UCLouvain

Mathematical modelling of epithelial growth, fission and lumen formation during embryonic thyroid development:

a combination of computational and experimental approaches

Doctoral dissertation presented by

Léolo Gonay

In partial fulfilment of the requirements for the degree of
Docteur en Sciences Agronomiques et Ingénierie Biologique

Members of the thesis committee :

Prof. Emmanuel Hanert (UCLouvain, BE), Supervisor
Prof. Christophe Pierreux (UCLouvain, BE), Supervisor
Prof. Marnik Vanclooster (UCLouvain, BE), Chairperson
Prof. René Rezsohazy (UCLouvain, BE), Secretary
Prof. Xavier Draye (UCLouvain, BE)
Prof. Frédéric Lemaigre (UCLouvain, BE)
Dr. Paul Van Liedekerke (INRIA, Paris, FR)
Dr. Alexander Fletcher (University of Sheffield, UK)

May 2021

Préface

Septembre 2016, année du début de mon parcours doctoral, l'UCLouvain met à l'honneur la profession de "saventurier" dans son année thématique consacrée à l'aventure scientifique. Quatre ans (et quelques mois) plus tard, en avril 2021, j'écris les dernières lignes de cette thèse, sorte de journal de bord de mon épopée scientifique. Mais si c'est mon nom qui est indiqué en première page de cet ouvrage, il serait à la fois naïf et illusoire de penser que j'ai effectué seul cette longue traversée en eaux (presque) inconnues. J'aimerais profiter de ces quelques pages pour remercier toutes les personnes qui, de près ou de loin, ont permis à ce navire d'arriver à bon port.

Je voudrais tout d'abord remercier mon papa et ma maman, qui, depuis mes premiers pas et mes premiers mots ("la lune", paraît-il) m'ont toujours soutenu, poussé quand il le fallait et dorloté quand j'en avais besoin, et ce, depuis maintenant 24 ans derrière (et un peu devant) les bancs de l'école. Mais ce n'était pas pour rien, vous avez maintenant un livre de plus dans votre bibliothèque !

Je tiens également à remercier mes promoteurs, Emmanuel et Christophe. Emmanuel qui sans le vouloir (ni le savoir, d'ailleurs), a été le catalyseur de mon envie de me lancer dans ce fameux "master Matlab", et finalement dans cette thèse, dès ma première année de BAC. S'il y a un cours que je ne manquais pas, c'était bien celui du Professeur Hanert ! Christophe, lui, est arrivé plus tard, après mon mémoire, mais a finalement accepté lui aussi de m'accueillir dans son laboratoire. Moi qui passais mon temps sur mon ordinateur à faire des choses qu'il ne comprenait pas, lui qui devait se demander ce que je faisais quand je n'étais pas là. Merci à vous deux pour votre confiance et votre encadrement précieux. On nous avait dit un jour qu'il était plus important d'avoir un bon promoteur qu'un bon sujet. J'en ai eu deux, et je ne pouvais pas rêver mieux !

Évidemment, je ne pourrais pas continuer sans remercier mes deux autres mamans, celles qui sont venues me chercher un jour au fond du bureau des mémorants, alors que je pensais encore pouvoir passer ces 4 ans en solitaire, seul et bien tranquille dans ma petite barque... Je ne vous remercierai finalement jamais assez d'être venu me sortir de force et de m'avoir installé à vos côtés. Pour ça, Charly, Spourky : "mercy". Merci de m'avoir permis d'être cette quatrième roue d'un carrosse qui roulait pourtant déjà si bien, au côté de ce cher Jona chez qui je pouvais trouver le soutien masculin dont j'avais parfois besoin. Jona, dont le départ, si émouvant fût-il, a été comblé par l'arrivée d'Ophé, pour qui on aurait presque besoin d'un chapitre entier ! Un énorme merci à vous 4, et à tous mes collègues, de Woluwé comme de Louvain-la-Neuve, qui m'ont fait me sentir chez moi chaque minute passée dans ces couloirs, labos, bibliothèques et autres cafèts, ainsi que pour les discussions stimulantes un peu plus sérieuses qu'on a pu avoir (eh oui, on est quand même là aussi pour travailler). Merci à Claude pour son soutien et son attention depuis le premier jour de mon mémoire et

jusqu'après même son dernier jour dans les laboratoires de Frédéric. Merci au marché de Noël.

Merci à tous mes amis, au WeKT, aux bwau's, aux swagmen, à Minou et Lilou, et à tous les autres, qui se demandent ce que je fais depuis 4 ans (peut-être un peu moins maintenant ?) mais qui ont toujours été là pour m'aider à garder le cap quand ça tanguait un peu plus. Mention spéciale pour Manon et Adri, avec qui j'ai cohabité pendant 3 ans et sans qui je serais probablement mort soit d'overdose de Nalu, soit de fatigue, soit de faim, soit d'ennui. Vous êtes les meilleurs colocs du monde, et je n'ai pas besoin de 4 ans de recherches pour le prouver. Merci à Charlotte et Hugo d'avoir été ma bouée de sauvetage et ma bonbonne d'oxygène. Un grand "thank you" à Marie pour son aide lors de la relecture.

A huge thanks to all the Chaste team for their availability and the help that would (almost) always get me out of the mud I would have stuck myself into. A special thank you to Alex, who must have received at least twice as many e-mails as my two supervisors combined. Thanks, Alex, it would have been virtually impossible without you.

Merci aux membres de la faculté pour leur aide précieuse (parce que oui une thèse, c'est aussi beaucoup d'administratif), merci à Matthieu et Ludovic pour leur excellent travail, merci à tous ceux dont j'ai pu croiser la route le temps d'une discussion, d'une tasse de thé, d'un verre ou d'une course, et qui ont rendu la mienne un peu plus navigable. Merci à tous les membres de mon jury pour le temps et l'attention mis à la révision de ce travail.

Ce travail a d'ailleurs été réalisé au sein de la faculté des Bioingénieurs et j'aimerais pour clôturer également mentionner quelques figures de la science moderne, sources d'inspiration un peu plus lointaines, sans qui le terme "bioingénieur" n'existerait probablement pas. On a coutume d'ailleurs de dire que le bioingénieur "n'excelle en rien" mais est "bon en tout". Parfait, s'il s'agit de mélanger l'expérimental avec le théorique (ou en d'autres termes de présenter une équation sur la même slide qu'une immuno), est-il une meilleure porte d'entrée ? Bien sûr, je suis loin d'être le premier à mêler physique et biologie. Déjà en 1917, D'Arcy Wentworth Thompson aborde les questions de l'origine de la forme des plantes et des animaux d'un point de vue théorique (oserais-je dire mathématique) dans son livre "On Growth and Form" [1]. Alan Turing, après avoir mis son génie au service de la paix durant la Seconde Guerre mondiale, publie un article intitulé "The Chemical Basis of Morphogenesis" [2], présentant un des premiers modèles théoriques décrivant les fondements chimiques de l'acquisition des formes et structures particulières de certains organes (ou "Turing patterns"). En 1944, le célèbre physicien Erwin Schrödinger (plus connu, il est vrai, pour son chat que pour ses travaux en biologie) publie lui un essai de vulgarisation scientifique, basé sur une série de conférences données au Trinity College de Dublin, et modestement intitulé "What is life ?" [3]. N'est-ce pas finalement la question que nous nous posons tous ? Évidemment, les chapitres qui suivent n'ont pas la prétention de répondre à cette question, mais s'inscrivent, d'une certaine façon, dans cette grande, fascinante et stimulante aventure scientifique.

"Mieux comprendre pour mieux guérir"
(Christian de Duve, 1917 - 2013)

Contents

Préface **III**

Abstract **1**

List of abbreviations **3**

1 Introduction **7**

2 Biological Framework **9**
2.1 Anatomy and histology of the thyroid 9
2.2 Embryonic development 12
2.2.1 Extrinsic factors regulating thyroid morphogenesis 20

3 Modelling Preliminaries **27**
3.1 Cell-based models 27
3.2 Vertex models of epithelia 39

4 Aim of the work **47**

5 Material and Methods **49**

6 Epithelial Proliferation and Initial Conditions **51**
6.1 Introduction 51
6.2 Methods 51
6.3 Results 53

7 VEGF Diffusion and Endothelial Motility **61**
7.1 Introduction 61
7.2 Methods 61
7.3 Results 64

8 Endothelial Boundary Conditions and Density **69**
8.1 Introduction 69
8.2 Methods 69
8.3 Calibration 72
8.4 Results 74

9 Cell polarisation and Lumen Generation **77**
9.1 Introduction 77
9.2 Methods 78

9.3 Calibration 81
9.4 Results 84
9.4.1 Epithelial polarisation 84
9.4.2 Lumen generation 85

10 Model Validation **89**
10.1 Epithelial cell surface area distribution 89
10.2 Lumen surface area distribution 91
10.3 Epithelial cells to lumen surface area ratio 92

11 Discussion **95**
11.1 Vertex-based model and Initial conditions 95
11.2 Epithelial cell proliferation 96
11.3 Epithelial adhesion and cohesion 97
11.4 Morphogen diffusion and endothelial density 103
11.5 Computational instabilities 106

12 Conclusion and Perspectives **109**

13 Appendix **113**
Tables 113
Figures 115

14 Publications **115**
Modelling of epithelial growth, fission and lumen formation during embryonic thyroid development : a combination of computational and experimental approaches 115
Dynamics and predicted drug response of a gene network linking dedifferentiation with beta-catenin dysfunction in hepatocellular carcinoma 117

Bibliography **142**

Abstract

Organogenesis is the phase of embryonic development leading to the formation of fully functional organs. It can be divided into two main processes: morphogenesis (acquisition of a shape and cellular organisation) and differentiation (acquisition of cellular function). The thyroid, a highly vascularised endocrine gland, derives from the endoderm and develops into its typical architecture (a multitude of tightly packed independent spherical follicular units) starting from a 3D mass of epithelial thyroid progenitors. Follicular organisation is unique and essential for thyroid function, *i.e.* thyroid hormone (T3 and T4) production. Previous *in vivo* studies showed that endothelial cells , besides their nutritive function, play a central role in thyroid gland morphogenesis. However, the precise mechanisms and biological parameters controlling the organisation of the thyroid epithelial progenitors into a multitude of single-layered thyroid follicles remain unknown. Animal experimentation to improve our knowledge of organogenesis is time- and money-consuming and has clear limitations. Here, we developed and used a 2-D vertex model of thyroid growth, angiogenesis and folliculogenesis, within the open-source Chaste framework. Our in *silico* model, based on initial *in vivo* images, correctly simulates the differential growth and proliferation of (central and peripheral) epithelial progenitors, as well as the morphogen-driven migration of endothelial cells, consistent with our experimental data. We show that endothelial invasion, associated with changes in epithelial cell adhesion, are key in driving the fission of the 3D epithelial mass into independent pre-follicular units. We also found that endothelial cells abundance and proximity to epithelial cells could induce their polarisation and subsequent opening of a cavity in the centre of the follicle. Our study illustrates how constant dialogue between theoretical and experimental approaches helped us better understand the role of cell migration, adhesion and polarisation in the embryonic development of the thyroid. We anticipate that the use of *in silico* models such as the one developed here can push forward the fields of developmental, regenerative or even medical biology.

List of abbreviations

Δt	Simulation time-step
ATC	Anaplastic Thyroid Cancer
BC	Boundary Conditions
BrdU	5-bromo-2'-deoxyuridine
CA	Cellular Automata
CAM	Cell Adhesion Molecules
CB	Center-Based models
CBM	Cell-Based Models
CHASTE	Cancer, Heart and Soft Tissue Environment
CPM	Cellular Potts Models
DCM	Deformable Cell Models
DNA	Deoxyribonucleic Acid
E	Embryonic day
EC	Endothelial Cells
ECM	Extracellular Matrix
EdU	5-Ethynyl-2´-deoxyuridine
FTC	Follicular Thyroid Cancer
IC	Initial Conditions
LGCA	Lattice Gas Cellular Automata
MCS	Monte Carlo Step
MDCK	Madin-Darby Canine Kidney

MTC	Medullary Thyroid Cancer
NIS	Sodium-Iodine Dymporter
OS	Overlapping-Spheres
PDE	Partial Differential Equation
PTC	Papillary Thyroid Cancer
RNA	Ribonucleic Acid
SEM	Subcellular Element Models
t	Simulation time
T3	Triiodothyronine
T4	Thyroxin
TG	Thyroglobulin
TPO	Thyroperoxidase
TSH	Thyroid-Stimulating Hormone
TSI	Thyroid-Stimulating Immunoglobulin
UBB	Ultimobranchial Body
VEGFA	Vascular Endothelial Growth Factor A
VM	Vertex-based Model
VT	Voronoi Tesselation

1 | Introduction

As of 2015, more than 3.2 million people were affected by thyroid cancer. The thyroid, a highly vascularised endocrine gland deriving from the endoderm, regulates various hormonal systems by producing triiodothyronine (T3), thyroxine (T4) and calcitonin. These hormones are crucial in human growth and development, particularly in young children. A healthy thyroid develops into its typical architecture, a multitude of tightly packed independent spherical follicular units, during a process called embryogenesis. This follicular organisation is essential for thyroid function. These functioning units are surrounded by a complex network of capillaries which, besides their nutritive function, are known to play a central role in the reorganisation of this epithelial tissue from a multi-layered mass into a multitude of mono-layered structures.

More and more studies are conducted tackling this issue, using the latest experimental techniques available in developmental biology. However, the precise mechanisms and biological parameters controlling this particular morphological reorganisation are still unclear. Moreover, animal experimentation to improve our knowledge of morphogenesis can be time- and money-consuming. There is therefore a growing need in the introduction of new fields, using different approaches, to give a fresh and different view on this kind of issue.

Today, mathematical modelling appears as a powerful complementary tool aiming to unlock, among others, complex biological systems. More and more models, varying in scale and complexity, are being developed in numerous fields of developmental, regenerative or even medicine biology. In this context, we seek to develop a 2D vertex-based model of murine thyroid growth, angiogenesis (formation of blood capillaries) and folliculogenesis (development of the individual thyroid follicles), in the hope that mathematical and computational biology could help to unravel the mysteries behind thyroid morphogenesis. *In silico* models can in fact be of great interest when it comes to testing and confronting various hypotheses regarding the complex processes at stake, without the limitations of *in vitro* experiments. We

therefore aimed to use this model, developed step-by-step at the edge of experimental and computational approaches, to investigate the role of endothelial cells, as well as the effect of epithelial cells cohesion and adhesion, in the formation and development of functional thyroid follicles.

2 | Biological Framework

Human embryogenesis is a long and complex process leading to the development and formation of a multicellular human embryo starting from a single cell. The seventeenth day marks the beginning of gastrulation, a phase during which the single-layered blastula is transformed into a multi-layered structure; the gastrula. These layers are known as the ectoderm (which will eventually give rise to the nervous system and the epidermis), the mesoderm (precursor of the vascular and locomotor systems) and the endoderm. Organs such as the lungs, the liver, the thyroid, the pancreas and the intestine initially develop from this embryonic endodermal layer.

2.1 Anatomy and histology of the thyroid

In humans, as in most mammals, the adult thyroid consists of two lobes (left and right) positioned anteriorly in the lower-neck region, close to the trachea (Fig. 2.1, (A)). The two lobes are connected by a thin layer of median tissue called the isthmus. Humans also usually possess four smaller parathyroid glands at the back of the thyroid gland. Notable differences can be observed regarding thyroid shapes among species. In cartilaginous fishes and some mammals, the thyroid shows as a single mass. In amphibians and birds, two separated lobes form bilateral glands [4].

The functional units of the thyroid are the follicles. These are spherical structures composed of a single layer of epithelial follicular cells called the thyrocytes, resting on a continuous basement membrane and surrounding a central cavity. This cavity, or lumen (plural lumina), is called the colloidal space and contains a homogeneous non-crystalline substance consisting of large iodinated molecules that constitute the thyroid hormone reservoir. This basement membrane that surrounds the follicular cells is a thin, non-cellular matrix that separates the epithelium from the underlying connective tissue. Its composition may vary spatially (between organs) or temporally (between developmental stages). However, it is mainly composed

of collagens[1] fulfilling a structural role and laminins[2] ensuring the adhesion with the epithelial surface. Beneath the basement membrane, the connective tissue contains fibroblasts, blood vessels and different other secreted proteins (that derive from the neural crest [5]). Together with the basement membrane, these proteins form the extra-cellular matrix (ECM). Note that the thyroid also contains other interstitial cells such as macrophages and mast cells [6].

Thyrocytes, the most abundant cell type in the thyroid gland are differentiated, polarised endocrine cells that produce the thyroid hormones, triiodothyronine (T3) and thyroxin (T4). Each thyrocyte possesses an apical surface facing the colloid, a basal surface facing the basement membrane and the connective tissue and a lateral pole interacting with other thyrocytes [9]. These cells are thus fully polarised and form a tight barrier between the colloidal space of the follicle and the extra-follicular environment or connective tissue. The cohesion of the thyrocytes monolayer is ensured by different junctional complexes. Adherens junctions (or *zonula adherens*) are formed by cadherins, mainly E-cadherins in the case of thyrocytes. Cadherins are type-1 transmembrane proteins requiring calcium ions to interact with other cadherins on adjacent cells. Intracellular portions of cadherins are linked to the actin cytoskeleton of the cells [7]. Besides their adhesion role, cadherins also play a key role in establishing and maintaining cell polarity [10, 11]. A second population of epithelial cells, the parafollicular hormone-producing cells, or C-cells, is also present in the thyroid. C-cells are located between the follicular cells and their basement membrane, which means that they have no direct contact with the lumen. They also derive from the endoderm and synthesise and secrete calcitonin, a hormone that regulates calcium level in the bloodstream. Thyroid follicles are surrounded by a dense network of capillaries that ensure a nutritive role (including iodine delivery to the follicles) and the distribution of thyroid hormones and calcitonin in the body (Fig. 2.1, (B)).

The main role of mature thyrocytes is the secretion of thyroid hormones T3 and T4. Thyroid hormones have various roles in organ development and homeostatic control of critical physiological mechanisms [12]. The production of these two essential hormones requires various proteins and occurs in several steps [9]. First, iodine is entrapped in the thyroid by active transport through sodium-iodine symporter (NIS) located at the basolateral membrane of the thyrocytes. Thyroglobulin (TG), which is synthesised and secreted apically into the

[1]Collagens are proteins assembled in a triple helix that further form fibers (type-1 collagen) or networks (type-4 collagen).

[2]Laminins are high-molecular-weight proteins composed of three chains (α,β and γ) forming a cross-shaped and multi-adhesive structure.

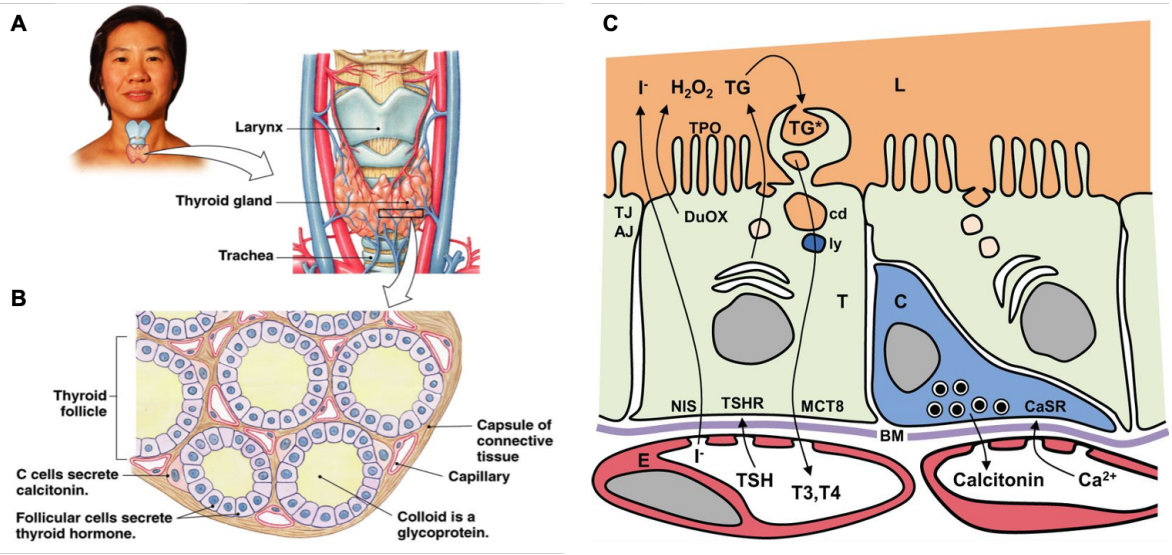


Figure 2.1: Schematic representation of the anatomy and histology of the thyroid gland. (A) The thyroid gland is positioned anteriorly in the lower-neck region, close to the trachea. (B) The thyroid parenchyma consists of follicles formed by single layers of thyrocytes enclosing a lumen filled with colloid. The thyroid follicles are surrounded by blood vessels and connective tissue. Figures adapted from [7]. (C) Principal reactions leading to the formation and secretion of thyroid hormones. Iodide (I^-) is entrapped in the thyroid by active transport, using the sodium-iodine symporter (NIS) expressed in the basolateral membrane of thyrocytes (T, in green). Thyroglobulin (TG) is synthesized and secreted apically into the follicular lumen (L, orange). It is iodinated by thyroid peroxidase (TPO) in the presence of hydrogen peroxide (H_2O_2), generated by an NADPH oxidase (DuOX). The same mechanism is used to couple iodotyrosines into iodothyronines that remain covalently bound (a chemical bond that involves the sharing of electron pairs between atoms, as opposed to ionic bonding) to the TG molecule. Iodinated TG (TG*) is then internalised by endocytosis and degraded *via* the fusion of endosomes with lysosomes (ly), releasing T3 and T4 into the circulation. These reactions are mainly regulated by the thyroid-stimulating hormone (TSH) produced in the pituitary gland, *via* its receptor (TSHR). Thyroid C cells (C, blue) are scattered close to the follicles. They are activated *via* calcium-sensing receptors (CaSR) when extracellular Ca^{2+} concentration is high, thus releasing calcitonin stored in dense-core granules. AJ, adherens junction; TJ, tight junction; cd, colloid droplet; E (red), endothelium (fenestrated capillaries); BM, basement membrane. Figure adapted from [8].

lumen, undergo a series of iodination reactions (addition of iodine to its tyrosine) by thyroid peroxidase (TPO) in the presence of hydrogen peroxyde, eventually leading to the formation of thyroid hormones T3 and T4. They both remain bound to thyroglobulin until the latter complex (TG*) is internalised by endocytosis, and then digested upon fusion with the lysosomes, eventually liberating free T3 and T4 that can diffuse out of the lysosome, through the basolateral plasma membrane of the cell and finally into the bloodstream (Fig. 2.1, (C)).

2.2 Embryonic development

Stages of thyroid morphogenesis

Progenitors of the thyrocytes and C-cells lineage originate from different endodermic structures. Thyrocytes progenitors bud from a midline anlage (or primordium) in the pharyngeal floor while C-cells arise from the fourth lateral pharyngeal pouches [13]. As for its shape, thyroid development and formation processes vary among species. In most mammals, five stages can be distinguished during thyroid morphogenesis: specification, budding, migration, bilobation and folliculogenesis [14]. Note that the morphogenetic stages of human and mouse thyroid development are roughly similar, but with different timing. Around embryonic day (E) 8.5 (in mice, around the 4th week in human), the thyroid placode is formed close to the base of the tongue (specification). At E9.5, this placode converts into a bud (budding) and at E10.5, precursors cells from this bud migrate caudally and detach themselves from the pharyngeal floor. Around E12.5-E13.5 (7th – 8th week in human), the thyroid primordium bifurcates and migrates laterally on both sides of the trachea and merge with the C-cells precursors (know as the ultimobranchial body, or UBB), eventually leading to the formation of the two lobes (bilobation) of the future thyroid gland [8]. The lobes increase in size by rapid proliferation and progressively, thyroid epithelial cells will reorganise and polarise to form the follicles (Fig. 2.3, B). Cell labelling has revealed that the proliferation rate of thyroid precursor cells vary spatiotemporally throughout all stages of thyroid development [15]. It has been shown that the initial growth of the thyroid placode (*i.e.* specification and budding) likely occurs by cell recruitment from outside its own specified domain. In fact, Carré and colleagues [16] suggested that cell proliferation was indeed repressed to enhance this incoming flux of progenitors cells. In fact, they showed that thyroid progenitors expressed cyclin-dependent kinase (protein kinases playing a role in regulating the cell cycle) inhibitors when present in the placode but not during the next development stages. Later

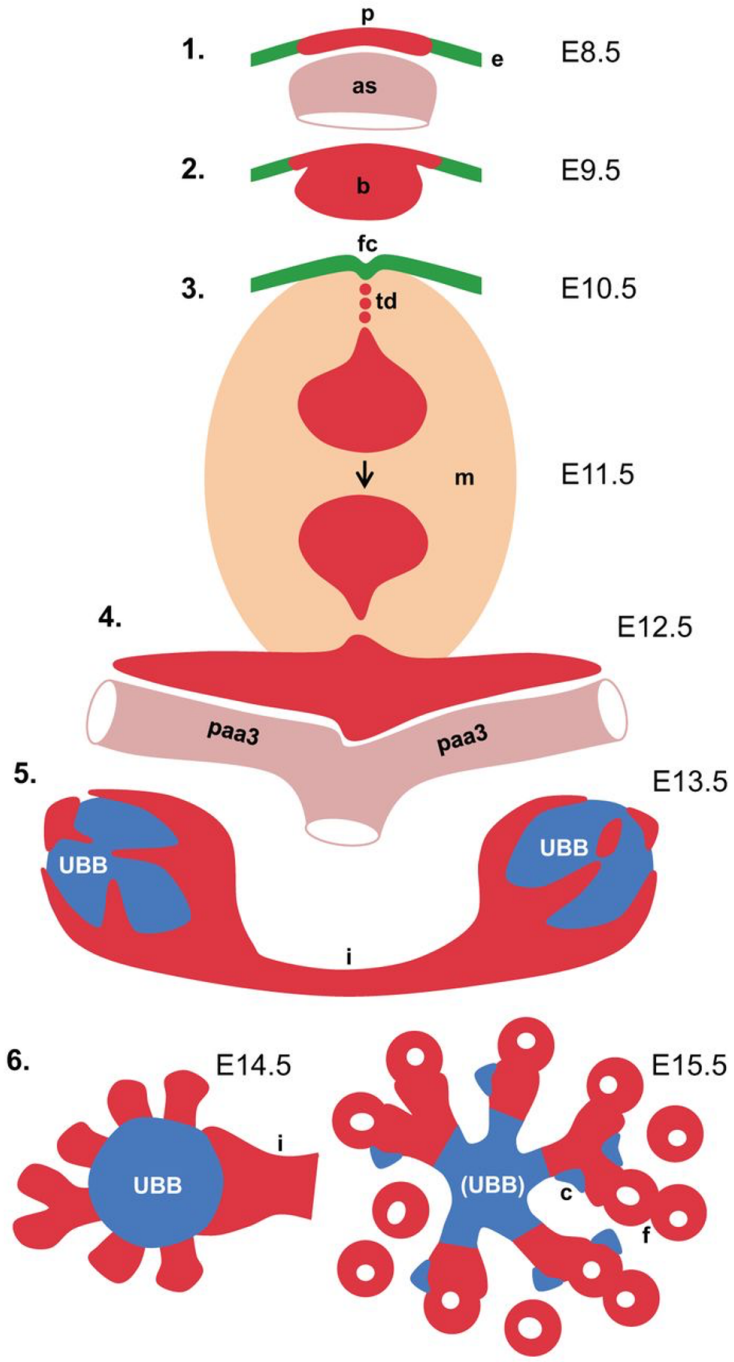


Figure 2.2: Thyroid morphogenesis in the mouse can be broken down into six discernible stages. (1) Specification of thyroid progenitors in a restricted area, known as the placode (p), in the pharyngeal floor of the endoderm (e) and in close proximity to the aortic sac (as). (2) Budding of the thyroid primordium (b) by adjacent cell recruitment. (3) Detachment and migration of the thyroid primordium into the mesenchyme tissue (m). (4) Elongation of the thyroid primordium, extending along the third pharyngeal arch arteries (paa3). (5) Bilobation and fusion of the thyroid primordium with the UBBs, with formation of the isthmus (i). (6) Lobe growth, initiation of folliculogenesis (transition from a mass to a multitude of microfollicles (f)) and mixing of the C cells. In mice, the entire process takes 1 week, after which enlargement of the gland is attributed to the gradual increase in the size of individual follicles. The morphogenetic stages of human thyroid development are nearly identical to those in mice, although much prolonged. Figure adapted from [8].

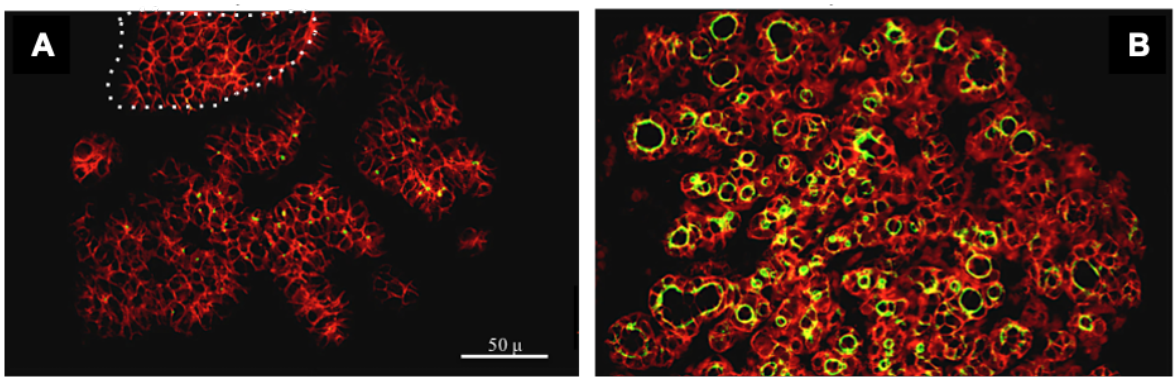


Figure 2.3: Reorganisation of the thyroid epithelial cell mass into a multitude of single-layered follicles and acquisition of the apical-basal polarity. (A) Immunostaining of a mouse thyroid gland at E14.5. Epithelial cells are marked in red (E-cadherin or CDH1) and display small ezrin-positive signal in green. The dotted white line outlines the parathyroid gland. (B) Immunostaining of a mouse thyroid gland at E18.5. The epithelial cells are now organised as single layers. The apical surface is marked in green (ezrin) and defines the lumen of the follicles. Figure adapted from [17].

on, bilateral elongation of the thyroid primordium shows a general increase of progenitor cell proliferation, but after fusion with the UBB, the growth of the two lobes would involve a region-dependent proliferation of epithelial cells [15].

At E14.5, epithelial cells located at the periphery of the mass are in contact with the extracellular matrix (ECM) and are thus partially polarised (notable presence of a basal pole). Starting from this stage onward, microlumina will appear between thyroid epithelial cells (Fig. 2.3, A). It is proposed that these microlumina form by fusion of pre-existing intracellular vesicles with the membrane opposite to the basal pole [18]. Around E16.5 - E17.5 (12th week in human), most of the cells are fully polarised with an apical pole facing the microlumen and a basal pole resting on a basement membrane. Thyroid follicles will continue to grow postnatally and reach and maintain their mature organisation by oriented cell divisions and filling of the colloidal space *via* apical delivering of TG-containing intracellular vesicles (Fig. 2.3, B).

Folliculogenesis, the last stage of embryonic thyroid development, consists in the formation of a central lumen filled with thyroglobulin. Distinct mechanisms of lumen formation have been described in several biological systems [19]. For instance, lumina can be formed by collective cell migration and folding of an epithelial sheet into a tube, *e.g.* the neural tube [20]. A lumen can also appear by creating a new space in a tight cellular mass, where no space existed before. This mechanism is known as *de novo* lumen formation. Four major mechanisms of *de novo* lumen formation have been described; cord hollowing, cell hollowing,

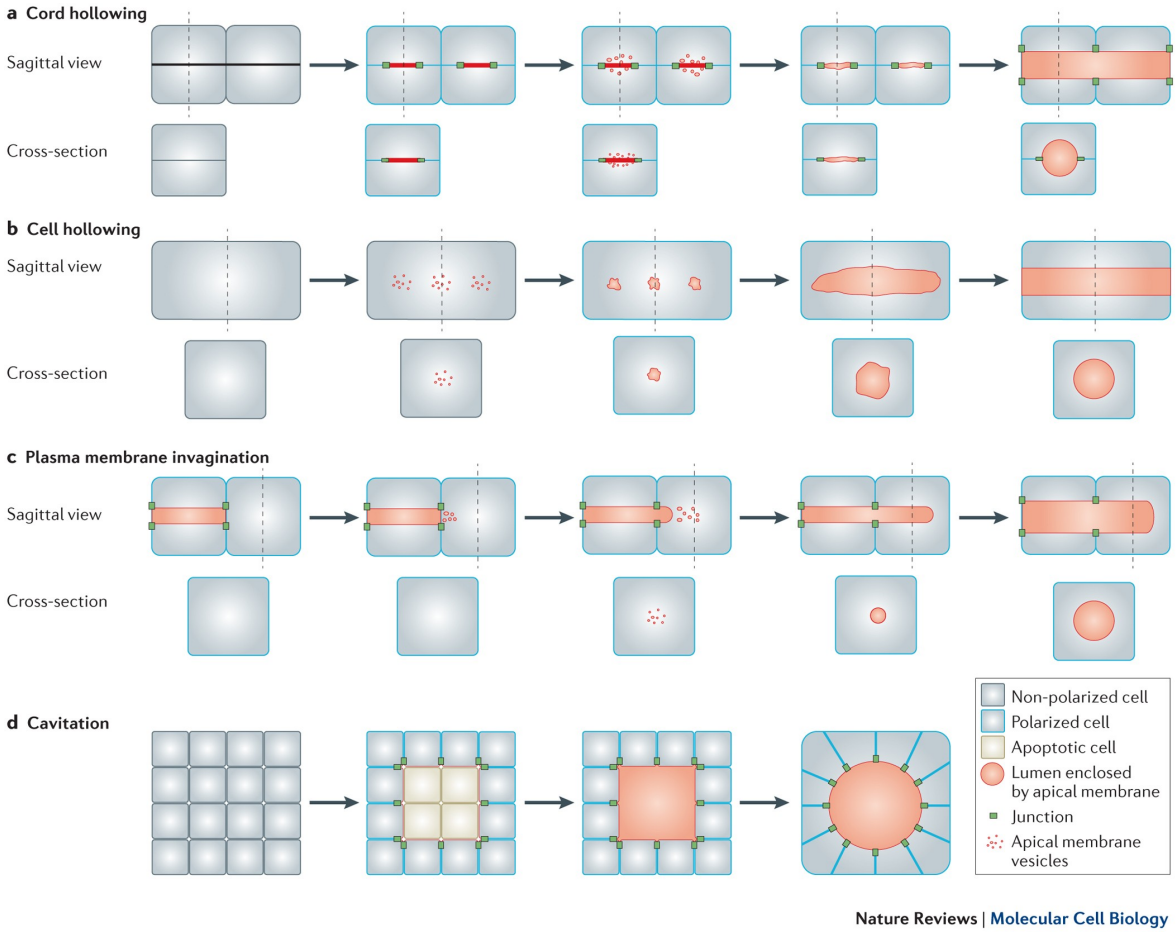


Figure 2.4: Lumina can develop *de novo* by four different morphological mechanisms. (a) In cord hollowing, secreted vesicle fuse at the apical contact surface between two cells. Continuous vesicular delivery contributes to the increase of membrane surface and expansion of the lumen by delivering fluids and other materials. As this expansion occurs, the newly formed lumen can merge with other adjacent lumina to eventually create an elongated lumen or tube. (b) Cell hollowing is rather similar to cord hollowing, in the sense that vesicle fusion is the major process leading to lumen formation. The main difference is that these vacuoles fuse together inside the cytoplasm, forming a large intracellular lumen before fusing to the plasma membrane of the cell. (c) In plasma membrane invagination, lumina are formed and extended into the adjacent cell by addition of membrane material to an already-existing apical membrane through vesicular delivery. (d) In cavitation, the lumen forms when cells in the centre of a cell mass are eliminated by apoptosis triggering surrounding cell shape changes. Figure adapted from [19].

plasma membrane invagination and cavitation (figure 2.4). In cord hollowing, lumina are formed when intracellular vesicles or vacuoles fuse with the membrane of a polarised cell, thereby creating an extracellular space and an apical pole. Delivery of new material and fluids will increase the membrane surface area and as the newly formed lumen expands, it may connect to other adjacent growing lumina within the cord of cells thus forming a continuous lumen or a tube. Cell hollowing is quite similar to cord hollowing in the difference that here, vesicles first fuse together within the cytoplasm rather than with a membrane, and form an intercellular lumen. The lumen of the tube is thus surrounded by only one cell. In plasma membrane invagination, lumina are formed by addition of membrane material to this already-existing membrane, also by vesicular delivery. Hence, the lumen will extend into the adjacent cell. Finally, in cavitation, lumen forms following apoptosis of cells in the centre of a mass of cells. Although the end result of cavitation looks like thyroid follicles, apoptosis is not observed during folliculogenesis and the mechanism governing lumen formation or folliculogenesis rather depends on the fusion of intracellular vesicles (as in the cord hollowing mechanism, see Fig. 2.4, b). The main difference is that in the thyroid, a space rather than an open tube is formed. Tight junctions[3] are established between cells (2.5, A), sealing the newly formed lumina from the extra-cellular environment. Note that intracellular vesicles used to form the lumen contain thyroglobulin that simultaneously and progressively fills the colloidal space as the lumen grow. It has been observed that around E15 (in mice), a significant number of microfollicles already present a small lumen, suggesting that folliculogenesis is a directed and synchronous process at this stage [8, 15, 17]. In mice, this entire process of thyroid morphogenesis takes approximately one week. In humans, it lasts several weeks and first evidence of hormone synthesis do not appear before the eleventh week of gestation [21].

In adults, the main regulator of thyrocyte proliferation and thus thyroid growth is the thyroid-stimulating hormone TSH, also known as thyrotropin or thyrotropic hormone. TSH is a glycoprotein hormone produced by thyrotrope cells in the anterior pituitary gland. On the contrary, embryonic thyroid cells proliferate independently of TSH. It is not before thyrocytes have differentiated and started producing thyroid hormones that the switch to a TSH-dependent proliferation occurs (figure 2.6, top). This could potentially indicate that proliferation during thyroid morphogenesis relies entirely on local signalling *via* intrinsic or extrinsic factors [23].

[3]Tight junctions, or *zonula occludens*, are extremely tight multiprotein junctional complexes, located at the apical side of the cell. They are composed of transmembrane proteins such as occludins, claudins and other junctional adhesion molecules.

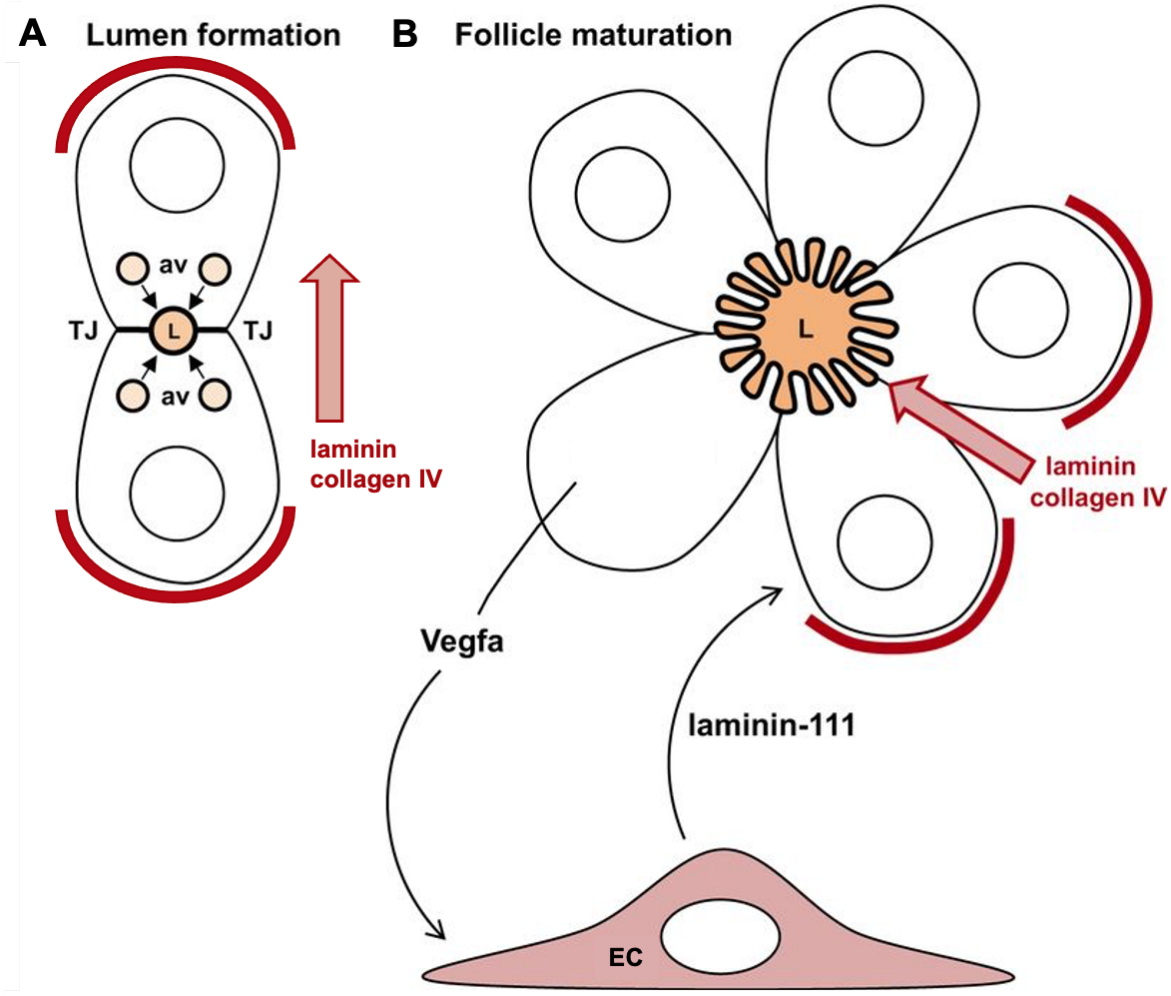


Figure 2.5: Cellular dynamics during thyroid bud formation and folliculogenesis. (A) Initiation of follicle formation between two progenitor cells that shows first signs of basal-apical polarity. Note that in this thesis, we place the role of EC upstream of the first signs of basal-apical polarisation. A microlumen (L) is formed by fusion of vesicles (av) that accumulate in the cytoplasm of both cells. These vesicles express apical markers such as ezrin and also contain thyroglobulin that is simultaneously secreted as the lumen forms. Strong tight junctions (TJs) are established between the two cells, segregating the lumen from the external environment. (B) Maturation of microfollicles in mature functional hormone-producing units requires a proper vascularisation. Endothelial cells (ec) ingrowth is triggered by secretion and diffusion of the growth factor VEGFA by thyroid cells. Other local factors (such as laminin-111) released by EC consecutively promote further maturation of follicles. Deposition of laminin and collagens into a nascent basement membrane then ensures follicles maturation and growth. Figure adapted from [8] and based on a model proposed by [22] and [17].

Intrinsic factors regulating thyroid morphogenesis

Many transcription factors[4] play a role in thyroid development. The exhaustive list will not be discussed and a detailed summary of their contribution is presented by Fernandez and colleagues in [24]. However, four of them seem to emerge as a signature network active in thyroid development (figure 2.7). In fact, Hhex, Nkx2-1, Pax8 and Foxe1 are all co-expressed in, and only in, the thyroid placode and at all the subsequent stages of thyroid development. It has been shown that the deletion of any of these genes leads to hypoplasia of the thyroid [13].

In mice, Hhex is expressed in the foregut endoderm as early as from E4.5 and its expression remains high until thyroid maturation. While it is still not clear how and if it may affect the positioning of the initial thyroid bud [25], it has been shown that Hhex has an essential role in later stages, by maintaining the expression of Nkx2-1, Pax8 and Foxe1 [25, 26]. In fact, Hhex null[5] mutants seem to display an appropriate thyroid primordium location and budding. Moreover, expression level of Hhex is not restored in these mutants, indicating that it is neither involved in thyroid primordium specification nor in its budding [25]. Nkx2-1 has been identified as a strong regulator of glandular differentiation into a follicular thyroid while having little impact on the earlier morphogenetic stages [27]. Similarly, studies have shown that Pax8 transcription factor is needed to ensure thyroid precursors survival as they disappear at around E11.5 in Pax8 null mouse embryos despite showing a comparatively normal anlage morphology [25]. These transcription factors also act as Foxe1 regulators, which in turn promote the migration of thyroid precursor cells (see previous section) [25]. Interestingly, all but Hhex have been shown to control the expression of adhesion proteins [8]. Moreover, it has been demonstrated that co-expression of Pax8 and Nkx2-1 plays a role in thyroid follicular cell differentiation, again supporting the importance of this regulatory network [28]. Finally, it has been shown how concomitant induced expression of Pax8 and Nkx2-1 in embryonic stem cell triggers differentiation towards thyroid progenitors able to form follicles that exhibit typical thyroid functions upon culture in 3D collagen gel [29].

[4]Transcription factors are proteins that control the rate of transcription of genetic information from DNA to messenger RNA, by binding to specific DNA sequences.

[5]Induced mutation (change) in a gene that leads to the absence of its mRNA and/or translated functional protein product.

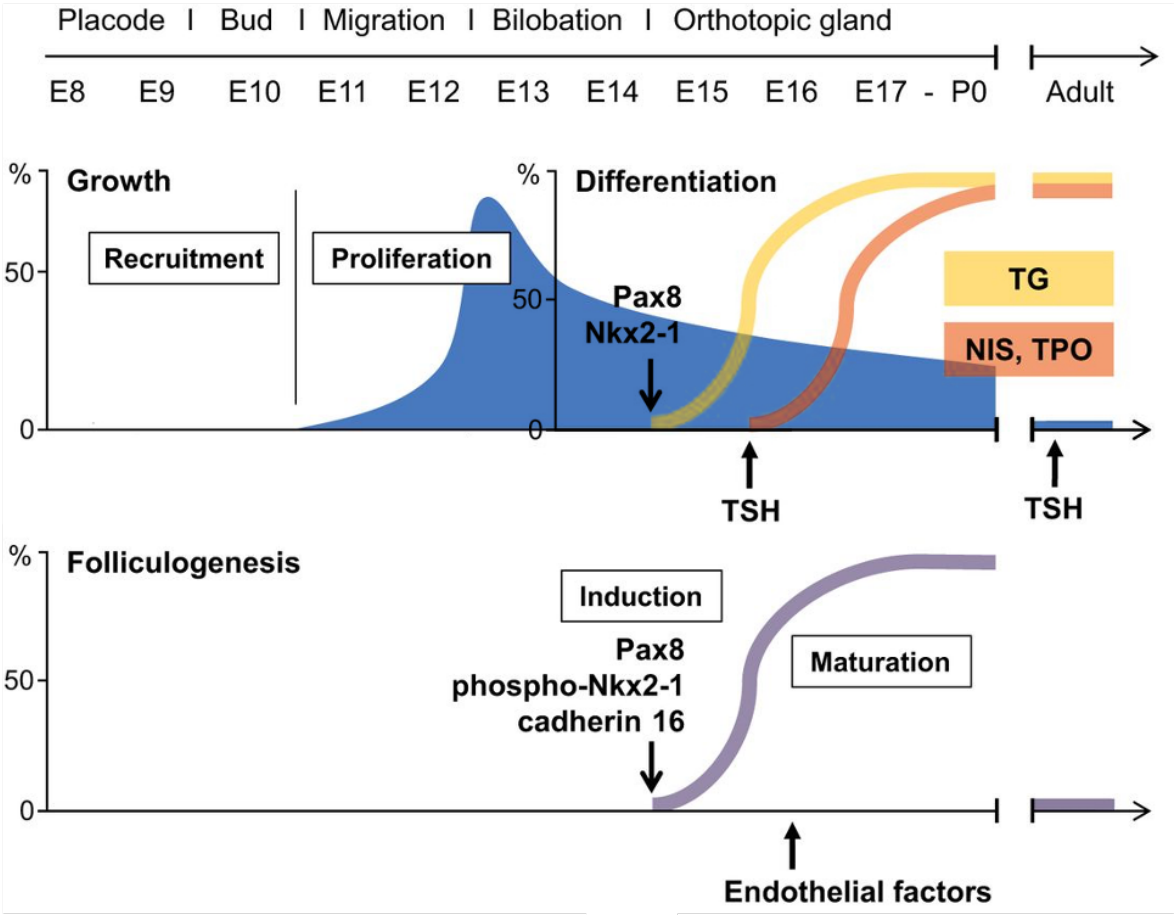


Figure 2.6: Dynamics of primordium growth and development and functional differentiation in the embryonic thyroid. In mouse embryos, the proliferation of Nkx2-1 positive progenitors (top, in blue) is low until the end of the thyroid primordium migration and the beginning of the bilobation process. In late development, thyroid progenitors proliferation causes embbryonic thyroid to enlarge in its final position. TSH does not influence thyroid growth until after the maturation of the thyroid follicles (top, blue). The differentiation of thyroid progenitors (top, in red and yellow) is concurrent with the TSH-independent expression of thyroglobulin (TG) and, 1 or 2 days later, the TSH-dependent expression of sodium-iodide symporter (NIS) and thyroid peroxidase (TPO). Transcription factors Pax8 and phosphorylated Nkx2-1, as well as the cadherin 16, are required for differentiation of progenitors and for follicle formation (bottom, purple). Not that folliculogenesis is also occuring synchronously with TSH expression but the precise factors instituting this simultaneity are as yet unidentified. Further maturation of maturing follicles relies on reciprocal interactions with microvessels and hence endothelial-derived factors (bottom). Figure adapted from [8]. (see also Fig. 2.5).

