ (in region D); (3) activator and inhibitor signals in keratinocytes are initialized to be the same as in epidermal wounds (see above); (4) fibrin clot is initialized as described in Chapter 2.

The end time of our model simulations is when re-epithelialization ends. To ensure that the re-epithelialization simulation results are stable, we usually continue simulations for another 1.5 cell cycles of the basal cells.

2.5.13 Model of heterogeneity in basal epidermal cells during wound healing

Modeling heterogeneity with respect to distance

Center of the basal cell nearest to the wound defines the position of the leading epithelial tongue. We used the distance (d) between cell center and leading tongue to calculate cell's self-renewal probability and frequency, A and I production rates, and migration speed. When $d > d_0$, cell acts as a cell in normal skin. When $d < d_0$, heterogeneity (2.1)-(2.3) is modeled

with respect to d. d_0 is set to be $10d_{\text{cell}}$, where d_{cell} represents average diameter of a cell.

Increased basal cell self-renewal probability and frequency

During wound re-epithelialization, basal cells increase their self-renewal probability (pro_0) and cell cycle (T_0), described by

$$T_0 = \begin{cases} (1+\xi_T)T_{0c} \cdot (1 - g\dfrac{d_0 - d}{d_0}), \text{during re-epithelialization and } d < d_0 \\ (1+\xi_T)T_{0c}, \text{otherwise} \end{cases} \tag{2.19}$$

$$\text{pro}_0 = \begin{cases} 1 - (1 - \text{pro}_{0c}) \cdot \dfrac{d}{d_0}, \text{during re-epithelialization and } d < d_0 \\ \text{pro}_{0c}, \text{otherwise} \end{cases} \tag{2.20}$$

where d is the distance between cell and leading tongue during wound re-epithelialization, and ξ_T is white noise term.

Active movement of basal cells

Velocity ($v0$) of active movement of basal cells toward the wound center and along the basement membrane (interface between regions D and E), is described by:

$$v_0 = \begin{cases} (1+\xi_v)v_{0c} \cdot (1 - \dfrac{d_0 - d}{d_0}), \text{during re-epithelialization and } d < d_0 \\ 0, \text{otherwise} \end{cases} \tag{2.21}$$

Increased production of signals in basal cells

During wound re-epithelialization, a cell's production of A and I depends on its distance (d) from the leading epithelial tongue. This is described by:

$$E_{pj} = \begin{cases} E_{pjmin} + (E_{pjmax} - E_{pjmin}) \cdot (\frac{d_0 - d}{d_0})^4, \text{during re-epithelialization and } d < d_0 \\ E_{pjmin}, \text{otherwise} \end{cases} \tag{2.22}$$

where $E_{pjmin}(j = 1, 2)$ is the production rate of $I(j = 1)$ and $A(j = 2)$ by keratinocytes in normal skin.

2.5.14 Scalability of the model

As a proof of principle of the model's scalability, we simulated deep and wide wounds with dense fibrin clot with a version of the model that contains two types of activators: A_1 ($p_{2\text{ECM}}$) that stimulates ECM production only and A_2 (p_{2F}) that stimulates fibroblast proliferation only (Figure 2.15). In this version of the model, the function of activator on fibroblast proliferation $f_F(F, p_2)$ is replaced with $f_F(F, p_{2F})$, whereas the function of activator on ECM production $f_C(F, p1, p2)$ is replace with $f_C(F, p1, p_{2\text{ECM}})$.

2.6 Supplementary tables and figures

Parameter name	Parameter value	Parameter name	Parameter value
One simulated day	6000	ξ_{g3}	$N(0, 10)$
d_{cell}	0.15	Td_c	100
pro_{0c}	0	ξ_d	100
pro_1	0.35	r_{contact}	0.3
g	0.5	$C_{\text{diffusion·diff}}$	1.6
T_{0c}	6000	$C_{\text{adhesion·diff}}$	0.4
ξ_{T0}	$N(0, 800)$	$C_{\text{repulsion·same}}$	1
T_{1c}	1000	$C_{\text{adhesion·same}}$	1
ξ_{T1}	$N(0, 150)$	U_{intra}	0.099
T_{2c}	1000	V_{intra}	0.1782
ξ_{T2}	$N(0, 150)$	$\zeta_{1\text{intra}}$	0.12
T_{g0c}	100	$\zeta_{2\text{intra}}$	0.36
ξ_{g0}	$N(0, 10)$	U_{inter}	0.0505
T_{g1c}	100	V_{inter}	0.04173
ξ_{g1}	$N(0, 10)$	Z	0.009
T_{g2c}	100	ϵ	0.01
ξ_{g2}	$N(0, 10)$	v_{0c}	13
T_{g3c}	100	ξ_v	$N(0, 0.1)$

Table 2.1: Parameter values for epidermal discrete model. Parameters are included for subcellular element method (SEM), cell cycle, cell self-renewal probability, cell growth rate, cell decay rate, selective adhesion, and cell migration.

Parameter name	Parameter value	Parameter name	Parameter value
L_x	7	n_{p2}	2
L_y	14	b_{cp1}	1
L_D	7	a_{cp1}	1
E_{p1max}	36	gm_{p1c}	0.0625
E_{p1min}	1	b_{cp2}	1
E_{p2max}	31.5	a_{cp2}	1
E_{p2min}	1	gm_{p2c}	0.0625
F_{in}	1	c_{Fp1}	0.1254
C_{in}	1	c_{Fp2}	0.1188
p_{1in}	0.03856	b_{p1}	1
p_{2in}	0.02407	a_{p1}	1
v_F	0.35	gm_{p1}	3.66
c_F	4.14	gm_{p2}	3.66
D_F	1.225	D_M	0.3537
gm_F	$N(0, 10)$	d_M	0.1
b_F	0.45	M_0	3.3
a_F	4	s_{chem}	20
gm_{p2F}	0.0625	s_{diff}	1
d_F	1	s_c	1
n_{p2F}	2	IM	0.5
D_C	0.35	n_{im}	2
c_c	7.3	p_{lim}	0.02777
n_{p1}	2		

Table 2.2: Parameter values for dermal continuum model. Parameters for fibroblasts, ECM, immune cells, activator, inhibitor and fibrin clot density are included.

Wound width	Wide	Narrow	
d_{wound}	1.4	0.7	
Wound depth	Deep	shallow	Very shallow
α_d	1.75	0.875	0.4375
Initial fibrin clot density	High	Low	
M_0	3.3	0.1	

Table 2.3: Parameter values for wound size and fibrin clot density.

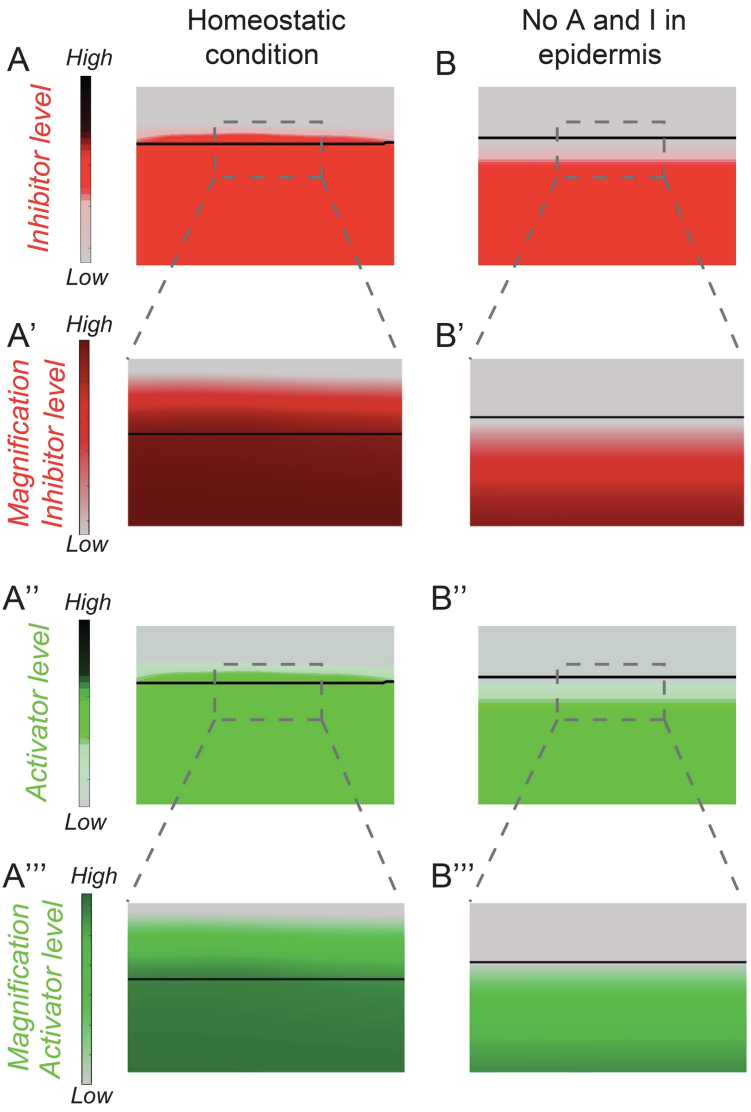

Figure 2.5: Magnified view of inhibitor and activator levels.
(A, A") Overall inhibitor and activator levels under homeostatic condition. (A', A"') Magnified view of inhibitor and activator levels near the basement membrane (black line). (B, B") Overall inhibitor and activator levels under the condition with no A/I in epidermis. (B', B"') Magnified view of inhibitor and activator levels near the basement membrane (black line).

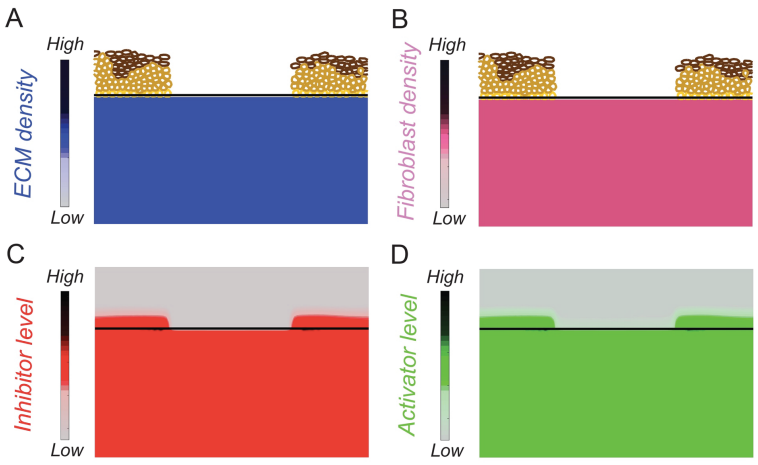

Figure 2.6: Initial conditions for epidermal abrasion wound model.
(A) Initial ECM density in the dermis. (B) Initial fibroblast density in the dermis. (C) Initial inhibitor level. (D) Initial activator level. Black lines mark position of the basement membrane.

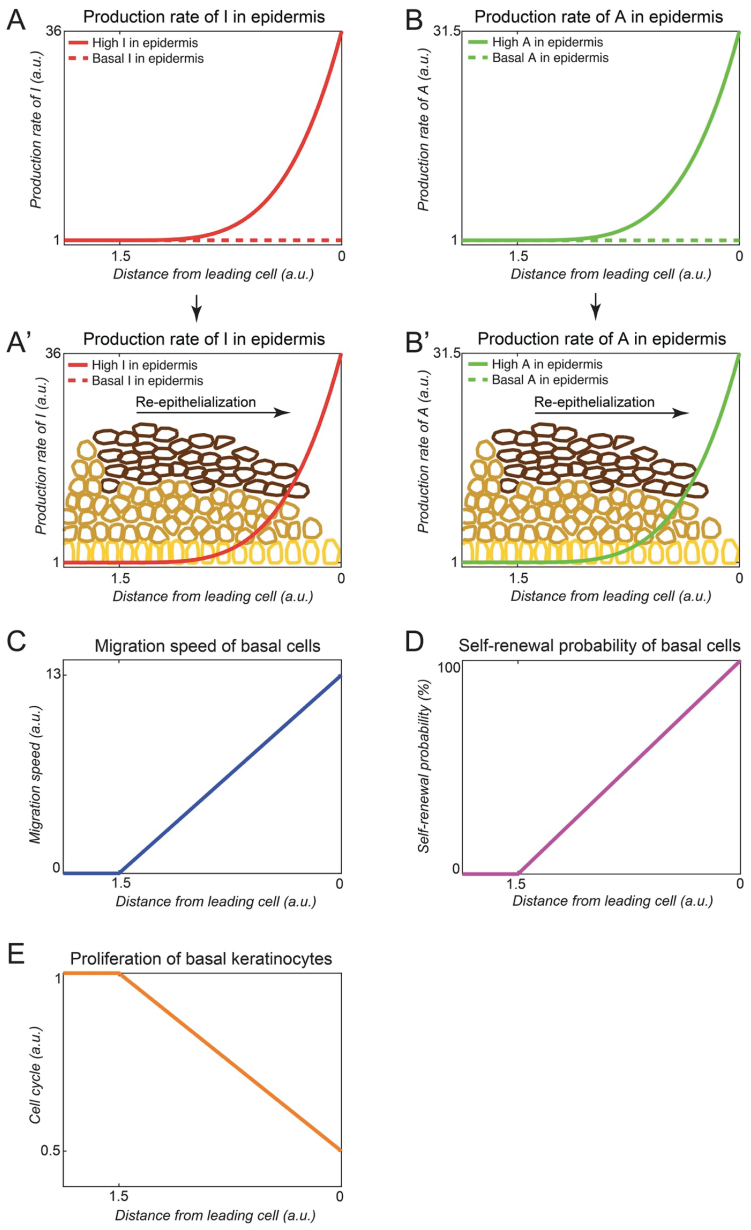

Figure 2.7: Modeling assumptions in basal epidermal cells after wounding. (A, B) High (solid line) and basal production rates (dashed line) of I and A. x-axis represents distance from the leading cell in arbitrary units (a.u.). y-axis represents production rate of I and A (in a.u.). (A', B') Overlay of I and A production rates onto modeled wound epidermis. (C) Migration speed of basal cells. x-axis represents distance from the leading cell (in a.u.), y-axis – migration speed (in a.u.). (D) Self-renewal probability of basal cells. x-axis represents distance from the leading cell (in a.u.), y-axis – self-renewal probability (in a.u.). (E) Proliferation of basal cells. x-axis represents distance from the leading cell (in a.u.), y-axis – cell cycle state (in a.u.).

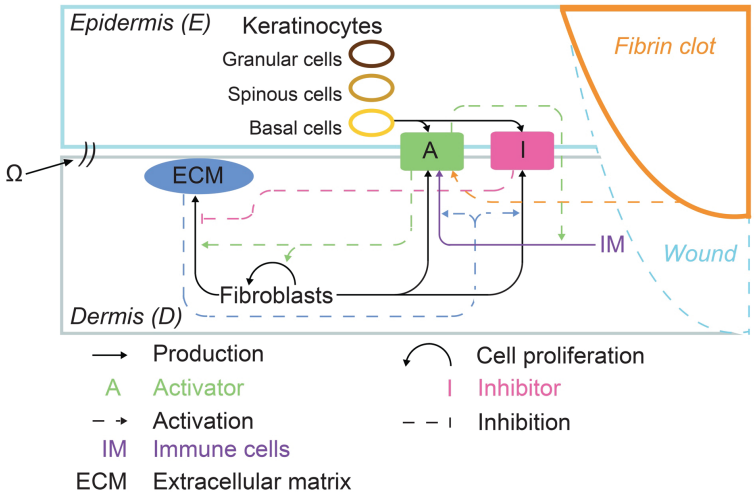

Figure 2.8: Schematic of two-dimensional multiscale hybrid model of wound healing regulated by fibrin clot.
The modeling domain is separated into epidermis (E) (light blue) and dermis (D) (gray) by a dynamic interface (Ω) to mimic the basement membrane. In E, keratinocytes are modeled discretely as basal (yellow), spinous (light brown) and granular (dark brown) cells. Basal cells are set to produce activator A (shaded green box) and inhibitor I (shaded dark pink box). In D, fibroblasts, ECM (shaded blue oval), immune cells (IM) and fibrin clot (orange) are modeled in a continuum, where fibroblasts produce A and I, and these processes are directly activated by ECM. Additionally, fibrin clot and immune cells (IM) produce A. Fibroblast proliferation is activated by A, and ECM production is activated by A and inhibited by I. A and I can also diffuse across Ω.

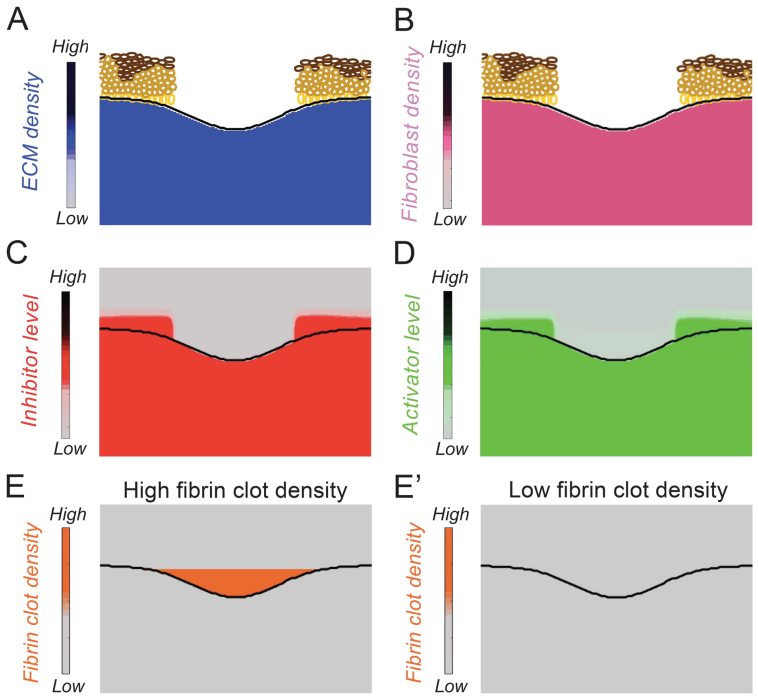

Figure 2.9: Initial conditions for dermal wound model.
(A) Initial ECM density in the dermis. (B) Initial fibroblast density in the dermis. (C) Initial inhibitor level. (D) Initial activator level. (E) High initial fibrin clot density condition. (E') Low initial fibrin clot density condition. Black lines mark position of the basement membrane.

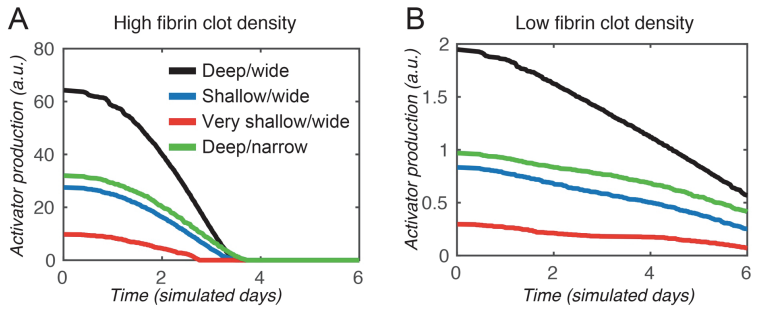

Figure 2.10: Activator production dynamics from fibrin clot.
(A) Activator production from high density fibrin clot. (B) Activator production from low density fibrin clot. x-axis represents time (simulated days).y-axis represents activator production from fibrin clot in arbitrary units (a.u.). Different colors correspond to distinct wound dimensions.

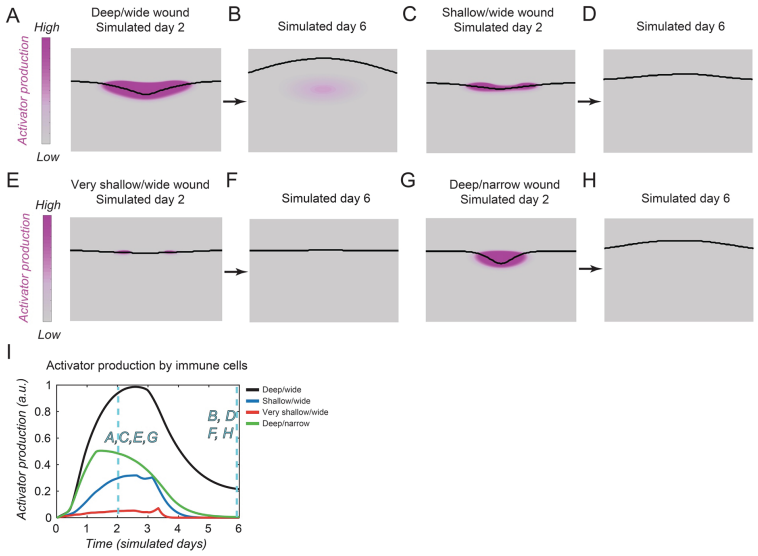

Figure 2.11: Activator production dynamics by immune cells.
(A, C, E, G) Activator production levels by immune cells in deep/wide, shallow/wide, very shallow/wide and deep/narrow wounds on simulated day 2, respectively. (B, D, F, H) Corresponding activator production levels by immune cells on day 6. (I) Temporal changes in activator production by immune cells. Color definitions for different wound geometries are as follows: deep/wide (black), shallow/wide (blue), very shallow/wide (red) and deep/narrow (green). x-axis represents time (simulated days). y-axis represents activator production in arbitrary units (a.u.).

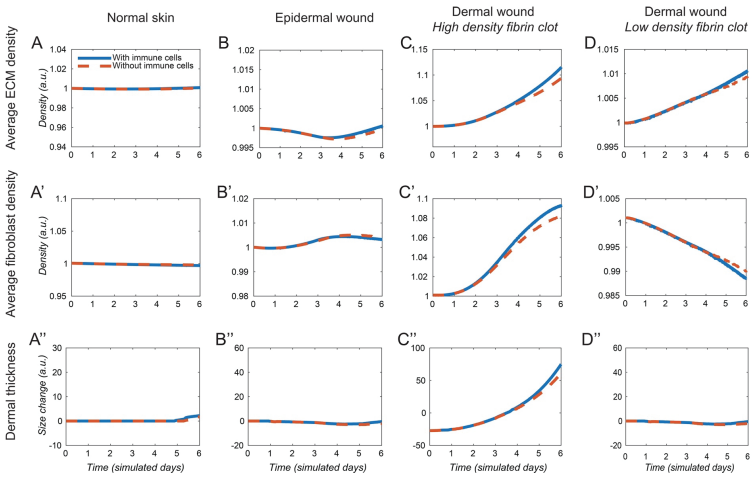

Figure 2.12: Comparison of wound healing outcomes in the model with and without immune cells.
(A-A") Temporal dynamics in average ECM density (A), fibroblast density (A') and dermal thickness (A") under homeostatic conditions between the model with immune cells (solid blue line) and without immune cells (dashed orange line) (see main Figure 2.2A-A"'). No significant differences are observed. (B-B") Temporal dynamics in average ECM density (B), fibroblast density (B') and dermal thickness (B") in epidermal abrasion wounds (see main Figure 2.3A-B"'). No significant differences are observed. (C- C") Temporal dynamics in average ECM density (C), fibroblast density (C') and dermal thickness (C") in dermal wounds with high density fibrin clot (see main Figure 2.4A-B"'). All three parameters are increased in the model with immune cells. (D- D") Temporal dynamics in average ECM density (D), fibroblast density (D') and dermal thickness (D") in dermal wounds with low density fibrin clot. ECM density moderately increased and fibroblast density moderately decreased in the model with immune cells. Values along x-axis are in simulated days and values along y-axis are in arbitrary units (a.u.).

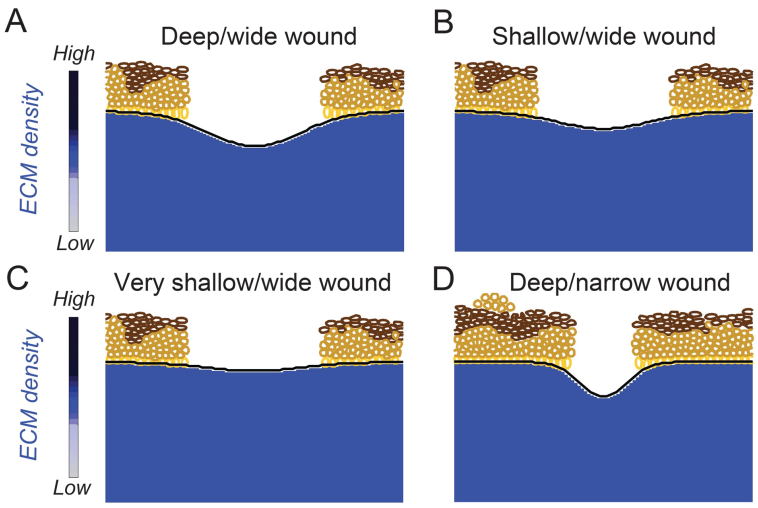

Figure 2.13: Initial assumptions about dermal wound geometry.
(A) Deep and wide wound. (B) Shallow and wide wound. (C) Very shallow and wide wound. (D) Deep and narrow wound. Initial ECM density in the dermis is shown (blue). Black lines mark position of the basement membrane.

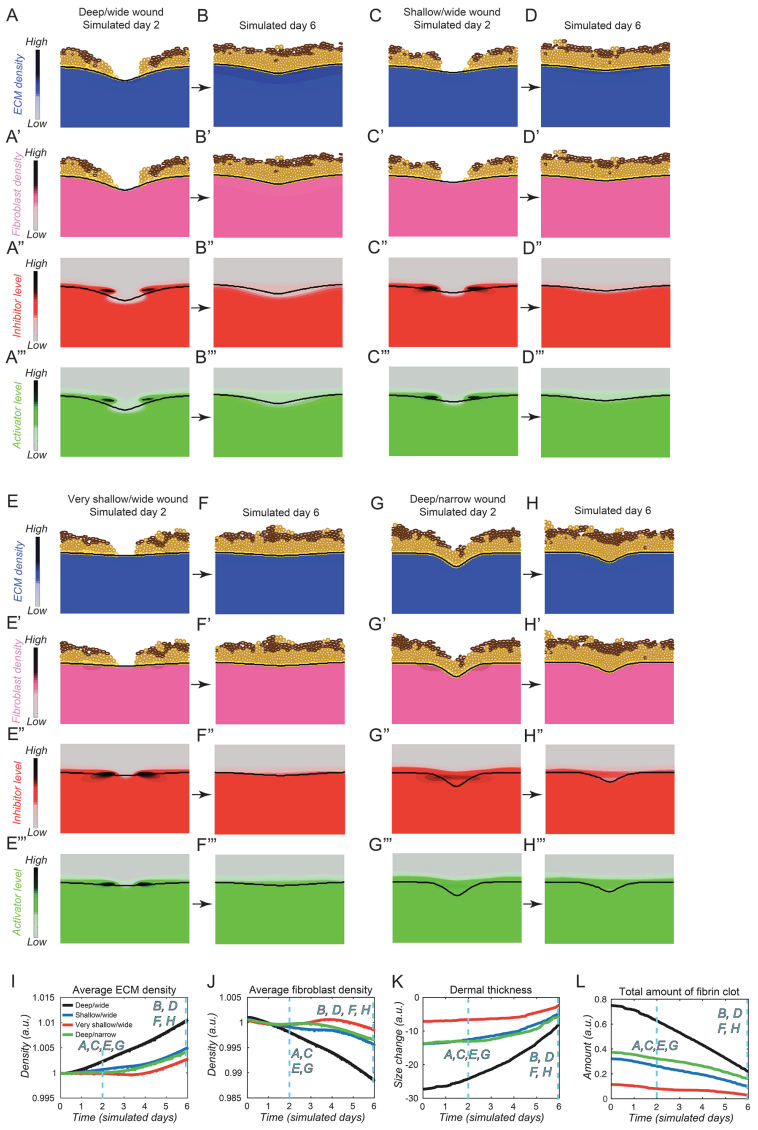

Figure 2.14: Healing outcomes of dermal wounds depend on wound geometry and fibrin clot density.

All dermal wound healing simulations shown on this figure were performed using the low-density fibrin clot condition and included immune cells. The following wound geometries were compared: (A-B"') deep and wide wounds, (C-D"') shallow and wide wounds, (E-F"') very shallow and wide wounds, and (G-H"') narrow and deep wounds. For each simulated condition, ECM density (blue), fibroblast density (pink), inhibitor level (red) and activator level (green) are shown. For each simulation, two consecutive snapshots are shown: day 2 and day 6. Black lines mark position of the basement membrane. (I) Temporal changes in average ECM density across modeling time. (J) Temporal changes in average fibroblast density across modeling time. (K) Temporal changes in dermal scar thickness across modeling time. (L) Degradation dynamics of the low-density fibrin clot across modeling time. Color definitions for each line on I-L are provided on the figure. Values along x-axis are in simulated days and values along y-axis are in arbitrary units (a.u.).

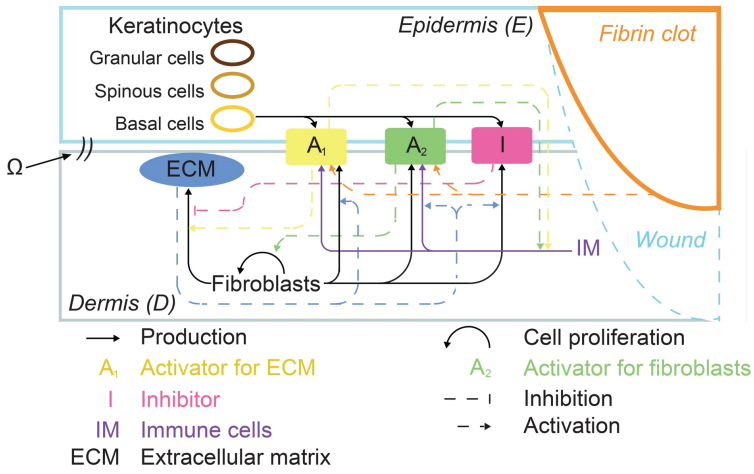

Figure 2.15: Schematic of two-dimensional multiscale hybrid model of wound healing with two activators.
The modeling domain is separated into epidermis (E) (light blue) and dermis (D) (gray) by a dynamic interface (Ω) to mimic the basement membrane. In E, keratinocytes are modeled discretely as basal (yellow), spinous (light brown) and granular (dark brown) cells. Basal cells are set to produce activators A_1 (shaded yellow box) and A_2 (shaded green box), and inhibitor I (shaded dark pink box). In D, fibroblasts, ECM (shaded blue oval), immune cells (IM) and fibrin clot (orange) are modeled in a continuum, where fibroblasts produce A_1, A_2 and I, and these processes are directly activated by ECM. Additionally, fibrin clot and immune cells produce A_1 and A_2. Fibroblast proliferation is activated by A_2, and ECM production is activated by A_1 and inhibited by I. A_1, A_2 and I can also diffuse across Ω.

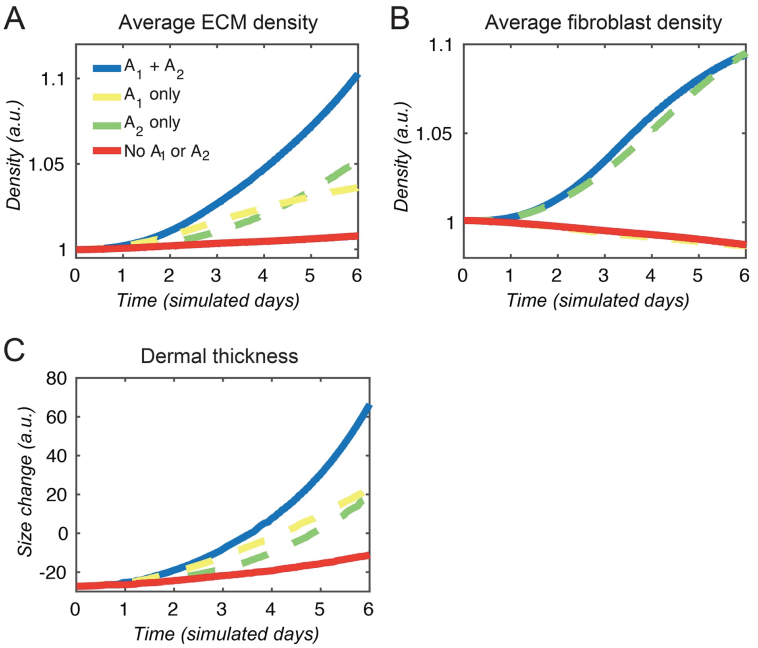

Figure 2.16: Comparison of wound healing outcomes between models with different activators.
The following four conditions were modeled: (i) Fibrin clot produces both A_1 (activator for ECM production) and A_2 (activator for fibroblast proliferation) (blue solid line); (ii) Fibrin clot produces A_1 only (yellow dashed line); (iii) Fibrin clot produces A_2 only (green solid line); (iv) Fibrin clot does not produce activators (red solid line). (A) Difference in average ECM density. (B) Difference in average fibroblast density. (C) Difference in dermal thickness. Values along x-axis are in simulated days and values along y-axis are in arbitrary units (a.u.).

Chapter 3

A multiscale model via single-cell transcriptomics reveals robust patterning mechanisms during early mammalian embryo development

During early mammalian embryo development, a small number of cells make robust fate decisions at particular spatial locations in a tight time window to form inner cell mass (ICM), and later epiblast (Epi) and primitive endoderm (PE). While recent single-cell transcriptomics data allows scrutinization of heterogeneity of individual cells, consistent spatial and temporal mechanisms the early embryo utilize to robustly form the Epi/PE layers from ICM remain elusive. Here we build a multiscale three-dimensional model for mammalian embryo to recapitulate the observed patterning process from zygote to late blastocyst. By integrating the spatiotemporal information reconstructed from multiple single-cell transcriptomic datasets, the data-informed modeling analysis suggests two major processes critical to the formation of Epi/PE layers: a selective cell-cell adhesion mechanism (via EphA4/EphrinB2)

for fate-location coordination and a temporal attenuation mechanism of cell signaling (via Fgf). Spatial imaging data and distinct subsets of single-cell gene expression data are then used to validate the predictions. Together, our study provides a multiscale framework that incorporates single-cell gene expression datasets to analyze gene regulations, cell-cell communications, and physical interactions among cells in complex geometries at single-cell resolution, with direct application to late-stage development of embryogenesis.

3.1 introduction

In mammals, the first two developmental events that occur are 1) the formation of the trophectoderm (TE) and inner cell mass (ICM) followed by 2) specification of the ICM into the primitive endoderm (PE) and epiblast (Epi). While both of these processes lead to the specification of primitive epithelial-like structures (the TE and PE) that wrap the future embryo (the Epi), the process that gives rise to the PE and TE are markedly different. While both are highly regulated processes, formation of the PE is both highly dynamic and stochastic by comparison. This raises the question, how can such a dynamic and stochastic process proceed robustly and reproducibly.

These first two developmental events lead to the formation of early multi-cellular structures that differ in both their gene expression and their location within the embryo. In the TE/ICM case, a monolayer shell of Cdx2 expressing TE cells surrounds an inner core of Oct4 expressing cells. In the Epi/PE case, an aggregate of Nanog expressing cells [64, 26] forms the Epi, which is surrounded by PE, a monolayer of Gata6 expressing cells [93] that separates the Epi from embryonic cavity (blastocoel). These specification processes have a number of similarities. A tristable gene regulation circuit controls differentiation from an uncommitted state to one of two differentiated states in both cases [11, 10, 15, 40]. Both also yield similar physical structures, an aggregate of cells surrounded by a monolayer.

Formation of the Epi and PE is however a distinctly more stochastic and dynamic process than TE/ICM formation. During formation of the TE/ICM, cells appear to choose their fate based on positional information (exterior cells become TE and interior cells become ICM). That is, cells differentiate in a mostly deterministic fashion and the TE and ICM structures are essentially constructed as a result of differentiation itself. While cells have been observed to move within the embryo, as few as 5% are exchanged between the TE and ICM [5]. Epi and PE cells on the other hand asynchronously (in time) differentiate to initially form a stochastically organized salt and pepper spatial distribution [48, 91, 8, 81] that later evolves into the canonical Epi and PE structures through cellular motions.