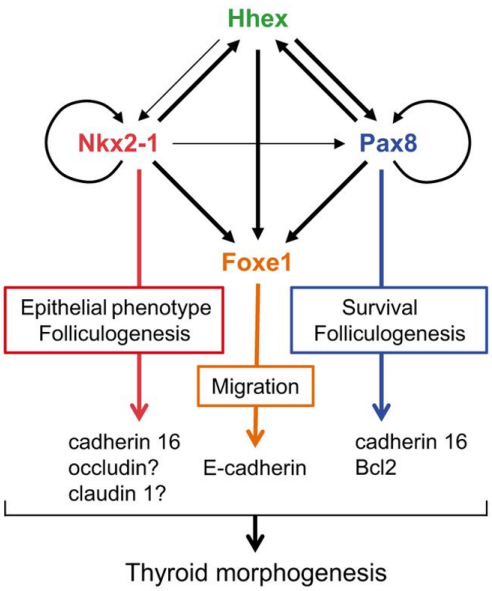


Figure 2.7: Transcription factors network involved in early thyroid development. Hhex, Nkx2-1, Pax8 and Foxe1 are co-expressed as from the thyroid placode stage. As the thyroid bud forms and grows, each factor but Foxe1 regulates the expression of the others (black arrows). Autoregulation of Nkx2-1 and Pax8 expression has also been shown in cultured thyroid cells. Moreover, all factors except Hhex are known to control the expression of adhesion proteins such as cadherin 16 and E-cadherin. Note that the regulation of occludin and claudin 1 (tight-junction proteins) by Nkx2-1 has also been shown in lung epithelial cells [30]. Figure adapted from [8].

2.2.1 Extrinsic factors regulating thyroid morphogenesis

During its development, the embryonic thyroid migrates through different micro environments before reaching its mature position. It is therefore not surprising that its development is regulated by various extrinsic factors coming from these embryonic tissues. For instance, the pharyngeal floor of the endoderm (from which thyroid progenitors derived) is located close to the cardiac tract (Fig. 2.2, 1). Cardiac mesoderm and other vascular signalling could thus regulate thyroid morphogenesis. In fact, significant correlation has been found between thyroid dysgenesis and cardiovascular malformation [31]. Other studies suggested that the neural crest cells could play a role in the thyroid development. For instance, Maeda and colleague [14] showed that in chicks embryos, the removal of the cranial neural crest results in malformation of the thyroid and parathyroid gland as well as the cardiovascular system.

Formation of the thyroid also coincides with the appearance of abundant endothelial cells (EC) [15]. EC are generally recruited by epithelial organs by secretion of growth factor[6] such as VEGFA (Vascular Endothelial Growth Factor), and are induced to form blood vessels in a process called angiogenesis. The combined process of production, diffusion and degradation of growth factors creates a gradient across the width of the epithelial tissue [36–38]. This gradient causes the activation of some endothelial cells (which will differentiate and become tip cells) in neighbouring blood vessels. Tip cells thus change shape and migrate

[6] A growth factor is a substance secreted by cells and able to stimulate various processes such as cell proliferation or migration. Many growth factors are protein or steroid hormones, acting as signalling molecules.

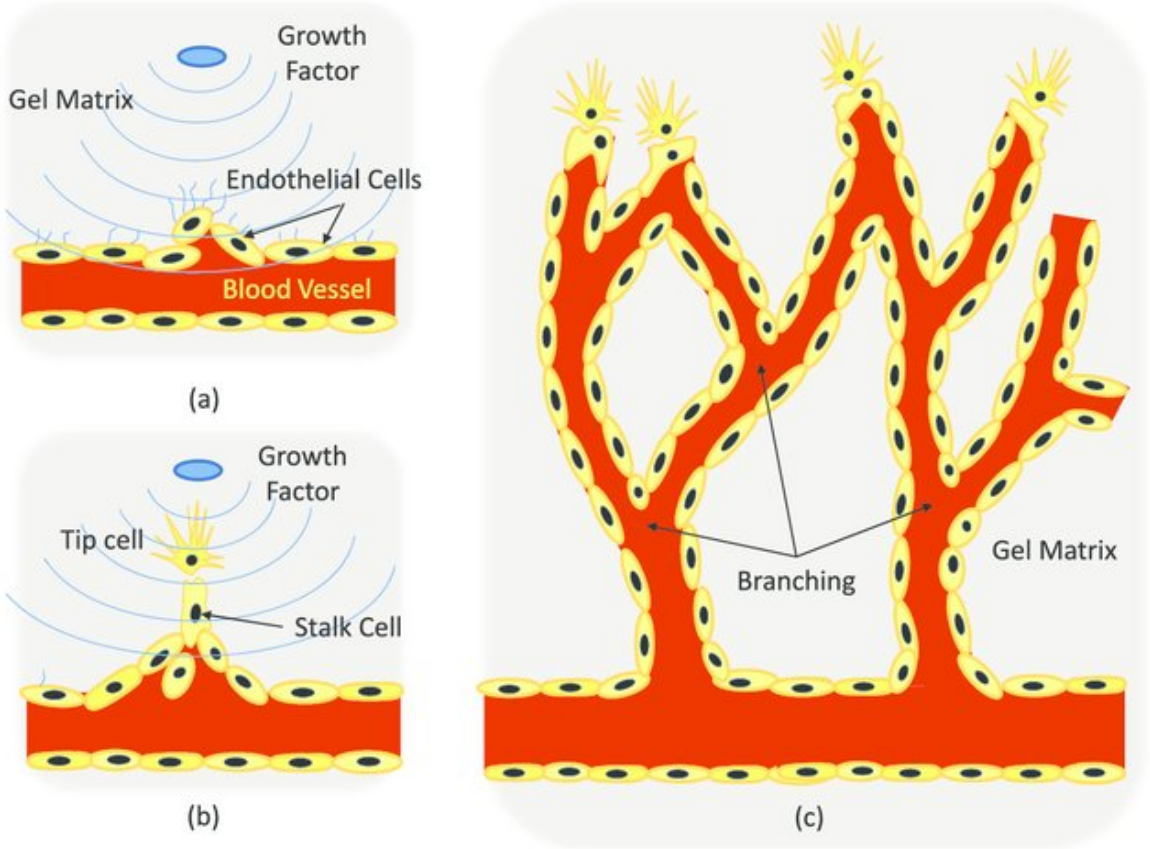


Figure 2.8: Angiogenic sprouting process. (a) EC residing in a blood vessel sprout out in response to growth factor molecules, such as VEGFA and FGF (for Fibloblast Growth Factors, a signalling protein promoting the proliferation of fibroblasts, EC and smooth muscle cells) or other chemical stimulators such as MMPs (matrix metalloproteinases, essential for the degradation of the ECM allowing the growth of endothelial cells). VEGFA is an isoform of the Vascular Endothelial Growth Factor, a secreted glycoprotein stimulating the formation of blood vessels by chemoattracting endothelial cells and increasing vascular permeability [32]. Since they are lined by a single layer of endothelial cells, this increase is essential for the prospective blood vessel of capillary to allow the flow of small molecules such as nutrients, water or ions. VEGFs are typically secreted when tissue need oxygen. (b) A leading non-proliferating cell, called a tip cell, detects gradients in certain growth factors and migrates toward the source while stalk cells follow behind the tip. This detection occurs by biding of VEGFA and activation of tyrosine kinase receptors on the cell membrane. Tip cells will extend long filopodia, a projection of the cell membrane supported by actin microfilament, which will guide directed cell motion hence migration. Following endothelial cells differentiate into stalk cells. Those cells are not able to detect growth factors but will proliferate under tip cell pulling force [33], bridging the gap between the latter and the parent blood vessel. Note that only the stalk cells will later generate the blood vessel lumen in a process called lumenogenesis (not showed). (c) New branches are occasionally formed by the migration of newly differentiated tip cells. A combination of sprouting, branching and connecting processes will eventually lead to the formation of a vascular network. Note the individual cell behavior is believed to be stochastic [34] meaning that sprouting and branching occurs randomly at various sites, yet resulting in a highly regulated network. Figure adapted from [35].

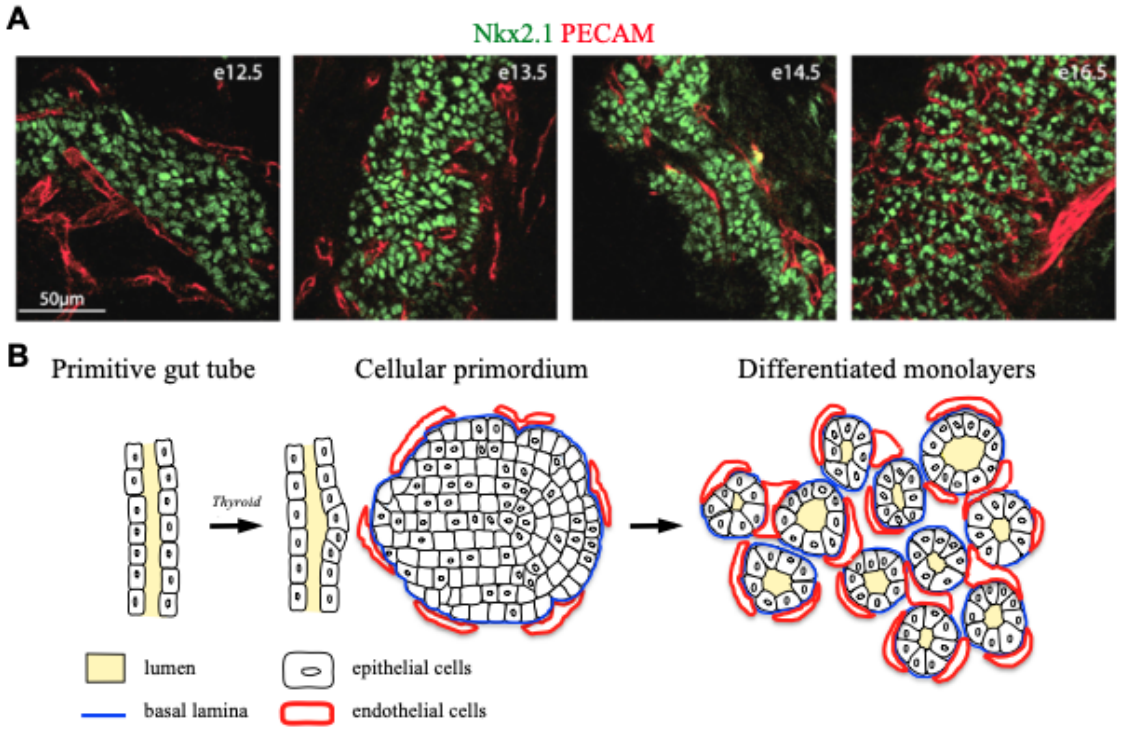


Figure 2.9: Interactions of endothelial cells with thyroid cells during the reorganisation of the thyroid primordium into single-layered follicles. (A) Interactions between epithelial cells (nucleus marked by transcription factor Nkx2.1, in green) and endothelial cells (marked by PECAM, a cellular adhesion molecule, in red). (B) Schematic of the reorganisation of the thyroid primordium into single-layered follicles. Figure adapted from [22].

in the direction of the gradient, closely followed by other endothelial cells from the blood vessel, that will be called the stalk cells. The combination of tip cell migration and strong cell-cell (tip-stalk) adhesion eventually leads to the formation of new endothelial branches. A schematic of the key concepts can be found in Fig. 2.8. Interestingly, studies illustrated the contribution of blood vessels to thyroid folliculogenesis, involving mutual interactions driven *via* VEGFA production by thyroid cells themselves [22]. Thyroid interactions with EC are observed as early as E8.5 (in mice) when precursor cells of the thyroid primordium associate with the aortic bulb[7] as well as pharyngeal arches[8] during the first stages of morphogenesis (*i.e.* budding, migration and bilobulation) [15]. Immunohistological analysis revealed that EC are continuously in contact with thyroid progenitors during embryonic development and folliculogenesis (Fig. 2.9). The thyroid epithelial bud is indeed rapidly surrounded by

[7]the primordial vascular channel from which the aortic arches arise

[8]series of six paired embryological vascular structures that give rise to the great arteries of the neck and head.

small capillaries during development. EC will progressively invade the epithelial mass to form a dense endothelial network inside and around the two lobes. At E16.5-E17.5, each follicle is closely surrounded by a rich network of capillaries [22]. The role of embryonic endothelial vessels in thyroid development was first considered by Fagman and colleagues [39]. They suggested that the pharyngeal arch arteries could serve as leading track for epithelial growth and development into a bilobed gland. Our laboratory [17, 22] demonstrated that the recruitment of endothelial cells during thyroid development depends on the expression of VEGFA by epithelial cells. Indeed, they showed that a decrease of more than 90% in VEGFA expression lead to a significant decrease in endothelial density at E17.5 (Fig. 2.10). Moreover, they demonstrated that this decrease in endothelial density had a strong impact on the reorganisation of epithelial cells into pre-follicular structures at E15.5 (Fig. 2.10, A and A') and eventually into large polarised epithelial monolayers (follicles) around birth (Fig. 2.10, B and B'). They finally showed that the assembly of the basement membrane and the opposite apical pole, partly controlled by endothelial cell invasion in the thyroid, was crucial for epithelial cell orientation, polarisation and thus proper follicle development (Fig. 2.10, C and C'). In line with the previous statements, they proposed a model in which thyroid cells produce VEGF, which induce EC recruitment and invasion into the thyroid primordium. Endothelial cell invasion would then induce changes in epithelial cell adhesion and cohesion and trigger epithelial production and deposition of basement membrane proteins. Interactions of epithelial cells with the basement membrane proteins *via* their specific basal receptors, integrins, brings an orienting clue to the epithelial cells and orient subsequent delivery of intracellular vesicles to the opposite pole of the cell, leading to the formation of lumina. Several *in vitro* studies support this model and the role of ECM proteins. For instance, it has been illustrated how *in vitro* cultures of MDCK[9] epithelial cells in collagen-enriched gel trigger the assembly of a basement membrane which promotes formation of cystic structures highly similar to thyroid follicles [40]. Ingeson-Carlson and Nilsson [41] also showed that dissociated thyroid cells embeded into a collagen matrix spontaneously form follicles without any other requirement.

Various features linking folliculogenesis with functional differentiation of the thyroid cells have also been observed. In human fetuses, TG is expressed before the first signs of follicle formation and thus endogenous hormone secretion [21]. Moreover, intracellular lumina have been observed in dispersed thyroid cells (Fig. 2.3 and [17]) suggesting that these apical vac-

[9]Madin-Darby Canine Kidney (MDCK) cells are a model mammalian cell line used in biomedical research.

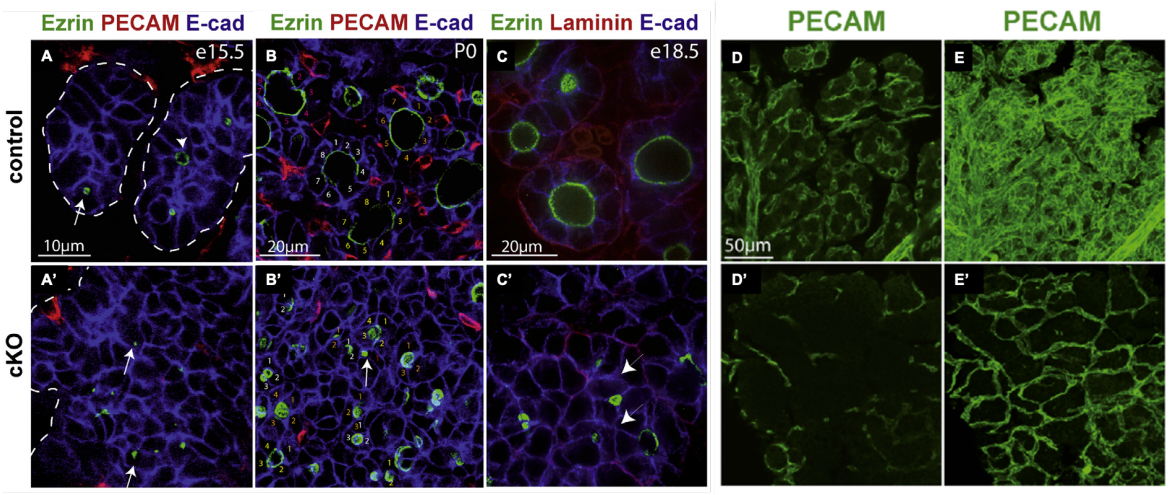


Figure 2.10: Epithelial reorganisation and follicle formation impairment induced by *in vivo* reduction in vascular density. ((A)-(B')) Epithelial cells are shown in blue (E-cad), the apical membrane of polarised epithelial cell in green (ezrin) and endothelial cells in red (PECAM). Tr, trachea; Oe, oesophagus. ((C),(C')) Epithelial cells are shown in blue (E-cad) with their basal pole in red (laminin) and their apical pole in green (ezrin). ((D)-(E')) Endothelium is showed in green (PECAM). (A) Control at E15.5. Epithelial masses (dotted lines) start to form pro-follicular structures and display ezrin-positive structures (in green). (A') VEGF KO at E15.5. The epithelium remains a mass of non-polarised cells and displays small random ezrin-positive signals. (B) Control around birth. Thyrocytes are organised into follicles enclosing a well-defined apical lumen (in green). (B') VEGF KO around birth. Thyrocytes are still organised as a multilayered mass. Few small ezrin-positive structures are scattered in this mass. (C) Control at E18.5. Localisation of the basal (in red) and apical (in green) of fully polarised epithelial cells (in blue). (C') VEGF KO at E18.5. A minority of epithelial cells display ezrin at one pole with no laminin at the opposite side (and vice-versa). (D) and (E) Control at E17.5. Endothelium is showed in green (2D section in (D), 3D reconstitution in (E)). (D') and (E') VEGF KO. The endothelial density has significantly decreased due to the KO in VEGF expression ((2D section in (D'), 3D reconstitution in (E')). Figures adapted from [22].

uoles may have a developmental role. Note that these observations support the mechanism of *de novo* follicle formation by fusion of intracellular lumina with a membrane and synchronised secretion of TG, in line with our previous statements. Importantly, this is also linking the last step of thyroid development (being folliculogenesis) to functional differentiation of the thyroid gland.

In view of all the above, it is now clear that thyroid morphogenesis, and more particularly folliculogenesis is regulated by a set of both intrinsic and extrinsic factors. In fact, follicular units in close proximity with the vascular system appear as key elements. Hence, understanding the morphogenetic processes that lead to the formation of mature and fully functional follicles closely associated with blood vessels is not only important for basic knowledge but

also for thyroid diseases affecting its organisation. While significant findings and progress are made day by day towards a better understanding of the developmental mechanisms of thyroid morphogenesis, there is still room for the development of new approaches.

Box 1 : Thyroid diseases

Several diseases can affect the thyroid gland, and can emerge due to various factors. The most common disorder of the thyroid gland is the Multinodular Goiter (MNG). MNG appears following iodine deficiency inducing an abnormal TSH-mediated growth of the thyroid follicles. It can also eventually lead to a hyperthyroidism condition after a few years, due to an overexpression of TG and an overproduction of thyroxin. Hyperthyroidism can overstimulate the patient's body's metabolism resulting in weight loss and irregular heartbeat. Note that the majority of the cases of hyperthyroidism are in fact a result of another pathology called Grave's disease. This autoimmune disease is characterised by an irregular enlargement of the thyroid caused by excessive production of thyroid hormones, this time due to the action of an antibody called thyroid-stimulating immunoglobulin (TSI). Those diseases most probably involved a combination of genetic (DNA mutation) but also external factors such as a poor diet, stress or smoking.

Thyroid cancer represents 2% of all cancers worldwide (3.2 million people in 2015 [42]). The five-year survival rate is close to 98% and up to 95% for 10 years. Note that the number of cases seems to have slightly increased this past decade, probably due to better detection and awareness. This crucially reminds us that, on the other hand, fewer cases detected does not automatically mean fewer patients affected by this or any disease. This is something worth acknowledging, especially at this particular time. There are four main types of thyroid cancer; papillary (PTC), follicular (FTC), medullary (MTC) and anaplastic (ATC). PTC is by far the most common type of thyroid cancer (up to 80% of all thyroid cancers) and develops from follicular cells. FTC also affects follicular cells, but in only about 10% of the cases. ATC, while being a rare form of thyroid cancer (1-2%), is also more aggressive, with a way poorer prognosis (10% 5-years survival rate). It is characterised by uncontrolled growth of thyroid cells that no longer resemble the original cells and exhibit very low differentiation markers. ATC is often the consequence of a late or non-diagnosed FTC or PTC. Unlike the first mentioned, MTC (3% of all the thyroid cancers) is a differentiated cancer that derives from C cells. Treatments for all forms of thyroid cancer often require radiation therapy (including radioactive iodine), but also chemotherapy or even surgery.

3 | Modelling Preliminaries

In the fields of theoretical biology, plant physiology and even medicine, mathematical modelling appears today as a powerful complementary tool aiming to unlock complex biological systems. Keeping in mind that they remain an abstraction of the reality, mathematical (or computational) models allow the investigation of the role of underlying cell mechanics and interactions in morphogenetic processes. In this sense, they can serve a broad diversity of role, such as testing hypothesis or generating predictions that are likely to be validated experimentally [43]. To this end, a wide range of models has been developed and several analyses of how these approaches can compare when applied to the same biological issue have been performed [44]. Within the context of this particular study, the choice has been made to focus on one specific cell-based modelling approach; the vertex model. This model seemed to be the most suitable to tackle the various aspects of thyroid morphogenesis.

3.1 Cell-based models

As opposed to continuum models describing the behaviour of a continuous mass (*e.g.* in fluid mechanics), cell-based models (CBM), also called agent-based or individual-based models, treat cells as discrete entities. Today, the majority of spatial models of tissue growth or development tackles this issue with a continuous point of view, treating a tissue as a continuum and working on length scales that are much larger than a typical cell length. Processes such as adhesion, motility or cell-cell signalling can therefore be hard to incorporate. Individual-based models can be used as a rather simple and computationally low demanding approach for simulating the mechanical behaviour of cells constituting a tissue [45]. In cell-based models, each cell is treated individually and characterized by a so-called state-vector that includes information regarding its position, velocity and internal state. The inputs to such models, as stated by Palm and Merks [46], are "all the behavioral rules that cells follow", both regarding internal processes and interactions with neighbours.

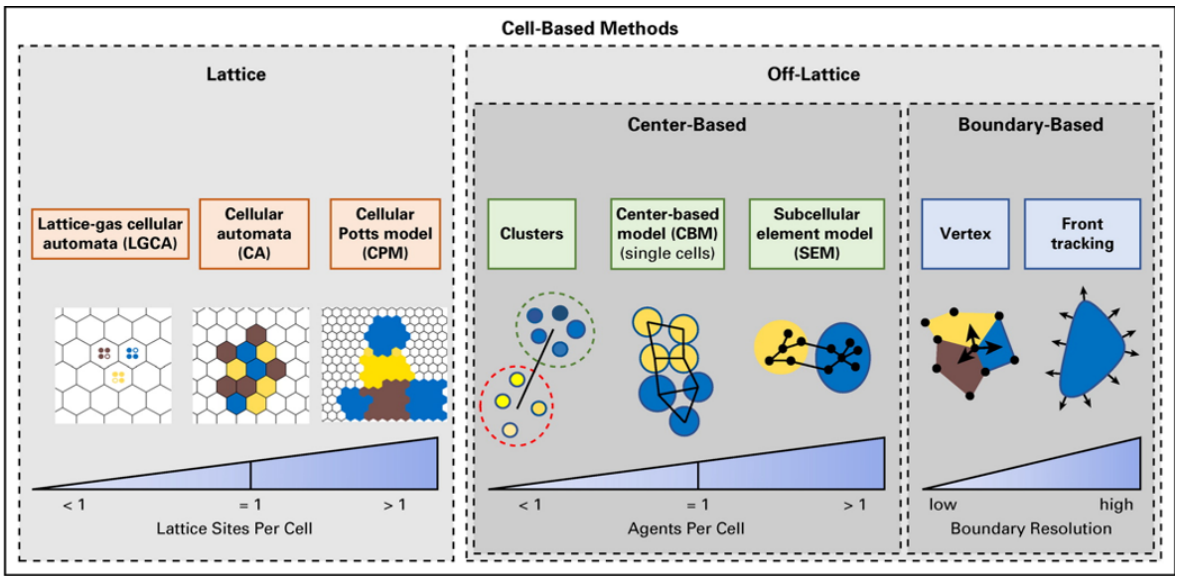


Figure 3.1: Schematic of Cell-based approaches. In Lattice models (left), cells are constrained to lie on an artificial grid. Regular structured meshes are often used (using squared or hexagonal-shaped lattice sites) and can be differentiated by the number of lattice per cell. In lattice gas CA (LGCA) models, a single lattice site can contain multiple cells. The number of cells that move through channels between individual lattice sites is tracked rather than the motion of each individual cell. In cellular automata (CA), each cell is represented by a single lattice site. The state of each cell is determined by a set of deterministic or stochastic rules applied on the previous state of the system. In cellular Potts models (CPM), each cells is formed by several lattice sites. Off-lattice models (right) are free from this fixed grid geometry and can be differentiated into centre-based and boundary-based models. In centre-based models, each cell is represented by one or many points free to move in space. Distinction can be made according to the number of agents per cell. Cells can be organised into clusters that are simulated as agents that interact by mechanical forces or other rule-based motions. In centre-based models single cells (or Overlapping-Spheres, OS), only the cell centre is tracked over time and each cell is therefore represented by a single point and associated with a specific radius, defining a sphere (in 3D). In Subcellular Element Model (SEM), each cell is constructed from subcellular elements (allowing more complex cell shape), with intracellular bindings between elements of a given cell and intercellular bindings connecting adjacent elements from neighbouring cells. Finally, in vertex-based models, each cell is represented by a set of points (vertices) linked together by edges while front-tracking models couple a continuous description of a viscous incompressible fluid with the dynamics of separate elastic cells. Figure adapted from [47].

As in every modelling processes, we rely on assumptions and simplifications, notably regarding the shape of the cells [44]. However, cell-based models are still able to capture the complexity of tissue organisation emerging from cell's collective behaviour [45]. Importantly, they also naturally account for the stochastic "noise" as well as the intrinsic heterogeneity of multicellular tissues, both in small two-dimensional populations of highly detailed cells and in larger three-dimensional systems of thousands of simpler cells [47]. Moreover, they are relatively easy to couple with other continuous models, such as reaction-diffusion models (*e.g.* simulating morphogen transport or cell-cell signaling) and therefore cover a wide range of processes active in various biological systems. [48]. Multiple cell-based modelling approaches exist, and they can be roughly categorised as on- and off-lattice models whether or not cells are constraints to lie on an artificial grid, or lattice (Fig. 3.1).

On-lattice models

Lattice-based models can either use regular structured meshes (*e.g.* based on a two or three-dimensional Cartesian system) or more complex unstructured meshes [49]. On-lattice models can further be differentiated by their spatial resolution.

In Cellular-automata (CA), the simplest agent-based model, each cell is represented by a single lattice site. The state of each cell is determined by a set of deterministic or stochastic rules applied to the previous state of the system [50] and the cell movement is defined in term of cell moving from one lattice site to another. The best-known example of the use of CA models is the so-called "Game of Life", developed in 1970 by the mathematician John Conway [51]. The domain of the Game of Life is an infinite two-dimensional grid of squared lattice site. Each site can be in two possible states, live or dead. These sites, or cells, communicate with their eight neighbours and at each time step, the transition from one state to an other occurs according to a set of 4 simple rules. Starting from an initial condition of live and dead cells, the first generation is created by simultaneously applying the 4 rules to each site. Several pattern can emerge from such game, such as "still lifes" (no change from one generation to another), "oscillators" (returns to its initial state after a certain number of generations) or "spaceships" (translate accross the domain). In developmental biology, CA can be useful to understand the emergence of tissue patterning during development. Cohen and colleagues [52] developed a model of lateral inhibition-mediated patterning and used a simple asynchronous[1] CA model to analyse the role of noise in this process. Lateral inhibition

[1] An asynchronous CA updates each cell independently, the new state of a cell hence affecting the calculation

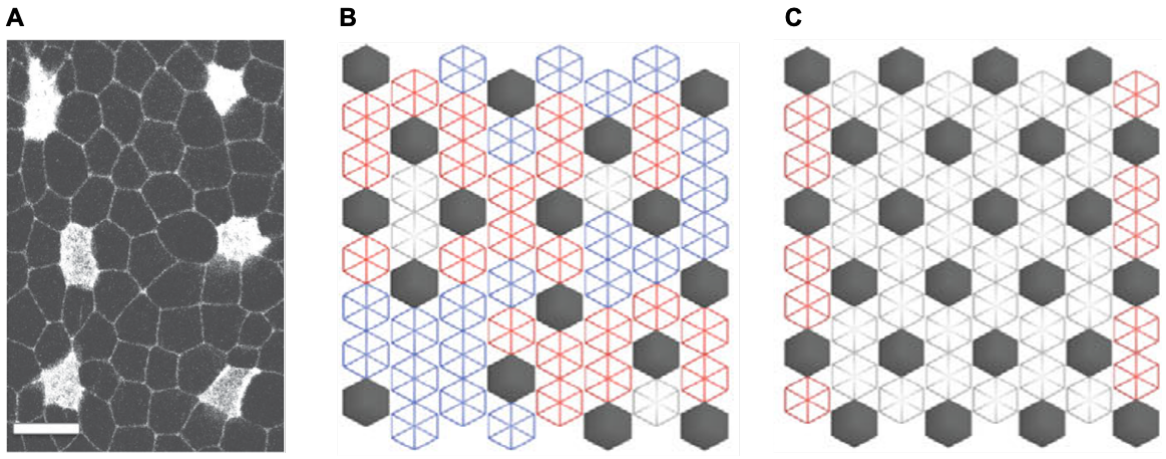


Figure 3.2: (A) Apical section of epithelial cells in an area of the notum close to the fly midline, at 14 h after pupae formation. Cells destined to become sensory organs (white) express Neuralized-Gal4. Epithelial cells are visible through E-cad staining. (B) Simulations of inhibitory signalling with spatial noise executed in an 8 x 8 hexagonally packed grid of cells. Active cells (expressing Neur-Gal4) in dark grey, inactive cells in red, blue or light grey depending on the number of active neighbours. Initial condition in the middle, simulation results after 100 steps (optimised state) on the right. Figures adapted from [52].

patterns emerge when a homogeneous group of cells compete to express an inhibitory signal [52]. Therefore, each cell either expresses this inhibitory signal or is inhibited by a signal expressed in one or more neighbouring cells, which eventually describes a binary state system (with the cells being either active or inactive). The rules governing such system are quite simple in the sense that a cell with no active neighbour will transition to an active state while a cell with at least one active neighbour will become (or stay) inactive. Note that this model constitutes a discrete version of lateral-inhibitions continuous models relying on the solving of differential equations to determine cell state [53, 54]. They showed that the lateral inhibition patterns emerging from such model remain fixed and irregular. They also demonstrated how patterns of spots and stripes can be obtained by adding noise (that is to say using this time stochastic rules and inhibition threshold) to the system. Such contact-inhibition patterns can notably be observed in the sensory organs in the notum of *Drosophilia* or in the fly pupae (Fig. 3.2, A). More recently, Gatenby and Smallbone [55] used simple cellular-automata to examine the role of hypoxia and acidosis in transition from in situ to invasive ductal carcinoma.

Cellular-Potts models (CPM) represent each cell as several lattice sites and therefore allow for more realistic cell shapes and integration of more complex processes such as cell

of the state of neighbouring cells.

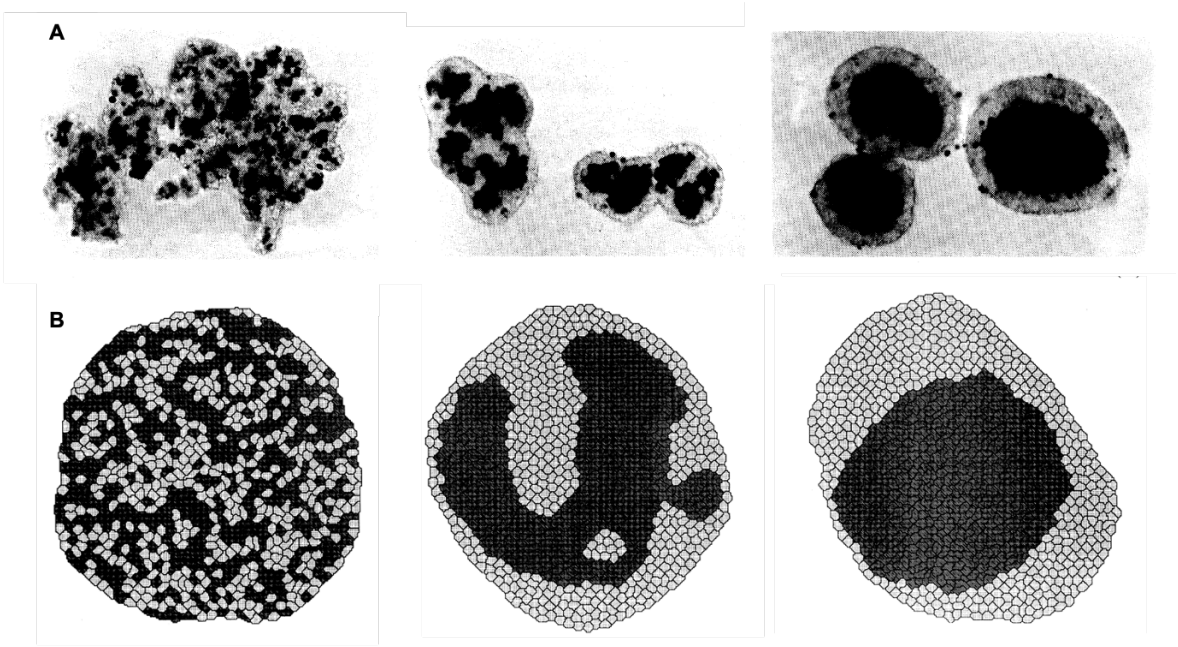


Figure 3.3: (A) Observations by P. Armstrong [56] of cell sorting between pigmented (dark) and neural (light) retinal cells after 5h (random mixing, left), 19h (partial cell sorting, middle) and 48h (complete cell sorting, right). (B) Cell sorting simulations using CPM. Figures adapted from [56] and [57].

membrane tension or cell volume conservation. Here, the resolution of the system is a bit more complex. In fact, the system evolves by attempting to minimise a total internal energy[2]. After applying all the rules governing processes such as cell growth, division or motility, one lattice site and its neighbouring sites are selected and the total energy resulting from copying the site properties to the second is calculated, with a certain probability referred as the temperature (T). In other words, the higher T, the more freely the cell moves. This iteration is repeated for a maximum of $N \times N$ times with N being the number of lattice sites. After each time step, the system configuration with the lowest energy is saved. This method, also known as a Monte Carlo Step (MCS) allows the model to simulate cells random motility and their ability to form and break connections with adjacent cells. Parameters such as adhesion parameters or the ability for a cell to reach its target area (natural or typical area of the cell) are taken into account in the definition of this energy function thus impacting the final configuration of the system. For instance, as early as 1993, Glazier and Graner [57] showed that differential adhesion with fluctuations is sufficient to explain a wide variety of cell rearrangement, using Cellular-Potts modelling. In fact, they showed that simulations of

[2]Note that this method is usually implemented in the majority of cell-based models and will be further detailed in Sec. 3.2.

large heterogeneous (*i.e.* with different cell types) cells population (up to 1000 cells with an area of 40 lattice sites) can exhibit different patterns corresponding to various biological cases only by varying some area constraints and the adhesion between cell types. These biological phenomena range from cell sorting of neural and pigmented retinal cells in chicken embryos (Fig. 3.3) to cell engulfment. Similarly, Walker and colleagues [58] showed that a simple rule-based model of cell behavior, such as contact inhibition of proliferation and migration, was sufficient to qualitatively predict the calcium-dependent pattern of wound closure observed *in-vitro*. In Lattice-gas cellular automata, each lattice site can contain multiple cells. The model tracks the number of cells passing through channels between individual lattice sites rather than the motion of each individual cell (as in the latter two).

Off-lattice models

Off-lattice simulations, with the suppression of a fixed grid geometry, allows for an even more detailed definition of mechanical interactions between cells. In center-based models, each cell is represented by (a) point(s) free to move in space. However, a distinction can be made between these models, once again, according to their spatial resolution.

In Center-based models (CB), also called Overlapping-Spheres (OS) models, only the cell centre is tracked over time and each cell is therefore represented by a single point associated with a specific radius. Cells are viewed as spherical in the absence of neighbours but can deform upon interaction with other cells. In fact, in OS models, two cells are considered as being neighbours if the distance between the two cell centres is smaller than an interaction radius and are overlapping if this distance is smaller than the natural separation between two cells, being the sum of the two radii. As the size of the cells changes over time, the interaction force (*i.e.* the adhesion force) between those cells also varies over time, according to the distance between their centres. Evidently, if this distance is larger than the interaction radius, the expression of their interaction is simply equal to zero [59]. Differential adhesion is implemented here by adding a constant in the adhesion force definition that depends on the nature of the cell interaction (homo- or heterotypical). Smeets and colleagues [60] used this method to develop a cell-based modelling platform for simulating the mechanics of cell expansion in microcarriers[3] (Fig. 3.4). In this study, contact forces (cell-cell or cell-substrate) and external forces on each cell are explicitly determined as previously demonstrated. The motion (*i.e.* the evolution of the cell position over time) of each cell thus depends on these

[3]A microcarrier is a small support matrix allowing for the growth of adherent cells.

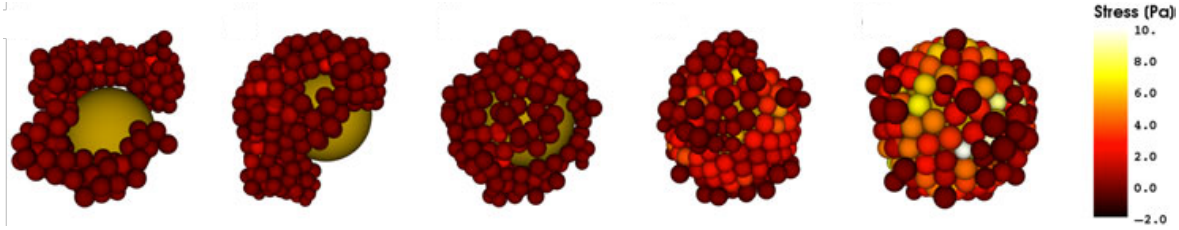


Figure 3.4: OS model used to simulate microcarrier cells (shade of red) expansion on microbeads (central dark yellow sphere) with different cell–substrate adhesion values at the same time point at the end of the simulation. Adhesion parameters value going from $2.5\text{x}10^{-6} J/m^2$ (left) to $40\text{x}10^{-6} J/m^2$ (right). The spatial distribution of the adherent cells as well as the compressive stress change with respect to the value of the adhesion parameter value. Figure adapted from [60].

forces, but also on cell division and on a Brownian motion[4] term. Note that since cells are slowly moving in a viscous environment, characterised by a very small Reynolds number [61], inertial terms are usually neglected in such models. Their model was able to quantify how the mechanical micro-environment of the cells changes during the time of confluency (*i.e.* the covering of the carrier by adherent cells) and their simulations how, even in the absence of biological heterogeneity, geometric properties of the microcarriers and the whole culture system was sufficient to generate a large heterogeneity in mechanical stress.

As for CA or CPM, centre-based models can also be of great use in developmental biology and more particularly in the study of multicellular monolayers. Mosaffa and colleagues [62] proposed a cell-centred model for the simulation of a planar and curved multicellular tissue (Fig. 3.5). They developed an OS model that includes cytoskeleton remodelling (*i.e.* intracellular changes) and stress relaxation due to cell reorganisation (*i.e.* changes in intercellular connectivity). This cell-cell connectivity is computed using a Delaunay triangulation (see Box 2) while cell boundaries are represented using a standard Voronoi tesselation. Voronoi-Tesselation (VT) models show each cells as a set of points, defined by the cell centre and all the associated points, closer to the latter than to every other cell centre. They showed that even with a purely elastic material, the tissue tends to relax due to intercalation process[5] into this layered structure. It is true that the relationship between tissue relaxation and cell intercalation is already experimentally well studied [63]. In fact, it is known that biological tissues exhibit two contradictory but essential properties that are robustness and plasticity, mainly ensured by the cell's cytoskeleton and acto-mysoin complex (see Sec. 3.2 and Fig. 3.10). Robustness allows the tissue to resist external stress while elasticity is es-

[4]Brownian motion is the random motion of particles suspended in a medium.
[5]Reversible inclusion of material, such as molecules of even cells.

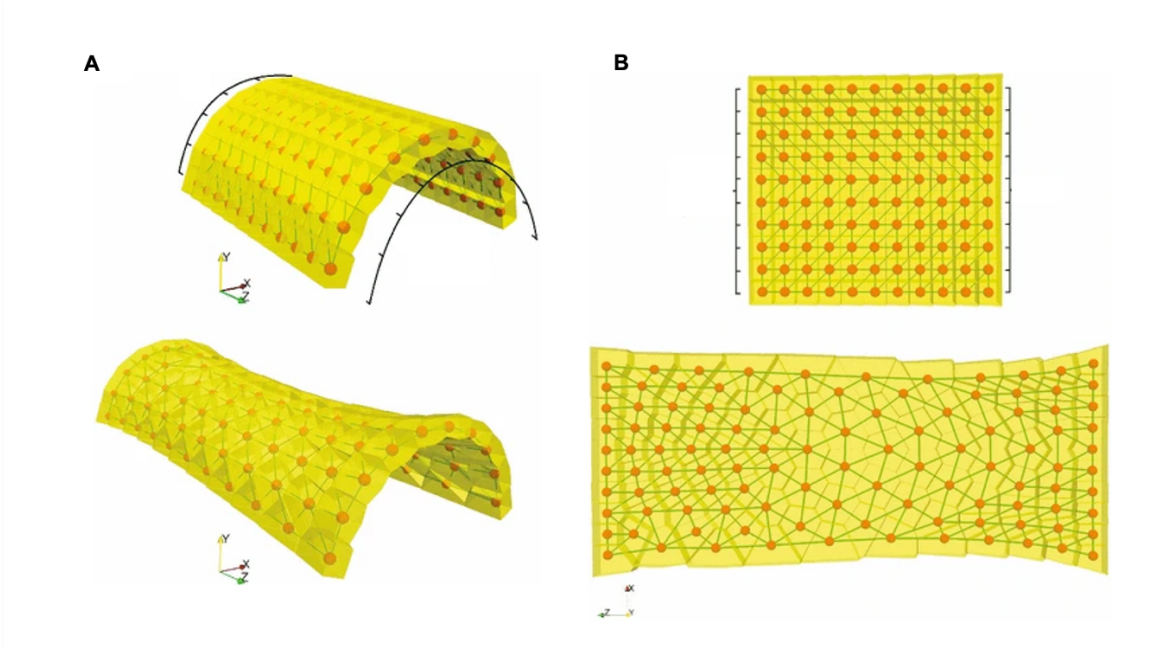


Figure 3.5: Examples of flat and curved monolayers in a VT model. Cells (light yellow) are generated using a Delaunay triangulation for as set of points (orange) and a further Voronoi tesselation for the cell boundary (dark yellow) (A) Flat monolayer. Initial geometry on the top, deformed geometry generated by their numerical simulations on the bottom. (B) Curved monolayer. Initial geometry on the top, deformed geometry generated by their numerical simulations on the bottom. Figures adapted from [62].

sential for tissue remodelling during development or regeneration. Nevertheless, their model could help to quantify this relationship. Moreover, they showed that significant differences can be observed whether the model is applied to planar or curved monolayers. Note that other off-lattice models exist, such as a cluster of Subcellular Element Model (see Fig. 3.1), but are significantly less used and will not be detailed here.

Finally, in boundary-based models, cells are this time defined by their boundary. In Vertex-based models (VM), each cell is represented by a polygon formed by a set of points (vertex, plural vertices) linked together by edges. The shape and position of the cell are hence determined by the position of each vertex over time. Once again, the system evolves (*i.e.* the vertices moves) by minimising a total internal energy. Here, adhesion parameters depend on the length of the edges connecting the vertices and forming the cells. See Sec. 3.2 for an extensive description of VM. Honda and Eguchi (1980) were the first to use these models to study epithelial sheet deformations [65], and for the past decades, an extensive variety of vertex-based models of epithelium dynamics have been developed [66–69]. One of the most notorious example of the latter is the work of Farhadifar and colleagues [70] on epithelial

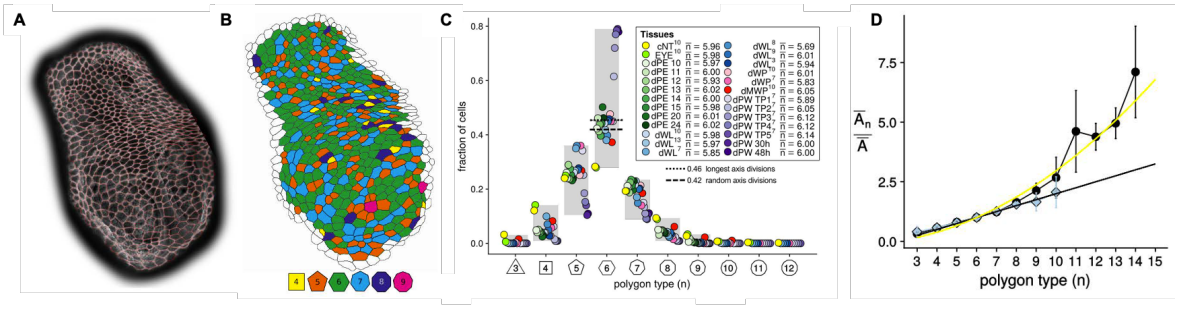


Figure 3.6: Typical epithelial packing from the eye disc peripodial epithelium of *Drosophilia* in vitro (A) and in the VM (B). Cells are coloured according to their number of neighbours. (C) Wide variability in neighbour numbers frequency (shaded grey). Model results are in dashed and dotted line, according to the division axis. Data points for different tissues. (D) Relative average apical area for different neighbour numbers in *Drosophila* wing disc. Control in blue, mutant in black. Lewis' law is illustrated by the black line and its quadratic generalisation by the yellow line. Figures adapted from [64].

packing. They used a vertex-based model to investigate the influence of cell mechanics, interactions and proliferation of epithelial reorganisation. In fact, they showed that their model was accounting both qualitatively and quantitatively for the packing of epithelial cells in the wing disc of *Drosophila*. Moreover, they showed that this model correctly simulates the response to perturbation by laser ablation observed experimentally. Similarly, Iber and colleagues [64] highlighted the emergence of compelling regularities in epithelial cell packing, notably the expression of Lewis's law. This law states that "the average apical cell area is linearly related to the number of neighbours". In other words, they proved that epithelial packing tends to minimise the average perimeter of the cells apical surfaces regardless of the cell shapes and areas. In fact, due to growth and division, cell areas and shapes differ. In VM, this is illustrated by the distribution of polygon types and therefore the number of neighbours per cell, since all the cells are directly sharing their edges with their neighbouring cells. They developed and used the software LIBIBCell [71] to simulate the tissue dynamics and showed that the observed epithelial organisation was indeed driven by a minimisation of the lateral contact between cells, hence their perimeters. They were also able to predict that the linear law could be generalised to a quadratic law with an increase in cells apical areas variability. They were able to experimentally confirm their predictions by generating heterogeneity in cell growth in the *Drosophilia* wing disc.

For an even more precise definition of the cell boundary, front-tracking (FT) models couple a continuous description of a viscous incompressible fluid (representing the cytoplasm of the cell) with the dynamics of separate elastic cells [72]. The boundaries are represented

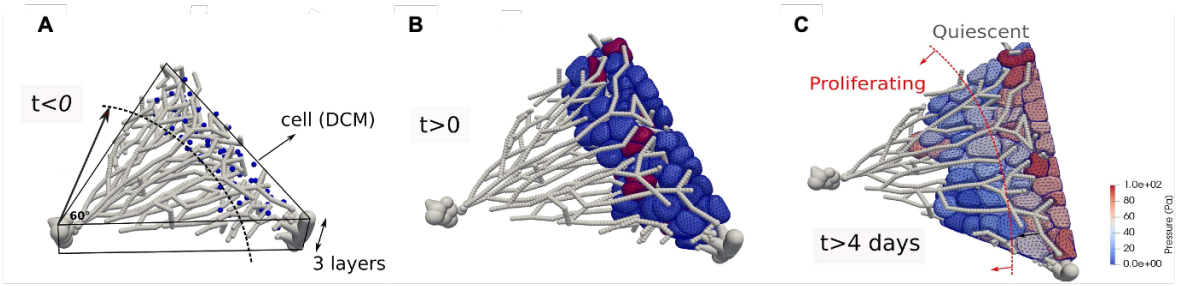


Figure 3.7: Simulations of lesion closure in the liver using a 3D hybrid DCM model. (A) Initialisation, 3 layers of cells growing to a realistic size. (B) Proliferation of hepatocytes. Dividing cells are shown in red. (C) Regeneration process. Cells are shown according to their internal pressure (in Pa). Figures adapted from [73].

by sets of points, connected together to form close polygons, and that move like an elastic membrane immersed in a fluid [71].

Besides these agent-based models, multiscale hybrid models, that consider both cellulars (cell-based) and microenvironmental (continuous) factors simultaneously and interactively are also more and more developed. For instance, they have been used to characterise and simulate tumour growth and invasion in a different environment [74]. More recently, Van Liedekerke and colleagues [73] developed a high-resolution 3D hybrid model to quantitatively understand the impact of biomechanical effects on the closure of lesions in the liver. They consider an example of tissue regeneration by cell division and migration in drug-induced depletion of hepatocytes (cell constituting the main parenchymal tissue of the liver). They actually used a Deformable Cell Model (DCM), which can be considered as a high-definition OS model. In fact, this particular agent-based model represents each cell as a complex 3D object whose surface is triangulated whose nodes are connected to each other by edges (Fig. 3.7). Each node moves accordingly to the equation of motion, in response to internal or external forces. The major difference with center-based model is that all the interactions between cells (or between cells and other simulation objects) are exerted between those surface nodes and not between cell centres, giving each cell in DCM the ability to deform whereas cells in CB tend to keep a more rigid spherical shape. This cell elasticity is critical to simulate the squeezing and migration of liver cell between the dense blood vessel network of the liver (the liver sinusoids, see Fig. 3.7). This hybrid model is therefore designed to simulate the interactions of these individual agents with any other mathematical objects, typically a continuous network of blood vessels. Note that they also coupled the DCM with a classic centre-based model. The latter is useful to simulate large cell population while CB can be used to simulate chosen parts of this system with a higher spatial resolution.