Epi and PE differentiation is regulated by two mutually antagonistic factors, Nanog and Gata6 [64, 26, 93]. Prior to the 32-cell stage (E3), these factors are co-expressed in almost all cells. By E3.5-E4, they are mutually exclusively expressed [8] in a salt and pepper distribution of Epi and PE cells. The proposed cause of this salt and pepper distribution is Fgf signaling [102, 114, 50], which is secreted by differentiated Epi cells and promotes expression of PE markers in neighbor cells. Interestingly, Epi/PE specification is not a bang-bang process at the population level. Instead, cells asynchronously differentiate at different times. While this could be viewed as a simple result of stochasticity, Saiz et al. [92] proposed the incremental commitment in conjunction with Fgf signaling is functionally important for controlling the proportions of PE and Epi cells. Numerical simulations verified this mechanism could robustly produce a salt and pepper distribution with proper cell proportions [102].

This still leaves the question of how these cells organize into canonical Epi and PE structures. Intercalation of cells into the PE layer due to blastocoel expansion contributes to PE formation [81]. Differential adhesion mediated sorting is also thought to play a crucial role [49, 70]. While this idea sounds enticing however, adhesion factors that would facilitate this sorting have, to our knowledge, not been previously identified in the morula or

early blastocyst stage mammalian embryos. Further, the presence of double positive (DP) cells expressing both Nanog and Gata6 has not previously been considered. Though these mechanisms have been investigated in isolated stages, a coherent understanding of how they corporate through multiple stages is still lacking.

A number of phenomenological modeling studies have been performed to study this developmental time frame. Non-spatial studies have helped identify the minimal gene interaction networks that regulate differentiation [11, 10, 15, 4]. Spatial models have been used to study how mechanical factors such as cell-cell interactions [41] or cell contractility [59] influence development. Others have more comprehensively spatially modeled physical and regulatory processes [40, 102, 49, 70]. Each of these studies have however been phenomenological in that they have largely integrated and been compared to imaging data and have not utilized the type of single cell RNA sequencing (scRNA-seq) data that has become available in recent years. Also, to the best of our knowledge, there have not been any three-dimensional models that comprehensively couple regulatory processes, spatial cell soring during Epi/PE separation, and single cell data to study Epi/PE formation.

Inspired by the promising results of data-centric approaches, several data-informed models have been introduced, for example, temporal models based on temporal RNA-sequencing data [43, 7], temporal models using time series of neural activity data [97], and spatiotemporal models calibrated with morphologic data [44]. These data-driven models depend on group average data without individual cell resolution. Recently the single-cell gene expression profiles become available for early mammalian embryo [56]. This opens up the opportunity of utilizing data in modeling at a resolution of individual cells [57, 58]. A single-cell qPCR dataset quantified 48 selected genes in mouse embryo from 1-cell stage to 64-cell stage [33]. Several recent scRNA-seq datasets on mouse early embryo provide an unbiased gene edxpression profiles of transcriptomics across different developmental stages, including E3.5-E6.5 [65], E5.25-E6.5 [9], and E6.5-E8.5 [79]. These scRNA-seq datasets allow us to explore the

heterogeneity among individual cells considering the expression levels of all genes. However, the spatial information is lost in scRNA-seq data, hindering the examination of communications among cells, which are crucial in cell fate decision. Further, currently no single dataset covering the time course from zygote to late blastocyst, and different datasets might be obtained using different techqniues.

Here we develop a data-informed, three-dimensional multiscale model of development of the mammalian embryo from the 1 to 128-cell stage to study how the dynamic interaction between the differentiation and sorting processes influences PE and Epi organization. This model couples 1) a model of gene regulation, 2) a family of models of adhesion mediated cell-cell interactions based directly from expression data for Eph/Ephrin pairs from single cell data, and 3) a 3D physical model of the embryo at a subcellular resolution. We use this model to study two essential questions. First, can observed Eph/Ephrin ligand receptor pairings found in the single cell data provide the appropriate adhesion conditions ensure formation of the PE and Epi. Second, what conditions must be met for Fgf signaling to effectively control allocation of cells to the PE and Epi fates.

Results demonstrate that while there are two candidate ligand receptor pairs observed in data, EphA4/EphrinB2 pairing is more likely to lead to adhesion differences that will facilitate sorting. This provides a specific hypothesis that could be tested through genetic manipulation of these candidate molecules. We further find that while cell allocation and organization appears to be relatively insensitive to the timing of Fgf signaling onset, continued signaling into the late 128-cell stage actually impairs proper organization. Further analysis of the scRNA-seq data confirms this hypothesis.

3.2 Results

3.2.1 A multiscale three-dimensional model from fertilization (1 cell) to late blastocyst (128 cells) stage

We constructed a 3D, multiscale spatial-temporal model for the development of the multicellular blastocyst from 1-cell to 128-cell stage (Fig 3.1a). The model couples two developmental processes that are critical to this early phase of development. 1) Regulation of cell fate specifying genes (Oct4/Cdx2 for TE/ICM and Nanog/Gata6 for Epi/PE) is modeled using ordinary differential equations (ODE) (Fig 3.1b,c). 2) Physical cell-cell interactions (including selective adhesion, Fig 3.1d) with subsequent cell migrations are modeled using the subcellular element method. This multiscale model (Fig 3.1e) recapitulates phenomelogically correct developmental process of the embryo from zygote to blastocyst (Fig 3.1f). We first briefly describe the gene regulatory model used and then subsequently describe the 3D modeling framework used to model the whole early embryo.

Gene regulatory dynamics associated with the TE/ICM and Epi/PE formations are modeled separately. Mutually antagonistic and self-activation dynamics of Oct4/Cdx2 modulated by cell contact (similar to [40]) are used to model TE/ICM formation (See Section 3.5 for detailed gene network equations). Similar mutually antagonistic dynamics between Nanog/Gata6 describe Epi/PE formation. In this case, cell-cell communication occurs via Fgf4/Fgfr2 regulation of the Erk signaling pathway (for simplicity, we refer to this as Fgf4/Fgfr2 or just Fgf signaling). The detailed equations are listed in Eq (3.1)-(3.6).

To integrate these regulatory dynamics with the mechanical and morphological aspects of embryo development, we developed a 3D spatial model where the embryo is modeled by a collection of discretely represented cells constrained in a spherical geometry (inspired by [40, 49]). The model was implemented in the framework of subcellular element method

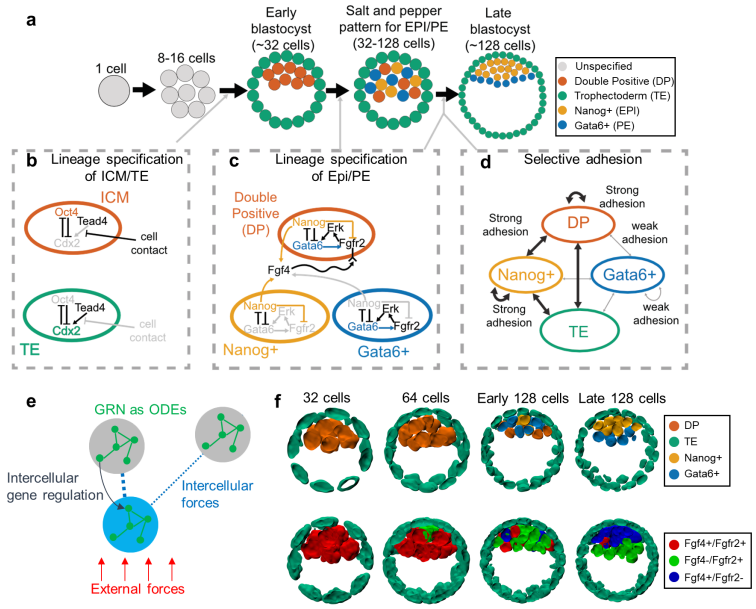

Figure 3.1: A multiscale model for early embryo development from fertilization to late blastocyst stage.
(a) Embryo shapes in different stages. The circle color indicates the cell type: Yellow for unspecified cell; Green for Trophectoderm (TE) cell; In Inner Cell Mass: Purple for double positive of Nanog and Gata6 (DP); Red for Nanog high and Gata6 low (Nanog+); Blue for Gata6 high and Nanog low (Gata6+). (b) Gene regulation models for TE/ICM specification before early blastocyst stage. The grey color represents weak cell contact or weak gene expression. (c) Gene regulation models for Nanog+/Gata6+ specification during early to late blastocyst stage. (d) Modeled selective adhesion between different cell types through early to late blastocyst stage. (e) A schematic illustration of the multiscale model containing spatial and gene expression dynamics of the cells. The correspondence to equations are as follows: GRNs, Eqs (3.1)-(3.5); intercellular forces, Eq (3.7); intercellular gene regulation, Eqs (3.5), (3.6); external forces, Section 3.5 Eqs (3.12), (3.13). (f) Simulated embryo with Nanog/Gata6 and Fgf4/Fgfr2 expression at different stages.

(SEM) [69] which represents a cell by a collection of elements (particles in 3D space). The spatial dynamics of the cells depicted by these elements partially depend on the modeled gene expressions and, in turn, provide a spatial reference for modeling intercellular gene regulations [40, 107, 68]. The following aspects are accounted for: cell-cell interactions; selective adhesion; cell division; confinement of cells by the zona pellucida; and cavity formation (see Eq (3.7) and Section 3.5 for details).

The cell divisions are modeled by splitting the elements of a mother cell into two subsets representing the two daughter cells. The cell divisions are scheduled as follows: 1) from 1- to 32-cell stage, all cells divide at the same time, 2) from 32- to 64-cell stage and from 64- to 128-cell stage, the cell cycle for each cell is modeled by a random variable uniformly distributed over a time window.

Among the aspects modeled in the spatial model, we are especially interested in the consequences of selective adhesion to evaluate how heterogeneous cellular adhesion mechanism impacts the pattern formation in early embryo development. We represent the selective adhesion mechanisms by assigning adhesion scores (AS) for different cell type pairs. A high AS means a strong adhesion and a low AS means a weak adhesion. The AS is implemented as the parameter α in Eq (3.7). To quantify AS and model selective adhesion, we will take two approaches. First, we will use single-cell RNA sequencing data [65] to quantify expression levels of adhesion related molecules (Eqs (3.8), (3.9)) and assign AS based on data. Second, we will explore the effectiveness of different phenomenological models of adhesion (encoded in the AS) to determine how different types of selective adhesion influence organization (see Section 3.5 for exact values used).

3.2.2 Integrative data and model analysis reveals selective adhesion differences driven by EphA4/EphrinB2 heterogeneity promotes proper sorting of the PE and Epi

In the following results, we classify the inner cell mass during lineage specification as suggested in [92] into Nanog+ (cells expressing high Nanog and low Gata6 committed to Epi), Gata6+ (cells expressing high Gata6 and low Nanog committed to PE), double positive (DP, cells expressing high Nanog and Gata6) and double negative (DN, cells expressing low Nanog and Gata6). We call a simulated pattern successful if a single aggregation of Epi cells is formed and is attached to the TE shell covered by a PE layer. A simulated pattern is classified as partial success if multiple Epi aggregates form but are still attached to the TE shell and covered by a PE layer. A simulated pattern is classified as failure if the embryo stays in salt-and-pepper pattern or Epi and PE form separate clusters. We have also developed a loss score to quantify the divergence of a simulated pattern from an ideal pattern (see Materials and Methods: Embryo pattern loss score).

After initial cell fate specification, Nanog+ and Gata6+ cells form a salt and pepper configuration. Biased active cell movement mediated by intercellular interactions was suggested [81] has been suggested to lead to organization of the resulting PE and Epi structures. Specifically, differential or selective adhesion has been proposed to sort cells of differing fates in a number of scenarios [25]. In the context of the embryo, prior phenomenological modeling demonstrated that selective adhesion between Epi/PE/TE can induce Epi/PE separation [49]. This study did not however account for the presence of DP cells (expressing both Nanog and Gata), which were recently shown to co-exist with Nanog+ and Gata6+ cells [92]. Further, candidate molecules that facilitate this selective adhesion have not been identified to our knowledge. We thus first investigate this sorting process accounting for the additional presence of DP cells and use single cell data to identify and test candidate

adhesion molecules that may drive selective adhesion.

To incorporate a data-informed mechanism for cell sorting into the model, we first quantified expression levels of adhesion-repulsion associated genes for different cell types from single-cell RNA sequencing data [65]. A family of ligand-receptor pairs, Eph/Ephrin, has been shown to contribute to selective adhesion, and different pairs may lead to strengthened or weakened adhesion [96, 83]. In particular, EphB2/EphrinB2 triggers repulsion [53], and EphA4/EphrinB2 increases adhesion [82]. Moreover, it has been shown that Eph/Ephrin pairs contribute to somite formation [36, 29]. We thus use single-cell RNA-seq data on mouse early embryo (E3.5 to E6.75) to quantitatively assess expressions of different Eph/Ephrin pairs [65].

We quantified the combined ligand/receptor expression level in the ICM using the logarithm of multiplication of ligand and receptor expression levels, based on the EphA4/EphrinB2 expression level in scRNA-seq data [65]. Quantification of these genes in the ICM shows that EphA4/EphrinB2 forms a two-mode Gaussian mixture distribution pattern (Fig 3.2a, green curve), which is a mixture of one with high adhesion gene expression (Fig 3.2a, blue curve) and one with low adhesion gene expression (red curve). Further analysis shows that different fractions of Nanog+/Gata6+/DP cells are in the high/low adhesion gene expression states (which is quantified through binarization based on the Gaussian mixture) (Fig 3.2b). Fewer Gata6+ cells highly express adhesion genes compared to Nanog+/DP cells, which suggests a stronger adhesion among Nanog+/DP cells than Gata6+ cells.

To test if this distribution of EphA4/EphrinB2 can drive correct spatial patterning, we first derived a data-informed selective adhesion model based on these expression levels. The selective adhesion is modeled by calculating adhesion scores (AS) directly from expression of ligand-receptor pairs in scRNA-seq data [33, 34] as described in Eq (3.8) and (3.9).

We focus primarily on the pair EphA4/EphrinB2 at E4.5. We also analyzed the EphB2/EphrinB2

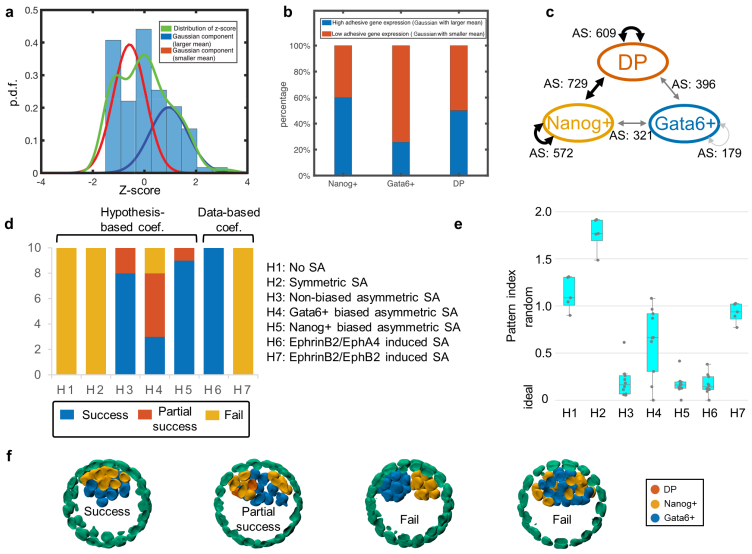

Figure 3.2: Data-informed selective adhesion model leads to successful cell arrangement at 128 cell stage.
(a) Histogram of z-score of summation of log([EphA4] + 1) and log([EphrinB2] + 1) at E4.5 stage. The green curve shows the distribution of z-score. The red and blue curves show two components of Gaussian Mixture to fit the distribution. (b) Percentage of high/low adhesive gene expression levels in Nanog+/Gata6+/DP cells at E4.5 stage. (c) An EphA4/EphrinB2 driven selective adhesion mechanism between Nanog+/Gata6+/DP cells, where higher adhesion score (AS) indicates stronger adhesion, and a positive AS means strengthened adhesion. (d) Success rate for Nanog+/Gata6+ cells arrangement in simulations with different selective adhesion hypotheses: (H1) no selective adhesion; (H2) symmetric selective adhesion where adhesion between Nanog+/Nanog+, Gata6+/Gata6+ and DP/DP cells are the same; (H3) asymmetric selective adhesion where DP cells have same adhesion with Nanog+ and with Gata6+; (H4) asymmetric selective adhesion where DP cells have stronger adhesion with Nanog+ cells than with Gata6+ cells; (H5) asymmetric selective adhesion where DP cells have stronger adhesion with both Nanog+ and Gata6+ cells; (H6) the Epha4/Efnb2 driven selective adhesion; (H7) the Ephb2/Efnb2 driven selective adhesion. (e) Pattern loss score of the simulations. Each data point corresponds to one simulation. A loss score of 0 indicates a perfect pattern and random cell type assignments have an expected loss score of 1. (f) Representative terminal Nanog+/Gata6+ cell arrangements for successful, partially successful, and failed cases.

combination (Fig 3.2d,e, model H7). However we are less confident in this data since EphB2 was unidentified in most cells, potentially due to dropout in the scRNA-seq data. Since our Eph/Ephrin quantification is limited to ICM cells due to the exclusion of TE cells in scRNA-seq data [65], we assumed adhesion between TE and Epi cells is stronger than that between PE and TE cells (following [49]). We also validated this mechanism (Fig 3.6) by demonstrating that without this interaction, organization fails.

For simulations of this Eph/Ephrin based adhesion model, we allowed the embryo to develop in silico to the 64-cell stage and then turned on selective adhesion (further examination of the effects of this starting time are discussed in the next section). For EphA4/EphrinB2, the normalized adhesion strength (AS) is calculated using Eq (3.9) (Fig 3.2c). In the simulations, EphA4/EphrinB2 driven selective adhesion was able to generate correct spatial pattern (Fig 3.2d,e, model H6). See Fig 3.2f for examples of success, partial success and failure cases of pattern formation. Based on the limited data available, results suggest the EphB2/EphrinB2 model may be insufficient to achieve organization (though better single cell data is needed here). These simulation results suggest an EphA4/EphrinB2 driven model of selective adhesion is sufficient to organize the Epi/PE structures.

This is of course not the only possible model of selective adhesion. We thus further studied the space of potential selective adhesion possibilities using phenomenological adhesion models that simply assign cell type dependent adhesions. By studying a range of different cell-type dependent adhesion models (Fig 3.2 and Fig 3.7), we identified two models that lead to effective organization (Fig 3.2d,e, models H3 and H5). In both effective models, there is a stronger adhesion among Nanog+ cells than that among Gata6+ cells and between Nanog+/Gata6+ cells. In model H5, DP cells have a stronger adhesion to Nanog+ cells whereas in H3, DP cells exhibit unbiased adhesion. Notably, model H5 recapitulates the qualitative dynamics predicted by the model using single cell expression of EphA4/EphrinB2.

In conclusion, the model suggests that EphA4/EphrinB2 distributions observed in scRNA

data from the embryo are sufficient to promote Epi/PE sorting. Further analysis of the space of possible selective adhesion models reveals two potential models that could in principle lead to proper organization. One of these two models has the exact qualitative structure found from the adhesion score analysis of EphA4/EphrinB2 interactions. Taken together, these results suggest this is a candidate ligand/receptor pair that drives sorting of the Epi/PE.

3.2.3 Selective adhesion mechanism occurrence before 128-cell stage ensures correct Epi/PE pattern formation.

We now consider the effect of timing of selective adhesion onset on embryonic organization. In the prior section, we artificially implemented this to occur at the 64-cell stage. It could however potentially take effect either earlier or later. Thus, we tested the EphA4/EphrinB2 driven mechanism with three different initiation times: immediately after the system reaches 32-cell stage, 64-cell stage or 128-cell stage. In the simulations, all three models give similar Nanog+/Gata6+/DP ratios round 40%/60%/0%, which are consistent with experimental data [92] (Fig 3.3a). These results suggest that the ratio between Nanog+/Gata6+/DP and spatial pattern are robust to the selective adhesion occurrence time. On the other hand, all the simulations with the selective adhesion starting time at 64- or 32-cell stage achieved correct final pattern (Fig 3.3b) while three out of ten simulations with the selective adhesion starting from 128-cell stage are only partially successful. In these partially successful cases, some Gata6+ cells aggregate near the TE (Fig 3.3c). This defect is potentially due to the absence of a time window overlap between the selective adhesion and the cell type transition from DP to Nanog+/Gata6+. These results suggest that while the Nanog+/Gata6+/DP ratio is robust to selective adhesion occurrence time, the corporation between selection adhesion and the cell fate regulation dynamics is crucial to the formation of correct spatial pattern.

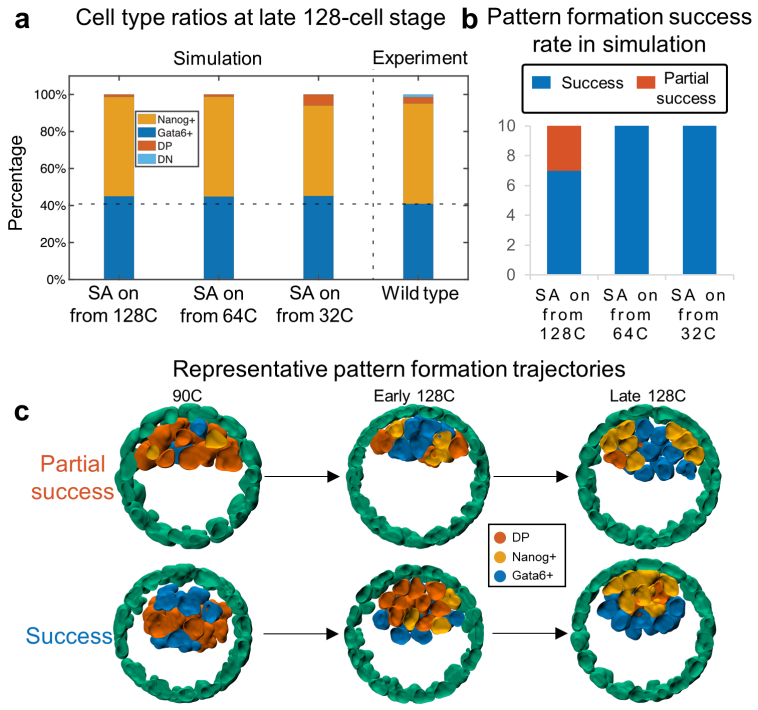

Figure 3.3: Sensitivity of selective adhesion starting time.
(a) ratio of Nanog+/Gata6+/DP/DN at 128C stage for different selective adhesion (SA) starting time; (b) Success rate for embryo development at 128C for different SA starting time. (c) Spatial pattern of simulation where selective adhesion starts from 128 cell stage (a partially successful case) and simulation where selective adhesion starts from 64 cell stage (a successful case).

We also evaluated the impact of cell movement randomness level (implemented as a Gaussian noise on subcellular element movement) on spatial pattern formation. Simulations were carried out with less random movement (1/10 of the amount of Gaussian noise compared to baseline simulation) and high random movement (5 times the amount of Gaussian noise compared to baseline simulation). Both Nanog/Gata6 expression levels and population ratio of Nanog+/Gata6+/DP were similar between these simulations and the baseline model (Fig 3.8). With lower movement randomness, the inner cell mass had a flatter shape, and there were some misplaced Gata6+ cells, whereas, with higher movement randomness, the inner cell mass has a rounder shape with few misplaced cells. In summary, the population ratio between Nanog+/Gata6+/DP is relatively robust the level of random movement but sufficiently large randomness of motion is required to reduce the instances of misplaced cells and ensure proper organization.

3.2.4 Attenuation of Fgf signaling after Epi/PE formation is required to maintain organization

Nanog and Gata6 are key specification factors for Epi/PE cells [74]. Image data shows that Nanog/Gata6 expression levels are both high at early blastocyst (32-cell stage) and become mutually exclusively expressed in Epi/PE cells at late blastocyst (128-cell stage) with Nanog+/Gata6+ cells maintaining a relatively stable ratio in late blastocyst stage: 55%-60% for Gata6+ cell and 40%-45% for Nanog+ cell [92]. Experimental evidence suggests this dynamics of Nanog/Gata6 expression is regulated by Fgf4/Fgfr2 signaling. Nanog+ cells secrete Fgf4 signal. Fgfr2 receptors, which are activated by Fgf4, promote the expression of Gata6 and antagonize Nanog expression [74]. A mathematical model with this Fgf/Fgfr/Erk signaling modulating Nanog/Gata6 expression was shown to generate appropriate fractions of Nanog+ and Gata6+ cells from the initial DP pool [4]. However, cell division and cell movement were excluded from this model, which are key processes in embryo development.

We incorporated Fgf/Fgfr2/Erk signaling into a spatial model (including cell division and cell movement) to evaluate the role of Fgf signaling on Nanog+/Gata6+ specification and maintenance during embryo development.

First, we evaluated whether Fgf signaling regulation on Nanog/Gata6 is necessary for Nanog+/Gata6+ lineage specification. Simulations with only Nanog/Gata6 mutual cross inhibition but without Fgf signaling cannot lead to the cell fate separation into Nanog+ and Gata6+. These simulations lead to incorrect Nanog/Gata6 expression patterns (DN dominated, DP dominated, Nanog+ dominated), regardless of the basal expression rate or their cross inhibition strength (Fig 3.9). Thus, as suggested by others, Fgf signaling regulation of Nanog/Gata6 likely ensures a correct Nanog+/Gata6+ population ratio. However, it is important to recall that the Fgf mediated specification process is occurring coincident with cell sorting.

We thus evaluated how the time period during which Fgf signaling is active influences cell ratios while sorting is occurring. We first test whether the timing of Fgf signaling termination influences organization. We carried out simulations where Fgf signaling either 1) is attenuated at 128-cell stage or 2) persists through the simulation. The control parameter ϵ_t in Eq (3.1) and (3.2) is used to modulate the effect of Fgf signaling mathematically (See Section 3.5 for details). Fgf attenuation leads to a Nanog+/Gata6+ cell type separation at the128-cell stage (Fig 3.4a) with cell type ratios consistent with experimental data [16] (Fig 3.4b). In contrast, when Fgf signaling remains persistent, there are more uncommitted cells (Fig 3.4d,e) and the Nanog+/Gata6+ ratio is inconsistent with experimental data (Fig 3.4b). Further results show that when Fgf signaling is persistent, 18.6% of Nanog+ cells highly express Gata6 (Fig 3.4d) and Gata6 expression is higher overall in these Nanog+ cells than when Fgf signaling is attenuated at the 64 cell stage (Fig 3.4e). Based on these results, we propose that the function of Fgf4/Fgfr2 is likely to be attenuated at 128-cell stage.

To test this hypothesis, we quantified the Fgf4/Fgfr2 expression levels in Nanog+/Gata6+ cells at different embryo developmental stages using scRNA-seq data [65] (Fig 3.4c, Fig 3.10).

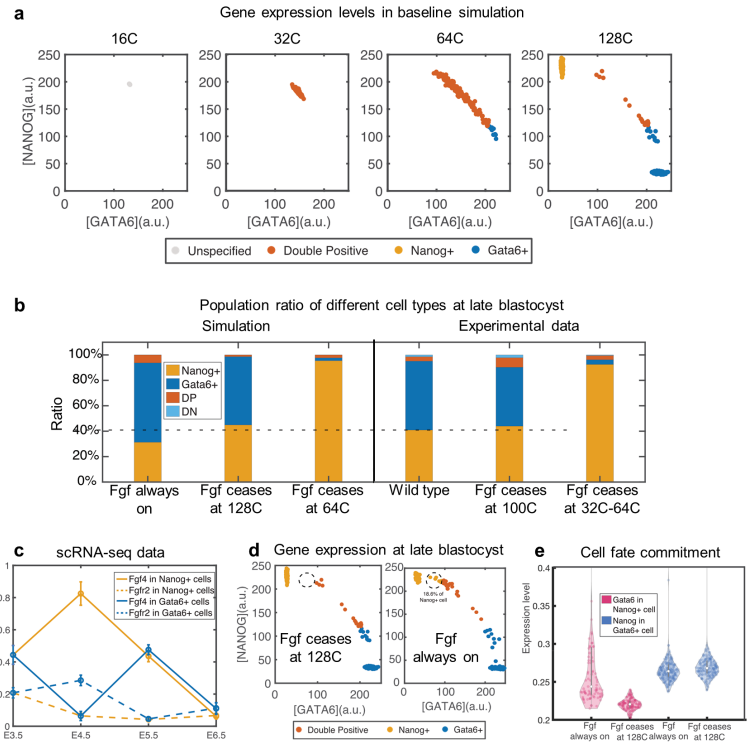

Figure 3.4: Nanog/Gata6 pattern for different Fgf4 ceasing time.
(a) Single cell gene expression level in different cell stages in baseline simulation. (b) Cell number percentage of DP/Nanog+/Gata6+/DN cells at different 128C (simulation)/ 100C-150C (experiment) stages with different Fgf ceasing time in simulations and experiments. The horizontal dashed line shows the percentage of Nanog+ cells in wild type experiment. (c) Fgf4/Fgfr2 expression levels (log1p transformed) in Nanog+/Gata6+ cells in different cell stages from scRNA-seq data. The vertical bar shows the standard deviation. (d) Gene expression level of Nanog and Gata6 at 128 cell stage if Fgf ceases at 128C and if Fgf is always on. (e) Violin plot of Gata6 level of Nanog+ cell (red) and Nanog level of Gata6+ cell (blue) for simulation where Fgf4 is always on and simulation where Fgf4 ceases at 128 cell stage. The hollow circle shows the median.

Results reveal that Fgf4/Fgfr2 expression levels are maximally different between Nanog+ and Gata+ cells at E4.5. After this stage, expression levels are essentially the same in the two populations, suggesting Fgf signaling is no longer modulating Nanog or Gata6 in a cell type-dependent fashion. This is consistent with the model hypothesis that Fgf signaling should attenuate after organization is achieved. Experiments also show that the percentages of Nanog+/Gata6+ cells are similar in wild type and the mutants with Fgf/Fgfr inhibitor added at 128-cell stage [26, 92, 3], further indicating the loss of function of Fgf/Fgfr after the 128-cell stage.

We next tested the scenario where Fgf signaling ceases at the 64-cell stage. The result shows there are only Nanog+ cells present at late 128-cell stage, consistent with experimental data (Fig 3.4b). These results suggest that Fgf signaling must become active at or prior to the 64-cell stage, but must cease functioning as final organization (late 128-cell stage) is reached. Fgf signaling early is required to ensure the salt and pepper distribution forms with proper cell ratios. However, Fgf signaling in this setting always tries to produce a salt and pepper distribution. Thus as the PE and Epi form through sorting, the Fgf signaling must be attenuated to ensure cells do not erroneously differentiate to reform that salt and pepper distribution.

Having shown that Fgf signaling is likely to attenuate after 128-cell stage, we now investigate the beginning time of Fgf signlaing. Prior to cell fate divergence, both Nanog and Gata6 are highly expressed at the beginning of early blastocyst. This motivates us to explore whether Fgf signaling is necessary for the high expression of Nanog/Gata6 at the beginning of early blastocyst stage. We tested four scenarios where Fgf signaling begins to regulate Nanog/Gata6 expression from 1-cell, 16-cell, 32-cell, and 64-cell stages (and ceases at 128-cell stage). The expression patterns are similar between simulations with Fgf signaling beginning at 1-cell and 16-cell stages. However, if Fgf signaling does not begin until the 32-cell or 64-cell stage, cell expressions are biased to the Nanog+ state prior to Fgf onset.

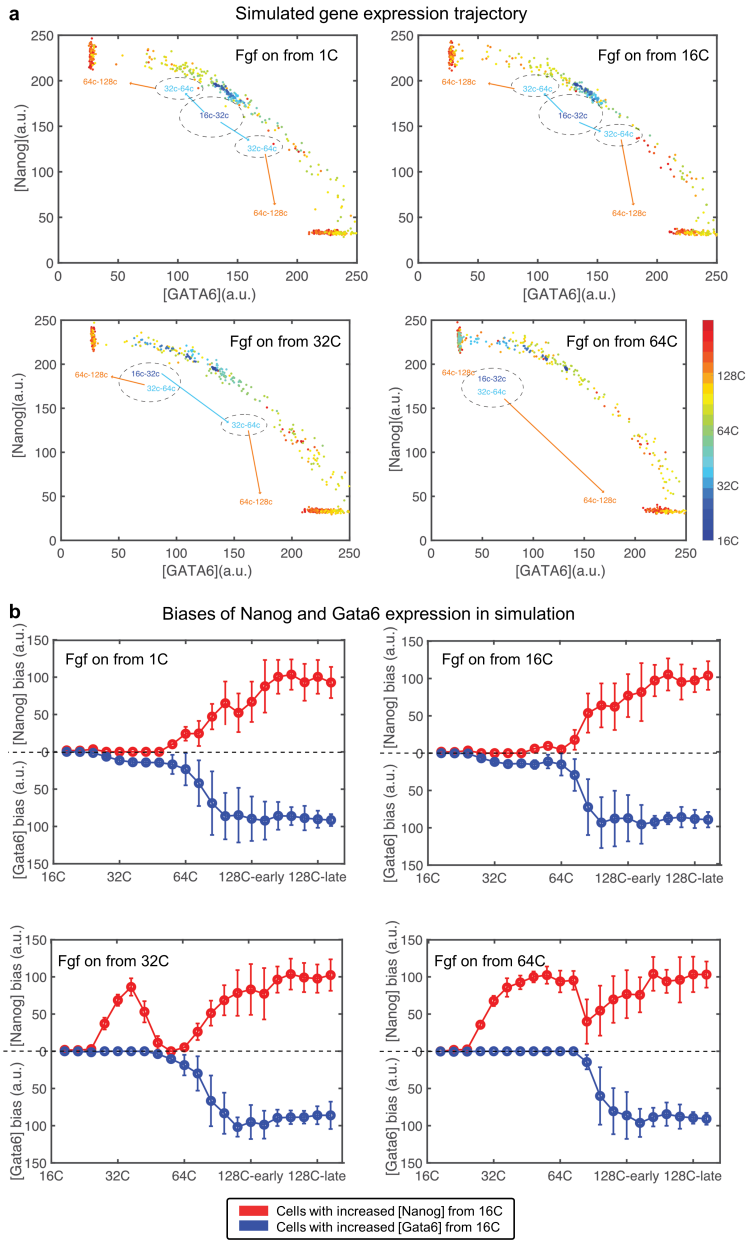

Figure 3.5: Sensitivity of Fgf on-time.
(a) Trajectory of Nanog/Gata6 expression over time for different Fgf on-time. Each dot represents one cell. The color of the cell represents the stage of the cell. Dashed circles represent the range of Nanog/Gata6 expression levels for different stages: dark blue for 16C-32C, little blue for 32C-64C and orange for 64C-128C. (b) Bias of Nanog or Gata6 gene expression level from the initial state (16C) in simulation with different Fgf on-time. Red curve shows the cells whose Nanog expression is higher than cells in initial state; Blue curve shows the cells whose Gata6 expression is higher than cells in initial state. The vertical line shows the error bar with one standard deviation.

Even in these cases though, gene expression becomes similar to baseline simulations after Fgf signaling commences (Fig 3.5). This suggests that while an early starting Fgf signaling is required to form DP cells at 32-64 cell stage, a later starting Fgf signaling can still induce the formation of Nanog+/Gata6+ separation even if the system has been biased to Nanog due to the absence of early Fgf signaling. Recently, experimental studies with inhibition of Fgf signaling at different stages [26, 92, 3] have shown that only Epi (Nanog+) cells are present if Fgf signaling is inhibited before early blastocyst (32-cell) stage but that PE (Gata6+) cells can be recovered if Fgf signaling is on at or after 32-cell stage. This result supports our finding that only Nanog+ cells exist when Fgf signaling is off, and some of Nanog+ cells can transfer into Gata6+ cells driven by Fgf signaling. Thus while Fgf signaling appears to be present prior to blastocyst formation, it may not be required at that stage to ensure Epi/PE formation.

Sufficiently strong gene expression noise was shown to be required to form the correct ICM/TE pattern [40]. We thus study the impact of noise levels on the expression of Nanog/Gata6 during Epi/PE formation. We considered scenarios where Nanog/Gata6 expression noise is lower (1/10 the strength) or higher (2 times the strength) than baseline levels. In the simulation, a higher noise level leads to fewer Gata6+ cells with commensurately more Nanog+ cells at 128-cell stage compared to the baseline simulation (Fig 3.11. This suggests that a high noise level on Nanog/Gata6 expression can lead to an incorrect Nanog+/Gata6+ ratio.