Box 2 : Delaunay triangulation and Voronoi diagram

In mathematics and computational geometry, a Delaunay triangulation for a set of points P in a plane (2D) is a partition to triangles $DT(P)$ that covers the convex Hull of P. The Convex Hull of a shape formed by a set of points is the smallest convex region that contains it. This triangulation is such that no point in P is inside the circumscribed circle[a] of any triangle in $DT(P)$. Different triangulations exist for the same set of points (exponentially increasing with the number of points) but the Delaunay triangulation maximizes the minimum angle of all the triangles, hence trying to make each triangle as close a possible to an equilateral triangle. In other words, the smallest angle in the Delaunay triangulation will always be at least as large as the smallest angle in any other triangulation.

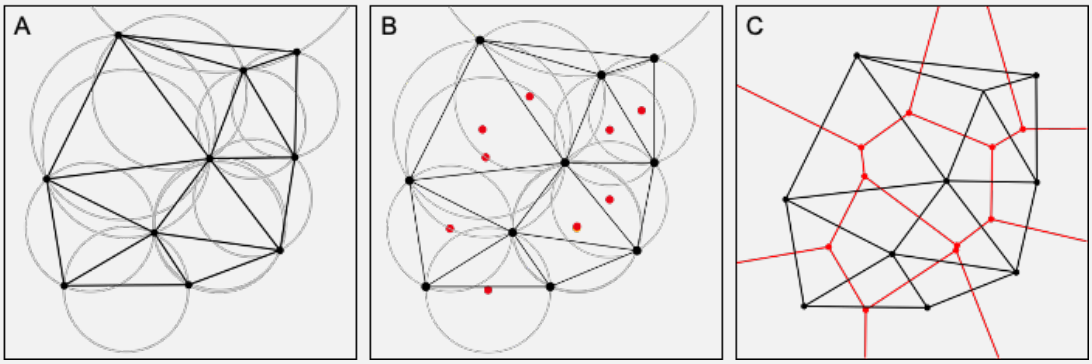


Figure 3.8: Delaunay diagram and Voronoi tesselation. (A) Delaunay triangulation for a set of points (black) with the circumscribed circle (grey) for each Delaunay triangle. (B) Centre of the circumscribed circle (red). (C) Dual Voronoi graph, with each Voronoi cell (red) formed by connecting the circumscribed circle centres.

A Voronoi diagram (or Voronoi tesselation) is a partition of a plane into regions that consist of all points that are closer to a given object in a set than to any other object of this set. In the simplest case, these objects are points on a 2D plane and are called seeds or generators. A Voronoi tesselation eventually generates a set of cells formed by vertices connected together by edges. The Delaunay triangulation of a P corresponds to the dual graph of the Voronoi diagram for P. In fact, the centres of each circumscribed circle in a Delaunay triangulation are vertices of the corresponding Voronoi diagram. The final tesselation shows a set of cells whose edges delimit the region for which any point is closer to a corresponding initial point of P than to any other point of P.

[a]Circle that passes through all the points of a polygon.

We showed here how cell-based modelling can help to understand critical processes involved in various fields such as tissue engineering, cancerology and developmental biology. Moreover, we illustrated how these models can help to design experiments or treatments that can directly benefit patients, highlighting the importance of a constant discussion between theoretical and experimental approaches. It goes without saying that each type of cell-based modelling approach comes with its advantages and inconveniences. *In fine*, one must always weigh the costs and benefits of each model regarding of the context of his study and research issue[6]. For instance, the immersed boundary method allows for a fairly detailed representation of the cell boundary, but is way more computationally costly than any other model and require a larger set of parameters, adding complexity to its calibration. As stated earlier, cellular automata (CA) and cellular Potts models (CPM) are quite efficient when it comes to computation time but can lead to lattice artifacts (*i.e.* when the emerging growth architecture does depend on the lattice symmetry [49]) and can lack in bio-mechanical realism, notably regarding cell division and motility. In CA, one cell can indeed only move a multiple of its size (one site) and it becomes really difficult to integrate processes such as cell size changes. In contrast to CA, CPM consider each cell to cover several lattice sites and therefore could tend to overcome the previous issue. However, there is no clear links between real time and the Monte-Carlo timestepping method used in CPM, and a further calibration of the temperature is usually needed to fit the model evolution with the observed time-scale. Off-lattice models constitute a clear improvement regarding cell shape and position since they allow a continuous evolution of the latter. In fact, VMs explicitly simulate cell adhesion and rearrangement, allowing a wide range of cell shapes and tissue morphologies which is crucial to be physically realistic and required when studying organ morphogenesis. On the contrary, centre-based models are not appropriate to simulate cells whose edges are in contact [75], which is yet the case in most epithelial tissues. And even if Voronoi tessellation (VT) models lean toward more realistic cell shapes than in OS, they still do not allow an explicit control of cell area or volume. In the case of epithelial monolayers modelling, the most commonly used framework is unsurprisingly the two-dimensional vertex model (VM) [61]. In light of these elements, we decided to use vertex-based modelling.

[6]Considerations on five types of cell-based models are summarised by Osborne and colleagues ([44]).

3.2 Vertex models of epithelia

Two-dimensional vertex-based models are off-lattice models in which each cell is approximated by a polygon whose vertices are free to move in response to forces due to growth, interfacial tensions or hydrostatic pressure within each cell (Fig. 3.9).

Equation of motion and differential adhesion

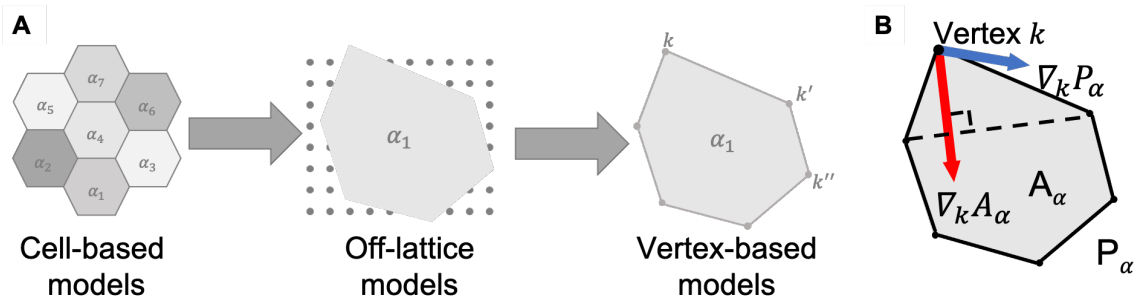


Figure 3.9: Schematic of Vertex Models. (A) Vertex models are cell-based off-lattice models where each cell is defined by a polygon whose vertices can freely move. (B) Forces acting on a vertex k of a cell α with a surface A and perimeter P.

In this framework, edges represent the plasma membrane while vertices biologically correspond to adherens junction where three or more cells meet. Here, we used a 2D VM to simulate the evolution of a 2D cell sheet, representing a cross-section in the developing thyroid. Note that a two-dimensional approximation is usually valid since the forces typically governing cells shapes are known, as shown earlier, to originate from molecules (e.g. actin or cadherins, see Fig. 3.10) located near the apical surface [70]. Therefore, representing only the cell surface in 2D rather than a full cell volume in 3D is sufficient to capture the physical dynamics of the cell. Moreover, 3D models demand more computational resources and 3D images are more complex to obtain. The evolution of a cell shape and position over time is due to a change in the position of each of its vertices. Assuming that random cell movement is reduced by their strong adhesion [76], the evolution over time of the position $\vec{r_k}$ of a vertex k is given by :

$$\eta \frac{d\vec{r_k}}{dt} = \vec{F_k}(t) \tag{3.1}$$

with η being the drag coefficient and $\vec{F_k}(t)$ the total force acting on vertex k at time t. In VM, forces acting on each vertex are derived from the minimisation of the total energy stored in the whole tissue [77]. In other words, this approach evaluates the energy stored

both inside the cells and along their membranes. This energy function (Eq. (3.2)), which is a function of the location of each vertex and closely related to the work needed to deform the cellular network, is used to define an equation of motion that must be integrated over time for each vertex :

$$E = \sum_{\alpha=1}^{N} \frac{K}{2}(A_\alpha - A_0)^2 + \sum_{j:boundaries} (\Lambda_j L_j) + \sum_{\alpha=1}^{N} \frac{\Gamma}{2} P_\alpha^2 \tag{3.2}$$

with K being the elasticity parameter, P_α and A_α respectively the perimeter and area of the cell α, A_0 the cell target area (can either be fixed manually or vary according to specific internal cell processes), Λ the line tension parameter, L the distance between two vertices and Γ the contractility parameter. Considering only the first and third term in Eq. 3.2, their sum gives rise to a force pointing outward due to the elasticity of the cell opposed to contractile forces (pointing inward). It also implies the existence of constraints related to the ability of each cell to undergo elastic deformation emerging from cellular cohesion (Fig. 3.10) since the second term defines a force related to the line tension emerging from cell-cell contacts. Any configuration in a stable steady state corresponds to a local minimum of the energy function defined earlier. Moreover, for a given energy function E, the resulting force acting on each vertex is the opposite of the energy gradient :

$$\vec{F}_k = -\vec{\nabla}_k E \text{ , for each vertex } k \text{ at } x_k \tag{3.3}$$

By computing the gradient (or a representation of how E varies in space), we obtain an explicit definition of the force acting on each vertex:

$$\vec{F}_k = \sum_{\alpha \ni S_k} (-K_\alpha (A_\alpha - A_0)\vec{\nabla}_k A_\alpha - \Gamma_\alpha \vec{\nabla}_\alpha P_\alpha) - \sum_{j \ni S_k} \Lambda_j \vec{\nabla}_k L_j \tag{3.4}$$

where the first term applies to the cells associated with vertex k and the second term with vertices sharing an edge with the latter. One advantage of this model over others is that it permits us to simulate differential adhesion by adding a different weight (*i.e.* different values for Λ according to the cell types, see λ in Table 13.1 in the Appendix) on each force component according to the type of the cell considered and its neighbours. In other words, it allows us to give more or less importance to the force balance between two homotypical cells than, for instance, two heterotypical cells.

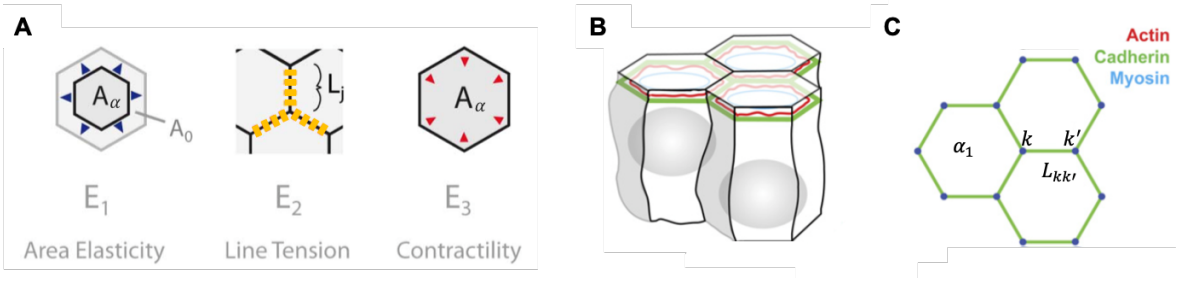


Figure 3.10: (A) Energetic contributions considered in the model. E_1 represents the plasma membrane elasticity, or its ability to withstand deformation constraints (turgor pressure, etc.) and its tendency to reach its target area (A_0, depending on how cellular growth is integrated into the model and often considered equal for all cells). E_2 represents the emerging tensions along the cell-cell boundaries. E_3 represents the membrane contractility, or its ability to withstand contraction constraints due to the activity of the contractile ring (Actin-Myosin complex) [78].(B-C) Schematic of cells in an epithelial monolayer. Cells are bound together by a network of tight junctions. Molecules involved in these intercellular junctions, such as Actin, Cadherins or Myosin, are usually located in the apical surface of the cells. Figure adapted from [70].

Rearrangement, growth and division

While vertex models can easily be related to foam models, there is however a typical and crucial feature of biological structures that is missing: the ability of the cells to grow, to divide and to die. Taking into account these processes is an essential step in the development of mathematical models of such systems [61].

Moreover, cells usually have a number of properties that must be integrated into our simulations. These properties usually correspond to biological or physical quantities (e.g. their surface area, age, or concentration of a certain chemical) that can be computed or initially set and stored. Likewise a cell can be allocated a certain cell type (i.e. a specific label referring to the specificities of the internal behaviour of the cell). More importantly, a cell needs to be assigned with a particular cell cycle, that defines when and how the cell will grow, divide, and die. They stand at the most basic level of the simulation, being one of the core concepts of cell-based models. A wide range of cell-cycle can be applied. Usually, a cell cycle consists of a list or sequence of rules defining when the cell is ready to divide and differentiate. Finally, a proliferative type is also to be applied to the cell object. The cells can be grouped into three general categories with respect to cell proliferation (known as proliferative type); stem cells, transit cells and differentiated cells [80]. Differences in the definition of the cell cycle may apply according to the latter.

Independently of the definition of the force model used for vertex motion, it is critical to

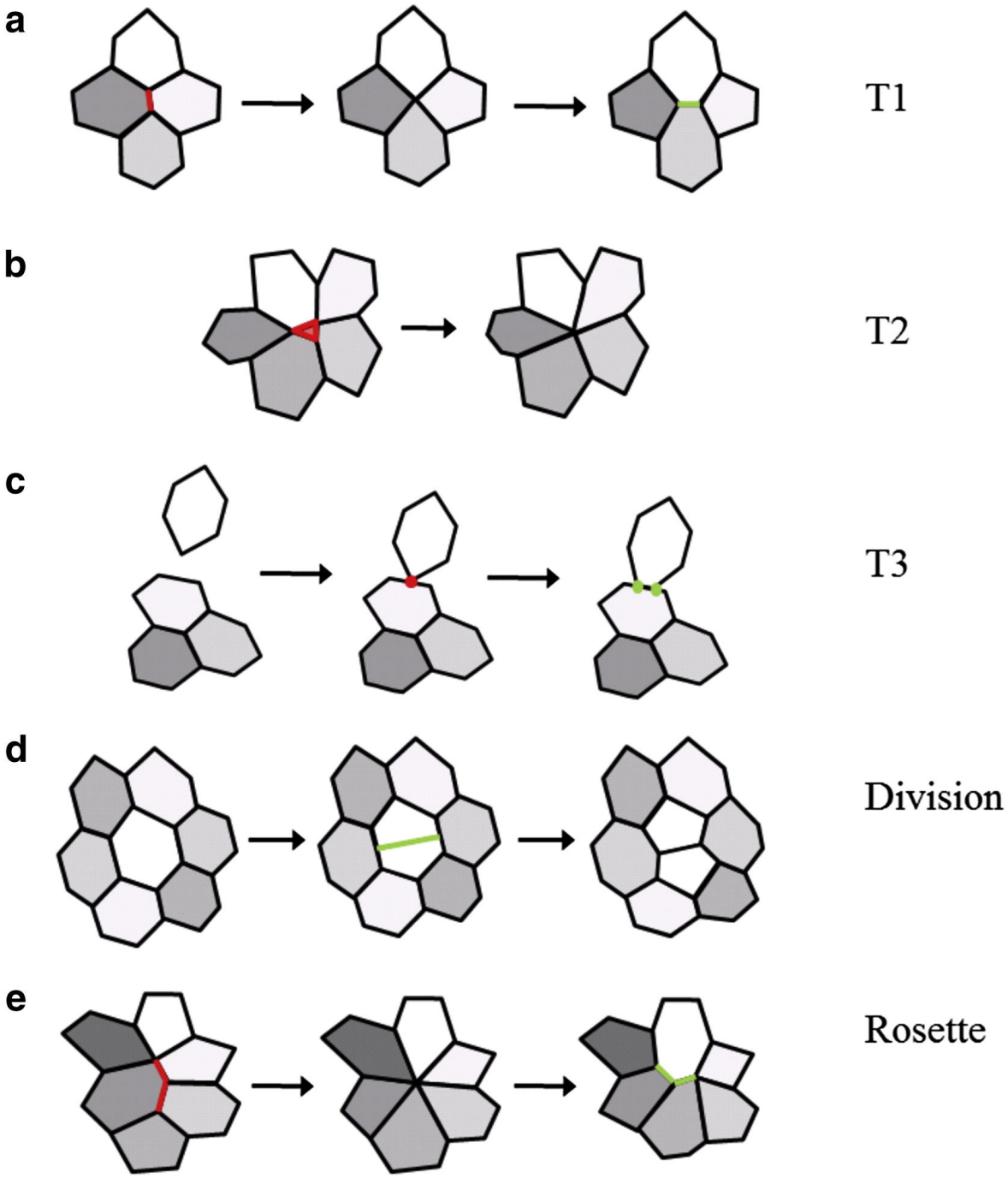


Figure 3.11: Schematic diagram of junctional rearrangements, cell division, and cell removal in vertex models. (a) A T1 transition, in which two vertices sharing a short edge merge into a single vertex, which then decomposes into two new vertices such that the local network topology is changed. (b) A T2 transition, in which a cell shrinks to zero area and is removed, corresponding to delamination and/or apoptosis. (c) Example of a T3 transition, in a vertex/edge intersection is avoided by replacing the approaching vertex with two new vertices that are associated with the element. (d) Cell division, in which a parent cell is divided (in this case through its short axis) through the addition of two new vertices, resulting in two daughter cells. (e) Formation and resolution of multicellular rosettes, a generalisation of T1 transitions involving a larger number of cells. (a–e) Edges or vertices that are removed (red); edges or vertices that are created (green). Figure adapted from [79].

ensure that the elements (or associated cells) never intersect and that they are able to form and break bonds with other elements. Five types of operations, called junctional rearrangements, can be performed on vertices and/or edges to avoid this (Fig. 3.11). T1 swaps, or edge rearrangements, occur when two vertices are located too close to each other, *i.e.* when the distance between those two vertices is smaller than a threshold value. These vertices will be moved and placed at a specific distance $d_{sep} > d_{min}$. This operation ensures that any vertex (that is not on the boundary of the mesh) is only contained in 3 different cells. If this vertex is located on the boundary of the mesh, it must be contained in a minimum of one and a maximum of two elements. T2 swaps, or element removal, allow a particular triangular element to be removed from the mesh if its area becomes smaller than a threshold value. T3 swaps, or element intersection, allow elements to merge together. When a vertex intersects a particular edge, it will then merge with the element containing this edge. This operation can be particularly needed in simulation with voids (such as wound healing) or free moving boundaries. Element division is essential to simulate cell mitosis leading the cell to divide into two daughter cells of equal area. This is achieved by defining a dividing line (being for instance the shortest axis through the cell centroid) and generating two new vertices at the intersection between this dividing line and the mother cell edges, generating two new cells. In some notable developmental processes, one can observe the formation of so-called multicellular rosettes [81]. To allow the formation of this particular event, cell neighbour exchange in T1 swap is replaced by a vertex merging operation when any cell edge length becomes smaller than a specific threshold.

Model implementation in Chaste

All of our following models have been implemented within Chaste , an open-source C++ library developed at the University of Oxford that provides a computational framework for the simulation of (amongst others) cell-based models [82]. It allows the development of models such as those described earlier, as well as the numerical solution of ordinary or partial differential equations (ODEs/PDEs).

Chaste is developed using a test-driven approach, which means that a "test-file" is written in order to perform all the functions written in the "source-code" files. In simple words, the test-file can be seen as a long complex recipe in which each line needs the information on how and when to be executed that are stored in a particular book (a source-code file) in a large library (the source code). This is a one-way reading process, only the test-file has access to

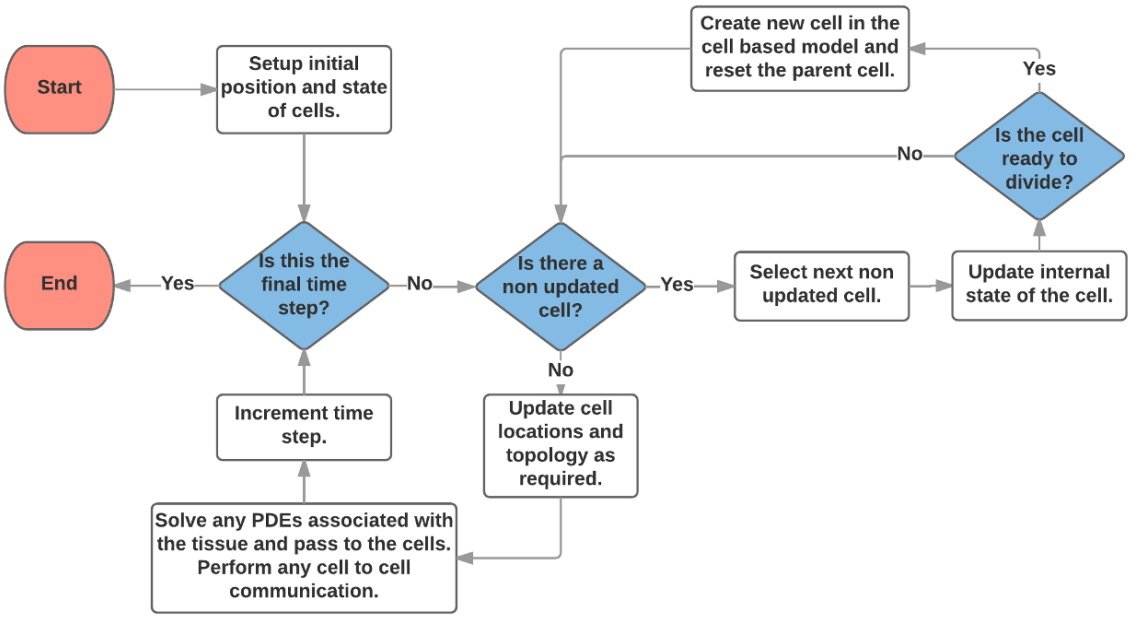


Figure 3.12: Flow chart for cell-based simulation algorithm in Chaste. Figure adapted from [44].

the source-code, but the same source file can be used in different test-files. Test-files usually start by generating a mesh, which comprises a set of "elements" formed by connected vertices. Each element in this mesh is then associated with a cell that possesses a unique index and particular internal properties, all defined in different files. All those cells are associated to form a cell population. A wide variety of source files can then be included. For instance, specific code will be needed to compute the force acting on each vertex, as well as any "active" force (such as motile force). PDE's, used to simulate reaction-diffusion processes, are also defined and solved in specific files. Note that this library is written in a hierarchical way. An initial file is written, defining a general process (in what is called a "class" in C++ language). Each class inheriting from this initial class can use the information already defined in it. For instance, a first-class called "CellCycle" can be written, defining the different cell-cycle stages and the timing of cell division. A second class called "DifferentialCellCycle", one step lower in this particular hierarchy, would then define a different duration for each stage regarding, for instance, the type of the cell.

Two important code components are "writers" and "modifiers". Writers give the ability to the computer to write information generated by the model outside the simulation (*i.e.* in a text file that can be used for post-processing needs). Modifiers allow the files to communicate with each other. As mentioned earlier, only the test file can access the information contained

and generated by source files (except for inheriting classes), but some of the processes defined in the latter may need information generated by other files. For instance, a cell-cycle class using the concentration of a certain morphogen in each cell could be needing the information relative to its concentration, itself generated by another totally independent class. Modifiers are particular classes that will gather the information generated by a particular file and make them available by any other class calling them.

When all the classes are called and all the inputs are given to the simulation, the evolution of the cell population is then simulated by solving the equation of motion over a certain simulation time, and using a defined time-step, for each vertex. A flow chart for cell-based simulation can be found in Fig. 3.12. In practice, the following steps are performed at each time step:

1. The internal (i.e. biological) state of each cell is updated;
2. Rearrangement rules and/or cellular division are implemented;
3. Equation of motion is solved simultaneously for each vertex;
4. The position of all vertices are updated simultaneously.

Importantly, the time taken to perform all these steps scales with the number of cells in the simulated population [61]. Junctional rearrangements are usually the limiting factor in term of computation time when it comes to simple simulations of a growing monolayer. Coupling such models with more demanding processes (*e.g.* diffusion or motility) can however significantly lengthen the simulation duration.

The result of the simulation, *i.e.* the state of the cell population, is exported every n time-step(s) (n being the output multiple) and stored in a specified output directory. They can be then opened in platforms such as ParaView for an interactive visualisation and analysis.

Box 3 : Why hexagons ?

Hexagonal shape is the most common shape used to represent cells in mathematical models. For that matter, the majority of simulations in Chaste (obviously excepting center-based simulations) are initialised with a honeycomb cell population. But this assumption does not come out of the blue. In fact, various examples of hexagonal shapes can be observed in nature (Fig. 3.13) , and it has a mathematical explanation. A sphere (or a circle in 2D) is naturally the geometric form that is the most efficient. In other words, it minimises the surface area for a given volume. That's why bubbles are spherical, or why water takes a spherical shape in a zero-gravity environment. But what happens when you pack bubbles together ? Their membranes fuse and rearrange such as their junctions will always form 120° angles, the same angle that defines a hexagon. Hexagons are indeed the form that will ensure a minimum perimeter for a maximum total area in a single layer packing. In the 19th century, Joseph Plateau (a Belgian physicist) showed that a 120° was also the most mechanical stable arrangement, where tensions on each intersection are all in balance.

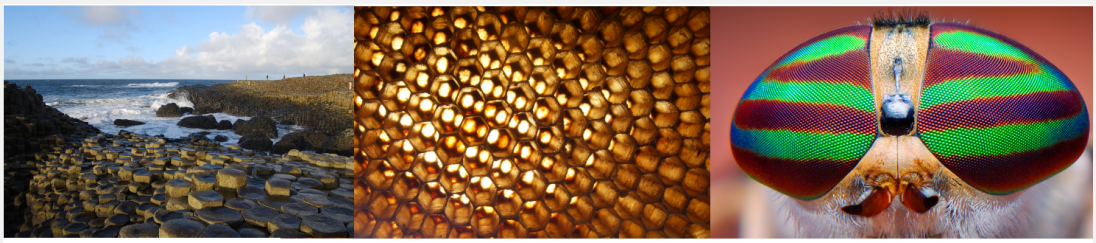


Figure 3.13: Exemples of hexagonal shapes in nature. (A) The Giant's Causeway in Ireland. (B) Closeup of an abandoned Apis florea nest. (C) Dichopic compond eyes of *Tabanus lineola*.

These laws explain why basalt, formed from cooling lava, cracks in hexagonal shapes to release tension in the most efficient way (Fig. 3.13, A). This also explains why honeycomb (Fig. 3.13, B) and fly eyes (Fig. 3.13, C) are hexagonal, using less material to respectively ensure an optimal space needed to grow larvae or maximal light sensing area. In vertex models, among others, the same assumptions are made to represent epithelial packing as hexagonal. Variability in cell area, differential adhesion and external forces will however eventually generate various cell shapes (see 3.1) and the final population hexagonal fraction.

4 | Aim of the work

Despite numerous studies conducted in the field of developmental biology and organogenesis, the precise mechanisms and biological parameters controlling organ formation are still unclear. Recently, Pierreux's laboratory showed that the recruitment and expansion of ECs during thyroid development depend on the expression of VEGFA by epithelial cells [22]. They further showed that the assembly of the basement membrane, partly controlled by endothelial cell invasion in the thyroid, was crucial for epithelial cell orientation, acquisition of polarity and follicle formation [17].

However, hypothesis-driven experimentation using animal models are known to be time-consuming, frequently technically challenging and sometimes not feasible at all. There is thus a growing need to employ complementary approaches to improve our understanding of tissue morphogenesis and identify new parameters. One such approach, computational modelling, offers a way of integrating static observations within a dynamic framework, and has proven a powerful tool aiming to help unlock, among others, complex biological systems [43, 83]. In silico models can be of great interest when it comes to testing and confronting various hypotheses regarding the complex processes at stake, without the limitations of *in vitro* or *in vivo* experiments. To the best of our knowledge, such a mathematical model has not yet been developed for thyroid morphogenesis.

In this context, we sought to develop a 2D vertex-based model of murine thyroid growth, fission, angiogenesis and folliculogenesis. We used computational tools to develop our model step-by-step, focusing on four main processes involved in thyroid morphogenesis and follicles development; cell proliferation, angiogenesis, epithelial cell polarisation and lumen formation. We worked at the interface between experimental and computational biology with these two fields constantly feeding each other. Our *in vivo* data and observations were used to calibrate our model while the results of our simulations highlighted new questions and challenges. We expect that this model could be further used to investigate the mechanisms and biological

parameters controlling the reorganisation of the multi-layered thyroid epithelial primordium into a multitude of single-layered follicles.

5 | Material and Methods

Animals

All mice were of the CD1 strain. Mice were raised and treated according to the NIH Guide for Care and Use of Laboratory Animals and experiments were approved by the University Animal Welfare Committee, Université Catholique de Louvain (2016/UCL/MD/005 and 2020/UCL/MD/011).

EdU injection

Mice were injected with a solution of EdU (Invitrogen) 25mg/ml at embryonic stages E13.5, E15.5 and E17.5. The solution was heated up to 70°C for 1 min to completely dissolve the compound. When at room temperature, the solution was intraperitoneally injected into the mice (100 μg EdU/g of mouse). Mice are sacrificed after 30 min by cervical dislocation. Mouse embryos are microdissected as described in [84].

Embryos fixation, section, and immunolabelling

Mouse embryos were fixed overnight at 4°C by paraformaldehyde 4% before paraffin embedding. Embryos sections were performed every 6 m, collected on Superfrost+ slides and kept overnight at 37°C. For immunolabelling, sections were first re-hydrated by successively xylene (3 x 5 min), isopropyl alcohol (30 sec), ethanol (90%, 70%, 30, 3 min each) and water. Sections were treated to allow target accessibility (un-masking) by citrate buffer 0.01 M at pH 6 while being heated up. Sections are then permeabilised by a solution PBS / Triton X-100 0.3% on an agitator plate. Finally, a blocking solution PBS - / Triton X-100 0.3% / BSA 10% / milk 3% was applied for 60 min. Antibodies, dilutions and immunolabeling protocol are described in [17, 22, 84] (See 13.2in Supplementary Material) . EdU was revealed using the Click-iTTM Plus EdU Alexa FluorTM 488 Imaging Kit (Invitrogen). Sections were scanned with a Panoramic 250 Digital Slide Scanner (3DHistech). For quantification of

EdU-positive and ezrin-positive cells, ezrin-labelled structures and endomucin-positive area, we used the image analysis platform HALO with its appropriate in-built algorithms; Indica Labs CytoNuclear FL for cytoplasmic and nuclear immunofluorescence and Indica Labs Area Quantification for endothelial density quantification. Quantification of the number of epithelial cells per lumen was performed by manual counting on HALO. These data were used in our calibration and validation processes.

All the experimentations described here before as well as the imaging have been carried out by C. Spourquet, P. Lemoine with the help of M. Baudoin and myself, and under the supervision of C. Pierreux.

Computational and statistical analysis

Statistical analysis and data pre-processing for the generation of our initial conditions were performed in MATLAB (MathWorks®). Data are presented as mean ±s.d. Our model is written in C++ using the open-source libraries Chaste. Simulation results are visualised in the open-source, multi-platform data analysis and visualisation application ParaView.

All the data management and analysis, as well as the model development, have been performed by myself, with the help of M. Baudoin and L. Lepers, under the supervision of E. Hanert.

6 | Epithelial Proliferation and Initial Conditions

6.1 Introduction

One of the crucial features of VM of epithelia is the ability for cells to grow and divide. These operations modify tissue connectivity and tend to increase the computational complexity but are nonetheless essential. In VM, when a cell undergoes division, it splits into two cells of equal area [85]. This is achieved by generating a mitosis axis and placing two new vertices at the intersection of this dividing line and the mother cell perimeter (Fig. 3.11, d). A large majority of VM that includes cell division, also called cell proliferation, is usually based on the assumption that cell division occurs stochastically [61]. For instance, one could assign a random time drawn from a normal distribution around a specific division time, and allow the cell to divide when this time is reached. Some models use more detailed description of the cell cycle, in which the duration of the different cell phases are modulated, for instance, by the concentration of certain proteins. Here, we experimentally quantified thyroid epithelial cell proliferation using EdU staining on mice embryo sections at E13.5, E15.5 and 17.5 to determine a specific epithelial cell cycle duration. We also used computational tools to precisely map proliferating cells on the sections and verify the existence of two "populations" of dividing cells, that are the central and peripheral cells[1].

6.2 Methods

We initially used BrdU (5-bromo-2'-deoxyuridine) staining to detect proliferating cells but rapidly switched to EdU (5-ethynyl-2'-deoxyuridine). While they are both thymidine analogs,

[1]Some early studies suggested the emergence of differential proliferation pattern between central and peripheral cells in the pancreas [86], the thyroid [87, 88] or the bovine corneal epithelial [89, 90].

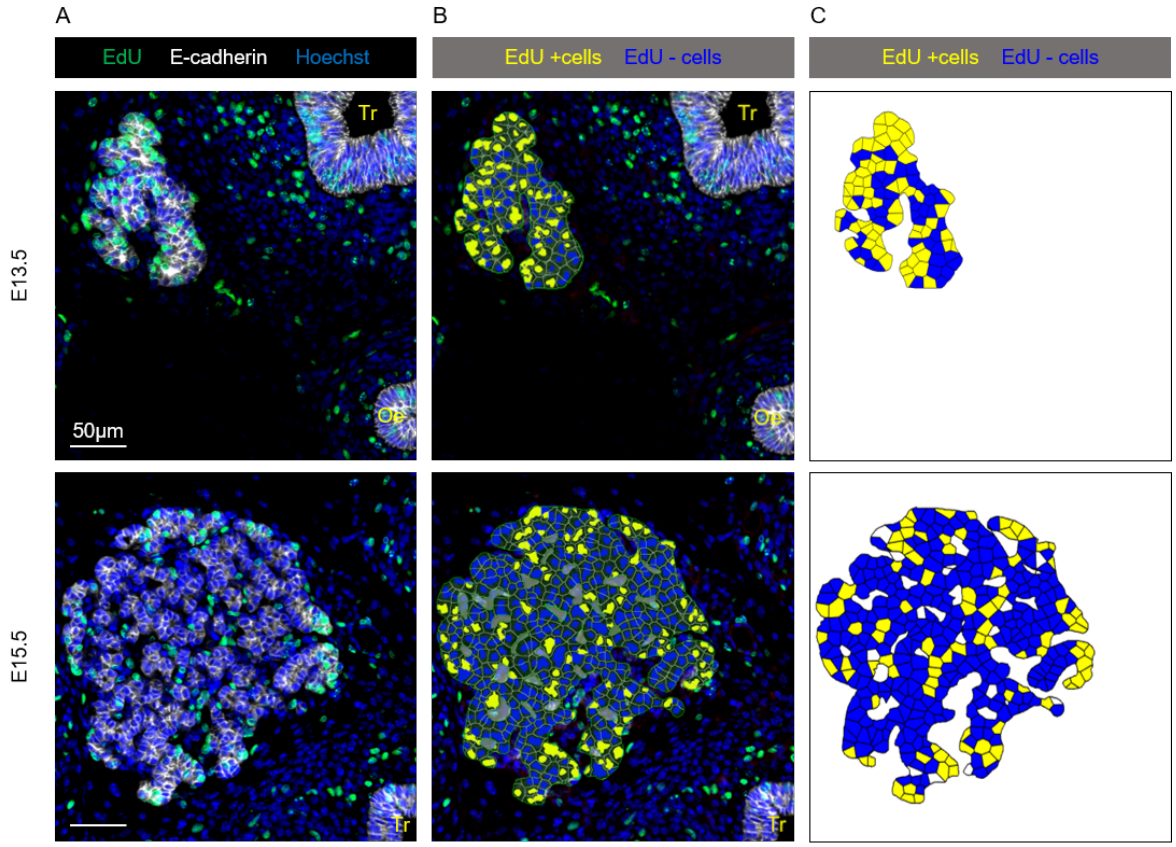


Figure 6.1: Proliferating cell detection in a thyroid section of a mouse embryo at E13.5 (top) and 15.5 (bottom) (A) Epithelial cells are shown in white (E-cadherin), nuclei in blue (Hoechst) and proliferating cells in green (EdU). Tr, trachea; Oe, oesophagus. (B) EdU^+ cells are detected and marked in yellow, EdU^- cell in blue. The green annotation delimits the analysis region. (C) The analysed tissue is recreated using quantitative data generated by HALO. This will allow us to differentiate central cells from peripheral cells for further investigation. As in B, EdU^+ cells are marked in yellow, EdU^- cells in blue.

BrdU tended to generate blurry images with high background noise. Several reasons could explain this. First, the detection mode differs between the two compounds; BrdU is revealed using sequentially a primary and a secondary antibody, with high molecular weight, whereas EdU is detected by a fluorescent azide 500 times smaller than the antibodies [91]. The tissue penetration of the antibodies is therefore significantly lower than for the azides. Moreover, the size of the antibodies could generate steric hindrance situations leading to a loss of antibodies detection. Finally, the fluorescent azide is highly specific to EdU, therefore limiting interference with other similar molecules.

Images from different embryos ($n = 3$) were analysed at different embryonic stages with the software HALO (Fig. 6.1). We used a Hoechst staining to detect all cell nuclei, E-cad for the epithelial cells, endomucin for the endothelial cell and EdU for proliferating

cells. HALO uses nuclei detection (Hoechst staining) in order to detect the cells and further uses the E-cad staining to determine the cell outline. Manual annotations allow specifying a particular region to analyse. Cells are further marked by HALO according to their EdU staining level (either positive or negative). A prior calibration of the stain detection threshold is however needed to ensure that the information given by the detection algorithm matched the observations. This calibration is made once on a single image for each embryonic stage. The potential bias is therefore only due to the subjective judgment of the user and considered acceptable. Moreover, if there is a bias, it would remain consistent for all analysis. Note that for quantitative data regarding the number of cells detected by HALO, a manual verification has also been performed. Four colleagues were asked to manually count the number of epithelial cells on a thyroid section at E13.5. The number of cells detected by HALO was inside the confidence interval ($\alpha = 0.05$) computed from the four manual counts.

6.3 Results

The results obtained with HALO allow for a more quantitative description of epithelial cell proliferation in the thyroid (Fig. 6.2). A notable decrease can be observed in the fraction of proliferating cells (EdU^+) over time. At E13.5, around 45% of the cells are proliferating. At E15.5, this percentage drops to around 20% and, at E17.5, only slightly higher than 10% of the cells are proliferating. These observations are coherent with the dynamics of thyroid development shown in Fig. 2.6. One can assume that the proliferation rate will keep on decreasing until birth, which happens around E19 in mice when cells are almost fully differentiated. In fact, our computational analysis revealed that proliferation rate tends to follow an exponential decrease over time (Fig. 6.2, B), reaching values close to zero at birth. Post-natal EdU injection should be performed to confirm this trend.

We further investigated the potential proliferation patterns emerging between what we call "central" and "peripheral" cells. Peripheral cells are all the cells that are in immediate contact with the periphery of the epithelial mass (*i.e.* the basement membrane and the ECM). Not surprisingly, central cells are all the remaining cells in the centre of the mass. In order to achieve such analysis, we needed to differentiate peripheral from central cells in the scanned images analysed by HALO. Besides quantitative information on cell staining, it also generates information relative to the position of each cell in space. More precisely, it gives the minimum and maximum location (in the two-dimensional space defined by the

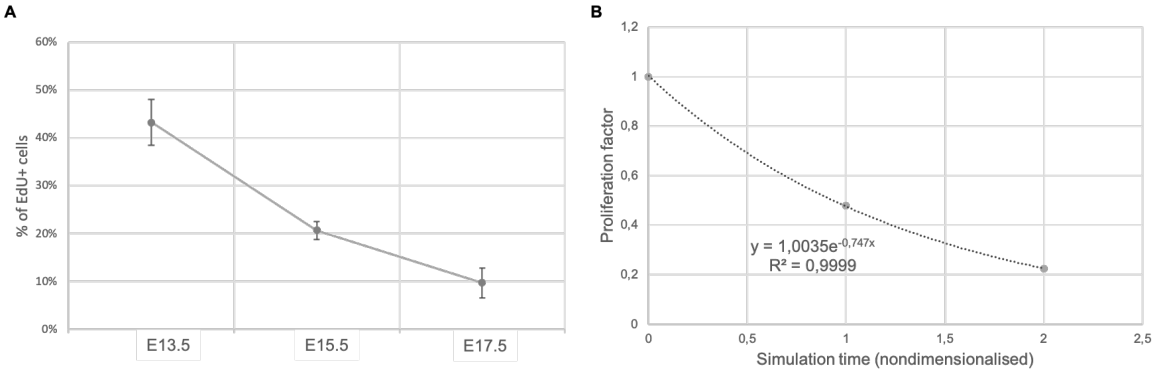


Figure 6.2: Evolution of epithelial cell proliferation over time. (A) Fraction of proliferating cells in the thyroid of the mouse embryo at different embryonic stages ($n = 3$). (B) Cell growth dynamics of epithelial cells in the embryonic thyroid. Both axis are nondimensionalised by respectively dividing the percentage of EdU^+ cells by the highest value (here at E13.5) and replacing the embryonic days by the time elapsed between two EdU injections (here 48h).

whole section) of the four edges of the rectangle enclosing each cell (Fig. 6.3, A). These data (along with all the other quantitative information) can therefore be exported and further used to recreate every rectangle associated with every cell (Fig. 6.3, B). The centre of each rectangle is calculated and subsequently used as Voronoi seeds (see Box 2 and Fig. 6.3, C). Even so, Voronoi diagrams still lack the ability to close the peripheral cells if no boundary is given to the algorithm. To do so, the position of the annotations previously used to define the whole analysed region is here given as boundary for the Voronoi tessellation (Fig. 6.3, C). Note that similar annotations can be used to determine excluded regions (*i.e.* regions where no analysis is to be performed) around endothelial cells or inter-epithelial spaces. Here, peripheral cells will be defined as an element (a cell) with at least one of its edges (linking two vertices together) that is not shared with any other cell, i.e. only belonging to that cell. A central cell will be surrounded by other epithelial cells with whom it will share all of its edges. In this way, the tissue used in our experimental immunolocalisation analysis can be entirely recreated, while keeping all the information generated by HALO (Fig. 6.1, C).

These data were used to distinctly determine the cell proliferation dynamics in central and peripheral cells. We observed that the dynamics of the two cell populations were rather similar. On the contrary, the fraction of cells proliferating differs between the two populations (Table 6.1). In fact, more than 60% of peripheral cells are proliferating at E13.5 compared to less than 40% for the central cells. At E15.5, it is almost 30% of peripheral cells that proliferate as compared to less than 10% in the centre. At E17.5, the proliferation rate is a bit above 10%. Note that we followed the assumption that at E17.5, all the cells are

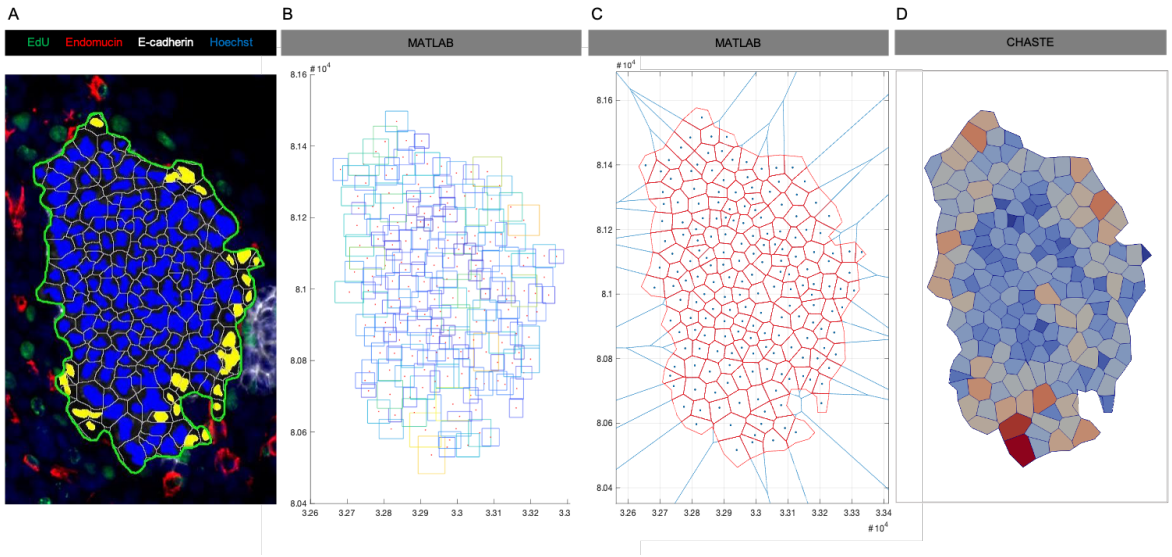


Figure 6.3: Process of initial condition generation from microscopic imagery. (A) E13.5 thyroid section labelled for E-cadherin (white), Endomucin (red), EdU (green) and nuclei (blue). Epithelial cell nuclei were coloured in blue and endothelial cell nuclei in yellow using HALO. The area of analysis is delineated by a green line (B) Rectangles enclosing epithelial and endothelial cells identified in (A) were recreated in MATLAB and the centres of each of these rectangles were generated (in red). Contours of rectangle are coloured according to their area (increasing from blue to yellow). (C) Voronoi diagram was generated in MATLAB using cell centres from (B) as seeds. Original tesselation in blue, final tesselation using the manual annotation of the area of analysis (A, in green) as boundary in red. (D) Visualisation of the final cell population in Chaste that will be further used as initial conditions in our simulations. Cells are coloured according to their area (increasing from blue to red).

considered as peripheral as practically all the mature follicles are formed and the cells are therefore supposedly all in contact with the ECM, explaining the fraction of 1 for peripheral cells in Table 6.1. This assumption is confirmed by our *in vivo* observations. Altogether, these results indicate that proliferation preferentially occurs in the periphery of the thyroid mass during folliculogenesis. Those results also illustrate the crucial transition from a mass to a multitude of monolayers previously mentioned, as the fraction of peripheral cell rapidly increases from 0.4 to 1.

Importantly, this data eventually allowed us to calculate a typical cell cycle length for the two cell populations, that is to say the time elapsed between two cell divisions, further used in our simulations. In order to do so, we also needed an estimation of the total number of cells in the embryonic thyroid at the different stages. We used a method derived from Cavalieri's principle. This method is defined in Box 4. The total number of cells is calculated for both lobes and those values are added up to obtain the final grand total (Table 6.2). To reach the number of cells observed after a certain period of time, cells need to proliferate.

Table 6.1: Ratio of proliferating cells in peripheral and central populations in three embryonic stages of the mouse thyroid ($n = 3$). In italic : fraction determined according to our initial hypothesis.

	Embryonic day		
	E13.5	**E15.5**	**E17.5**
Peripheral cells (1 = whole tissue)	0.39 ±0.06	0.55 ±0.09	*1*
Total EdU+ cells (1 = whole tissue)	0.43 ±0.05	0.21 ±0.02	0.11 ±0.03
Peripheral/EdU_+ cells (total peripheral cells)	0.63 ±0.11	0.29 ±0.05	0.11 ±0.03
Central/EdU_+ cells (total central cells)	0.38 ±0.07	0.09 ±0.02	*0*

Knowing that a single cell will always divide into two daughter cells, the cell population will grow exponentially as follows:

$$N_f = N_i 2^n \tag{6.1}$$

with N_i being the initial number of cells (here at E13.5), n the average number of cell generations and N_f the final number of cells (here after 48h, at E15.5).

Using the assumption that the initial cell population can be isolated in two distinct populations, we have :

$$2^{n_p} R_p N_i + 2^{n_c} R_c N_i = N_f \tag{6.2}$$

and therefore :

$$2^{n_p} R_p + 2^{n_c} R_c = \frac{N_f}{N_i} \tag{6.3}$$

with R_p and R_c being the initial fraction of peripheral and central cells (here at E13.5), n_p and n_c the number of peripheral cell generations. We also know the fraction of peripheral (hence central) cells and their respective proliferation rate from our experimental results, which can be defined as follows:

$$n_p = \left(\frac{R_{p,EdU^+}}{R_{c,EdU^+}}\right) \cdot n_c \tag{6.4}$$

In fact, this fraction is the mathematical expression of how much more (or less) the peripheral cells proliferate compared to central cells.

Eq. (6.3) can therefore be re-written as follows:

$$2^{\left(\frac{R_{p,EdU^+}}{R_{c,EdU^+}}\right) n_c} R_p + 2^{n_c} R_c = \frac{N_f}{N_i} \tag{6.5}$$

Knowing R_p, R_{p,EdU^+}, R_c, N_i and N_f, we can numerically solve Eq. 6.5 to calculate n_c (and then n_p).

We can then compute the duration (in hours) of the typical cell cycle for both peripheral and central cell populations, respectively t_p and t_c, by solving the following equations:

$$t_p = \frac{t_e}{n_p} \tag{6.6}$$

$$t_c = \frac{t_e}{n_c} \tag{6.7}$$

with t_e being the time elapsed between N_i and N_f, here 48h. Eventually, we obtained a typical cell cycle duration of 11 hours and 20 minutes for peripheral cells and 31 hours and 52 minutes for central cells.

These estimations have been used in our simulation as a typical cell cycle duration for the two epithelial "cellular types". It is important to remember that in this framework, the cell "type" will always refer to a "computational" label assigned to the latter and not only to its morphological or phenotypical type. A single cell can therefore be assigned multiple labels. For instance, a peripheral epithelial cell will be labelled as "epithelial", but also "peripheral". A chart illustrating the various cell types in our simulations can be found in Fig. 6.4.

In our model, all epithelial cells are characterised by a phase-dependent cell cycle. This means that each cell will only divide when it has reached the end of the mitotic phase. The duration of each phase is set as a fraction of the total cell cycle duration, considering that for a rapidly proliferating human cell with a total cycle time of 24h, the G_1 phase (first growth phase) typically lasts about half of the cycle, the S phase (DNA replication) about a third of the total duration, the G_2 phase (gap phase between DNA replication and cell division) about 4h ($\frac{1}{6}$th) and the M phase (mitosis) about 1 hour. This cell cycle is defined in Chaste and is applied to any epithelial cell [80], according to their position in the cell population. In fact, at each time step, each cell position (*i.e.* peripheral or central) is updated (using the same method as defined earlier) and the correct label is then applied. The total cycle time was calculated using quantitative data gathered at E13.5 (and E15.5 for N_f). However, we showed

Table 6.2: Estimated total number of cells in the embryonic thyroid at three embryonic stages using the modified Cavalieri method (n=3).