To evaluate whether our model is able to capture correct experimental pattern in different initial conditions of Nanog/Gata6 expression, we used the Nanog/Gata6/Fgf4/Fgfr2 scRNA-seq expression data as the initial condition at 32- or 64- cell stage for simulations starting at the corresponding stages. Thus far, all initial conditions have been spontaneously determined by the gene expression ODE's. Here, we initiated simulations from 32-cell stage and 64-cell stage (Fig 3.12) using the corresponding single-cell qPCR data [33, 34] as initial con-

ditions respectively. In both simulations, correct spatial patterns formed and were similar to the imaging data at 64-cell and 128-cell stages [92] (Fig 3.12). Besides, the ratio between Nanog+/Gata6+ cells was qualitatively consistent with the image data [92] when different initial conditions were used. Despite the demonstrated robustness, the model was also able to reflect the differences in initial conditions in the simulation results. A Nanog+/Gata6+ population ratio more biased to Gata6 was obtained when the initial conditions contained more Gata6+ cells (Fig 3.12).

In conclusion, the model reveals that Fgf signaling is likely to attenuate after 128-cell stage, the correct Nanog+/Gata6+ population at 128-cell stage is robust to the time of Fgf signaling onset, organization can be achieved with biologically observed cell distributions of Nanog/Gata6 at the 32 and 64-cell stages, and that the noise level in Nanog/Gata6 gene expression has the potential to alter the proportion of Nanog+ and Gata6+ cells.

3.3 Discussion

We developed a first of its kind (to our knowledge) multiscale model of development of the early embryo from the 1-cell to 128-cell stage. This three-dimensional model couples intracellular and intercellular gene regulation with selective adhesion in a realistic embryonic geometry to model the spatiotemporal development trajectory from the oocyte (1-cell stage) to blastocyst (128-cell stage). Further, we use scRNA-seq data to inform both the regulatory and the adhesion interaction parameters of this model.

Using this model, we study the roles of Fgf signaling on gene regulation, along with selective adhesion mediated cell motions on organization of the PE and Epi. This model makes two essential predictions. First, that the specific ligand-receptor pair EphA4/EphrinB2 is likely a major driving force behind the selective adhesion necessary to sort the initial salt and pepper

distribution of cells into the PE and Epi structures. While models of embryonic development have been previously used to demonstrate that selective adhesion may facilitate formation of the Epi/PE, those studies had two shortcomings. First, they did not account for the presence of double-positive cells expressing both Nanog and Gata6. Second, they were phenomenological in nature and the adhesion mechanisms were not supported by experimental data. Here, used expression of Eph/Ephrin pairs from scRNA-seq data directly parameterize our model. Results show that EphA4/EphrinB2 distributions among Nanog+ and Gat6+ cells would be sufficient to direct sorting. Thus while prior models have demonstrated the necessity of cell sorting, our results demonstrate the sufficiency specifically of EphA4/EphrinB2 in directing that sorting. This provides the first directly testable hypothesis for the presence of adhesion mediated cell sorting in the embryo.

Second, we demonstrate that attenuation of Fgf signaling is necessary after organization of the PE/Epi is achieved. Fgf signaling is known to modulate cell fate decisions during the formation of the initial salt-and-pepper distribution of cells [102, 114, 50]. We explore when this timing should be active. Results show that proper organization is not sensitive to the timing of signaling initiation. When Fgf is not present early on, the model accurately predicts a reduced presence of DP cells commonly observed at the 64 cell stage. However, as long as signaling is present by the 64 cell stage, cell allocations are correct as is organization. We do however find that organization strongly depends on when Fgf signaling is terminated. Fgf signaling must be attenuated after organization is achieved. Consistent with this, further single cell analysis of Fgf/Fgfr expression demonstrates that they are homogeneously distributed prior to E3.5 and after E5.5, suggesting it is not effecting gene expression in a cell type dependent fashion. Thus, it appears that Fgf may only be actively influencing gene expression during the initial formation of the salt and pepper distribution.

Both Fgf signaling and selective adhesion have been shown to be successful in regulating cell type specification and pattern formation, respectively. However, the corporation between

them is mostly unknown. The model revealed that a temporal overlap of these two mechanisms could increase the robustness of early embryo development. Specifically, the model occasionally led to PE cells aggregated between TE layer and Epi cells if there is no time overlap between Fgf signaling and selective adhesion.

Having shown that Fgf signaling is likely to attenuate after 128-cell stage, question follows that how is Fgf downregulated. We hope to explore this direction in the future based on this modeling framework utilizing the emerging data resources to identify potential regulators of Fgf signaling and study how they participate in later embryo developments. Moreover, it is unknown how is Fgf signaling initiated such that its activity significantly increases during a short period from E3.5 to E4.5. These questions can potentially be addressed using the emerging spatial transcriptomics techniques [89, 20, 99] and integrating machine learning techniques with modeling. In general applications, modern spatial transcriptomics data will allow the construction and validation of complex models with numerous genes and interactions. The data-informed model allows the accommodation of realistic 3D geometries complementary to the spatial transcriptomics data which are mostly in 2D. Moreover, the data-informed models will, in turn, serve as a sandbox to predict the outcomes from perturbations of the mechanisms which can generate numerous in silico spatial data under different conditions and mechanisms.

In the temporal direction, it is worth extending the data-driven model to later developmental stages. This longtitudinal extension can be informed and validated with the available scRNA-seq data at later temporal points, such as the datasets upto E8.5 [79, 71] or between E9.5 and E13.5 [6]. Moreover, gene regulatory networks could be extracted from scRNA-seq data in an unsupervised manner to promote the automation of the modeling framework. The potential simulated embryos beyond late blastocyst can be further validated using spatial gene expression data [86].

While this work partially relies on knowledge of gene regulatory networks and candidate

regulators of selective adhesion, it is also possible to unsupervisedly integrate modeling with data to uncover novel gene-gene interactions and unknown selective adhesion genes. The data has helped us improve the model, it is worth exploring the other direction of using model to improve data analysis. For example, one can systematically screen signaling through different ligand-receptor pairs after predicting the spatial arrangements of scRNA-seq data using the model. In addition to the interactions and communications between cells, it is still an open question of how hydrodynamics in the blastocoel influences Epi/PE patterning. Since our model resembles the geometry in real biological tissues, it is well suited for the future integration of a fluid mechanics model of the blastocoel. Finally, due to the efficient implementation that harnesses GPU computing, this model could be extended to study later developmental stages involving more cells.

3.4 Materials and Methods

3.4.1 Model equations and simulations

The model consists of three major components: 1) a gene regulatory network model addressing cell type specification, 2) a subcellular element model describing spatial dynamics of cells, and 3) a data-informed adhesion force model.

Gene regulatory network model

The gene expression dynamics of Nanog/Gata6/Fgf are represented by the following non-dimensional stochastic ODE system:

$$\frac{d[Nan]_i}{dt} = vsn_0 + \underbrace{\frac{vsn_1 \cdot Kin_i^u}{Kin_1^u + \epsilon_t[Erk]_i^u}}_{\text{Inhibition by Erk}} + \underbrace{\frac{vsn_2 \cdot [Nan]_i^v}{Kan^v + [Nan]_i^v}}_{\text{Self amplification}} \cdot \underbrace{\frac{Kin_2^w}{Kin_2^w + [Gat]_i^w}}_{\text{Inhibition by Gata6}} - \underbrace{k_N[Nan]_i}_{\text{Degradation}} + \underbrace{\sigma_N[Nan]_i \cdot \eta_N}_{\text{Noise}} \tag{3.1}$$

$$\frac{d[Gat]_i}{dt} = vsg_0 + \underbrace{\frac{vsg_1 \cdot \epsilon_t[Erk]_i^r}{Kag_1^r + \epsilon_t[Erk]_i^r}}_{\text{Promotion by Erk}} + \underbrace{\frac{vsg_2 \cdot [Gat]_i^s}{Kag_2^v + [Gat]_i^s}}_{\text{Self amplification}} \cdot \underbrace{\frac{Kig^q}{Kig^q + [Nan]_i^q}}_{\text{Inhibition by Nanog}} - \underbrace{k_g[Gat]_i}_{\text{Degradation}} + \underbrace{\sigma_G[Gat]_i \cdot \eta_G}_{\text{Noise}} \tag{3.2}$$

$$\frac{d[Fr]_i}{dt} = vsfr_1 \underbrace{\frac{Kifr^x}{Kifr^x + [Nan]_i^x}}_{\text{Inhibition by Nanog}} + vsfr_2 \underbrace{\frac{[Gat]_i^y}{Kafr^y + [Gat]_i^y}}_{\text{Promotion by Gata6}} - \underbrace{k_{Fr}[Fr]_i}_{\text{Degradation}} + \underbrace{\sigma_{Fr}[Fr]_i \cdot \eta_{Fr}}_{\text{Noise}} \tag{3.3}$$

$$\frac{d[Fs]_i}{dt} = vsf \underbrace{\frac{[Nan]^z}{Kaf^z + [Nan]_i^z}}_{\text{Promotion by Nanog}} - \underbrace{k_{Fs}[Fs]_i}_{\text{Degradation}} + \underbrace{\sigma_{Fs}[Fs]_i \cdot \eta_{Fs}}_{\text{Noise}} \tag{3.4}$$

$$\frac{d[Erk]_i}{dt} = va[Fr]_i \underbrace{\frac{[Fp]_i}{Kd + [Fp]_i} \cdot \frac{1 - [Erk]_i}{Ka + 1 - [Erk]_i}}_{\text{Promotion by perceived Fgf4}} - \underbrace{k_{Erk}\frac{[Erk]_i}{Ki + [Erk]_i}}_{\text{Degradation}} + \underbrace{\sigma_{Erk}[Erk]_i \cdot \eta_{Erk}}_{\text{Noise}} \tag{3.5}$$

where $[Nan]_i, [Gat]_i, [Fr]_i, [Fs]_i$ and $[Erk]_i$ represent secreted Nanog, Gata6, Fgfr2, Fgf4, and Erk in cell i . The perceived Fgf4 from neighboring cells for cell i is described by

$$[Fp]_i = \sum_{j:|r_i - r_j| < r_{contact}} (1 + \gamma_j)\frac{[Fs]_j}{N_j} \tag{3.6}$$

where $r_{contact}$ is a cutoff determining if cells located at r_i and r_j are neighboring cells N_j is

the number of neighbors of cell j , and γ_j is a Gaussian noise.

Cell spatial dynamics

Every cell in the model is represented by a collection particles (elements) in space. The movement of element j of cell i is governed by the following differential equation:

$$\frac{d[Fs]_i}{dt} = vsf \underbrace{\frac{[Nan]^z}{Kaf^z + [Nan]_i^z}}_{\text{Promotion by Nanog}} - \underbrace{k_{Fs}[Fs]_i}_{\text{Degradation}} + \underbrace{\sigma_{Fs}[Fs]_i \cdot \eta_{Fs}}_{\text{Noise}} \tag{3.7}$$

where $r_{i,j}$ is its position, α is a parameter depicting the intercelluar adhesion strengths, and F_{exter} is the external forces driving zona pellucida confinement and cavity formation (see Section 3.5 for detailed equations).

Connecting with scRNA-seq data

The adhesion score is estimated from the expression levels of ligand-receptor pairs in scRNA-seq data:

$$AS_0(CT_i, CT_j) = (\frac{[L]_i[R]_j + [L]_j[R]_i}{2})^{n_{force}} \tag{3.8}$$

$$AS(CT_i, CT_j) = AS_0(CT_i, CT_j)/AS_0(Nanog+, Nanog+) \tag{3.9}$$

where $[L]_i$ and $[R_i$ are the average ligand and receptor expression levels among cells of type i (CT_i) obtained from single-cell data. The relative adhesion score in Eq (3.3.3.9) is assigned to the parameter α in Eq (3.7). The parameter n_{force} corresponds to the type of adhesion modification such that $n_{force} = 1$ for ligand-receptor pairs that strengthens adhesion (e.g. EphA4/EphrinB2) and $n_{force} = -1$ for ligand-receptor pairs that weaken adhesion (e.g. EphB2/EphrinB2).

Embryo pattern loss score

We use a pattern loss score to quantitatively describe the successfulness of the resulting patterns in simulations. Let $x_i, 1 \leq i \leq n_{\text{icm}}$ be the positions of the inner cell mass at the last frame of the simulation and t_i^{mdl} be the cell type of the i th which is either Epi or PE. Assuming the embryo is centered at the origin, we assign a reference location called the ICM pole at $r\bar{x}/||\bar{x}||$ where $\bar{x}$ is the average position of inner cell mass and r is the average distance from TE cells to the embryo center. We then assign an ideal cell type t_i^{ideal} that best replicates the known embryo pattern by assigning the n_{epi} cells closest to the ICM pole as Epi type and the rest as PE type. From these two cell type assignments, we define two index collections, $J_1 = \{i : t_i^{\text{mdl}} = \text{Epi}, t_i^{\text{ideal}} = \text{PE}\}$ and $J_2 = \{i : t_i^{\text{mdl}} = \text{PE}, t_i^{\text{ideal}} = \text{Epi}\}$. The difference between these two cell type assignments is quantified by $d_{\text{mdl}} = \inf_{y \in \Gamma}$ where Γ is the collection of all bijections from J_1 to J_2. Similarly, we also quantify the difference between randomly assigned cell types and the ideal cell type assignment, and repeat this process to derive a empirical expected difference $\bar{d}_{\text{rand}}$. Finally, we use $d_{\text{mdl}}/\bar{d}_{\text{rand}}$ as a normalized loss score quantifying the successfulness of the formed pattern. A simulated pattern identical to the ideal cell type assignment has a loss of 0.

Simulation

The model was implemented in C programming language. Parallel computing was used for the movement of elements through OpenCL. The simulations were carried out on the High Performance Computing Cluster at University of California Irvine.

3.4.2 Data analysis

Two different single-cell gene expression datasets were used: 1) a scRNA-seq dataset with 721 cells and 24484 genes at four temporal points from E3.5 to E6.5 [65], and 2) a single-cell qPCR dataset measuring 48 selected genes in 442 cells from the 1-cell stage up to 64-cell stage [33]. The spatial imaging data at early and late blastocyst consists of measurement of cell type marker genes Oct4, Cdx2, Gata6, and Nanog [48, 92]. See Section 3.5 for data processing details.

3.5 Modeling details

3.5.1 Spatial model

Subcellular element method

Subcellular element method [69] represents a cell by a collection of elements in space. The movement of element of cell is governed by the following differential equation:

$$\frac{d\mathbf{r}_{i,j}}{dt} = \underbrace{-\nabla_{i,j}\sum_{k\neq j} V_{\text{intra}}\ (|\mathbf{r}_{i,j}-\mathbf{r}_{i,k}|)}_{\text{Intracellular forces}} \underbrace{-\nabla_{i,j}\sum_{k\neq i}\sum_{l}\alpha\cdot V_{\text{inter}}\ (|\mathbf{r}_{i,j}-\mathbf{r}_{k,l}|)}_{\text{Intercellular forces}} + \underbrace{F_{\text{exter}}\ (\mathbf{r}_{i,j})}_{\text{External forces}} + \underbrace{F_{\text{image}}\ (\mathbf{r},\mathbf{g}_{\text{mdl}}\ ,\mathbf{g}_{\text{spa}}\ )}_{\text{Image data-driven}} \tag{3.10}$$

where $\mathbf{r}_{i,j}$ is its position, α is a parameter depicting the intercellular adhesion strengths, and F_{exter} is the external forces driving zona pellucida confinement and cavity formation. The term F_{image} is the image data-driven part using the modeled expression $\mathbf{g}_{\text{mdl}}$ and expression in imaging data $\mathbf{g}_{\text{spa}}$ of some spatial reference gene. Lennard-Jones potential commonly used

to describe atomic interactions is used here for the potential functions (V_{intra} and V_{inter}) describing interactions between elements:

$$V(r) = \varepsilon \left[\left(\frac{r_{\text{m}}}{r} \right)^{12} - 2 \left(\frac{r_{\text{m}}}{r} \right)^{6} \right] \tag{3.11}$$

where r_m is the distance that minimizes the potential function.

Cell division

The simulation begins with 1280 elements representing a single cell and ends with 128 cells each represented by 10 elements. When dividing a cell into two descendants, a dividing plane of random orientation is placed where it could divide the elements into two groups of the same size. The timing of the division was scheduled as the following. There are six milestone time points when the system reaches 1, 2, 4, 8, 16, 32, 64 cells. At each milestone time point, the division time for each cell was drawn from a uniform distribution. This approach mimics the process that the cell divisions are partially synchronized where they do not divide at the same time but they likely share similar growth speed. Once a cell is divided, its gene expression is carried on by its two descendants.

External forces

The confinement of cells from zona pellucida and the formation of the inner cavity are realized by applying external forces (F_{exter}in Eq. 3.10) to the elements.

Here we model the zona pellucida as a sphere with a fixed volume centered at the origin.

For an element $\mathbf{r}_{i,j}$, its adhesion to the membrane is modeled by

$$F_{\text{exter-zona}}\ (\mathbf{r}_{i,j}) = \frac{\mathbf{r}_{i,j}}{r_{\text{embryo}}} \min\left\{F_{\max}, F_0/\left(r_{\text{embryo}}\ - |\mathbf{r}_{i,j}|\right)\right\} \tag{3.12}$$

where r_{embryo} is the fixed radius of the embryo, $F_{\max}$ is a cutoff value to stabilize the movements, and F_0 is a coefficient for the force magnitude. In the implementation, if the position $\mathbf{r}_{i,j}$ becomes outside of the sphere after a simulation step, it is moved back into the sphere by setting its position to $(r_{\text{embryo}}\ - \epsilon)\frac{\mathbf{r}_{i,j}}{|\mathbf{r}_{i,j}|}$, where ϵ is a positive parameter with a small value.

After compaction, cavity is known to form by fluid inside the embryo. Many factors contribute to the formation of cavity such as the sophisticated salt and osmotic transport. Since our focus is not on the cavity fluids, for simplicity, we use a few phantom elements to occupy the space of cavity so that the regular elements representing cells could not enter the cavity. Specifically, when introducing the cavity, the cells should have differentiated into TE and ICM. We place the first phantom element on the zona pellucida that is farthest away from the current geometric center of ICM. Then, four more phantom elements are placed on the zona pellucida around this first one such that they are equally spaced and form a plane which is perpendicular to the direction specified by the first phantom element with a distance to the sphere center equal to 2/3 of the sphere radius. The forces due to these phantom elements are applied to elements $\mathbf{r}_{i,j}$ in ICM and are defined as

$$F_{\text{exter-cavity}}\ (\mathbf{r}_{i,j}) = -F_{\text{rep}} \sum_k \frac{\mathbf{r}_{i,j} - \mathbf{r}_{\text{phantom},k}}{|\mathbf{r}_{i,j} - \mathbf{r}_{\text{phantom},k}|} \tag{3.13}$$

where $\mathbf{r}_{\text{phantom},k}$ is the position of the k th phantom element and F_{rep} is a coefficient for the repulsion force that pushes the ICM away from the cavity.

3.5.2 Gene network model

Specification of TE and ICM

A mutual inhibition model of Oct4/Cdx2 is used to model the specification of TE and ICM [40]. The model is implemented as ODEs:

$$\frac{d[Oct]_i}{dt} = k_O \underbrace{\left(b_O + a_O \frac{[Oct]_i^n}{\theta_o^n + [Oct]_i^n}\right)}_{\text{Self amplification}} \underbrace{\left(1 + I_{CO} \frac{\theta_C^n}{\theta_C^n + [Cdx]_i^n}\right)}_{\text{Inhibition by Cdx 2}} - \underbrace{d_O[Oct]_i}_{\text{Degradation}} + \underbrace{\sigma_O[Oct]_i \cdot \eta_O}_{\text{Noise}} \tag{3.14}$$

$$\frac{d[Cdx]_i}{dt} = k_C \underbrace{\left(b_C + S_i + a_C \frac{[Cdx]_i^n}{\theta_C^n + [Cdx]_i^n}\right)}_{\text{Self amplification}} \underbrace{\left(1 + I_{OC} \frac{\theta_O^n}{\theta_o^n + [Oct]_i^n}\right)}_{\text{Inhibition by Cdx2}} - \underbrace{d_C[Cdx]_i}_{\text{Degradation}} + \underbrace{\sigma_C[Cdx]_i \cdot \eta_C}_{\text{Noise}} \tag{3.15}$$

where $[Oct]_i$ and $[Cdx]_i$ represent the relative expression levels of Oct4 and Cdx2 in cell i . The term η is a noise term of zero mean and unit standard deviation. In Eq. 3.15, S_i represents the cell contact impact on Cdx expression. The cell contact parameter S_i is defined as 1.5(number of outer elements)/(number of all elements). An element in a cell is regarded an outer element if its distance from the embryo boundary is shorter than 1. This parameter describes the contact intensity between a cell the embryo boundary.

Specification of Epi and PE

The specification of Epi and PE is modeled by a mutual inhibition model of Nanog/Gata6 mediated by Fgf signaling. The ODEs for the model are defined as:

$$\frac{d[Nan]_i}{dt} = vsn_0 + \underbrace{\frac{vsn_1 \cdot Kin_1^u}{Kin_1^u + \varepsilon_t[Erk]_i^u}}_{\text{Inhibition by Erk}} + \underbrace{\frac{vsn_2[Nan]_i^v}{Kan^v + [Nan]_i^v}}_{\text{Self amplification}} \cdot \underbrace{\frac{Kin_2^w}{Kin_2^w + [Gat]_i^w}}_{\text{Inhibition by Gata6}} - \underbrace{k_N[\text{Nan}]_i}_{\text{Degradation}} + \underbrace{\sigma_N[Nan]_i \cdot \eta_N}_{\text{Noise}} \tag{3.16}$$

$$\frac{d[Gat]_i}{dt} = vsg_0 + \underbrace{\frac{vsg_1 \cdot \varepsilon_t[Erk]_i^r}{\text{Kag}_1^r + \varepsilon_t[Erk]_i^r}}_{\text{Self amplification}} + \underbrace{\frac{vsg_2[Gat]_i^s}{Kag_2^v + [Gat]_i^s}}_{\text{Promotion by Erk}} \underbrace{\frac{Kig^q}{Kig^q + [Nan]_i^q}}_{\text{Inhibition by Nanog}} - \underbrace{k_G[Gat]_i}_{\text{Degradation}} + \underbrace{\sigma_G[Gat]_i \cdot \eta_G}_{\text{Noise}} \tag{3.17}$$

$$\frac{d[Fr]_i}{dt} = \underbrace{vsfr_1 \frac{Kifr^x}{Kifr^x + [Nan]^x}}_{\text{Inhibition by Nanog}} + \underbrace{vsfr_2 \frac{[Gat]^y}{Kafr^y + [Gat]^y}}_{\text{Prmotion by Gata6}} - \underbrace{k_{Fr}[Fr]_i}_{\text{Degradation}} + \underbrace{\sigma_{Fr}[Fr]_i \cdot \eta_{Fr}}_{\text{Noise}} \tag{3.18}$$

$$\frac{d[Fs]_i}{dt} = \underbrace{vsf \frac{[Nan]^z}{Kaf^z + [Nan]_i^z}}_{\text{Promotion by Nanog}} - \underbrace{k_{F_s}[Fs]_i}_{\text{Degradation}} + \underbrace{\sigma_{Fs}[Fs]_i \cdot \eta_{F_s}}_{\text{Noise}} \tag{3.19}$$

$$\frac{d[Erk]_i}{dt} = va[Fr]_i \underbrace{\frac{[Fp]_i}{Kd + [Fp]_i} \cdot \frac{1 - [Erk]_i}{Ka + 1 - [Erk]_i}}_{\text{Promotion by perceived Fgf4}} - \underbrace{k_{Erk}\frac{[Erk]_i}{Ki + [Erk]_i}}_{\text{Degradation}} + \underbrace{\sigma_{Erk}[Erk]_i \cdot \eta_{Erk}}_{\text{Noise}} \tag{3.20}$$

where $[Nan]_i, [Gat]_i, [Fr]_i, [Fs]_i$ and $[Erk]_i$ represent secreted Nanog, Gata6, Fgfr2, Fgf4, and Erk in cell i . The perceived Fgf4 from neighboring cells for cell i is described by

$$[Fp]_i = \sum_{j:|r_i - r_j| < r_{contact}} (1 + \gamma_j)\frac{[Fs]_j}{N_j} \tag{3.21}$$

where $r_{contact}$ is a cutoff determining if cells located at r_i and r_j are neighboring cells, N_j

is the number of neighbors of cell j , and γ_j is a Gaussian noise. The cutoff distance for cell-cell contact is defined as $r_{contact} = \frac{\sqrt{3}}{2}(r_i + r_j)$, where r_i and r_j are the average radius of the cell types of cell i and j.

3.5.3 Variants of baseline multiscale model

We describe how the different timing of the processes and the selective adhesion mechanisms are implemented in the hypothesis-driven model.

Controlling timing of processes

The simulation is carried out numerically with 260000 temporal steps. The embryo reaches 2-cell, 4-cell, 8-cell, 16-cell, 32-cell, 64-cell, 128-cell states at temporal steps 10000, 20000, 30000, 40000, 50000, 95000, and 120000 respectively. In different in silico experiments, we set the on or off time of FGF signaling or selective adhesion to 16-cell, 32-cell, 64-cell, and 128-cell stages by beginning or terminating the processes at steps 39000, 75000, 95000, and 120000 respectively. These simulation steps correspond to the t_0 and t_1 values in Table 3.2 that controls the value of ϵ_t.

Selective adhesion models

The potential function governing the intercellular interactions will generate repulsion forces if the cells are too close or attraction forces otherwise. We implement different cell-type dependent selective adhesion models by multiplying the generated forces with coefficients depending on the force direction.

Model interaction with data

Data informed components in models: 1) In models H6 and H7, the parameters of the selective adhesion mechanism are inferred using scRNA-seq data [65] based on the expression of Nanog, Gata6, EphA4, EphB2, EphrinB2. 2) In the simulation in Fig 3.12, the scRNA-seq data [65] and single-cell qPCR data [34] (Nanog/Gata6 expression) were used as initial conditions for the models. Validation of the model simulations: 1) The spatial pattern and cell type composition at late blastocyst are validated by the 3-dimensional imaging data [92]. 2) The observed attenuation of Fgf signaling activity from the model was supported by the analysis of Fgf expression levels in scRNA-seq data [65]. 3) The single-cell qPCR data [34] is used to validate the early stage of simulations (1C to 32C stages).

3.5.4 Data processing

scRNAseq data

The original data [65] was downloaded using the GEO accession code GSE100597. The count matrix was preprocessed with log1p operation. The cell types (Nanog+, Gata6+, double positive, or double negative) were determined based on the expression levels of Nanog/Gata6.

Single-cell qPCR data

The single-cell qPCR data [34] was downloaded as supplementary data of the original publication. We followed the original publication to assign cell types for 32C and 64C stage cells. A PCA was first done on the set of 64C cells. We then used k-means clustering with k=3 to split the cells into three clusters. The marker gene expressions were used to assign the cell types (TE: Cdx2, Epi: Nanog, PE: Gata6). The 32C cells were projected to the PCA

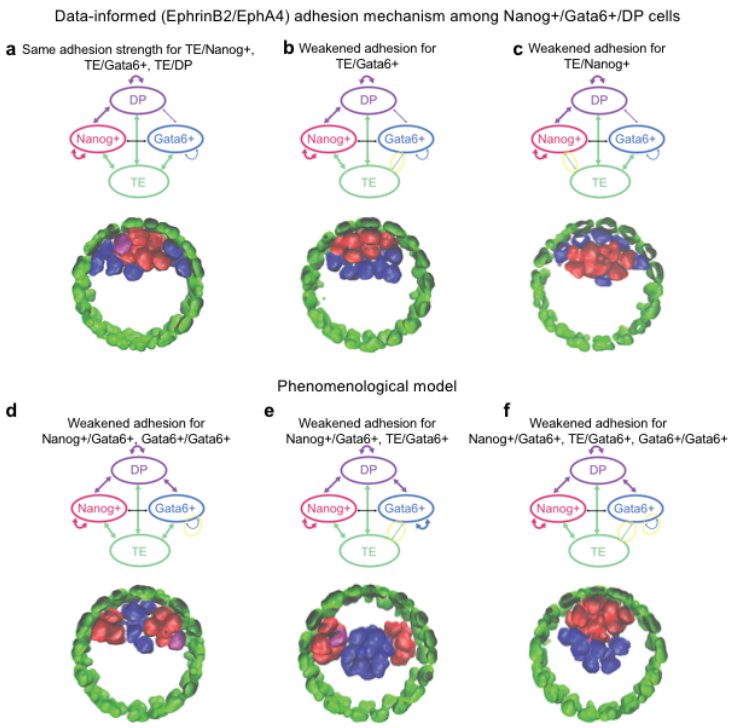

Figure 3.6: Adhesion mechanisms involving TE cells and the simulation results. a-c. Different adhesion mechanisms between TE cells and other cell types. The adhesion mechanism among Nanog+/Gata6+/DP cells is inferred from data based on the ligand-receptor pair EphrinB2/EphA4. d-f. Phenomenological models about selective adhesion involving TE cells.

space of 64C cells. The cells were separated into two clusters using k-means clustering with k=2. The marker gene expressions were then used to assign the cell types to the clusters (TE: Cdx2, ICM: Oct4).

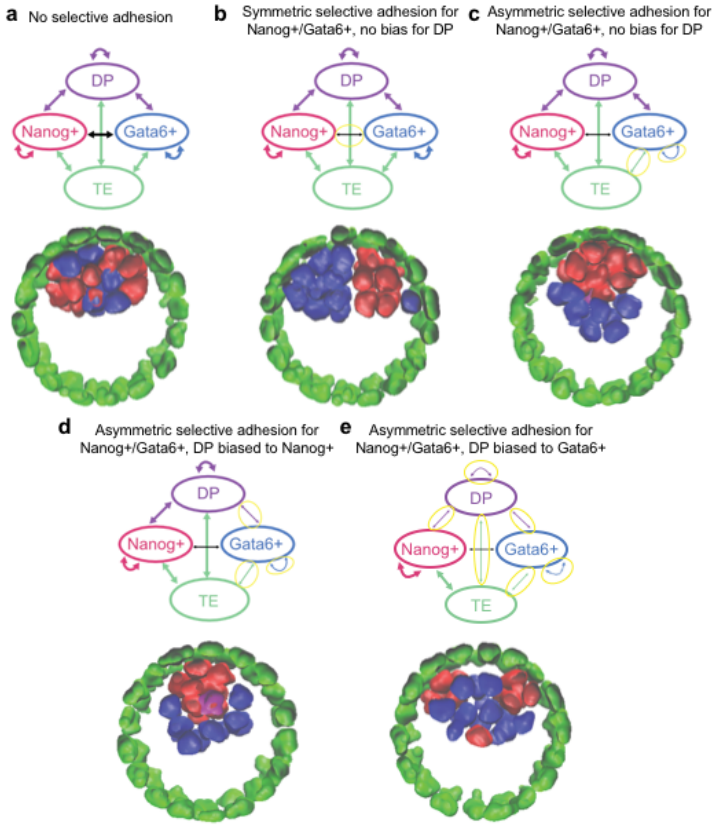

Figure 3.7: Adhesion mechanisms involving DP cells and the simulation results. a-c. Different selective adhesion mechanisms without bias for DP. d, e. Different selective adhesion mechanisms with biased adhesion for DP.

Parm.	Value	Parm.	Value	Parm.	Value
k_o	0.32	θ_o	0.5	σ_O	1.0
k_C	0.32	θ_c	0.5	σ_C	1.0
b_o	2.0	I_{CO}	1.5	η_O	$\sim \mathcal{N}(0,1)$
b_C	0.7	I_{OC}	1.5	η_C	$\sim \mathcal{N}(0,1)$
a_O	1.0	d_o	0.4	n	4
a_C	1.0	d_C	0.4		

Table 3.1: Parameters in Eq. (3.14) and Eq. (3.15). The parameters for the dimensionless equations are taken from ref. [40].

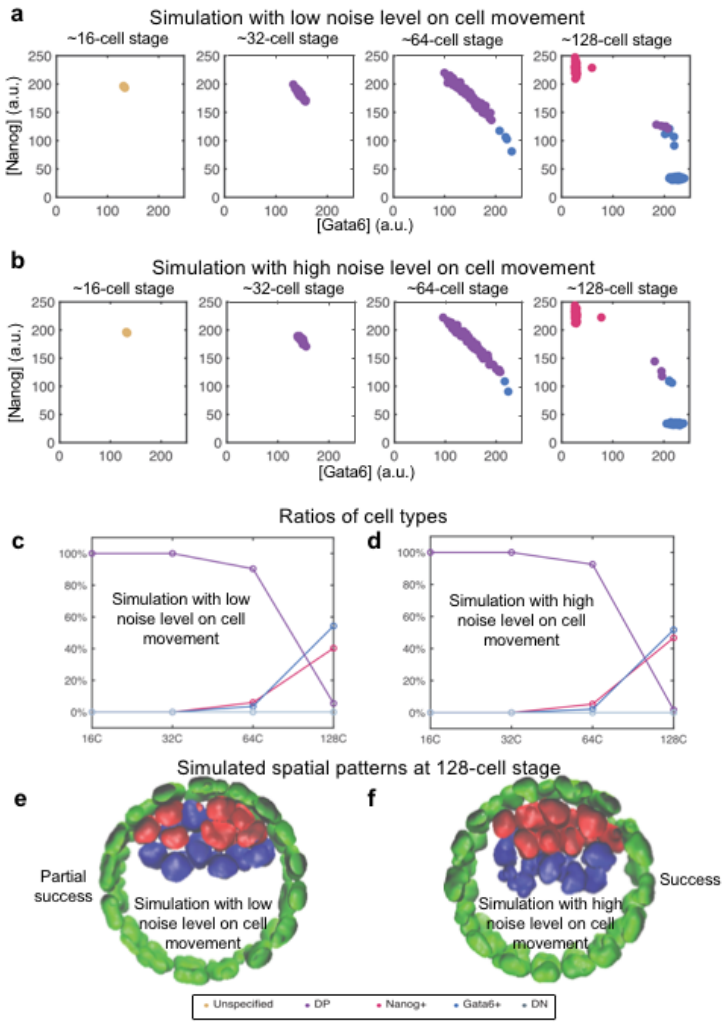

Figure 3.8: Simulation results with different noise levels on cell movement. a,b. The simulated Nanog/Gata6 expression levels with low or high cell movement noise level. c,d. The ratios of cell types during simulation with low or high cell movement noise level. e,f. Representative simulated spatial patterns at 128-cell stage.

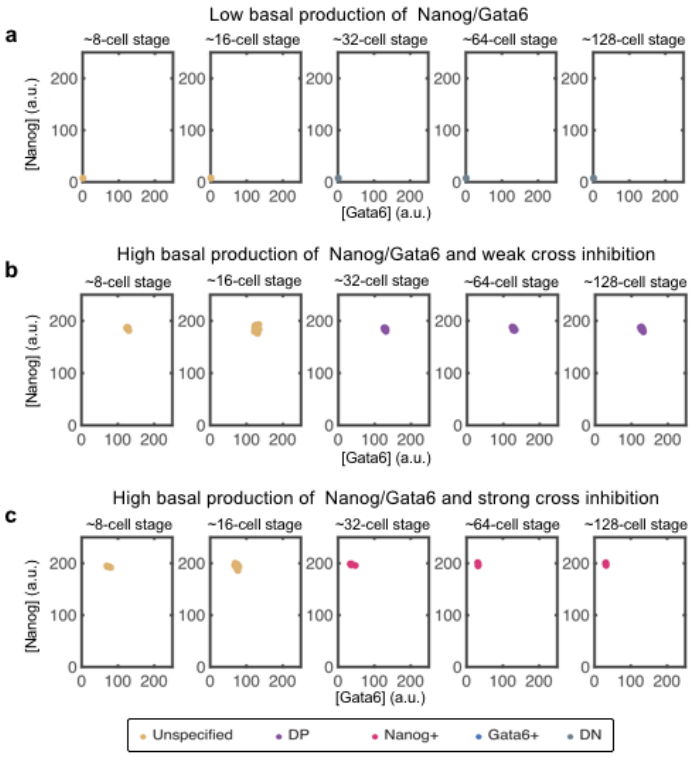

Figure 3.9: Simulation results with only mutual inhibition between Nanog and Gata6.