	Embryonic day		
	E13.5	**E15.5**	**E17.5**
Total number of cells	8073 ±2822	61491 ±5484	216830 ±6843

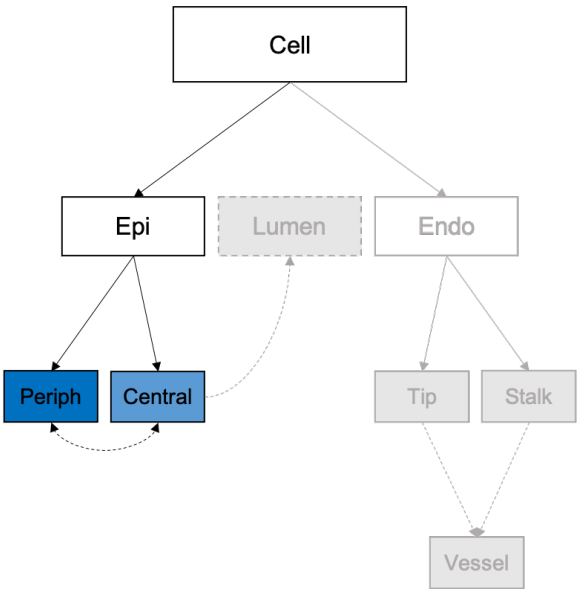


Figure 6.4: Hierarchical chart illustrating the different cell types in our simulation. Simple arrows represent the course of cell type assignment upon division (or at $t = 0$). At the start of the simulation, each element is assigned to a cell object. This cell is immediately given a cell type label, either "Epi" (for epithelial cells) or "Endo" (for endothelial cells). Similarly, a label referring to its positions is added for each epithelial cells, either "Periph" (for peripheral cells) or "Central" (for central cells). As cells move in the population, their position may change during the course of the simulation. Cells can therefore switch from peripheral to central (and vice-versa) throughout the course of the simulation. Plain lines corresponds to cells types defined *pre*-simulation, dotted lines to types allocated *in*-simulation.

that cell proliferation rapidly drops throughout embryonic thyroid development. To account for this exponential decrease, a "generation factor" was added to the cell-cycle definition. Each time a cell divides, a new cell-cycle duration is applied to the latter, which is multiplied by this generation factor, calculated using the non dimensionalized epithelial growth dynamics (Fig. 6.2, B). This ensures that for each new generation, cells will proliferate less and less rapidly until they differentiate and stop dividing. Note that our simulations typically stop at E19.5, *i.e.* around birth.

Another parameter governing the dynamics of cell growth and division in our model is what is called the "target area", that is to say the size (area in 2D, volume in 3D) that cells will tend to reach (A_0 in Eq. 3.2). The area of both the cells and their nuclei is given by HALO, and can therefore be used for analysis. A summary of this information can be found in Table 6.3. We can see that the mean area of epithelial cells seems to be rather constant through development even if a slight decrease can be observed between E13.5 and E15.5, both in peripheral and central cells. This could suggest that thyroid growth is mainly due to cell proliferation rather than cell growth, which is again consistent with the dynamics of primordium growth and development described in section 2.2. Despite this slight variation in the mean cell area, we made the assumptions that the target area in our model would be equal for all epithelial cells and that it would remain constant during the whole simulation. An interesting note can however be made regarding the position and proliferative status of a cell

Table 6.3: Mean area of epithelial cells in the embryonic thyroid at three embryonic stages ($n = 3$).

	Embryonic day		
	E13.5	**E15.5**	**E17.5**
Mean epithelial cell area [μm^2]	77.68 ±2.90	73.97 ±3.21	82.30 ±5.6
Mean peripheral cell area [μm^2]	83.31 ±2.5	82.7 ±2.48	82.30 ±5.6
Mean central cell area [μm^2]	74.13 ±2.22	63.14 ±1.89	n/a ±n/a
Mean peripheral EdU$^+$ cell area [μm^2]	84.74±2.51	96.49 ±2.89	89.16 ±6.07

at a particular embryonic stage. In fact, peripheral cells appear to be larger than central cells at both E13.5 and E15.5. This could be explained by the simple fact that cells at the periphery of the epithelial mass are in a less densely populated area (*i.e.* with fewer neighbours) and hence are likely to take up more space than central cells. Moreover, proliferating cells are larger on average than non-proliferating cells, which is also rather excepted. In fact, EdU only marks cells that are undergoing DNA replication (S phase) which is the last phase before mitosis and cell division. It is also fair to assume that the majority of the cells that are not EdU$^+$ are in the growing (G) phase. In fact, Ki-67 labelling performed in our lab showed that all epithelial cells are positive for this marker, and are thus in an active phase of the cell cycle.

As mentioned earlier, one can use simple honeycomb-shaped initial cell populations (Fig. 7.2, A) for the simulations, generated by in-built Chaste functions. However, more complex initial conditions (IC) can also be used. In Chaste, it is provided as a list of nodes associated with their positions, and a list of cells associated with the corresponding nodes. Our images were therefore also used to generate realistic IC for our simulations. In fact, the function used to generate the bounded Voronoi tesselation returns two distinct structures. The first gathers the information relative to the cells. Each cell is linked to a specific ID as well as to a list of all the cell's corresponding vertices IDs. The second structure simply lists all the nodes (vertices) with their location in space (in 2D). Note that post-processing is needed before exporting the final files. In fact, the cells generated can contain a large number of vertices. While it is perfectly fine to represent cells within this software, Chaste does not cope well with a large number of vertices close to each other. Despite the implementation of specific functional rearrangements (see Section 3.2), it could still lead to simulation errors. To avoid this, a number of vertices that are not at the intersection between two or more cells are deleted and the cells and nodes structures are updated. These tables are then written in text files that can be used in Chaste to generate realistic initial conditions (Fig. 6.3, D).

Box 4 : How to estimate the total cell numbers in the tissue ?

The method we used to determine the number of cells in the total tissue at each embryonic stages is based on a principle of geometry called the "Cavalieri's principle". It states as follows: "Suppose two regions in 3D space (solids) are included between two parallel planes. If every plane parallel to these two planes intersects both regions in cross-sections of equal area, then the two regions have equal volumes". The Cavalieri method is often used as a stereological tool to determine tissue volume or absolute cell count [92]. Here, we used the different sections as parallel planes and knowing the number of cells in each section and considering a typical cell height of 6μm, we could determine the number of cells in the whole tissue. Concretely, this total number of cells can be computed using the following formula:

$$\text{total number of cells} = \sum_{i=1}^{n} \left(\frac{(n_i + n_{i+1}}{2} \cdot \frac{a_i}{b_i} \cdot \frac{b_i}{H}\right) \tag{6.8}$$

with n_i being the number of cells on section i, n_{i+1} the number of cells on the subsequent section $(i+1)$, a_i the distance between section i and $i+1$ (here 60 μm), b_i the width of a section (here 6 μm) and H the typical cell height. In other words, we calculate the mean number of cells between two sections and integrate this value over the whole tissue (*i.e.* all the sections), taking into account the width of a section and the distance between two sections (or "how many sections do I have between two measurements") as well as the typical height of an epithelial cell (or "how many cells do I typically have within the width of one section"). This method obviously relies on several assumptions, notably that epithelial cells take the form of a cuboid and that all our sections would always cut epithelial cells alongside their apical surface.

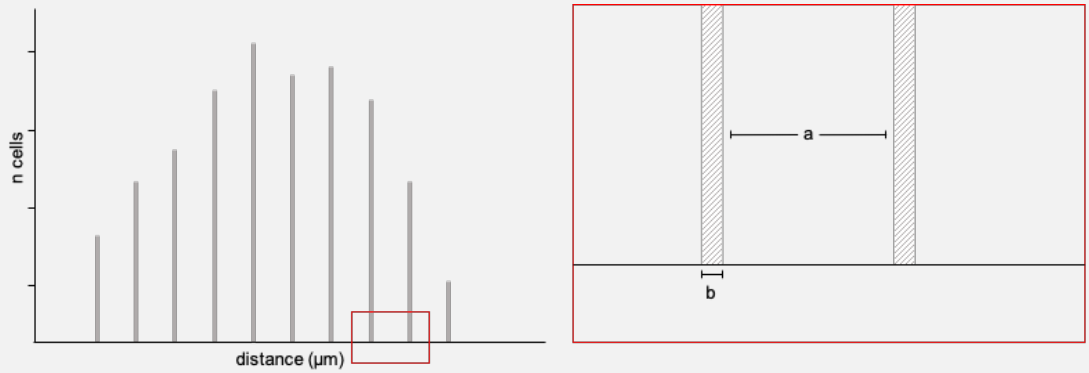


Figure 6.5: Illustration of the modified Cavalieri method used to determine the total number of cells in the embryonic thyroid at different stages.

7 | VEGF Diffusion and Endothelial Motility

7.1 Introduction

We showed how endothelial cells have a central role in embryonic thyroid morphogenesis and it was therefore critical to include them when simulating thyroid morphogenesis. Moreover, in order to mimic the dynamic remodelling of endothelial cells during thyroid morphogenesis, our model must be able to simulate endothelial motility. It has been previously detailed how the diffusion of the growth factor Vascular Endothelial Growth Factor (VEGF) can induce endothelial tip cell migration hence new blood vessel formation. In our model, this was achieved by solving partial differential equations (PDEs). PDEs are used to mathematically describe the advection, diffusion and reaction of a component over time and space. In fact, it has been demonstrated how morphogen diffusion can be described by the processes of molecule production in an initial layer of cells, molecule spreading non-directionally and degrading at a certain rate, and this independently of the actual biological mechanisms involved [93].

7.2 Methods

Here, we used a PDE solver defined in Chaste to solve a continuous model of morphogen transport [94] with a source term. This differential equation allows us to calculate the concentration of a morphogen throughout the tissue and at any time, according to a specific production and degradation rates as well as to a diffusion coefficient. The model equation reads :

$$\frac{\partial c}{\partial t} + \vec{w} \cdot \vec{\nabla} c - \vec{\nabla} \cdot (D\vec{\nabla} c) = f(\vec{x}) - k_d c \text{ in } \Omega(t) \tag{7.1}$$

with $c(\vec{x}, t)$ being the concentration of morphogen at location $\vec{x}$ and time t, the vector field $\vec{w}$ designating the velocity of the cells moving in the tissue, $f(\vec{x})$ a function defining the rate of production of c and D and k_d being respectively the diffusion and degradation coefficients. This equation is solved on the domain $\Omega(t)$.

Concretely, its solution is calculated at each time step for each point on this domain. In our simulation, this domain is defined by a computational mesh corresponding to a triangulation of the vertex mesh, that is to say by dividing each polygon of this mesh (*i.e.* each cell) into smaller triangles, formed by the vertices defining this polygon and its centre (Fig. 7.2, D). In other words, this PDE is solved on each vertex of the VM mesh together with the centre of each cell moving at a speed $\vec{w}$, and at each time step. Note that these parameters for the VEGF kinetics were not calibrated since it is experimentally difficult to measure how quickly VEGF diffuses through the tissue (notably due to the existence of different VEGF isoforms with different ECM binding capacities). However, the VEGF concentration quickly converges towards a steady state where there is a balance between production at the cell level and loss on the tissue boundary. We assumed here that the morphogen is initially produced by all epithelial cells and that the boundary conditions (*i.e.* the imposed solution at every point on the boundary of the domain) applied are in the form $c(t) = 0$ (Dirichlet boundary condition) to ensure the formation of a gradient. This amounts to saying that the morphogen leaves the tissue (or is degraded) on the boundary, which causes its concentration to drop to zero. The solution is approximated using a Forward Euler discretisation in time and a linear finite element approximation in space. The morphogen gradient is also computed for each cell (being the average gradient at each node defining the cell) and stored. This gradient is precisely used to induce and direct endothelial cell motility. In our framework, we defined two cell types in addition to an "endothelial" type: "tip" cells and "stalk" cells (see Fig. 7.1). The implementation of cell motility also requires the definition of an "active" force, in contrast with the "passive" cell force resulting from solving the equation of motion. In Chaste, forces are defined as vectors with 2 components (in 2D). Here, this active force is applied on every vertex (k) of any cell labelled as a "tip" cell. This vector has the exact same direction as the direction of the average gradient on every cell neighbouring a given tip cell (*i.e.* towards the highest VEGF concentration). For a tip cell α and with a morphogen gradient $\vec{\nabla}_\alpha c$, this force can be expressed as follows :

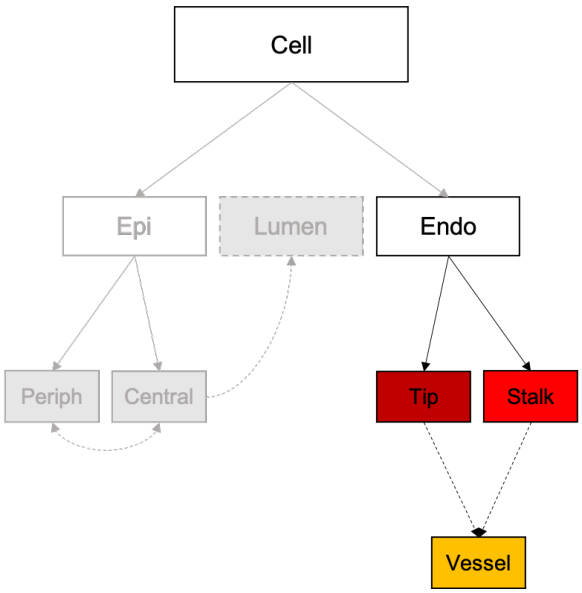


Figure 7.1: Hierarchical chart illustrating the different cell types in our simulation. Simple arrows represent the course of cell type assignment upon division (or at $t = 0$). At the start of the simulation, each element is assigned to a cell object. This cell is immediately given a cell type label, either "Epi" (for epithelial cells) or "Endo" (for endothelial cells). Similarly, a label referring to its motile status is added for each endothelial cells, either "Tip" or "Stalk". As two tip cells meet each other, they lose their ability to move, hence their "tip" type and are instantly labelled, along with all the following stalk cells, as "vessel", allowing to generate stable endothelial networks. Plain lines corresponds to cells types defined *pre*-simulation, dotted lines to types allocated *in*-simulation.

$$\vec{F_k} = \frac{m\vec{\nabla}_\alpha c}{\left\|\vec{\nabla}_\alpha c\right\| + \epsilon} \tag{7.2}$$

with $\vec{F_k}$ being the total force applied on a given vertex k of a cell α in both dimension, m the magnitude of the force, $\vec{\nabla}_\alpha c$ the value of the resulting gradient of morphogen computed for a cell α along both axis , $\left\|\vec{\nabla}_\alpha c\right\|$ is the norm (or magnitude) of $\vec{\nabla}_\alpha c$ and ϵ is an extremely small positive number used to avoid division by zero, *i.e.* when $\vec{\nabla}_\alpha c = 0$.

This force definition implies that cells move according to the direction of the gradient. On the other hand, the magnitude of this force, further translated into a velocity, only depends on the value of m. As mentioned earlier, no quantitative or spatial calibration could be performed on VEGF expression level and diffusion; instead we used an empirical calibration. In fact, in our simulations, the value of m has been set so as to ensure simulation stability first (large enough to trigger cell movement but small enough to avoid instabilities), but also to satisfy our experimental observations. The final value of this parameter can be found in the Appendix (Tab. 13.1).

It has been mentioned that only the tip cell has the ability to move according to active forces. On the other hand, stalk cells are expected to move only as a result of strong adhesion between endothelial cells (*i.e.* with the tip cells). In our model, tip cells are considered differentiated (they do not have the ability to divide) while stalk cells were still defined as "transit" type. Moreover, a different cellular type is to be applied to stalk cells as they do

not proliferate in the same way as epithelial cells. Here, stalk cells are expected to divide to make up for the space generated by tip cell migration, *i.e.* between the original and final tip cell position. We therefore defined a new cell cycle in which stalk cells divide when exceeding a certain elongation ($\text{elong}_{\text{max}}$, see Table 13.1 in the Appendix) caused by strong adhesion with neighbouring endothelial cells (tip or stalk). This elongation is characterised by what is called an "elongation shape factor" computed at each time step for each cell, a dimensionless quantity (*i.e.* with no physical dimension) that numerically describes the shape of every cell in the population (independent of its size). Mathematically, it is defined as the ratio of the two second moments i_1 and i_2 of a particle around its two principal axes [95]:

$$f_{elong} = \sqrt{\frac{i_2}{i_1}} \tag{7.3}$$

Once again, no quantitative experimental calibration has been made on this endothelial cell elongation shape factor. Indeed, ECs are elongated and flat cells which are difficult to fully visualise on tissue sections. In this manner, the elongation threshold has been set to allow the stalk cells to grow and divide as expected while avoiding incoherent elongation and potential rupture in the network caused by too high values of the latter.

Using the same IC as earlier, we manually labelled four endothelial tip cells and four corresponding stalk cells (Fig. 7.3, B). Here, an important note is to be made on the external environment in our simulations. In this framework, we focus on the epithelial primordium and various cellular mechanisms or reorganisations occurring within the latter. Consequently, existing blood vessels arising from earlier angiogenetic growth and sprouting at a distance from the epithelial primordium are not explicitly defined in our model. We therefore made the assumption that endothelial cells in contact with the epithelial primordium at $t = 0$ can be considered as tip cells. Insights on how to account for this external blood vessel network will be further discussed.

7.3 Results

As a first test, we initially considered a model simulating the evolution of a basic epithelial sheet of 72 cells previously generated by the growth of a regular hexagonal lattice of 3×3 cells (Fig. 7.3, A). All cells initially produce some morphogen and all the parameters used are taken from [94]. All cells are labelled as epithelial and the adhesion parameters are set to ensure stable and realistic cell growth. Considerations on epithelial adhesion parameter inference

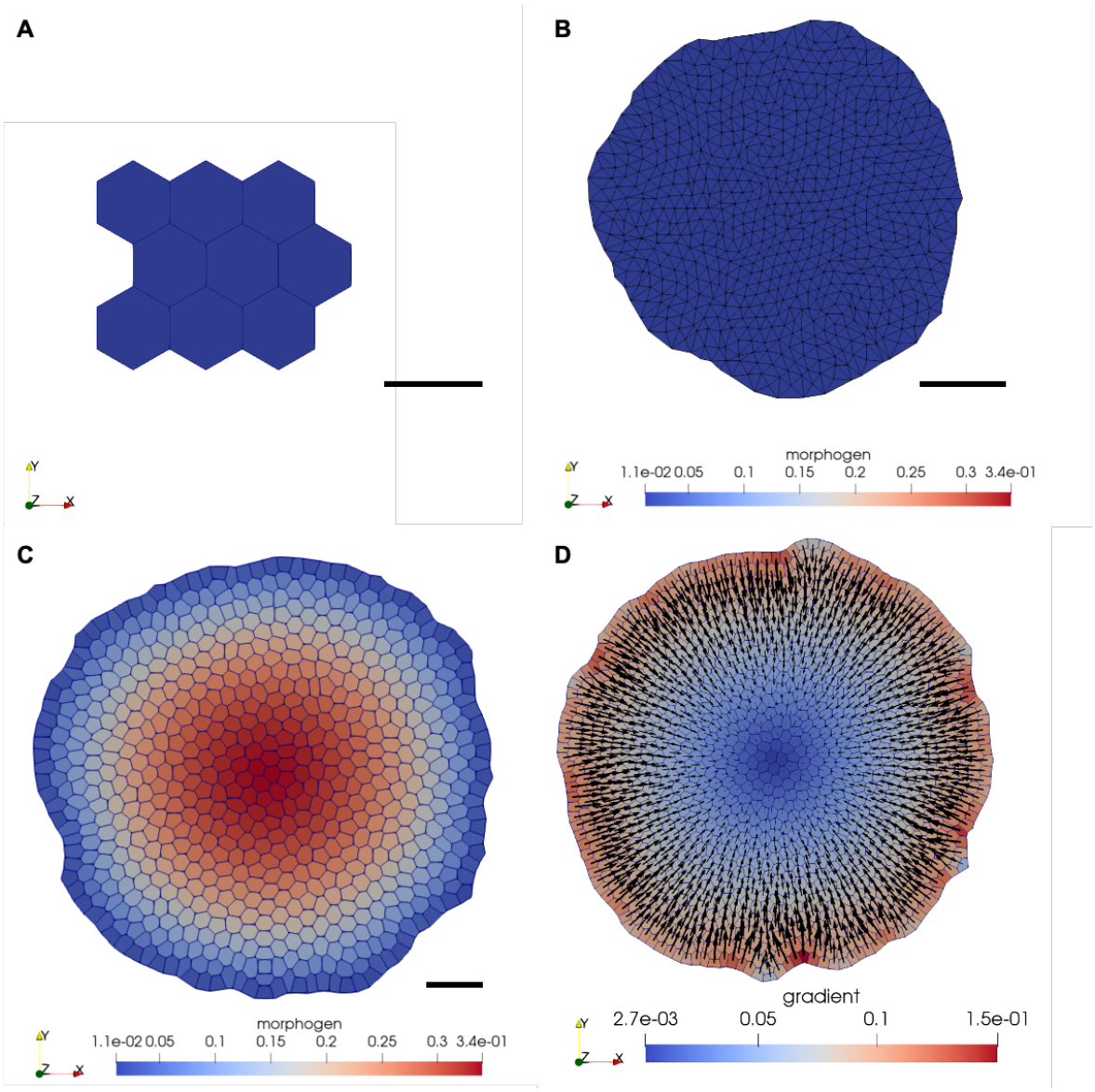


Figure 7.2: Simulation of morphogen gradient formation in a VM induced by a coupled process of diffusion and degradation. Discretisation of the space using a triangulated mesh created by all the vertices and the cell centres. (A) Initial configuration used to generate the initial conditions : hexagonal packing of 3×3 cells (« honeycomb » configuration). (B) Snapshot of the cell population at simulation time t_0. Morphogen concentration ("morphogen") is equal to 0 for all cells. The underlying triangular mesh is also shown in black. (C) Snapshot at t= 12h. Morphogen is produced on all the nodes of every epithelial cell, except on boundary nodes where $c(x,t) = 0$ (Dirichlet boundary conditions). A morphogen concentration gradient is formed. (D) Snapshot at $t = 48$h. Cells are coloured according to their morphogen gradient. The vector field is also shown, illustrated by an arrow for each cell.

are made in Sec. 11. The simulation lasted for 24h with a time step of $\Delta t = 0.001$h. We showed that our preliminary model with parameters from the literature is able to convincingly simulate the formation of a morphogen gradient within a growing epithelial sheet (Fig. 7.2, C-D). The multiple cell divisions and rearrangements do not affect the numerical solution of the PDE even though the mesh is constantly reshaped[1]. While this may not be an issue with short test simulations, simulation of large populations of cells over a longer period of time can require significantly longer computational time. To avoid this, the initial solver was therefore modified to approximate a new solution if and only if $modulo(\frac{t}{10\Delta t}) = 0$, with t being the time elapsed and Δt the time step. In other words, the time step used in the resolution of the diffusion equation is 10 times larger than the one used in the global simulation, the solution of this equation being therefore computed only every 10 "global" time steps. Importantly, this triggers a change in the numeric value of the morphogen concentration over time, but will not change the spatial dynamics of its spreading. Since morphogen diffusion is only needed to control the direction of migrating endothelial cells and not their velocity, only the direction of the resulting gradient is actually used.

We then ran the simulation including endothelial cell motility and proliferation, for $t =$ 48h and $\Delta t = 0.001$h. All the parameters used in this model can be found in the Appendix (Tab. 13.1). We show that our model and assumptions can successfully simulate a simplified process of angiogenesis or endothelial network formation and that the combination of endothelial cell motility and proliferation causes the segmentation of the initial epithelial mass. The tip cell moved inside the cell sheet at a slightly variable velocity, in response to the balance[2] with passive adhesion/cohesion forces acting on each vertex. Stalk cells tended to follow their corresponding tip cell and fill the void created by tip cell progression by cellular division. As expected, stalk cells divided when they reached a critical elongation (Fig. 7.3 C). When a tip cell came into contact with any other endothelial cell, it lost its ability to move, fused with the endothelial cell and triggered the formation of an endothelial network (Fig. 7.3 D). Interestingly, when the simulation attained its relaxed configuration, the newly formed endothelial network seemed to also have reached a certain equilibrium. In fact, the endothelial cells located at the border of the epithelial cell sheet, in contact with the external environment, appeared to maintain this contact and remain at their position rather

[1]Note that a fixed mesh could also be used, which could significantly reduce computation time but would be rather useless in order to compute the gradient of morphogen in each cell. In fact, it is precisely the change in cell shape and population architecture that will induce this particular gradient dynamic.

[2]A cell will actively move if the force applied on it outbalance the passive forces on at least one of its vertex.

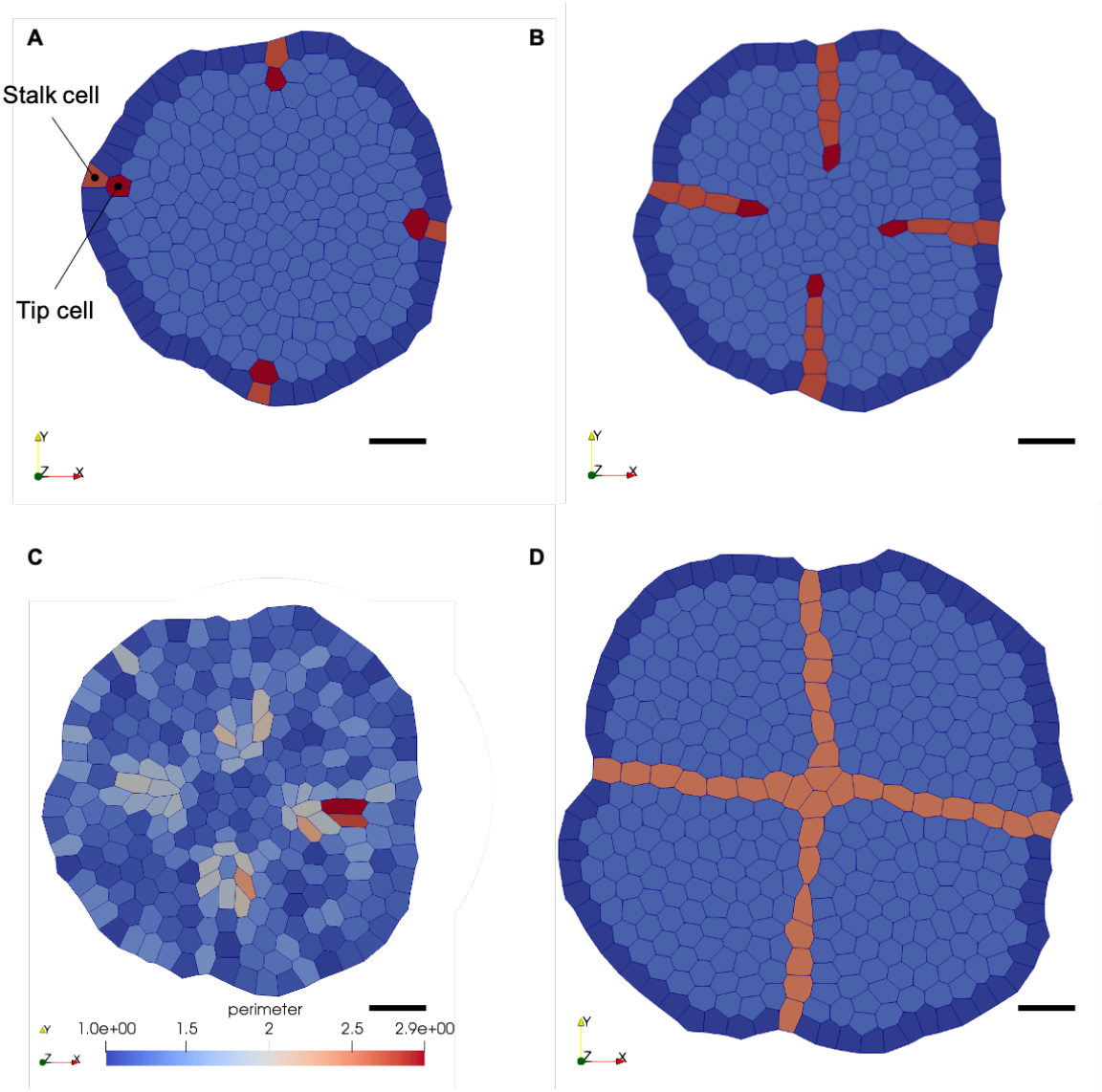


Figure 7.3: Snapshots of a VM simulation of an basic epithelial sheet growth with simple endothelial network formation. (A) Configuration at initial time. Endothelial tip cells are labelled in red, endothelial stalk cells in dark orange. Each tip cell is directly followed by a stalk cell. (B) Configuration at simulation time $t = 0.35$h. Endothelial tip cells are moving following the morphogen gradient, stalk cells are following by strong adhesion and division. (C) Configuration at simulation time $t = 0.35$h. Cells are coloured according to their elongation factor ("perimeter"). This illustrates how endothelial stalk cell division is mainly active on the cells directly following the tip cell. (D) Configuration at simulation time $t = 12$h. All the tip cells have reached the centre of the mass. Every endothelial cells are now labelled as "vessel" (light orange) cells and form a stable and coherent network.

than invade the epithelium. This is due to both a small value of the corresponding adhesion parameter (hence a high adhesion) and a sufficient rate of cell division. Similarly, the vessel-like[3] endothelial branches of the network managed to stay coherent.

[3]Vessels are here represented as a simple succession of endothelial cells, far from the *in vivo* tube-forming structures.

8 | Endothelial Boundary Conditions and Density

8.1 Introduction

We previously explained how the model primarily focuses on intercellular mechanisms within the epithelial primordium. However, this multi-cellular mass lies in a more complex environment and is surrounded by various tissues such as, for instance, existing blood vessels. We made the choice that only the endothelial cells migrating inside this mass would be represented. The impact of the whole endothelial network, both on a mechanistic and chemical level, is nevertheless to be taken into account. Moreover, we rapidly observed that using more realistic, and thus complex, IC would eventually lead to a break and engulfment of endothelial cells, due to epithelial cell proliferation. Here, in order to account for the presence of existing blood vessel outside the initial boundary of our simulations (defined by the epithelial mass), we developed a set of new boundary conditions (BC) for both endothelial and epithelial cells, thereby generating a "no man's land" behind every peripheral initial stalk cell.

8.2 Methods

In our first test simulation, we showed how epithelial cells would proliferate and induce the growth of the initial primordium without the implementation of any boundary conditions. However, the use of IC generated from our image analysis (see Sec. 12) emphasised some first limitations of our model. In fact, the combination of this initial population configuration and epithelial cell proliferation seemed to rapidly lead to endothelial enfoldment (*i.e.* enfoldment of said endothelial cells inside the epithelial mass). This can be explained by the complexity of our initial configuration, in comparison to the one used in our first test. Some endothelial

cells are indeed initially located already "inside" the epithelial mass (Fig. 8.4, A) and were therefore rapidly enfolded upon the proliferation of peripheral epithelial cells. Moreover, the extremely strong cohesion between endothelial cells seemed to sometimes cause the initial stalk cell to follow the motile tip cell rather than adhering with the periphery of the cell population and diving according to their previously described cell cycle model. To avoid this, we implemented specific boundary conditions that can be broken down into two distinct mechanisms.

First, every boundary node owned by a cell labelled as "Endo" and "Stalk" is to be fixed, while the other nodes keep their ability to move (Fig. 8.3, B, red dots). This should ensure the coherent growth and division of stalk cells while avoiding the latter being dragged down by the active migration of the tip cells. Secondly, peripheral epithelial cells are resting on a basement membrane preventing them from adhering to another epithelial *via* their basal pole. There is therefore a need to consider the physical space taken up by the non-visible endothelial cells in our simulations. In Chaste, boundary conditions are defined as a set of rules applied to any nodes at each time step. Here, it takes the form of two straight lines starting from any endothelial stalk cells, delimiting the boundary of our simulation (Fig. 8.3, B, red dashed lines). In other words, as soon as a given node is supposed to be located between those two lines (after the 4 steps described in Sec. 3.2), it is relocated to the closest point on one of the two boundary lines (by orthogonal projection). The most general equation for a straight line is of the form $ax + by + c = 0$[1]. Here, we compute a, b and c in order to define two straight lines starting from each fixed peripheral stalk cell. The equations for those boundary lines read as follows :

$$\begin{cases} f_{l,\alpha} \equiv ax_\alpha + by_\alpha = l \\ f_{r,\alpha} \equiv ax_\alpha + by_\alpha = r \end{cases} \tag{8.1}$$

with f_l and f_r being respectively the left and right boundary corresponding to a stalk cell α, x_α and y_α the position of the centre of the same endothelial cell α.

Concretely, the parameters a, b are computed using the equation of the straight line passing through the centre of α and the centre of the mass while l and r are set for $f_{l,\alpha}$ and $f_{l,\alpha}$ to cross the two boundary vertices of each peripheral stalk cell. This line is supposedly representing the course of the tip cell towards the source of the gradient and the assembly of a basement membrane by the epithelial cells. However, since those parameters are estimated

[1]In this case for instance, the line would be horizontal if $a = 0$ and vertical if $b = 0$.

at $t = 0$, no diffusion is yet observed. We therefore made the assumptions that at early stages of embryonic development, the thyroid primordium takes a rather spherical form and the centre of mass can act as an acceptable source point, which is consistent with our *in vivo* observations. Obviously, these boundaries need to be defined only for the physical space behind the stalk cell, considering that every vessel of the endothelial network has the same "direction" of its corresponding tip cell. Only the nodes located in the semi-infinite plane defined by two boundary lines and the boundary edge of one peripheral stalk cell should be affected by these boundary conditions. To achieve this and for any epithelial boundary node i considered and a stalk cell α, two vectors are generated at each time step; a vector $\vec{v_{\alpha,i}}$ between the centre of α and i Fig. 8.3, B, in green) and one between the center of α and the centre of mass $\vec{v_{\alpha,m}}$ (Fig. 8.3, B, in white). The boundary condition is applied on i if and only if the dot product $\vec{v_{\alpha,i}} \cdot \vec{v_{\alpha,m}} < 0$, that is to say if the two vectors have an opposite direction. This ensures that the virtually infinite straight boundary lines only affect the nodes that are on the same side as the peripheral stalk cell in our simulations as well as the following vessel that is not explicitly represented, with respect to the source of the VEGF gradient driving endothelial cell motility.

In our simulations, endothelial cells are initially manually labelled according to our observations (see Fig. 6.3, A) and the evolution of the endothelial population hence only depends on the dynamics of those initial endothelial cells. We showed how boundary conditions could help to account for the pre-existing blood vessels. However, there is still a need to consider the constant development of the endothelial network inside and on the boundary of the epithelial mass. Here, we integrate the formation of new endothelial branches, by the addition of newly differentiated tip cells. We also experimentally quantified endothelial density using endomucin staining on mice embryo sections at E13.5, E15.5 and 17.5. In order to simulate the migration of endothelial cell initially not considered in the model, we therefore developed a method that triggers the addition of new endothelial cells on the boundary of the epithelial mass. In order to do so, the most efficient way is to simply transform a randomly selected epithelial peripheral cell into a stalk cell as well as to label one of its neighbours (this time an epithelial central cell) as a "tip" cell. Note that this random selection is nevertheless governed by a few rules. These are implemented to avoid, for instance, a new stalk cell developing too close (or even adjacently) to existing vessels. Moreover, one and only one of its neighbouring epithelial central cell is also to be transformed into a tip cell. The newly labelled stalk and tip cells possess the exact same properties as the cells manually labelled in our initial configu-

ration. The position of every new stalk cell as well as the parameters of its corresponding BC straight lines is also added to the simulation. While this cell type transformation is obviously not observed *in vivo*, to simulate the arrival or recruitment of new blood vessels around the thyroid leading to the growth of the endothelial network. One must remember that the cell types described in our model are in fact not strictly corresponding to actual morphological or phenotypical cell types.

8.3 Calibration

Images from different embryos ($n = 3$) were analysed at different embryonic stages (Fig. 8.1). We used a Hoechst staining to detect all cell nuclei, E-cad for the epithelial cells and endomucin for the endothelial cells. A prior calibration of the stain detection threshold was performed for the three states. We used a specific area detection algorithm within the software HALO to quantify the endothelial density. We defined this endothelial density as the fraction of the surface occupied by endothelial cells to the total surface (*i.e.* the sum of all the epithelial and endothelial cells within the analysis region).

These results allow for a more quantitative description of the dynamics of endothelial growth within the embryonic thyroid (Table 8.1). We observed that endothelial cells represent less than 2% in area at E13.5. However, the majority of endothelial cells seem to be located outside the epithelial primordium at the particular state and the endothelial density at that state could therefore be considered as almost nil. On the contrary, endothelial cells occupy almost 7% of the space at E15.5 and this value rises to more than 12% at E17.5. Our results suggest a linear evolution of the relative endothelial surface over time (Fig. 8.2). However, it is obvious that this increase is not infinite and that endothelial growth is expected to stop concurrently with the growth and maturation of thyroid follicles. Nevertheless, these data were used to calibrate our model of endothelial network development.

Table 8.1: Endothelial density in the embryonic thyroid at three embryonic stages ($n = 3$).

	Embryonic day		
	E13.5	**E15.5**	**E17.5**
Relative endothelial surface [%]	1.95 ±0.04	4.83 ±0.28	10.33 ±1.4

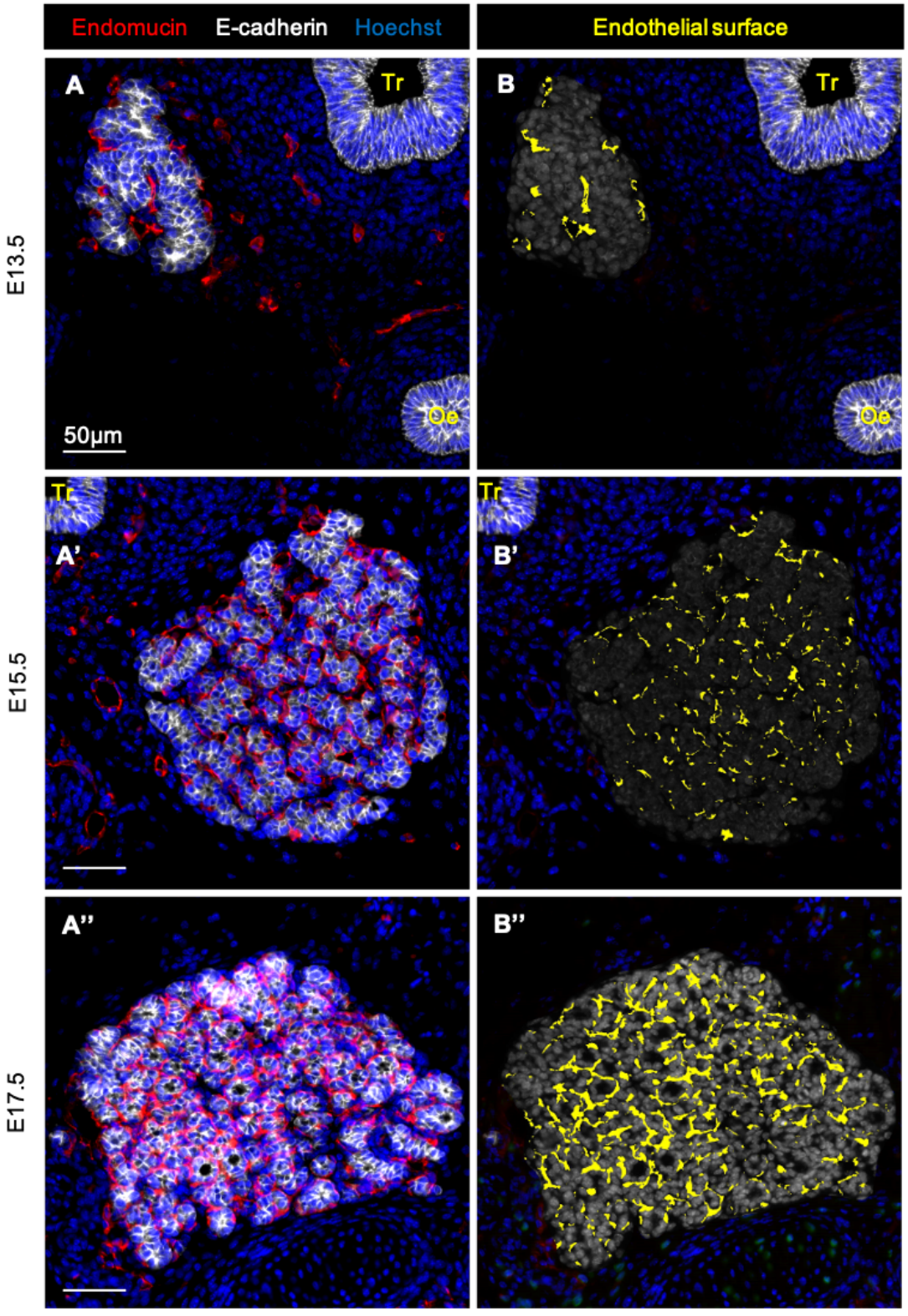


Figure 8.1: Visualisation and quantification of endothelial cells in developing thyroid at three different embryonic stages. (A-A") Immunolabelling of epithelial cells in white (E-cadherin), and endothelial cells in red (Endomucin) at three embryonic stages. Nuclei are stained in blue (Hoechst). Tr, trachea; Oe, oesophagus. (B-B") Endothelial cell within the analysed region of interest (in grey) were coloured in yellow using HALO.

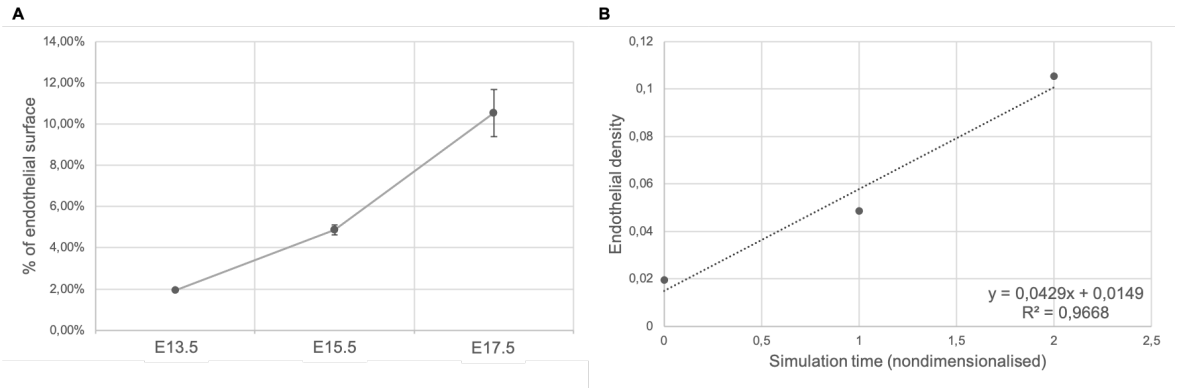


Figure 8.2: Evolution of endothelial cell density over time. (A) Fraction of endothelial cell in the tissue, in % ($n = 3$). (B) Endothelial density dynamics in the embryonic thyroid. Both axis are nondimensionalised by respectively using a cell area density (dimensionless) and replacing the embryonic days by the time elapsed between two measures (here 48h).

8.4 Results

We ran the simulation with new IC and BC and for $t = 24$h and $\Delta t = 0.001$h. All the other parameters are the same as the one used in the test simulation and can be found in the Appendix (Tab. 13.1). We manually labelled four endothelial tip cells and four corresponding stalk cells, according to our observations in E13.5 thyroid sections. We show that our BC was able to simulate the mechanical impact of endothelial vessels on epithelial growth and rearrangement. Tip cells migrated towards the gradient of VEGF. As expected, peripheral endothelial cells stayed relatively fixed while the following stalk cells divided at a sufficient rate to ensure the formation of a coherent endothelial network. Importantly, our simulation also shows signs of a budding-like process induced by epithelial growth limited by the boundary conditions generated by every peripheral stalk cell. We showed that the endo-BC were able to simulate the mechanical impact of endothelial vessel invasion, the permanent epithelial-epithelial cell disruption by ECM deposition, the presence of outward vessels, and epithelial growth on thyroid mass fragmentation (Fig. 8.4, C). As expected, peripheral endothelial cells remained relatively fixed at their original position (Fig. 8.4, D inset), while stalk cells divided at a sufficient pace to fill the space behind the migrating tip cells and ensure the formation of a coherent endothelial virtual vessel. Importantly, our simulation also shows signs of a budding-like process generated by epithelial proliferation at the periphery limited by the presence of endothelial cells (Endo-BC) and ECM (Fig. 8.4, D). Epithelial cells proliferate along the boundaries (accounting for the presence of endothelial cells that would connect with distant vessels and the deposition of ECM preventing epithelial cell-cell interactions)

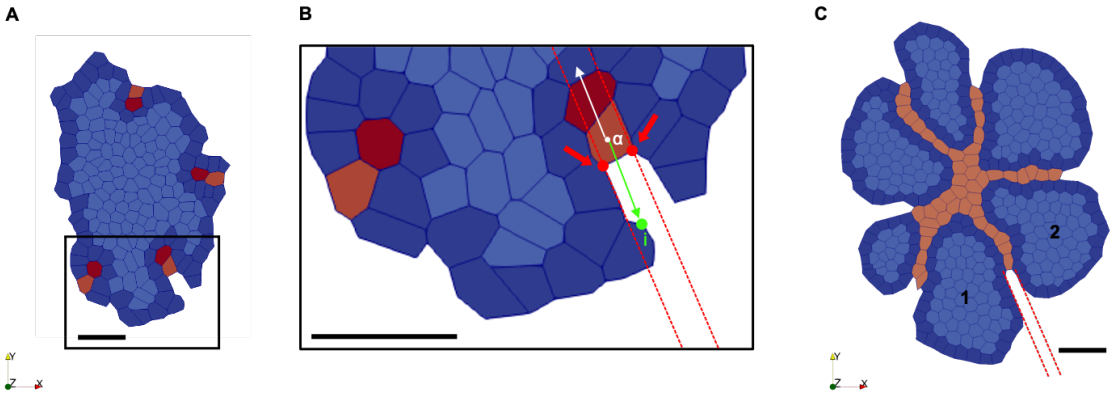


Figure 8.3: Schematic of the endo-BC implementation. (A) Initial configuration of the thyroid VM (t = 0h). Epithelial cells are labelled in blue (dark at the periphery and light in the centre), endothelial tip cells in red, and endothelial stalk cells in orange. Positions of a new stalk cell and a new tip cell are indicated. (B) Close-up view of (A) with an illustration of the endo-BC. The two boundary nodes from the endothelial stalk cell are fixed (red dots, indicated by the red arrows). The two boundary lines (dashed red lines) are parametrised to pass through the centre of each of the two fixed nodes as well as through the centre of the initial mass. (C) Configuration at simulation time t = 18h. The two boundary lines (dashed red lines) defined a "no man's land" between the periphery of the newly formed islets (1 and 2). Epithelial cells proliferate and smoothly move along those two lines.

and eventually lead to the formation of smaller individual mass separated by newly formed endothelial virtual vessels.

Moreover, our model is also able to simulate the constant migration of endothelial cells by adding new endothelial stalk cells (and their corresponding tip cells) on the periphery of the epithelial population, according to the dynamic of endothelial density discussed earlier. These stalk cells are instantly fixed as expected. Endothelial BC are also updated accordingly. Newly formed endothelial tip cells also correctly migrated towards the gradient, at the same velocity as the initially labelled tip cell (given that only the direction of the VEGF gradient drives endothelial motility), and stalk cells still divided according to the same cell cycle model. This induces a continuous evolution of the endothelial network as tip cell eventually fuse with existing vessels (Fig. 8.4 B, C and D).

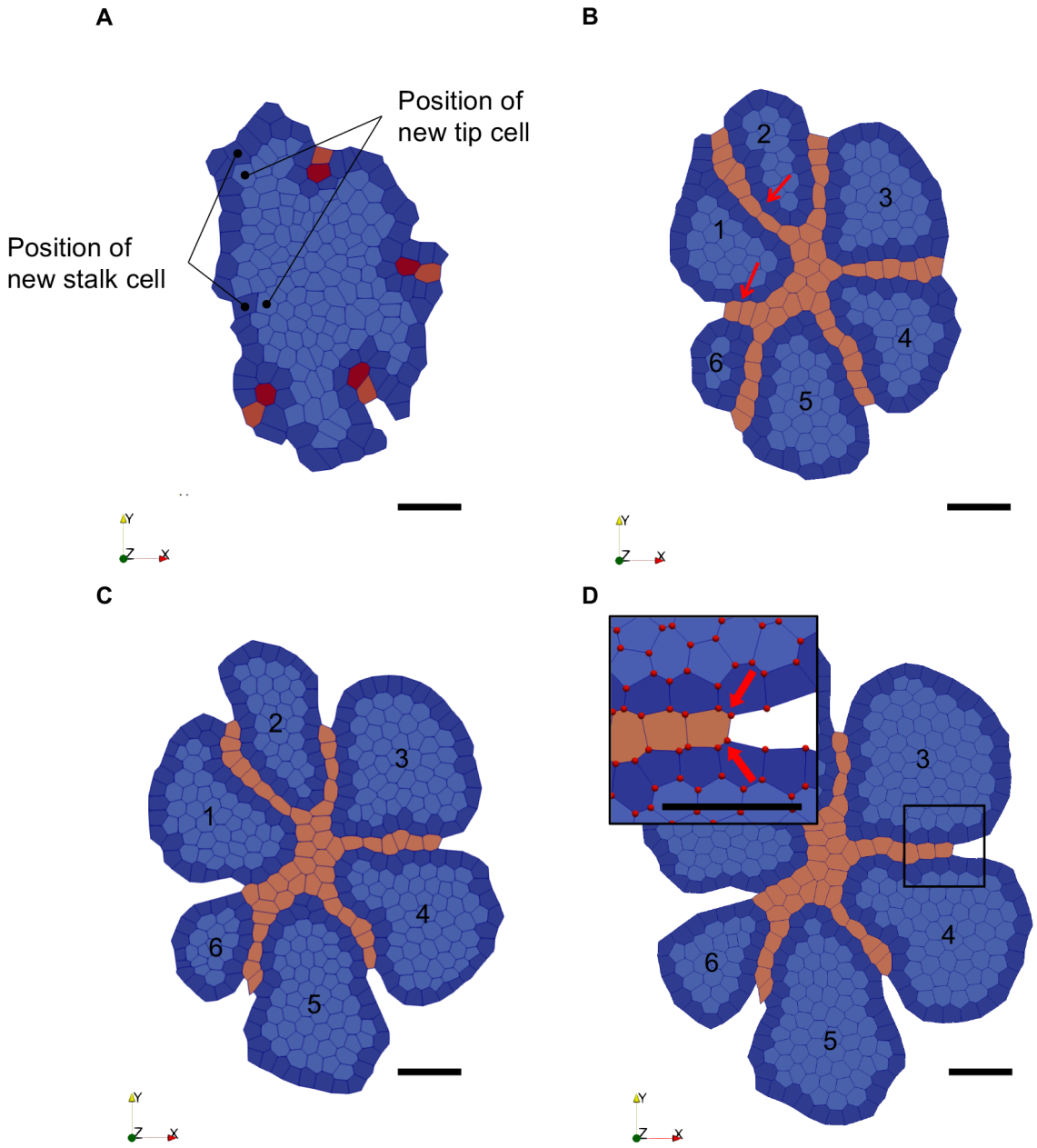


Figure 8.4: Snapshots of the VM simulation of thyroid epithelial sheet growth and endothelial invasion within the epithelial sheet (A) Initial configuration of the thyroid VM (t = 0h). Epithelial cells are labelled in blue (dark at the periphery and light in the centre), endothelial tip cells in red, and endothelial stalk cells in orange. Positions of a new stalk cell and a new tip cell are indicated. (B) Configuration at simulation time t = 10h with endothelial boundaries (Endo-BC) created for the most peripheral stalk cells. New tip and stalk cells have been added randomly and migrated towards the gradient, forming new vessels (red arrows). Emerging bud-like structures are numbered (from 1 to 6). (C) Configuration at simulation time t = 18h. Endothelial tip cells have formed a simple but coherent network of endothelial vessels (orange) and epithelial cells continued to proliferate and grow outwardly. (D) Configuration at simulation time t = 24h. Epithelial cells of new sheets continued to grow and show clear signs of bud-like structures formation. The «virtual» endothelial vessels induced by the Endo-BC are now visible between the new epithelial islets. The close-up view shows the vertices represented as red dots. Fixed vertices are indicated by the two red arrows.