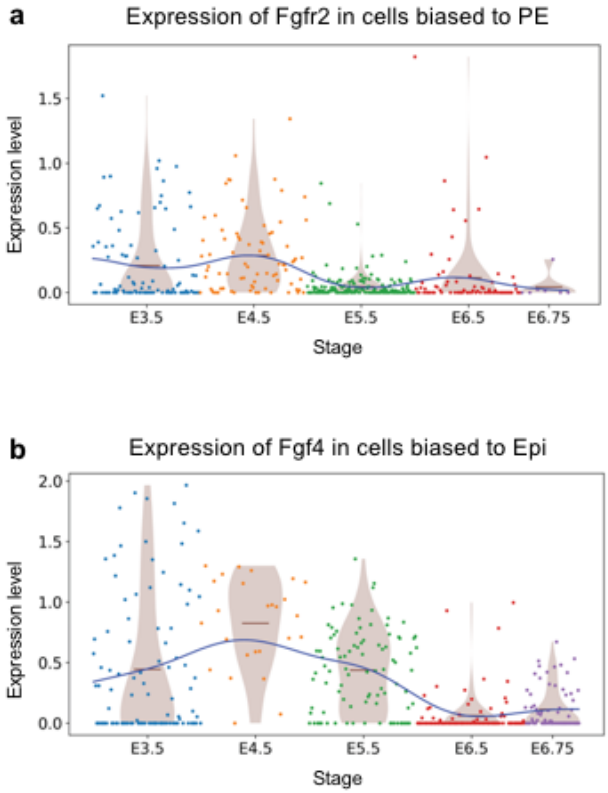

Figure 3.10: Expression of cell fate regulating genes in scRNA-seq data at different stages. a. The expression levels of Fgfr2 in cells biased to PE ([Nanog] $\leq$ [Gata6]) The curve is obtained from Gaussian process regression showing the trend and the bars in the violin plots show the mean values. Values after log1p transform of the original counts are used. b. Similar to (a) but for Fgf expression in cells biased to Epi ([Nanog ¿ [Gata6]).

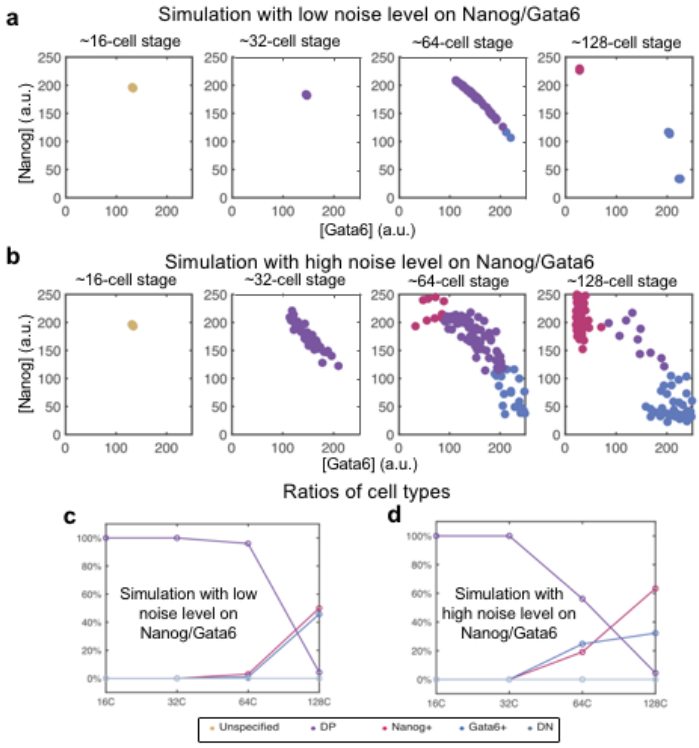

Figure 3.11: Simulation results with different noise levels on Nanog and Gata6 expression. a,b. The simulated Nanog/Gata6 expression levels with low or high noise level on Nanog and Gata6 expression. c,d. The ratios of cell types during simulation with low or high noise level.

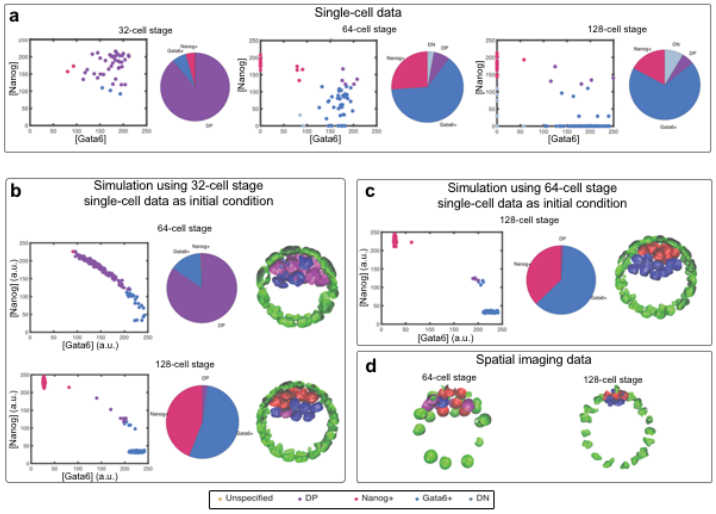

Figure 3.12: Baseline hypothesis-driven model simulation results using single-cell data as initial conditions.
a. Expression of Nanog and Gata6 of ICM cells in the single-cell datasets at different stages. The 32-cell and 64-cell stage data are taken from the single-cell qPCR dataset and the 128-cell stage data is taken from the scRNA-seq data. Expression values are normalized to match the scale in the model. b. Simulation result of the baseline hypothesis-driven model using the 32-cell stage single-cell qPCR data as the initial condition. c. Simulation result of the baseline hypothesis-driven model using the 64-cell stage single-cell qPCR data as the initial condition. d. The spatial pattern of Nanog+ and Gata6+ cells in spatial imaging data.

Parm.	Value	Parm.	Value	Parm.	Value
$^*vsn_0$	1.1041	Kag_2	0.55	$^*\sigma_{Fs}$	1.0
vsn_1	0.0051	$^*\sigma_G$	1.0	va	20.0
Kin_1	0.28	k_G	0.2	Kd	2.0
vsn_2	0.321	q	4	Ka	0.7
Kan	0.55	r	3	Ki	0.7
$^*\sigma_N$	1.0	s	4	k_{Erk}	3.3
k_N	0.2	Kig	1.6	$^*\sigma_{\text{Erk}}$	1.0
u	3	vsfr_1	2.8	x	1
v	4	Kifr	0.5	y	1
w	4	$\text{vsf}\, r_2$	2.8	z	4
ϵ_t	$1 \cdot I_[t_0, t1]$	$Kafr$	0.5	$^*\eta_N$	$N(0,1)$
Kin_2	1.604	k_{Fr}	1.0	$^*\eta_G$	$N(0,1)$
$^*vsg_0$	0.04103	$^*\sigma_{Fr}$	1.0	$^*\eta_{Fr}$	$N(0,1)$
vsg_1	0.04032	vsf	0.6	$^*\eta_{Fs}$	$N(0,1)$
Kag_1	0.28	Kaf	5.0	$^*\eta_{Erk}$	$N(0,1)$
vsg_2	0.321	k_{Fs}	0.077	γ_j	$N(0,1)$
ki	0.7				

Table 3.2: Parameters in Eq. (3.17-3.21). The parameters marked with * are calibrated and other parameter values are taken from ref. [4]

	H1	H1	H2	H2	H3	H3	H4	H4
	Rep.	Attr.	Rep.	Attr.	Rep.	Attr.	Rep.	Attr.
Epi-Epi	0.25	4	0.25	4	0.25	4	0.25	4
PE-PE	0.25	4	0.25	4	1	0.75	1	0.75
TE-TE	1	1	1	1	1	1	1	1
DP-DP	0.25	4	0.25	4	0.25	4	0.25	4
Epi-PE	0.25	4	2	1	2	1	2	1
Epi-TE	1	1	1	1	1	1	1	1
Epi-DP	0.25	4	0.25	4	0.25	4	0.25	4
PE-TE	1	1	1	1	6	0.1	6	0.1
PE-DP	0.25	4	2	1	0.25	4	2	1
TE-DP	1	1	1	1	1	1	1	1

	H5	H5	H6	H6	H7	H7
	Rep.	Attr.	Rep.	Attr.	Rep.	Attr.
Epi-Epi	0.25	4	0.266	3.756	1.19	0.84
PE-PE	1	0.75	0.850	1.176	0.765	1.308
TE-TE	1	1	1	1	1	1
DP-DP	1	0.75	0.25	4	0.25	4
Epi-PE	2	1	0.474	2.108	0.614	1.628
Epi-TE	1	1	1	3	1	3
Epi-DP	2	1	0.209	4.788	0.538	1.86
PE-TE	6	0.1	6	0.1	6	0.1
PE-DP	1	0.75	0.384	2.6	0.304	3.288
TE-DP	6	0.1	1	1	1	1

Table 3.3: Parameters for selective adhesion models. H1: no SA; H2: Symmetric SA; H3: Non-biased asymmetric SA; H4: Asymmetric SA with DP behavior biased to Gata6+; H5: Asymmetric SA with DP behavior biased to Nanog+; H6: EphrinB2/EphA4 induced SA; H7: EphrinB2/EphB2 induced SA.

Bibliography

[1] E. Anitua, R. Prado, M. Azkargorta, E. Rodriguez-Suárez, I. Iloro, J. Casado-Vela, F. Elortza, and G. Orive. High-throughput proteomic characterization of plasma rich in growth factors (prgf–endoret)-derived fibrin clot interactome. *Journal of tissue engineering and regenerative medicine*, 9(11):E1–E12, 2015.

[2] S. Barrientos, O. Stojadinovic, M. S. Golinko, H. Brem, and M. Tomic-Canic. Growth factors and cytokines in wound healing. *Wound repair and regeneration*, 16(5):585–601, 2008.

[3] S. Bessonnard, S. Coqueran, S. Vandormael-Pournin, A. Dufour, J. Artus, and M. Cohen-Tannoudji. Icm conversion to epiblast by fgf/erk inhibition is limited in time and requires transcription and protein degradation. *Scientific reports*, 7(1):1–12, 2017.

[4] S. Bessonnard, L. De Mot, D. Gonze, M. Barriol, C. Dennis, A. Goldbeter, G. Dupont, and C. Chazaud. Gata6, nanog and erk signaling control cell fate in the inner cell mass through a tristable regulatory network. *Development*, 141(19):3637–3648, 2014.

[5] M. Bischoff, D.-E. Parfitt, and M. Zernicka-Goetz. Formation of the embryonic-abembryonic axis of the mouse blastocyst: relationships between orientation of early cleavage divisions and pattern of symmetric/asymmetric divisions. *Development*, 135(5):953–962, 2008.

[6] J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, 566(7745):496–502, 2019.

[7] Y. H. Chan, J. Intosalmi, S. Rautio, and H. Lähdesmäki. A subpopulation model to analyze heterogeneous cell differentiation dynamics. *Bioinformatics*, 32(21):3306–3313, 2016.

[8] C. Chazaud, Y. Yamanaka, T. Pawson, and J. Rossant. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the grb2-mapk pathway. *Developmental cell*, 10(5):615–624, 2006.

[9] S. Cheng, Y. Pei, L. He, G. Peng, B. Reinius, P. P. Tam, N. Jing, and Q. Deng. Single-cell rna-seq reveals cellular heterogeneity of pluripotency transition and x chromosome dynamics during early mouse development. *Cell reports*, 26(10):2593–2607, 2019.

[10] V. Chickarmane and C. Peterson. A computational model for understanding stem cell, trophectoderm and endoderm lineage determination. *PLoS one*, 3(10):e3478, 2008.

[11] V. Chickarmane, C. Troein, U. A. Nuber, H. M. Sauro, and C. Peterson. Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput Biol*, 2(9):e123, 2006.

[12] S. Christley, B. Lee, X. Dai, and Q. Nie. Integrative multicellular biological modeling: a case study of 3d epidermal development using gpu algorithms. *BMC systems biology*, 4(1):107, 2010.

[13] B. D. Cumming, D. McElwain, and Z. Upton. A mathematical model of wound healing and subsequent scarring. *Journal of The Royal Society Interface*, 7(42):19–34, 2010.

[14] J. T. Daub and R. M. Merks. A cell-based model of extracellular-matrix-guided endothelial cell migration during angiogenesis. *Bulletin of mathematical biology*, 75(8):1377–1399, 2013.

[15] L. De Mot, D. Gonze, S. Bessonnard, C. Chazaud, A. Goldbeter, and G. Dupont. Cell fate specification based on tristability in the inner cell mass of mouse blastocysts. *Biophysical journal*, 110(3):710–722, 2016.

[16] R. R. Driskell, B. M. Lichtenberger, E. Hoste, K. Kretzschmar, B. D. Simons, M. Charalambous, S. R. Ferron, Y. Herault, G. Pavlovic, A. C. Ferguson-Smith, et al. Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature*, 504(7479):277–281, 2013.

[17] H. Du, Y. Wang, D. Haensel, B. Lee, X. Dai, and Q. Nie. Multiscale modeling of layer formation in epidermis. *PLoS computational biology*, 14(2):e1006006, 2018.

[18] S. A. Eming, T. Krieg, and J. M. Davidson. Inflammation in wound repair: molecular and cellular mechanisms. *Journal of Investigative Dermatology*, 127(3):514–525, 2007.

[19] S. A. Eming, P. Martin, and M. Tomic-Canic. Wound repair and regeneration: mechanisms, signaling, and translation. *Science translational medicine*, 6(265):265sr6–265sr6, 2014.

[20] C.-H. L. Eng, M. Lawson, Q. Zhu, R. Dries, N. Koulena, Y. Takei, J. Yun, C. Cronin, C. Karp, G.-C. Yuan, et al. Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. *Nature*, 568(7751):235–239, 2019.

[21] B. L. Eppley, J. E. Woodell, and J. Higgins. Platelet quantification and growth factor analysis from platelet-rich plasma: implications for wound healing. *Plastic and reconstructive surgery*, 114(6):1502–1508, 2004.

[22] V. Falanga, L. Zhou, and T. Yufit. Low oxygen tension stimulates collagen synthesis and col1a1 transcription through the action of tgf-β1. *Journal of cellular physiology*, 191(1):42–50, 2002.

[23] M. W. Ferguson and S. O'Kane. Scar–free healing: from embryonic mechanisms to adult therapeutic intervention. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1445):839–850, 2004.

[24] K. W. Finnson, S. McLean, G. M. Di Guglielmo, and A. Philip. Dynamics of transforming growth factor beta signaling in wound healing and scarring. *Advances in wound care*, 2(5):195–214, 2013.

[25] R. A. Foty and M. S. Steinberg. The differential adhesion hypothesis: a direct evaluation. *Developmental biology*, 278(1):255–263, 2005.

[26] S. Frankenberg, F. Gerbe, S. Bessonnard, C. Belville, P. Pouchin, O. Bardot, and C. Chazaud. Primitive endoderm differentiates via a three-step mechanism involving nanog and rtk signaling. *Developmental cell*, 21(6):1005–1013, 2011.

[27] W. L. Garner. Epidermal regulation of dermal fibroblast activity. *Plastic and reconstructive surgery*, 102(1):135–139, 1998.

[28] D. Gay, O. Kwon, Z. Zhang, M. Spata, M. V. Plikus, P. D. Holler, M. Ito, Z. Yang, E. Treffeisen, C. D. Kim, et al. Fgf9 from dermal γ δ t cells induces hair follicle neogenesis after wounding. *Nature medicine*, 19(7):916–923, 2013.

[29] J. A. Glazier, Y. Zhang, M. Swat, B. Zaitlen, and S. Schnell. Coordinated action of n-cam, n-cadherin, epha4, and ephrinb2 translates genetic prepatterns into structure during somitogenesis in chick. *Current topics in developmental biology*, 81:205–247, 2008.

[30] A. Gord, W. R. Holmes, X. Dai, and Q. Nie. Computational modelling of epidermal stratification highlights the importance of asymmetric cell division for predictable and robust layer formation. *Journal of the Royal Society Interface*, 11(99):20140631, 2014.

[31] N. S. Greaves, K. J. Ashcroft, M. Baguneid, and A. Bayat. Current understanding of molecular and cellular mechanisms in fibroplasia and angiogenesis during acute wound healing. *Journal of dermatological science*, 72(3):206–217, 2013.

[32] C. F. Guerrero-Juarez, P. H. Dedhia, S. Jin, R. Ruiz-Vega, D. Ma, Y. Liu, K. Yamaga, O. Shestova, D. L. Gay, Z. Yang, et al. Single-cell analysis reveals fibroblast heterogeneity and myeloid-derived adipocyte progenitors in murine skin wounds. *Nature communications*, 10(1):1–17, 2019.

[33] G. Guo, M. Huss, G. Q. Tong, C. Wang, L. L. Sun, N. D. Clarke, and P. Robson. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Developmental cell*, 18(4):675–685, 2010.

[34] G. Guo, M. Huss, G. Q. Tong, C. Wang, L. L. Sun, N. D. Clarke, and P. Robson. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Developmental cell*, 18(4):675–685, 2010.

[35] G. C. Gurtner, S. Werner, Y. Barrandon, and M. T. Longaker. Wound repair and regeneration. *Nature*, 453(7193):314–321, 2008.

[36] S. D. Hester, J. M. Belmonte, J. S. Gens, S. G. Clendenon, and J. A. Glazier. A multi-cell, multi-scale model of vertebrate segmentation and somite formation. *PLoS Comput Biol*, 7(10):e1002155, 2011.

[37] T. Hillen and K. J. Painter. A user's guide to pde models for chemotaxis. *Journal of mathematical biology*, 58(1-2):183, 2009.

[38] J. R. HINSHAW and E. R. MILLER. Histology of healing split-thickness, full-thickness autogenous skin grafts and donor sites. *Archives of Surgery*, 91(4):658–670, 1965.

[39] T. Hiragun, E. Morita, T. Tanaka, Y. Kameyoshi, and S. Yamamoto. A fibrogenic cytokine, platelet-derived growth factor (pdgf), enhances mast cell growth indirectly via a scf-and fibroblast-dependent pathway. *Journal of investigative dermatology*, 111(2):213–217, 1998.

[40] W. R. Holmes, N. S. R. de Mochel, Q. Wang, H. Du, T. Peng, M. Chiang, O. Cinquin, K. Cho, and Q. Nie. Gene expression noise enhances robust organization of the early mammalian blastocyst. *PLoS computational biology*, 13(1):e1005320, 2017.

[41] H. Honda, N. Motosugi, T. Nagai, M. Tanemura, and T. Hiiragi. Computer simulation of emerging asymmetry in the mouse blastocyst. *Development*, 135(8):1407–1414, 2008.

[42] R. Hosokawa, K. Nonaka, M. Morifuji, L. Shum, and M. Ohishi. Tgf-β3 decreases type i collagen and scarring after labioplasty. *Journal of dental research*, 82(7):558–564, 2003.

[43] J. Intosalmi, H. Ahlfors, S. Rautio, H. Mannerstöm, Z. J. Chen, R. Lahesmaa, B. Stockinger, and H. Lähdesmäki. Analyzing th17 cell differentiation dynamics using a novel integrative modeling framework for time-course rna sequencing data. *BMC systems biology*, 9(1):81, 2015.

[44] J. Intosalmi, A. C. Scott, M. Hays, N. Flann, O. Yli-Harja, H. Lähdesmäki, A. M. Dudley, and A. Skupin. Data-driven multiscale modeling reveals the role of metabolic coupling for the spatio-temporal growth dynamics of yeast colonies. *BMC molecular and cell biology*, 20(1):1–13, 2019.

[45] M. Ito, Z. Yang, T. Andl, C. Cui, N. Kim, S. E. Millar, and G. Cotsarelis. Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding. *Nature*, 447(7142):316–320, 2007.

[46] T. Iwayama, C. Steele, L. Yao, M. G. Dozmorov, D. Karamichos, J. D. Wren, and L. E. Olson. Pdgfrα signaling drives adipose tissue fibrosis by targeting progenitor cell plasticity. *Genes & development*, 29(11):1106–1119, 2015.

[47] D. Jiang, D. Correa-Gallegos, S. Christ, A. Stefanska, J. Liu, P. Ramesh, V. Rajendran, M. M. De Santis, D. E. Wagner, and Y. Rinkevich. Two succeeding fibroblastic lineages drive dermal development and the transition from regeneration to scarring. *Nature cell biology*, 20(4):422–431, 2018.

[48] M. Kang, A. Piliszek, J. Artus, and A.-K. Hadjantonakis. Fgf4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development*, 140(2):267–279, 2013.

[49] P. Krupinski, V. Chickarmane, and C. Peterson. Simulating the mammalian blastocyst-molecular and mechanical interactions pattern the embryo. *PLoS Comput Biol*, 7(5):e1001128, 2011.

[50] E. W. Kuijk, L. T. van Tol, H. Van de Velde, R. Wubbolts, M. Welling, N. Geijsen, and B. A. Roelen. The roles of fgf and map kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. *Development*, 139(5):871–882, 2012.

[51] S. H. Kwon and G. C. Gurtner. Is early inflammation good or bad? linking early immune changes to hypertrophic scarring. *Experimental dermatology*, 26(2):133–134, 2017.

[52] C. H. Lim, Q. Sun, K. Ratti, S.-H. Lee, Y. Zheng, M. Takeo, W. Lee, P. Rabbani, M. V. Plikus, J. E. Cain, et al. Hedgehog stimulates hair follicle neogenesis by creating inductive dermis during murine skin wound healing. *Nature communications*, 9(1):1–13, 2018.

[53] K.-T. Lin, S. Sloniowski, D. W. Ethell, and I. M. Ethell. Ephrin-b2-induced cleavage of ephb2 receptor is mediated by matrix metalloproteinases to trigger cell repulsion. *Journal of Biological Chemistry*, 283(43):28969–28979, 2008.

[54] Y. Liu, Y. Li, N. Li, W. Teng, M. Wang, Y. Zhang, and Z. Xiao. Tgf-β1 promotes scar fibroblasts proliferation and transdifferentiation via up-regulating microrna-21. *Scientific reports*, 6:32231, 2016.

[55] W.-C. Lo, C.-S. Chou, K. K. Gokoffski, F. Y.-M. Wan, A. D. Lander, A. L. Calof, and Q. Nie. Feedback regulation in multistage cell lineages. *Mathematical biosciences and engineering: MBE*, 6(1):59, 2009.

[56] M. Luecken and F. Theis. Current best practices in single-cell RNA-seq analysis: a tutorial. Molecular systems biology, 2019.

[57] M. D. Luecken and F. J. Theis. Current best practices in single-cell rna-seq analysis: a tutorial. *Molecular systems biology*, 15(6):e8746, 2019.

[58] P. Macklin. Key challenges facing data-driven multicellular systems biology. *GigaScience*, 8(10):giz127, 2019.

[59] J.-L. Maître, H. Turlier, R. Illukkumbura, B. Eismann, R. Niwayama, F. Nédélec, and T. Hiiragi. Asymmetric division of contractile domains couples cell positioning and fate specification. *Nature*, 536(7616):344–348, 2016.

[60] C. D. Marshall, M. S. Hu, T. Leavitt, L. A. Barnes, H. P. Lorenz, and M. T. Longaker. Cutaneous scarring: basic science, current treatments, and future directions. *Advances in wound care*, 7(2):29–45, 2018.

[61] S. McDougall, J. Dallon, J. Sherratt, and P. Maini. Fibroblast migration and collagen deposition during dermal wound healing: mathematical modelling and clinical implications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1843):1385–1405, 2006.

[62] S. N. Menon, J. A. Flegg, S. W. McCue, R. C. Schugart, R. A. Dawson, and D. S. McElwain. Modelling the interaction of keratinocytes and fibroblasts during normal and abnormal wound healing processes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741):3329–3338, 2012.

[63] S. N. Menon, J. A. Flegg, S. W. McCue, R. C. Schugart, R. A. Dawson, and D. S. McElwain. Modelling the interaction of keratinocytes and fibroblasts during normal and abnormal wound healing processes. *Proceedings of the Royal Society B: Biological Sciences*, 279(1741):3329–3338, 2012.

[64] K. Mitsui, Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami, K. Takahashi, M. Maruyama, M. Maeda, and S. Yamanaka. The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and es cells. *Cell*, 113(5):631–642, 2003.

[65] H. Mohammed, I. Hernando-Herraez, A. Savino, A. Scialdone, I. Macaulay, C. Mulas, T. Chandra, T. Voet, W. Dean, J. Nichols, et al. Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. *Cell reports*, 20(5):1215–1228, 2017.

[66] H. Murata, L. Zhou, S. Ochoa, A. Hasan, E. Badiavas, and V. Falanga. Tgf-β3 stimulates and regulates collagen synthesis through tgf-β1-dependent and independent mechanisms. *Journal of Investigative Dermatology*, 108(3):258–262, 1997.

[67] A. M. Nelson, S. K. Reddy, T. S. Ratliff, M. Z. Hossain, A. S. Katseff, A. S. Zhu, E. Chang, S. R. Resnik, C. Page, D. Kim, et al. dsrna released by tissue damage activates tlr3 to drive skin regeneration. *Cell stem cell*, 17(2):139–151, 2015.

[68] A. Nematbakhsh, M. Levis, N. Kumar, W. Chen, J. J. Zartman, and M. Alber. Epithelial organ shape is generated by patterned actomyosin contractility and maintained by the extracellular matrix. *bioRxiv*, 2020.

[69] T. J. Newman. Modeling multi-cellular systems using sub-cellular elements. *arXiv preprint q-bio/0504028*, 2005.

[70] S. B. Nissen, M. Perera, J. M. Gonzalez, S. M. Morgani, M. H. Jensen, K. Sneppen, J. M. Brickman, and A. Trusina. Four simple rules that are sufficient to generate the mammalian blastocyst. *PLoS biology*, 15(7):e2000737, 2017.

[71] S. Nowotschin, M. Setty, Y.-Y. Kuo, V. Liu, V. Garg, R. Sharma, C. S. Simon, N. Saiz, R. Gardner, S. C. Boutet, et al. The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature*, 569(7756):361–367, 2019.

[72] A. T. Nurden, P. Nurden, M. Sanchez, I. Andia, and E. Anitua. Platelets and wound healing. *Frontiers in bioscience: a journal and virtual library*, 13:3532, 2008.

[73] S. O'Kane and M. W. Ferguson. Transforming growth factor βs and wound healing. *The international journal of biochemistry & cell biology*, 29(1):63–78, 1997.

[74] E. Oron and N. Ivanova. Cell fate regulation in early mammalian development. *Physical biology*, 9(4):045002, 2012.

[75] S. Osher and R. P. Fedkiw. Level set methods: an overview and some recent results. *Journal of Computational physics*, 169(2):463–502, 2001.

[76] J. Ovadia and Q. Nie. Numerical methods for two-dimensional stem cell tissue growth. *Journal of scientific computing*, 58(1):149–175, 2014.

[77] I. Pastar, O. Stojadinovic, N. C. Yin, H. Ramirez, A. G. Nusbaum, A. Sawaya, S. B. Patel, L. Khalid, R. R. Isseroff, and M. Tomic-Canic. Epithelialization in wound healing: a comprehensive review. *Advances in wound care*, 3(7):445–464, 2014.

[78] C. Philippeos, S. B. Telerman, B. Oulès, A. O. Pisco, T. J. Shaw, R. Elgueta, G. Lombardi, R. R. Driskell, M. Soldin, M. D. Lynch, et al. Spatial and single-cell transcriptional profiling identifies functionally distinct human dermal fibroblast subpopulations. *Journal of Investigative Dermatology*, 138(4):811–825, 2018.

[79] B. Pijuan-Sala, J. A. Griffiths, C. Guibentif, T. W. Hiscock, W. Jawaid, F. J. Calero-Nieto, C. Mulas, X. Ibarra-Soria, R. C. Tyser, D. L. L. Ho, et al. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*, 566(7745):490–495, 2019.

[80] M. V. Plikus, C. F. Guerrero-Juarez, M. Ito, Y. R. Li, P. H. Dedhia, Y. Zheng, M. Shao, D. L. Gay, R. Ramos, T.-C. Hsi, et al. Regeneration of fat cells from myofibroblasts during wound healing. *Science*, 355(6326):748–752, 2017.

[81] B. Plusa, A. Piliszek, S. Frankenberg, J. Artus, and A.-K. Hadjantonakis. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development*, 135(18):3081–3091, 2008.

[82] D. M. Poitz, G. Ende, B. Stütz, A. Augstein, J. Friedrichs, C. Brunssen, C. Werner, R. H. Strasser, and S. Jellinghaus. Ephrinb2/epha4-mediated activation of endothelial cells increases monocyte adhesion. *Molecular immunology*, 68(2):648–656, 2015.

[83] A. Poliakov, M. Cotrina, and D. G. Wilkinson. Diverse roles of eph receptors and ephrins in the regulation of cell migration and tissue assembly. *Developmental cell*, 7(4):465–480, 2004.

[84] C. S. Potten and M. Loeffler. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. lessons for and from the crypt. *Development*, 110(4):1001–1020, 1990.

[85] W. M. Reichert. *Indwelling neural implants: strategies for contending with the in vivo environment*. CRC Press, 2007.

[86] L. Richardson, S. Venkataraman, P. Stevenson, Y. Yang, N. Burton, J. Rao, M. Fisher, R. A. Baldock, D. R. Davidson, and J. H. Christiansen. Emage mouse embryo spatial gene expression database: 2010 update. *Nucleic acids research*, 38(suppl_1):D703–D709, 2010.

[87] Y. Rinkevich, G. Walmsley, M. Hu, Z. Maan, A. Newman, M. Drukker, M. Januszyk, G. Krampitz, G. Gurtner, H. Lorenz, et al. Skin fibrosis. identification and isolation of a dermal lineage with intrinsic fibrogenic potential. science 348, aaa2151. *This study identifies a dermal lineage with high fibrogenic potential. Ablation of these CD26+ cells attenuates cutaneous fibrosis*, 2015.

[88] A. B. Roberts, M. B. Sporn, R. K. Assoian, J. M. Smith, N. S. Roche, L. M. Wakefield, U. I. Heine, L. A. Liotta, V. Falanga, and J. H. Kehrl. Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. *Proceedings of the National Academy of Sciences*, 83(12):4167–4171, 1986.

[89] S. G. Rodriques, R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, C. R. Vanderburg, J. Welch, L. M. Chen, F. Chen, and E. Z. Macosko. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434):1463–1467, 2019.

[90] E. Rognoni, A. O. Pisco, T. Hiratsuka, K. H. Sipilä, J. M. Belmonte, S. A. Mobasseri, C. Philippeos, R. Dilão, and F. M. Watt. Fibroblast state switching orchestrates dermal maturation and wound healing. *Molecular systems biology*, 14(8):e8174, 2018.

[91] J. Rossant, C. Chazaud, and Y. Yamanaka. Lineage allocation and asymmetries in the early mouse embryo. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1436):1341–1349, 2003.

[92] N. Saiz, K. M. Williams, V. E. Seshan, and A.-K. Hadjantonakis. Asynchronous fate decisions by single cells collectively ensure consistent lineage composition in the mouse blastocyst. *Nature communications*, 7(1):1–14, 2016.

[93] N. Schrode, N. Saiz, S. Di Talia, and A.-K. Hadjantonakis. Gata6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst. *Developmental cell*, 29(4):454–467, 2014.

[94] J. A. Sherratt and J. D. Murray. Models of epidermal wound healing. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 241(1300):29–36, 1990.

[95] B. A. Shook, R. R. Wasko, G. C. Rivera-Gonzalez, E. Salazar-Gatzimas, F. López-Giráldez, B. C. Dash, A. R. Muñoz-Rojas, K. D. Aultman, R. K. Zwick, V. Lei, et al. Myofibroblast proliferation and heterogeneity are supported by macrophages during skin repair. *Science*, 362(6417), 2018.

[96] A. Singh, E. Winterbottom, and I. O. Daar. Eph/ephrin signaling in cell-cell and cell-substrate adhesion. *Frontiers in bioscience (Landmark edition)*, 17:473, 2012.

[97] A. Skupin, H. Kettenmann, and M. Falcke. Calcium signals driven by single channel noise. *PLoS Comput Biol*, 6(8):e1000870, 2010.

[98] A. Stachowicz, J. Siudut, M. Suski, R. Olszanecki, R. Korbut, A. Undas, and J. R. Wiśniewski. Optimization of quantitative proteomic analysis of clots generated from plasma of patients with venous thromboembolism. *Clinical Proteomics*, 14(1):38, 2017.

[99] P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294):78–82, 2016.

[100] B. K. Sun, Z. Siprashvili, and P. A. Khavari. Advances in skin grafting and treatment of cutaneous wounds. *Science*, 346(6212):941–945, 2014.

[101] J. J. Tomasek, G. Gabbiani, B. Hinz, C. Chaponnier, and R. A. Brown. Myofibroblasts and mechano-regulation of connective tissue remodelling. *Nature reviews Molecular cell biology*, 3(5):349–363, 2002.

[102] A. Tosenberger, D. Gonze, S. Bessonnard, M. Cohen-Tannoudji, C. Chazaud, and G. Dupont. A multiscale model of early cell lineage specification including cell division. *NPJ systems biology and applications*, 3(1):1–11, 2017.

[103] T. Velnar, T. Bailey, and V. Smrkolj. The wound healing process: an overview of the cellular and molecular mechanisms. *Journal of International Medical Research*, 37(5):1528–1542, 2009.

[104] F. Vermolen and E. Javierre. A finite-element model for healing of cutaneous wounds combining contraction, angiogenesis and closure. *Journal of mathematical biology*, 65(5):967–996, 2012.

[105] H. Wang and S. K. Dey. Roadmap to embryo implantation: clues from mouse models. *Nature Reviews Genetics*, 7(3):185–199, 2006.

[106] K.-Y. Wang, S. Yamada, H. Izumi, M. Tsukamoto, T. Nakashima, T. Tasaki, X. Guo, H. Uramoto, Y. Sasaguri, and K. Kohno. Critical in vivo roles of wnt10a in wound healing by regulating collagen expression/synthesis in wnt10a-deficient mice. *PLoS One*, 13(3):e0195156, 2018.

[107] Q. Wang, W. R. Holmes, J. Sosnik, T. Schilling, and Q. Nie. Cell sorting and noise-induced cell plasticity coordinate to sharpen boundaries between gene expression domains. *PLoS computational biology*, 13(1):e1005307, 2017.

[108] Y. Wang, C. F. Guerrero-Juarez, Y. Qiu, H. Du, W. Chen, S. Figueroa, M. V. Plikus, and Q. Nie. A multiscale hybrid mathematical model of epidermal-dermal interactions during skin wound healing. *Experimental dermatology*, 28(4):493–502, 2019.

[109] Z. Wang, Y. Wang, F. Farhangfar, M. Zimmer, and Y. Zhang. Enhanced keratinocyte proliferation and migration in co-culture with fibroblasts. *PloS one*, 7(7):e40951, 2012.

[110] J. Wei, D. Melichian, K. Komura, M. Hinchcliff, A. P. Lam, R. Lafyatis, C. J. Gottardi, O. A. MacDougald, and J. Varga. Canonical wnt signaling induces skin fibrosis and subcutaneous lipoatrophy: a novel mouse model for scleroderma? *Arthritis & Rheumatism*, 63(6):1707–1717, 2011.

[111] S. Werner, T. Krieg, and H. Smola. Keratinocyte–fibroblast interactions in wound healing. *Journal of Investigative Dermatology*, 127(5):998–1008, 2007.

[112] D. Whitby and M. Ferguson. The extracellular matrix of lip wounds in fetal, neonatal and adult mice. *Development*, 112(2):651–668, 1991.

[113] M. Xue and C. J. Jackson. Extracellular matrix reorganization during wound healing and its impact on abnormal scarring. *Advances in wound care*, 4(3):119–136, 2015.

[114] Y. Yamanaka, F. Lanner, and J. Rossant. Fgf signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development*, 137(5):715–724, 2010.