9 | Cell polarisation and Lumen Generation

9.1 Introduction

Another crucial process needed to be described in our model of thyroid morphogenesis: folliculogenesis. This last step of thyroid embryonic development relies on the production and fusion of intracellular vesicles filled with apical material (Tg), with the developing apical pole of epithelial cells. However, VM represents cells as polygons. Vertices are linked together by edges to form cells, generating a coherent population of cells and the spatial resolution of VM, which unfortunately does not allow the definition of such mechanisms. Moreover, the very definition of VM prevents the existence of stable "voids" in the simulation. In order to get around these model limitations, lumina were defined here as a distinctive cell type ("Lumen", see Fig. 9.1). Each lumen is here represented by one cell, with completely different properties (target area, adhesion, no proliferation). The easiest way to generate new cells in VM is simply by cell division. The addition of new vertices and cells is technically possible, but become somewhat tricky when it comes to generating new cells inside the epithelial cell population. Lumina are therefore generated by the division of an epithelial mother cell, with one daughter cell staying epithelial and the other one becoming a lumen (hence labelled as such). Here, we discuss the implementation of the different computational mechanisms leading to the generation of such lumina in Chaste. We also experimentally quantified the dynamics of epithelial cell polarisation as well as lumen formation and growth at three different embryonic stages, in order to further calibrate the various parameters involved in our lumen generation model.

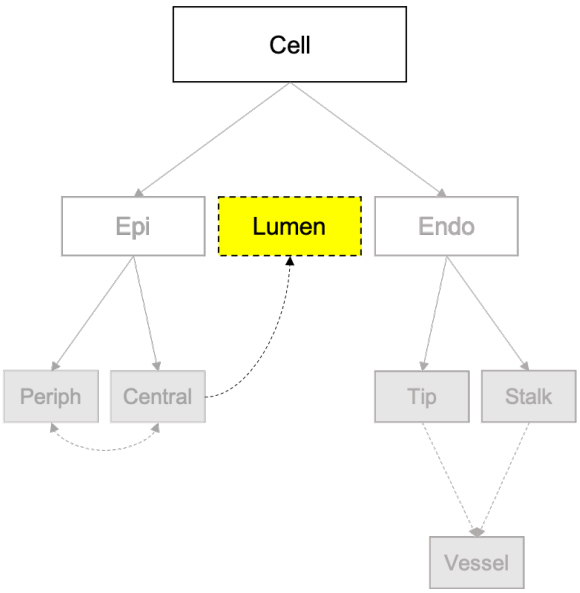


Figure 9.1: Hierarchical chart illustrating the different cell types in our simulation. Simple arrows represent the course of cell type assignment upon division (or at $t = 0$). At the start of the simulation, each element is assigned to a cell object. This cell is immediately given a cell type label, either "Epi" (for epithelial cells) or "Endo" (for endothelial cells). Similarly, a label referring to its positions is added for each epithelial cells, either "Periph" (for peripheral cells) or "Central" (for central cells). Plain lines corresponds to cells types defined *pre*-simulation, dotted lines to types allocated *in*-simulation. "Lumen" is a particular label used in our simulation to represent thyroid lumina. Upon division, any new central epithelial cell can transform into a lumen cell if reaching certain criteria defined hereafter.

9.2 Methods

We described in the introduction (see Sec. 2.2) how cell polarisation plays a central role in lumen formation during embryonic thyroid development. It was therefore crucial for our model to be able to integrate this process. However, cellular polarisation is rather difficult to quantify *per se*. In other words, there is no real insight on "How much a cell is polarised". In fact, polarity can be defined as the asymmetric distribution of macromolecules and other organelles in a cell. However, despite the existence of a few quantitative studies tackling the issue of polarity establishment and quantification ([96–98]), VM still does not allow the simulation of intra-cellular objects hence the spatial reorganisation eventually triggering cell polarisation. Moreover, existing polarity models often rely on the incorporation of cell signalling pathways and feedbacks [99] that have not been studied here [100]. Cell polarity was thus abstracted by the implementation of a polarity vector $\vec{p}$ in every epithelial cell. Concretely, each epithelial cell has been assigned a polarity vector, initially null. This vector is characterised by a direction, accounting for the basal-apical polarity (*i.e.* pointing toward the apical pole) of the cell, as well as a magnitude. This magnitude is accounting for the so-called polarisation level of the cell (hence answering the question of "how much a cell is polarised").

We described how *in vivo*, the migration of endothelial cell induces epithelial cell polarisation by deposition of basal membrane ([17]). Epithelial cells in contact with the basement

membrane are oriented and produce small intracellular vesicles that can fuse with the cell membrane opposite to the one in contact with the basement membrane to form an apical lumen. The formation of this lumen causes a disconnection of the junctions between the apical pole of this oriented and polarised cell and its neighbour (second line). Eventually, this cell with an apical pole will fully polarise by assembling a basement membrane that can in turn trigger the polarisation of a third line neighbouring epithelial cell.

In our model, we thus suggested that polarisation could propagate across the epithelial cells by "contagion". Both the magnitude and direction of the polarity vector of any epithelial cell are therefore constantly updated as a result of the interaction with the neighbouring endothelium, with "lumen" and epithelial cells and with the boundary of the epithelium.

Concretely, the magnitude and direction of any polarity vector $\vec{p}$ are computed at each time step and for each epithelial cell α, in a series of sequential well-defined steps. First, the vector is reduced by a value proportional to its actual magnitude, ensuring that it will not grow infinitely and eventually reach an equilibrium , as follows :

$$\frac{d\vec{p}}{dt} = -\delta_{pol}\vec{d} \tag{9.1}$$

with δ_{pol} being the decrease parameter[1]. This manipulation only change the magnitude of $\vec{p}$ and not its direction.

Then, we compute the impact of endothelial (hence ECM) or lumen cells on $\vec{p}$ at each time step. Considering an epithelial cell α, the definition of the polarity vector $\vec{p}$ of the epithelial cell read as follows :

$$\frac{d\vec{p}}{dt} = \rho_{pol}\vec{d} \tag{9.2}$$

with $\frac{d\vec{p}}{dt}$ being the variation of the polarity vector $\vec{p}$ over time, ρ_{pol} the polarisation rate, which directly denotes the impact of endothelial (ρ_{endo}, see Table 13.1 in the Appendix) or lumen (ρ_{lumen}, see Table 13.1 in the Appendix) cells on the polarity vector and $\vec{d}$ is a normalised vector ($\left\|\vec{d}\right\| = 1$) pointing from the centre of all neighbouring cells to the centre of α . Note that this normalisation ensures that the numeric value of $\vec{d}$ does not impact the amplitude of $\vec{p}_\alpha$ but only its direction.

Here, a positive value of ρ_{pol} for endothelial cells will induce a polarisation of the epithelial cell in the direction opposed to the endothelial cell (*i.e.* from the "basal" edge to the "apical"

[1]Multiplying by the time-step (Δt) allow to keep an identical decrease regardless of the time-stepping used.

edge). On the contrary, a negative value of ρ_{pol} for lumina will trigger the polarisation of the cell towards the lumen again, from the "basal" edge to the "apical" edge). The implementation of the impact of epithelial cells on $\vec{p}$ slightly differs from Eq. 9.2. In fact, not only will the proximity of the latter cells be taken into account here, but also their own polarity vector. For two neighbouring epithelial cells, with their respective polarity vector $\vec{p}$, and with an angle θ between the two polarity vector, the new definition of $\vec{p}$ reads as follow :

$$\frac{d\vec{p}}{dt} = \pm\epsilon_{pol}\cos\theta \tag{9.3}$$

for all cells, with ϵ_{pol} being the polarisation rate component related to epithelial cells (similar to ρ_{pol} in eq. 9.2) and θ the angle between α and any neighbouring epithelial cell. This time, the sign of ϵ_{pol} depends on the direction of the neighbouring cell vector. Once again, only the direction of this vector is taken into account.

In our model, lumina are represented as cells. We explained that the most efficient way to generate new cells is by cell division. However, when a mother cell divides, it gives rise to two identical daughter cells. Here, the polarity vector is used to drive cell division and hence lumen generation. The fate of the two daughter cells is determined by the position along the polarity vector: the proximal cell in contact with the basement membrane remains epithelial and takes the place of the mother cell with a polarity vector, while the distal daughter cell becomes a lumen cell. Mathematically, a first vector is computed, linking the centre of the two daughter cells α_1 and α_2 (Fig. 9.2, $\vec{v_1}$ in red). Two vectors are then generated, the first (Fig. 9.2, $\vec{r_1}$ in green) being the resulting vectors from the addition of $\vec{v_1}$ and the polarity vector $\vec{p}$ (both daughter cells have the same polarity vector upon division) and the second (Fig. 9.2, $\vec{r_2}$ in purple) being the resulting vectors from the subtraction of $\vec{v_1}$ and the polarity vector $\vec{p}$. If the $\vec{r_2}$ is bigger than $\vec{r_1}$, α_2 is considered as distal and becomes a lumen. Moreover, since epithelial cells in our model normally grow and divide according to their cell cycle duration (see Sec. 6), this had to be bypassed in order to generate lumen cells when the polarity vector was gaining in amplitude and imposing cell division. Therefore, the cell cycle has been modified to allow any epithelial cell α to divide according to two conditions :

1. The magnitude of $\vec{p}_\alpha$ is larger than a fixed threshold ;
2. It is old enough to divide (*i.e.* its age is larger than a fixed threshold).

As the lumen is generated from an existing cell, it loses all the information related to the "epi" label while gaining new characteristics specific to its new "lumen" label. However, it

still keeps the same size as the initial daughter cell. The target area is therefore instantly set to a smaller value. This simplification ensures that the lumen starts growing from a reasonably small size as seen in thyroid embryonic sections. This target size will nevertheless be constantly updated throughout the simulation. In fact, if the neighbouring epithelial cell influences the generation of new lumina, it also has a direct impact on its size. In order to account for the constant production of small vesicles by epithelial cells observed *in vivo*, the target area of any lumen cell is incremented at each time step by a value proportional to the norm of the polarity vectors of the neighbouring epithelial cells. In other words, any lumen will grow at a rate depending on the "level of polarity" of its neighbouring epithelial cells. Concretely, the target area A_λ of any lumen cell λ is computed at each time step an vary as follows :

$$\frac{dA_\lambda}{dt} = \nu \sum_{\alpha=1}^{i} \|\vec{p_\alpha}\| - \upsilon A_\lambda \tag{9.4}$$

with ν and υ being the parameters for respectively the impact of neighbouring epithelial cells of λ and its actual target area, $\|\vec{p_\alpha}\|$ the magnitude of the polarity vector of the neighbouring cell α and i the number of neighbouring cells. The negative sign ensures to reach an equilibrium (when both terms cancel each other out). Note that this incrementation is applied only if its value is positive, to avoid lumen to potentially decrease in size and eventually shrink which would not be consistent with *our in vivo* observations.

9.3 Calibration

Obviously, these parameters do not refer to biological measures. This means that they cannot be directly measured *in vitro* or *in vivo*. No experimental data was therefore available on cell

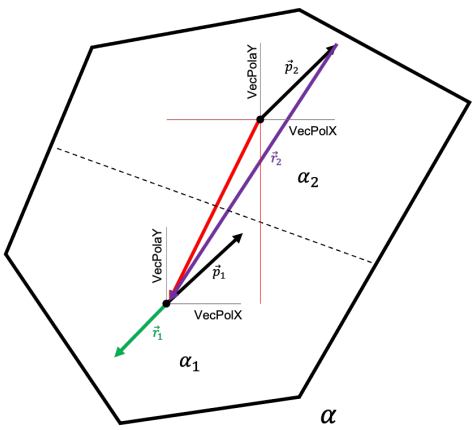


Figure 9.2: Schematic of the vector generation process leading to lumen generation upon epithelial cell division. A first vector ($\vec{v_1}$, in red) in calculated between the two daughter cells centres. The fate of the two daughter cells is determined by their polarity vectors (respectively $\vec{p}_{\alpha 1}$ and $\vec{p}_{\alpha 2}$). The resulting vectors from the addition of $\vec{p}_{\alpha 1}$ and $\vec{v_1}$ ($\vec{r_1}$, in green) and the subtraction of $\vec{p}_{\alpha 2}$ and $\vec{v_1}$ ($\vec{r_2}$, in purple) are calculated. If $\|\vec{r_2}\| > \|\vec{r_1}\|$, the daughter cell α_2 becomes a lumen (as shown), and vice versa.

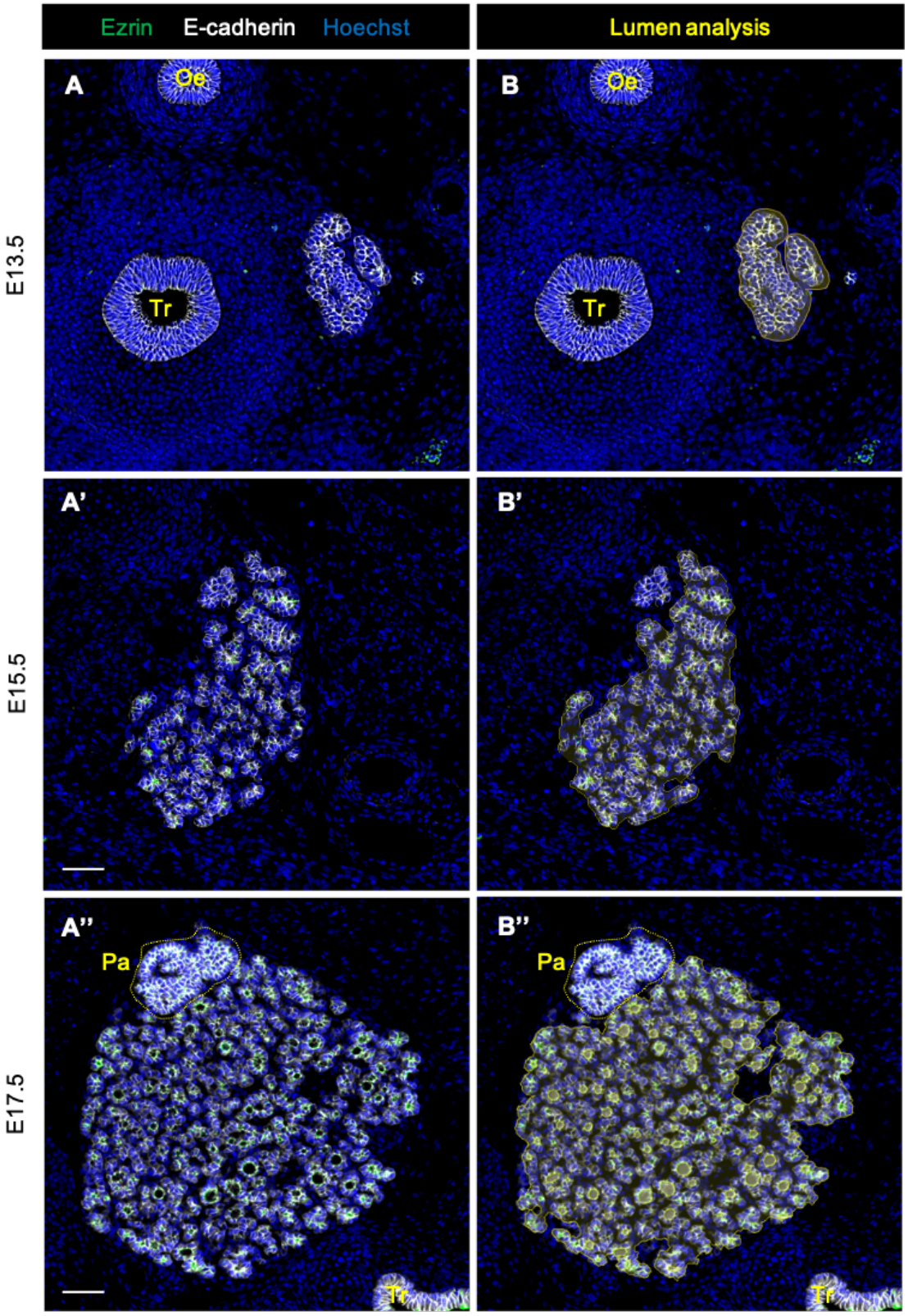


Figure 9.3: Visualisation and quantification of lumen formation in developing thyroid at three different embryonic stages. (A-A") Immunolabelling of epithelial cells in white (E-cadherin) and apical pole of epithelial cells in green (Ezrin) at three different embryonic stages. Nuclei are stained in blue (Hoechst). Tr, trachea; Oe, oesophagus; Pa, parathyroid. (B-B") Epithelial lumen within the analysed region of interest (in grey) were coloured in yellow using HALO (mostly visible in B' and B").

polarisation. Consequently, we used various indicators, that can be evaluated both in our simulations and experimentally, in order to quantitatively calibrate the parameters involved in our model. These indicators are (i) the ratio of the total lumen surface to the epithelial total surface, (ii) the ratio of the number of epithelial cells to the total number of lumina and (iii) the variation of the number of lumina. This calibration was performed concomitantly with the calibration of our lumen generation model described hereafter.

Images from different embryos ($n = 1$) were analysed at two[2] different embryonic stages (E15.5 and E17.5) with the software HALO (Fig. 9.3). Note that ezrin actually marks the intracellular vesicles (too small to be visualised at the magnification showed) and the apical membrane of epithelial cells directly facing the lumen. Lumina are hence manually annotated. Epithelial cells are further marked by HALO according to their ezrin staining level (either positive or negative).

Our experimental data allows for a more quantitative description of the dynamics of lumen formation and growth (Table 9.1). As expected, a notable increase can be observed in the absolute number of lumina between E15.5 and E17.5, which tripled over 48 hours. More interestingly, the epithelial cells to lumen number ratio is consistent over the same period of time, with approximately 23 times more epithelial cells than lumina at E15.5 and around 22.5 at E17.5. This would mean that starting from E15.5, lumina are continuously and homogeneously forming and growing as epithelial cells proliferate. Note however that these results are summed over two lobes for each thyroid section and that a relatively high variability can be observed between the two lobes, notably regarding the lumen to epithelial cell ratio. This could be explained by the incidence of the thyroid sections, leading to different results for the two lobes. Moreover, various studies have shown that morphological asymmetry can be observed in the thyroid of healthy patients [101, 102]. However, the applicability of these results to murine thyroid was not investigated here. On the other hand, the lumen to epithelial cells total surface ratio is significantly higher at E17.5 : 5.14% as compared to 0.2% at E15.5, which indicates that lumina expansion occurs between E15.5 and E17.5. The lack of quantitative data at intermediate states (say E14.5 and E16.5) does not allow for a precise description of the dynamics of folliculogenesis but these results are consistent with previous observations presented in Sec. 2.2 and in [17].

We also observed a net increase in the fraction of ezrin$^+$ epithelial cells. In fact, only traces of ezrin can be detected at E13.5. These correspond mainly to ezrin located in small

[2]Thyroid at E13.5 does not show sufficient sign of cell polarisation.

Table 9.1: Ratio of lumen to epithelial surfaces in embryonic thyroid sections at two embryonic stages ($n = 1$).

	Embryonic day	
	E15.5	**E17.5**
Total number of lumina	38	125
Total lumen surface [µm²]	276.22	12468.9
Total number of epithelial cells	868	2817
Total epithelial surface [µm²]	136285.4	242197
Epithelial cells to lumen number ratio	22.9	22.5
Lumen to epithelial cells total surface ratio	0.2%	5.14%
Ezrin$^+$ epithelial cell ratio	43%	79%

intracellular vesicles. At E15.5, just under half of the epithelial cell is marked by ezrin and at E17.5% this value rises to almost 80%. We would expect this fraction to reach values close to 100% at that time of development since all the epithelial cells are supposedly polarised and should therefore show a clearly marked apical pole. However, the high sensibility of the detection threshold in HALO can explain the variation of the ezrin$^+$ cells and a small change in the latter actually induces a significant change in this fraction. Moreover, one must consider that the apical pole constitutes a very small domain of a polarised epithelial cell, and that histological sections may not pass through this ezrin$^+$ domain. Finally, the exactness of the embryonic stages actually analysed cannot be guaranteed, and it is common to observe a variation of around 12 hours between two sections of different embryos[3].

These results were used to calibrate the different parameters involved in our lumen generation model. This calibration was performed by running different simulations with different values for the three parameters (*i.e.* the impact of endothelial, epithelial and lumen cells on epithelial cell polarity hence lumen formation and growth) and computing an error function based on the indicators described in Table 9.1. This allowed us to find an acceptable set of parameters to be further used in our simulations (see Sec. 9.4 and 10).

9.4 Results

9.4.1 Epithelial polarisation

We ran the simulation with the same IC and endothelial BC used earlier, and for $t = 24$h and $\Delta t = 0.001$h. All the parameters value can be found in the Appendix (Tab. 13.1). We manually labelled four endothelial tip cells and four corresponding stalk cells. We observed

[3]Note however that this particular data were not used *stricto senso* in our calibration but still once again illustrates the dynamics of epithelial cell polarisation in the embryonic thyroid.

here that the polarisation level of any epithelial cell, represented by its polarity vector (Fig. 9.4, D) vary according to its surroundings. Endothelial cells have a large impact on epithelial polarisation, while the impact of the boundary is slightly lower. Moreover, we can clearly see how epithelial cells tend to polarise, as expected, rather orthogonally (although not in a strictly mathematical way due to the cell's shape and type heterogeneity) and in the opposite direction of their neighbouring endothelial cell and/or the epithelial mass periphery. Our model thus satisfactorily simulates the dynamic establishment of the basal-apical polarity of epithelial cells upon endothelial invasion. Note that the impact of lumen cells on cell polarisation is not shown as lumen generation is not implemented yet. The contagion behaviour of cell polarity can also be observed as the polarisation level of various epithelial cells located internally from the endothelium and/or the periphery is rising significantly although towards lower values (Fig. 9.4, C).

9.4.2 Lumen generation

We ran the simulation with the same IC and endothelial BC used earlier, and for $t = 72$h and $\Delta t = 0.001$h. All the parameters value can be found in the Appendix (Tab. 13.1). All the previous modules (*i.e.* epithelial proliferation, VEGF diffusion, angiogenesis, cell polarisation and lumen generation) were implemented. We manually labelled four initial endothelial tip cells and four corresponding stalk cells. Our model was able to successfully generate lumina as a result of endothelial migration and resulting epithelial cell polarisation. Epithelial cells started to generate small lumina after around 24 hours of simulation. At that point, endothelial cells have formed a small vessel network and have segmented the initial thyroid primordium in smaller individual mass (Fig. 9.5, A), due to a combination of peripheral and central epithelial cells proliferation and endothelial boundary conditions. At around 48h, lumina have grown, due to the effect of neighbouring polarised epithelial cells. They have reached a significant size and are scattered in the cell population but clearly develop predominantly in close proximity with the infiltrating endothelium or at the periphery, which is consistent with the cell polarisation dynamics established earlier (Fig. 9.5, B). Epithelial cells polarisation propagated across the mass triggering lumina to develop homogeneously in the growing epithelium. After around 72h, various lumina seemed to have reached a size equilibrium and appeared in a rather discoid shape (Fig. 9.5, C). Obviously, the very definition of a cell in our model prevents the lumina taking on a more circular shape, that also highly depends on the number of epithelial cell surrounding the latter.

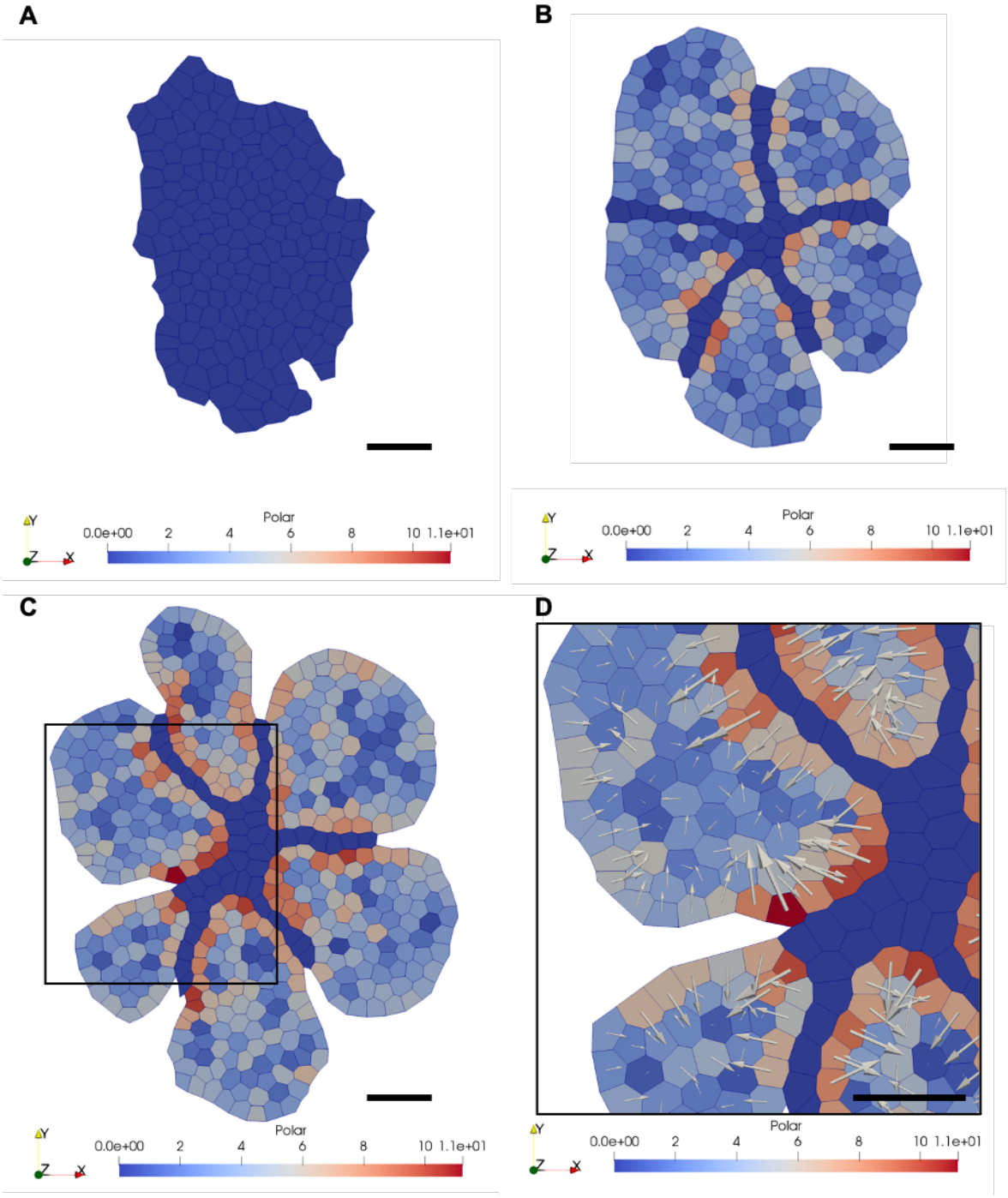


Figure 9.4: Snapshots of a VM simulation of the epithelial primordium with realistic conditions. (A) Configuration at initial time. No cell is yet polarised. (B) Configuration at simulation time $t = 12$h. Cells neighbouring the endothelium as well as the periphery show first sign of polarisation. Cells are coloured according to their polarisation level, ranging from non polarised (navy blue) to fully polarised (red).(C) Configuration at simulation time $t = 24$h. Epithelial cells polarisation is now clearly visible. (D) Close-up view of (C). Cell polarity is represented by an arrow (polarity vector) for each epithelial cell.

Consideration on lumen surface area and corresponding number of epithelial cells are made in chapter 10. The thyroid primordium still showed a coherent endothelial network and a significant number of lumina (a bit less than 100 after 72h). The largest lumina appeared to be 2 to 3 times larger in area than neighbouring epithelial cells. A very important note is nevertheless to be made regarding the final simulation time. In fact, the initial goal was to run a simulation over a period of 96 hours, similar to the timing of our experimental analysis (E13.5 to E17.5). Note however that despite all the rules implemented in our polarisation mode, a very small amount of lumina appeared to be neighbouring each other. This should not happen (as it is not biologically coherent) and while a few attempts to solve this issue (such as cell deletion of merging) have been made, working on this aspect should be a priority for further developments. Moreover, the combination of all these more or less complex processes, unfortunately, ended up causing instabilities that our model could not handle, triggering the simulation to break and stop. A large section of our discussion (see chap. 11) is dedicated to this issue.

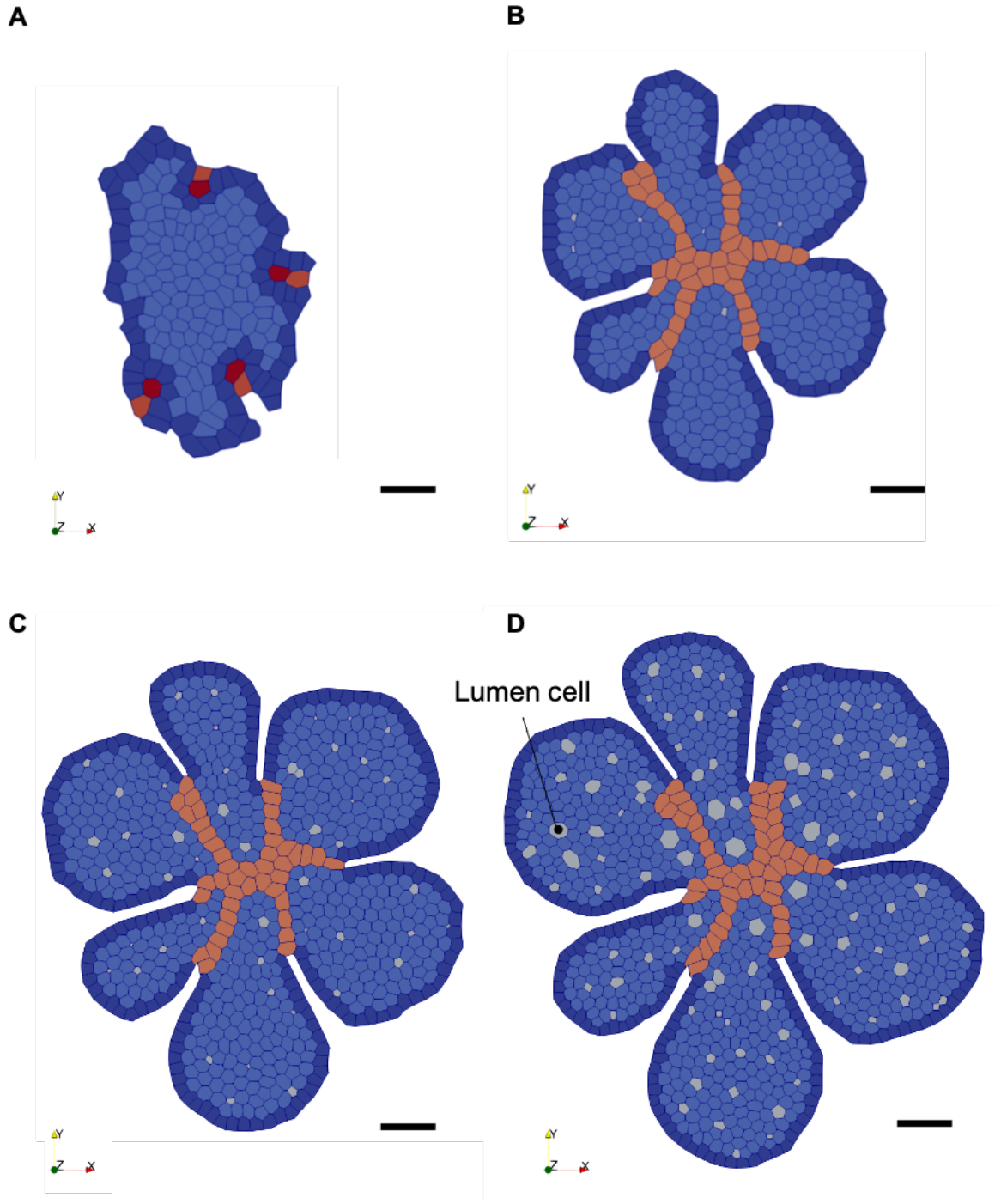


Figure 9.5: Snapshots of the VM simulation of epithelial polarisation and lumen formation in the growing thyroid epithelial sheet invaded by endothelial cells. (A) Initial configuration of the thyroid VM (t = 0h). No cell is polarised at that stage. (B) Configuration at simulation time $t = 24$h. Endothelial vessel network is in dark blue. Degree of epithelial cell polarisation can be visualised with the colour scale (blue, not polarized to red, polarized). (C) Close-up view of (B) with polarity vectors (white arrows) representing polarisation state for each epithelial cell. (D) Configuration at $t = 48$h. Small lumina can be observed within each of the epithelial islets. Epithelial cells keep on proliferating despite a progressive reduction in proliferation rate. (E) Configuration at $t = 72$h. More lumina have developed within the fragmented epithelial sheet, and lumina size continued to expand with more epithelial cells reaching a polarized state. On the contrary, epithelial cell proliferation continued to decline, and expansion is difficult to visualise.

10 | Model Validation

The various parameters involved in our model have been calibrated using experimental data gathered from murine embryonic thyroid sections at E13.5, E15.5 and E17.5. In order to validate our model, we used three new indicators: the distribution of epithelial cell surface area, the distribution of lumen surface and the evolution of the lumen surface area with respect to the number of epithelial cells surrounding the lumina. The experimental values of these indicators were estimated using the scanned images analysed by HALO. In our model, specific writers (see Sec. 3.2) allowed us to compute the different values of these indicators at the end of the simulation.

10.1 Epithelial cell surface area distribution

Thyroid sections from embryos allowed us to have a relatively accurate 2D representation of organs that are actually in three dimensions. However, tissues such as the thyroid are more than a simple stacking of cubic cells. Sections can thereby cut through cells in various ways. A particular section, say the one located at the centre of the organ (as used here) can in fact show cells differently than they actually are and the profiles of epithelial cells observed on the scanned images are not always oriented in an identical way. As a consequence, using the mean value of epithelial cells as a validation indicator would be of little interest for such validation. Here, we compared the epithelial cell surface area distribution both *in vivo* and *in silico* as a more quantitative evaluation of our simulation results (Fig. 10.1). Note however that all the values from our model come from our simulation after only 72h.

At first glance, cell surface area distribution seems to be relatively similar. A first note is to be made on size units. In fact, while HALO measures the cell area in μm^2, the position of the rectangles enclosing each cell (Fig. 6.3, B) is given by the software in pixels. These values are also standardised in our IC generation process for convenience and visualisation purposes. There is therefore a difference in the numerical value of cell surface between our

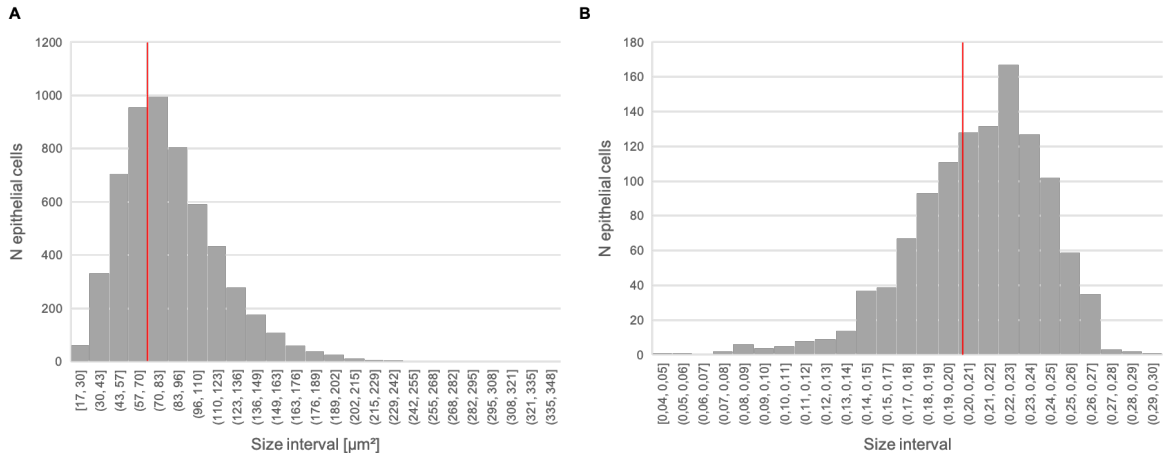


Figure 10.1: Epithelial cells size distribution in the embryonic thyroid comparison. (A) Epithelial cells distribution in the embryonic thyroid at E17.5 ($n = 3$). Each bar shows the number of lumen within the size interval indicated on the x-axis. The mean is indicated by the red vertical line. (B) Epithelial cells size distribution in our simulation at simulation time $t = 72$h.

experimental observation and our model results. However, we only worked using ratio, both for calibration and validation, making this difference non-relevant. Both distributions are non-normal (as proven by a Chi-Squared Goodness-of-Fit test), preventing any generic statistical tests (such as Student or ANOVA). However, for non-normal distributions of large samples, a Kolmogorov-Smirnov test can be performed to compare the two distributions. Unfortunately, this statistical analysis barely characterises these two samples as not similarly distributed at the 95% confidence level. Regardless of this result, a few comments can be made regarding some descriptive statistics. We can see for instance that the shape of the distribution is somehow inverted between the two cell population. In fact, at E17.5, a significant fraction of epithelial cells showed surface areas two to three times larger than the mean (μ_{exp}. In fact, around 4% of the cells had an area larger than $\mu_{exp} + 2\sigma_{exp}$ (σ_{exp} being the standard deviation), while 0% in the $\mu - 2\sigma_{exp}$ interval.

On the other hand, in our model, we noticed a non-negligible amount of small cells (4%) in the $\mu_{sim} - 2\sigma_{sim}$ while almost no cell had a surface area over $\mu_{sim} + 2\sigma_{sim}$. This shift could be explained by the fact that in our model, cells have a certain (imposed) tendency to reach a specified target area and have no real reason to grow above this value, while the global force equilibrium can cause cells to stay relatively small. This phenomenon is less observed *in vivo* where cells are more likely to occupy a larger space, and are more subject to intrinsic variability, even if they are still constrained to inherent cell-control mechanisms [103]. The high relative value of σ_{exp} (almost 40% of μ_{exp} in comparison to less than 20 % for

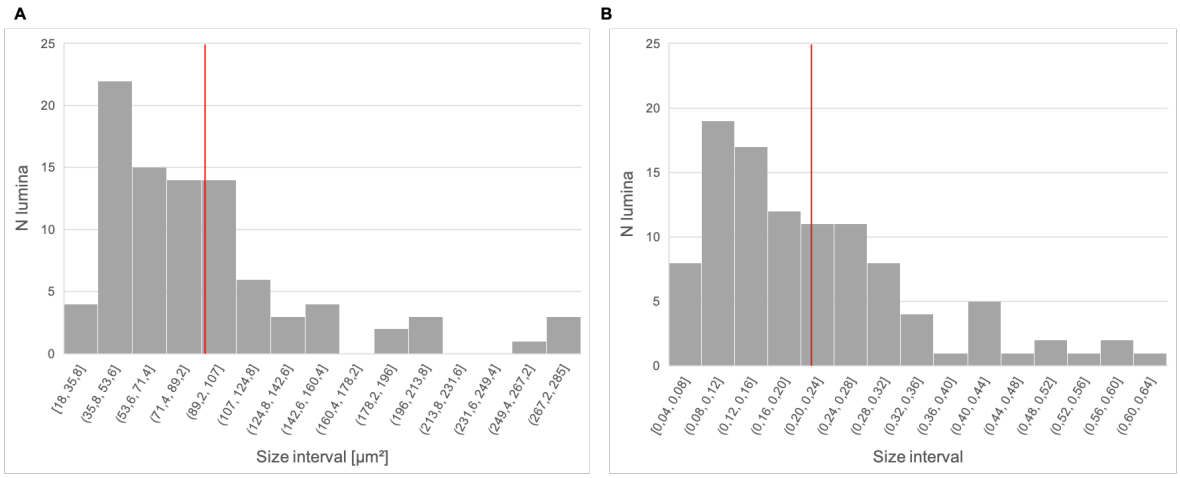


Figure 10.2: Lumen size distribution comparison. (A) Lumen size distribution in the embryonic thyroid at E17.5 ($n = 3$). Each bar shows the number of lumen within the size interval indicated on the x-axis. The mean is indicated by the red vertical line. (B) Lumen size distribution in our simulation after 72h.

μ_{sim}) supports this hypothesis. This could also explain why a statistical test would refute the comparison of these two distributions. Note for instance that in both cases, almost two-thirds of the cell populations showed surface area within the $\mu \pm \sigma$ interval, and similarly that 96% are within the $\mu \pm 1.96\sigma$ interval (which is precisely expected in a normal distribution).

10.2 Lumen surface area distribution

In the same way, lumen surface area distributions have been compared (Fig. 10.2). Once again, the units differ for the same reason as explained above. This time, they clearly do not follow a normal distribution. Nevertheless, both distributions showed similar behaviours. For instance, both distributions exhibit a maximum number of lumina in the second bin[1]. Moreover, between 20% (*in silico*) and 30% (*in vivo*) of the lumina have a surface area distributed around the mean (in the two bins around the latter). Note however that the ratio of the total epithelial cell number to the total lumen number is somewhat lower in our simulation (a bit less than 15, as compared to around 22 *in vivo*), meaning that our model generates slightly more lumina than it should. This could mean that, while the dynamics of lumen growth is successfully implemented, the rate of cell generation should be revised to be more consistent with our *in vivo* observations[2]. Even if few quantitative comparisons can be made on these lumen surface area distributions, it gives encouraging results regarding the

[1]Here, the distribution has been split into 15 bins, but similar results appear with, for instance, 25 bins.
[2]This issue is further discussed in chapter 11

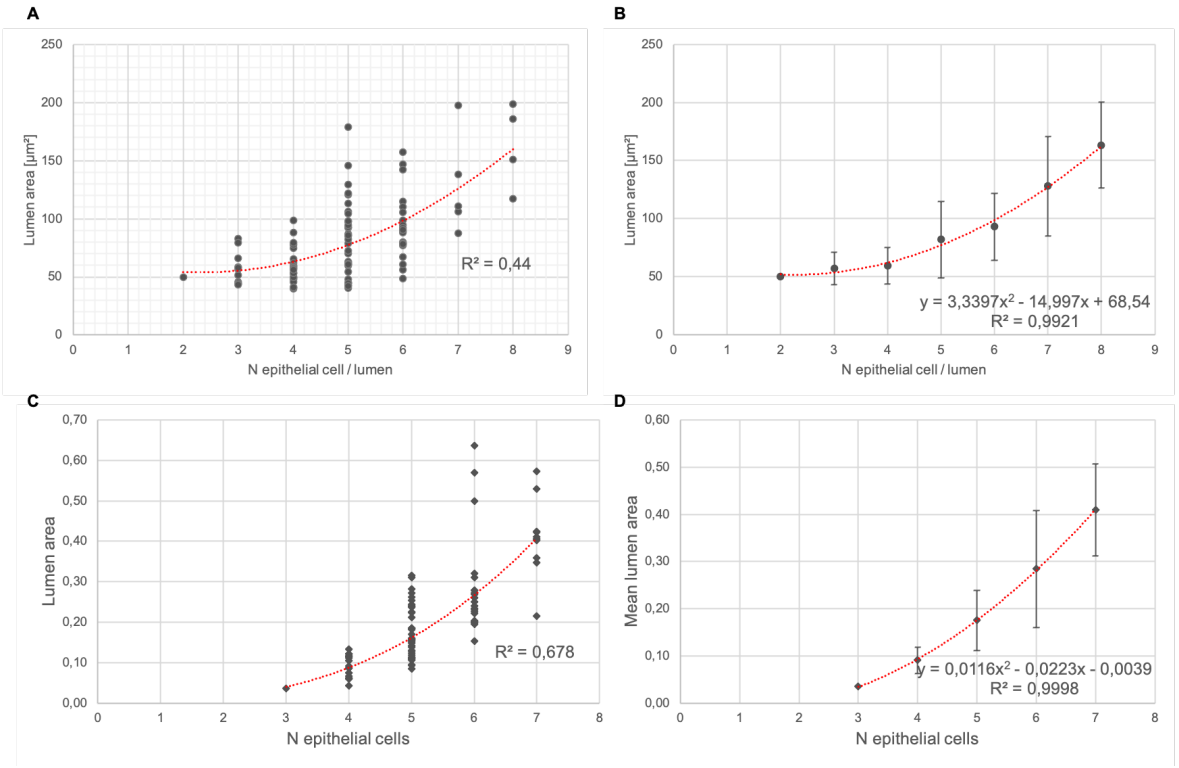


Figure 10.3: Epithelial cell per lumen in the embryonic thyroid dynamic comparison. (A) Evolution of the lumina areas observed in thyroid sections a E17.5 according to their respective number of neighbouring epithelial cells (n = 1). Each dot represents a lumen, quadratic regression showed in red. (B) Evolution of the mean lumen area observed in thyroid sections a E17.5 with respect to the number of neighbouring epithelial cells. (C) Evolution of the lumina areas in our model at simulation time $t = 72$h according to their respective number of neighbouring epithelial cells. Each dot represents a lumen, quadratic regression showed in red. (D) Evolution of the mean lumen area in our model at simulation time $t = 72$h with respect to the number of neighbouring epithelial cells.

dynamics of lumen formation and growth in our model.

10.3 Epithelial cells to lumen surface area ratio

An interesting feature that was not mentioned earlier is the evolution of the number of epithelial cells with respect to the surface area of the lumina. We explained in the introduction (see Sec. 2.1) that the follicular units of the mature thyroid take a spherical form. This particular architecture requires the adaptation of the shape of the cells themselves, in order to correctly organise around the spherical colloidal space. We analysed sections of embryonic thyroid at E17.5 and reported the number of epithelial cells surrounding each visible lumen, as well as its surface area (Fig. 10.3).

These observations suggested that the mean surface area of the lumina evolves quadrati-

cally with the number of epithelial cells surrounding it (Fig. 10.3, B). This dynamic is also observed *in silico* (Fig. 10.3 C and D) and can thus be compared with our *in vivo* observations. Note that other regressions were also applied to our data and the quadratic regression indeed appeared to be the best fit. Our model seems to correctly represent this quadratic evolution, even though the slope of the regression is slightly higher but still comparable (one needs to remember that the scales of the axes are different). Note that since lumina are represented as cells, adhesion parameters between epithelial and lumen cell types are also to be set accordingly. No calibration has been carried out which could explain the difference observed between the two quadratic regressions. All these results nevertheless suggest once again that our model seems to adequately simulate the dynamics of lumen formation.

11 | Discussion

In this work, we created a bridge between experimental science and mathematical modelling, transforming qualitative and descriptive biological information into quantitative data that can be used to drive a process-based morphogenesis model. Specifically, we used the developing thyroid as a biological system in which a mass of non-polarised epithelial cells is transformed into a multitude of polarised monolayers forming the follicles, surrounded by blood vessels. The mathematical model, developed from *in vivo* images, correctly simulated the differential growth and proliferation of central and peripheral epithelial cells, and the VEGFA-induced invasion of endothelial cells. We showed that fission of the multi-layered epithelial mass into independent pre-follicular units was controlled by endothelial cell invasion, combined with reduced epithelial cell-cell adhesion. The model also revealed that endothelial cell abundance and proximity to epithelial cells promoted basal-apical polarisation of the latter and subsequent follicular lumen formation. However, computational instabilities prevented our model to further simulate the constant migration of new endothelial cells, leading to a crucial lack of endothelial vessels. This was a deception because we expect that new tip cells should appear from the outside of the epithelial mass and from the newly formed virtual vessels to invade the newly formed epithelial islets. As a consequence, our model didn't succeed to perfectly represent the formation of individual single-layered follicles.

11.1 Vertex-based model and Initial conditions

While a wide range of model types are used in computational biology, in the case of epithelial monolayers, the most commonly used framework is the two-dimensional vertex model [61]. VMs are indeed often considered as highly appropriate when it comes to cell proliferation, adhesion and long-range signalling [44] and Chaste appeared as a suitable framework to develop our model. In most VMs, initial configurations are usually set as simple honeycomb-shaped cellular sheets. Here, we considered initial conditions closer to reality and directly

derived initial conditions of the model from microscopic images of developing thyroid. We believe that our approach is a fast, easy but precise way to obtain realistic initial conditions for mathematical models of morphogenesis or tumour development.

Note that the actual effect of IC on the results of our simulations is not assessed here. While this could indeed be done, for instance, by running simulations with various IC from different embryos, one must remember that animal use should be reduced. Another strategy would be to assess the model sensitivity to the initial conditions by running stochastic simulations. That amounts to add some stochastic variability (*i.e.* noise) characterised by a specified s.d. to the initial cell distribution. Various simulations could be run where each would start from a slightly different initial cell distribution, and the variability in the final cell distributions could be assessed and compared with the initial variability. The same sensitivity analysis could be performed by adding variability to the initial cell size distribution or to any other model component (adhesion, abundance of endothelial cells, etc.) that could influence the evolution of the tissue.

11.2 Epithelial cell proliferation

We successfully characterised and computed the proliferation dynamics of what we called "peripheral" and "central" epithelial cells in the murine embryonic thyroid, and used these data in the development of our cell-cycle models. Our observations of cell growth dynamics in the embryonic thyroid are also consistent with other studies using BrdU [15, 88]. Moreover, Liang et al. [87] reported and proposed a branching growth for the thyroid, similar to the pancreas or other branching organs, and found more proliferation in subpopulations of Sox9+/Fgfr2+ cells localised at the periphery of developing thyroid lobes, as compared to more centrally located cells, thereby generating branch- or cord-like growth [87]. Although the proliferation kinetics of adult thyroid epithelial cells have been well studied [104–106], to the best of our knowledge this is the first report providing quantitative and spatiotemporal information on the proliferation rate and pattern of embryonic thyroid epithelial cells and proposing a calculation for the cell cycle duration of peripheral and central cell populations. Our observations that the majority of EdU+ cells are located at the periphery of the thyroid primordium, combined with those of Fagman [15] and Liang [87], suggest that thyroid epithelial growth could rely on the existence of proliferating epithelial tip cell subpopulations as described in the pancreas [107].

11.3 Epithelial adhesion and cohesion

Besides ECs expansion and invasion, epithelial cell adhesion appeared as a second critical parameter for the fragmentation, or fission, of the thyroid lobes. Note that as explained in the introduction (see Sec. 3.1), this feature is not explicitly defined here but adhesion weight parameters are rather specified to describe the potentially differential adhesion and cohesion between different cell types, as well as between cells and the periphery (ECM). Here, we mostly used values inspired from the literature [77] and adapted them when needed in order to satisfy our observations. In fact, these parameters do not refer to actual biological values. Even though, measuring forces acting between cells *in vivo* is no easy task. Methods such as laser ablation or forces inferences by estimation of the tensions and pressure between cells from microscopic imagery [108] exist but were not used here.

As explained earlier, the cohesion of the thyrocytes monolayer is mainly ensured by junctional complexes formed by E-cadherins (see Section 2.1). However, besides the classical epithelial (E) cadherin, another cadherin molecule, R-cadherin (Renal-specific cadherin), now called cadherin-16, has been reported to be induced in the thyroid epithelium and maintained in the adult thyroid [109]. Furthermore, a role for cadherin-16 was proposed in the developing thyroid when it was observed that E-cadherin (Cdh1) knockout had normal follicles and only caused mild defects to the organ [110]. Cadherin-16 has also been identified as a downstream target of the transcription factor Pax8, expressed in the thyroid and kidney [111]. More recently, using thyroid cells in culture, Pax8 was found to control apico–basal follicular polarisation and follicle formation through Cdh16 [112]. In addition, silencing of Cdh16 expression led to the formation of defective follicular structures characterized by very low laminin expression at the follicle–matrix interface [112]. Interestingly, in Smad1/5 double knockout thyroid that presents defective thyrocyte polarisation and follicle formation [17], expression of Cdh16 was also dramatically reduced at E14.5 (Villacorte and Pierreux, unpublished).

To compare the abundance of these two cadherins, we therefore measured the absolute number of mRNA molecules for E-cadherin (Cdh1) and R-cadherin (Cdh16) in developing thyroid (Spourquet and Pierreux, unpublished). Although we measured 16× more Cdh1 as compared to Cdh16 mRNA molecules at E13.5, we found 1.2× more Cdh16 than Cdh1 molecules at E17.5 (Fig. 11.1, A). In other words, and if compared to epithelial cell abundance, while at E13.5, each epithelial cell contained 50 Cdh1 and 3 Cdh16 mRNA molecules,

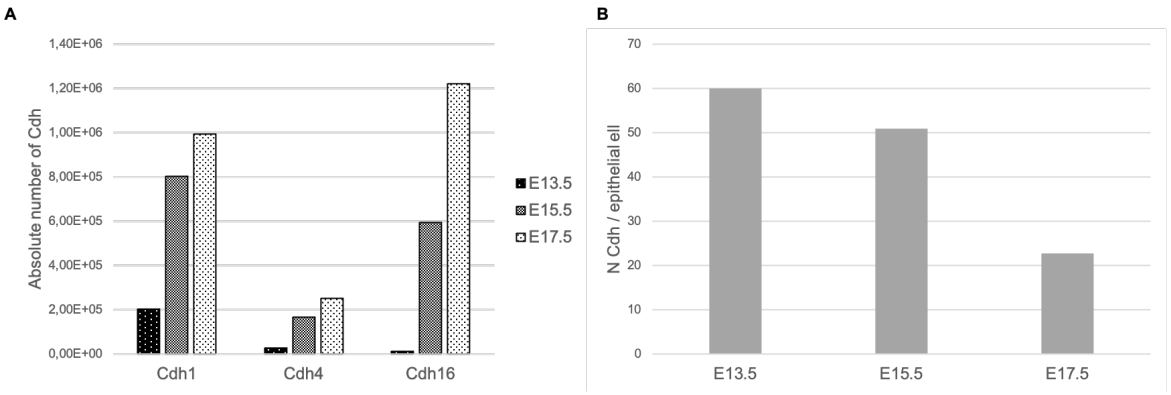


Figure 11.1: Quantification of the major epithelial cadherin families in the embryonic thyroid at three different embryonic stages. (A) Absolute number of cadherin proteins at E13.5 (black dotted), E15.5 (grey dotted) and E17.5 (white dotted). (B) Number of cadherin molecules per epithelial cell in the embryonic thyroid (per thyroid lobe) at three different embryonic stages. This number of cells is computed using the modified Cavalieri method explicited in Box 4.

at E17.5 each epithelial cell contained 9 Cdh1 and 11 Cdh16 mRNA molecules. Thus, fewer Cdh molecules per cell and a different ratio of these two cadherins (Fig. 11.1, B). Absolute numbers were obtained by qPCR while the relative number, *i.e* the number of Cdh molecules per epithelial cell, were computed using the total number of epithelial cells estimated by the modified Cavalieri method. These data on Cdh16 suggests that Cdh16 expression at the time of folliculogenesis could reduce epithelial cell-cell adhesion by interfering with Cdh1. This could be explained by the fact that in order to reorganise into single-layered follicles during morphogenesis, epithelial cells need to modify their cell-cell contacts (Fig. 13.1, A). In fact, a transition from a mass to a monolayer induces a decrease in the number of cell neighbours for each epithelial cell, hence a potential decrease in the number of Cdh per cell, as suggested by our quantification results.

It is however still not clear how a number of cell adhesion molecules (CAM) measured *in vivo* could infer a proportional variation in homotypical cells cohesion. In their computational model of monoclonal conversion in the colonic crypt, Mirams and colleagues [113] hypothesised that a ten-fold in β-catenin observed *in vivo* would correspond to a similar ten-fold increase in what they refer as the adhesion scaling parameter (*i.e.* parameter characterising the differential adhesion between wild-type and mutant cells). This parameter increase is nonetheless a potential overestimation of the real change in cell adhesion as other factors, such as the availability of certain other adhesion proteins, may be taken into account [113]. The observations discussed here were therefore not used as such in our model calibration.

Even so, to test this hypothesis, we ran our simulations for a simulation time of $t =$ 1h and with the same IC, this time with different values for epithelial cell to epithelial cell adhesion parameter (Fig. 11.2). We observed that with small values for this parameter (hence a high cohesion), and with all the other parameters kept constant, tip cell were not able to migrate inside the epithelial mass (Fig. 11.2, A and B). On the other hand, with higher values of cell-cell adhesion parameter (λ_{e-e}), endothelial cells velocity was increased, resulting in an altered vessel network configuration 11.2, A and D). Moreover, the global force balance seemed to induce a change in the shape and size of epithelial cells. In fact, with small values of the cell-cell adhesion parameter, the differential adhesion forces model tend to maximise the length of edges between epithelial cells, hence their surface area. A high value of this parameter had an opposite impact, triggering an apparent change in cell-ECM contact despite the fact that this parameter was not modified. Altogether, these results illustrate the variation of both epithelial and endothelial cells behaviour resulting from a change in the adhesion forces. This could suggest that the observed loss of cohesion (or at least a loss of relative adhesion proteins) between epithelial cells *in vivo* is needed to allow the invasion of endothelial cell in the thyroid primordium and its subsequent fragmentation.

This would not be the first example of morphogenesis processes induced by dynamical changes in cell cohesion. For instance, Ko and Lobo [114] showed, using a continuous model, that the regulation of cell adhesion molecules expression by the diffusion of the morphogen Nodal[1] induces specific pattern and cellular behaviour resulting in the involution of the mesendodermal germ layer in the zebrafish. Similarly, Rui Zhen Tan and Keng-Hwee Chiam developed a CPM to demonstrate that chemotaxis and differential adhesion were needed for robust cell sorting in the zebrafish neural tube [115].

The same method was applied for the study of integrins, which are transmembrane receptors involved this time in cell-ECM adhesion. They are formed by the association of two subunits : one molecule of integrin α (Itga) and one molecule of integrin (Itgb). We quantified the absolute (Fig. 11.1, A) and relative (Fig. 11.1, B) number of Itga proteins in the embryonic thyroid at three different stages. Once again, absolute numbers were obtained by qPCR while the relative number, *i.e* the number of Itga molecules per epithelial cell, were computed using the total number of epithelial cells estimated by the modified Cavalieri method. We observed a similar behaviour as for the Cdh quantification, with a significant drop in the number of integrin molecules between E13.5 and E17.5. This may come as a

[1]Nodal is a protein encoded by the gene NODAL. It belongs to the Transforming Growth Factor beta (TGF-) superfamily.

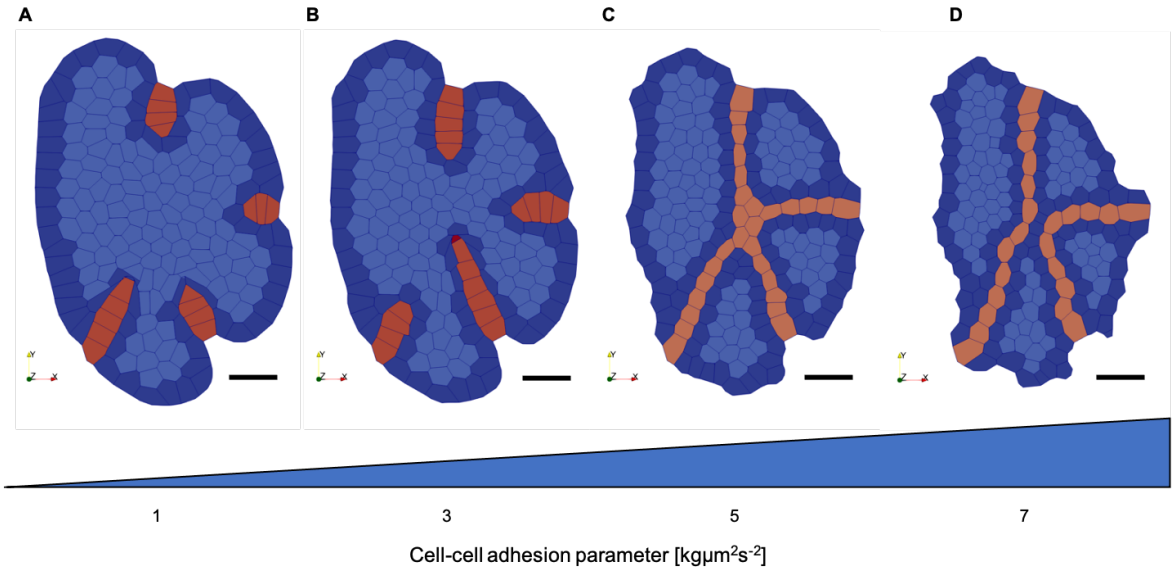


Figure 11.2: Snapshots of a VM simulation of the epithelial primordium at simulation time t = 1h and for different values for the epithelial cell-cell adhesion parameter (with every other parameter values kept the same). (A) Final configuration with cell-cell adhesion parameter = 1. Tip cells are not able to migrate despite the implementation of a motile force. (B) Final configuration with cell-cell adhesion parameter = 1. Some tip cells manage to slowly migrate towards the gradient but do not reach the center after 1h. (C) Final configuration with cell-cell adhesion parameter = 5 (normal configuration). (D) Final configuration with cell-cell adhesion parameter = 7. Low adhesion between epithelial cells favours a faster migration of endothelial tip cells. Epithelial cell appears smaller due to the global force balance.

surprise since, according to our initial hypothesis (Fig. 13.1, A), we would have expected an increase in the relative number of Itga molecules since the very definition of a monolayer implies that every epithelial cell has contact with the ECM. However, mature thyroid follicles are spherical, and as demonstrated in the introduction (see Box 3), a sphere (or a disc in 2D) is the most energy-efficient configuration. These follicles obviously do not lie in an empty space as foam bubbles would, but we can assume that epithelial cells would induce a decrease in their relative number of Itga molecules, hence a loss in cell-ECM adhesion, to facilitate the transition to individual spherical monolayers. In addition, models of epithelial cell polarisation based on *in vitro* data suggest that Itg are expressed continuously by non-polarised cells and would reside on all the faces of the cuboidal cell [116]. Upon polarisation, integrin would relocate on one pole, the basal pole, of the epithelial cell. Lastly, in contrast to the three cadherins selected (Cdh1, 4 and 16, see Fig. 11.1), these integrin molecules are also expressed by invading endothelial cells. The observed decrease (Fig. 11.3, B) may thus be exaggerated.

We therefore used these data to develop a model of dynamic cell-ECM adhesion applied

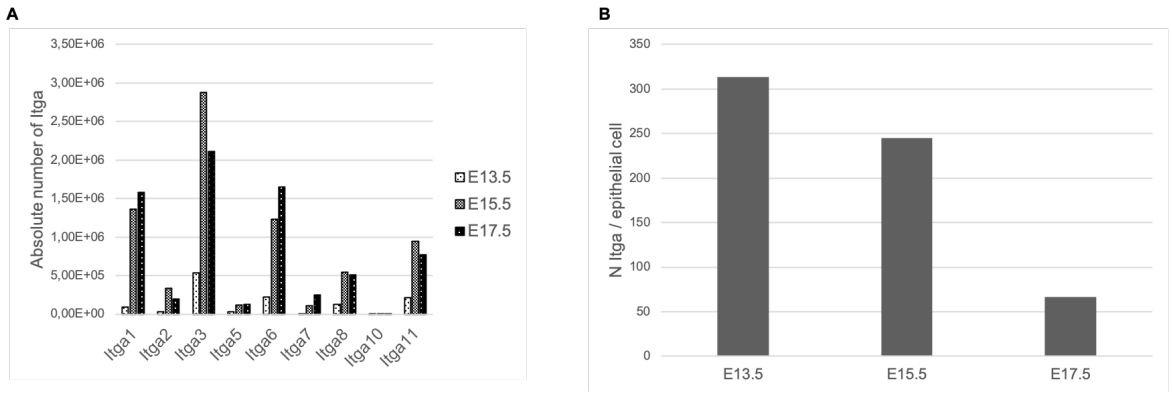


Figure 11.3: Quantification of the major integrin families in the embryonic thyroid at three different embryonic stages. (A) Absolute number of Itga proteins at E13.5 (white dotted), E15.5 (grey dotted) and E17.5 (black dotted). (B) Number of Itga molecules proportional to the total number of epithelial cells in the embryonic thyroid (per thyroid lobe) at three different embryonic stages. This number of cells is computed using the modified Cavalieri method explicited in Box 4.

to every peripheral epithelial cell in order to potentially validate our last hypothesis. We used our model to run a simplified simulation of initial 4×4 honeycomb cell population, for simulation time $t = 96$h and with our modified differential adhesion force (Fig. 11.4). We observed that this modified adhesion model had a notable effect on the epithelial cell organisation. The continuously decreasing cell-ECM (or here cell-boundary) contacts indeed lead to the formation of two more circular individual epithelial structures organising around the lumen (Fig. 11.4, B).

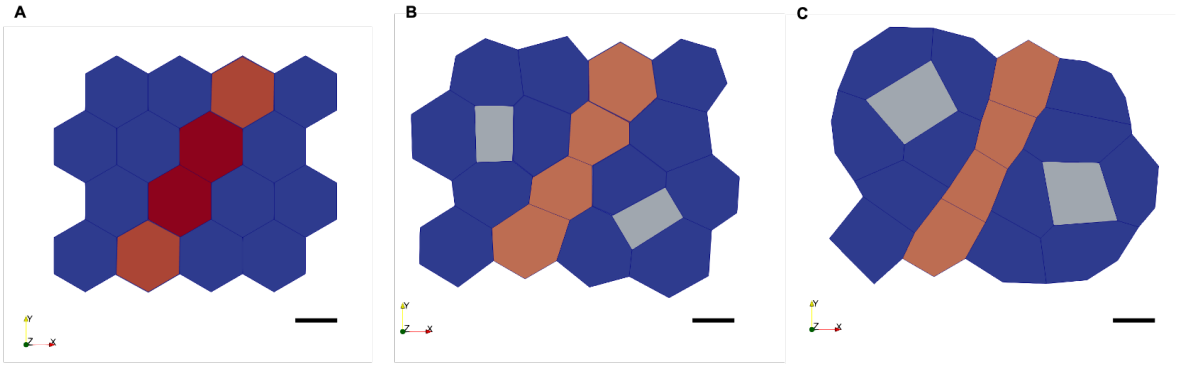


Figure 11.4: Snapshots of a VM simulation of two simple follicles formation. (A) Simple initial honeycomb configuration. Two cells are labelled as tip cells and two cells as stalk cells. (B) Configuration at simulation time $t = 96$h without cell-ECM dynamic adhesion. Two lumen are formed and epithelial cells organise around this lumen, taking rather hexagonal shape. Endothelial motility is implemented but not visible since a vessel is already formed. (C) Configuration at simulation time $t = 96$h with implementation of the cell-ECM dynamic adhesion model. Two lumen are also formed. This time, epithelial cells organise around this lumen to form rather spherical structures due to a decrease in their adhesion with the ECM.

An interesting parallel can be drawn between the number of epithelial cells to lumen surface area ratio and the change in cell-ECM adhesion regarding the final shape of these epithelial cells. In fact, if we consider the lumen as a regular polygon (as we actually assume in our model) of n sides with n being the number of epithelial neighbour cells, we can make the assumption that the surrounding epithelial cells would take a hexagonal or pentagonal form in order to best fill the remaining space. However, we illustrated how the follicles eventually formed spheres (or discs in 2D). We can therefore legitimately assume that these spheres (or discs) would have a radius proportional to the colloidal space that is also spherical (as we can clearly see in Fig. 9.3). The most efficient way for the cells to form a sphere while staying coherent is therefore to adopt a trapezoidal configuration, with a small (curved) edge (*i.e.* their apical membrane) facing the colloidal space and a larger curved base-line (*i.e.* their basal membrane). Interestingly, this architecture also allows the cells to keep a rather constant surface area no matter the size of the lumen.

Crosstalk between adhesion and cohesion has been demonstrated at a cellular level in cancer invasion and metastasis, in which the dynamics between the two lead to an increased plasticity of tumour cells [117]. However, the role of this interplay is still unclear in organogenesis processes, such as thyroid morphogenesis. Our model results eventually suggest that a dynamic decrease in both the cell-cell and cell-ECM adhesion could favour the segmentation of the thyroid primordium by endothelial cells as well as the development of spherical units around growing lumina.

While this is illustrated for a more simplified initial condition and for a significantly smaller number of cells, we could expect our model to behave similarly, provided the implementation of a more stable angiogenesis model. However, this was unfortunately not particularly visible in the results of our simulations. So far, the model indeed fails to simulate the eventual formation of individual follicles as observed *in vivo*. This can be mainly explained by the fact that our angiogenesis model, implementing the dynamic formation of new endothelial branches, was not stable enough in the long run (hence not implemented) to allow the formation of a dense and more complex endothelium network. This has multiple implications in our final simulation results as well as in our validation process. First, all the data computed *in silico* therefore corresponds to an earlier embryonic stage than E17.5, for which our experimental quantification and image analyses have been made. This could imply that all our validation results should be slightly adapted in order to fit with our actual *in vivo* data, despite satisfactorily reproducing the overall dynamics. Secondly, the formation

of spherical units relies, as we have seen, on a change in the cell-ECM contacts. This change is only implementable for peripheral cells , that is to say the cells that are supposed to be resting on a basement membrane. In our model, this cell type is applied to every cell in contact with either the periphery of the cell population or with the endothelial cells, as they partly control the assembling of this basement membrane. With fewer endothelial cell, it was inevitable that a significantly smaller fraction of epithelial cells would indeed acquire this property, preventing the formation additional of spherical-shaped individual units. Finally, this also explains why, while epithelial cells adequately gathered around the lumina, we did not observe the expected change in their shape.

11.4 Morphogen diffusion and endothelial density

Our angiogenesis model was able to correctly simulate the formation of an endothelial network within the epithelial mass in response to a morphogen-dependent gradient. Despite the computational instabilities preventing the dynamic evolution of this network by the addition of new tip cells (see following section), an interesting note can be made regarding the representation of endothelial cells in VMs. In fact, while VMs are generally accepted as a powerful tool to simulate epithelial cell sheets, endothelial cells exhibit notable differences in their shape and proliferation behaviour. Various mathematical and computational models tackling the issue of angiogenesis exist, ranging from complex continuous 3D models such as "Microvessel Chaste" [118] to more simple cell-based models [119–121]. On the other hand, various models of cell polarisation have been developed [100]. For instance, Hazelwood and Hancock used a cell-based model to predict that three particular proteins are required to induce intrinsic cell polarity in *Drosophilia* wings [122]. Moreover, they showed that two of them were additionally required for transmitting this cell polarity. An interesting parallel can be drawn with our model or cell polarisation. In fact, they treated each cell as a dipole (with their specific vector pointing respectively from and towards these poles), dynamically coordinating its angle with its neighbouring homotypical cell. This is therefore in line with our definition of cell polarity being considered as a "contagion" process between neighbouring homotypical cells. However, this does not integrate the mechanical interactions of endothelial vessels with the epithelium. In our model, we described endothelial and epithelial cells in the same framework in order to study the various impacts of the growing capillaries network on epithelial cell polarisation and thyroid follicles formation. We showed that cell motility,

proliferation and adhesion were somehow acceptably simulated in our model, despite some inaccuracies (notably regarding endothelial cell size[2] and density).

Interestingly, our model can be used to consider the effect of the suppression of morphogen production (hence diffusion), in order to illustrate *in silico* the potential effect of a VEGF KO *in vivo*. We ran the simulation for $t = 24$h and $\Delta t = 0.001$h. All stalk cells were initially fixed (or fixed upon addition) and the production term in the PDE equations was set to 0 for all epithelial cells.

We observed that while epithelial cells proliferated inducing the growth of the epithelial sheet, tip cells did not migrate and stayed relatively fixed at their initial position. New endothelial cells are still added, but with no subsequent migration. No endothelial network was therefore formed and the endothelial density after 48h was far lower than in the normal condition (Fig. 11.5, B). However, the few endothelial cells still had a polarisation effect on epithelial cells, as well as the periphery of the epithelial cell population (Fig. 11.5, C and D). This polarisation level is nevertheless too low to generate any lumen, except at a very few locations, and always at close proximity to initial endothelial cells. These results suggest that no morphogen production, leading to a loss in endothelial density, would prevent the polarisation of epithelial cells and thus the formation of lumina, which is consistent with our initial hypothesis (Fig. 2.9 and 13.1, B). The similarity between this simulation and the phenotype of the VEGFA KO indicates that the various processes and parameters are correctly implemented. Moreover, this once again highlights the crucial role of endothelial cells, and consequent basement membrane assembling, in the development of thyroid follicles.

In our results presented in chapter 9, our model illustrated how this polarisation could propagate in the epithelial mass, due to contact between cells with opposite polarity or with lumina. However, our simulation results showed the development of lumina rather homogeneously in the mass (Fig 9.5, D). While the lumina located in the center of the developing individual epithelial structures are significantly smaller, we would expect that the lack of sufficient endothelial vessels prevents the growth and expansion of lumina at such location. This could suggest that certain parameter value, notably the effect of epithelial cells polarity on neighbouring epithelial cells, should be revised. We could indeed expect these lumina to grow even larger in the missing last hours in our simulations. A further calibration,

[2]Endothelial cell size and shape were not quantitatively characterised here but it is well know that endothelial cells usually take a more flat shape, with a larger perimeter than epithelial cells, as observed on our scanned images.

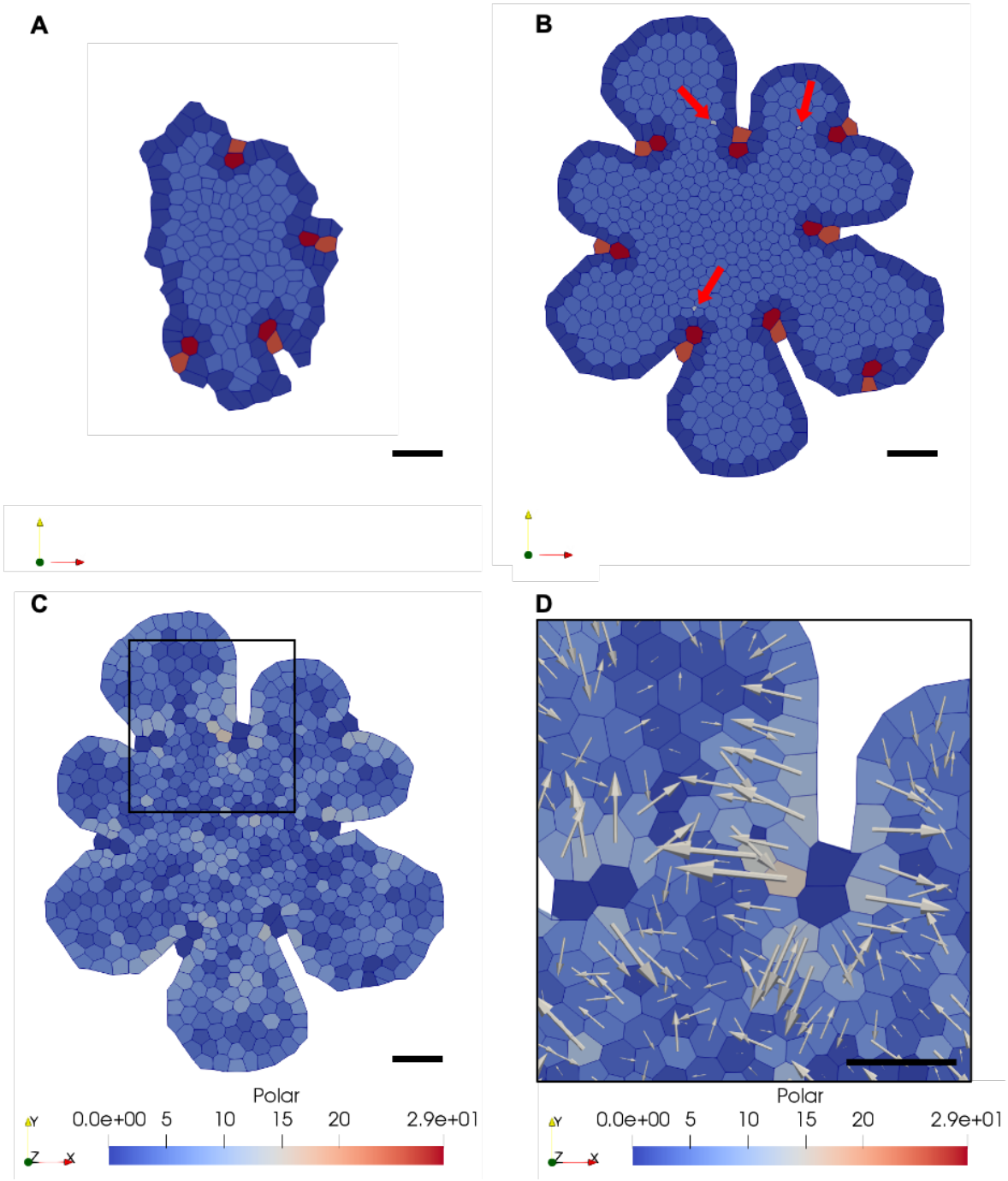


Figure 11.5: Snapshots of a VM simulation of the epithelial primordium with no morphogen diffusion. (A) Configuration at initial time. Endothelial tip cells are labelled in red, endothelial stalk cells in dark orange. Each tip cell is directly followed by a stalk cell. (B) Configuration at simulation time $t = 48$h with no morphogen diffusion. Stalk and tip cells stay fixed at their initial position but do not migrate. New endothelial cells are still added throughout the simulation. Small lumina (indicated by a red arrow) can be observed due to the polarisation of epithelial cells close to existing endothelial cells and/or the periphery of the cell population. (C) Configuration at simulation time $t = 48$h with now morphogen diffusion. Cells are coloured according to their polarisation level. (D) Close-up view of (C). Cell polarity is represented by an arrow (polarity vector) for each epithelial cell. Arrows are scaled according to the magnitude of the polarity vector.

providing the implementation of a stable angiogenesis model, could allow for a more accurate definition of our model parameters. Moreover, this could also lead to the formation of a more adequate number of final lumina, as expressed in chapter 9.

11.5 Computational instabilities

Criticism of this model can be levelled at various points. One of its major flaws is obviously the lack of stability most probably emerging from our angiogenesis model. In fact, the addition of new stalk cells coupled with the implementation of endothelial BC always triggered instabilities resulting in the failure of our test simulation. This could be explained by the fact that we tried here to assemble various modules, each simulating specific and somewhat very different processes, in one substantial simulation. In fact, the model is expected to deal with different processes, at different scales and for different cell types.

But the true challenge here resided in the definition of endothelial BC. In fact, transformation of the initial epithelial sheet into smaller islets was only possible when the issue of endothelial boundaries was solved by imposing "no man's land" regions behind the most peripheral ECs. These endo-BC were essential to simulate the connection with distant peripheral vessels. Indeed, these connections are not visible in 2D histological sections and were thus absent from the initial condition of our vertex model. We faced two other problems with ECs in our 2D VM that we could not solve and that led to under-representation of ECs. First, ECs coming from the third dimension, *i.e.* from above or below the analysed plane or epithelial sheet, could not be modelled. Secondly, due to stability issues, appearance of new tip cells was possible during the first 24h of our simulation but not beyond. This was a deception because we expect that new tip cells should appear from the newly formed virtual vessels to invade the newly formed epithelial islets. Moreover, anchoring specific vertices (i.e. boundary vertices belonging to the most external stalk cells) adds another level of complexity to the model. However, without anchoring these vertices, the ECs would simply behave like any other epithelial cells, even with different adhesion parameter values, and would always relax to a spherical configuration with no fragmentation of the initial sheet (as shown in Fig. 7.3). Yet by anchoring these vertices, we had to ensure that the epithelial cells belonging to different islets would not fuse and eventually lead the simulation to similar results hence the development of specific endo-BC. However, junctional rearrangements are still performed on endothelial cells and can lead to the motion of vertices that we would expect not to move,

typically for peripheral stalk cells. This can lead to very small, if not too small edges, that our simulations can no longer handle.

We initially postulated that VM would be the most suitable framework in the case of our particular study [44]. And while it gave, in our opinion, satisfactory results, the development of various modules however highlighted the limitations of such models. For instance, our model, as all VM, represent cells as polygons. Vertices are linked together by edges to form cells, generating a coherent population. However, the very definition of VM prevents the existence of stable "voids" in the simulation. This has been overcome by defining lumen as particular cells. The easiest way to generate new cells in VMs is simply by cell division. Adding vertices and cells is technically possible but difficult when it comes to generating new cells inside the epithelial cell population. Lumina were therefore generated by the division of an epithelial mother cell in two daughter cells. One daughter cell remained epithelial and the other one became a lumen. We developed our polarisation model so that the lumen cell would develop at the "apical edge" of the mother cell (i.e. at the extremity of the mother cell polarity vector). Lumina are also known to have a spherical form. It is obviously not the case in our simulation since lumen is represented by a particular polygonal cell, defined by vertices and edges. A circular form will therefore never be achieved but lumina grew in size as a function of the number and polarity of epithelial cells surrounding each of them. Despite what we considered as satisfactory results, other models such as overlapping-spheres models (OS, off-lattice models where cells are represented as a sphere with a single node as cell-centre) or cellular Potts models (CPM, lattice-based models where each cell is represented by several lattice sites) could have potentially handled this issue more efficiently. In fact, in OS, cells are not explicitly linked together in a coherent population and could allow the formation of "voids" or intercellular spaces. However, this also directly implies that OS are less appropriate when it comes to simulate processes such as differential adhesion or engulfment and that they offer a poor representation of the cell shape in an epithelial tissue (unlike VMs). In CPM, lattice sites can be "empty" and therefore also considered as such. Note however that the formation and growth of the lumen cells in our model depend on our polarisation model, itself based on the definition of polarity vector. Providing that the implementation of such model is technically feasible, this framework should also be considered here. Moreover, many of the instabilities in fact emerged from our model not being able to handle small edges and/or fixed vertices. Simple models such as CPM do not depend on the motion of such computational objects and the definition of endothelial BC could therefore be

simplified. Moreover, processes such as diffusion and motility can also be implemented quite straightforwardly in this framework. In fact, in off-latice models (such as VM), the solving of Reaction-Advection-Diffusion equations is easily implemented using the mesh created by the nodes (or vertices). And while simple models such as CA are clearly not designed to deal with this level of complexity, in CPM, this discretisation of the spatial domain is made by calculating the constrained Delaunay triangulation of the centres of the cells [44]. Even though this is relatively more computationally demanding due to the cost of calculating this mesh, CPM is still considered as an appropriate framework for simulation long-range signalling. Moreover, in our simulations, the VEGFA gradient is evolving with modelling time; and upon fragmentation of the initial sheet, and with the appearance of new epithelial islets, secondary VEGFA gradients should develop in every islet, thereby redirecting ECs migration. This also could be achieved in CPM as the tesselation in fact changes over time as the tissue grows and changes its shape. In light of all these considerations, CPM should therefore be considered as a viable alternative for the development of 2D models of thyroid morphogenesis. However, we mentioned earlier how a major drawback related to the use of CPM is the time evolution. In that extend, more complex models such as Deformable Cell Model (DCM), that can arguably be seen as an extended VM, could emerge as a more suitable framework [123].

12 | Conclusion and Perspectives

Developing such a complex model was undoubtedly a real challenge. If the initial question was rather well-defined, the potential ways to eventually provide elements of answer were more clouded. However, after four years, we managed to clear the path and to enlighten some of the crucial mechanisms at stake in thyroid morphogenesis. We used a combination of experimental and computational approaches to build, calibrate and validate our *in silico* 2D model of thyroid growth, angiogenesis and folliculogenesis.

We showed that so-called central and peripheral cells exhibit different profiles of proliferation and integrated this in our model to accurately simulate the differential growth and proliferation of epithelial cells. Using EdU labelling from E13.5 to E17.5, we precisely quantified and localised cells in the S-phase of the mitotic cycle. These experimental data confirmed increased proliferative potential at the periphery as compared to the centre, decreased thyrocyte proliferation from E13.5 to E17.5 and allowed to compute the cell cycle duration of central (31h52) and peripheral (11h20) epithelial cells. This topological and quantitative information was instrumental for deriving key model parameters. We also developed computational algorithms to allow us to use initial conditions based on *in vivo* images. Our model also integrates the morphogen-driven migration of endothelial cells. We showed that a combination of tip cell motility and stalk cell proliferation was sufficient to generate a network of coherent endothelial vessels. In fact, our model correctly simulated the formation of an endothelial network within the epithelial mass (here modelled as a sheet), in response to a VEGFA-dependent gradient. We indeed imposed the expression of VEGFA to the epithelial cell type forming our initial condition. This resulted in the formation of a centro-peripheral gradient of VEGFA that induced migration and invasion of endothelial cells towards the centre of the epithelial sheet. Peripheral to central ECs invasion, concomitant with outward expansion of the epithelial sheet, allowed fragmentation or fission of the initial sheet in smaller growing epithelial islets. Our model also simulated apico-basal polarisation of epithelial cells

and subsequent follicular lumen formation. We developed and used an epithelial polarisation model by assigning a polarity vector to every epithelial cell. This allowed an explicit definition of the polarisation degree of epithelial cells required before generating a lumen. Our model can also be used to test the importance of biological parameters. Knowing the *in vivo* effect of VEGFA KO on thyroid folliculogenesis, our model was validated by suppressing the VEGFA morphogen production (hence diffusion), *in silico*. In the absence of VEGFA production in the epithelial sheet, endothelial tip cells do not migrate and remain at their initial position. The endothelial density after 48h was significantly lower than in the presence of a VEGFA gradient. The few endothelial cells present had a polarisation effect on the surrounding epithelial cells, as well as the periphery of the epithelial sheet. This polarisation effect was nevertheless too low to generate large and numerous lumina. The similarity between this simulation and the phenotype of the VEGFA KO indicates that the various processes and parameters are correctly implemented. Moreover, we used our model to test the potential effects of dynamical changes in epithelial cell-cell and cell-ECM contacts on endothelial migration and epithelial reorganisation. Data on Cdh16 previously acquired in our lab combined with our observation that epithelial cell-cell adhesion should be reduced to allow ECs invasion highlighted the crucial role of CAM in thyroid folliculogenesis.

Nevertheless, our model still has room for improvement. The lack of robustness is significantly holding back its further development and use. However, we managed to make to most out of what we did with what we had. Further calibration or validation, for instance using more data at different embryonic stages, may be required to enhance the accuracy of such model. Moreover, we highlighted the strengths, but also the potential weaknesses of vertex models in the context of this particular issue. While processes such as cell proliferation and adhesion can be correctly simulated, the variety of cellular types (epithelial, endothelial, lumen) coming with their own shape and dynamics definitely added some complexity to our model, that our VM model could eventually not handle. Considering other frameworks, such as Cellular-Potts models or Deformable Cell Model could therefore be useful in order to best suit our biological questions and hypothesis.

Models should not be judged by their complexity or level of details, but by what one can learn from them. Here, the definition and implementation of numerous modules, accounting for various processes involved in the morphogenesis of the embryonic thyroid, altogether gave encouraging results. By using a model, we advanced our understanding of the thyroid gland morphogenesis. Such a study would have been much more difficult to achieve by following

only a classical *in vivo/in vitro* experimental approach. "All models are wrong, but some can be useful". With that in mind, we firmly believe that *in silico* approaches will keep on developing and that our results paved the way for more advanced studies in the field of mathematical biology.

13 | Appendix

Table 13.1: Model parameters.

Parameter	Description	Value	Dimension
K	Deformation energy parameter	55.0	[Force.Length^{-1}]
λ_{e-e}	Epithelial cell-cell adhesion parameter	5.0	[Force]
λ_{ed-ed}	Endothelial cell-cell adhesion parameter	1.0	[Force]
λ_{l-l}	Lumen cell-cell adhesion parameter	5.0	[Force]
λ_{e-ed}	Epi-Endo cell adhesion parameter	5.0	[Force]
λ_{e-l}	Epi-Lumen cell adhesion parameter	6.0	[Force]
λ_{ed-l}	Endo-lumen cell-cell adhesion parameter	35.0	[Force]
λ_{e-bnd}	Epi-boundary adhesion parameter	10.0	[Force]
$\lambda_{end-bnd}$	Endo-boundary adhesion parameter	1.0	[Force]
λ_{l-bnd}	Lumen-boundary adhesion parameter	8.0	[Force]
Γ	Contractility parameter	1.0	[Force.Length^{-1}]
D	Morphogen diffusion coefficient	0.1	[Length.Time^{-1}]
k_d	Morphogen degradation coefficient	9.0	[Time^{-1}]
$\text{elong}_{\text{max}}$	Endothelial elongation shape factor threshold	2.3	[/]
m	Endothelial active force magnitude	10.0	[Force]
δ_{pol}	Polarisation decrease rate	0.075	[Time^{-1}]
ρ_{endo}	Polarisation impact rate of endothelial cells	0.12	[Time^{-1}]
ρ_{lumen}	Polarisation impact rate of lumen cells	-0.015	[Time^{-1}]
ρ_{bnd}	Polarisation impact rate of the boundary	0.12	[Time^{-1}]
ϵ_{epi}	Polarisation impact rate of epithelial cells	-0.04	[Time^{-1}]
ν	Epithelial cells impact factor on lumen target area	0.002	[Time^{-1}]
υ	Lumen impact factor on lumen target area	2.10^{-4}	[Time^{-1}]

Table 13.2: Antibodies and dilutions.

Primary antibodies			Secondary antibodies
Target	**Species**	**Dilution**	
Endomucin	Rat (SantaCruz)	1/1000	Alexa-568 Goat Anti-Rat IgG (Invitrogen)
E-cad	Mouse IgG2a (BD Bioscience)	1/500	Alexa-647 Goat Anti-Mouse IgG2a (Invitrogen)
Ezrin	Mouse IgG1 (Thermo Scientific)	1/400	Alexa-488 Goat Anti-Mouse IgG1 (Invitrogen)
Laminin	Rabbit (Sigma)	1/100	Alexa-568 Goat Anti-Rabbit IgG (Invitrogen)

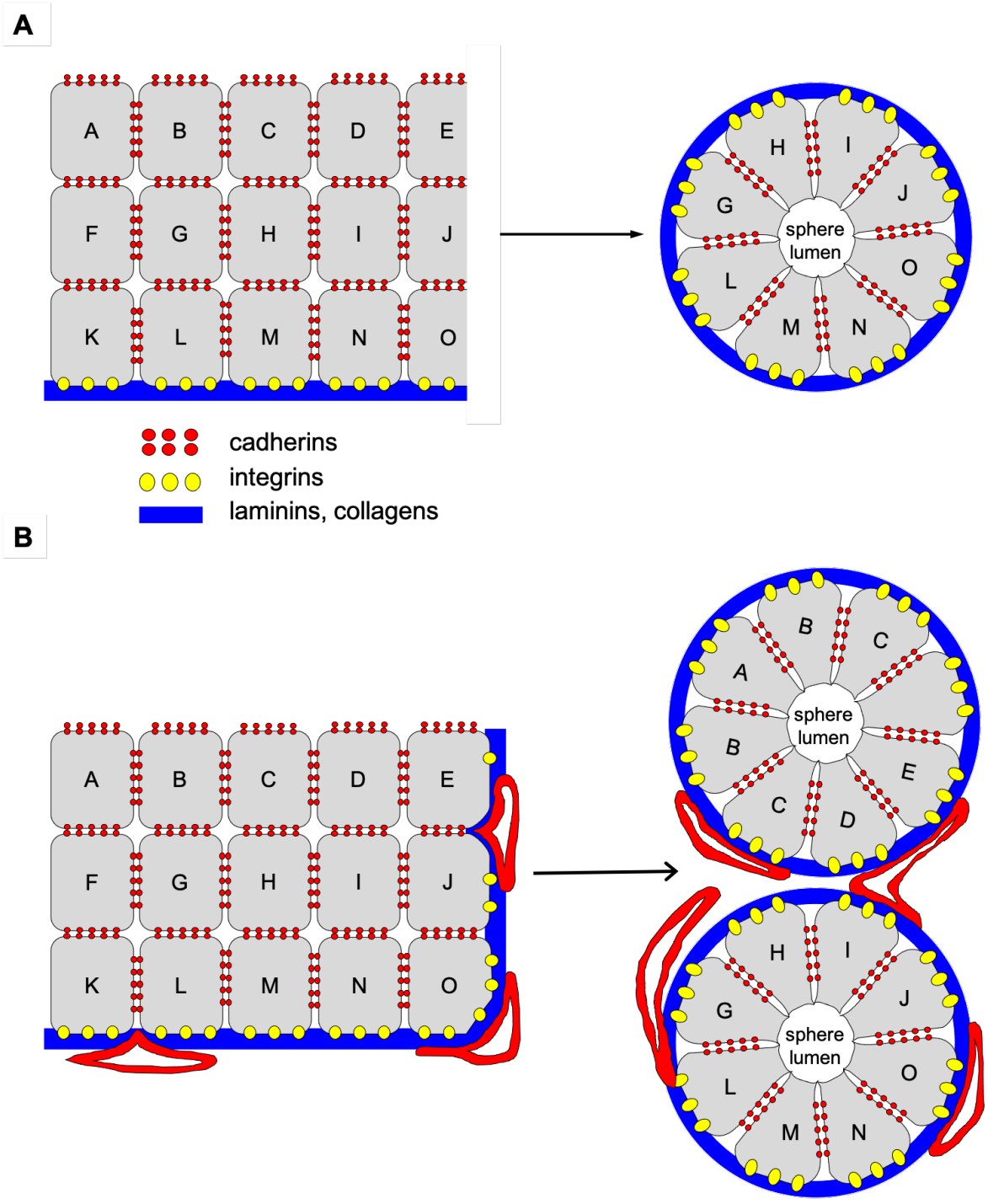


Figure 13.1: (A) Illustration of our initial hypothesis regarding the changes in cell-cell and cell-ECM contacts potentially favouring the formation of thyroid follicles. During the reorganisation of the thyroid epithelial cell mass into a multitude of single-layered follicles, each epithelial cell is expected to reduce its number of neighbours, decreasing its number of cell-cell cohesion proteins (cadherins, in red). Each epithelial cell is also expected to eventually be in contact with the ECM (in blue), illustrated by a relative increase in the number of cell-ECM adhesion proteins (integrins, in yellow). (B) Illustration of the role of endothelial cells (in red) in the fission of the initial epithelial mass.

14 | Publications

Modelling of epithelial growth, fission and lumen formation during embryonic thyroid development : a combination of computational and experimental approache

Léolo Gonay, Catherine Spourquet, Matthieu Baudoin, Ludovic Lepers, Pascale Lemoine, Alexander Fletcher, Emmanuel Hanert and Christophe E Pierreux

A paper gathering the major results presented in this work has been submitted, reviewed and accepted in "Frontiers in Endocrinology". Publication expected by May 2021.

frontiers in Endocrinology

ORIGINAL RESEARCH
published: 07 June 2021
doi: 10.3389/fendo.2021.655862

Modelling of Epithelial Growth, Fission and Lumen Formation During Embryonic Thyroid Development: A Combination of Computational and Experimental Approaches

Leolo Gonay[1,2], *Catherine Spourquet*[2], *Matthieu Baudoin*[1,2], *Ludovic Lepers*[1,2], *Pascale Lemoine*[2], *Alexander G. Fletcher*[3], *Emmanuel Hanert*[1†] *and Christophe E. Pierreux*[2*†]

[1] Earth and Life Institute, UCLouvain, Louvain-La-Neuve, Belgium, [2] de Duve Institute, UCLouvain, Woluwé-Saint-Lambert, Belgium, [3] School of Mathematics and Statistics, University of Sheffield, Sheffield, United Kingdom

OPEN ACCESS

Edited by:
Mikael Nilsson, University of Gothenburg, Sweden

Reviewed by:
Robert Opitz, Charity University Medicine Berlin, Germany
Lisanne Rens, Delft University of Technology, Netherlands

***Correspondence:**
Christophe E. Pierreux
christophe.pierreux@uclouvain.be

[†]These authors share senior authorship

Specialty section:
This article was submitted to Thyroid Endocrinology, a section of the journal Frontiers in Endocrinology

Received: 19 January 2021
Accepted: 22 April 2021
Published: 07 June 2021

Citation:
Gonay L, Spourquet C, Baudoin M, Lepers L, Lemoine P, Fletcher AG, Hanert E and Pierreux CE (2021) Modelling of Epithelial Growth, Fission and Lumen Formation During Embryonic Thyroid Development: A Combination of Computational and Experimental Approaches. Front. Endocrinol. 12:655862. doi: 10.3389/fendo.2021.655862

Organogenesis is the phase of embryonic development leading to the formation of fully functional organs. In the case of the thyroid, organogenesis starts from the endoderm and generates a multitude of closely packed independent spherical follicular units surrounded by a dense network of capillaries. Follicular organisation is unique and essential for thyroid function, *i.e.* thyroid hormone production. Previous *in vivo* studies showed that, besides their nutritive function, endothelial cells play a central role during thyroid gland morphogenesis. However, the precise mechanisms and biological parameters controlling the transformation of the multi-layered thyroid epithelial primordium into a multitude of single-layered follicles are mostly unknown. Animal studies used to improve understanding of organogenesis are costly and time-consuming, with recognised limitations. Here, we developed and used a 2-D vertex model of thyroid growth, angiogenesis and folliculogenesis, within the open-source Chaste framework. Our *in silico* model, based on *in vivo* images, correctly simulates the differential growth and proliferation of central and peripheral epithelial cells, as well as the morphogen-driven migration of endothelial cells, consistently with our experimental data. Our simulations further showed that reduced epithelial cell adhesion was critical to allow endothelial invasion and fission of the multi-layered epithelial mass. Finally, our model also allowed epithelial cell polarisation and follicular lumen formation by endothelial cell abundance and proximity. Our study illustrates how constant discussion between theoretical and experimental approaches can help us to better understand the roles of cellular movement, adhesion and polarisation during thyroid embryonic development. We anticipate that the use of *in silico* models like the one we describe can push forward the fields of developmental biology and regenerative medicine.

Keywords: modelling, morphogenesis, thyroid, epithelial, vertex model

Dynamics and predicted drug response of a gene network linking dedifferentiation with beta-catenin dysfunction in hepatocellular carcinoma

Claude Gérard, Mickaël Di-Luoffo, Léolo Gonay, Stefano Caruso, Gabrielle Couchy, Axelle Loriot, Darko Castven, Junyan Tao, Katarzyna Konobrocka, Sabine Cordi, Satdarshan P Monga, Emmanuel Hanert, Jens U Marquardt, Jessica Zucman-Rossi, Frédéric P Lemaigre

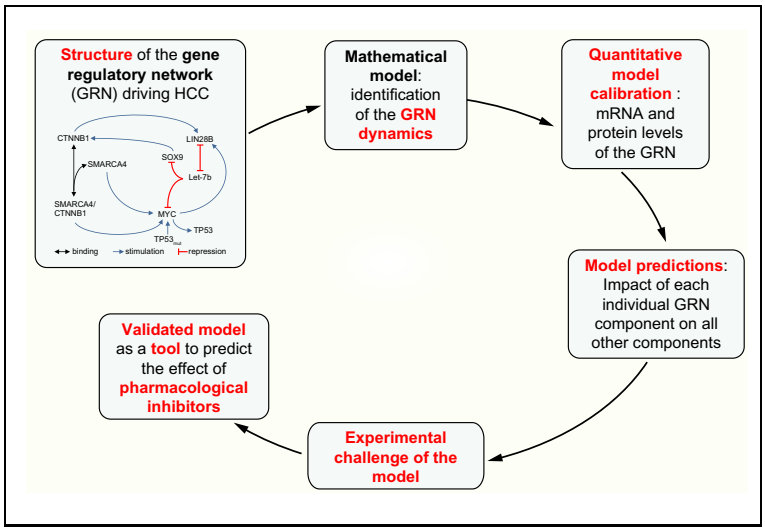
During the first year of my thesis, I collaborated with Dr. Claude Gérard and colleagues, in the lab of Prof. Frédéric Lemaigre, on the development of a computational model of a gene network linking dedifferentiation with beta-catenin dysfunction in hepatocellular carcinoma. I particularly oriented my work in the implementation of user-friendly platforms designed to run and confront the model to external data. The paper resulting from this collaboration is enclosed hereafter.

Research Article
Cancer

JOURNAL
OF HEPATOLOGY

Dynamics and predicted drug response of a gene network linking dedifferentiation with beta-catenin dysfunction in hepatocellular carcinoma

Graphical abstract



Highlights

- We identified a gene regulatory network (GRN) involved in a subset of hepatocellular carcinomas.
- Expression of GRN members and targets correlates with proliferation and prognosis.
- A quantitative mathematical model of the GRN revealed the network dynamics.
- This model is a potential tool to assess the impact of pharmacological inhibitors.

Authors

Claude Gérard, Mickaël Di-Luoffo, Léolo Gonay, ..., Jens U. Marquardt, Jessica Zucman-Rossi, Frédéric P. Lemaigre

Correspondence

frederic.lemaigre@uclouvain.be
(F.P. Lemaigre)

Lay summary

Hepatocellular carcinoma (HCC) is a heterogeneous disease driven by the concomitant deregulation of several genes functionally organized as networks. Here, we identified a gene regulatory network involved in a subset of HCCs. This subset is characterized by increased proliferation and poor prognosis. We developed a mathematical model which uncovers the dynamics of the network and allows us to predict the impact of a therapeutic agent, not only on its specific target but on all the genes belonging to the network.

http://dx.doi.org/10.1016/j.jhep.2019.03.024
© 2019 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved. J. Hepatol. 2019, 71, 323–332

Research Article
Cancer

JOURNAL OF HEPATOLOGY

Dynamics and predicted drug response of a gene network linking dedifferentiation with beta-catenin dysfunction in hepatocellular carcinoma

Claude Gérard[1], Mickaël Di-Luoffo[1], Léolo Gonay[1,2], Stefano Caruso[3,4], Gabrielle Couchy[3,4], Axelle Loriot[1], Darko Castven[5], Junyan Tao[6], Katarzyna Konobrocka[1], Sabine Cordi[1], Satdarshan P. Monga[6], Emmanuel Hanert[2], Jens U. Marquardt[5], Jessica Zucman-Rossi[3,4], Frédéric P. Lemaigre[1,*]

[1]*de Duve Institute, Université catholique de Louvain, Brussels, Belgium;* [2]*Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium;* [3]*Centre de Recherche des Cordeliers, Sorbonne Universités, Inserm, UMRS-1138, F-75006 Paris, France;* [4]*Functional Genomics of Solid Tumors, USPC, Université Paris Descartes, Université Paris Diderot, Université Paris 13, Labex Immuno-Oncology, équipe labellisée Ligue Contre le Cancer, F-75000 Paris, France;* [5]*Department of Medicine I, Johannes Gutenberg University, Mainz, Germany;* [6]*Department of Pathology, Medicine and the Pittsburgh Liver Research Center, University of Pittsburgh, School of Medicine and University of Pittsburgh Medical Center, Pittsburgh, PA, USA*

Background & Aims: Alterations of individual genes variably affect the development of hepatocellular carcinoma (HCC). Thus, we aimed to characterize the function of tumor-promoting genes in the context of gene regulatory networks (GRNs).

Methods: Using data from The Cancer Genome Atlas, from the LIRI-JP (Liver Cancer – RIKEN, JP project), and from our transcriptomic, transfection and mouse transgenic experiments, we identify a GRN which functionally links LIN28B-dependent dedifferentiation with dysfunction of β-catenin (*CTNNB1*). We further generated and validated a quantitative mathematical model of the GRN using human cell lines and *in vivo* expression data.

Results: We found that LIN28B and CTNNB1 form a GRN with SMARCA4, Let-7b (*MIRLET7B*), SOX9, TP53 and MYC. GRN functionality is detected in HCC and gastrointestinal cancers, but not in other cancer types. GRN status negatively correlates with HCC prognosis, and positively correlates with hyperproliferation, dedifferentiation and HGF/MET pathway activation, suggesting that it contributes to a transcriptomic profile typical of the proliferative class of HCC. The mathematical model predicts how the expression of GRN components changes when the expression of another GRN member varies or is inhibited by a pharmacological drug. The dynamics of GRN component expression reveal distinct cell states that can switch reversibly in normal conditions, and irreversibly in HCC. The mathematical model is available via a web-based tool which can evaluate the GRN status of HCC samples and predict the impact of therapeutic agents on the GRN.

Conclusions: We conclude that identification and modelling of the GRN provide insights into the prognosis of HCC and the mechanisms by which tumor-promoting genes impact on HCC development.

Lay summary: Hepatocellular carcinoma (HCC) is a heterogeneous disease driven by the concomitant deregulation of several genes functionally organized as networks. Here, we identified a gene regulatory network involved in a subset of HCCs. This subset is characterized by increased proliferation and poor prognosis. We developed a mathematical model which uncovers the dynamics of the network and allows us to predict the impact of a therapeutic agent, not only on its specific target but on all the genes belonging to the network.

© 2019 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

Introduction

Various etiologies are associated with HCC, leading to heterogeneity in clinical outcome, histology, transcriptomic profile and mutational spectrum.[1–5] Such heterogeneity causes a variable response to therapeutic agents, as in mouse models with *Ctnnb1*-induced HCCs which show heterogeneous sensitivity to CTNNB1 inhibitors.[6] Thus, designing novel therapeutic strategies against HCC requires the identification of inhibitors of individual tumor-promoting genes and also the characterization of the molecular networks in which those genes exert their functions.

Dedifferentiation of hepatic cells contributes to HCC progression.[7–9] In this context, poorly differentiated HCC develop as a result of forced induction of LIN28B, an RNA-binding protein which is repressed during normal hepatic cell differentiation. LIN28B is re-expressed in a subset of human HCCs characterized by high serum levels of α-fetoprotein,[10,11] thereby associating dedifferentiation, HCC progression and LIN28B expression.[12] In parallel, *CTNNB1* is one of the most frequently mutated genes in HCC.[13] Therefore, we explore the possibility that HCC progression depends on a gene regulatory network (GRN)

Keywords: Hepatocellular carcinoma; Mathematical model; Gene regulatory network; HGF/MET pathway; Principal component analysis; MicroRNA; LIN28B; CTNNB1; Personalized medicine.
Received 16 August 2018; received in revised form 15 March 2019; accepted 21 March 2019; available online 4 April 2019
* Corresponding author. Address: de Duve Institute, Université catholique de Louvain, Avenue Hippocrate 75 (B1-75.03), 1200 Brussels, Belgium. Tel.: +32 2 764 7583 (direct)/+32 2 764 7550 (secretariat); fax: +32 2 764 7573.
E-mail address: frederic.lemaigre@uclouvain.be (F.P. Lemaigre).

linking LIN28B-dependent dedifferentiation with CTNNB1 dysfunction.

Several approaches can be used to identify GRNs.[14] Here, we selected a method which captures the biological logic of gene networks. The identification of the network is literature-based. The number of interacting components in the network is limited but perfectly compatible with the characterization of the network's dynamics by our mathematical modelling approach.[15,16] We identified an HCC-promoting GRN comprising several members connecting LIN28B with CTNNB1 via Let-7b (*MIRLET7B*), MYC, SMARCA4 (also called BRG1), TP53 and SOX9. We further investigated the system-level dynamics of the GRN using a quantitative mathematical model, which was calibrated and validated using mRNA and protein expression data from HCC cell lines, primary HCC cells, patient databanks and mouse models.

Materials and methods

Data normalization and statistical analysis

Data normalization is described in the supplementary information. Measured data are means ± SD. Significance was assessed by Student's *t* test (*p* value). For RNA sequencing data of the HCC cohort from The Cancer Genome Atlas (TCGA), we calculated the adjusted *p* value using a Benjamini-Hochberg correction (p^* value) on the entire transcriptome.

RNA sequencing and miRNA sequencing

Data were from TCGA (http://firebrowse.org/), and from LIRI-JP (RIKEN) of the International Cancer Genome Consortium (ICGC; https://dcc.icgc.org/projects/LIRI-JP). For the TCGA cohort, we converted the "scaled_estimate" in the "illuminahiseq_rnaseqv2_unc_edu_Level_3_RSEM_genes" file into TPM by multiplying by 10^6.

Plasmids and microRNAs

pCDNA3.1, pcDNA3-MYC, pCl-neo β-Catenin (CTNNB1^{S33Y}), pBABE-BRG1 were from Thermo Fisher Scientific (Waltham, MA, USA), Wafik El-Deiry (Addgene plasmid # 16011), Bert Vogelstein (Addgene plasmid # 16519) and Robert Kingston (Addgene plasmid # 1959) respectively. The miR mimic control (MIMAT0000039: UCACAACCUCCUAGAAAGAGUAGA) and miR mimic hsa-let-7b: hsa-let-7b-5p (MIMAT0000063: UGAGGUAGUAGGUUGUGUGGUU), hsa-let-7b-3p (MIMAT0004482: CUAUACAACCUACUGCCUUCCC) were from Dharmacon (Lafayette, CO, USA).

Cell culture

Human Huh7, HepG2, Hep3B cell lines and primary human HCC cells HCC31[17] were grown in DMEM (Lonza, Leusden, Netherlands), 10% fetal bovine serum (5% for HCC31; Merck, Darmstadt Germany), L-Glutamine (2 mM) (Thermo Fisher Scientific), Penicillin-Streptomycin (50 U/ml and 50 µg/ml) (Gibco™, Waltham, MA, USA) and Amphotericin B (Gibco™) (2.5 µg/ml). Cells were grown in 60 mm dishes and transfected with 3 µg plasmid and 120 nM miRmimic, using jetPRIME® (Polyplus-Transfection, Illkirch-Graffenstaden, France) for 48, 72 or 96 h, in at least 3 independent experiments. DNA was transfected 24 h after plating the cells. For LIN28 inhibition, Huh7 and HCC31 cells were grown in the presence of 120 µM (*N*-Methyl-*N*-[3-(3-methyl-1,2,4-triazolo[4,3-*b*]pyridazin-6-yl) phenyl]acetamide) for 6 days; the medium with inhibitor was changed every day.

Western blotting

Protocol and antibodies are described in the supplementary information.

RNA extraction and analysis

Total RNA was isolated from Huh7 cells using Trizol (#1029602, Invitrogen, Life technologies). cDNA synthesis was performed with MMLV reverse transcriptase (#28025-13, Invitrogen, Life technologies) according to manufacturer's protocol. MicroRNA expression (Let-7a and Let-7b) was quantified by reverse transcription quantitative PCR (RT-qPCR) using Kapa SYBR Fast 2X Universal Master Mix (#KK4601, Sopachem, Ochten, Netherlands). Specific stem-loop primers were used for reverse transcription, and RT-qPCR was performed using a specific forward primer and a common universal reverse primer. *Let-7a* Fwd/Rev: ACACTCCAGCTGGGTGAGGTAGTAGGTTG/CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACTATACA; *Let-7b* Fwd/Rev: ACACTCCAGCTGGGTGAGGTAGTAGGTTGT/CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACCACAC; *ACTB* Fwd/Rev: TCCTGAGCGCAAGTACTCTGT/CTGATCCACATCTGCTGGAAG. Each ΔCt between the measured transcripts and the housekeeping genes was normalized to their control conditions using the $2^{-\Delta\Delta Ct}$ method.

For further details regarding the materials used, please refer to the CTAT table and supplementary information.

Results

Identification of a gene regulatory network driving hepatocellular carcinoma

We selected an approach in which GRN members meet stringent functional criteria. First, their role in tumor promotion must be validated by animal experimentation and/or high-throughput sequencing data from patients. Second, GRN members must be connected by direct or indirect functional links characterized by protein-protein and protein-DNA interactions, or epistatic relationship identified in loss- and gain- of function analyses. By combining data from the literature, we first reconstituted a GRN comprising 7 cross-regulating components: the miRNA *Let-7b*, the RNA-binding protein LIN28B, the ATP-dependent helicase SMARCA4, and the transcription factors SOX9, MYC, CTNNB1 and TP53 (Fig. 1A; supplementary information, section 1). When interactions have been identified in non-hepatic cells, we verified whether they also occurred in cultured HCC cell lines (see below).

Data from TCGA (Fig. 1A) showed that expression of *Let-7b* is reduced in HCC compared to adjacent non-tumor tissue. This is consistent with the concomitant overexpression of LIN28B, a repressor of *Let-7b*. All other GRN components, except MYC, were increased in HCC. Further, a principal component analysis (PCA), which considers only the expression of the 7 GRN components in 50 non-tumor controls and 368 HCCs revealed that non-tumor tissue and HCC tumors clustered separately (Fig. 1B and Fig. S1A). In addition, we clustered the HCC cohort in 2 groups of 100 samples based on the low versus high expression levels of each GRN component. This analysis showed that, except for *Let-7b*, high expression of a component is associated with high expression of the other GRN components. (Fig. S2; supplementary information, Section 2).

PCA investigating multiple targets of the GRN components (Table S1) identified separate clusters for non-tumor tissue and HCC tumors, indicating that the components of the GRN

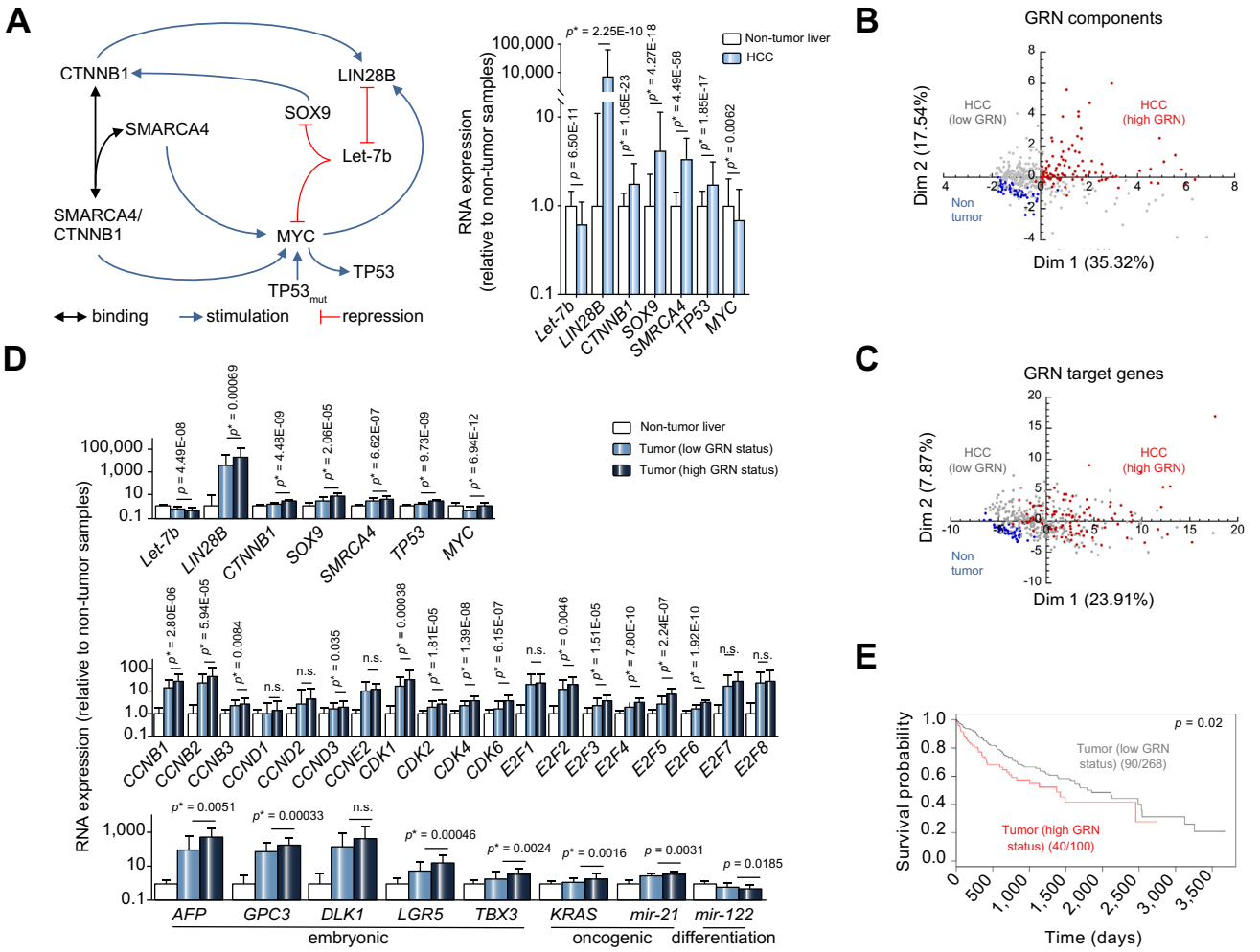


Fig. 1. Identification of a GRN involved in HCC. (A) Structure of the GRN (left), and RNA levels of GRN components in non-tumor tissue (n = 50) and HCC tumors (n = 368) from the HCC cohort in TCGA (right). (B) PCA plots based on the expression of the GRN components, and (C) on selected GRN targets (Table S1) in control and HCC samples. (D) Tumors were ranked according to low or high GRN status. Expression of the GRN components correlated well with high GRN status (top); and with expression of proliferation (middle), embryonic and oncogenic markers (bottom), and with lowest levels of the differentiation marker miR-122 (bottom). (E) Kaplan-Meier curves (overall survival) showed that patients with high GRN status (median time = 1,372 days, 40 deceased patients amongst 100) exhibited a lower survival probability than patients with low GRN status (median time = 1,852 days, 90 dead patients amongst 268) (Wilcoxon test, $p = 0.02$). $p = p$ value (Student's t test); $p^* = p$ value (Student's t test with Benjamini-Hochberg correction on the entire transcriptome). GRN, gene regulatory network; HCC, hepatocellular carcinoma; PCA, principal component analysis; TCGA, The Cancer Genome Atlas. (This figure appears in colour on the web.)

were not only dysregulated in HCC but also that they were actively controlling their targets (Fig. 1C and Fig. S1B). Therefore, we considered that concomitant overexpression of the GRN components LIN28B, SMARCA4, SOX9, CTNNB1 and TP53, and downregulation of Let-7b is indicative of GRN activity.

Out of the 368 HCCs from the TCGA cohort whose GRN component expression had been analyzed by PCA we selected the 150 samples with the highest Dimension 1 to which *CTNNB1*, *SMARCA4*, *SOX9, MYC* and *TP53* expression contribute the most, and out of the latter we selected the 100 samples with highest Dimension 2 to which *LIN28B* and *Let-7b* contribute the most (Fig. S1A and red dots in Fig. 1B). These 100 samples were defined hereafter as displaying high GRN status and were compared with the other 268 tumors, which were defined as low GRN status, and with non-tumor liver tissue. The expression of the GRN components differed significantly between the tumors with low and high GRN status (Fig. 1D). Moreover, the samples with high GRN status displayed the highest expression of proliferation, embryonic and oncogenic markers, and the lowest levels of differentiation markers (Fig. 1D).

HCCs are divided in proliferative and non-proliferative classes.[13] GRN status, defined as above, correlated with increased MET signaling, a pathway whose activation characterizes a subset of the proliferative HCC class.[18] Indeed, HCCs with low or high GRN status clustered separately from the non-tumor samples in a PCA plot when considering the expression of 110 HGF/MET target genes (Fig. 2A). In addition, the 100 tumors with highest HGF/MET target gene expression displayed increased expression of *CTNNB1*, *SOX9*, *SMARCA4*, *MYC* and *TP53* compared to the other 268 tumors (Fig. 2B).

Earlier transcriptomic analyses classified patients with HCC into 6 subgroups, G1 to G6.[19] PCA analysis using 216 randomly selected genes that are differentially expressed in G1 to G6, revealed that tumors from the TCGA cohort with high GRN status cluster the farthest from the non-tumor tissue (Fig. 2C). Among the 216 genes differentially expressed between the G1

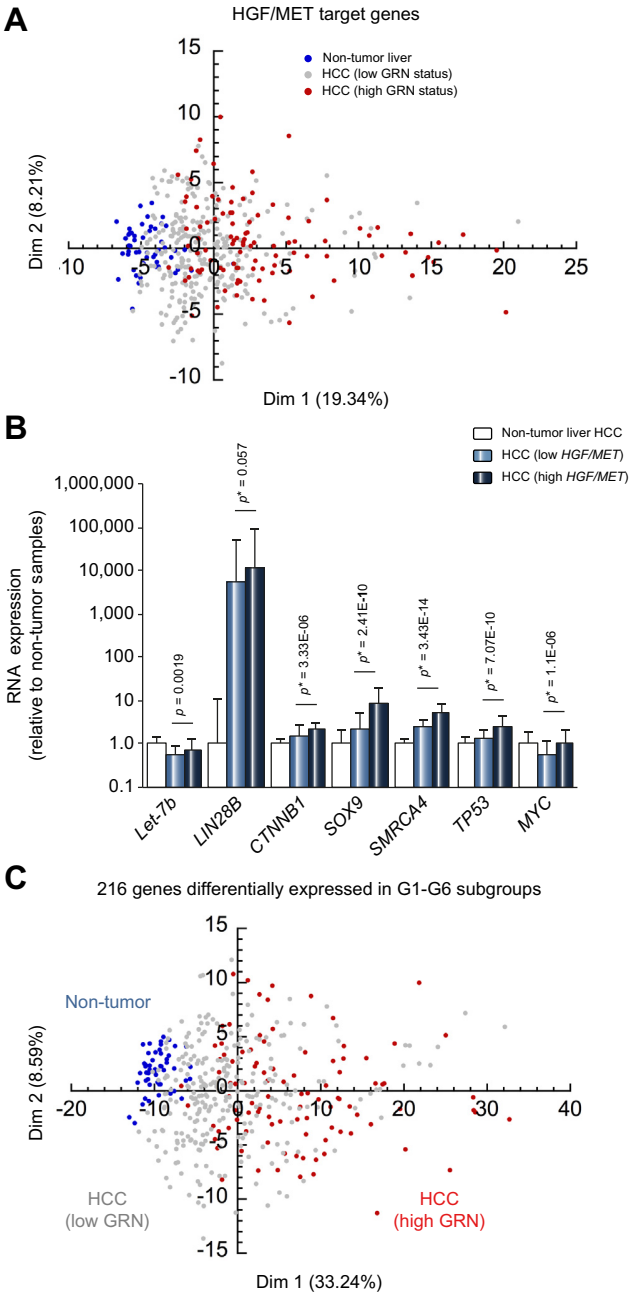


Fig. 2. The gene regulatory network promotes a proliferative HCC phenotype. (A) Non-tumor samples (blue; n = 50) and HCCs (n = 368) cluster separately in a PCA analysis based on the expression of 110 HGF/MET target genes (Table 1 in Ref. 18). (B) Expression of GRN components in non-tumor samples (blue, n = 50) and in HCC with low (grey; n = 268) or high HGF/MET target expression (red; n = 100). 100 tumor samples with high HGF/MET signaling were selected by PCA analysis based on the expression levels of 110 HGF/MET target genes (panel A). (C) Clustering of normal (blue) *vs.* HCC with low (grey) and high GRN status (red) in a PCA analysis based on the expression of 216 differentially expressed genes of the 6-group HCC classification. (B) Data are means ± SD. GRN, gene regulatory network; HCC, hepatocellular carcinoma; PCA, principal component analysis. p and p^* calculated as in Fig. 1. (This figure appears in colour on the web.)

to G6 subgroups, the 50 genes that contribute the most to the Dimension 1 in the PCA plot were all upregulated in G1, G2 and G3 and were mainly involved in cell proliferation (Table S2), supporting the idea that GRN functionality may contribute to confer a transcriptomic profile typical of the HCC proliferation class.

We also investigated whether GRN status correlated with prognosis. The survival probability was lower in the 100 patients with HCC and a high GRN status than in the 268 patients with HCC and a low GRN status (Fig. 1E). When performing the same analysis with individual components of the GRN, the expression of *Let-7b*, *MYC* and *TP53* did not correlate with survival, while expression of *CTNNB1*, *SOX9*, *SMARCA4* and *LIN28B* were inversely correlated with survival (Fig. S3). Therefore, the status of the GRN is a good marker of prognosis.

To determine if the characteristics of HCC with high GRN status are valid for a distinct patient cohort, we analyzed the GRN in the LIRI-JP (Riken) cohort from the ICGC. As in the TCGA cohort, the HCC with low or high GRN status clustered separately (Fig. 3A). The GRN is also functional in LIRI-JP since HCC with high GRN status overexpress the GRN target genes (Fig. 3A). HCC with high GRN status also overexpress proliferation markers and the HGF/MET target genes (Fig. 3B-C), and negatively correlate with prognosis (Fig. 3D).

In conclusion, we have identified a GRN of functionally interacting partners which are misexpressed in HCC. The status of the GRN, characterized by consistent and concomitant misexpression of its components and targets, correlates with proliferation, dedifferentiation and prognosis.

Mathematical model of the GRN

To develop a tool that determines how variation of an individual GRN component affects the expression of all the others, we built a quantitative mathematical model, using a set of 20 kinetic equations describing the expression of each network component (mRNA and protein) as a function of time (supplementary information, sections 3–4). We quantitatively calibrated the model using the mRNA levels of all GRN components available in TCGA and in a dataset from 34 human HCC cell lines[20] (Fig. S4A-C). Two sets of parameter values and initial conditions were determined for normal tissues on one hand, and for HCC tumors and cell lines on the other hand (Tables S3-7).

To quantitatively calibrate the model with protein expression values we transiently overexpressed GRN components in cultured HCC cell lines. These experiments validated in HCC the cross-regulating pairs of the GRN which had originally been identified in non-hepatic tumors. Indeed, overexpression of MYC in HepG2 or Huh7 cells stimulated TP53 and *LIN28B* expression, thereby validating in HCC the MYC → *TP53* and MYC → *LIN28B* interactions (Fig. 4A and Fig. S5A). Similarly, overexpression of constitutively active CTNNB1^{S33Y} and of SMARCA4 in HepG2 or Huh7 cells validated the CTNNB1 → *LIN28B* and CTNNB1 → *MYC* and SMARCA4 → *MYC* interactions in HCC (Fig. 4B and Fig. S5B-C). Finally, transient transfection of *Let-7b-5p* mimic RNA repressed LIN28B, MYC and SOX9, confirming the regulatory links in the GRN (Fig. 4C).

The duration of the transient transfections in Fig. 4 (48 h, 72 h, 96 h) reflected the minimal time-lapse required to monitor a significant change in TP53, LIN28B, MYC or SOX9 protein levels following overexpression of their respective stimulator or repressor. Therefore, both the level of protein induction and the timing required to observe a significant change in protein level were used to further calibrate the mathematical model, while maintaining the quantitative calibration of all mRNA expression levels. We then simulated the transient transfection conditions by increasing the transcription rate constants of

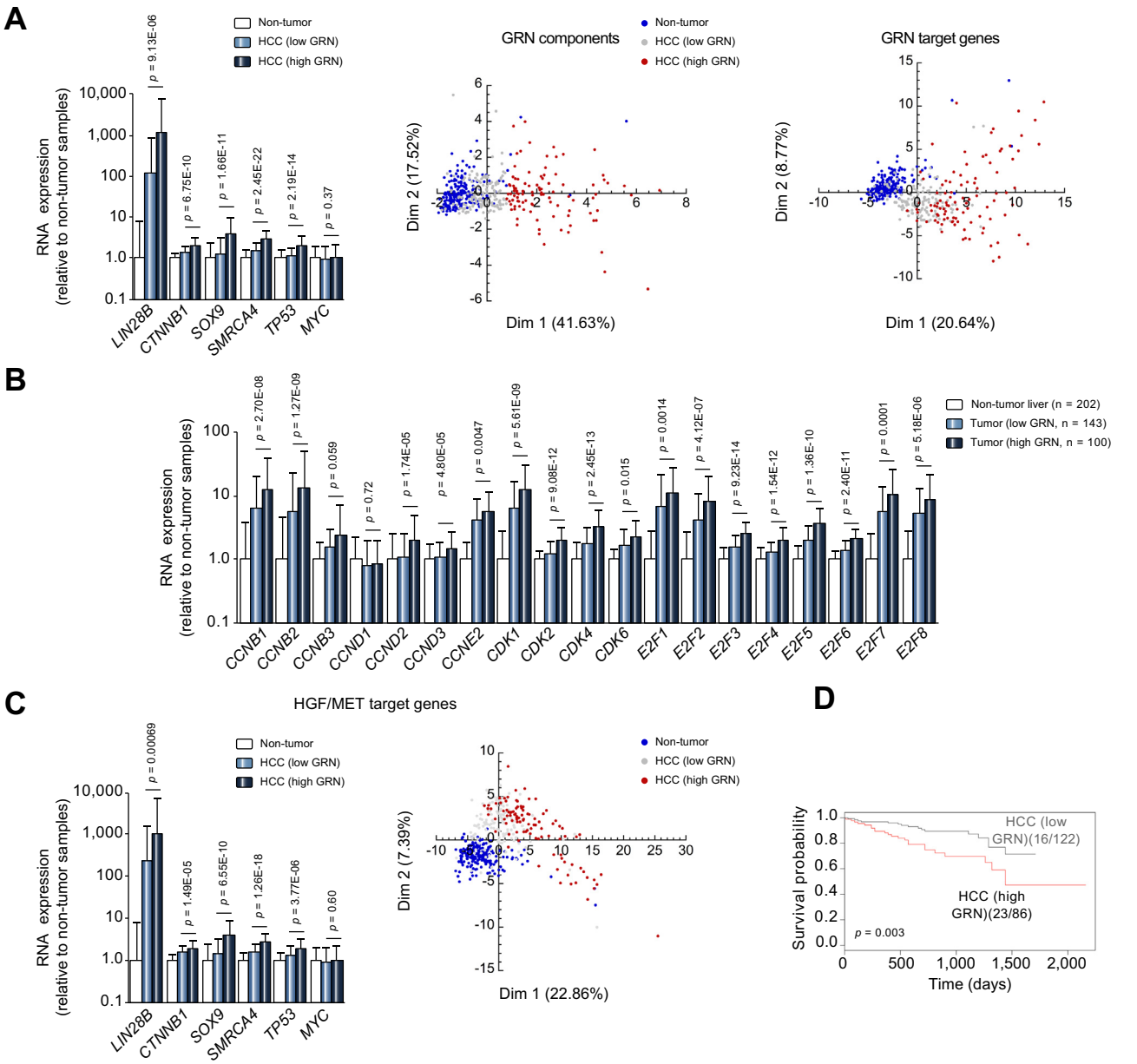


Fig. 3. Validation of the GRN in the LIRI-JP HCC cohort. (A, left) mRNA levels of GRN components in non-tumor tissue (n = 202) and HCC tumors (n = 243) from the HCC cohort (LIRI-JP) in ICGC. (A, middle) PCA plots based on the expression of the GRN components, and (A, right) on GRN targets (listed in Table S1) in control (blue), HCC low (grey) or high GRN status (red). (B) Expression of the GRN components correlated well with expression of proliferation markers. (C, left) Expression of GRN components in non-tumor samples (blue, n = 202) and in HCC with low (grey; n = 142) or high HGF/MET target expression (red; n = 100). (C, right) Non-tumor samples (blue; n = 202), HCC low GRN (grey; n = 142) and HCC high GRN status (red; n = 100) cluster separately in a PCA analysis based on the expression of 110 HGF/MET target genes. Data are means ± SD; p and p^* calculated as in Fig. 1. (D) Kaplan-Meier curves (overall survival) showed that patients with high GRN status (23 deceased patients amongst 86) exhibited a lower survival probability than patients with low GRN status (16 dead patients amongst 122) (Wilcoxon test, $p = 0.003$). GRN, gene regulatory network; HCC, hepatocellular carcinoma; ICGC, International Cancer Genome Consortium; PCA, principal component analysis; TCGA, The Cancer Genome Atlas. (This figure appears in colour on the web.)

MYC, CTNNB1 or *Let-7b*, starting at t = 0 h (Fig. 4A-C, right). The results showed that simulating a 10-fold increase in MYC, which mimics the observed 10-fold increase in transfected MYC protein, predicted a ~ 2-fold increase in LIN28B and TP53 after 72 h; this predicted increase in LIN28B and TP53 matched closely the measured values (compare white and blue bars in Fig. 4A). Similarly, simulating an 8-fold increase in CTNNB1^{S33Y} predicted a ~ 3- and 11-fold induction of LIN28B and MYC after 48 h; these inductions again matched closely the measured 4- and 9-fold induction of LIN28B and MYC, following an 8-fold increase in transfected CTNNB1^{S33Y} (Fig. 4B). Moreover, following a simulated 8-fold induction of *Let-7b* the model reproduced quantitatively the experimentally measured impact of *Let-7b* induction on LIN28B, MYC, and SOX9 (Fig. 4C). Finally, Let-7b induction in Huh7 cells did not affect *CTNNB1* expression, which fitted with the simulation (Fig. 4C). We concluded that the mathematical model faithfully recapitulates the expression of GRN components.

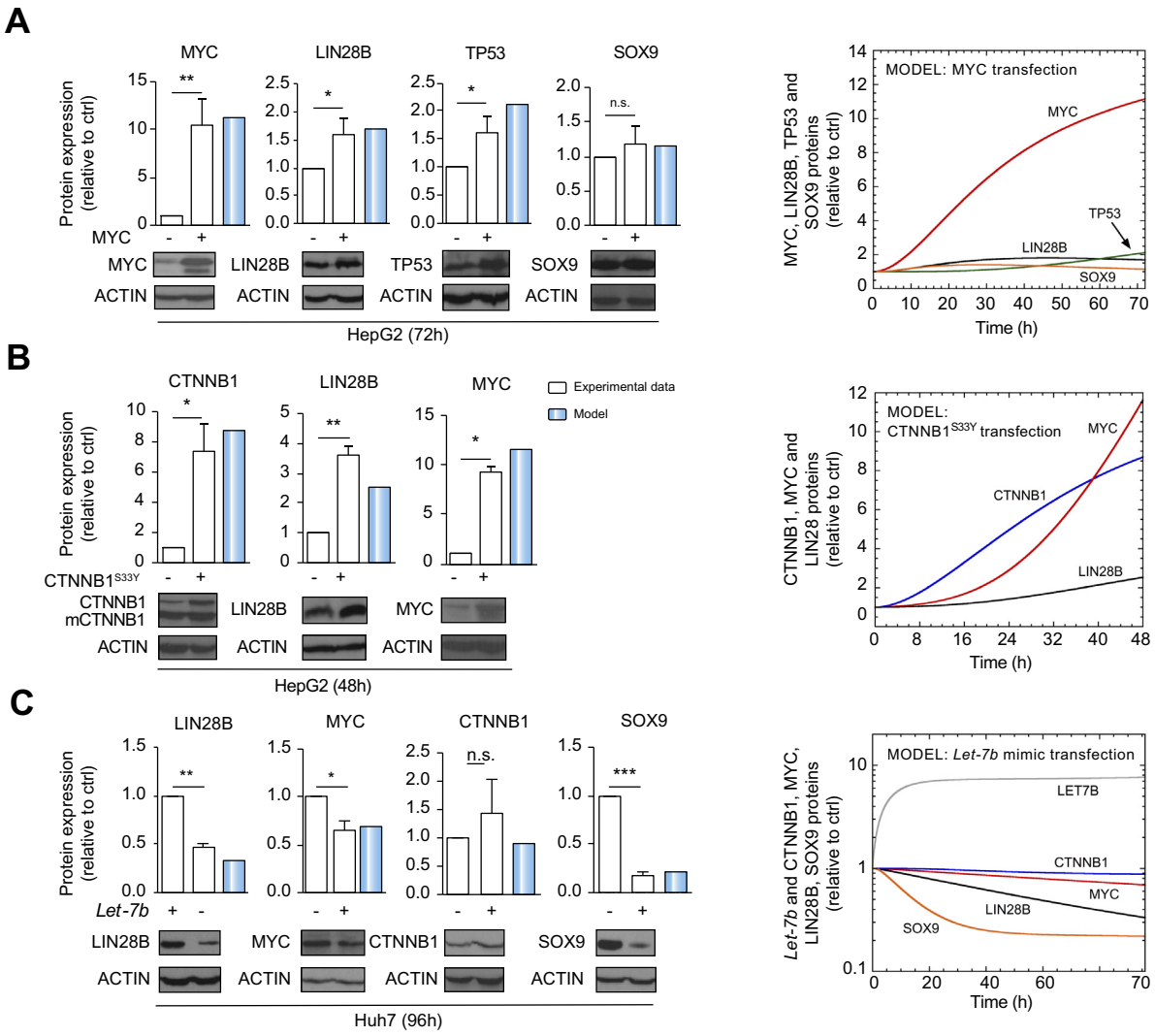


Fig. 4. Calibration of the mathematical model on protein expression in human HCC cell lines. (A) Expression of MYC, LIN28B, TP53 and SOX9 protein following MYC overexpression in HepG2 cells. (B) Expression of CTNNB1, LIN28B and MYC protein following CTNNB1^{S33Y} overexpression in HepG2 cells. mCTNNB1 corresponds to endogenous mutant CTNNB1 resulting from inframe deletion of exon 3. (C) Expression of LIN28B, MYC, CTNNB1 and SOX9 in the presence of *Let-7b-5p* mimic RNA in Huh7 cells. Data in bar graphs are means ± SD, n ≥ 3. $^{*}p < 0.05$; $^{**}p < 0.01$ and $^{***}p < 0.001$ (Student's *t* test). Blue bars correspond to protein levels calculated in the mathematical model with (A) overexpression of MYC, (B) overexpression of CTNNB1^{S33Y} and (C) overexpression of *Let-7b*. The right panels in A-C show the simulated temporal evolution, with the appropriate time scale, of the GRN proteins following MYC, CTNNB1^{S33Y}, or *Let-7b* mimic transfection. See supplementary information for parameter values and initial conditions. GRN, gene regulatory network; HCC, hepatocellular carcinoma. (This figure appears in colour on the web.)

Validation of the mathematical model in mouse and human

To validate the model, we challenged it by predicting the impact of a stabilizing CTNNB1 mutation on *LIN28B* and *Let-7b* expression and compared the prediction with values from TCGA. To this end we first partitioned the HCC samples of the TCGA database into 2 cohorts characterized by the presence or absence of a stabilizing CTNNB1 mutation, the most frequent *CTNNB1* mutations being missense mutations in exon 3. This revealed that *CTNNB1* mRNA levels were higher in the presence of a mutation (Fig. 5A). Therefore, to simulate CTNNB1 stabilization in the model we increased the *CTNNB1* mRNA synthesis rate, increased the activation rate of the inactive (complexed) CTNNB1 form, and decreased the inactivation rate of the active (stabilized) form. Under these conditions, when *CTNNB1* mRNA values from the mutated and non-mutated samples were introduced into the model (Fig. 5A, blue bars), it faithfully predicted the reduction in *Let-7b* observed in non-mutated HCC samples as well as the slightly larger decrease of *Let-7b* seen in HCC samples with *CTNNB1* mutation. The model also predicted the increased *LIN28B* expression in HCC which was similar whether *CTNNB1* was mutated or not (compare blue and white bars in Fig. 5A).

A second validation was obtained by considering previous results from a mouse model of MYC-induced liver cancer.[12] We simulated a 4,000-fold increase in *Myc* mRNA, which corresponds to the experimental induction of *Myc* in mice. Our simulation predicted a MYC-induced 7,000-fold overexpression of *Lin28b* mRNA and a 0.4-fold reduction in *Let-7b*. This prediction fitted well with the experimentally measured induction of *Lin28b* and reduction of *Let-7b* (Fig. 5B).

Third, as a proof of concept that the mathematical model can predict the impact of a pharmacological inhibitor of a GRN

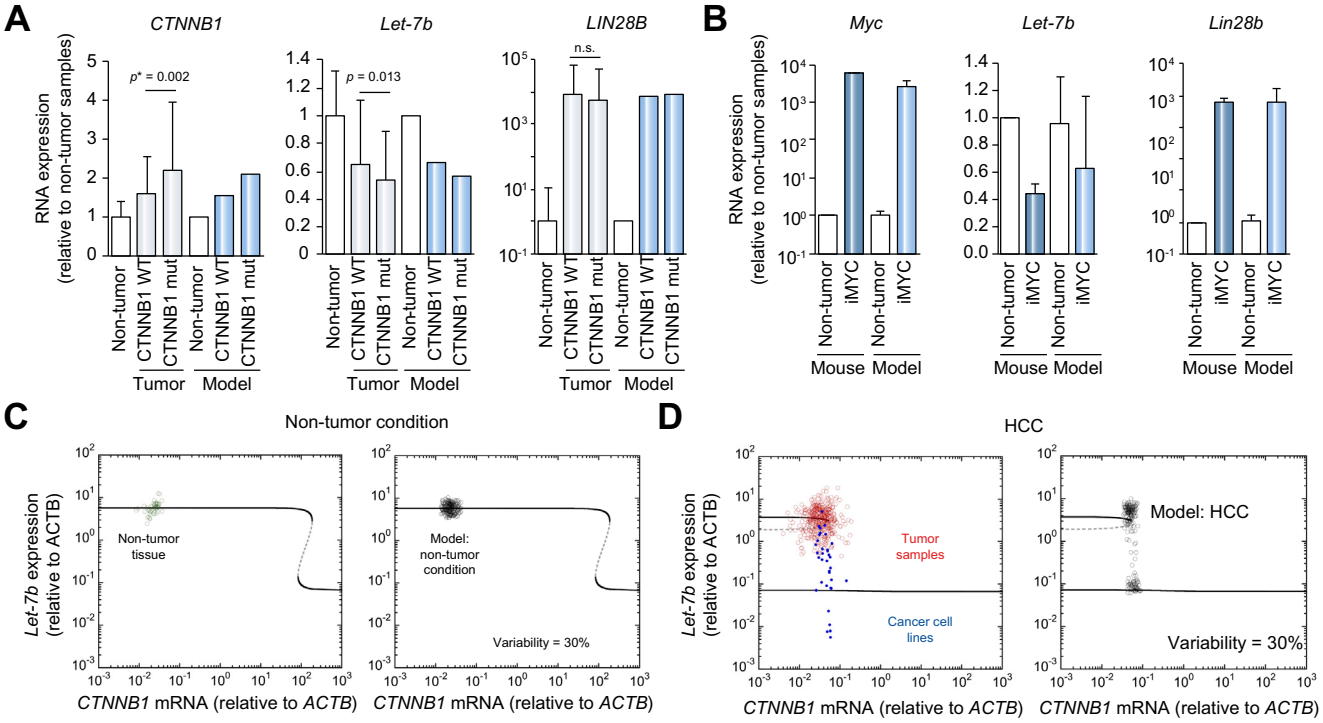


Fig. 5. Validation of the model and GRN dynamics in HCC. (A) Expression of stabilized CTNNB1 mutant (n = 98) and wild-type CTNNB1 (n = 270) in HCC compared to normal tissue (n = 50). The predicted levels of *Let-7b* and *LIN28* following a simulation of increased expression and stabilization of CTNNB1 fit with the experimental observations. (B) Measured and simulated impact of MYC induction on *LIN28B* mRNA and *Let-7b* expression levels, in a mouse model of MYC-induced HCC. (A, B) Results are means ± SD; p and p^* values calculated as in Fig. 1. (C) Modelling *Let-7b* levels as a function of *CTNNB1* mRNA in normal conditions (left). Values in non-tumor tissue (n = 50; green circles) are superimposed on the model curve. Solid curves, stable steady states; dashed curves, unstable steady states. Modelling 30% random variations around the basal value of each parameter in 200 cells resulted in a distribution of *Let-7b*/*CTNNB1* values in the upper stable state (right) which recapitulated the heterogeneity measured in normal samples. (D) Modelling *Let-7b*/*CTNNB1* mRNA values in HCC conditions (left). *Let-7b*/*CTNNB1* mRNA values in 368 HCC tumors (red circles) or 34 cell lines (blue dots) are superimposed on the model curve. Random variation of parameters like in (C) recapitulated the heterogeneity measured in patients with HCC and cell lines (right). See supplementary information for parameter values and numerical simulations. GRN, gene regulatory network; HCC, hepatocellular carcinoma. (This figure appears in colour on the web.)

component, we simulated the inhibition of LIN28B in HCC: simulating 60% inhibition of LIN28B protein predicted an increase in *Let-7b* and a decrease in SOX9 and MYC protein; CTNNB1 remained unaffected (Fig. 6). We then evaluated the results of the simulation in a HCC cell line, Huh7 (Fig. 6A), and in HCC31 primary liver cancer cells[17] (Fig. 6B), by growing these cells in the presence of LIN28 inhibitor (*N*-Methyl-*N*-[3-(3-methyl-1,2,4-triazolo[4,3-*b*]pyridazin-6-yl)phenyl]acetamide).[21] The inhibitor reduced cell proliferation (not shown), and, as expected, induced the expression of *Let-7b* and *Let-7a*. Importantly, the changes of SOX9, MYC and CTNNB1 protein levels measured by western blot matched well with the simulations (Fig. 6), thereby validating the mathematical model as a tool to predict the impact of a pharmacological inhibitor on the GRN.

GRN dynamics are characterized by a bistable switch

To characterize the dynamic properties of the GRN and determine whether they identify distinct HCC cell states, we modelled the steady-state levels of *Let-7b*, selected as a representative variable, as a function of *CTNNB1* mRNA. In normal conditions, high *Let-7b* was associated with low *CTNNB1* mRNA, and *vice versa* (Fig. 5C, left). The system exhibits a reversible bistable switch from high to low levels of Let-7b, at supraphysiological levels of *CTNNB1* mRNA. Green circles in Fig. 5C represent measured *Let-7b*/*CTNNB1* mRNA values in control samples from TGCA. However, simulating *Let-7b* as a function of *CTNNB1* mRNA in HCC conditions revealed an irreversible bistable switch occurring at low *CTNNB1* mRNA levels (Fig. 5D). Indeed, rising *CTNNB1* mRNA from low to supraphysiological levels in HCC would induce a switch from high to low *Let-7* expression, but reverting from supraphysiological *CTNNB1* to low *CTNNB1* levels would not allow high *Let-7* levels to be restored. The red circles in Fig. 5D reflect tumor-to-tumor heterogeneity of the HCCs and are predominantly positioned around the upper branch of the bistable switch. Blue dots in Fig. 5D represent *Let-7b*/*CTNNB1* values in 34 human HCC cell lines and reflect the cell line-to-cell line heterogeneity. These dots are predominantly positioned around the lower branch of the bistable switch, suggesting that culture conditions of cell lines promote switching of the GRN to a state distinct from the mean cell state in patient tumors. In non-tumor conditions (Fig. 5C), the expression of the GRN components is well below that of the threshold at which the bistable switch occurs. Therefore, a bistable switch is unlikely to contribute to the initiation of the disease. In contrast, in HCC, the expression level of GRN components in tumors and cell lines is around the threshold of the bistable switch, suggesting that the bistable switch can contribute to HCC progression (Fig. 5D).

The distribution of measured *Let-7b*/*CTNNB1* values in Fig. 5-C-D reflects inter-sample heterogeneity. When modelling a normal heterogeneous cell population with 30% of random variations around the basal value of each parameter, the

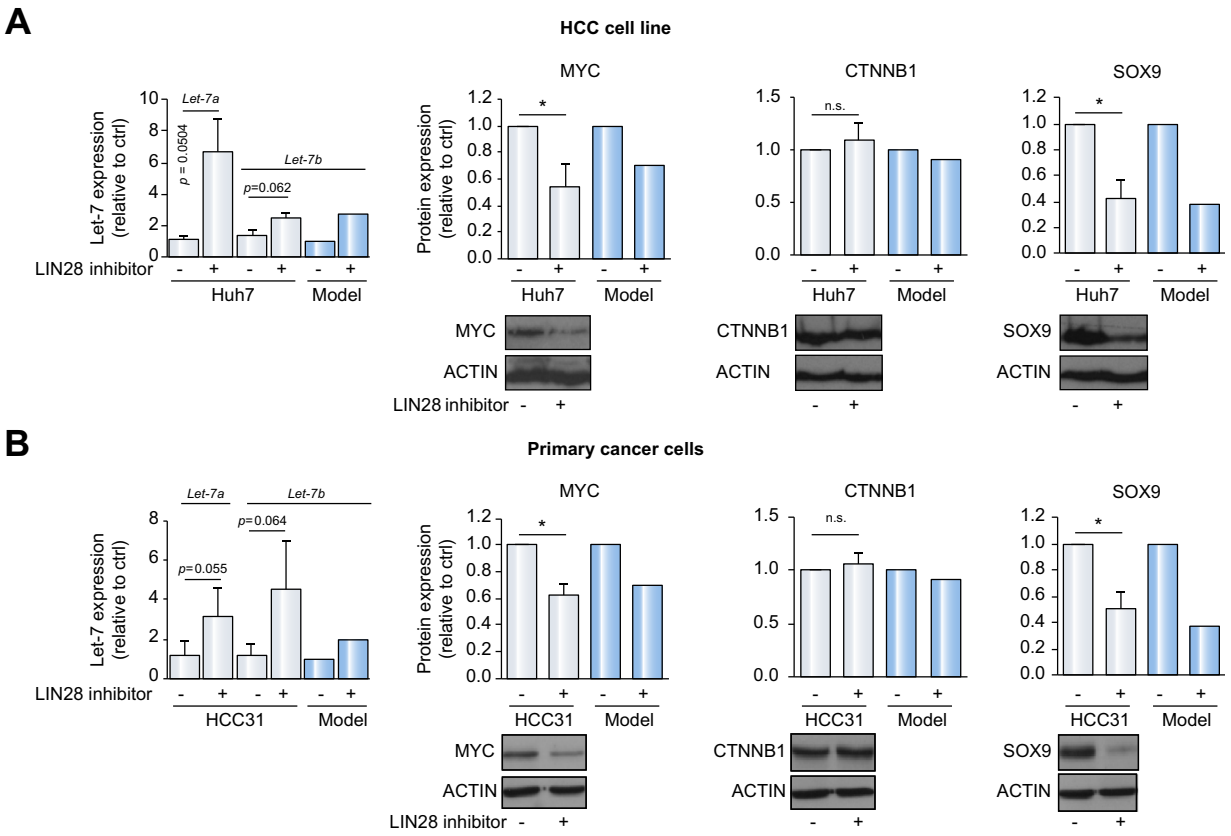


Fig. 6. Mathematical model as a tool to predict the impact of LIN28 inhibition on the GRN dynamics. (A, B) Simulating the absence (K_{ILET7} = 21) or presence (K_{ILET7} = 82) of LIN28 inhibition for 6 days (144 h) predicts expression levels of *Let-7b*, MYC, CTNNB1 and SOX9 which are close to the levels measured following (A) treatment of Huh7 cells and (B) HCC31 primary liver cancer cells with 120 μM of LIN28 inhibitor.[21] (B) *Let-7* expression levels were simulated and measured after 96 h. (A, B) Results are means ± SD; n ≥ 3; *p < 0.05 (Student's t test). GRN, gene regulatory network; HCC, hepatocellular carcinoma.

Let-7b|*CTNNB1* values remained clustered in the upper stable steady state (Fig. 5C, right). Modelling an HCC cell population with the same % of parameter variation generated a distribution of *Let-7b*|*CTNNB1* values similar to that seen in HCC patients and cell lines (Fig. 5D, right), indicating that the model accounts for the sample's heterogeneity.

We analyzed the robustness of the bistable switch dynamics in normal and HCC conditions towards variation in the value of different parameters (Fig. S6-7). These results showed that GRN dynamics rest on a robust reversible bistable switch that should not occur in normal conditions. Meanwhile, in HCC conditions, GRN dynamics exhibit a sensitive irreversible bistable switch that could cause heterogeneity among tumor samples.

GRN status in gastrointestinal cancers

We next verified if the GRN might be activated in cancers distinct from HCC. We looked at RNA expression in TCGA cohorts of cholangiocarcinoma, stomach and esophageal carcinoma, colorectal adenocarcinoma, thyroid carcinoma, kidney renal clear cell carcinoma, breast carcinoma, bladder urothelial carcinoma, and lung adenocarcinoma (Fig. 7A and Fig. S8). Significant and consistent induction of all tumor-promoting GRN components was detected in cholangiocarcinoma, stomach and esophageal carcinoma carcinoma, and colorectal adenocarcinoma, *i.e.* in gastrointestinal cancers. *Let-7b* did not show the expected reduction in those tumors, yet depending on the tumor type, other *Let-7* family members were downregulated, like in cholangiocarcinoma where *Let-7c* is strongly repressed. The other cancer types did not show a consistent increase in GRN components, suggesting that the GRN does not display tumor-promoting activity in these tumors.

Tu and coworkers developed a mouse model of colorectal cancer with a similar gene expression pattern as in human cancer,[22] which offered the opportunity to validate the mathematical model in this cancer type. The published mouse data showed that induction of *Lin28b* for 6 months triggered cell proliferation, downregulation of *Let-7* and upregulation of *Sox9*, the latter being illustrated by immunostaining.[22] The subsequent removal of *Lin28b* for 2 months partially restored the expression of *Let-7* and proliferation markers. As in the experiments,[22] the mathematical model predicted that a 6-month induction of *Lin28b* causes downregulation of *Let-7b* and upregulation of *Sox9*, and that those expression levels partially revert to near normal values 2 months after *LIN28* LIN28B removal (Fig. 7B), thereby validating the GRN model in colorectal cancer. The mathematical model also showed that the expression of *MYC* (Fig. 7B) and *CTNNB1* mRNAs (not shown) is not affected by *LIN28B* induction, which fitted well with the observations.[22]

Discussion

Here we identified a GRN which combines CTNNB1, SMARCA4, SOX9, LIN28B, Let-7b, TP53 and MYC. MYC expression is not strongly affected but it plays a role in mouse models of HCC.[12] Also, when comparing the HCC samples with low MYC expression with those expressing high levels of MYC, we found that

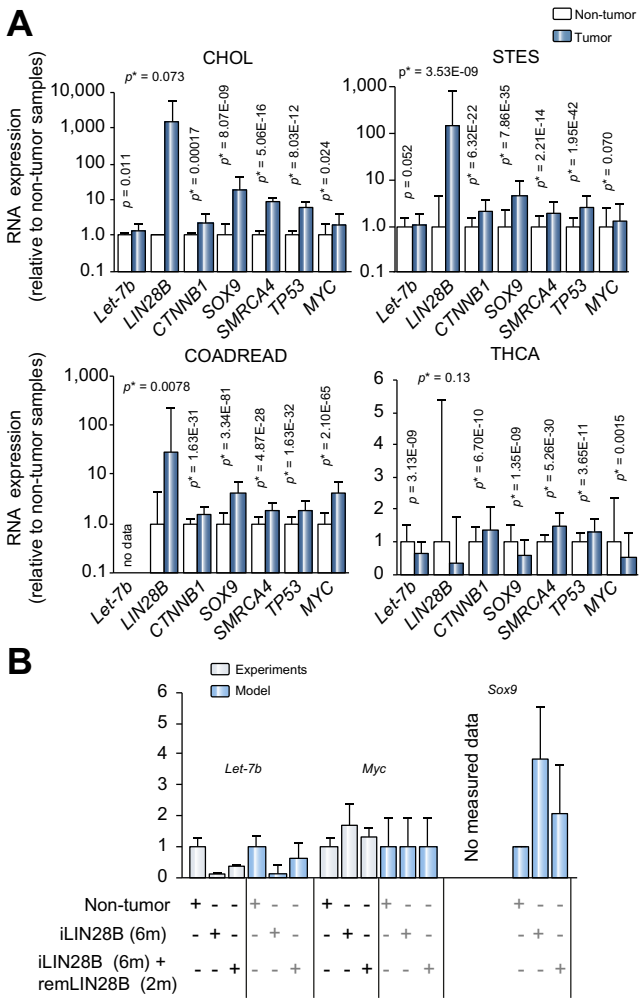


Fig. 7. Cancer-type specificity of GRN functionality and validation of the model in colorectal cancer. (A) Tumor-promoting components of the GRN are induced in CHOL, STES, and COADREAD. THCA had a distinct gene expression profile. Data (mean ± SD) are from TCGA; the number of normal and tumor samples is 9 and 36 (CHOL), 50 and 600 (STES), 51 and 624 (COADREAD), and 59 and 501 (THCA). (B) RNA levels of *Let-7b*, *Myc* and *Sox9* in experiment (grey) and in the model (blue) in non-tumor condition, after 6 months of *Lin28b* induction (LIN28B), and after 6 months of *Lin28b* induction followed by 2 months of partial *Lin28b* removal (remLIN28B) (see supplementary information for parameter values). CHOL, cholangiocarcinoma; COADREAD, colorectal adenocarcinoma; GRN, gene regulatory network; STES, stomach and esophageal carcinoma; TCGA, The Cancer Genome Atlas; THCA, thyroid carcinoma. *p* and *p** values calculated as in Fig. 1.

MYC expression positively correlates with nearly all GRN components, namely CTNNB1, SOX9, SMARCA4 and TP53 (Fig. S2).

Since the mathematical model resorts to 2 sets of parameters, one for normal conditions and one for tumors and cell lines, it accounts for the GRN dynamics in each state, but not for the dynamical transition between normal and HCC states. Indeed, to properly calibrate the model on mRNA expression, the transcription rates of *LIN28B*, *SOX9*, mutant *TP53*, *CTNNB1* and *SMARCA4* had to be increased in the HCC conditions ('*tumor*' parameter in Tables S4-5). This indicates that the interactions between GRN components are not sufficient to account for the normal-to-HCC transition. The parameters implicitly integrate the impact of external regulators of the GRN. Yet, the modelling strategy cannot integrate the full spectrum of regulations, and a number of regulators might not be known. Since modelling the transition from normal to HCC requires adaptation of the transcription rates of *LIN28B*, *SOX9*, mutant *TP53*, *CTNNB1* and *SMARCA4*, we suggest that the mechanisms controlling the expression of those genes warrant further investigation.

Classifying HCCs identified proliferative versus non-proliferative classes, or subdivided patients with HCC into subgroups G1 to G6.[19] Consistently, the GRN is most activated in the G1, G2 and G3 subgroups which are associated with poor differentiation, severe prognosis, and overexpression of genes regulating cell proliferation.[19] The targets of CTNNB1 are heterogeneous as CTNNB1 induces progenitor-type genes in tumor cells such as cyclin-D1 or VEGF-A, but also regulates expression of antioxidant, pro-survival and pro-hepatocyte differentiation genes in normal hepatocytes.[23] The G5 and G6 subgroups are associated with activation of CTNNB1 targets typical of mature hepatocytes, whereas the G1, G2, and G3 subgroups show predominant activation of progenitor-type targets. Interestingly, high GRN status was associated with high expression of progenitor-type CTNNB1/TCF/LEF targets (Fig. S9 and Table S8). *TERT* promoter mutations are found in 60% of HCCs.[13,24] However, out of the 368 HCC samples of the TCGA cohort, the 100 samples with the highest *TERT* expression did not consistently misexpress GRN components compared to the 268 HCC samples with lower *TERT* expression levels (Fig. S9C).

The GRN comprises positive feedback loops which are at the origin of bistable switches. To our knowledge, our work identified the first potential irreversible bistable switch involved in a specific subset of HCC tumors. The identification of irreversible states in HCC provides evidence that targeting specific GRN components using drugs might be therapeutically ineffective when cancer cells are in a locked state with regard to the function of the drug target. Also, our comparison of patient tumors and cultured HCC lines indicates that cell lines may be in a state distinct from that of tumors with regard to the function of the GRN.

To facilitate the analysis of the GRN, we set up a web-platform allowing researchers to check the GRN status in new samples of HCC and gastrointestinal tumors. This platform presents a graphical user interface that integrates the expression levels of the GRN components of tumor samples, and which implements the mathematical model to test which component of the network is the best target to modulate the network dynamics (http://biomodelling.eu/apps.html).

In conclusion, we anticipate that identification and dynamical modelling of a set of GRNs composed of interacting components that are often mutated and/or dysregulated in HCC, as is the case here for CTNNB1 and *Let-7*/LIN28,[6,10,12,25] will contribute to provide a global picture of tumor-promoting gene function in HCC. Our study presents a concept and tools that could help in the design of bespoke therapies for the treatment of HCC.

Financial support

The work of FPL was supported by the Interuniversity Attraction Pole Programme (Belgian Science Policy, PVII-47), the D.G. Higher Education and Scientific Research of the French Community of Belgium (ARC 15/20-065), the F.R.S.-FNRS (Belgium: Grants T.007214 and J.0058.15), and the Belgian Foundation Against Cancer (grant 2014-125). J.Z-R's group is supported by INSERM (France), the Ligue Nationale contre le Cancer (France, Equipe Labellisée), Labex OncoImmunology (France, investissement d'avenir), Coup d'Elan de la Fondation Bettencourt-

Shueller (France), the SIRIC CARPEM (France) and Fondation Mérieux (France). J.U.M. is supported by grants from the German Research Foundation (MA 4443/2-2) and the Volkswagen Foundation (Germany, Lichtenberg program).

Conflict of interest

The authors declare no conflicts of interest that pertain to this work.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

C.G. and F.P.L. conceived and designed the study. C.G. performed transcriptomic data analysis and mathematical modelling. M.D-L., K.K. and S. Cordi performed transfection experiments and quantified RNA and proteins. L.G. created the web-based platform under supervision of E.H. A.L. did the survival analysis and helped with transcriptomic data analysis. S. Caruso, G.C. and J.Z-R. provided transcriptomic data and human cell lines and contributed to the interpretation of the modelling results. D.C. and J.U.M. provided primary human HCC cells. J.T. and S. P.M. shared and analysed data from transgenic mice. C.G., M. D-L. and F.P.L. wrote the manuscript with input from all authors.

Acknowledgements

The authors thank P. Jacquemin, C. Pierreux, L. Desmet and the Lemaigre laboratory members for help, and L. Nguyen and H. Zhu for sharing information.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhep.2019.03.024.

References

Author names in bold designate shared co-first authorship

[1] Calderaro J, Couchy G, Imbeaud S, Amaddeo G, Letouze E, Blanc JF, et al. Histological subtypes of hepatocellular carcinoma are related to gene mutations and molecular tumour classification. J Hepatol 2017;67:727–738.
[2] Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhori G, et al. A pathology atlas of the human cancer transcriptome. Science 2017;357, pii: eaan2507.
[3] **Schulze K**, **Imbeaud S**, **Letouze E**, Alexandrov LB, Calderaro J, Rebouissou S, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet 2015;47:505–511.
[4] Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell 2017;169, 1327–1341 e1323.
[5] Breuhahn K, Gores G, Schirmacher P. Strategies for hepatocellular carcinoma therapy and diagnostics: lessons learned from high throughput and profiling approaches. Hepatology 2011;53:2112–2121.
[6] Tao J, Zhang R, Singh S, Poddar M, Xu E, Oertel M, et al. Targeting beta-catenin in hepatocellular cancers induced by coexpression of mutant beta-catenin and K-Ras in mice. Hepatology 2017;65:1581–1599.
[7] **Zeng X**, **Lin Y**, **Yin C**, Zhang X, Ning BF, Zhang Q, et al. Recombinant adenovirus carrying the hepatocyte nuclear factor-1alpha gene inhibits hepatocellular carcinoma xenograft growth in mice. Hepatology 2011;54:2036–2047.
[8] Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. Nat Genet 2002;31:339–346.
[9] **Jung KH**, **Zhang J**, Zhou C, Shen H, Gagea M, Rodriguez-Aguayo C, et al. Differentiation therapy for hepatocellular carcinoma: Multifaceted effects of miR-148a on tumor growth and phenotype and liver fibrosis. Hepatology 2016;63:864–879.
[10] Takashima Y, Terada M, Udono M, Miura S, Yamamoto J, Suzuki A. Suppression of lethal-7b and miR-125a/b maturation by Lin28b enables maintenance of stem cell properties in hepatoblasts. Hepatology 2016;64:245–260.
[11] Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, et al. Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet 2009;41:843–848.
[12] **Nguyen LH**, **Robinton DA**, **Seligson MT**, Wu L, Li L, Rakheja D, et al. Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models. Cancer Cell 2014;26:248–261.
[13] Zucman-Rossi J, Villanueva A, Nault JC, Llovet JM. Genetic landscape and biomarkers of hepatocellular carcinoma. Gastroenterology 2015;149, 1226–1239 e1224.
[14] Wilkinson AC, Nakauchi H, Gottgens B. Mammalian transcription factor networks: recent advances in interrogating biological complexity. Cell Syst 2017;5:319–331.
[15] **Rata S**, **Suarez Peredo Rodriguez MF**, Joseph S, Peter N, Echegaray Iturra F, Yang F, et al. Two interlinked bistable switches govern mitotic control in mammalian cells. Curr Biol 2018;28, 3824–3832 e3826.
[16] Ryu H, Chung M, Dobrzynski M, Fey D, Blum Y, Lee SS, et al. Frequency modulation of ERK activation dynamics rewires cell fate. Mol Syst Biol 2015;11:838.
[17] Castven D, Becker D, Czauderna C, Wilhelm D, Andersen JB, Strand S, et al. Application of patient-derived liver cancer cells for phenotypic characterization and therapeutic target identification. Int J Cancer 2019;144:2782–2794.
[18] Kaposi-Novak P, Lee JS, Gomez-Quiroz L, Coulouarn C, Factor VM, Thorgeirsson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. J Clin Invest 2006;116:1582–1595.
[19] **Boyault S**, **Rickman DS**, **de Reynies A**, Balabaud C, Rebouissou S, Jeannot E, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. Hepatology 2007;45:42–52.
[20] Rebouissou S, La Bella T, Rekik S, Imbeaud S, Calatayud AL, Rohr-Udilova N, et al. Proliferation markers are associated with MET expression in hepatocellular carcinoma and predict tivantinib sensitivity in vitro. Clin Cancer Res 2017;23:4364–4375.
[21] **Roos M**, **Pradere U**, Ngondo RP, Behera A, Allegrini S, Civenni G, et al. A small-molecule inhibitor of Lin28. ACS Chem Biol 2016;11:2773–2781.
[22] Tu HC, Schwitalla S, Qian Z, LaPier GS, Yermalovich A, Ku YC, et al. LIN28 cooperates with WNT signaling to drive invasive intestinal and colorectal adenocarcinoma in mice and humans. Genes Dev 2015;29: 1074–1086.
[23] Nejak-Bowen KN, Monga SP. Beta-catenin signaling, liver regeneration and hepatocellular cancer: sorting the good from the bad. Semin Cancer Biol 2011;21:44–58.
[24] Nault JC, Mallet M, Pilati C, Calderaro J, Bioulac-Sage P, Laurent C, et al. High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. Nat Commun 2013;4:2218.
[25] **Wu L**, **Nguyen LH**, Zhou K, de Soysa TY, Li L, Miller JB, et al. Precise let-7 expression levels balance organ regeneration against tumor suppression. Elife 2015;4 e09431.

Bibliography

[1] D. W. Thompson. "On growth and form." English. In: *On growth and form.* (1942). Publisher: Cambridge Univ. Press.

[2] A. M. Turing. "The chemical basis of morphogenesis". en. In: *Bulletin of Mathematical Biology* 52.1 (Jan. 1990). ISSN: 1522-9602. DOI: 10.1007/BF02459572.

[3] Erwin Schrodinger. *What is Life?: With Mind and Matter and Autobiographical Sketches.* Canto Classics. Cambridge: Cambridge University Press, 2012. ISBN: 978-1-107-60466-7. DOI: 10.1017/CBO9781107295629.

[4] Aubrey Gorbman. "Some Aspects of the Comparative Biochemistry of Iodine Utilization and the Evolution of Thyroidal Function". en. In: *Physiological Reviews* 35.2 (Apr. 1955). ISSN: 0031-9333, 1522-1210. DOI: 10.1152/physrev.1955.35.2.336.

[5] Yoko Kameda et al. "Expression of the Epithelial Marker E-Cadherin by Thyroid C Cells and Their Precursors During Murine Development". en. In: *Journal of Histochemistry & Cytochemistry* 55.10 (Oct. 2007). ISSN: 0022-1554, 1551-5044. DOI: 10.1369/jhc.7A7179.2007.

[6] Carla Visciano et al. "Tumor-Associated Mast Cells in Thyroid Cancer". en. In: *International Journal of Endocrinology* 2015 (2015). ISSN: 1687-8337, 1687-8345. DOI: 10.1155/2015/705169.

[7] DU Silverthorn. *Human Physiology: An Integrated Approach, 4th Edition.* Pearson. 2007. ISBN: 978-0-8053-6851-2.

[8] Mikael Nilsson and Henrik Fagman. "Development of the thyroid gland". en. In: *Development* 144.12 (June 2017). ISSN: 0950-1991, 1477-9129. DOI: 10.1242/dev.145615.

[9] Bernard Rousset. *Chapter 2 Thyroid Hormone Synthesis And Secretion (in: Endotext).* MDText.com, Inc. 2000.

[10] A. Z. Wang, G. K. Ojakian, and W. J. Nelson. “Steps in the morphogenesis of a polarized epithelium. I. Uncoupling the roles of cell-cell and cell-substratum contact in establishing plasma membrane polarity in multicellular epithelial (MDCK) cysts”. en. In: *Journal of Cell Science* 95.1 (Jan. 1990). Publisher: The Company of Biologists Ltd Section: Journal Articles. ISSN: 0021-9533, 1477-9137.

[11] María Fernanda Izaguirre and Victor Hugo Casco. “E-cadherin roles in animal biology: A perspective on thyroid hormone-influence”. en. In: *Cell Communication and Signaling* 14.1 (Dec. 2016). ISSN: 1478-811X. DOI: 10.1186/s12964-016-0150-1.

[12] C Maenhaut. *Ontogeny, Anatomy, Metabolism and Physiology of the Thyroid (in: Endotext)*. MDtext.com, Inc. 2015.

[13] Mario De Felice and Roberto Di Lauro. “Thyroid Development and Its Disorders: Genetics and Molecular Mechanisms”. en. In: *Endocrine Reviews* 25.5 (Oct. 2004). ISSN: 0163-769X, 1945-7189. DOI: 10.1210/er.2003-0028.

[14] Kazuhiro Maeda et al. “Postotic and preotic cranial neural crest cells differently contribute to thyroid development”. en. In: *Developmental Biology* 409.1 (Jan. 2016). ISSN: 0012-1606. DOI: 10.1016/j.ydbio.2015.10.026.

[15] Henrik Fagman, Louise Andersson, and Mikael Nilsson. “The developing mouse thyroid: Embryonic vessel contacts and parenchymal growth pattern during specification, budding, migration, and lobulation”. en. In: *Developmental Dynamics* 235.2 (Feb. 2006). ISSN: 1058-8388, 1097-0177. DOI: 10.1002/dvdy.20653.

[16] Aurore Carre et al. “Hes1 Is Required for Appropriate Morphogenesis and Differentiation during Mouse Thyroid Gland Development”. en. In: *PLOS ONE* 6.2 (2011). Publisher: Public Library of Science. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0016752.

[17] Mylah Villacorte et al. “Thyroid follicle development requires Smad1/5- and endothelial cell-dependent basement membrane assembly”. In: *Development (Cambridge, England)* 143.11 (June 2016). ISSN: 0950-1991. DOI: 10.1242/dev.134171.

[18] Kazunori Ishimura and Hisao Fujita. “Development of cell-to-cell relationships in the thyroid gland of the chick embryo”. en. In: *Cell and Tissue Research* 198.1 (Apr. 1979). ISSN: 1432-0878. DOI: 10.1007/BF00234830.

[19] Sara Sigurbjörnsdóttir, Renjith Mathew, and Maria Leptin. "Molecular mechanisms of de novo lumen formation". en. In: *Nature Reviews Molecular Cell Biology* 15.10 (Oct. 2014). ISSN: 1471-0072, 1471-0080. DOI: 10.1038/nrm3871.

[20] Sumire Ishida et al. "Epithelial Sheet Folding Induces Lumen Formation by Madin-Darby Canine Kidney Cells in a Collagen Gel". en. In: *PLOS ONE* 9.8 (2014). Publisher: Public Library of Science. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0099655.

[21] Gabor Szinnai et al. "Sodium/Iodide Symporter (NIS) Gene Expression Is the Limiting Step for the Onset of Thyroid Function in the Human Fetus". en. In: *The Journal of Clinical Endocrinology & Metabolism* 92.1 (Jan. 2007). Publisher: Oxford Academic. ISSN: 0021-972X. DOI: 10.1210/jc.2006-1450.

[22] Anne-Christine Hick et al. "Reciprocal epithelial:endothelial paracrine interactions during thyroid development govern follicular organization and C-cells differentiation". en. In: *Developmental Biology* 381.1 (Sept. 2013). ISSN: 00121606. DOI: 10.1016/j.ydbio.2013.04.022.

[23] M. P. Postiglione et al. "Role of the thyroid-stimulating hormone receptor signaling in development and differentiation of the thyroid gland". en. In: *Proceedings of the National Academy of Sciences* 99.24 (Nov. 2002). Publisher: National Academy of Sciences Section: Biological Sciences. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.242328999.

[24] Lara P. Fernández et al. "New Insights into FoxE1 Functions: Identification of Direct FoxE1 Targets in Thyroid Cells". en. In: *PLoS ONE* 8.5 (May 2013). Ed. by Alfredo Fusco. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0062849.

[25] Rosanna Parlato et al. "An integrated regulatory network controlling survival and migration in thyroid organogenesis". en. In: *Developmental Biology* 276.2 (Dec. 2004). ISSN: 00121606. DOI: 10.1016/j.ydbio.2004.08.048.

[26] Clifford W Bogue et al. "Hex expression suggests a role in the development and function of organs derived from foregut endoderm". en. In: (July 2000).

[27] Daniel Silberschmidt et al. "In vivo role of different domains and of phosphorylation in the transcription factor Nkx2-1". en. In: *BMC Developmental Biology* 11.1 (2011). ISSN: 1471-213X. DOI: 10.1186/1471-213X-11-9.

[28] Risheng Ma, Rauf Latif, and Terry F. Davies. "Human Embryonic Stem Cells Form Functional Thyroid Follicles". In: *Thyroid* 25.4 (Apr. 2015). ISSN: 1050-7256. DOI: 10.1089/thy.2014.0537.

[29] Francesco Antonica et al. "Generation of functional thyroid from embryonic stem cells". In: *Nature* 491.7422 (Nov. 2012). ISSN: 0028-0836. DOI: 10.1038/nature11525.

[30] E. Aaron Runkle et al. "Occludin Is a Direct Target of Thyroid Transcription Factor-1 (TTF-1/NKX2–1)". en. In: *Journal of Biological Chemistry* 287.34 (Aug. 2012). Publisher: American Society for Biochemistry and Molecular Biology. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M112.367987.

[31] H. Devos et al. "A search for the possible molecular mechanisms of thyroid dysgenesis: sex ratios and associated malformations". eng. In: *The Journal of Clinical Endocrinology and Metabolism* 84.7 (July 1999). ISSN: 0021-972X. DOI: 10.1210/jcem.84.7.5831.

[32] Jd Ramsden. "Angiogenesis in the thyroid gland". en. In: *Journal of Endocrinology* 166.3 (Sept. 2000). ISSN: 0022-0795, 1479-6805. DOI: 10.1677/joe.0.1660475.

[33] D. H. Ausprunk and J. Folkman. "Migration and proliferation of endothelial cells in preformed and newly formed blood vessels during tumor angiogenesis". eng. In: *Microvascular Research* 14.1 (July 1977). ISSN: 0026-2862. DOI: 10.1016/0026-2862(77)90141-8.

[34] C. L. Stokes, D. A. Lauffenburger, and S. K. Williams. "Migration of individual microvessel endothelial cells: stochastic model and parameter measurement". eng. In: *Journal of Cell Science* 99 (Pt 2) (June 1991). ISSN: 0021-9533.

[35] Levi Wood, Roger Kamm, and Harry Asada. "Stochastic modeling and identification of emergent behaviors of an Endothelial Cell population in angiogenic pattern formation". In: *The International Journal of Robotics Research* 30.6 (May 2011). Publisher: SAGE Publications Ltd STM. ISSN: 0278-3649. DOI: 10.1177/0278364910394690.

[36] Holger Gerhardt et al. "VEGF guides angiogenic sprouting utilizing endothelial tip cell filopodia". en. In: *Journal of Cell Biology* 161.6 (June 2003). ISSN: 1540-8140, 0021-9525. DOI: 10.1083/jcb.200302047.

[37] Holger Gerhardt. "VEGF and endothelial guidance in angiogenic sprouting". en. In: *Organogenesis* 4.4 (Oct. 2008). ISSN: 1547-6278, 1555-8592. DOI: 10.4161/org.4.4.7414.

[38] S. R. McDougall et al. "A Hybrid Discrete-Continuum Mathematical Model of Pattern Prediction in the Developing Retinal Vasculature". en. In: *Bulletin of Mathematical Biology* 74.10 (Oct. 2012). ISSN: 0092-8240, 1522-9602. DOI: 10.1007/s11538-012-9754-9.

[39] Henrik Fagman et al. "Genetic Deletion of Sonic Hedgehog Causes Hemiagenesis and Ectopic Development of the Thyroid in Mouse". English. In: *The American Journal of Pathology* 164.5 (May 2004). Publisher: Elsevier. ISSN: 0002-9440, 1525-2191. DOI: 10.1016/S0002-9440(10)63745-5.

[40] David M. Bryant et al. "A molecular network for de novo generation of the apical surface and lumen". In: *Nature cell biology* 12.11 (Nov. 2010). ISSN: 1465-7392. DOI: 10.1038/ncb2106.

[41] Camilla Ingeson-Carlsson and Mikael Nilsson. "Dual contribution of MAPK and PI3K in epidermal growth factor-induced destabilization of thyroid follicular integrity and invasion of cells into extracellular matrix". en. In: *Experimental Cell Research* 326.2 (Aug. 2014). ISSN: 0014-4827. DOI: 10.1016/j.yexcr.2014.04.006.

[42] Theo Vos et al. "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015". English. In: *The Lancet* 388.10053 (Oct. 2016). Publisher: Elsevier. ISSN: 0140-6736, 1474-547X. DOI: 10.1016/S0140-6736(16)31678-6.

[43] G. Wayne Brodland. "How computational models can help unlock biological systems". en. In: *Seminars in Cell & Developmental Biology* 47-48 (Dec. 2015). ISSN: 10849521. DOI: 10.1016/j.semcdb.2015.07.001.

[44] James M. Osborne et al. "Comparing individual-based approaches to modelling the self-organization of multicellular tissues". en. In: *PLOS Computational Biology* 13.2 (Feb. 2017). Ed. by Qing Nie. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005387.

[45] Bruno Hay Mele et al. “Cell-Based Models in Plant Developmental Biology: Insights into Hybrid Approaches”. en. In: *Frontiers in Environmental Science* 3 (Nov. 2015). ISSN: 2296-665X. DOI: 10.3389/fenvs.2015.00073.

[46] Margriet M. Palm and Roeland M. H. Merks. “Large-Scale Parameter Studies of Cell-Based Models of Tissue Morphogenesis Using CompuCell3D or VirtualLeaf”. In: *Tissue Morphogenesis*. Ed. by Celeste M. Nelson. Vol. 1189. Series Title: Methods in Molecular Biology. New York, NY: Springer New York, 2015. ISBN: 978-1-4939-1163-9 978-1-4939-1164-6. DOI: 10.1007/978-1-4939-1164-6_20.

[47] John Metzcar et al. “A Review of Cell-Based Computational Modeling in Cancer Biology”. en. In: *JCO Clinical Cancer Informatics* 3 (Nov. 2019). ISSN: 2473-4276. DOI: 10.1200/CCI.18.00069.

[48] D. C. Walker and J. Southgate. “The virtual cell–a candidate co-ordinator for ’middle-out’ modelling of biological systems”. en. In: *Briefings in Bioinformatics* 10.4 (July 2009). ISSN: 1467-5463, 1477-4054. DOI: 10.1093/bib/bbp010.

[49] P. Van Liedekerke et al. “Simulating tissue mechanics with agent-based models: concepts, perspectives and some novel results”. en. In: *Computational Particle Mechanics* 2.4 (Dec. 2015). ISSN: 2196-4378, 2196-4386. DOI: 10.1007/s40571-015-0082-3.

[50] G.Bard Ermentrout and Leah Edelstein-Keshet. “Cellular Automata Approaches to Biological Modeling”. en. In: *Journal of Theoretical Biology* 160.1 (Jan. 1993). ISSN: 00225193. DOI: 10.1006/jtbi.1993.1007.

[51] Martin Gardner. “Mathematical Games”. en. In: *SCIENTIFIC AMERICAN* (1970).

[52] Michael Cohen, Buzz Baum, and Mark Miodownik. “The importance of structured noise in the generation of self-organizing tissue patterns through contact-mediated cell-cell signalling”. eng. In: *Journal of the Royal Society, Interface* 8.59 (June 2011). ISSN: 1742-5662. DOI: 10.1098/rsif.2010.0488.

[53] J. R. Collier et al. “Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling”. eng. In: *Journal of Theoretical Biology* 183.4 (Dec. 1996). ISSN: 0022-5193. DOI: 10.1006/jtbi.1996.0233.

[54] Sarah J. Bray. “Notch signalling: a simple pathway becomes complex”. en. In: *Nature Reviews Molecular Cell Biology* 7.9 (Sept. 2006). Number: 9 Publisher: Nature Publishing Group. ISSN: 1471-0080. DOI: 10.1038/nrm2009.

[55] R A Gatenby et al. “Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer”. en. In: *British Journal of Cancer* 97.5 (Aug. 2007). ISSN: 0007-0920, 1532-1827. DOI: 10.1038/sj.bjc.6603922.

[56] Peter B. Armstrong. “Cell Sorting Out: The Self-Assembly of Tissues In Vitro”. In: *Critical Reviews in Biochemistry and Molecular Biology* 24.2 (Jan. 1989). Publisher: Taylor & Francis _eprint: https://doi.org/10.3109/10409238909086396. ISSN: 1040-9238. DOI: 10.3109/10409238909086396.

[57] James A. Glazier and François Graner. “Simulation of the differential adhesion driven rearrangement of biological cells”. en. In: *Physical Review E* 47.3 (Mar. 1993). ISSN: 1063-651X, 1095-3787. DOI: 10.1103/PhysRevE.47.2128.

[58] D.C. Walker et al. “Agent-Based Computational Modeling of Wounded Epithelial Cell Monolayers”. en. In: *IEEE Transactions on Nanobioscience* 3.3 (Sept. 2004). ISSN: 1536-1241. DOI: 10.1109/TNB.2004.833680.

[59] F. A. Meineke, C. S. Potten, and M. Loeffler. “Cell migration and organization in the intestinal crypt using a lattice-free model”. en. In: *Cell Proliferation* 34.4 (Aug. 2001). ISSN: 0960-7722, 1365-2184. DOI: 10.1046/j.0960-7722.2001.00216.x.

[60] Bart Smeets et al. “Quantifying the mechanical micro-environment during three-dimensional cell expansion on microbeads by means of individual cell-based modelling”. In: *Computer Methods in Biomechanics and Biomedical Engineering* 16.10 (Oct. 2013). ISSN: 1025-5842. DOI: 10.1080/10255842.2013.829461.

[61] Alexander G. Fletcher et al. “Vertex Models of Epithelial Morphogenesis”. en. In: *Biophysical Journal* 106.11 (June 2014). ISSN: 00063495. DOI: 10.1016/j.bpj.2013.11.4498.

[62] Payman Mosaffa et al. “Cell-centred model for the simulation of curved cellular monolayers”. en. In: *Computational Particle Mechanics* 2.4 (Dec. 2015). ISSN: 2196-4386. DOI: 10.1007/s40571-015-0043-x.

[63] Thomas Lecuit and Pierre-François Lenne. “Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis”. en. In: *Nature Reviews Molecular Cell Biology* 8.8 (Aug. 2007). Number: 8 Publisher: Nature Publishing Group. ISSN: 1471-0080. DOI: 10.1038/nrm2222.

[64] Marco Kokic et al. *Minimisation of surface energy drives apical epithelial organisation and gives rise to Lewis' law*. en. preprint. Biophysics, Mar. 2019. DOI: 10.1101/590729.

[65] Hisao Honda and Goro Eguchi. "How much does the cell boundary contract in a monolayered cell sheet?" en. In: *Journal of Theoretical Biology* 84.3 (June 1980). ISSN: 00225193. DOI: 10.1016/S0022-5193(80)80021-X.

[66] Michael Weliky and George Oster. "The mechanical basis of cell rearrangement". en. In: ().

[67] L. Hufnagel et al. "On the mechanism of wing size determination in fly development". en. In: *Proceedings of the National Academy of Sciences* 104.10 (Mar. 2007). ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0607134104.

[68] Georgios Trichas et al. "Multi-Cellular Rosettes in the Mouse Visceral Endoderm Facilitate the Ordered Migration of Anterior Visceral Endoderm Cells". en. In: *PLoS Biology* 10.2 (Feb. 2012). Ed. by Charles D. Little. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.1001256.

[69] Guillaume Salbreux et al. "Coupling Mechanical Deformations and Planar Cell Polarity to Create Regular Patterns in the Zebrafish Retina". en. In: *PLoS Computational Biology* 8.8 (Aug. 2012). Ed. by Stanislav Shvartsman. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002618.

[70] Reza Farhadifar et al. "The Influence of Cell Mechanics, Cell-Cell Interactions, and Proliferation on Epithelial Packing". en. In: *Current Biology* 17.24 (Dec. 2007). ISSN: 09609822. DOI: 10.1016/j.cub.2007.11.049.

[71] Simon Tanaka, David Sichau, and Dagmar Iber. "LBIBCell: a cell-based simulation environment for morphogenetic problems". en. In: *Bioinformatics* 31.14 (July 2015). ISSN: 1367-4803, 1460-2059. DOI: 10.1093/bioinformatics/btv147.

[72] Katarzyna A. Rejniak. "An immersed boundary framework for modelling the growth of individual cells: An application to the early tumour development". en. In: *Journal of Theoretical Biology* 247.1 (July 2007). ISSN: 00225193. DOI: 10.1016/j.jtbi.2007.02.019.

[73] Paul Van Liedekerke et al. "A quantitative high-resolution computational mechanics cell model for growing and regenerating tissues". en. In: *Biomechanics and Modeling in Mechanobiology* 19.1 (Feb. 2020). ISSN: 1617-7959, 1617-7940. DOI: 10.1007/s10237-019-01204-7.

[74] Alexander R.A. Anderson et al. "Tumor Morphology and Phenotypic Evolution Driven by Selective Pressure from the Microenvironment". en. In: *Cell* 127.5 (Dec. 2006). ISSN: 00928674. DOI: 10.1016/j.cell.2006.09.042.

[75] H. Honda and T. Nagai. "Cell models lead to understanding of multi-cellular morphogenesis consisting of successive self-construction of cells". en. In: *Journal of Biochemistry* 157.3 (Mar. 2015). ISSN: 0021-924X. DOI: 10.1093/jb/mvu088.

[76] Y. Mao et al. "Planar polarization of the atypical myosin Dachs orients cell divisions in Drosophila". en. In: *Genes & Development* 25.2 (Jan. 2011). ISSN: 0890-9369. DOI: 10.1101/gad.610511.

[77] Tatsuzo Nagai and Hisao Honda. "A dynamic cell model for the formation of epithelial tissues". en. In: *Philosophical Magazine B* 81.7 (July 2001). ISSN: 1364-2812, 1463-6417. DOI: 10.1080/13642810108205772.

[78] Philipp Spahn and Rolf Reuter. "A Vertex Model of Drosophila Ventral Furrow Formation". en. In: *PLoS ONE* 8.9 (Sept. 2013). Ed. by Alexandre J. Kabla. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0075051.

[79] Alexander G. Fletcher et al. "Implementing vertex dynamics models of cell populations in biology within a consistent computational framework". en. In: *Progress in Biophysics and Molecular Biology* 113.2 (Nov. 2013). ISSN: 00796107. DOI: 10.1016/j.pbiomolbio.2013.09.003.

[80] GM Cooper. *The Cell: A Molecular Approach*. 2nd ed. Sinauer Associates, 2000.

[81] J. Todd Blankenship et al. "Multicellular Rosette Formation Links Planar Cell Polarity to Tissue Morphogenesis". en. In: *Developmental Cell* 11.4 (Oct. 2006). ISSN: 1534-5807. DOI: 10.1016/j.devcel.2006.09.007.

[82] Joe Pitt-Francis et al. "Chaste: A test-driven approach to software development for biological modelling". en. In: *Computer Physics Communications* 180.12 (Dec. 2009). ISSN: 00104655. DOI: 10.1016/j.cpc.2009.07.019.

[83] Claire J. Tomlin and Jeffrey D. Axelrod. “Biology by numbers: mathematical modelling in developmental biology”. en. In: *Nature Reviews Genetics* 8.5 (May 2007). Number: 5 Publisher: Nature Publishing Group. ISSN: 1471-0064. DOI: 10.1038/nrg2098.

[84] Jonathan Degosserie et al. “Extracellular vesicles from endothelial progenitor cells promote thyroid follicle formation”. In: *Journal of Extracellular Vesicles* 7.1 (June 2018). ISSN: 2001-3078. DOI: 10.1080/20013078.2018.1487250.

[85] G. Wayne Brodland and Jim H. Veldhuis. “Computer simulations of mitosis and interdependencies between mitosis orientation, cell shape and epithelia reshaping”. eng. In: *Journal of Biomechanics* 35.5 (May 2002). ISSN: 0021-9290. DOI: 10.1016/s0021-9290(02)00006-4.

[86] Qiao Zhou et al. “A multipotent progenitor domain guides pancreatic organogenesis”. eng. In: *Developmental Cell* 13.1 (July 2007). ISSN: 1534-5807. DOI: 10.1016/j.devcel.2007.06.001.

[87] Shawn Liang et al. “A branching morphogenesis program governs embryonic growth of the thyroid gland”. eng. In: *Development (Cambridge, England)* 145.2 (Jan. 2018). ISSN: 1477-9129. DOI: 10.1242/dev.146829.

[88] Sanjay Gawade et al. “Cell Growth Dynamics in Embryonic and Adult Mouse Thyroid Revealed by a Novel Approach to Detect Thyroid Gland Subpopulations”. In: *Thyroid* 26.4 (Feb. 2016). Publisher: Mary Ann Liebert, Inc., publishers. ISSN: 1050-7256. DOI: 10.1089/thy.2015.0523.

[89] M. Polak et al. “Early pattern of differentiation in the human pancreas.” en. In: *Diabetes* 49.2 (Feb. 2000). Publisher: American Diabetes Association. ISSN: 0012-1797, 1939-327X. DOI: 10.2337/diabetes.49.2.225.

[90] Peter Eggli, Mike Boulton, and John Marshall. “Growth characteristics of central and peripheral bovine corneal epithelial cells in vitro”. en. In: *Graefe’s Archive for Clinical and Experimental Ophthalmology* 227.3 (May 1989). ISSN: 1435-702X. DOI: 10.1007/BF02172760.

[91] Adrian Salic and Timothy J. Mitchison. “A chemical method for fast and sensitive detection of DNA synthesis in vivo”. en. In: *Proceedings of the National Academy of Sciences* 105.7 (Feb. 2008). Publisher: National Academy of Sciences Section: Physical Sciences. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0712168105.

[92] Christoph Schmitz et al. “Current automated 3D cell detection methods are not a suitable replacement for manual stereologic cell counting”. English. In: *Frontiers in Neuroanatomy* 8 (2014). Publisher: Frontiers. ISSN: 1662-5129. DOI: 10.3389/fnana.2014.00027.

[93] Anna Kicheva et al. “Kinetics of morphogen gradient formation”. eng. In: *Science (New York, N.Y.)* 315.5811 (Jan. 2007). ISSN: 1095-9203. DOI: 10.1126/science.1135774.

[94] Aaron M. Smith et al. “Incorporating chemical signalling factors into cell-based models of growing epithelial tissues”. en. In: *Journal of Mathematical Biology* 65.3 (Sept. 2012). ISSN: 1432-1416. DOI: 10.1007/s00285-011-0464-y.

[95] Robert W. Cahn. “Quantitative image analysis of microstructures. Edited by H. E. Exner and H. P. Hougardy. DGM Informations- gesellschaft Verlag, Oberursel, 1988. 235 pp., bound, DM 95, $ 68.—ISBN 3-88355-132-5 (English Edition)”. en. In: *Advanced Materials* 2.2 (1990). ISSN: 1521-4095. DOI: 10.1002/adma.19900020216.

[96] Davide Heller et al. “EpiTools: An Open-Source Image Analysis Toolkit for Quantifying Epithelial Growth Dynamics”. en. In: *Developmental Cell* 36.1 (Jan. 2016). ISSN: 1534-5807. DOI: 10.1016/j.devcel.2015.12.012.

[97] Dene L. Farrell et al. “SEGGA: a toolset for rapid automated analysis of epithelial cell polarity and dynamics”. en. In: *Development* 144.9 (May 2017). Publisher: Oxford University Press for The Company of Biologists Limited Section: TECHNIQUES AND RESOURCES. ISSN: 0950-1991, 1477-9129. DOI: 10.1242/dev.146837.

[98] Raphaël Etournay et al. “TissueMiner: A multiscale analysis toolkit to quantify how cellular processes create tissue dynamics”. In: *eLife* 5 (May 2016). Ed. by W James Nelson. Publisher: eLife Sciences Publications, Ltd. ISSN: 2050-084X. DOI: 10.7554/eLife.14334.

[99] Luke Martin McCaffrey and Ian G. Macara. “Signaling Pathways in Cell Polarity”. In: *Cold Spring Harbor Perspectives in Biology* 4.6 (June 2012). ISSN: 1943-0264. DOI: 10.1101/cshperspect.a009654.

[100] Katherine H. Fisher, David Strutt, and Alexander G. Fletcher. “Integrating planar polarity and tissue mechanics in computational models of epithelial morphogenesis”. en. In: *bioRxiv* (May 2017). Publisher: Cold Spring Harbor Laboratory Section: Confirmatory Results. DOI: 10.1101/138172.

[101] MUSTAFA YILDIRIM, ŞENOL DANE, and BEDRI SEVEN. "Morphological Asymmetry in Thyroid Lobes, and Sex and Handedness Differences in Healthy Young Subjects". In: *International Journal of Neuroscience* 116.10 (Jan. 2006). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00207450500513997. ISSN: 0020-7454. DOI: 10.1080/00207450500513997.

[102] Michael Ying and Dennis M. C. Yung. "Asymmetry of Thyroid Lobe Volume in Normal Chinese Subjects: Association with Handedness and Position of Esophagus". en. In: *The Anatomical Record* 292.2 (2009). ISSN: 1932-8494. DOI: https://doi.org/10.1002/ar.20803.

[103] Amanda A. Amodeo and Jan M. Skotheim. "Cell-Size Control". In: *Cold Spring Harbor Perspectives in Biology* 8.4 (Apr. 2016). ISSN: 1943-0264. DOI: 10.1101/cshperspect.a019083.

[104] J. Coclet et al. "Cell Population Kinetics in Dog and Human Adult Thyroid". en. In: *Clinical Endocrinology* 31.6 (1989). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2265.1989.tb01290.x. ISSN: 1365-2265. DOI: https://doi.org/10.1111/j.1365-2265.1989.tb01290.x.

[105] J. E. Dumont et al. "Physiological and pathological regulation of thyroid cell proliferation and differentiation by thyrotropin and other factors". en. In: *Physiological Reviews* (July 1992). Publisher: American Physiological Society Bethesda, MD. DOI: 10.1152/physrev.1992.72.3.667.

[106] M. Baptist, J. E. Dumont, and P. P. Roger. "Demonstration of cell cycle kinetics in thyroid primary culture by immunostaining of proliferating cell nuclear antigen: differences in cyclic AMP-dependent and -independent mitogenic stimulations". eng. In: *Journal of Cell Science* 105 (Pt 1) (May 1993). ISSN: 0021-9533.

[107] Corinna Cozzitorto et al. "A Specialized Niche in the Pancreatic Microenvironment Promotes Endocrine Differentiation". eng. In: *Developmental Cell* 55.2 (Oct. 2020). ISSN: 1878-1551. DOI: 10.1016/j.devcel.2020.08.003.

[108] S. Ishihara et al. "Comparative study of non-invasive force and stress inference methods in tissue". en. In: *The European Physical Journal E* 36.4 (Apr. 2013). ISSN: 1292-8941, 1292-895X. DOI: 10.1140/epje/i2013-13045-8.

[109] H. Fagman et al. “Expression of Classical Cadherins in Thyroid Development: Maintenance of an Epithelial Phenotype throughout Organogenesis”. en. In: *Endocrinology* 144.8 (Aug. 2003). Publisher: Oxford Academic. ISSN: 0013-7227. DOI: 10.1210/en.2003-0393.

[110] Gaetano Calì et al. “Conditional inactivation of the E-cadherin gene in thyroid follicular cells affects gland development but does not impair junction formation”. eng. In: *Endocrinology* 148.6 (June 2007). ISSN: 0013-7227. DOI: 10.1210/en.2006-1344.

[111] Tiziana de Cristofaro et al. “An essential role for Pax8 in the transcriptional regulation of cadherin-16 in thyroid cells”. eng. In: *Molecular Endocrinology (Baltimore, Md.)* 26.1 (Jan. 2012). ISSN: 1944-9917. DOI: 10.1210/me.2011-1090.

[112] Petrina Koumarianou, Gonzalo Goméz-López, and Pilar Santisteban. “Pax8 controls thyroid follicular polarity through cadherin-16”. en. In: *Journal of Cell Science* 130.1 (Jan. 2017). ISSN: 0021-9533, 1477-9137. DOI: 10.1242/jcs.184291.

[113] Gary R. Mirams et al. “Chaste: An Open Source C++ Library for Computational Physiology and Biology”. en. In: *PLoS Computational Biology* 9.3 (Mar. 2013). Ed. by Andreas Prlic. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002970.

[114] Jason M. Ko and Daniel Lobo. “Continuous Dynamic Modeling of Regulated Cell Adhesion: Sorting, Intercalation, and Involution”. en. In: *Biophysical Journal* 117.11 (Dec. 2019). ISSN: 0006-3495. DOI: 10.1016/j.bpj.2019.10.032.

[115] Rui Zhen Tan and Keng-Hwee Chiam. “Computational Modeling Reveals that a Combination of Chemotaxis and Differential Adhesion Leads to Robust Cell Sorting during Tissue Patterning”. en. In: *PLOS ONE* 9.10 (Oct. 2014). Publisher: Public Library of Science. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0109286.

[116] Jessica L. Lee and Charles H. Streuli. “Integrins and epithelial cell polarity”. In: *Journal of Cell Science* 127.15 (Aug. 2014). ISSN: 0021-9533. DOI: 10.1242/jcs.146142.

[117] M. Canel et al. “E-cadherin-integrin crosstalk in cancer invasion and metastasis”. en. In: *Journal of Cell Science* 126.2 (Jan. 2013). ISSN: 0021-9533, 1477-9137. DOI: 10.1242/jcs.100115.

[118] James A. Grogan et al. “Microvessel Chaste: An Open Library for Spatial Modeling of Vascularized Tissues”. In: *Biophysical Journal* 112.9 (May 2017). ISSN: 0006-3495. DOI: 10.1016/j.bpj.2017.03.036.

[119] F. D. Bookholt et al. “Mathematical modelling of angiogenesis using continuous cell-based models”. en. In: *Biomechanics and Modeling in Mechanobiology* 15.6 (Dec. 2016). ISSN: 1617-7940. DOI: 10.1007/s10237-016-0784-3.

[120] Amy L. Bauer, Trachette L. Jackson, and Yi Jiang. “A Cell-Based Model Exhibiting Branching and Anastomosis during Tumor-Induced Angiogenesis”. en. In: *Biophysical Journal* 92.9 (May 2007). ISSN: 0006-3495. DOI: 10.1529/biophysj.106.101501.

[121] Diego Guidolin, Piera Rebuffat, and Giovanna Albertin. “Cell-Oriented Modeling of Angiogenesis”. In: *The Scientific World Journal* 11 (Oct. 2011). ISSN: 2356-6140. DOI: 10.1100/2011/586475.

[122] Lee D Hazelwood and John M Hancock. “Functional modelling of planar cell polarity: an approach for identifying molecular function”. In: *BMC Developmental Biology* 13 (May 2013). ISSN: 1471-213X. DOI: 10.1186/1471-213X-13-20.

[123] Paul van Liedekerke et al. “Influence of cell mechanics in embryonic bile duct lumen formation: insight from quantitative modeling”. en. In: (2021).